Journal of Theoretical and Applied Information Technology

31st October 2023. Vol.101. No 20

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ENHANCING SATELLITE IMAGE CLASSIFICATION IN NOISY ENVIRONMENTS WITH DEEP CONVOLUTIONAL NEURAL NETWORKS

Y.VISHNU TEJ[1], M. JAMES STEPHEN[2] , P.V.G.D. PRASAD REDDY[3]

[1]Research Scholar, Department of CS & SE, Andhra University, Visakhapatnam, AP, INDIA.
[2]Professor, Dr. B.R. Ambedkar Chair Professor, Andhra University, Visakhapatnam, AP, INDIA.
[3]Sr. Professor, Department of CS & SE, Andhra University, Visakhapatnam, AP, INDIA.

Email:[1]vishnutej.br@gmail.com

ABSTRACT

Satellite Image Classification is an essential aspect of remote sensing image processing. Satellite imagery has emerged as a pivotal data source for diverse applications such as land cover classification, urban planning, environmental monitoring, and disaster management. Satellite images are often subjected to various sources of noise, which can degrade the performance of traditional image classification techniques. Deep Convolutional Neural Networks (DCNNs) have shown remarkable success in image analysis tasks and have demonstrated potential for handling noisy satellite images. This paper investigates the performance of three popular DCNN architectures, namely VGG-16, ResNet-50, and Inception V4, for classifying noisy satellite images. To create a diverse and challenging dataset, we introduce noise into the original high-resolution satellite images, simulating real-world noise scenarios. The RSI-CB dataset covers various geographic regions and land cover types, encompassing the challenges faced during satellite image analysis. It contains six categories with 33 sub-classes and over 24,000, 256 X 256 pixel images. This paper contributes to advancing the use of DCNNs for satellite image classification in noisy environments. The study offers valuable guidance for selecting appropriate architectures based on the noise characteristics of satellite image datasets, ultimately enhancing the accuracy and reliability of satellite-based applications in challenging real-world conditions.

Keywords: *Deep Convolution Neural Network, Remote Sensing Image Classification, VGG-16, ResNet-50, Inception V4*

1. INTRODUCTION

Remote sensing has transformed our ability to understand and monitor the Earth's surface, providing invaluable insights into land cover, environmental changes, urban development, and natural disasters. Classification of satellite imagery, which involves designating specific land cover categories to individual image pixels, is one of the fundamental tasks in remote sensing image processing. This classification is essential for numerous applications, including urban planning, environmental monitoring, agricultural management, and emergency response.

However, the inherent challenges presented by satellite imagery frequently disrupt the efficacy of conventional image classification techniques. These obstacles include varying atmospheric conditions, sensor noise, impulse noise, variations in illumination, and the complexities of diverse land cover types. Despite such obstacles, the need for reliable and precise classification has prompted the investigation of advanced machine learning techniques, particularly Deep Convolutional Neural Networks (DCNN).

DCNNs, a subset of deep learning algorithms, have demonstrated remarkable success in various computer vision tasks, including image classification, object detection, and segmentation [1-3]. Due to their ability to acquire hierarchical features directly from raw data and their capacity to capture complex spatial patterns, they are ideally adapted for remote sensing image analysis [4-6]. In addition, DCNNs have the potential to handle noise and variations in satellite imagery that conventional methods have difficulty mitigating.

The benchmark for remote sensing image classification (RSI-CB) is available for download at https://github.com/lehaifeng/RSI-CB consists of six categories and 33 subclasses containing over 24,000 images. RSI-CB currently comprises six categories: agricultural land, construction land and facilities, transportation and facilities, water and water conservation facilities, forestry, and other land uses [7]. Each of these categories is subdivided into several subcategories. The main contributions of this paper are as follows: (1) The system consists of six categories and thirty-five subclasses. This data set can be utilized to effectively develop new data-driven algorithms and advance state-of-the-art techniques due to its large number of images, diverse objects, and complex categories. (2) Classical DCNN models, such as VGG-16[8], Inception V4 [9], and ResNet-50 [10] are used to train the Dataset. DCNN models trained by RSI-CB have good performance and generalization ability. Experiments demonstrate that RSI-CB is a more applicable standard for remote sensing image classification in the era of big data and has numerous potential applications [11]. This paper explores the classification of large-scale remote sensing images using DCNNs.

By adding impulse noise to high-resolution satellite images, we simulate real-world noise scenarios to evaluate the robustness and adaptability of our chosen DCNN architectures.

Our contributions to this study extend beyond benchmarking the efficacy of deep convolutional neural networks (DCNNs) on remote sensing imagery. We provide insights regarding the selection of suitable architectures based on the noise characteristics inherent in satellite image datasets. Our research seeks to enhance the accuracy and dependability of satellite-based applications in challenging real-world conditions by addressing the challenges of noise environments. This research ultimately advances the use of deep convolutional neural networks (DCNNs) for satellite image classification, bridging the divide between cutting-edge deep learning techniques and the needs of practical remote sensing applications.

2. RELATED WORK

Basic Requirements for Remote Sensing Image Benchmark Using Deep Learning: Deep learning models, such as DCNN, have significantly advanced in various tasks, including image tracking and scene understanding. DCNN models are highly complex and contain millions of parameters; consequently, they are susceptible to overfitting on tiny benchmarks [8-10].

Incorporating advanced machine learning techniques has led to a revolutionary transformation in classifying remote sensing images. Traditional methods relied on hand-engineered features and pixel-based approaches and frequently needed help conveying satellite imagery's inherent complexity and variability. Deep learning, particularly Deep Convolutional Neural Networks (DCNNs), has propelled remote sensing image classification into new domains of precision and adaptability.

Initially designed for general image analysis tasks, deep convolutional neural networks (DCNNs) have effectively confronted remote sensing challenges. Several architecture families have been utilized for remote sensing image classification, each contributing distinctive advantages to the field.

The VGG architecture, introduced by Simonyan and Zisserman [8], emphasized the significance of network depth when learning complex features. The VGG-16 variant, with its uniform architecture and stacked 3x3 convolutional layers, demonstrated the effectiveness of deep neural networks in image analysis. VGG-16's capacity to understand complex features incrementally prepared the way for its adaptation to remote sensing [8]. Researchers recognized its potential to reveal the subtle land cover characteristics embedded in satellite images.

The ResNet family pioneered the concept of residual connections [10] by utilizing this foundation. Residual networks, such as ResNet-50, mitigated the vanishing gradients problem by incorporating shortcut connections that allowed the network to learn residual mappings. This innovation enabled the training of much deeper neural networks and sparked a revolution in the classification of remote sensing images [10]. The deep architecture of ResNet-50 demonstrated a remarkable aptitude for learning complex features, making it ideally suited for capturing the intricate patterns present in satellite images.

In contrast, the Inception architecture introduced a novel method for feature extraction by employing multiple filter sizes in parallel branches [13]. This parallel processing improved the network's ability to capture various image details. Utilizing the assets of both architectures, the subsequent development of Inception-v4, which incorporated elements from the ResNet architecture, enhanced its capabilities even further [13]. The inherent capacity of the architecture to capture features at multiple scales complemented the multiscale nature of remote sensing images.

In remote sensing image classification, these architectures have been intensively studied on various datasets, ranging from multispectral satellite images to high-resolution aerial photographs. Zhang

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

et al. [12] examined deep learning architectures, such as VGG-16, ResNet-50, and Inception-v3, for classifying land cover in high-resolution remote sensing images. Their findings underscored the importance of architectural depth in remote sensing applications by highlighting the benefits of deeper architectures in distinguishing complex land cover features.

Recent efforts in the field have broadened the investigation to include chaotic data, which is more representative of actual remote sensing conditions. Researchers have injected synthetic noise into remote sensing datasets [14] to simulate the unpredictability introduced by atmospheric conditions and sensor noise. This direction recognizes the practical difficulties encountered in satellite image analysis, where noise can induce uncertainty in classification results. This paper evaluates the performance of three prominent DCNN architectures (VGG-16, ResNet-50, and Inception V4) under both noisy and noise-free conditions. The study provides valuable guidance on selecting the most appropriate DCNN architecture based on noise characteristics, thereby improving the accuracy and reliability of satellite-based applications in challenging real-world conditions. It also addresses the critical issue of noise in satellite image analysis, bridging the gap between cutting-edge deep learning techniques and practical remote sensing requirements. This research is significant for satellite image classification because VGG-16's depth facilitates capturing intricate patterns, ResNet-50's skip connections assist in training deep networks, and Inception V4 excels at managing multiscale and complex features.

Despite these commendable advancements, there is a cavity in the literature regarding evaluating these architectures on large-scale remote sensing image classification tasks with chaotic data. To address this deficiency, this study aims to comprehensively assess the performance of the VGG-16, ResNet-50, and Inception V4 architectures on the "RSI-CB" dataset. By incorporating synthetic noise and including diverse land cover categories, the "RSI-CB" dataset provides a unique perspective for evaluating the adaptability of these architectures in a remote sensing context characterized by noise.

3. DATASET

3.1. Distribution Characteristics of RSI-CB:

The RSI-CB benchmark comprises 33 subcategories containing approximately 24,000 images, with an average of approximately 690 images per category. According to Table 1, the significant classes correspond to their subclasses [7,11]. Additional subcategories are within the main types of transportation and facilities, forests, and water and water conservation facilities. Table 2 shows the frequency of each category.

RSI-CB has the advantages of spatial resolution, quantity, and a novel database construction method. The construction of RSI-CB is based not only on the meaning of the database but also on the crowdsourced data-based method for its potential application value in weak-supervised learning to achieve automated data annotation and error correction [7].

Increasing the number of images in these categories may improve the chance of the inter-class feature response interval having an independent distribution, enhancing image classification accuracy.

The RSI-CB construction principles are as follows:

1. Each category must contain an abundance of data. Per category, RSI-CB contains approximately 690 patches.
2. The level of each category is intended to enhance the diversity and breadth of the standard. The data set consists of 33 leaf nodes (sub-classes) connected to six parent nodes (parent classes) in a two-level tree structure.
3. The primary substance of the object must be easily distinguishable to prevent semantic divergence in images.
4. Each class has various imaging angles, sizes, shapes, and colors to increase sample diversity, enhancing the model's generalization performance and robustness.

4. METHODOLOGY

Adding different levels of impulse noise to the images to see how well the models could classify them correctly. Introduced impulse noise with varying probabilities to the test images to simulate real-world scenarios with different levels of image distortion. And trained three popular DCNN architectures – VGG-16, ResNet 50, and Inception V4 on the dataset and subsequently tested these models using the noisy images to assess their classification accuracy and performance under noisy conditions.

Data organization comprises three primary aspects: selection of the training, validation, and test sets for RSI-CB; data augmentation; and data organization for model transfer performance.

1. Randomly selecting data: The training, validation, and test sets are chosen randomly according to a

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

particular proportion, and labeling is disrupted further to reflect the randomness and objectivity of the data.
2. Data augmentation: We augment all RSI-CB data for each image by cutting a fixed-size segment from the center and upper-left, upper-right, lower-left, and lower-right corners of each image and flipping them before feeding them into the DCNN. Thus, the original data is multiplied by ten.
3. Data organization for model transfer performance: We examine the transferability of the RSI-CB training model to other data sets and the efficacy of its transferability.

Table 1. The sub-categories corresponding to the large categories in RSI-CB

Large Class	Subclass
Agricultural land	green_farmland, dry_farm, bare_land
Woodland	Artificial_grassland, sparse_forest, forest, mangrove, river_protection_forest, shrubwood, sapling
Transportation and facility	airport_runway, avenue, highway, marina, parking lot, crossroads, bridge, airplane
Water area and facility	coastline, dam, Hirst, lakeshore, river, sea, stream
Construction land and facility	city_building, container, residents, storage_room, pipeline, town
Other land	desert, snow_mountain, mountain, sand beach

Table 2. Different sub-classes and the number of images in each

Categories	Number	Categories	Number	Categories	Number
airplane	351	dry_farm	1309	river	539
airport_runway	678	Forest	1082	river_protection_forest	524
artificial_grassland	283	green_farmland	644	Sand beach	536
avenue	544	highway	223	sapling	879
bare_land	Hirst	hirst	628	sea	1028
bridge	469	lakeshore	438	Shrub wood	1331
city_building	1014	mangrove	1049	snow_mountain	1153
coast_line	459	marina	366	sparse_forest	1110
container	660	mountain	812	storage_room	1307
crossroads	553	parking lot	467	stream	688
dam	324	pipeline	198	town	355
desert	1092	residents	810		

4.1. Data Collection and Preparation

The "RSI-CB 256" dataset, an extensively curated collection of remote sensing images encompassing a wide variety of land cover types and geographic regions, is the foundation of our study. The dataset, consisting of over 24,000 high-resolution images with a resolution of 256x256 pixels, is a comprehensive benchmark for evaluating the performance of DCNN architectures in large-scale remote sensing image classification.

Impulse noise is introduced into the original high-resolution satellite images to make the dataset with the complexities encountered in real-world remote sensing scenarios. This noise synthesis method is inspired by atmospheric distortions and sensor-induced artifacts, thereby augmenting the dataset's realism and applicability to real-world scenarios. Noise simulation is used to examine how well the chosen DCNN architectures can perform in the presence of varying levels and types of noise.

4.2. Deep Convolutional Neural Network Architectures

Our investigation focuses primarily on three notable DCNN architectures: VGG-16, ResNet-50, and Inception V4. These architectures have attained much attention for their capabilities in image analysis tasks and have demonstrated remarkable success across various datasets [15-17].

4.2.1. VGG-16 architecture

VGG-16 has 16 layers, 13 convolutional layers, and three fully linked layers.

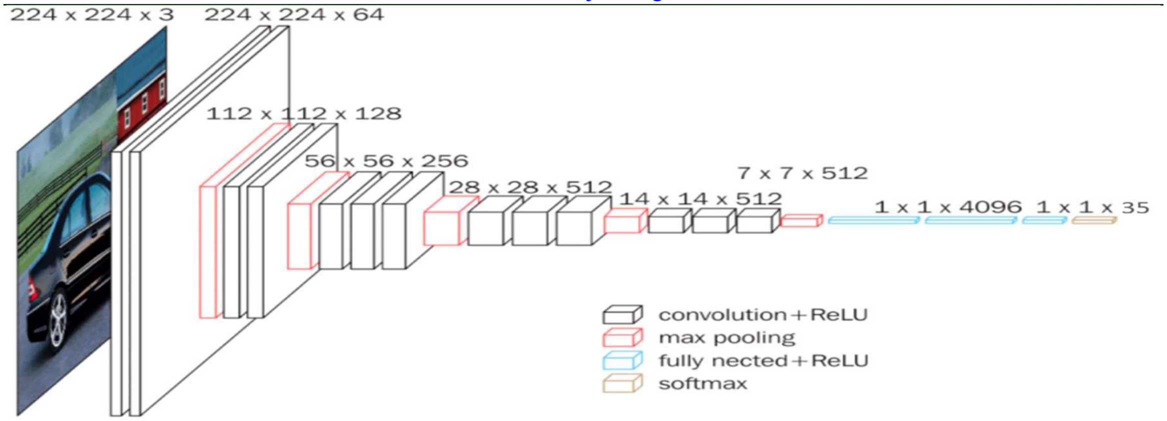


Figure 1. VGG-16 Architecture

Figure 1 shows the VGG-16 Architecture. All convolutional layers feature a 3x3 pixel receptive field and are followed by a Rectified Linear Unit (ReLU) activation function. The network architecture is summarized below:
VGG-16's input layer takes an RGB image with a resolution of 224x224 pixels. The image has three color channels (Red, Green, and Blue), each having a pixel value ranging from 0 to 255. Convolutional Layers: VGG-16 has 13 convolutional layers, each using a tiny 3x3 filter with a stride of 1. To extract features, these filters are applied to the input image. The first layer has 64 filters, followed by layers with 128, 256, 512, and 512 filters. As we go further into the network, the number of filters in each layer rises, enabling VGG-16 to learn more complicated patterns and higher-level features.
Activation Function: A Rectified Linear Unit (ReLU) activation function is applied element-by-element after each convolutional layer. The ReLU function adds non-linearity, allowing the network to learn and represent more complicated data connections.
Max Pooling Layers: A max pooling layer with a 2x2 window and a stride of 2 is applied after every two convolutional layers. Max pooling aids in reducing the spatial dimensions of feature maps while preserving the most relevant characteristics. It also adds some translation invariance, making the model more resistant to changes in the location of objects in the input image.
Fully Connected Layers: Three fully connected layers follow the convolutional layers. These traditional artificial neural network layers link all neurons from the previous layer to all neurons in the next layer. The first two completely connected layers contain 4,096 neurons, followed by a third fully connected layer with several neurons equal to the number of classes in the classification job (for example, 1,000 for ImageNet classification).
Softmax Activation: The VGG-16 output layer employs the softmax activation algorithm. The last layer outputs are converted into class probabilities using the softmax algorithm. Each output shows the likelihood that the input image belongs to a particular class. The projected class for the input image is the class with the most significant probability.
VGG-16 is distinguished by its deep architecture, which employs many convolutional and fully linked layers to learn hierarchical representations of input data. While it performed well on image classification tests, its most significant disadvantage is its high computational cost owing to its depth. Nonetheless, VGG-16 is a powerful model that laid the path for following advances in deep learning and computer vision.

4.2.2. ResNet-50 Architecture

ResNet-50 is a ResNet (Residual Neural Network) family deep convolutional neural network architecture. It was developed by Microsoft Research experts and has 50 layers. Using residual blocks and skip connections is the major innovation of ResNet-50. Figure 2 shows the ResNet-50 Architecture.
ResNet-50 receives an RGB image with 224x224 pixels as input. The image has three color channels (Red, Green, and Blue), and the pixel values are generally scaled to the range [0, 1].
Convolutional Layers: The ResNet-50 design starts with a conventional convolutional layer with 64 filters, a 7x7 kernel size, and a stride of 2. To minimize the spatial dimensions of the input image, this layer conducts a down sampled convolution on it.
Batch Normalisation and Activation: A batch normalization layer is added after the first convolutional layer, followed by a ReLU activation function. Batch normalization normalizes the layer

Journal of Theoretical and Applied Information Technology
31st October 2023. Vol.101. No 20

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

activations to stabilize and accelerate training. The activation of the ReLU brings non-linearity into the network.

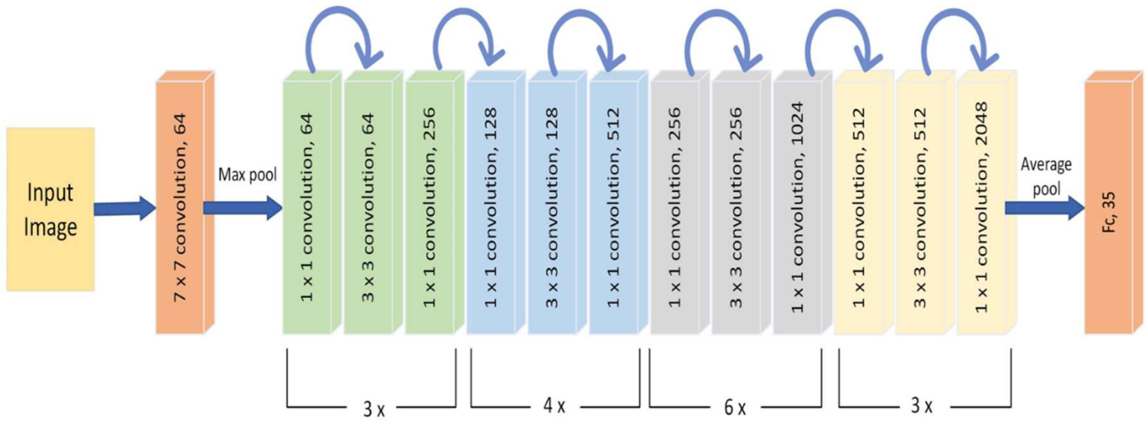


Figure 2. ResNet-50 Architecture

Max Pooling: A max pooling layer with a 3x3 window and a stride of 2 is applied after the initial batch normalization and ReLU layer. Max pooling minimizes the spatial dimensions of the feature maps, which aids in downsampling and increasing translation invariance.
Residual Blocks: The residual blocks are the main building blocks of ResNet-50. The network is divided into four phases, each comprising many residual blocks. Each residual block contains a sequence of convolutional layers with filter configurations of 1x1, 3x3, and 1x1. The 1x1 convolutions are used to lower and increase the number of filters, resulting in a more efficient network. Each step raises the number of filters by a factor of two (64, 128, 256, and 512).
Shortcut connections in the residual blocks are also responsible for "skipping" one or more convolutional layers. The output of the convolutional layers is added to the output of the shortcut connection element by element. This skip link allows the gradient to travel straight across the network, overcoming the vanishing gradient issue and aiding deep network training.
Global Average Pooling: A global average pooling layer is used after the last stage of residual blocks. The spatial dimensions of the feature maps are reduced to a single value per channel via global average pooling, resulting in a fixed-length feature vector.
Fully Connected Layers: The feature vector from the global average pooling layer is passed into fully connected layers for classification. ResNet-50 features a final fully connected layer with several units equal to the number of categorization classes.
The output layer employs the softmax activation function to generate the final class probabilities.

4.2.3. Inception V4 Architecture

Inception version 4 is a deep convolutional neural network architecture proposed by Google researchers. It is an extension and enhancement of the original Inception architecture (GoogLeNet), combining principles from the Inception and ResNet architectures. Figure 3 shows the Inception V4 Architecture.
The Inception-v4 architecture is highly complicated, with the following essential components and features:
Modules for initialization: The inception modules are the fundamental building pieces of Inception-v4. These modules use concurrent convolutional filters of varying sizes (1x1, 3x3, and 5x5) and max-pooling procedures. These parallel filters' outputs are then concatenated and supplied to the next layer. This approach enables the network to catch information at different sizes and learn multiple representations from input data.
Factorization and Bottleneck: Inception-v4 employs factorization and bottleneck methods to minimize computing complexity. Instead of a single 5x5 convolutional filter, it uses two 3x3 filters, reducing the number of parameters while preserving expressiveness. This "bottleneck" structure aids in the creation of a more complex network.
Stem and Reduction blocks: Inception-v4 adds specialized "stem" and "reduction" blocks at the network's beginning and middle, respectively. The stem block is in charge of initial feature extraction and lowering the spatial dimensions of the input. In contrast, the reduction block decreases the spatial

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

dimensions even more while increasing the filters. These blocks improve the network's capacity to learn hierarchical features.

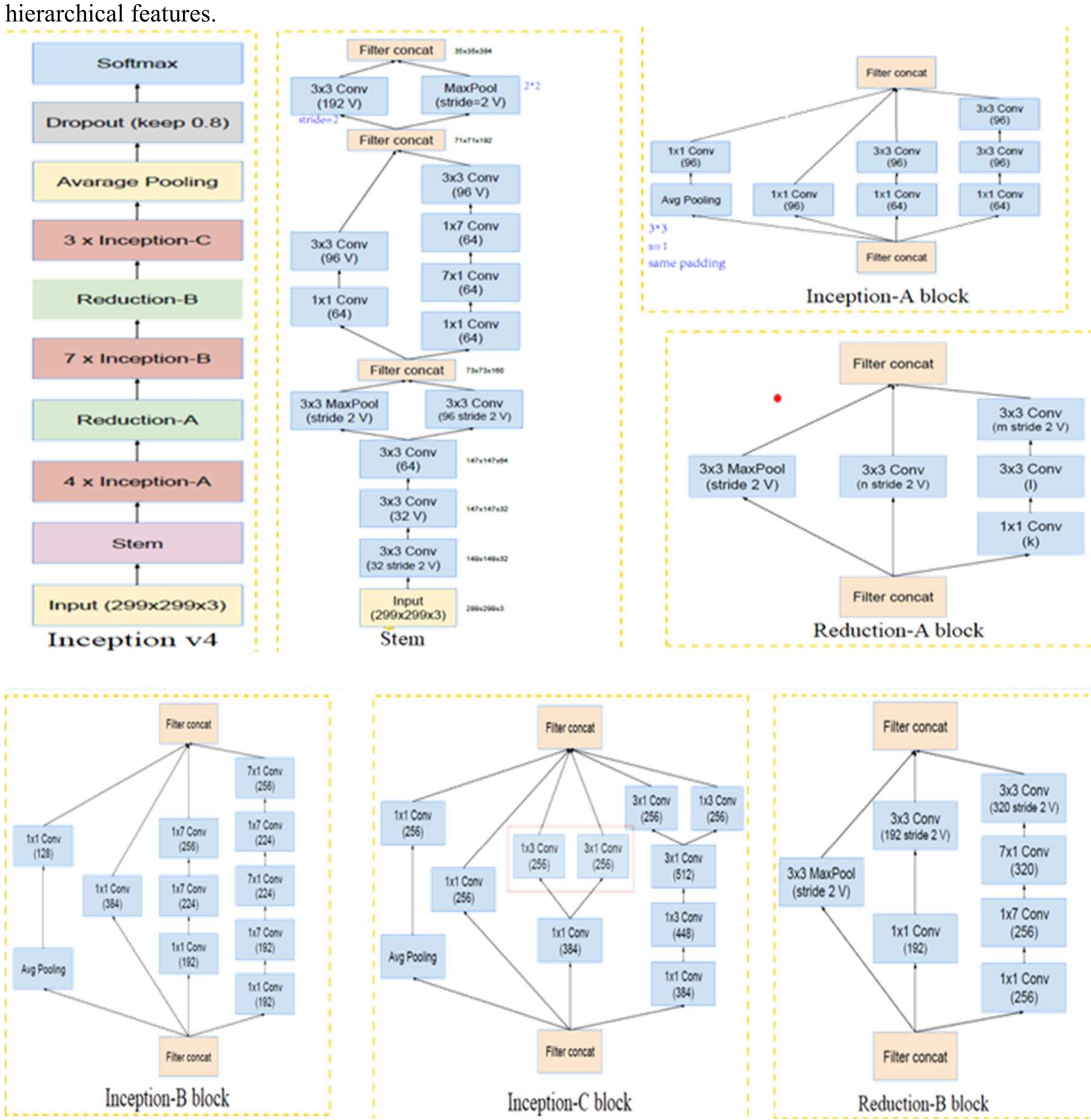


Figure 3. Inception V4 Architecture

Inception-ResNet: Inception-v4 adds residual connections (inspired by ResNet) into certain network regions, resulting in the Inception-ResNet hybrid design. These residual connections contribute to smoother optimization and improved gradient flow during training, hence minimizing the vanishing gradient issue.

Inception-v4 is distinguished by its breadth and depth, implying that it contains many filters (width) and a deep network structure. This enables the model to capture more complicated patterns and hierarchies in the input data, improving image recognition accuracy.

5. EXPERIMENTAL RESULTS

Accuracy is a fundamental performance metric used to assess the correctness of a classification model's predictions. It provides a straightforward measurement of how well the model's predictions correspond to the actual ground truth labels in the dataset. The accuracy metric is particularly useful when the classes in the dataset are approximately balanced, meaning that the number of instances in each class is roughly equal.

"Number of Correctly Classified Instances" refers to the number of instances for which the model's

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

predicted class matches the actual ground truth class in the accuracy formula. The "Total Number of Instances" indicates the extent of the complete dataset.

A higher accuracy value indicates that the model's predictions align well with the actual labels, whereas a lower accuracy suggests that the model has difficulty accurately classifying instances. While accuracy provides a comprehensive view of the model's performance.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

True Positive (TP) refers to the number of predictions where the classifier correctly predicts the positive class as positive.

True Negative (TN) refers to the number of predictions where the classifier correctly predicts the negative class as negative.

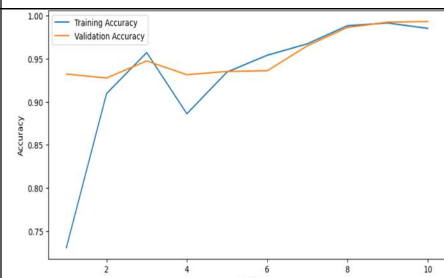
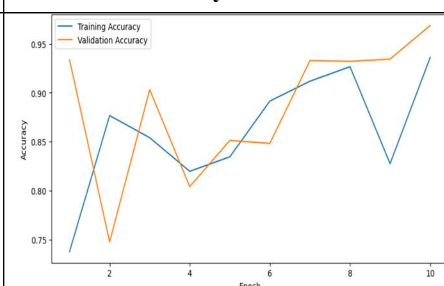
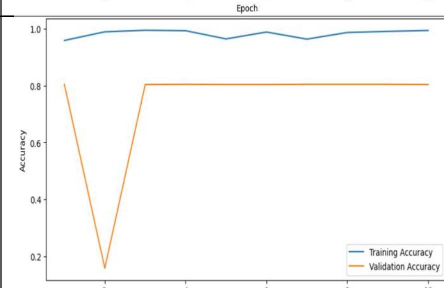
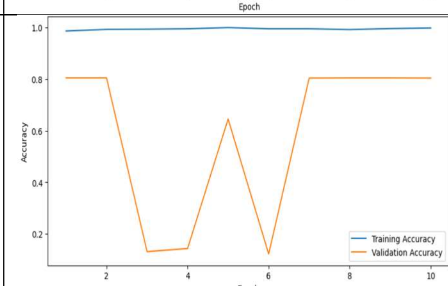
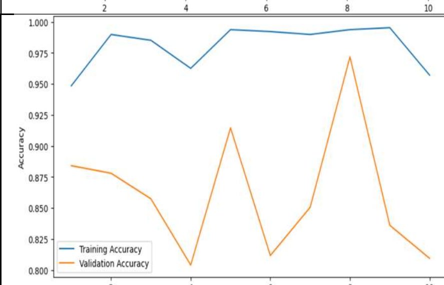
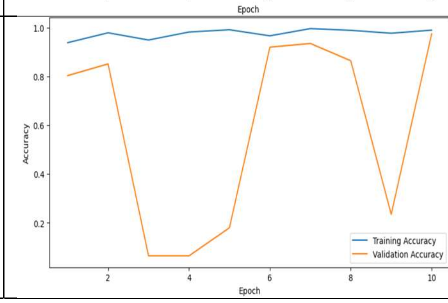
False Positive (FP) refers to the number of predictions where the classifier incorrectly predicts the negative class as positive.

False Negative (FN) refers to the number of predictions where the classifier incorrectly predicts the positive class as negative.

Table 3 Describes The Overall Performance Of Different Models Based On DCNN.

. Methods	Trained Data		Test Data	
	Accuracy (without Noise)	Accuracy (with Noise)	Accuracy (without Noise)	Accuracy (with Noise)
VGG-16	97.43%	95.32%	98.63%	93.46%
ResNet-50	98.42%	97.30%	80.44%	80.33%
Inception V4	98.28%	97.17%	96.88 %	94.30%

Table 4. Shows The Classification Accuracy With Noise And Without Noise

Method	Accuracy without Noise	Accuracy with Noise
VGG-16		
ResNet-50		
InceptionNet V4		

The test results of the models with noise and without noise of VGG-16 and Inception V4 for RSI-CB were more than 90%. Regarding Trained data accuracy without noise, ResNet-50 achieves the highest accuracy of 98.42%, followed by Inception V4 with 98.28% and VGG-16 with 97.43%. Regarding Test data accuracy without noise, VGG-16 achieves the highest accuracy of 98.63%, followed by Inception V4 with 96.88% and ResNet-50 with 80.44%. This indicates that ResNet-50 performs exceptionally well in accurately classifying the Trained data without noise, and VGG-16 Performs well in accurately classifying the test data outperforming the other two models. When noise is introduced in the Trained data, the performance of the models fluctuates. ResNet-50 maintains a high accuracy of 97.30%, showcasing its robustness against noise. Inception V4 model with an impressive accuracy of 97.17%, indicating its ability to handle noisy data effectively. In the Test data, Inception V4 maintains a high accuracy of 94.30%, and VGG-16 with an accuracy of 93.46%, showcasing its robustness against noise. However, VGG-16 experiences a significant drop in accuracy to 95.32% in the trained data, and ResNet-50 significantly drops to 80.33%, suggesting that it is more noise sensitive than the other two models.

It is worth noting that ResNet-50 demonstrates consistent performance across both scenarios, while Inception V4 shows a notable improvement in accuracy when noise is present. On the other hand, VGG-16 performs well in the absence of noise but struggles to maintain accuracy when noise is introduced.

6. CONCLUSION

This paper focuses on the challenges posed by noisy environments through the application of Deep Convolutional Neural Networks (DCNNs) for remote sensing image classification. This paper investigates the efficacy of three prominent DCNN architectures - VGG-16, ResNet-50, and Inception V4 on the RSI-CB 256 dataset, which consists of over 24,000 high-resolution images representing diverse land cover categories.

The paper highlights the importance of satellite image classification in various applications, including urban planning, environmental monitoring, and disaster management. It recognizes that satellite images are susceptible to multiple noise sources, which can hinder conventional classification techniques. DCNNs, which are renowned for their capacity to acquire complex patterns and hierarchical characteristics, show promise in dealing with noisy satellite images.

This paper aims to assess the suitability of DCNN architectures under noisy conditions and provide guidance for selecting architectures based on noise characteristics. To accomplish this, impulse noise is introduced into high-resolution satellite images to simulate actual noise scenarios. They illustrate the diversity of the RSI-CB dataset, which consists of six categories with 33 sub-classes and 256x256 pixel images.

This paper compares the performance of VGG-16, ResNet-50, and Inception V4 on both trained and test data sets through extensive experimentation. While all models exhibit high accuracy with clear data, their efficacy varies when noise is added. VGG-16 exhibits consistent performance in both scenarios, whereas Inception V4 demonstrates enhanced noise accuracy. ResNet-50 performs admirably on clear data but struggles to sustain precision when noise is present. This paper provides valuable insights into selecting suitable architectures based on noise characteristics, thereby augmenting the accuracy and dependability of satellite-based applications in difficult real-world conditions. The authors address the crucial issue of noise in satellite image analysis by bridging the divide between cutting-edge deep learning techniques and the requirements of practical remote sensing applications.

Limitations of this work, This paper provides valuable insights into the application of Deep Convolutional Neural Networks (DCNNs) for satellite image classification in noisy environments, with an emphasis on impulse noise. However, it provides numerous opportunities for more in-depth and exhaustive discussions. To provide a deeper comprehension of model performance, it could benefit from more detailed quantitative results, such as a broader range of evaluation metrics beyond accuracy. A comparative analysis with traditional or baseline models, discussions on generalization to other datasets and real-world applications, and an examination of the potential ethical implications of satellite image classification would be useful in enhancing the context.

REFERENCES:

[1] Girshick, R., 2015. Fast r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 1440-1448).

[2] Liu, Y., Zhang, Y.M., Zhang, X.Y. and Liu, C.L., 2016. Adaptive spatial pooling for image classification. Pattern Recognition, 55, pp.58-67.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[3] Zeng, X., Ouyang, W., Yang, B., Yan, J. and Wang, X., 2016. Gated bi-directional cnn for object detection. In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part VII 14 (pp. 354-369). Springer International Publishing.

[4] Nogueira, K., Penatti, O.A. and Dos Santos, J.A., 2017. Towards better exploiting convolutional neural networks for remote sensing scene classification. Pattern Recognition, 61, pp.539-556.

[5] Salberg, A.B., 2015, July. Detection of seals in remote sensing images using features extracted from deep convolutional neural networks. In 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (pp. 1893-1896). IEEE.

[6] Hu, F., Xia, G.S., Hu, J. and Zhang, L., 2015. Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. Remote Sensing, 7(11), pp.14680-14707.

[7] Li, H., Dou, X., Tao, C., Wu, Z., Chen, J., Peng, J., Deng, M. and Zhao, L., 2020. RSI-CB: A large-scale remote sensing image classification benchmark using crowdsourced data. Sensors, 20(6), p.1594.

[8] Simonyan, K. and Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[9] Shankar, K., Zhang, Y., Liu, Y., Wu, L. and Chen, C.H., 2020. Hyperparameter tuning deep learning for diabetic retinopathy fundus image classification. IEEE Access, 8, pp.118164-118173.

[10] He, K., Zhang, X., Ren, S. and Sun, J., 2015. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In Proceedings of the IEEE international conference on computer vision (pp. 1026-1034).

[11] Dorigo, M. and Stützle, T., 2019. Ant colony optimization: overview and recent advances (pp. 311-351). Springer International Publishing.

[12] He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[13] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2818-2826. doi:10.1109/CVPR.2016.308

[14] Jensen, D.J., 2019. Imaging Spectroscopy Applications for Mapping and Monitoring Environmental Change in Louisiana's Coastal Wetlands. University of California, Los Angeles.

[15] Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2017. Imagenet classification with deep convolutional neural networks. Communications of the ACM, 60(6), pp.84-90.

[16] Simonyan, K. and Zisserman, A., 2014. Two-stream convolutional networks for action recognition in videos. Advances in neural information processing systems, 27.

[17] Vogelmann, J.E., Howard, S.M., Yang, L., Larson, C.R., Wylie, B.K. and Van Driel, N., 2001. Completion of the 1990s National Land Cover Data Set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. Photogrammetric Engineering and Remote Sensing, 67(6).