OPEN ACCESS
International Journal of Innovations in Science & Technology

Nature Scene Classification Using Transfer Learning with Inception V3 on the Intel Scene Dataset

Muhammad Talha Jahangir[*1], Syed Yameen Ali[1], Muhammad Hamza Khan[2], Ahmad Ali Zain[1], Umair Arshad[1]

[1]Department of Computer Science, MNS UET, Multan, Pakistan.

[2]Department of Electrical Engineering, MNS UET, Multan, Pakistan.

[*]**Correspondence:** email: mtalhajahangir@mnsuet.edu.pk

Citation | Jahangir. M. T, Ali. S. Y, Khan. M. H, Zain. A. A, Arshad. U, "Nature Scene Classification Using Transfer Learning with Inception V3 on the Intel Scene Dataset", IJIST, Vol. 06 Issue. 03 pp 1537-1553, Sep 2024

Received | Aug 27, 2024 **Revised |** Sep 26, 2024 **Accepted |** Sep 27, 2024 **Published |** Sep 30, 2024.

Nature scene classification is vital for various applications, including environmental monitoring and autonomous systems need to develop efficient models that can sort out different scenes. This work proposes a new approach using state-of-the-art CNNs like InceptionV3, Xception and VGG19, to enhance the classification accuracy and generalization of nature scenes. We worked with six classes with 20,926 training images and 5,228 validation images and augmented the data to improve the model. Models were fine-tuned from the pre-trained models of ImageNet and early stopping and model checkpoints were used to avoid overfitting. The results indicated that the proposed InceptionV3 model achieved a training accuracy of 94.49% and validation accuracy of 92.81% which is higher than previous work and Xception model had a high accuracy of 95.52% but the model might be overfitting. During the comparison of the results, it was revealed that InceptionV3 provided the highest accuracy with the least standard deviation, which proved the effectiveness of the selected architecture for scene classification. These results indicate that the selection of the model and the technique for the classification of nature scenes is important. It is a good advancement in the field of nature scene classification and provides a reliable solution to enhance accuracy in real-world scenarios.

Keywords: Nature Scene; Xception Model; Inception-V3 Model; Convolutional Neural Network; Transfer Learning.

Introduction:

Nature scene classification is a simple computer vision task that involves categorizing images of outdoor environments into several predefined categories. This area has garnered significant attention due to its applications in environmental monitoring, autonomous navigation, and content-based image search. In recent years, deep learning and convolutional neural networks (CNNs) have revolutionized picture classification. These powerful models can acquire the hierarchical representations of the visual features and therefore can learn the low-level textural features as well as the higher-level semantic features. In nature scene classification, it has been demonstrated that CNNs yield outstanding performance and are even superior to hand-designed feature extraction methods. The Intel Image Classification dataset is an excellent opportunity to investigate and compare the various deep learning architectures for the classification of nature scenes. This dataset can be employed by researchers to develop new architectures for the network, investigate transfer learning techniques, and evaluate data augmentation techniques to improve the classification accuracy and generalization capability of the network.

Advancements in remote sensing technologies have enabled the effortless acquisition of substantial satellite data [1]. The high-resolution satellite images now accessible, have created both opportunities and challenges for those in the remote sensing community, including researchers and scientists. In recent years, researchers and scientists have relied on remote sensing data to investigate different semantic tasks such as road segmentation [2], land cover semantic segmentation [3] and classification [4], building extraction [5], farmland segmentation [6], and multiple geospatial objects detection [7]. Among these semantic tasks, the classification of land cover has received major attention from the research community because of its wide-ranging applications in different fields. Identifying the scene is the objective of land cover classification, given a remote sensing image. Such information may be applied in urban planning [8][9], disaster evaluation [10], landslide risks [11], ecosystem monitoring [12], depletion of groundwater [13], crop fields [14], and several other disciplines.

The primary challenge in classifying natural sceneries is in the multitude of objects and their variations present in outdoor settings. Outdoor landscapes are dynamic and may encompass multiple items, varying lighting conditions, and intricate textural patterns, complicating the algorithms' ability to identify specific characteristics. In addition, the difference between the various scene types may not be clear and this makes the categorization process even more difficult. In the case of image classification in remote sensing, objects are rotated within a view, and the background is more heterogeneous than in the case of aerial images [15].

To overcome these challenges, there is a need to establish a much-improved nature scene classification system. By leveraging breakthroughs in deep learning, our proposed methodology can significantly enhance scene recognition by enabling accurate classification across six diverse nature scene categories: building, forest, glacier, mountain, sea, and street. It gives a good structure to work with and it can accommodate real-life scenarios which improves the model's generalization ability and thus prevents overfitting. Further, our approach reduces the process of feature extraction since deep learning architectures can learn features from images. Since this approach can be applied to environmental monitoring, city planning, and landscape assessment, it is connected to the public interest and addresses major issues in various disciplines.

Novelty and Objectives of Study:

- Systematic evaluation of multiple pre-trained deep learning models (VGG19, MobileNetV2, ResNet, InceptionV3, Xception) to classify nature scenes, which has not been extensively compared in prior studies.
- Achieving higher accuracy in nature scene classification than reported in previous studies by optimizing model architectures and training approaches.

- Providing a detailed performance comparison of models using diverse metrics, offering insights into which model is best suited for nature scene classification tasks.

The objective of this research is to identify the most accurate and efficient deep learning model for classifying six distinct categories of nature scenes, surpassing the accuracy achieved in previous studies.

Literature Review:

In 2017, Drawing inspiration from the Inception-V3 model, Francois Chollet developed the Xception model to make better use of the parameters by replacing some of the Inception modules with depth-wise separable convolutions [16]. Scene classifications remain problematic, however, when limited to ImageNet data alone. A novel ML model for categorizing scenes using the Intel scene dataset, Rep Conv, was presented by Soudy, M., et al [17]. The suggested model outperformed the benchmark models in terms of efficiency and obtained accuracies of 93.55 ± 0.11 and 75.54 ± 0.14 in the multiclassification training and validation datasets, respectively.

An innovative approach to scene categorization has been proposed by Xizhi Wu et al., which utilizes transfer learning to operate on the Xception system trained on the data set of ImageNet. Also, we show that the model can spot mountain and glacier vistas that weren't in the pre-training photographs. Xception achieved a success rate of 91.20%, whereas Inception-V3 of transfer learned categorization had a success rate of 91.81%, according to the experimental data [18].

Dataset:

The data set used in this research is the Intel Image Classification data set, which is available on Kaggle [19]. It comprises approximately 25,000 images, each of size 150x150 pixels, categorized into six distinct scene types: building, forest, glacier, mountain, sea, and street. These categories are labeled numerically as follows: buildings (0), forest (1), glacier (2), mountain (3), sea (4), and street (5).

Table 1. Class distribution in both the training and test data

Class	Train	Test
Buildings	3434	437
Forest	3433	474
Glacier	3540	553
Mountain	3626	525
Sea	3436	510
Street	3522	501

Table 1 gives the distribution of the class for the training and test data set and the number of samples in each class. The training set includes about 14,000 images used to build and adjust the model. The test set comprises approximately 3,000 images that have never been used in the training process and can be used to assess the performance of the model. Last of all, the prediction set with about 7000 images is used for prediction when the model is built. Figure 1 presents a series of natural scenes designed to encompass the several attributes pertinent to the training of the picture classification model. Using this dataset, the goal is to develop a system that can accurately identify various scenes which can be applied in automated scene analysis and geographical information systems.

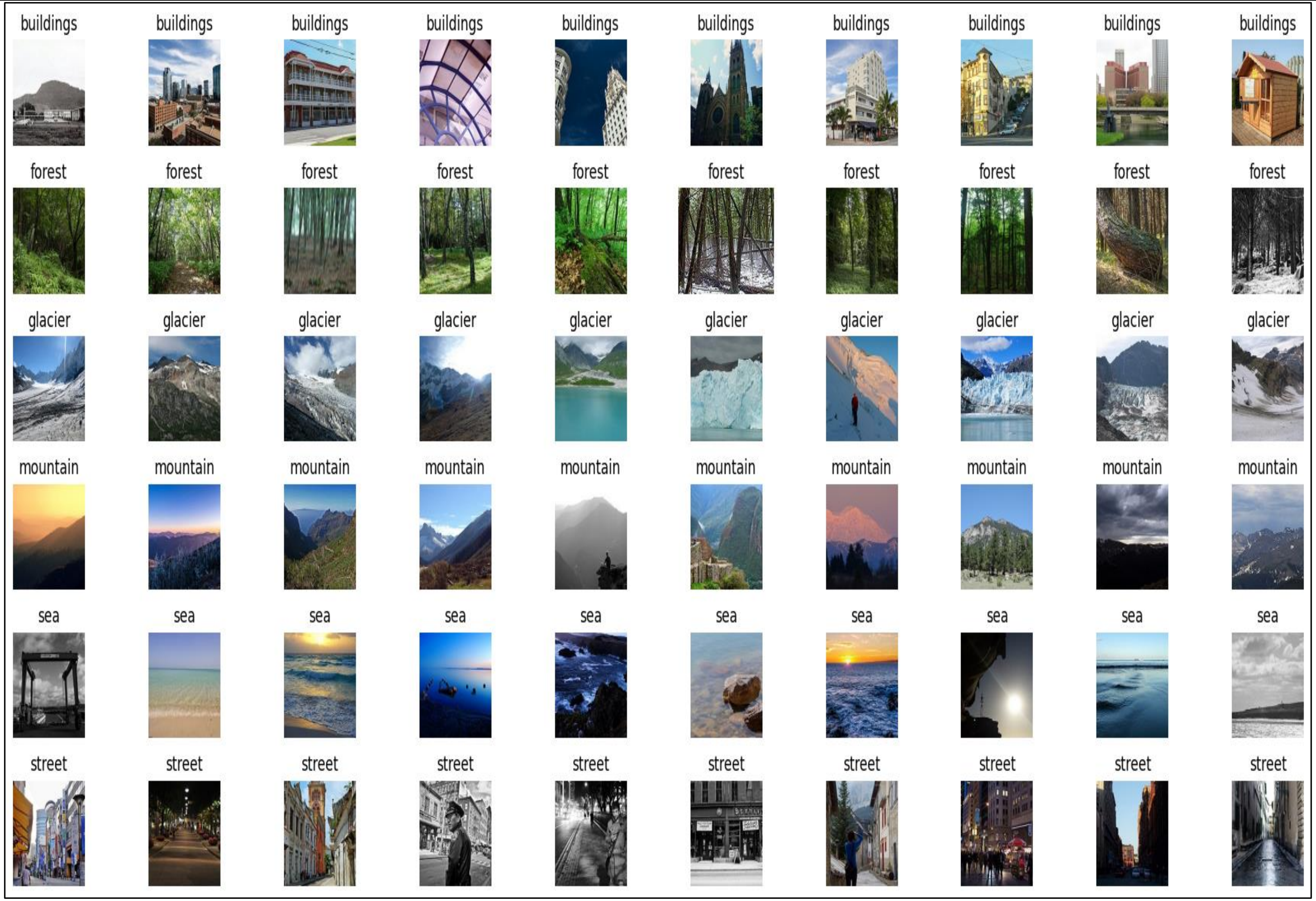

Figure 1. Samples of the images in each of the scene categories in the dataset.

The distribution of images across the different nature scene categories is shown in Figure 2.

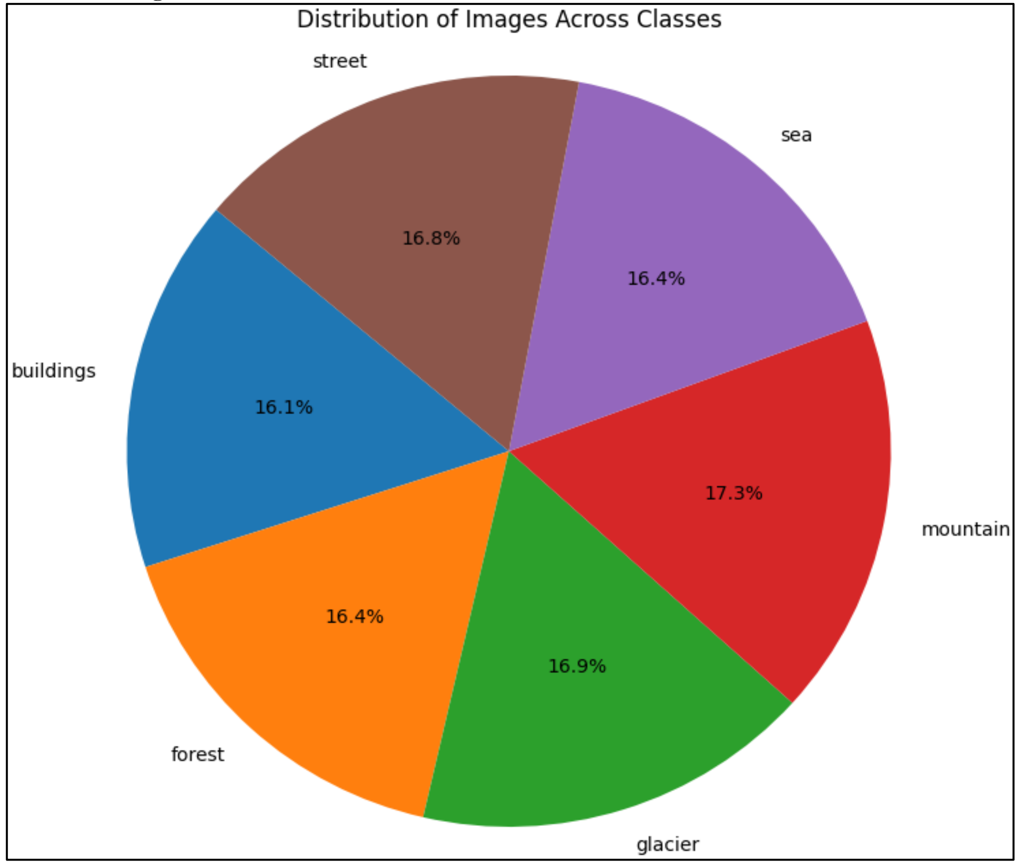

Figure 2. Distribution of Images Across Nature Scene Categories

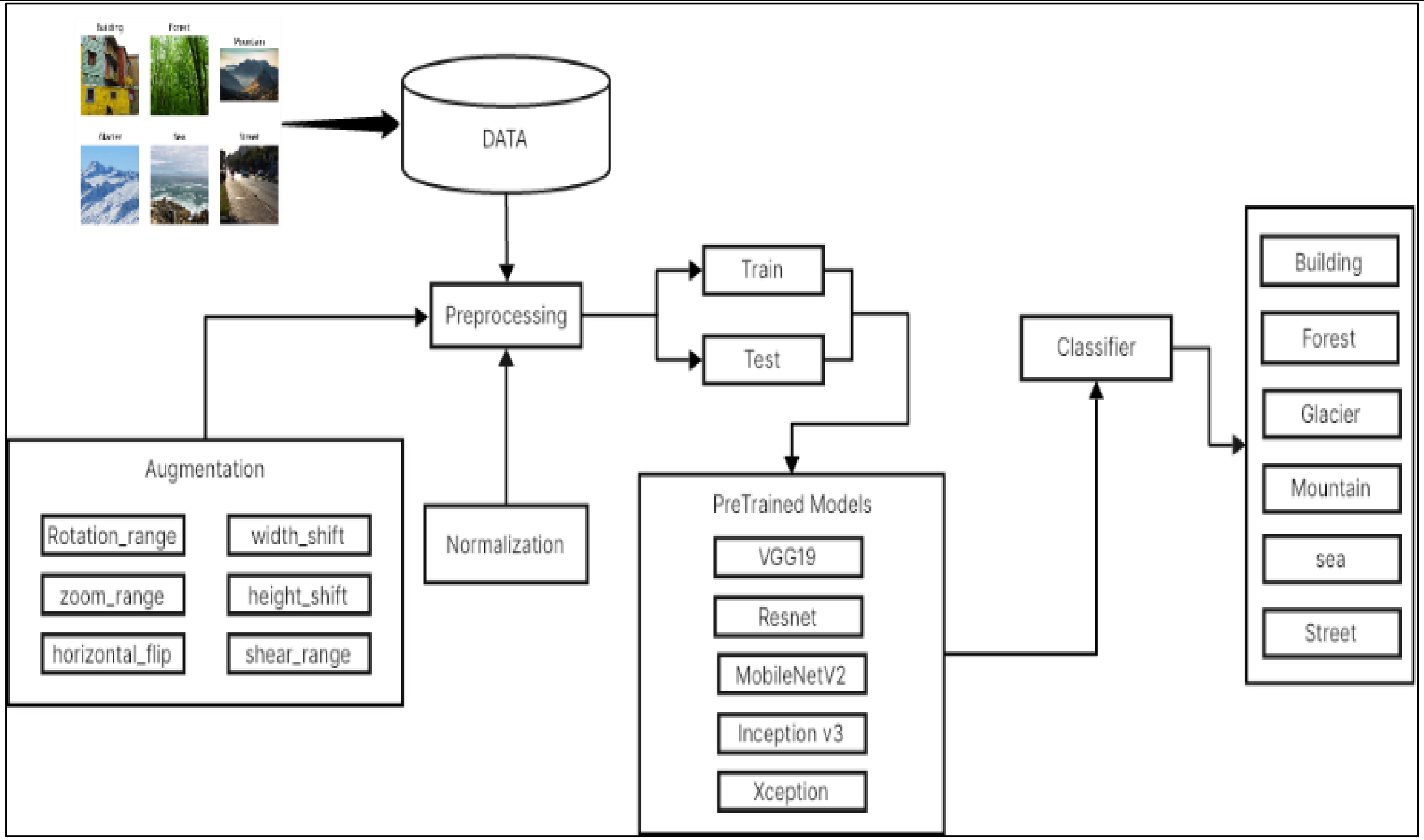

Figure 3. Workflow Diagram of Proposed Methodology

Materials and Methods:

This section provides a comprehensive overview of the system design and methodology as shown in Figure 3. To avoid model bias, we conducted an Exploratory Data Analysis (EDA) before training the model to ensure that the data set was balanced and there were no biases. This helped establish the strategies to employ in model selection and fine-tuning because we desired a fair comparison of the models in all the categories. In addition to the EDA, we employed other pre-trained state-of-the-art deep learning architectures including VGG19, MobileNetV2, ResNet50, InceptionV3, and Xception. These models were trained to improve their feature extraction abilities specifically to our data set. The final step of transfer learning was to replace the classification layers with the number of classes in the Intel dataset and train the models on data preprocessed by us. When comparing the results of the classification of nature scenes, we were able to systematically compare the efficiency of each of the models and provide an overall representation of the architectural performance.

In the evaluation of each model, we used some performance measures including Confusion Matrix, Training Time, Model Accuracy and Validation Accuracy. To further describe the model's behavior, we employed Accuracy and Loss graphs and Receiver Operating Characteristic (ROC) curves for the model's comparison.

Data Preprocessing:

We used data augmentation to improve the model's accuracy and to expand the quantity of training data. This study used the Keras Image Data Generator to apply various operations to the original photos, therefore expanding the Intel classification of pictures dataset. Vertical scaling, shearing, zooming, horizontal flipping, and random rotation up to 40 degrees were used for these operations. Horizontal scaling up to 20% was used. In addition, we made sure that no data was lost during the conversions by employing the "nearest" pixel fill option while creating the augmented photos.

Each of the six types of data including buildings, forests, glaciers, mountains, sea, and streets was enhanced. An increase in the number of images of the set for training and the introduction of variability in the form of 2000 enhanced photos for each category has helped the model understand more specific aspects and traits of the scene classifications. A directory was created to contain these enhanced photographs so they could be used during training. By training the model on several picture variants, this method effectively reduces overfitting and improves its ability to anticipate the results of unknown data.

Data Augmentation Parameter:

Data augmentation strategies are applied to increase the dataset and improve model performance. Using the 'Image Data Generator' class in Kera's, several techniques are used to boost data-set distinctiveness. Augmentation parameters include rotation range which randomly rotates images by up to 30 degrees; width_shift_range and height_shift_range which allow horizontal and vertical shifts up to 20% of the image dimensions; shear_range which applies shear transformations up to 20 degrees and zoom_range which enables zooming in or out by up to 20%.

Further, empty pixels are filled using fill_mode along with nearest pixel values whereas, images are turned horizontally using horizontal_flip. Data augmentation helps the models to generalize better across multiple settings, enhancing their performance and ability to handle a range of visual conditions. Table 2 highlights the typical parameters that are applied during the data augmentation to enhance the diversity of the training set.

Table 2. Augmentation Parameters

Augmentation Technique	Description	Value/Range
Fill_mode	Rotate Angle	0 to 30 degrees
Width_shift_range	Shift horizontally.	Up to 20% of width
Height_shift_range	Shift vertically	Up to 20% of height

Shear_range	Shear angle.	Up to 20 degrees
Zoom_range	Zoom in/out	20% zoom
Horizontal_flip	Flip horizontally	True
Fill Mode	Fill Empty	Nearest

Results and Discussion:

Experiment Setup:

The experiments were performed on a high-performance computer.

- Configuration with the following characteristics:
- **CPU:** AMD RYZEN 9 5900X
- **GPU:** NVIDIA GEFORCE RTX 4080 SUPER 16G
- VENTUS 3X OC
- **Memory:** 32 GB RAM

VGG19:

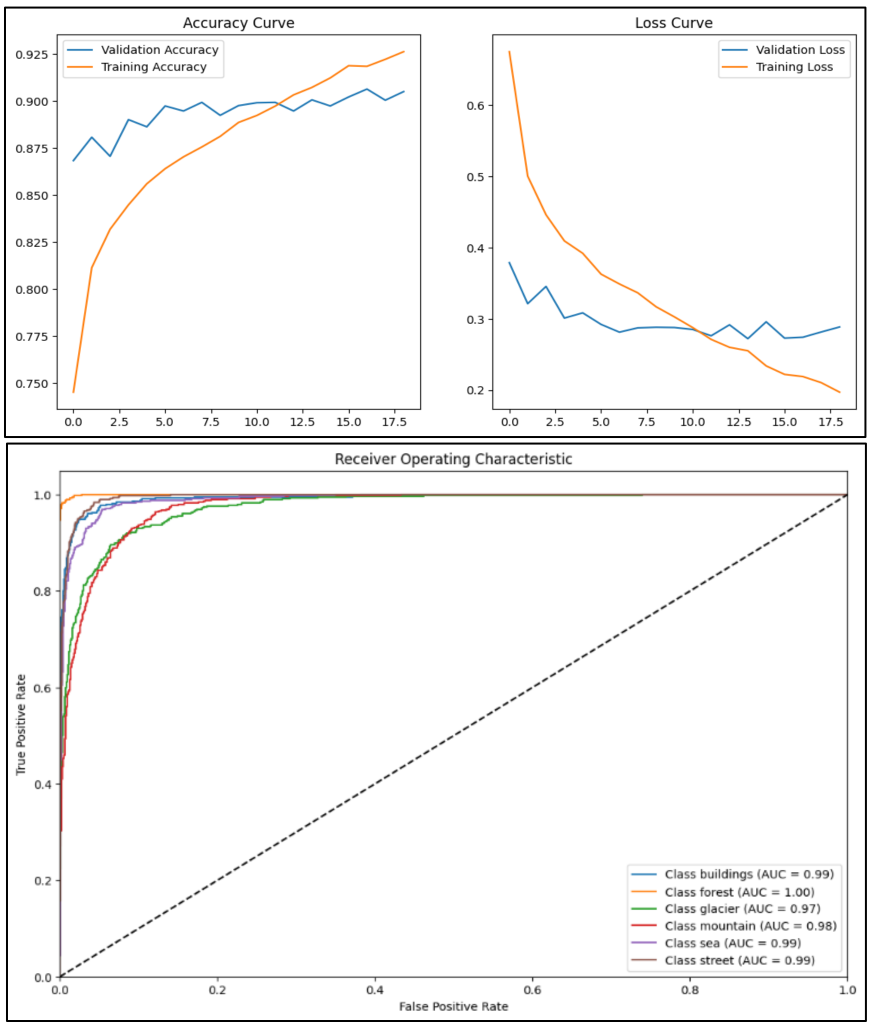

Figure 4. Performance Graphs of VGG19

In current research we utilized the VGG19 model for scene classification in nature, which had previously been trained in the ImageNet dataset. There was a total of 20,926 training photos and 5,228 testing images in the six-class dataset. Horizontal flipping and rescaling were used to enhance the model's ability to generalize to new data. In order to keep the features learnt from ImageNet in the VGG19 model's pre-trained convolutional layers, we frozen them. To avoid overfitting and ensure task flexibility, we initialized and trained the dense layers using dropout.

Adam optimizer was used to train this model for 20 epochs at a learning rate of 1e-4. Model checkpointing served to preserve the optimal model, while validation loss was employed to accomplish early halting. Training accuracy was improved from an initial 74.52% to a final validation accuracy of 86.74%. As of the 20th epoch, the validation error was 90.51% and the training accuracy was 92.64%. The accuracy of 90.05% on the objective test set of 3000 photos further demonstrates the model's excellent generalizability. In order to assess the model's efficacy across all classes, we acquired the classification report, the confusion matrix and ROC curves. These findings showed that the proposed model can effectively differentiate between various nature scenes.

Figure 4 shows the training accuracy and training loss of the VGG19 model through two graphs. The Accuracy Curve depicts a gradual rise in training accuracy as well as validation accuracy to about 92% and 90% respectively. On the other hand, the Loss Curve shows that the training loss and the validation loss has decreased drastically in the initial epochs, and while the training loss kept on declining, the validation loss was almost stagnant and rose slightly towards the end. These results indicate that VGG19 can learn from the provided data; however, the increasing gap between the training and validation metrics suggests a risk of overfitting in the later epochs.

The ROC curve for the multi-class classification model shows excellent performance across six classes, with AUC values ranging from 0.97 to 1.00. Notably, the forest class achieves a perfect AUC of 1.00, indicating outstanding discriminative ability in distinguishing different scene types.

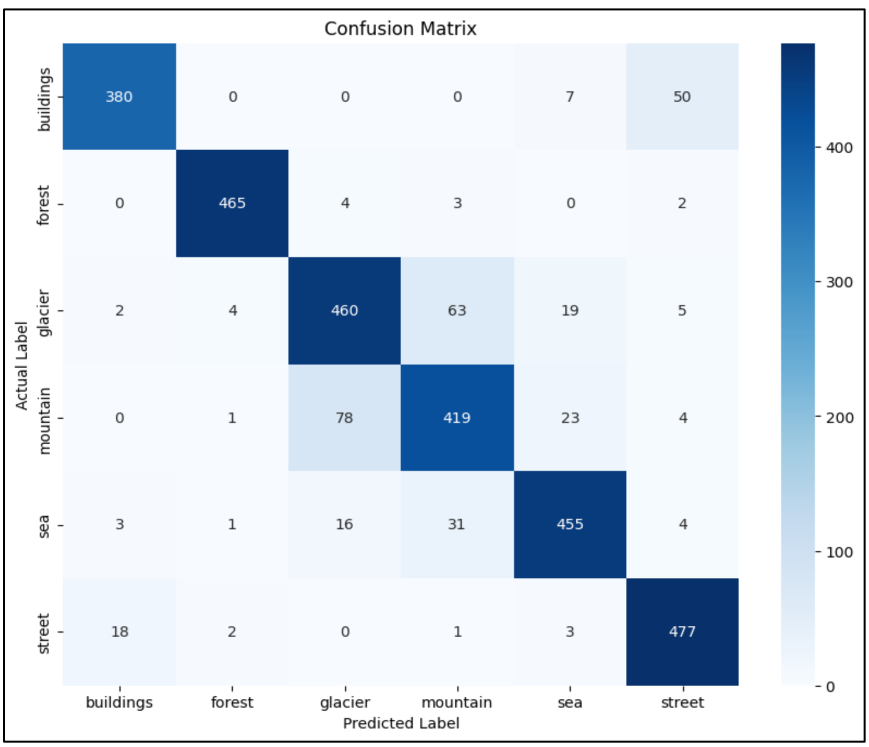

Figure 5. Confusion Matrix of VGG19

The confusion matrix in Figure 5 shows the performance of a classification model for six classes: buildings, forest, glacier, mountain, sea, and street. The diagonal elements represent correct predictions, with high numbers indicating good accuracy for each class. There are some misclassifications, particularly between glacier and mountain classes, and between buildings and street classes, but overall, the model performs well with most predictions falling on the diagonal.

Mobile Net V2: For nature scene classification, the Mobile Net V2 model was used, taking advantage of its small architecture and the efficient transfer learning provided by the pre-trained weights from ImageNet. To prevent overfitting, the design incorporates the following layers: the MobileNetV2 base model, an average around the world layer, a dense layer that has 1024 units of ReLU activation, plus a dropout layer that has a rate of 0.5. Utilizing the Adam optimizer and a learning rate of 1e-4, the model underwent 20 iterations of training. To ensure the greatest performance, it included early halting after 5 iterations and model checkpointing depending on validation loss.

Results showed a training accuracy of 95.54% and a validation accuracy of 92.67% after 18 epochs, demonstrating a strong ability to generalize to unseen data. Evaluation of test data revealed high classification capabilities across scene categories, supported by a classification report with high F1-scores and a confusion matrix showing in Figure 7 minimal class overlap. Receiver Operating Characteristic (ROC) curves demonstrated an AUC greater than 0.9 for each class, underscoring the model's robustness in classification tasks as shown in Figure 6.

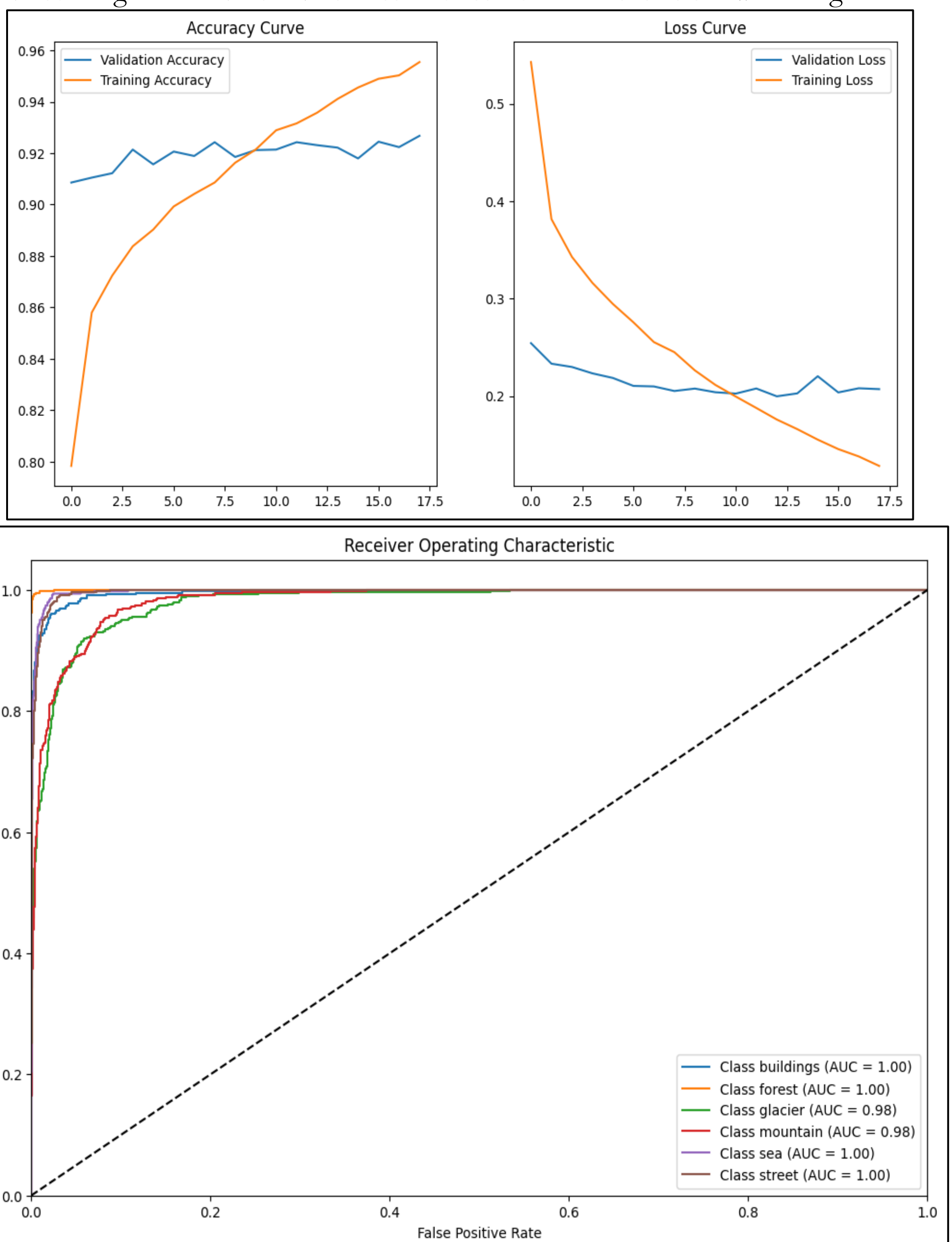

Figure 6. Performance Graphs of MobileNetV2

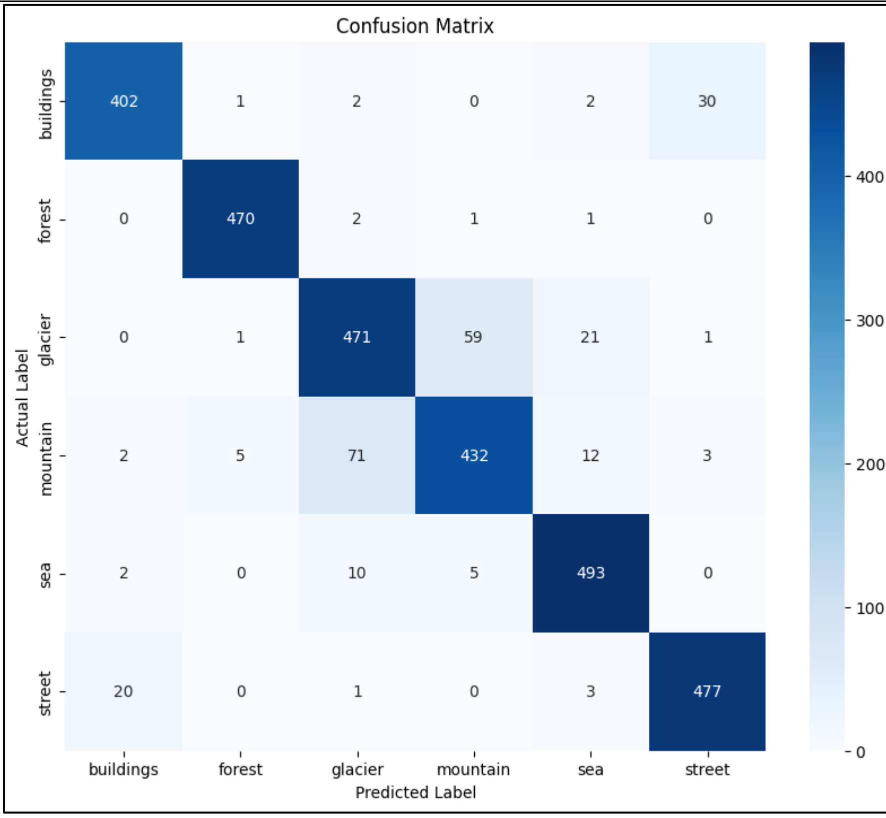

Figure 7. Confusion Matrix of MobileNetV2

ResNet50:

The ResNet50 model was used for classifying the nature scenes and the strong architecture of the model along with weights from ImageNet were used for the purpose of transfer learning. It consists of the ResNet50 base model, the global average pooling layer, the dense layer with 1024 neurons and ReLU activation function, and the dropout layer with the dropout rate of 0.5. The model was compiled using the Adam optimizer with a learning rate of 1e-4 and trained for 20 epochs, incorporating early stopping with patience of 5 epochs and model checkpointing based on validation loss to retain the best-performing model.

Results indicated a training accuracy of 56.98% and a validation accuracy of 65.34% after 20 epochs, demonstrating challenges in fully capturing the complexities of the nature scene classification task. Evaluation on test data highlighted robust classification capabilities across scene categories, as evidenced by a classification report showcasing high F1-scores and a confusion matrix reflecting minimal class overlap. ROC curves indicated an AUC greater than 0.9 for each class, emphasizing the model's effectiveness in classification tasks despite the lower-than-expected overall accuracy. In Figure 8, accuracy curves show an unusual pattern with validation outperforming training, indicating potential issues with data sampling or model fitting. Loss curves decrease over time, but the consistently lower validation loss compared to training loss is atypical and warrants investigation. The ROC curves for the six classes performed above the random baseline, with the forest class achieving the highest AUC of 0.98, while the glacier class had the lowest AUC at 0.86. While these AUC values seem good, the unusual accuracy and loss patterns suggest the model's performance and generalization ability may not be as robust as initially appears.

In Figure 9 The confusion matrix displays the performance of a classification model for six landscape categories, showing correct predictions and misclassifications. It reveals strengths in forest identification, and weaknesses in distinguishing glaciers from mountains, and highlights areas for improvement in the model's ability to differentiate between similar landscape types.

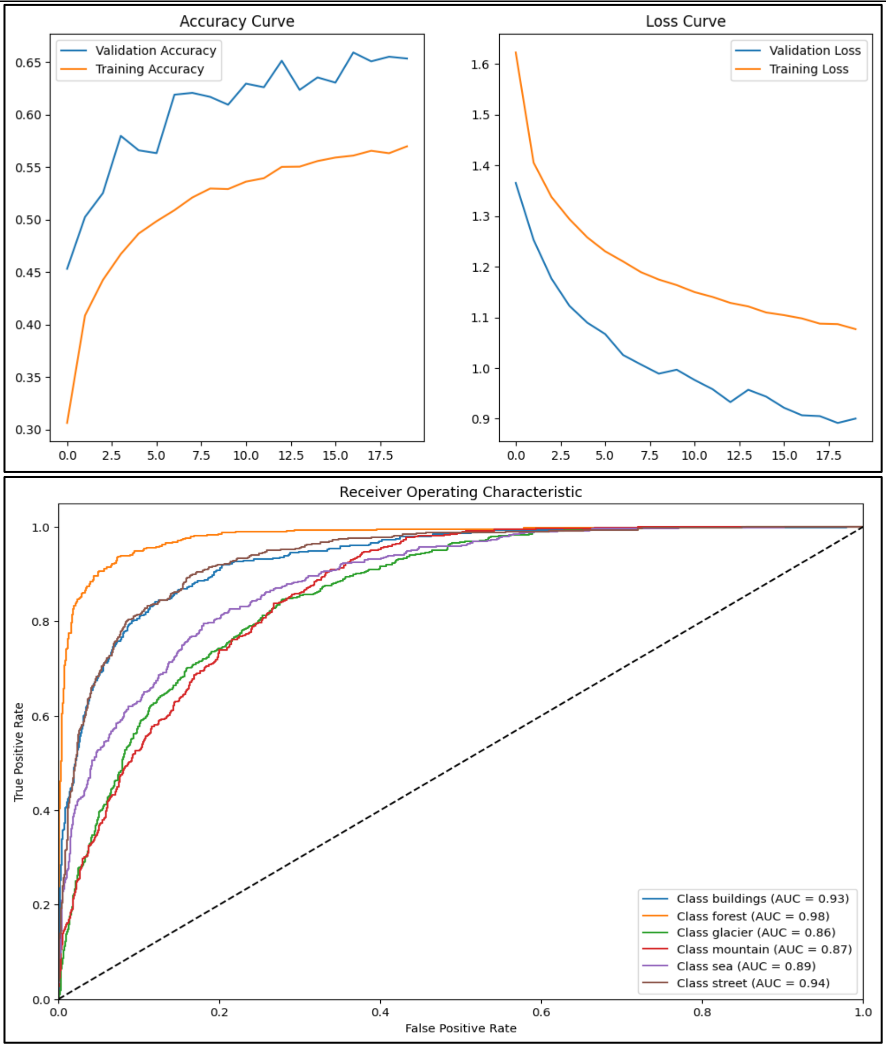

Figure 8. Performance Graphs of ResNet50

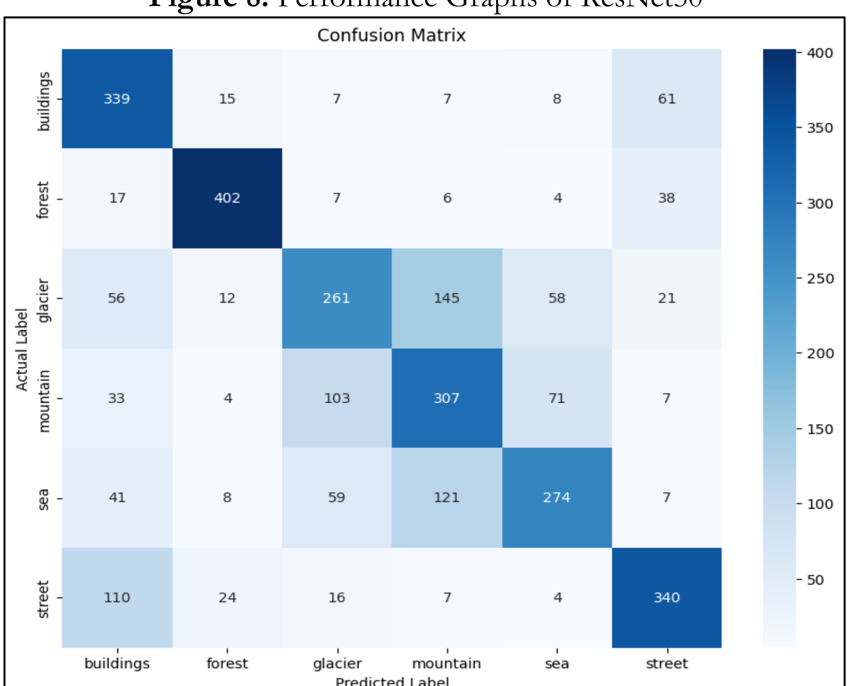

Figure 9. Confusion Matrix of ResNet50

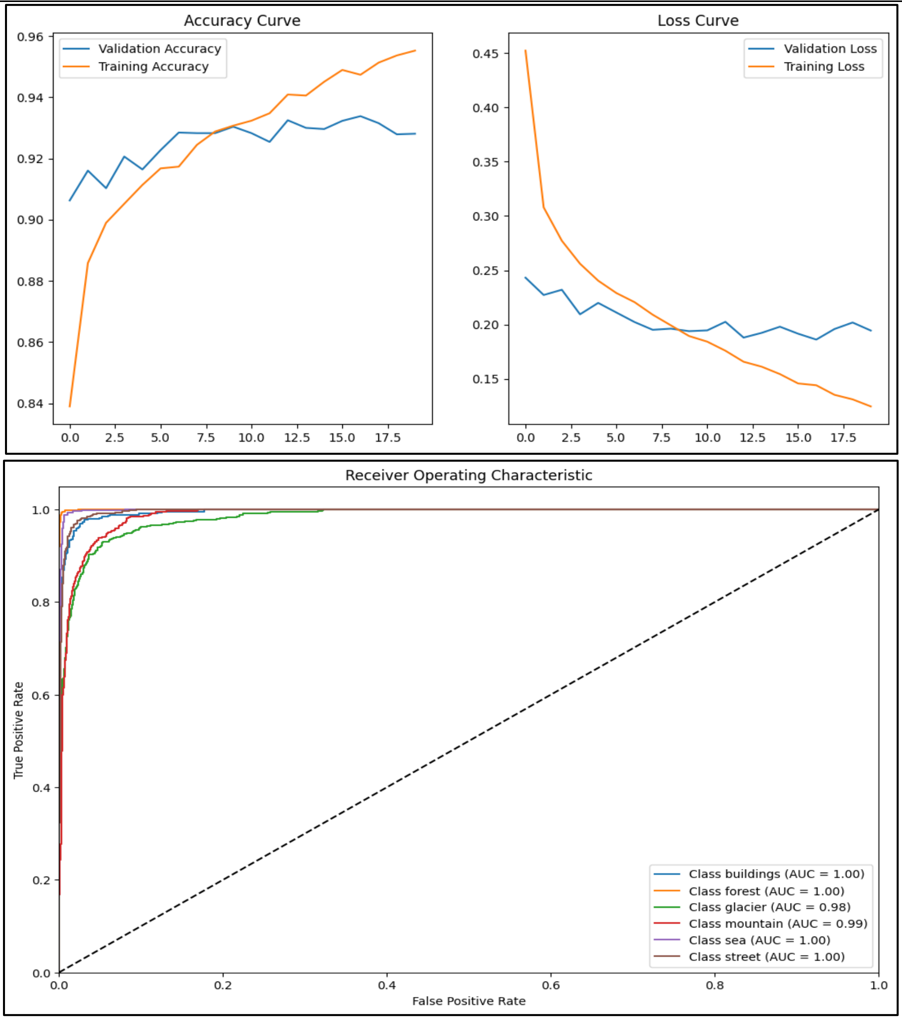

Figure 10. Performance Graphs of Xception

Figure 11 shows the confusion matrix of Xception.

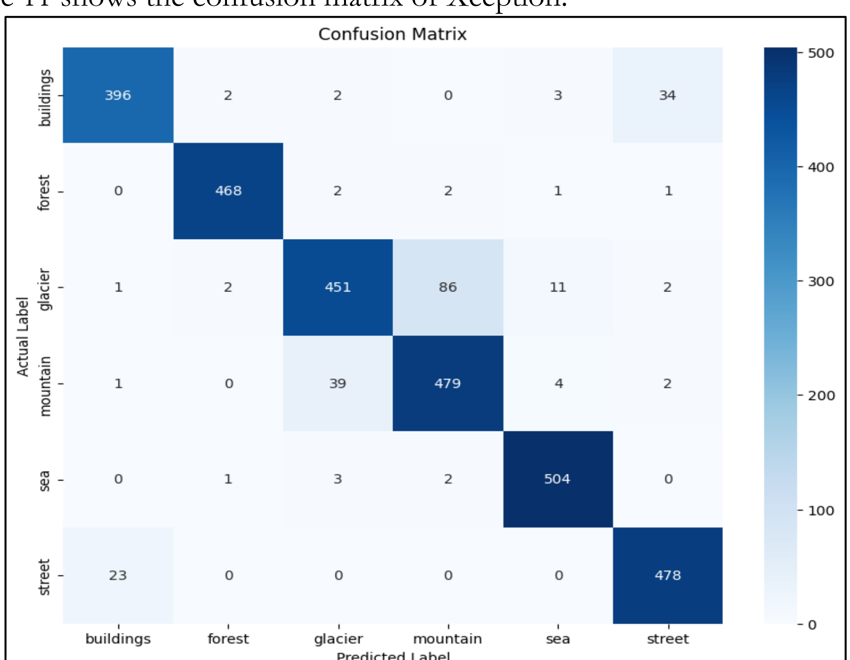

Figure 11. Confusion Matrix of Xception

Xception:

We evaluated the Xception CNN for classifying nature scenes into six categories. The model was trained on a preprocessed dataset with data augmentation, using transfer learning from ImageNet weights. Custom layers were added for our specific task, and the model was trained for 20 epochs with early stopping and checkpointing.

The training accuracy was 95.52% and the validation accuracy was 92.81% which indicated good generalization but there could be overfitting as shown in Figure 10. In the final assessment, a different set of test data was used, reported the accuracy of each class. While the high accuracy is good, the difference between the training and validation performance suggested that the model needs to be fine-tuned to generalize better, perhaps by using better forms of regularization or by using a more diverse set of training data.

InceptionV3:

To categorize the outside sceneries into six distinct kinds, we used the InceptionV3 CNN in this research. Data pre-processing and data augmentation were carried out to make a template more generalizable. We used 20 percent of the training data for internal validation after separating it from the validation data. We used Google Keras in the system model. We utilized a pre-trained model that had been developed with weights from the ImageNet dataset to refine the features we had learned. To make the design more suited for a classification challenge, the top layers were removed and replaced.

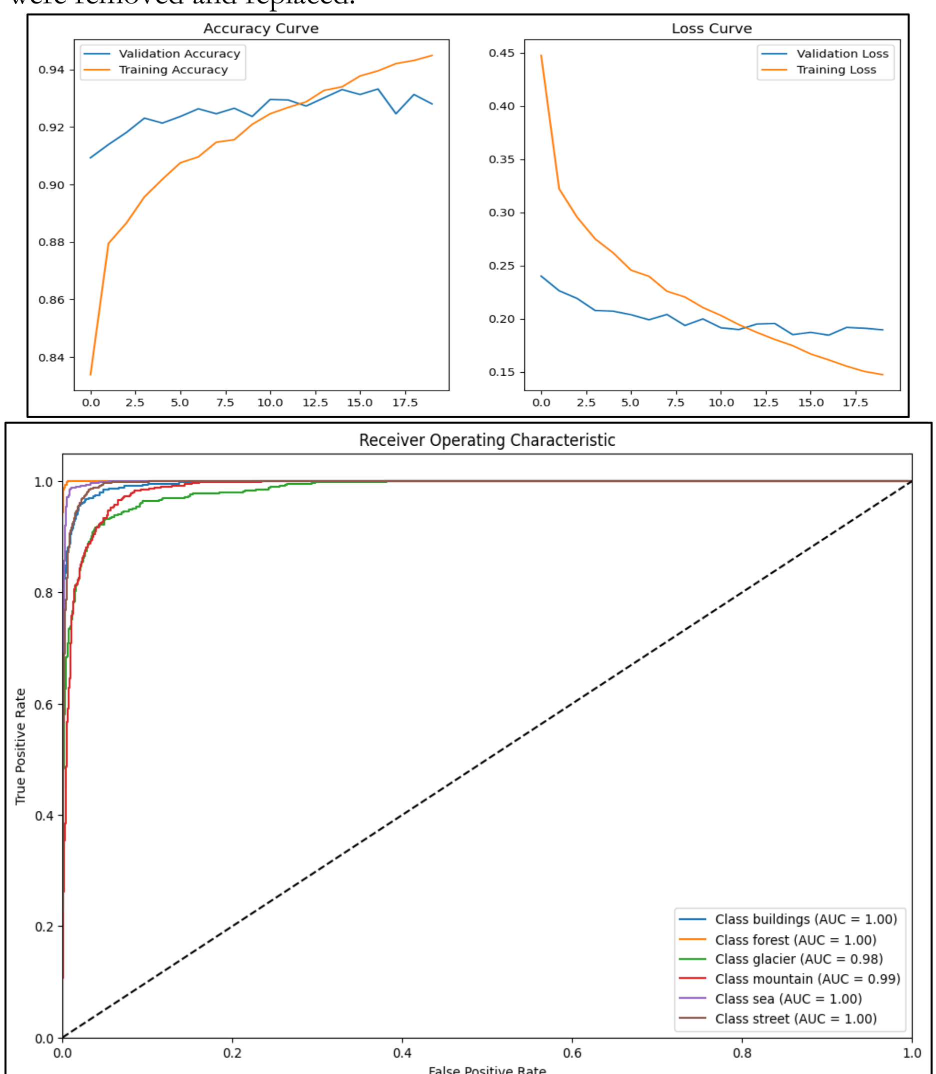

Figure 12. Performance Graphs of InceptionV3

As indicated in Table 3, the following layers were used: an average global pooling layer, a dense one with 1024 neurons that used ReLU induction, a dropout layer employing a rate of 0.5 to prohibit overfitting, and, finally, a final dense layer featuring softmax activation for multiclass classification. The Adam optimizer's hyperparameters were set at 1e-4 for the learning rate and a batch size of 32. We used validation loss to preserve the best model masses after 20 epochs of training, and we used callbacks to end training early.

At the end of the training, the model obtained a training accuracy of 94.49% and a validation accuracy of 92.81% which indicated that the model had learned well; however, the difference between the training and validation accuracy indicated overfitting. The classification report provided the precision, recall, and F1-score for each class and the confusion matrix provided information on how well the model performed for each class. The AUC for each class was also calculated to analyze the performance of the model to an extent further.

Table 3. Summary of Proposed InceptionV3 Model Architecture

Layer Type	Output Shape	Parameters
Input Layer	(299, 299, 3)	-
InceptionV3 (Base Model)	(None, 8, 8, 2048)	22,91,032
Global Average Pooling	(None, 2048)	0
Dense (ReLU)	(None, 1024)	2,099,712
Dropout (0.5)	(None, 1024)	0
Dense (Softmax)	(None, 6)	6,150
Total Parameters		25,032,894

Figure 12 illustrates the performance of the InceptionV3 model, showcasing strong results. The rising training accuracy and falling training loss indicated effective learning, while slight fluctuations in validation accuracy and a small rise in validation loss suggested potential overfitting. The ROC curve highlighted excellent classification performance, with near-perfect AUC scores (around 1.0) for most classes, except for the glacier class (AUC = 0.98), which was still very high. Overall, the InceptionV3 model is well-trained but may benefit from regularization techniques to enhance its generalization on unseen data.

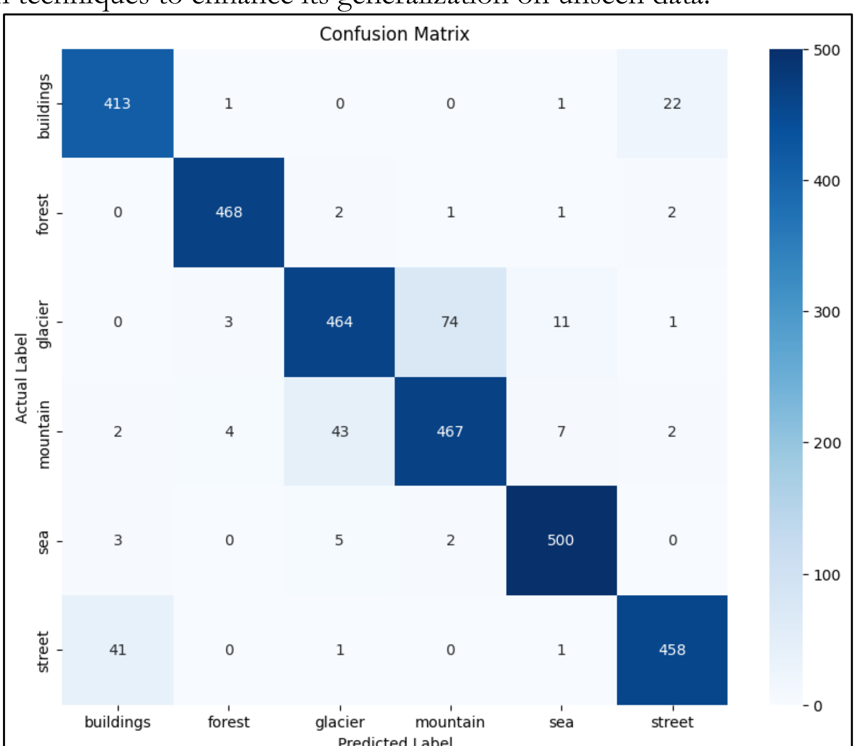

Figure 13. Confusion Matrix of InceptionV3

In Figure 13 confusion matrix shows that the InceptionV3 model performed best on the "sea" category with perfect classification (500 correct predictions), followed closely by strong performance on "forest" (468 correct out of 474) and "street" (458 correct out of 500). The "mountain" class also performed well (467 correct out of 525), but there is some confusion with the "glacier" class, where 43 instances were misclassified as "glacier." The biggest challenge for the model was distinguishing between "glacier" and "mountain," with 74 "glacier" images misclassified as "mountain." Overall, the model demonstrates high accuracy, especially in distinguishing man-made structures and water-based images, but struggles more with nature landscapes, particularly between glaciers and mountains.

Discussion:

In our experiments, different CNN architectures were compared for the nature scene classification task in terms of classification accuracy and generalization through transfer learning from the ImageNet weights. The choice of the VGG19 model for image detection is based on its simple design but relatively deep architecture that yields exceptional feature-extracting abilities and high accuracy rates for image classification. ResNet50 has considerably better performance that results from its residual connections which prevent vanishing gradients. This is because inception V3 maintains parallel convolutional filter features which make it cover multi-scale in efficient and best angle of all scenes. Xception systems successfully recognize complex features in images and use considerably less computational resources than traditional convolutional frameworks. Finally, MobileNetV2 was chosen owing to its faster and lightweight design coupled with reasonable classification accuracy while operating in ideal conditions for constrained environments. Among the models under consideration, InceptionV3 was the most accurate model with a training accuracy of 94.49% and a validation accuracy of 92.81% as indicated in Table 4. These results not only demonstrate that it is capable of generalizing to unseen data but also demonstrate a higher performance measure than previous studies. InceptionV3 architecture can handle the issues of nature scene classification while data augmentation enhances its performance and avoids overfitting. Xception achieved high accuracy, but the shortcomings related to overfitting reveal the flaws of the method. MobileNetV2 was also good with a training accuracy of 95.54% and a validation accuracy of 92.67%, but like Xception, MobilnetV2 overfit more than InceptionV3. On the other hand, ResNet50 has the lowest training and validation accuracy of 56.98% and 65.34% respectively which shows that ResNet50 has problems in learning the details of the classification task. Moreover, InceptionV3 had better accuracy than other models and the split between the training and validation set was better, so it is the best model to classify the nature scenes because of the better generalization and learning.

Table 4. Comparison of all the Results

Model	Training Accuracy	Validation Accuracy
InceptionV3	94.49%	92.81%
VGG19	92.64%	90.51%
MobileNetV2	95.54%	92.67%
Xception	95.52%	92.81%
ResNet50	56.98%	65.34%

Table 4 shows that our results are better than the previous study [18] particularly when using InceptionV3 and Xception. The training accuracy of the InceptionV3 model was 94.49% and validation accuracy was 92.81% which is higher than the training accuracy of 93.95% and validation accuracy of 91.81% of the previous study. Likewise, the Xception model yielded a training accuracy of 95.52 % and a validation accuracy of 92.81%, which is higher than the previous work done with 92.37% training accuracy and 91.20% validation accuracy. These improvements support the effectiveness of the proposed method in improving the nature scene classification.

Table 5. Comparison with previous study

	Our Results		Previous Study [18]	
Model	Training Accuracy	Validation Accuracy	Training Accuracy	Validation Accuracy
InceptionV3	94.49%	92.81%	93.95	91.81
Xception	95.52%	92.81%	92.37	91.20

Conclusion and Future Work:

In the present study, we compared various structures of CNN for nature scene categorization and found that InceptionV3 was the best model with a training accuracy of 94.49% and validation accuracy of 92.81% which is higher than the previous work. It was able to address classification issues, complemented by data augmentation methods and was seen to perform well when tested on new data. Other models such as Xception and MobileNetV2 also performed well in terms of accuracy, but overfitting was also observed implying that choosing the right model is crucial to avoid overcomplicated models. In the future, to regulate the large weight values we plan to continue using L2 regulation, and to expand the training data, various methods of data expansion will be used to improve the model's ability to generalize.

Acknowledgment: On behalf of the research team, the authors would like to express the deepest gratitude to all the academic institutions, mentors, and peers who provided their invaluable input during this research.

Author's Contribution: All authors contributed equally to the conceptualization, design, implementation, and writing of this manuscript.

Conflict of Interest. The authors declare no conflict of interest regarding the publication of this manuscript in IJIST.

References:

[1] Z. Dong, M. Wang, Y. Wang, Y. Zhu, and Z. Zhang, "Object Detection in High Resolution Remote Sensing Imagery Based on Convolutional Neural Networks with Suitable Object Scale Features," IEEE Trans. Geosci. Remote Sens., vol. 58, no. 3, pp. 2104–2114, Mar. 2020, doi: 10.1109/TGRS.2019.2953119.

[2] F. Bastani et al., "RoadTracer: Automatic Extraction of Road Networks from Aerial Images," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 4720–4728, Dec. 2018, doi: 10.1109/CVPR.2018.00496.

[3] S. D. Khan, L. Alarabi, and S. Basalamah, "Deep Hybrid Network for Land Cover Semantic Segmentation in High-Spatial Resolution Satellite Images," Inf. 2021, Vol. 12, Page 230, vol. 12, no. 6, p. 230, May 2021, doi: 10.3390/INFO12060230.

[4] S. Talukdar et al., "Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A Review," Remote Sens. 2020, Vol. 12, Page 1135, vol. 12, no. 7, p. 1135, Apr. 2020, doi: 10.3390/RS12071135.

[5] S. D. Khan, L. Alarabi, and S. Basalamah, "An Encoder–Decoder Deep Learning Framework for Building Footprints Extraction from Aerial Imagery," Arab. J. Sci. Eng., vol. 48, no. 2, pp. 1273–1284, Feb. 2023, doi: 10.1007/S13369-022-06768-8/METRICS.

[6] M. T. Chiu et al., "Agriculture-vision: A large aerial image database for agricultural pattern analysis," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 2825–2835, 2020, doi: 10.1109/CVPR42600.2020.00290.

[7] S. D. Khan, L. Alarabi, and S. Basalamah, "A Unified Deep Learning Framework of Multi-scale Detectors for Geo-spatial Object Detection in High-Resolution Satellite Images," Arab. J. Sci. Eng., vol. 47, no. 8, pp. 9489–9504, Aug. 2022, doi: 10.1007/S13369-021-06288-X/METRICS.

[8] L. Lin, L. Di, C. Zhang, L. Guo, and Y. Di, "Remote Sensing of Urban Poverty and Gentrification," Remote Sens. 2021, Vol. 13, Page 4022, vol. 13, no. 20, p. 4022, Oct. 2021, doi: 10.3390/RS13204022.

[9] A. Kazemzadeh-Zow, A. Darvishi Boloorani, N. N. Samany, A. Toomanian, and A. Pourahmad, "Spatiotemporal modelling of urban quality of life (UQoL) using satellite images and GIS," Int. J. Remote Sens., vol. 39, no. 19, pp. 6095–6116, Oct. 2018, doi: 10.1080/01431161.2018.1447160.
[10] M. A. A. Hoque, S. Phinn, C. Roelfsema, and I. Childs, "Tropical cyclone disaster management using remote sensing and spatial analysis: A review," Int. J. Disaster Risk Reduct., vol. 22, pp. 345–354, Jun. 2017, doi: 10.1016/J.IJDRR.2017.02.008.
[11] B. Zhao, Q. Dai, L. Zhuo, S. Zhu, Q. Shen, and D. Han, "Assessing the potential of different satellite soil moisture products in landslide hazard assessment," Remote Sens. Environ., vol. 264, p. 112583, Oct. 2021, doi: 10.1016/J.RSE.2021.112583.
[12] N. J. Murray et al., "The role of satellite remote sensing in structured ecosystem risk assessments," Sci. Total Environ., vol. 619–620, pp. 249–257, Apr. 2018, doi: 10.1016/J.SCITOTENV.2017.11.034.
[13] S. Mahato and S. Pal, "Groundwater Potential Mapping in a Rural River Basin by Union (OR) and Intersection (AND) of Four Multi-criteria Decision-Making Models," Nat. Resour. Res., vol. 28, no. 2, pp. 523–545, Apr. 2019, doi: 10.1007/S11053-018-9404-5/METRICS.
[14] P. L. Raeva, J. Šedina, and A. Dlesk, "Monitoring of crop fields using multispectral and thermal imagery from UAV," Eur. J. Remote Sens., vol. 52, no. sup1, pp. 192–201, Mar. 2019, doi: 10.1080/22797254.2018.1527661.
[15] B. Petrovska, T. Atanasova-Pacemska, R. Corizzo, P. Mignone, P. Lameski, and E. Zdravevski, "Aerial Scene Classification through Fine-Tuning with Adaptive Learning Rates and Label Smoothing," Appl. Sci. 2020, Vol. 10, Page 5792, vol. 10, no. 17, p. 5792, Aug. 2020, doi: 10.3390/APP10175792.
[16] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017, vol. 2017-January, pp. 1800–1807, Nov. 2017, doi: 10.1109/CVPR.2017.195.
[17] M. Soudy, Y. Afify, and N. Badr, "RepConv: A novel architecture for image scene classification on Intel scenes dataset," Int. J. Intell. Comput. Inf. Sci., vol. 22, no. 2, pp. 63–73, May 2022, doi: 10.21608/IJICIS.2022.118834.1163.
[18] X. Wu, R. Liu, H. Yang, and Z. Chen, "An Xception Based Convolutional Neural Network for Scene Image Classification with Transfer Learning," Proc. - 2020 2nd Int. Conf. Inf. Technol. Comput. Appl. ITCA 2020, pp. 262–267, Dec. 2020, doi: 10.1109/ITCA52113.2020.00063.
[19] "Intel Image Classification." Accessed: Oct. 01, 2024. [Online]. Available: https://www.kaggle.com/datasets/puneet6060/intel-image-classification