Classifying Land Cover of Ireland From Satellite Imagery using Deep Learning and Transfer Learning

MSc Research Project
MSc in Data Analytics

Prateek Mata
Student ID: x21125279

School of Computing
National College of Ireland

Supervisor: Dr. Catherine Mulwa

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Prateek Mata
Student ID:	x21125279
Programme:	MSc in Data Analytics
Year:	2022
Module:	MSc Research Project
Supervisor:	Dr. Catherine Mulwa
Submission Due Date:	15/12/2022
Project Title:	Classifying Land Cover of Ireland From Satellite Imagery using Deep Learning and Transfer Learning
Word Count:	7151
Page Count:	25

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Prateek Mata
Date:	1st February 2023

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Classifying Land Cover of Ireland From Satellite Imagery using Deep Learning and Transfer Learning

Prateek Mata
x21125279

Abstract

The use of satellite images for land cover classification has become increasingly important for governments like Ireland, where the population is increasing rapidly, to monitor changes in the environment and support urban development, disaster management, and sustainable development. However, implementing land cover and land use classification can be challenging due to the absence of a space program and the lack of labelled training data, which is expensive to generate. This research project implements two deep learning models trained on a novel land use and land cover classification dataset using transfer learning to classify a new dataset of land cover images of Ireland created using Sentinel-2A satellite images, out of which the ResNet-50 achieved a 73% accuracy. In addition, it also shows how data augmentation helps improve the accuracy of a land cover classification model where less labelled training data is available.A novel classification model was created based on the identified gaps in the results of the analyzed literature review, which is also presented.

1 Introduction

The features of the earth's surface, such as forests, water bodies, farmlands and developed areas, are constantly changing due to natural and human activities. It is, therefore, essential for the government of a country like Ireland to monitor these changes, which will support them in monitoring forests and natural disasters, support land use planning and agriculture, and help make better policies for sustainable development. The technology used for remote sensing is advancing rapidly; thanks to satellites, unmanned aerial vehicles (UAVs), and drones, a large amount of data about the surface of the Earth is now accessible. These satellite images have recently emerged as the most helpful instrument for classifying different types of land use and land cover (LULC)(Hütt et al.; 2016). Deep learning (DL) models have evolved and become one of the most commonly accepted emerging technologies as a powerful solution to approach a variety of machine learning (ML) challenges, such as satellite image classification, and can manage the expanding amount of earth observation data. This research focuses on using deep learning models to classify the land cover of unseen satellite images of Ireland using transfer learning from a novel land cover dataset.

Background and Motivation

The majority of Ireland's constantly growing population lives in major urban cities like Dublin, Cork, and Galway. To better plan for urban development, including land allocation, new transport routes, agricultural planning to meet the demands of the recent increase in population, and addressing the housing crisis that has been a significant issue in Ireland for the past couple of years, the government must monitor changes in the land cover in real-time as well as historical changes. The government's responsibility does not end with urban development; it also includes disaster management policy-making, planning for geological activities, and sustainable development. It will be challenging for a nation that does not have a space program to monitor these changes without expending a significant amount of effort and resources. Deep learning, more specifically, convolution neural networks (CNN), has become quite popular among researchers to solve image classification issues. 2010 was the year that Volodymyr and Hinton (2010) first utilised CNN to extract buildings and road networks from remote sensing data.

Deep learning models need a significant amount of data; however, labelled training data is exceptionally scarce, and the process of generating labelled data for each new recognition task is expensive, thus making it challenging to implement Convolutional Neural Networks (CNN) for satellite image classification tasks. To combat the problem of inadequate labelled training data, researchers have recommended using transfer learning, which employs existing pre-trained models built on natural image data sets like ImageNet and learns solely on the network's top, fully connected layers. Using transfer learning in conjunction with Data Augmentation to enlarge the training dataset artificially is the approach discussed further in the report.

1.1 Research Question and Objectives

Ireland's growing population creates a problem for the government to improve urban development and address issues like the housing crisis. Monitoring changes in land cover is important for planning and disaster management. However, the lack of a space program and scarce labelled training data for deep learning models make it difficult and expensive to do so.

RQ *" To what extent can deep learning models(ResNet-50 and VGG16) trained using a novel land cover and land use dataset and transfer learning help the Irish government classify Ireland's land cover into different classes from out-of-domain satellite images of Ireland from the European space agency's Sentinel 2A mission?"*

Sub RQ *"Can data augmentation techniques applied on the training datasets help improve the accuracy of the deep learning model where less labelled training data is available?"*

To solve these research questions, the research objectives mentioned Table1 were laid out and implemented.

Table 1: Research objectives

Objectives	Description
1	A critical review of literature on Land use and Land cover classification(2012- 2022)
2	Dataset creation and processing
2.1	Programmatically reclassify EuroSAT into desired classes
2.2	Image augmentation on EuroSAT to prevent class imbalance
2.3	Creating Irish dataset by generating image patches of Ireland's land cover using satellite images
3	Implementation, evaluation and results for classification of land cover of Ireland
3.1	Implementation, evaluation and results of ResNet-50
3.1(a)	Implementation of ResNet-50 using both datasets(augmented and non-augmented)
3.1(b)	Evaluation and results of ResNet-50 without augmentation
3.1(c)	Evaluation and results of ResNet-50 with augmentation
3.2	Implementation, evaluation and results of VGG-16
3.2(a)	Implementation of VGG-16 using both datasets(augmented and non-augmented)
3.2(b)	Evaluation and results of VGG-16 without augmentation
3.2(c)	Evaluation and results of VGG-16 with augmentation
4	Comparison of developed models

1.2 Project Contributions

This project's major contribution is to train and compare two different pre-trained deep learning models, ResNet-50 and VGG16, using transfer learning on a novel land cover dataset that can be efficiently utilised to classify Ireland's land cover. This study aims to assist the Irish government in monitoring its land cover and land use. The developed models will identify and classify land cover into several classes, which will aid in sustainable development and efficient urban planning.

The minor contribution of this project is to see how Data augmentation helps to increase classification accuracy where a significant amount of labelled training data is not present.

The rest of the technical report is structured as- Section 2 is a review of the related work in the domain of land use and land cover classification from 2012 to 2022. Section 3 represents the scientific methodology used for land cover classification using deep learning techniques; further, Section 4 describes the implementation, evaluation and results for the two deep learning models. Finally, section 5 concludes the study and proposes future work.

2 Literature Review on Land Use and Land Cover Classification of Satellite Images(2012-2022)

2.1 Introduction

This literature review examines the field of land use and land cover classification and helps determine how the accuracy can be improved even when there is a limited amount of training data that has been labelled. The following subsections make up this section's organisational structure- (2.1) Investigation of Traditional Methods for Land Use and Land Cover Classification (2.2) Critique of Deep Learning in Land Use and Land Cover Classification (2.2.1) Comparison of Pre-trained Deep Learning Models in Land Use and Land Cover Classification (2.3) Critical Review of Techniques for Less Training

Data (2.3.1) Review of Transfer Learning for Small Datasets (2.3.2) Critique of Data Augmentation to Increase Data Size (2.4) Conclusion and Identified Gaps.

2.2 Investigation of Methods for Land Use and Land Cover Classification

The 1970s saw the beginning of the process of land cover classification, with visual analysis serving as the initial form. Within the past 50 years, the field of remote sensing has undergone significant changes, including the introduction of many space programs by agencies such as the European Space Agency (ESA) and NASA, and this is reflected in the field of land cover classification.Ganasri and Dwarakish (2015) used the maximum likelihood classification(MLC) technique, which bases its decisions on the probability that a pixel belongs to a particular information class. Even though MLC produced an acceptable accuracy of 89.36%, it incorrectly categorised wasteland and fallow land as urban areas and did not accurately portray the spatial distribution of the urban regions.Acharya et al. (2016) used a minimum distance classifier which sorts pixels into several classes according to their relative positions the researcher compared minimum distance classifier and MLC using KOMPSAT 3A images in four different bands. Both perform extremely poorly in classification; they are more favoured in the classification of vegetation and built-up areas. In order to classify the same Landsat thematic mapper images acquired over Guangzhou City, Li et al. (2014) compared maximum likelihood classification, logistic regression, random forest and support vector machines algorithms. With well-tuned parameters and sufficiently representative training samples, the algorithms were capable of producing high classification accuracies. Maximum likelihood classification, logistic regression and Support vector machine all had lower accuracy when the size of the training set was small, but random forest wasn't affected by this and did better than all the others. When using traditional machine learning algorithms to achieve such high and consistent accuracy for classification on remote sensing imagery, the pixels per class should be high concluded Noi and Kappas (2017), which may not be the case when dealing with low to medium-resolution imagery.

2.3 Critique of Deep Learning in Land Use and Land Cover Classification

Carranza-García et al. (2019) suggested a general deep learning architecture for performing land use and land cover classification over remote sensing pictures from various sources, specifically radar and hyperspectral datasets. Convolutional neural networks were contrasted with machine learning methods such as random forest, support vector machine and k-nearest neighbors. Despite the diversity of the studied images' origins and characteristics, the CNN model outperformed the other ML models and showed promising results, demonstrating that deep learning is a powerful solution to the LULC classification problem. In a study using SPOT6 satellite data from a tropical island, Rousset et al. (2021) compared traditional machine learning methods to those of Deep Learning, concluding that the latter yielded superior results. One more deep learning model, SegNet, has shown itself to be a practical and quick option for LULC tasks (Sathyanarayanan et al.; 2020). Using Sentinel-2A multispectral data from a single observation, they found that deep learning, a rapidly developing field, holds promise for simplifying and scaling up the analysis of satellite images and demonstrating the potential for automating spectral data analysis in whole or in part.

Convolutional neural networks (CNNs) have shown remarkable performance in image classification when deep learning methods are taken into account. The results of the (activedeep) study demonstrated that CNNs have the potential to be utilised for LULC classification. Ma et al. (2019) looked at how to make better use of existing CNNs for the LULC classification. Two remote sensing datasets were used to evaluate full-trained, fine-tuned, and pre-trained (AlexNet and GoogLeNet) CNNs. Both AlexNet and GoogLeNet performed exceptionally well, and the results suggest that fine-tuning based on pre-trained CNNs is typically superior to using fully-trained CNNs. AlexNet and GoogLeNet showed an accuracy of 95.72 and 96.83 on the UC Merced land use dataset, with the latter performing better and the fine-tuned models outperformed fully trained models by 3% for AlexNet and 9% for GoogleNet. Thus proving that pre-trained deep learning models have great potential for land cover classification.

2.3.1 Comparison of Pre-trained Deep Learning Models in Land Use and Land Cover Classification

Table 2: Comparison of Pre-Trained Deep Learning Models for LULC

Author	Model	Dataset	Accuracy
(Naushad et al.; 2021)	ResNet-50	EuroSAT	99.04%
	VGG-16		98.14%
(Zhang et al.; 2020)	ResNet-50	UC Merced	95.95%
	VGG-16		92.50%
	Inception V4		91.73%
(Alem and Kumar; 2022)	ResNet-50	UC Merced	92.46%
	Inception V3		94.36%
	VGG-19		99.64%
(Mahamunkar and Netak; 2022)	ResNet-50	EuroSAT	96.57%
	VGG-16		97.68%
	VGG-19		95.72%

Alem and Kumar (2022) aimed to enhance LULC classification performance in remote sensing images by employing the transfer learning model. The training speed and accuracy of the model were enhanced by using the bottleneck feature extraction method on pre-trained models. The performance of the model is also apparent in all models, with accuracy results of 92.46% for Resnet50V2, 94.36% for Inception-V3, and 99.64% for VGG-19, respectively. Some classes, such as medium residential, dense residential, and golf courses, have low accuracy, whereas the remainder of the classes performs admirably. Naushad et al. (2021) looked into two different possible deep learning architectures, namely VGG16 and Wide ResNet-50, and fine-tuned them with the RGB bands of the EuroSAT dataset. The proposed methodology not only advanced the previous state of the art but also established a standard achieving an accuracy of 99.17% for the multispectral red-green-blue bands in the EuroSAT dataset.Mahamunkar and Netak (2022) made use of three pre-trained machine learning models with ResNet, Random Forest, and VGG classifiers. These pre-trained models are evaluated with EuroSAT. The VGG16 model surpassed the other models with a classification accuracy of 97.68%, and the ResNet50 model provided a 96.43% classification accuracy. Table Table 2 summarises and compares the accuracies achieved by the pre-trained models in these studies.

2.4 Critical Review of Techniques for Less Training Data

CNN, despite having a structure with numerous layers and being capable of feature extraction, is challenging to train with tiny datasets(Hu et al.; 2015). It requires a considerable amount of data, or it runs the risk of becoming highly specialised. In order to attain sufficient model accuracy, it needs to learn features from a large number of training data. However, the process of producing labelled data for each new recognition task is expensive.

2.4.1 Review of Transfer Learning for Small Datasets

The features learned by the layers across diverse datasets exhibit consistent characteristics, as shown by Xiangnan et al. (2017)Yosinski et al. (2014). As the model progresses through its training, the convolution operators at its base layer pick up on the broad features, and those at its upper layers make the shift to features that are more particular to the dataset used for training. They compared training a CNN to human beings. People with more background knowledge can pick up new skills more quickly than those with less. By utilising three different remote-sensing datasets and two different pre-trained models, de Lima and Marfurt (2020) was able to demonstrate that transfer learning is an effective method for remote-sensing scene classification. Their findings demonstrate that, despite the obvious differences between the two image sets, knowledge may be transferred from natural photos to remote sensing imagery. Alem and Kumar (2022) tackled LULC classification in remote sensing photos by use of transfer learning on deep learning models with bottleneck feature extraction. In contrast to other deep CNN models, which can take days to train from scratch, transfer learning's training time is far more efficient and can be completed in minutes. Their strategy enhanced the model's training speed and precision and yielded notable results overall.

2.4.2 Critique of Data Augmentation to Increase Data Size

A well-known method in deep and machine learning is data augmentation(Carranza-García et al.; 2019); it involves making tweaks to existing pictures in a training dataset in order to increase the total number of images available for usage in the dataset. CNNs' performance in the classification of remote sensing scenes was enhanced by Yu et al. (2017) usage of data augmentation. Data augmentation can considerably improve the diversity and completeness of data. When used to train deep learning models, the experimental results of the model which had augmentation operations applied to it surpassed the model architecture trained on the dataset with no augmentation. Remote sensing photos were subjected to random flipping, rotations, and translations by Stivaktakis et al. (2019). These procedures were chosen since they did not add to the satellite images' spectral or topological information, both of which are necessary for a reliable classification result.

2.5 Conclusion and Identified Gaps

Evidence for the need to create a model for satellite image classification using pre-trained models learned via transfer learning and trained on a different set of labelled satellite images is evident from the reviewed literature work, which reveals both gaps in the existing literature and clear evidence for the necessity of doing so and answers research question1.1. In addition, the evaluated literature provides a foundation for answering the sub-research question 1.1 related to data augmentation in remote sensing data and fulfills the research objective 1.

3 Scientific Methodology Approach and Project Design

3.1 Introduction

This chapter describes the scientific methodology and architectural design used to develop this project. CRISP-DM(Wirth and Hipp; 2000), which has been utilised in numerous pieces of research in this field, is the primary approach employed and modified to create a methodlogy for land cover and land use classification of satellite images of Ireland

3.2 Land Cover Classification of Ireland Methodology

This project's methodology is based on the Cross-Industry Standard Process for Data Mining (CRISP-DM), which provides a simple model for analysing the data in data mining projects. CRISP-DM is deemed suitable for this project because it includes the business aspect required for a comprehensive understanding of the project. Figure 1 describes the modified CRISP-DM method used to classify Ireland's land cover, followed by a thorough explanation of the method.

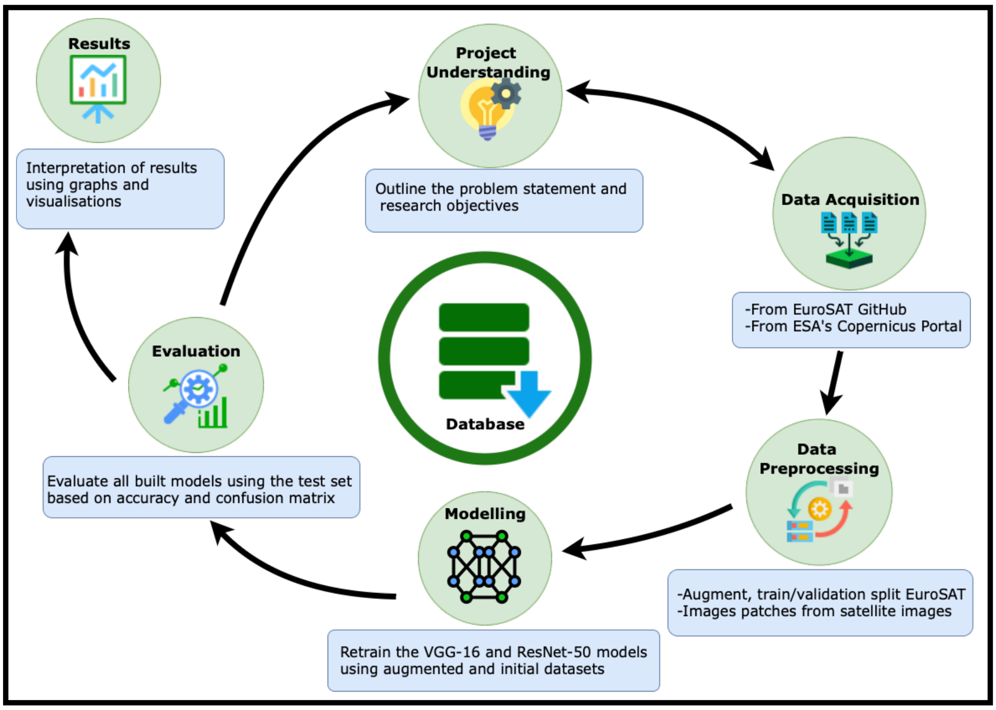

Figure 1: Land Cover Classification of Ireland Methodology

Project Understanding: The research project understanding stages examined the identified problem and outlined the requirements from a business perspective. The project objectives and problem statement were then created using this list of requirements.

Data Acquisition: This research project used two data sources. The first dataset was sourced from the official GitHub repository of EuroSAT, and the second dataset was created using Sentinel2A satellite images available publicly on ESA's online portal. Python and QGIS, a geographic information system, were used in this phase.

Data Acquisition: The acquired data was prepared for use in training and testing the models during this phase. Basic image augmentation was used on the EuroSAT and was divided into training and validation sets. The Ireland image patches were divided into different land cover classes and resized to 64x64 pixels for the model.

Modelling: Using the non-augmented and augmented images of EuroSAT, the models VGG-16 and ResNet-50 were developed as part of this project. Using the training and validation sets produced in the earlier stage, these were trained and validated. The performance of these retrained models was then evaluated using the test set made for the Ireland patches.

Evaluation: In this step, the effectiveness of developed models was assessed using metrics like accuracy and classification matrix. To determine whether the models could accomplish the specified business objectives, they were extensively examined and analysed.

Results: The project's final phase concentrated on the interpretation of the gained knowledge and findings, which were plotted on graphs and evaluated to see if they met the project's objectives.

3.3 Design Specification

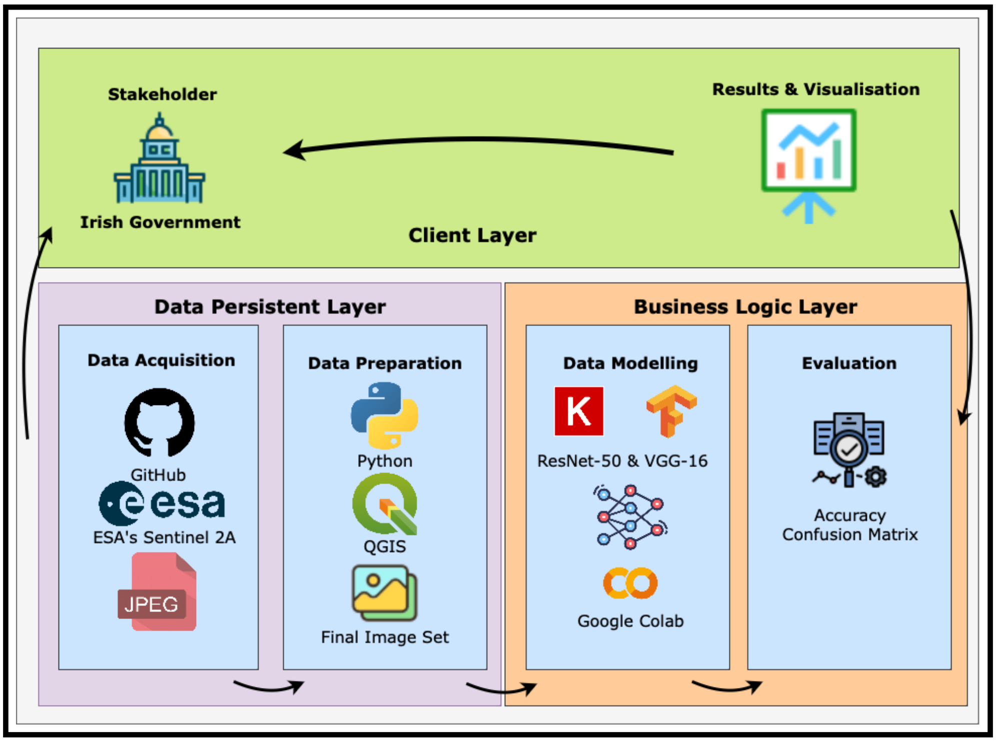

Figure 2: Design Specification

A two-tier architecture was used in the classification of this research project to classify land use and land cover using deep learning models. The project's fundamental implementation component is shown in the Figure 2. The client layer is the top tier of the architecture. The Irish government, the stakeholder, will receive the visualisations and results of the classification models from this layer.

The business logic layer, which makes up the second tier, is where classification model training and data collecting, preparation, and transformation take place. The datasets are obtained by this layer from GitHub and the ESA portal, where they are then enhanced and produced in the proper format for the models to be trained on. The models are then built upon these datasets and tested to get the accuracies on which they are evaluated.

3.4 Conclusion

In this research project, the scientific methodology has been adopted as the plan for implementation. The CRISP-DM methodology has been tailored to fit the requirements of the project. This modified methodology is used to create a process flow diagram, which guides the design of a three-tier architecture that illustrates the implementation of the project. The three tiers of architecture work together to provide a comprehensive solution for classifying Irish land cover.

4 Implementation, Evaluation and Results for Classification of Land Cover of Ireland

4.1 Introduction

In this section, we describe the models used to classify Ireland's land use and land cover from satellite images. We provide a detailed description of how the satellite image patches were obtained from the Copernicus web portal and discuss the implementation evaluation and results of the models. Accuracy is used as a metric to assess the models, and the confusion matrix is used to investigate the model-wise true or false rate prediction. Here we compare the different models that have been implemented and pick the one with the best accuracy.

4.2 Dataset Creation and Pre-processing

4.2.1 Data Augmentation and Collection of EuroSAT

In Section 2.2,the EuroSAT dataset was introduced as a valuable resource for land use and land cover research. Helber et al. (2017) compiled these satellite images from 34 different countries using the Sentinel 2A satellite. The dataset includes 27,000 images, categorised into ten different types of land cover, with 2,000 to 3,000 images per class. The distribution of images across these classes is shown in the accompanying Table 3.

For this research project, the dataset was downloaded from the official GitHub repository[1], and the ten land cover classes were re-categorized into five distinct classes of the land cover of interest. As a result of the re-categorization, there is a significant difference in the number of images between the five classes, this class imbalance can result

[1] https://github.com/phelber/EuroSAT

Table 3: Land Cover Classes and Number of Images in EuroSAT

Land Cover	Number of Images
Annual Crop	3000
Forest	3000
Sealake	3000
Residential	3000
River	2500
Permanent Crop	2500
Herbaceous Vegetation	3000
Pasture	2000
Highway	2500
Industrial	2500
Total	**27000**

in the classification models exhibiting bias towards the dominant class, and thus basic data augmentation techniques were applied. Randomised horizontal and vertical flips, rotations, zooming and shear adjustments were performed as part of the augmentation because they don't add to the satellite images' spectral or topological information, as demonstrated by Stivaktakis et al. (2019) and discussed in section 2.4.2. The Table 4 shows the recategorised classes and the number of images in these new classes before and after augmentation.

Table 4: Recategorised Classes and Number Images After Augmentation

Land Cover	Number of Images Before Augmentation	Number of Images After Augmentation
Developed	5500	10,500
Forest	3000	10,497
TransportLand	2500	10,498
Vegetation	10,500	10,500
WaterBodies	5,500	10,500

The image set was split into training and validation sets in a ratio of 80:20. Figure 3 depicts images present in each class.

4.2.2 Collection of Satellite Images for Ireland

Sentinel-2A, a European optical imaging satellite, was launched by the Copernicus Programme of the European Space Agency in 2015. The Copernicus program makes data with a 10m, 20m, and 50m resolution and 13 spectral bands available for research and commercial purposes. To make image patches for this project, red, green, blue and near-infrared patches were used at 10m resolution; GeoTIFF files for Dublin, Cork and Galway as the main focus and some of its surrounding counties were downloaded from the Copernicus online portal[2].

Over 300 Large Urban Zones and their surrounding areas are represented on the European Urban Atlas, which features accurate land use labels. GeoPackage format files were downloaded from the Urban Atlas 2018[3], which provided labels for several land

[2] https://scihub.copernicus.eu/dhus/#/home

[3] https://land.copernicus.eu/local/urban-atlas/urban-atlas-2018

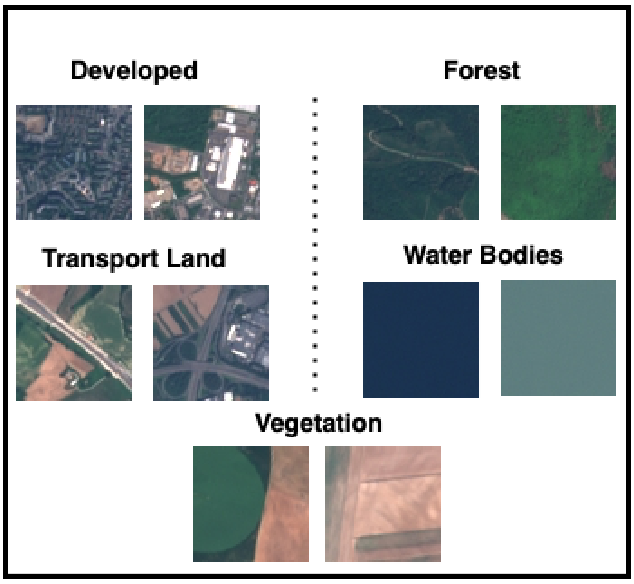

Figure 3: Sample Images from Each Class after Reclassifying

cover classes. These GeoTIFF and GeoPackage files were imported to QGIS, an open-source geographic information system where raster images for the land cover classes were exported. The Figure 4 shows the GeoTIFF file and the GeoPackage files for the Dublin area on top of one another to help demonstrate the process of generating these images.

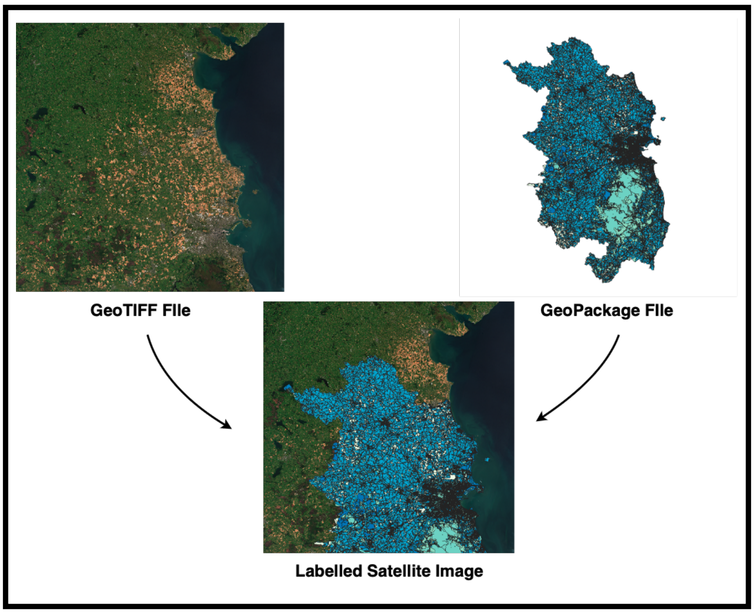

Figure 4: Generating Labelled Satellite Image for Ireland

After these images were extracted, patches of 64x64 pixels were created and divided into the same five land cover classes; these patches will be used as the test set for the classification models to be evaluated. The Figure 5 shows the number of images distributed in each class of the generated Irish dataset and a sample image for each class.

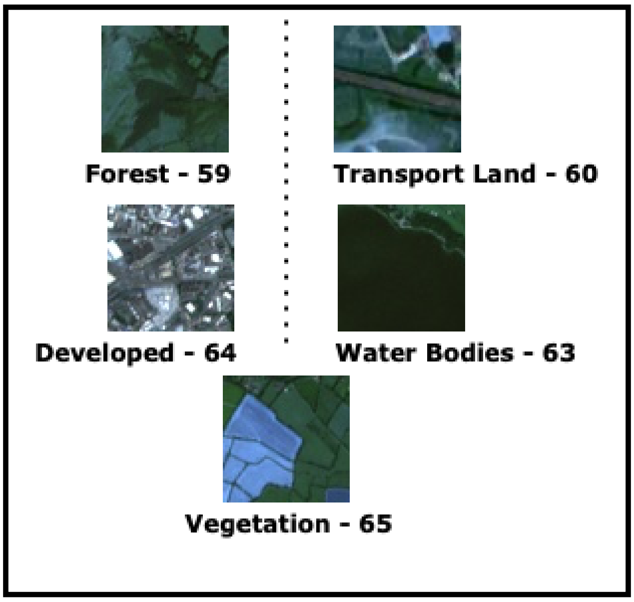

Figure 5: Sample Images and Number of Images in Each Class of Irish Set

4.3 Implementation, Evaluation and Results of ResNet-50

4.4 Implementation of ResNet-50

In the field of computer vision, the ResNet-50 deep learning model has seen extensive use. This model, developed by Microsoft's researchers, is a variation of the widely used ResNet model and has been shown to be effective at a number of tasks, including image classification, object detection, and semantic segmentation. The name "ResNet-50" comes from the fact that it is 50 layers deep, making it more complicated than many other deep learning models. Layers of convolution, ReLu, and batch normalisation make up a typical residual network block. ResNet50's strengths lie in its encouragement of feature usage, its reinforcement of feature propagation, and its drastic reduction in the number of parameters. The identity connection is a unique feature of ResNet that allows for deeper models by skipping connections between layers and instead adding the output of the preceding layer to the layer to which it is connected(He et al.; 2016).

A ResNet50 model with imagenet weights was imported using the Keras library. The top layers were frozen, rendering them untrainable, and three additional layers were added. The output of ResNet50 is flattened before being sent to a ReLu activation function layer and then a 5-class softmax classifier. The model was compiled with Adam optimiser with a 0.001 rate of learning, accuracy as the metrics, and categorical cross-entropy as the loss function. Using the image data generator function of the Keras preprocessing library, the input image was resized to 64x64 pixels, converted to the proper format, and a training and validation split of 0.2 was applied to the model. The models are evaluated using accuracy, and confusion matrix. Below are the outcomes and evaluations of both models with augmented and non-augmented images. Thus, objective 3.1 of Section 1.1 is met.

$$Accuracy = \frac{TruePositive + TrueNegative}{TruePositive + FalsePositive + TrueNegative + FalseNegative} \quad (1)$$

The equation 1 shows the formula used to calculate accuracy of a model

4.4.1 Evaluation and Results of ResNet-50 without Augmentation

First, the ResNet-50 model was trained on the reclassified image set without any augmentation using an 80/20 split between training and validation. The model was run for ten epochs, after which it achieved an accuracy of 99.24% on the training set and 93.26% on the validation set. The model was then evaluated on the Irish image patches test set, where the accuracy dropped significantly to 60.50%. The drop in accuracy was already predicted and could be the cause of the class imbalance.

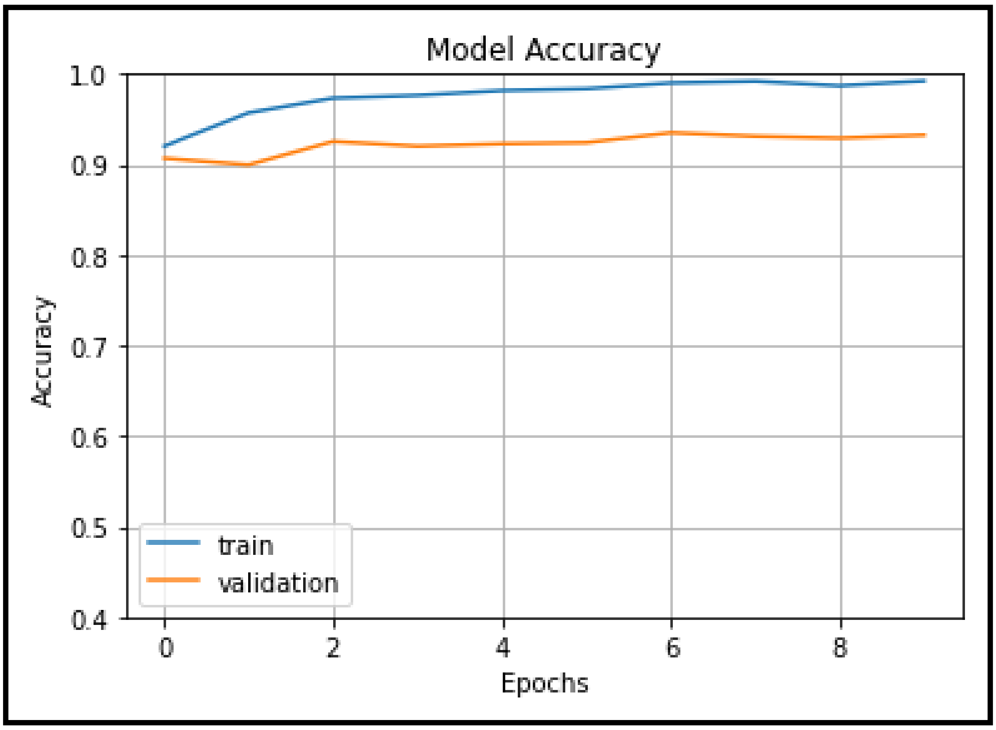

Figure 6: Training and Validation Accuracy of ResNet-50(No Augmentation) over 10 Epochs

The graph in the Figure 6 shows the training and validation accuracy for ten epochs, and the confusion matrix in Figure 7 and Figure 8 shows how the validation set was classified accurately, but the new test set was classified poorly, and the model classified most of the images in the vegetation class; the vegetation class has the maximum number of images in the training set which explains why this class imbalance may have caused the drop in accuracy for the test set.

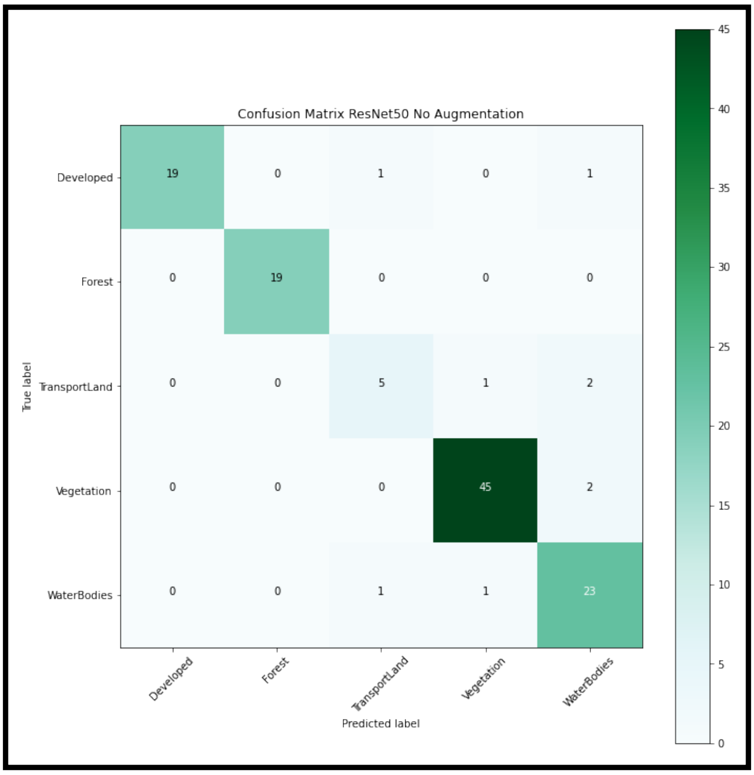

Figure 7: Confusion Matrix for ResNet-50(Without Augmentation) on Validation Set

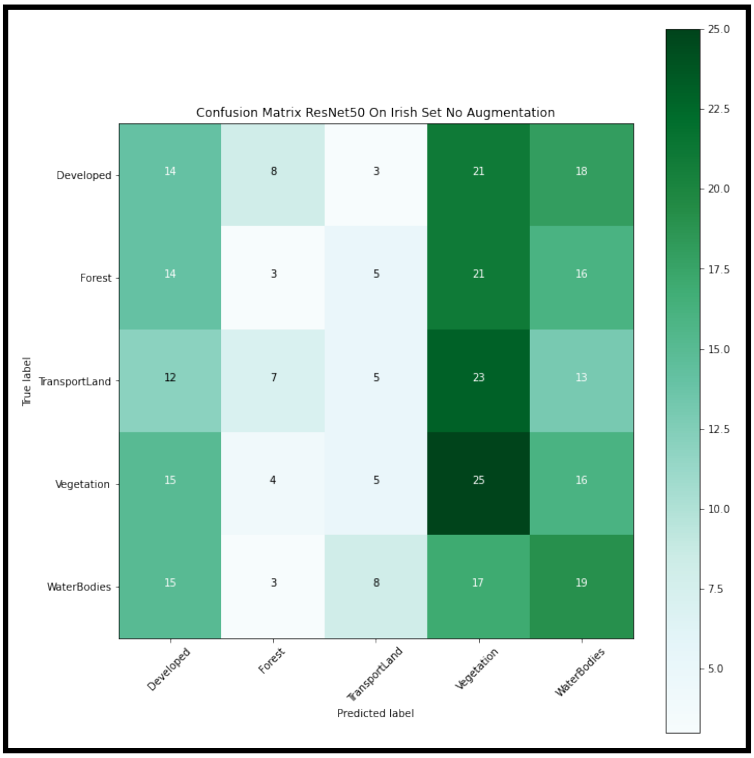

Figure 8: Confusion Matrix for ResNet-50(Without Augmentation) on Test Set

4.4.2 Evaluation and Results of ResNet-50 with Augmentation

4.4.2 Evaluation and Results of ResNet-50 with Augmentation The ResNet-50 model was trained using an 80/20 split between training and validation on the reclassified image set, with basic image augmentation applied to remove class imbalance. After ten epochs, the model achieved an accuracy of 99.18% on the training set and 91.17% on the validation set. The model was then tested on the Irish image patches test set, where an accuracy of 73% was achieved, which is an improvement of 13% as compared to the model trained with no augmentation.

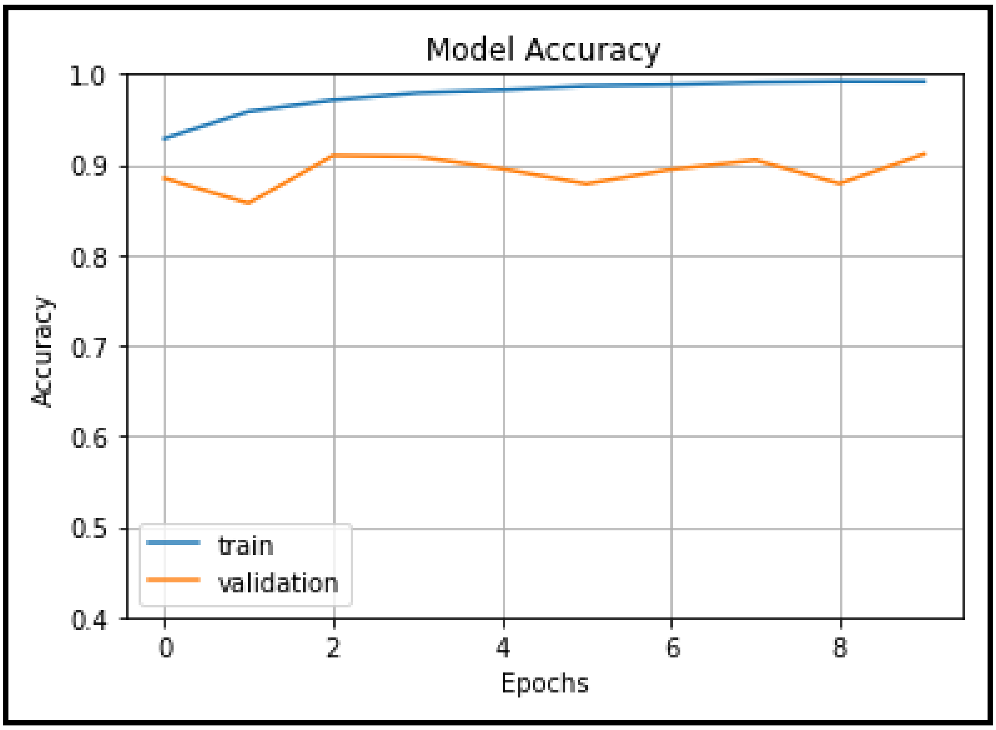

Figure 9: Training and Validation Accuracy of ResNet-50(With Augmentation) over 10 Epochs

The graph in the Figure 9 shows the validation and training accuracies over the epochs, and Figure 10 and 11 show the confusion matrix on the validation and testing sets. The number of true positives for the test set has improved significantly compared to the previous model. However, there are still a considerable number of images falsely classified and a large number of images belonging to the vegetation class wrongly classified as water bodies.

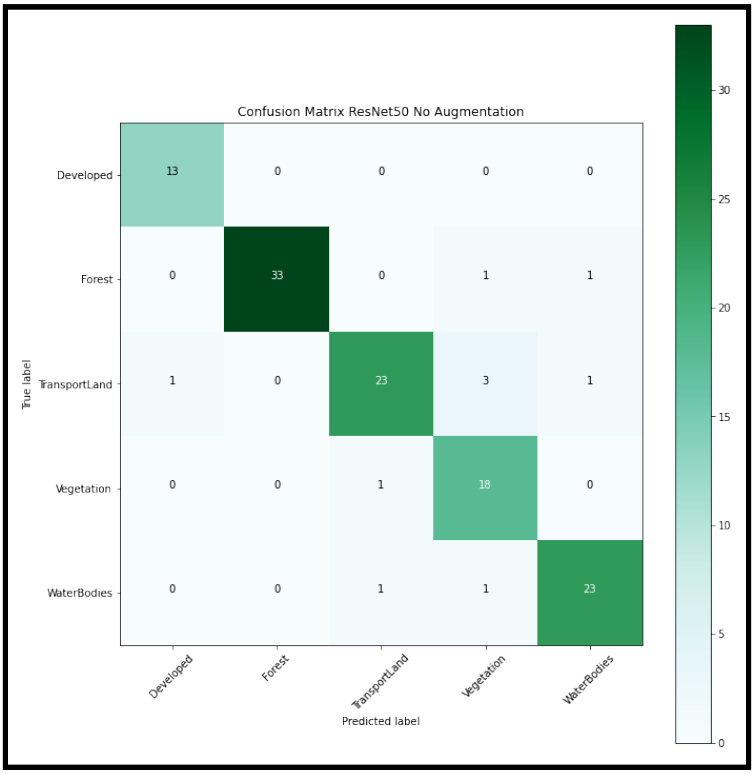

Figure 10: Confusion Matrix for ResNet-50(With Augmentation) on Validation Set

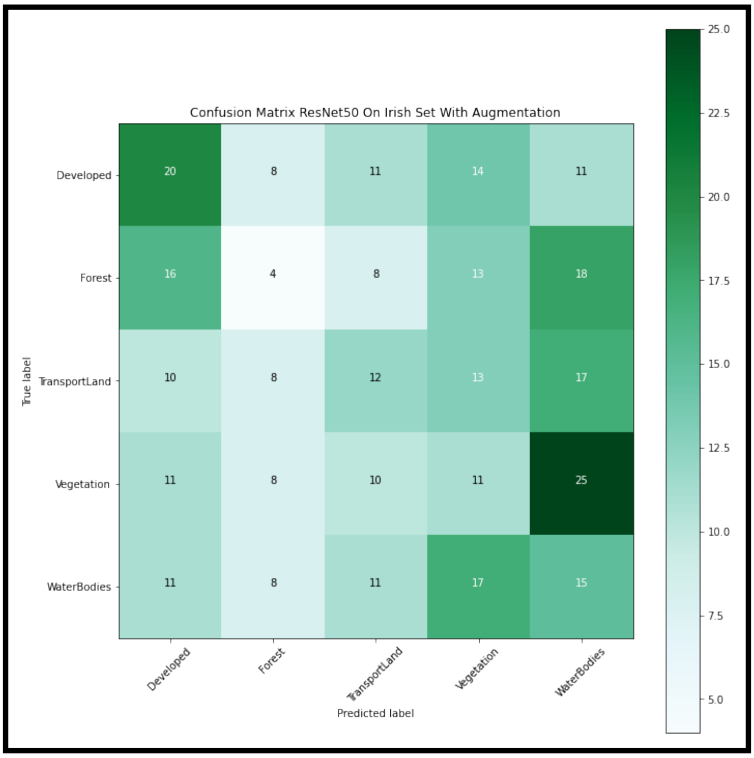

Figure 11: Confusion Matrix for ResNet-50(With Augmentation) on Test Set

4.5 Implementation, Evaluation and Results of VGG-16

4.5.1 Implementation of VGG-16

A Convolutional neural network model trained on the ImageNet dataset, which is frequently used for image classification tasks. "VGG" refers to the researchers at Oxford's Visual Geometry Group who developed the model, while "16" indicates the number of network layers. It was published in a 2014 paper bySimonyan and Zisserman (2014)). There are a total of sixteen layers, thirteen of which are convolutional and three of which are fully connected. The input images' features are extracted by the convolutional layers, while the classification is handled by the fully-connected layers. All of the convolutional layers in the model have a kernel size of 3x3 and a stride of 1. Initially, the number of filters used is relatively low, but as the network's depth grows, more filters are introduced to help the model pick up more complex features.

Similar to the ResNet model; using Keras, a pretrained VGG-16 model with imagenet weights was imported. The top three layers were frozen. VGG's output is flattened before being sent to ReLu and a 5-class softmax classifier. Adam optimiser was used to compile the model with a 0.001 learning rate, accuracy as the metric, and categorical cross-entropy as the loss function. The input set of images was resized to 64x64 pixels, and converted to the proper format, and a 0.2 training/validation split was applied to the model. Accuracy and confusion matrix are used to evaluate the models. Below are the results and evaluations of augmented and non-augmented models. Thus, Section 1.1 Objective 3.2 is met.

4.5.2 Evaluation and Results of VGG-16 With No Augmentation

Similar to the Residual Network model in the previous section, the VGG-16 model was trained using an 80-20 split of the reclassified image set. After the model was run for ten epochs, an accuracy of 98.63% was achieved on the training set and 93.43% on the validation set. The model was then evaluated using the test set, and an accuracy of 65% was achieved, which is relatively low and can result from the class imbalance as seen in the previous models.

Figure 12, 13 and 14 show the graph for the training and validation accuracies over ten epochs, confusion matrix for validation and test sets, respectively. The confusion matrix for the validation set shows that the model trained is capable of classifying most of the images accurately; the confusion matrix for the test set shows how the model fails to classify several images in the right class, and many images are classified into vegetation and waterbodies classes. The model is trained again with augmented images in the next section.

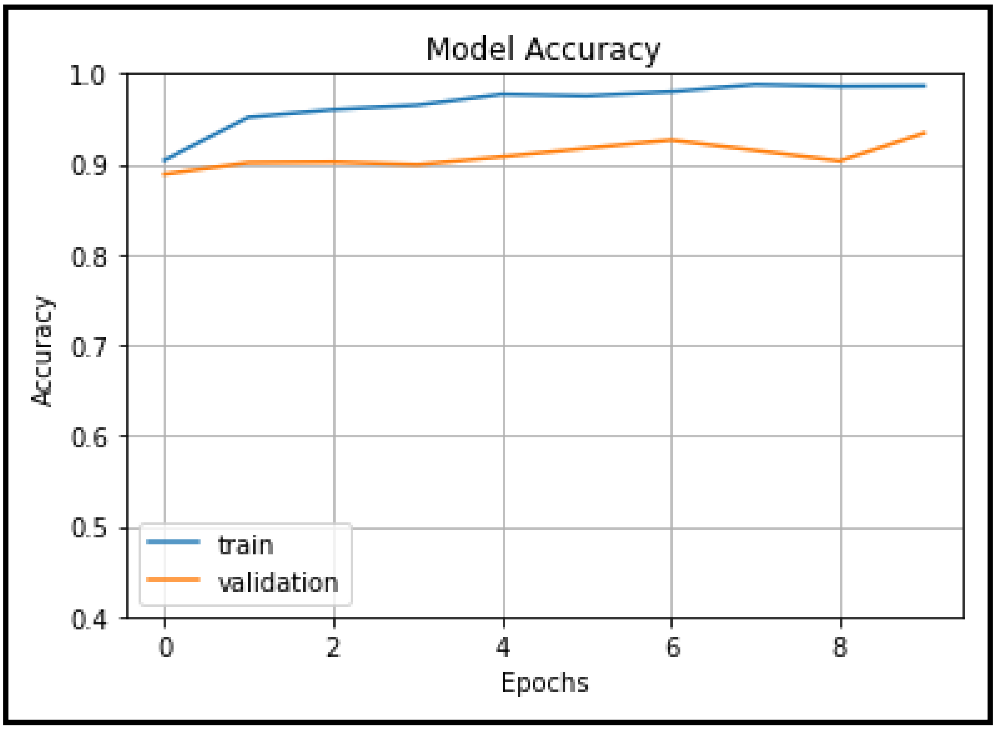

Figure 12: Training and Validation Accuracy of VGG-16(No Augmentation) over 10 Epochs

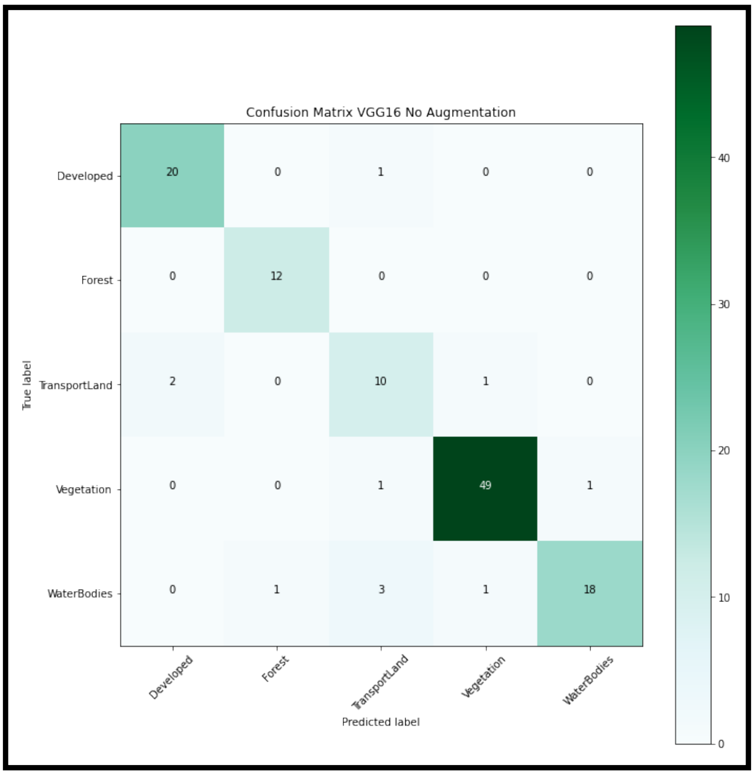

Figure 13: Confusion Matrix for VGG-16(No Augmentation) on Validation Set

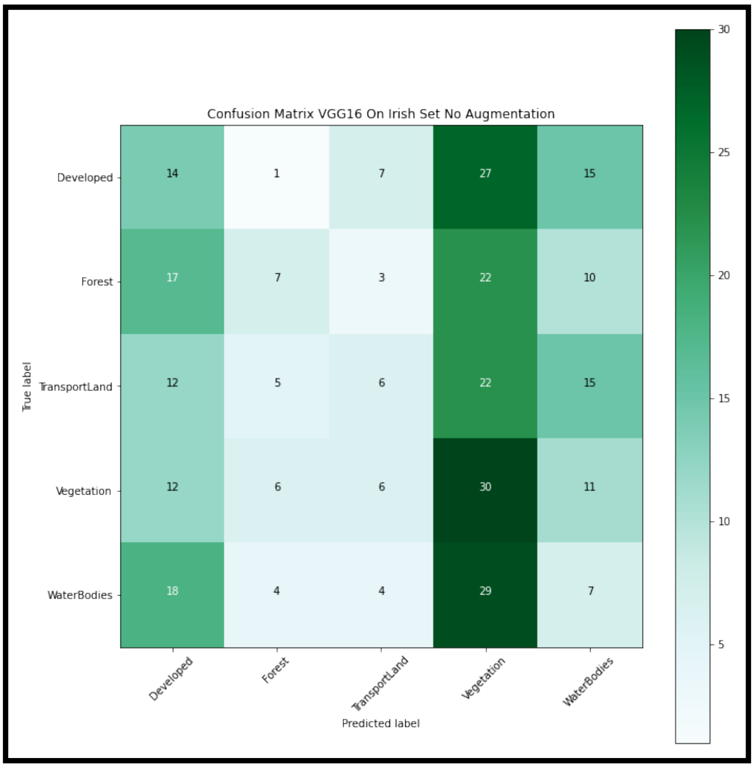

Figure 14: Confusion Matrix for VGG-16(No Augmentation) on Test Set

4.5.3 Evaluation and Results of VGG-16 With Augmentation

To try and improve the classification accuracy on the test set, the VGG-16 model was trained again using an 80-20 training and validation of the image set after applying basic augmentation techniques to it. The model was run for 10 epochs and had an accuracy of 98.51% on the training set and 90.04% on the validation set. Using the test set the model was then evaluated, achieving an accuracy of 65%, showing no improvement from the model with no augmentation applied to it.

Figure 15, 16 and 17 show the graph for the training and validation accuracies over ten epochs, confusion matrix for validation and test sets, respectively. The confusion matrices depict similar results as the model previously trained and show no improvement after using augmented images.

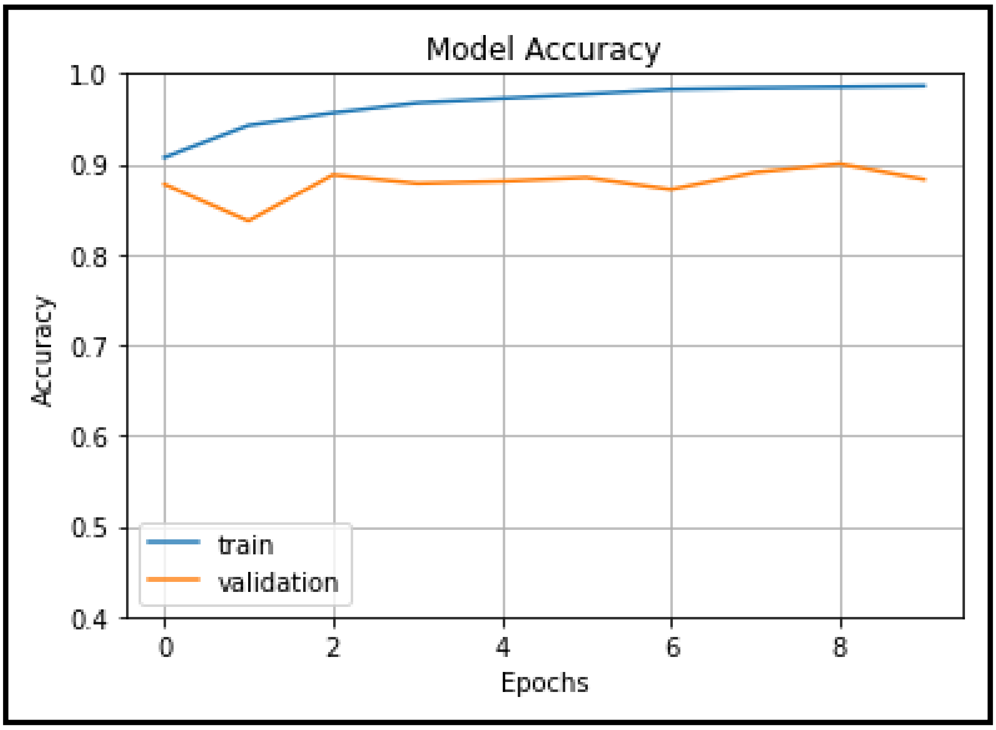

Figure 15: Training and Validation Accuracy of VGG-16(With Augmentation) over 10 Epochs

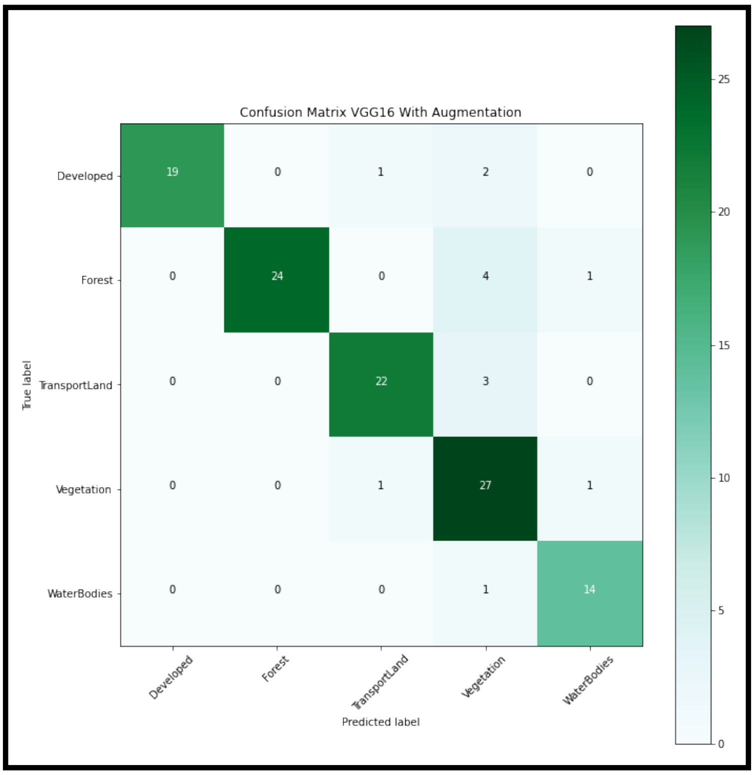

Figure 16: Confusion Matrix for VGG-16(With Augmentation) on Validation Set

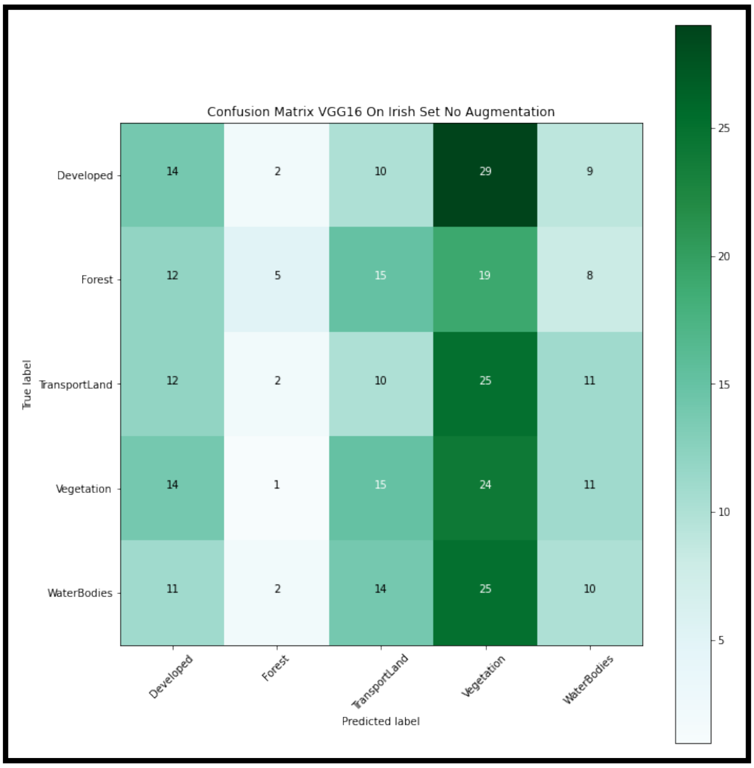

Figure 17: Confusion Matrix for VGG-16(With Augmentation) on Test Set

4.6 Discussion and Comparison of Developed Models

Figure 18 displays the comparison of the ResNet-50 and VGG-16 models' performance with and without the use of image augmentation. . In the first part of the project, the pre-trained models were used on the first set of the image dataset without augmentation. Figure 18 shows that all the base models accurately classified the land covers, with all models having similar training and validation accuracies achieving high accuracies. To evaluate the performance of the models on the Ireland satellite image set, each model was tested using the Irish set as the test set. The base ResNet model achieved a low accuracy but the accuracy improved when it was retrained using the training set with augmented images, providing insight and answering the research question and sub-research question in section 1.1. The base VGG model did not improve in accuracy after using the augmented image set and both the base and augmented VGG-16 models achieved a low accuracy in classifying the Irish satellite image set. This section marks the completion of objective 4 from section 1.1

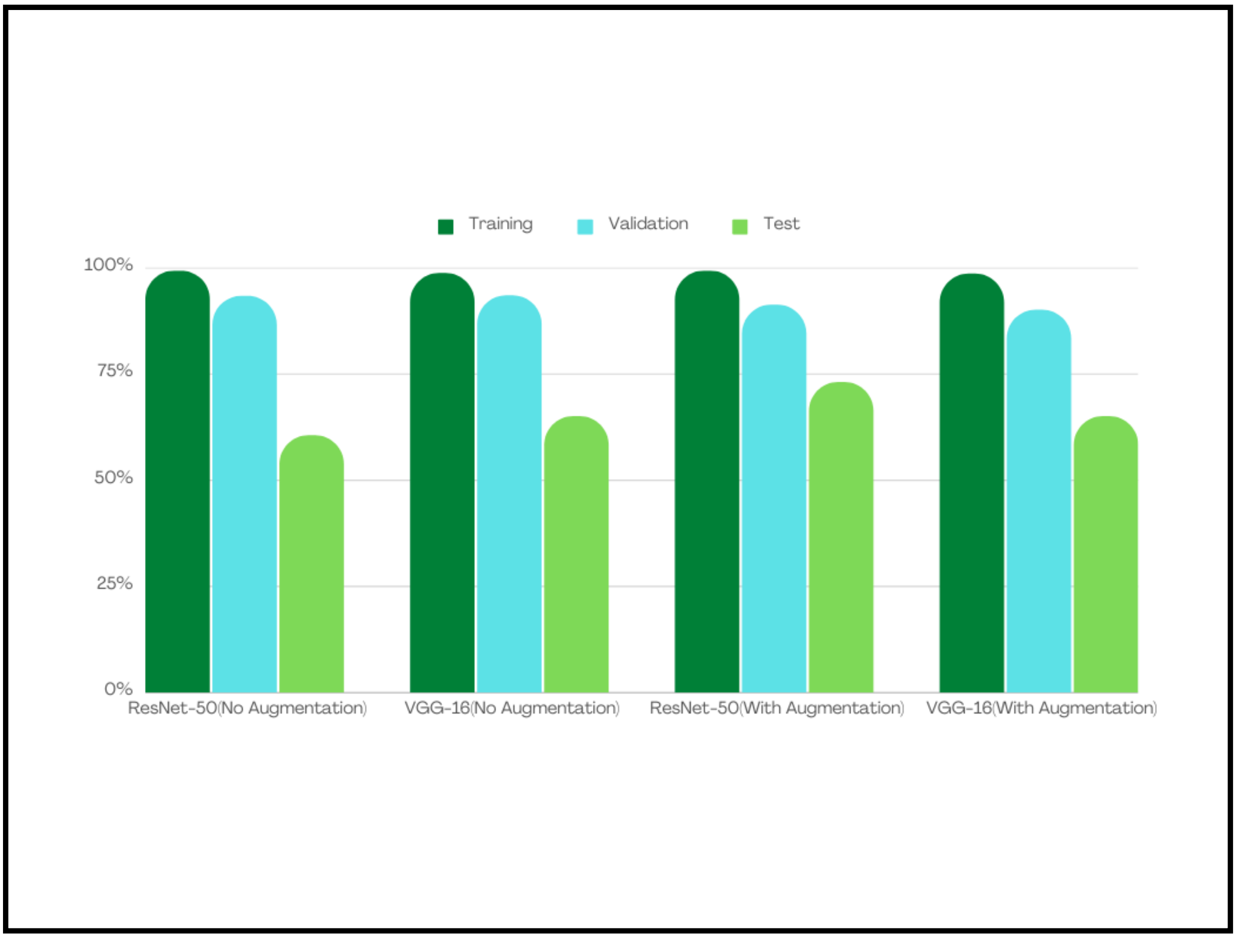

Figure 18: Comparison of The Developed Models

5 Conclusion and Future Work

After conducting a thorough literature review, this project formulates a methodology for categorising land use and land cover (LULC) of Ireland. After reviewing the available literature, the deep learning models VGG16 and ResNet50 were chosen for further investigation. This study set out to answer the questions of how well deep learning models (VGG16, ResNet50) perform when used to classify unseen out-of-domain satellite images and whether or not image augmentation improves the performance of the models when trained on a novel LULC dataset with several land cover classes. To find out, two sets of training datasets were made. One set was pre-processed, while the other set was added with rescaling, rotation, zoom, and horizontal/vertical flip operations. The two sets of data were used to train and test the models that were made. ResNet50 performed better on both datasets achieving an improved accuracy of 73% on the Irish dataset after data augmentation and had a boost of 13% in its accuracy from the base model. In contrast to ResNet50, VGG16 did not show any improvement in accuracy when used with augmented images, despite the fact that its performance on the first set of non-augmented data was comparable to that of ResNet50. The developed models performed well with out-of-domain images and resulted in a novel approach that organisations like the Irish government can put to use for the sustainable development of the country. This way, the research questions and objectives outlined in subsection 1.4 were accomplished. To the best of the candidate's knowledge, over the past few decades, insufficient research has been conducted on the classification of out-of-domain land cover using deep learning, thus filling the gaps in the existing body of knowledge discussed in subsection 1.5. During the course of this reserach project the candidate gained knowledge of how to use GeoTIFF and Geopackage files in QGIS to raster satellite images, how deep learning models function and how to fine-tune them, what image augmentation techniques are available and when to use them, and how a cloud-based deployment solution can shorten the training time of a model.

Future Work: Despite the fact that the developed models had a good performance and were able to contribute to the existing literature, they also had some limitations of their own. When applied to the Irish image set, the models achieved an acceptable level of classification accuracy; however, there were a number of incorrect classifications, especially in the categories of vegetation and water bodies. It's possible that this is because the terrain in the EuroSAT set used for training comes from 34 different countries. In the future, researchers can form a dataset with land cover images captured in a geographic area of interest use it to train the classifier model. To broaden the scope of the study, additional pre-trained models can be compared. The developed models can be utilised in industrial applications, and their performance can be monitored, thereby increasing the system's reliability.

Acknowledgement

I would like to express my sincere appreciation to Dr. Catherine Mulwa for her consistent and thorough guidance throughout the semester, as well as for her patience in providing appropriate and detailed feedback whenever it was required. I would also like to thank National college of Ireland for the library resources that assisted me in writing this technical report and conducting this research project's critical literature review using peer reviewed papers from reputable sources. Without my parent's support, blessings and confidence, I would not have been able to finish this research.

References

Acharya, T. D., Yang, I. T. and Lee, D. H. (2016). Land cover classification using a kompsat-3a multi-spectral satellite image, *Applied Sciences (Switzerland)* **6**.

Alem, A. and Kumar, S. (2022). Transfer learning models for land cover and land use classification in remote sensing image, *Applied Artificial Intelligence* **36**.

Carranza-García, M., García-Gutiérrez, J. and Riquelme, J. C. (2019). A framework for evaluating land use and land cover classification using convolutional neural networks, *Remote Sensing* **11**.

de Lima, R. P. and Marfurt, K. (2020). Convolutional neural network for remote-sensing scene classification: Transfer learning analysis, *Remote Sensing* **12**.

Ganasri, B. and Dwarakish, G. (2015). Study of land use/land cover dynamics through classification algorithms for harangi catchment area, karnataka state, india, *Aquatic Procedia* **4**: 1413–1420.

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

Helber, P., Bischke, B., Dengel, A. and Borth, D. (2017). Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification.
URL: *http://arxiv.org/abs/1709.00029*

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**: 14680–14707.

Hütt, C., Koppe, W., Miao, Y. and Bareth, G. (2016). Best accuracy land use/land cover (lulc) classification to derive crop types using multitemporal, multisensor, and multi-polarization sar satellite images, *Remote Sensing* **8**.

Li, C., Wang, J., Wang, L., Hu, L. and Gong, P. (2014). Comparison of classification algorithms and training sample sizes in urban land classification with landsat thematic mapper imagery, *Remote Sensing* **6**: 964–983.

Ma, J., Sun, Y., Deng, G., Huang, S., Tao, Y., Teng, Q. and Meng, X. (2019). Evaluation of different approaches of convolutional neural networks for land use and land cover classification based on high resolution remote sensing images; evaluation of different approaches of convolutional neural networks for land use and land cover classification based on high resolution remote sensing images.

Mahamunkar, G. S. and Netak, L. D. (2022). Comparison of various deep cnn models for land use and land cover classification, Vol. 13184 LNCS, Springer Science and Business Media Deutschland GmbH, pp. 499–510.

Naushad, R., Kaur, T. and Ghaderpour, E. (2021). Deep transfer learning for land use and land cover classification: A comparative study, *Sensors* **21**.

Noi, P. T. and Kappas, M. (2017). Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using sentinel-2 imagery, *Sensors (Basel, Switzerland)* **18**.

Rousset, G., Despinoy, M., Schindler, K. and Mangeas, M. (2021). Assessment of deep learning techniques for land use land cover classification in southern new caledonia, *Remote Sensing* **13**.

Sathyanarayanan, D., Anudeep, D. V., Das, C. A. K., Bhanadarkar, S., Uma, D., Hebbar, R. and Raj, K. G. (2020). A multiclass deep learning approach for lulc classification of multispectral satellite images, Institute of Electrical and Electronics Engineers Inc., pp. 102–105.

Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition, *arXiv preprint arXiv:1409.1556* .

Stivaktakis, R., Tsagkatakis, G. and Tsakalides, P. (2019). Deep learning for multilabel land cover scene categorization using data augmentation, *IEEE Geoscience and Remote Sensing Letters* **16**: 1031–1035.

Volodymyr, G. E. M. and Hinton (2010). Learning to detect roads in high-resolution aerial images, Springer Berlin Heidelberg, pp. 210–223.

Wirth, R. and Hipp, J. (2000). Crisp-dm: Towards a standard process model for data mining.

Xiangnan, Yin, C., Weihai, Wu, Xingming, Yue and Haosong (2017). Fine-tuning and visualization of convolutional neural networks.

Yosinski, J., Clune, J., Bengio, Y. and Lipson, H. (2014). How transferable are features in deep neural networks?, *Advances in neural information processing systems* **27**.

Yu, X., Wu, X., Luo, C. and Ren, P. (2017). Deep learning in remote sensing scene classification: a data augmentation enhanced convolutional neural network framework, *GIScience and Remote Sensing* **54**: 741–758.

Zhang, D., Liu, Z. and Shi, X. (2020). Transfer learning on efficientnet for remote sensing image classification, Institute of Electrical and Electronics Engineers Inc., pp. 2255–2258.