Deep Learning Framework for Land Use and Land Cover Classification and Change Detection

MSc Research Project
Data Analytics

Jawed Siddique
Student ID: x22232915

School of Computing
National College of Ireland

Supervisor: Paul Stynes, Musfira Jilani & Mark Cudden

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Jawed Siddique		
Student ID:	x22232915		
Programme:	MSc. Data Analytics	**Year:**	2023 - 2024
Module:	MSc. Research Project		
Supervisor:	Paul Stynes, Musfira Jilani & Mark Cudden		
Submission Due Date:	12/08/2024		
Project Title:	Deep Learning Framework for Land Use and Land Cover Classification and Change Detectio		
Word Count:	6024	**Page Count:**	22

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: Jawed Siddique

Date: 11/08/2024

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):	

Deep Learning Framework for Land Use and Land Cover Classification and Change Detection

Jawed Siddique
x22232915

Abstract

Land Use and Land Cover (LULC) classification and change detection have numerous environmental applications. However, the limited availability of high-quality labelled satellite data, high intra-class variability, and issues related to the noise in the available data LULC classification and change detection could be challenging. This study aims to address this challenge by performing LULC classification and detecting LULC change between the years 2018 and 2024 for the Greater Dublin Area by using the Sentinel-2 satellite imagery dataset and Deep Learning architecture. This study used data pre-processing and data transformation techniques to prepare a multi-spectral tiled satellite image dataset before employing supervised classification techniques. For classification purposes, a deep, multi-layered CNN architecture ResNet50 model was used and five major classes: Artificial Surfaces, Agricultural Areas, Forest and Seminatural Areas, Wetlands, and Water Bodies were identified. For LULC classification evaluation, the classification error matrix, accuracy, precision, recall, F-Score and kappa analysis were used. For LULC change detection, the differences in the areas covered by each land cover class between the periods are measured. According to the findings from the study, the Agricultural land cover class accounted for 74.67% in 2018 and 75.27% in 2024, making it the most extensive class. There is, however, a shift in the second-largest class cover from Forest and Seminatural Areas in 2018 to Artificial Areas in 2024. Additionally, this study identified a significant change in the Forest (-3.20%) and Artificial Surfaces (+1.24) from 2018 to 2024. Lastly, the classification methodology achieved overall accuracy, precision and Kappa statistics of 92.38%, 92.41, and 0.91 respectively.

Keywords: Land Use, Land Cover, Change Detection, Sentinal-2, Deep Learning, Convolutional Neural Networks, ResNet50, Accuracy Assessment

1. Introduction

Land Use and Land Cover (LULC) classification and change detection are critical for effective land resource planning and management (Esfandeh *et al.*, 2021; Yao *et al.*, 2021). This helps in informed urban planning (Hamidi and Ewing, 2014), research on global climate change (Coppin et al. 2004), and natural disaster management (Qingmu, Liao & Chen, 2021). According to the

Environmental Protection Agency, Ireland (EPA, Ireland), Land Use (LU) refers to the various ways humans utilize land, while Land Cover (LC) describes everything visible on the land surface (EPA, 2020a). Starting primarily in the 1970s, manual classification and statistical techniques were used for the study of LULC (Yan *et al.*, 2020), however, advancements in computing power and technology over the past three decades, have introduced machine learning and deep learning methods for LULC classification and change detection utilizing Remote Sensing and Geographic Information Systems (GIS) imagery (Wang *et al.*, 2022; Garcia, Gutierrez & Riquelme, 2019; Phiri & Morgenroth, 2017) with varying success which in turn can be attributed to the various challenges associated with the datasets and techniques used (Alshari & Gawali, 2021).

A major challenge in LULC classification and change detection is the low variability of the spectral profiles of different LULC types, high intra-class variability, and issues related to the noise in satellite data and the low spatial resolution of freely available satellite images (Lin et al., 2015; Rodriguez-Galiano et al., 2012). The study of land use and land cover for the Republic of Ireland also faces these challenges due to the limited availability of high-quality labelled training data (Brennan *et al.*, 2017). Additionally, as per the CSO 2022 data, the Republic of Ireland is experiencing major shifts with respect to urban and rural population and land uses. Currently, there is an urban population of 64.5% which is growing at an average annual rate of 0.39% (CSO, 2022). The agricultural sector, covering approximately 66% of the total land area, is also undergoing rapid transformations due to environmental policies and market demands (EPA, 2023b). Finally, the country's natural habitats, including forests and wetlands which make up about 10% and 2% of the land area respectively, are facing pressures from both urban expansion and climate change (EPA, 2023c). These dynamic and diverse land use changes necessitate robust and accurate LULC classification systems to inform sustainable development policies, environmental conservation efforts, and effective resource management. This research is motivated by this need to address these challenges and provide a method for LULC classification and change detection, thereby supporting informed decision-making, resource management and sustainable urban planning.

This research aims to investigate how well CNN-based deep learning architecture performs for LULC classification and change detection in the area of interest – the Greater Dublin Area, during the period 2018 - 2024.

The major contributions of this research are as follows:

1. A classification of the land use and land cover in the Greater Dublin Area with respect to the five most predominant classes, namely Agricultural Areas, Artificial Surfaces, Forest and Seminatural Areas, Water Bodies, and Wetlands.
2. An analysis of change in land use and land cover in the area of interest over six years from 2018 to 2024.

This paper discusses and performs a critical analysis of past studies and related work in Section 2 with a focus on LULC classification and change detection using deep learning techniques. Subsequently in Section 3, the research methodology is discussed. Section 4 discusses the design specification of the study and Section 5 discusses the implementation of the designed framework. The achieved results are presented and discussed in Section 6. Finally, the conclusion of this research and possible future works is presented in Section 7.

2. Related Work

LULC classification and change detection using machine learning was introduced in the 1970s (Alshari & Gawali, 2021) and the two most widely used classification methods are the Maximum Likelihood Classifier (MLC) and Minimum Distance Classifier (MDC) because of their dependability and simplicity (Khurana & Gupta, 2022; Srivastava *et al.*, 2012). Tripathi & Kumar (2017) achieved an accuracy of 85-90% when they utilized MLC for change detection in the forest category of the Himalayan region of India; Samanta & Paul (2016) used the MLC technique for coastal LULC classification and change detection; Hossen et. al (2018) achieved a high accuracy of 93.33% using MLC and MDC techniques to assess and predict future changes to the LULC of Manzala Lake; and Sampath & Radhakrishnan (2022) employed MLC and MDC techniques in a comparative study on soil erosion rates using the RUSLE model. Chugtai, Abbasi & Karas (2021) in their study showed that the MLC and MDC models gave good results in the classification of vegetation and forest areas, however, their accuracy was reduced when applied to the classification of other regions such as urban landscapes. Additionally, Noi & Kappas (2019), showed that in the absence of high-resolution images, the MLC and MDC techniques failed to achieve consistently high accuracy as these methods use pixel comparison for classification.

Over the past decade, advanced machine learning methods such as support vector machines (SVMs), random forest (RF) etc. and deep learning methods like neural networks have been used for LULC classification. For example, studies by Gislason, Benediktsson & Sveinsson (2006) and Mochizuki & Murakami (2012) focus on using machine learning and assessing the performance of various tree-based classifiers, including random forests (RF), classification and regression trees (CART), and other decision trees using bagging and boosting methods. These studies found that random forests consistently outperformed the other classifiers. Shao & Lunetta (2012), used support vector machines (SVMs) for land cover classification tasks. Srivastava *et al.* (2012) used deep learning techniques for land use and land cover classification and showed that Artificial Neural Networks (ANNs) show better performance over SVM in classifying croplands.

As both machine learning and deep learning techniques have been equally effective in LULC classification, various comparative studies have been performed to determine the most effective approach. Rousset *et al.* (2021) performed a comparative study on SPOT6 satellite data for the topical island using XGBoost machine learning techniques and deep learning methods. Their study found that while both approaches performed similarly for land cover (LC) classification, the deep learning method outperformed XGBoost for land use (LU) classification, achieving an accuracy of 61.45% compared to XGBoost's 51.56%. Similarly, Kussul et al. (2017) evaluated the performance of CNNs against random forests using remote sensing data from Landsat-8 and Sentinel-1A satellites. Their study concluded that CNNs outperformed the random forest methods. Garcia, Gutierrez & Riquelme (2019) performed a comparative study between deep learning and machine learning techniques using five publicly available datasets having hyperspectral (Indian Pines, Salinas, Pavia) and radar (San Fransisco and Flevolad) images. These datasets cover both urban and rural landscapes and have different sizes, spatial resolutions, and number of classes. Support vector machines (SVMs), k-nearest neighbours (1NN, 2NN and 5NN), random forest (RF) and CNNs techniques were applied to the datasets for classification.

Their findings showed that the CNN model had a better computation time (on both CPU and GPU) for four of the five datasets compared to the other machine learning models, consistently achieving accuracy greater than 95% and the highest average accuracy for all datasets. This demonstrated that deep learning – particularly CNNs, is capable of adapting to datasets of varying sizes and characteristics, making them a suitable choice for LULC image classification. The study however faced challenges in designing the CNN architecture and finding the optimal configuration and parameters.

This challenge was addressed in the research by Naushad *et. al* (2021), where they applied transfer learning instead of training CNNs from scratch. By applying transfer learning they fine-tuned pre-trained networks, namely, the Visual Geometry Group (VGG16) and Wide Residual Networks-50 (ResNet-50), on the red–green–blue (RGB) version of the EuroSAT dataset. Their study results showed that Wide Residual Networks-50 (ResNet-50) outperformed the previous best results in both computational efficiency and accuracy, achieving an accuracy of 99.17%. This study however also showed that although there was a difference in the accuracy of the class prediction, the learning pattern of the VGG and ResNet50 was similar. Mahamunkar & Netak (2022), compared random forests with two variants of ResNet and VGG models on the EuroSAT dataset. Their results too corroborated the findings by Naushad *et. al* (2021) and showed that pre-trained models achieved the highest accuracy, reinforcing that transfer learning is an effective approach for LULC classification using deep learning techniques.

Considering the performance of deep learning and transfer learning techniques in LULC classification, several studies have been carried out for different geographic regions on datasets. These studies have utilized datasets generated from a variety of satellite sources, including LandSAT, Sentinel-1, Sentinel-2, and others. Jagannathan & Divya (2021) used deep learning for LULC classification and change detection using deep convolutional neural networks for the region of Coimbatore, India. Sefrin, Reise & Keller (2021) performed image segmentation using a fully convolutional neural network (FCN) and long short-term memory (LSTM) network architectures to detect land cover change detection in the region around Klingenberg in the federal state of Saxony, Germany. Their approach followed generating a novel dataset by merging the ground truth dataset with the Sentinel-2 imagery and generating class labels before performing a training, validation and test split.

In conclusion, current research shows that various models such as ANNs, ResNet50, and VGG16 are used in the LULC classification and change detection for various geographic regions with very good performance. This includes performing the classification using the EuroSAT dataset as well as novel datasets generated from satellite imagery. However, to the best of our knowledge, no major study has been conducted on LULC classification and change detection for the Republic of Ireland from satellite imagery data. This research aims to address this gap by proposing a framework for LULC classification and change detection using the ResNet50 deep learning technique.

3. Methodology

The research methodology comprises five steps: data gathering, data pre-processing, data transformation, data modelling, and evaluation and results, as illustrated in Figure 1.

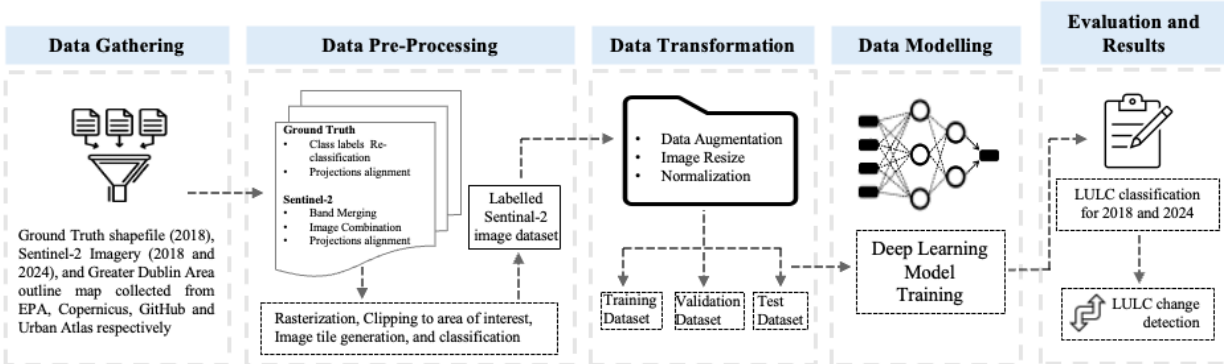

Figure 1: Research Methodology

3.1. Data Gathering

The first step in this study involves gathering the ground truth and Sentinel-2 input data for the years 2018 and 2024 for the area of interest. These three datasets are introduced in Section 3.1.1 and Section 3.1.2 respectively.

3.1.1. Ground Truth Data

The vector data produced by the EPA[1] covering the land cover classes for the Republic of Ireland is considered the ground truth for this study. From the ground truth data, the area of Greater Dublin is extracted. This area is located on the east coast of the Republic of Ireland and includes the capital city of Dublin and its surrounding counties of Meath, Kildare and Wicklow. The area of interest covers approximately 6,986 sq. kilometres and has a population of approximately 2.2 million (Eastern and Midland Regional Assembly, 2019). The ground truth land cover data consists of five major land cover classes: Agricultural Areas, Artificial Surfaces, Forest and Seminatural Areas, Water Bodies, and Wetlands. Figure 2 shows the ground truth data for the area of interest aggregated in 2018[1].

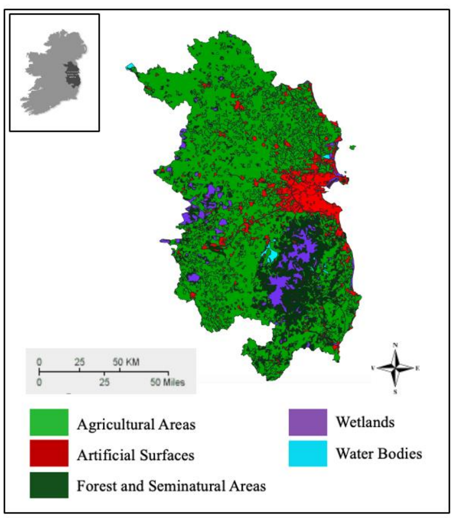

Figure 2: Visualization of Greater Dublin Area (area of interest); and the land cover ground truth data (2018) obtained from Environmental Protection Agency, Ireland.

3.1.2. Sentinel-2 Input Data

Sentinel-2, an optical imaging satellite launched in 2015 by the European Space Agency (ESA) as part of the Copernicus Programme provides multispectral imagery across 13 spectral bands ranging from visible and near-infrared (NIR) to short-wave infrared. The satellite program covers the Earth's surface every five days, generating imagery with spatial resolutions of 10m, 20m, and 60m depending on the specific spectral band (Drusch *et al.*, 2011). As input data for this study, Sentinel-2 MSI (MultiSpectral Instrument, Level-1C) remotely sensing imagery with a maximum of 5% cloud cover is downloaded from the Copernicus online portal[2] for the years 2018 and 2024. However, due to the extent of the area of interest, the data is not available in a single image and therefore, requires downloading separate images representing the North and South of the area of interest. Figure 3 shows the downloaded north and south images generated by the Sentinel-2 satellite for the years 2018 and 2024.

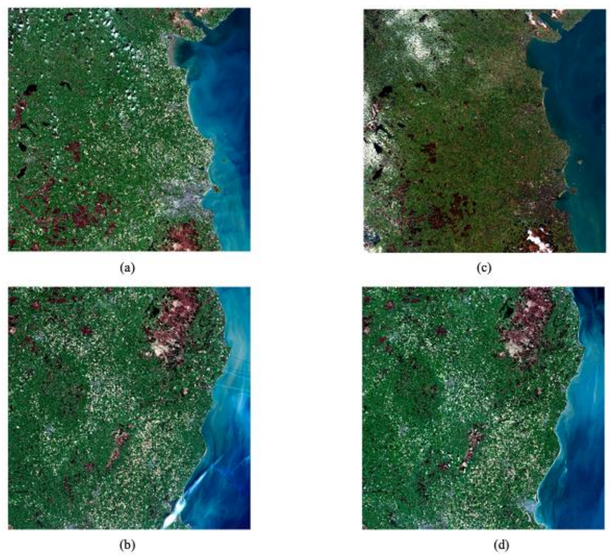

Figure 3: Downloaded Sentinel-2 images from the Copernicus portal. (a - b) North and South region of Greater Dublin Area for 2018; (c-d) North and South region of Greater Dublin Area for 2024

3.2. Data Pre-Processing

The second step involves applying pre-processing techniques to the gathered datasets and preparing them for LULC classification and change detection using deep learning. For the ground truth vector data, the pre-processing includes reclassification, clipping, rasterization, and reprojection; while for the Sentinel-2 the pre-processing involves band merging, image combination, and reprojection.

The ground truth data containing detailed classification is first reclassified into five high-level classes discussed in Section 3.1.1. Next, an outline map of the Greater Dublin Area is downloaded from the Copernicus Land Monitoring Service Urban Atlas[3]. Using this outline map the reclassified shapefile is clipped to the area of interest. This reclassified and clipped ground truth shapefile covering the area of interest is then rasterised by class labels to a spatial resolution of 10m. This resolution of the ground truth raster image is selected to match the resolution of the Sentinel-2 imagery.

1 https://gis.epa.ie/GetData/Download
2https://land.copernicus.eu/en/products/urban-atlas/urban-atlas-2018#download
3https://dataspace.copernicus.eu/explore-data/data-collections/sentinel-data/sentinel-2

As presented in Section 3.1.2., the Sentinel-2 data for the years 2018 and 2024 is downloaded as separate images representing the north and south sections of the area of interest and has information across multiple spectral bands. Therefore, first, the multiple spectral bands (B2, B3, B4 and B8), all at 10m resolution, are merged to generate a single multi-band image and the process is repeated for the downloaded image tiles representing the north and south region. The generated multi-band images were then mosaiced into a single image. Further, this combined image was clipped to cover the area of interest using the preprocessed ground truth vector file. Next, the ground truth raster data and combined Sentinel images were reprojected to the same Coordinate Reference System (EPSG: 32629 WGS 84 / UTM zone 29N). Figure 4 shows tiled and labelled 2018 Sentinel-2 sample images representing the five classes.

Lastly, tiled images of 64*64 pixels were created using the rasterised ground truth data for 2018 and the mosaiced satellite imagery for the years 2018 and 2024. For 2018, these tiled images were then classified according to the ground truth class labels resulting in a fully labelled 2018 satellite imagery dataset. The 2024 tiled images were unlabelled as the ground truth data is currently not available and thus these images were used without labels for prediction. Figure 4 shows tiled and labelled 2018 Sentinel-2 sample images representing the five classes. Figure 5 shows tiled sample images from the 2024 Sentinel-2 dataset.

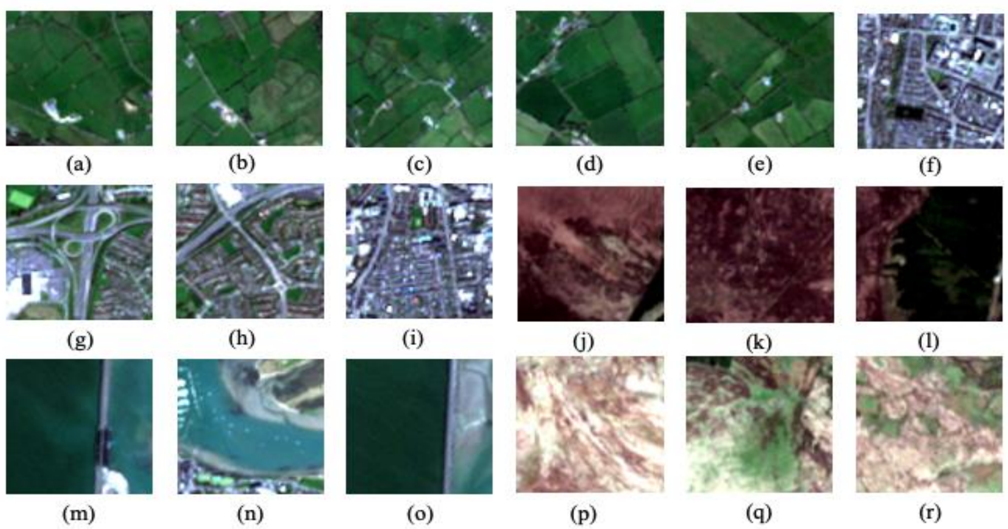

Figure 4: Land use and Land cover sample tiled images generated from Sentinel-2 satellite for 2018. (a-e) Agricultural Areas; (f-i) Artificial Surfaces; (j-l) Forest and Seminatural Areas; (m-o) Water Bodies; (p-r) Wetlands

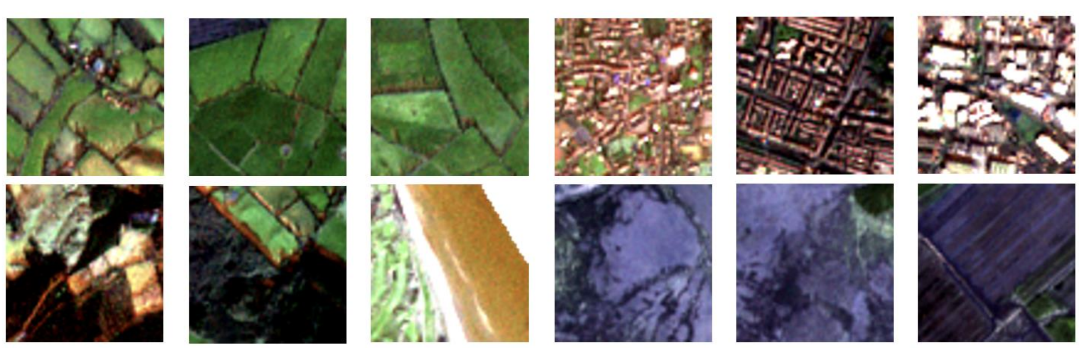

Figure 5: Land use and Land cover sample tiled images generated from Sentinel-2 satellite for 2024.

3.3. Data Transformation

The third step involves data augmentation, image resizing and normalization of the pre-processed Sentinel-2 dataset. Data Augmentation is the process of generating slightly modified samples from existing data that can help enrich the training dataset and improve model generalization and performance. This can be achieved through various techniques such as geometric transformations, colour space adjustments, and noise injection (Shorten & Khoshgoftaar, 2019). In this study, the collected dataset for 2018 contained 5 classes: Artificial Surfaces (12,778 images), Agricultural Areas (1,451 images), Forest and Seminatural Areas (1,944 images), Water bodies (80 images), and Wetlands (859 images). This imbalance in class could result in the model being biased towards the dominant class. As a result, several data augmentation techniques were applied during the training phase using the Pytorch 'transforms' module. 'RandomHorizontalFlip' and 'RandomVerticalFlip' were used to randomly flip the images along horizontal and vertical axes, 'RandomRotation' was applied to rotate the image by 20 degrees, and 'ColorJitter' was applied to generate samples with different brightness, contrast, saturation and hue. The random seed value was set to 20.

ResNet50 model is pre-trained on the ImageNet dataset, thus it is crucial to standardize the input dataset to match the data distribution of the pre-trained model. To achieve this, normalization technique was applied to the dataset. 'RandomResizedCrop' was used to randomly crop a portion of the image and resize it to the target size of 224x224 pixels. This also involved scaling the pixel value to a range defined by the mean and standard deviation of the ImageNet dataset. The mean and standard deviation values were set to [0.485, 0.456, 0.406] and [0.229, 0.224, 0.225] respectively[4].

The transformed 2018 dataset was split into a ratio of 60:20:20 as training, validation and test datasets. The 2024 unlabelled image dataset was used only as a test dataset. Data augmentation, resizing, and normalization were all applied to the training dataset, whereas only resizing and normalization were applied to the validation and test datasets. Figure 6 shows the data pre-processing and data transformation workflow.

3.4. Data Modelling

The fourth step consisted of training a deep, multi-layered CNN architecture ResNet50 using transfer learning. 'ResNet' stands for Residual Network and '50' refers to the number of layers in this network. This is a pre-trained model that has been trained on Google servers on the ImageNet dataset. The ImageNet dataset is a large-scale image dataset containing over 14 million images across 1,000 different classes. The ResNet50 architecture is divided into four main components: the convolutional layer, the identity block, the convolutional block and the fully connected layers. The convolutional layers extract features from the input image, while the identity and convolutional blocks process and transform these features. Finally, the fully connected layers are used to make the final classification (Medium, 2023). By leveraging the pre-trained weights from ImageNet, ResNet50 can effectively transfer learned features and improve the model's performance on new tasks. This approach significantly reduces the need for extensive training data and computational resources, making it a powerful tool for image classification tasks. Figure 7 shows the ResNet50 model architecture.

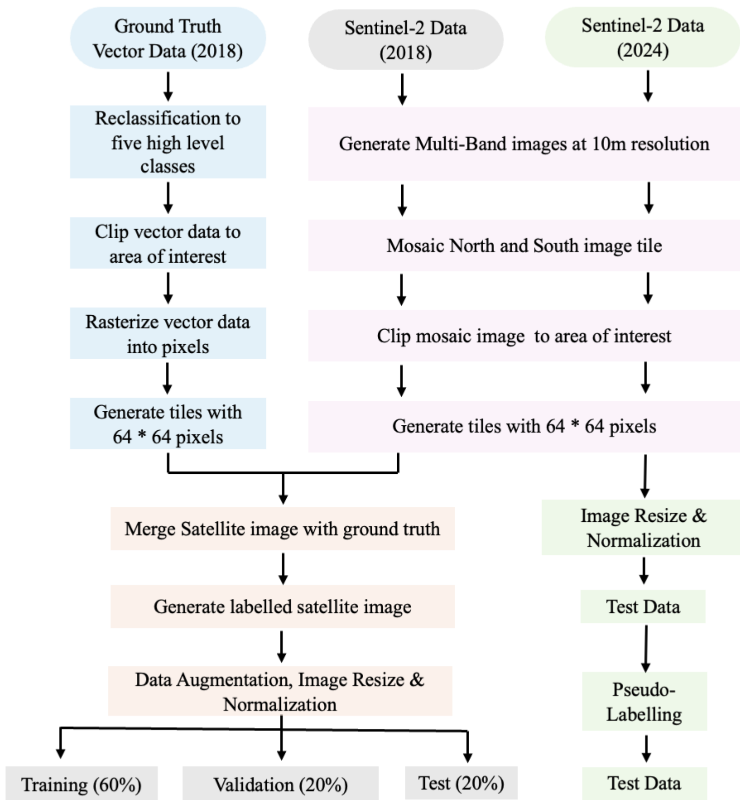

Figure 6: Data pre-processing and transformation workflow for the ground truth and Sentinel-2 datasets

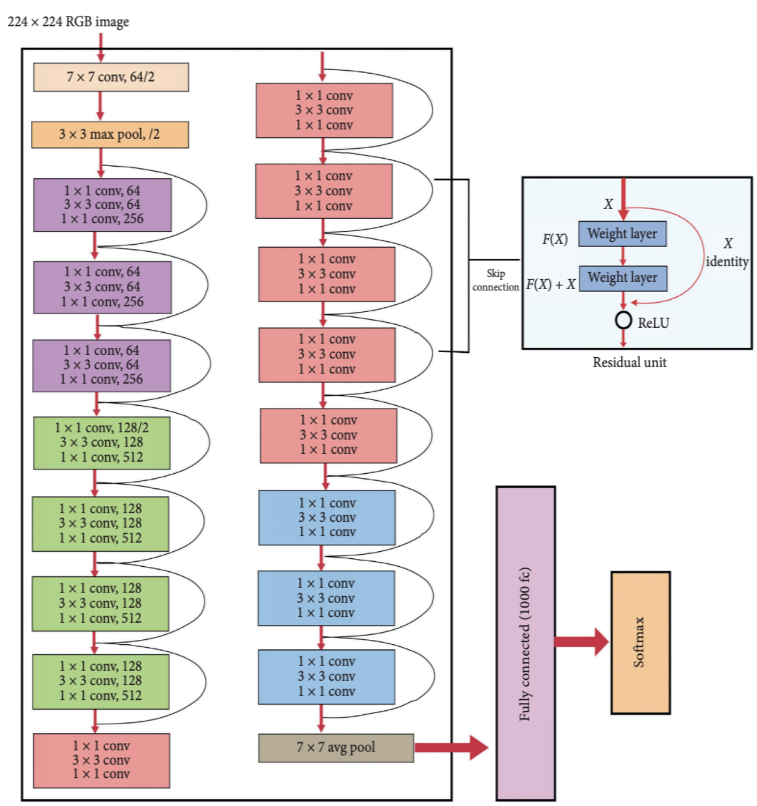

Figure 7: ResNet50 Model Architecture with internal layerwise detail (Shabbir *et al.*, 2021)

3.5. Evaluation

The last step in the research involved assessing the model's performance using key performance metrics and measuring the change in land use and land cover areas for each classification category between the years 2018 and 2024. This step allowed for a detailed comparison of the extent to which different land use types had changed over the specified period and also provided insights into the accuracy of the model's predictions.

4. Design Specifications and Implementation

This section outlines the modelling and evaluation techniques, the architecture utilized to implement these methods, and the associated requirements. Section 4.1 details the modelling techniques used, while Section 4.2 describes the evaluation methods employed to assess the model's performance. Section 4.3 discusses the architecture utilized for the implementation of the discussed techniques, while Section 4.4 discusses the associated requirements.

4.1. Modelling Techniques

This research uses the Convolutional Neural Network (CNN) with the ResNet-50 architecture version 2.0 pre-trained on the ImageNet dataset, which is known for its robustness in image classification tasks (Sharma, Jain & Mishra, 2018). Implemented in Pytorch v2.3.1, the pre-trained ResNet-50 model has 23,518,277 trainable parameters and utilizes residual learning through 50 layers and skip connections to effectively capture complex patterns and features in large-scale image data (He *et al.*, 2016), making it particularly effective for LULC classification compared to other CNN models (Dastour & Hassan, 2023). In this research, the pre-trained model layers were kept open to learn new features from the training dataset. The feature matrices extracted from the convolutional layers were supplied to the fine-tuned fully connected layer, specifically designed to match the number of output classes. The key parameters of the experimental setup, including the activation function, optimizer, learning rate, loss function, and number of epochs, are summarized in Table 1.

Table 1: ResNet50 parameters used in this research

Parameters	Values
Activation Function	Softmax
Optimizer	Stochastic Gradient Descent (SGD)
Learning Rate	0.001
Loss Function	Cross-Entropy
Number of Epochs	10

4.2. Evaluation Techniques

In this study, the ResNET50 model was used for the classification of both labelled and unlabelled datasets, as discussed in Section 3.3. Therefore, depending on the type of the dataset different evaluation techniques were employed to measure the model performance. Section 4.2.1. discusses the evaluation techniques used to measure the model's performance on labelled 2018 data.

Section 4.2.2. discusses the methods of evaluating the model's performance on unlabelled 2024 data. Section 4.2.3. discusses the techniques used for change detection.

4.2.1. Evaluation Techniques for Prediction of Labelled Data (2018)

In this case, the model efficacy was evaluated by analyzing the confusion matrix which detailed the predicted versus true class distribution. Figure 8 shows a 2*2 confusion matrix for a binary classification task.

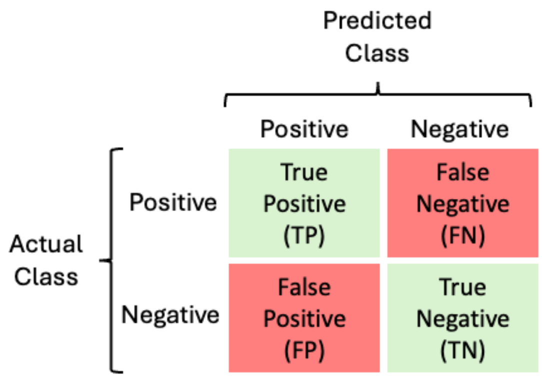

Figure 8: 2 * 2 confusion matrix for the binary classification

This distribution provides insights into the number of correct and incorrect predictions for each class, helping identify patterns of misclassification, and analysing key performance matrices such as overall accuracy (*OA*), class-wise accuracy, precision (*P*), recall (*R*), F-Score and Cohen's Kappa Coefficient.

The overall accuracy (*OA*) metric is a crucial measure for any CNN model as it represents the ratio of the total correct predictions divided by the total number of predictions. Precision (*P*) is the ratio of correctly classified images to the total number of images classified in a particular category, while recall (*R*) is the ratio of correctly classified images to the total number of relevant images present in the database. Furthermore, *F*-Score and Cohen's Kappa coefficient (κ) were also calculated to provide additional insights into the model's performance.nMathematically overall accuracy, precision, recall, F-Score, and κ can be calculated from the equations (1) – (5).

$$\text{Overall Accuracy } (OA) = \frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FP+FN}} \tag{1}$$

$$\text{Precision (P)} = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \tag{2}$$

$$\text{Recall } (R) = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \tag{3}$$

$$F\text{-Score} = \frac{2 \text{ x (Precision * Recall)}}{\text{(Precision + Recall)}} \tag{4}$$

$$\mathrm{k} = \frac{\mathrm{OA - q}}{1 - \theta}; \text{ where } \theta \text{ is the probabilities for each class in the observed data.} \tag{5}$$

Additionally, class-wise performance metrics such as classwise accuracy, precision, recall and F-Score were calculated. This measures the model's performance in correctly predicting instances for each specific class individually which helps to analyze any potential imbalance in prediction accuracy and also assess the model performance not only overall but also at an individual class level.

4.2.2. Evaluation Techniques for Prediction of Unlabelled Data (2024)

To evaluate the model's effectiveness in predicting unlabelled 2024 satellite data, two distinct techniques were employed to evaluate the model's effectiveness in predicting unlabeled data.

1. Manual Verification Technique: This technique involved selecting a small, random subset of images from the dataset and manually labelling them to create a reference set. By comparing the model's predictions for these manually labelled images with the predicted labels, key performance metrics like accuracy and precision can be calculated (Yakimovich *et al.*, 2021). Although this approach is limited, it is practical and provides a method to check the model's overall performance. For this research, 10% of the total images belonging to each class were selected and manually labelled for 2024. After model prediction the predicted class labels were compared to the manually labelled ground truth to evaluate the model's performance.

2. Pseudo Labeling technique: Pseudo labelling is a semi-supervised learning technique where the model's predictions on unlabeled data are used as 'pseudo-labels' to augment the labelled training dataset (Medium, 2022). After the model prediction is completed on the unlabelled dataset, an evaluation is done using the 10% sampling method, as explained in 4.2.2.1. The sampled manually labelled data is then added to the existing 2018 labelled data used to train the model earlier. The model is then re-trained using this new 'pseudo-labelled' augmented dataset. This newly trained model is then used to make the predictions for the unlabelled dataset. These new predictions are re-evaluated using the 10% manual verification technique as explained in 4.2.2.1. Figure 9, shows the workflow for implementing the pseudo-labelling technique for unlabelled dataset.

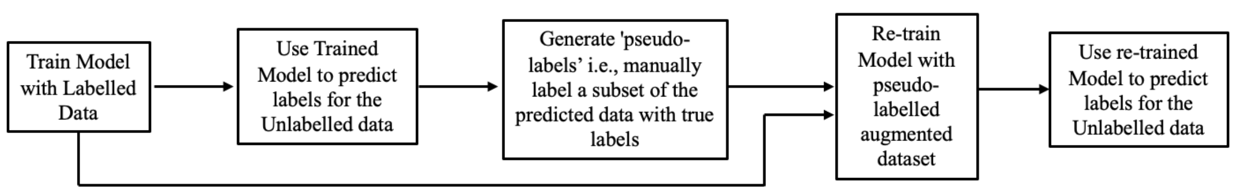

Figure 9: Workflow for implementing 'pseudo-labelling' technique for unlabelled data

4.2.3. Evaluation Techniques for Change Detection

To assess the change in LULC for the area of interest, first the land area covered by five different classes was calculated for the years 2018 and 2024 using Sentinel-2 satellite imagery. The area for each class was then compared by subtracting the 2018 values from the 2024 values to determine the percentage change. Following this technique gives the extent and direction of land cover changes over the six years.

4.3. Utilized Architecture for Implementation

For the implementation of the proposed workflow - data collection, processing, transformation, modelling, and evaluation, this research utilizes various software tools. For data collection, pre-processing, and transformation, this research used the QGIS version 3.28 and ArcGIS Pro version 2.9 software, along with Python GeoPandas (gpd), Rasterio, Pandas and Numpy libraries. The modelling technique is implemented using PyTorch version 2.3.1; while the evaluation was conducted using the Python scikitlearn ('sklearn') library. All the experiments were performed on an Apple MacbookPro with an M1 chip, 8-core CPU, 8 GB RAM, and 512 GB storage, running on macOS Sonoma version 14.5. The Python code was developed and executed within the Jupyter Notebook environment, version 6.4.3. Furthermore, Python libraries 'matplotlib', and 'seaborn' were utilized for visualization purposes.

4.4. Associated Requirements

The associated requirements for this research included accessing data from the official EPA, Ireland, and Copernicus websites. Additionally, UTM was required to run a Windows virtual machine on a Mac to work with the ArcGIS software and GitHub repositories were used to understand ResNET50 architecture.

5. Evaluation and Result Analysis

The following section presents the main findings from the research using the evaluation techniques discussed in Section 4.2. Sections 5.1 and 5.2 focus on the classification results for the years 2018 and 2024. Section 5.3 discusses the change in Land Use and Land Cover for the same period.

5.1. Classification Results with 2018 Satellite Data

The ResNet50 model was trained on a 60% subset and validated on a 20% subset of the augmented 2018 satellite data, as detailed in Section 3.3. After 10 epochs, the model achieved an accuracy of 92.37% on the training dataset and 92.01% on the validation dataset. The training and validation loss were 0.20 and 0.20 respectively. Figures 10a and 10b display the accuracy and loss metrics for the training and validation dataset over 10 epochs.

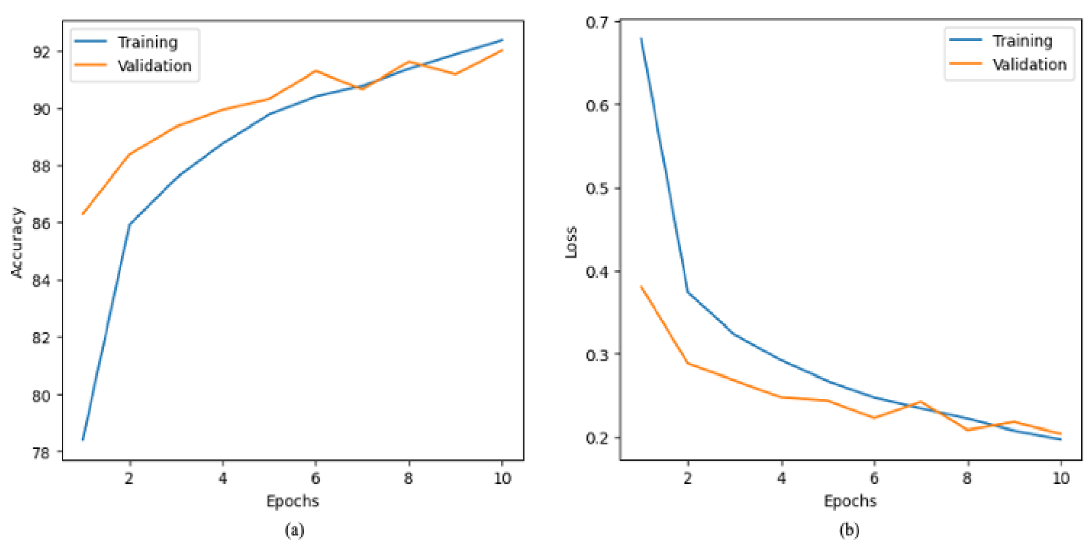

Figure 10: Demonstration of training and validation samples (a) accuracies and (b) the loss values

When the model was evaluated on the test dataset, it showed an overall accuracy of 92.48% and a loss value of 0.19. Additionally, as discussed in Section 4.2.1 the confusion matrix and other performance parameters such as precision, recall, F-score and kappa coefficients were also evaluated. The precision, recall, *F*-score and κ values are 92.41%, 92.51%, 92.45% and 0.91 respectively. Figure 11 shows the confusion matrix for the 2018 image classification across five classes. Table 2 provides a summary of the various performance metrics overall.

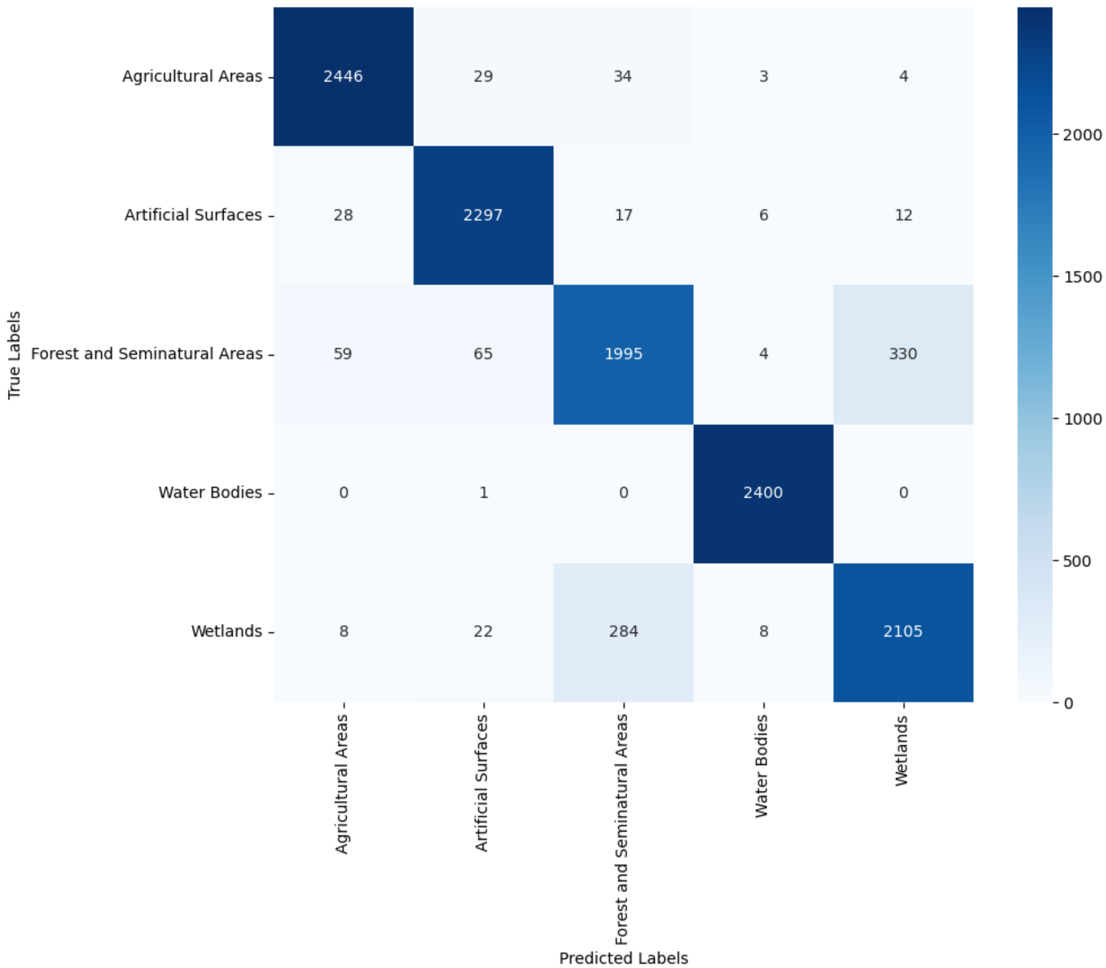

Figure 11: Confusion matrix for 2018 augmented test dataset

Table 2: Summary of model performance on the 2018 test dataset

Performance Metrics	Scores
Overall Accuracy	92.48%
Precision	92.41%
Recall	92.51%
F-Score	92.45%
Cohen's Kappa Coefficient	0.91

The high values of overall accuracy and balanced precision and recall, as shown in Table 2 and Figure 9, demonstrate the strong overall predictive performance of the proposed model. The F-score, which balances precision and recall, is also high, confirming that the model has a good trade-off between avoiding false positives and false negatives. However, for a more detailed evaluation of a multi-class problem, class-wise metrics are needed as they show the percentage of correctly classified samples for each class. Table 3 summarises classwise performance metrics in the labelled 2018 image dataset.

Table 3: Classwise performance for the 2018 test dataset

Class Name	Accuracy (%)	Precision (%)	Recall (%)	F-Score (%)
Agricultural Areas	96.26	96.26	97.22	96.74
Artificial Areas	95.15	95.15	97.33	96.23
Forest and Seminatural Areas	85.62	85.62	81.33	83.42
Water Bodies	99.13	99.13	99.96	99.54
Wetlands	85.88	85.88	86.73	86.31

The classwise performance, as shown in Table 3, further validates the overall effectiveness of the model as indicated in Table 2. The model performs well in classifying Agricultural Areas, Artificial Areas and Water Bodies as it shows high accuracy and balanced precision and recall. However, the performance falls for the prediction of Forest and Seminatural Areas and Wetlands, as the accuracy and F-scores are ~85%, indicating few challenges in correctly classifying these more complex or variable land cover types. The relatively lower recall for these classes also suggests the model may struggle with false negatives in these classes, potentially missing some true instances.

5.2. Classification Results with 2024 Satellite Data

The trained model was used to classify 2024 tiled images into five land cover classes, namely Agricultural Areas, Artificial Surfaces, Forest and Seminatural Areas, Water Bodies, and Wetlands. To assess the model's prediction, the manual verification technique was used, as discussed in Section 4.2.2. A randomly selected 10% subset of this unlabelled data was manually labelled to establish a ground truth for comparison with the model's predictions. Table 4 presents the classwise performance for the model's predictions on the 10% manually labelled dataset before pseudo-labelling.

Table 4: Classwise performance for 10% predicted 2024 dataset before pseudo-labelling

Class Name	Accuracy (%)	Precision (%)	Recall (%)	F-Score (%)
Agricultural Areas	97.85	97.86	74.05	84.31
Artificial Areas	27.41	27.41	86.21	41.61
Forest and Seminatural Areas	42.6	42.61	55.37	48.16
Water Bodies	31.81	31.82	87.5	46.67
Wetlands	51.09	52.01	14.77	23.01

The results show strong performance in predicting the 'Agricultural Areas' class. However, the performance metrics suggest that the model failed to predict the other classes, namely Artificial Surfaces, Forest and Seminatural Areas, Water Bodies, and Wetlands, with the same level of accuracy. Additionally, the results also showed very low values of precision, recall and F-score indicating challenges in the model's ability to consistently and accurately identify these other classes, likely leading to a high rate of misclassification. To address this challenge, the pseudo-labeling technique was applied, and the model's predictions were re-evaluated, as discussed in Section 4.2.2.2. Table 5 presents the classwise performance for the model's predictions after applying the pseudo-labelling.

Table 5: Classwise performance for 10% predicted 2024 dataset after pseudo-labelling

Class Name	Accuracy (%)	Precision (%)	Recall (%)	F-Score (%)
Agricultural Areas	95.73	95.73	96.78	96.26
Artificial Areas	71.25	71.26	89.47	79.33
Forest and Seminatural Areas	78.73	78.72	58.42	67.07
Water Bodies	99.99	99.98	50.01	66.67
Wetlands	62.16	62.16	66.35	64.19

The results indicate that after applying pseudo-labeling, the model's performance showed a significant improvement in predicting the five land use and land cover classes across various metrics. Furthermore, there were fewer misclassifications compared to the initial prediction done on the data without pseudo-labelling, particularly in the more challenging classes of Forest and Seminatural Areas and Wetlands.

5.3. Change Detection

Table 6 and Table 7 present the spatial area covered by the five different classes and the percentage of the area of interest for 2018 and 2024. Table 8 indicates the change in the spatial coverage of the land cover classes.

Table 6: Land cover classes with their spatial coverages and percentages of the area of interest for 2018

Class Name	Spatial Coverage (sq. km)	Percentage of the area of interest
Agricultural Areas	5233.87	74.67%
Artificial Areas	594.33	8.48%
Forest and Seminatural Areas	796.26	11.36%
Water Bodies	32.77	0.47%
Wetlands	351.85	5.02%

Table 7: Land cover classes with their spatial coverages and percentages of the area of interest for 2024

Class Name	Spatial Coverage (sq. km)	Percentage of the area of interest
Agricultural Areas	5276.06	75.27%
Artificial Areas	681.16	9.72%
Forest and Seminatural Areas	571.8	8.16%
Water Bodies	26.21	0.37%
Wetlands	453.84	6.48%

Table 8: Change in spatial coverage of Land cover classes between the years 2018 and 2024

Class Name	Change in Spatial Coverage (sq. km)	Percentage change of the area of interest
Agricultural Areas	42.19	0.60%
Artificial Areas	86.83	1.24%
Forest and Seminatural Areas	-224.46	-3.20%
Water Bodies	-6.56	-0.09%
Wetlands	101.99	1.46%

The results show that in both 2018 and 2024, the 'Agricultural Areas' class represents the largest proportion of spatial coverage in the Greater Dublin Area. In 2018, the Agriculture class covered 74.67% of the total land cover and covered 75.27% of the total land cover in 2024. In 2018 the second largest class was 'Forest and Seminatural Areas' covering 11.36% of the land. There is however a notable shift with 'Artificial Areas' surpassing Forest and Seminatural Areas to become the second-largest class, covering 9.72% of the area.

The results from Table 8 show that 'Agricultural Areas' saw a slight increase of 42.19 km^2, or 0.60% growth. Artificial Areas experienced a more substantial expansion, growing by 86.83 km^2 or 1.24%, indicating increased urban and built environments. This result seems consistent with the data released by Central Statistics Office Ireland, 2023. In contrast, Forest and Seminatural Areas faced a significant reduction, with a decrease of 224.46 km^2 (3.20%), suggesting a notable loss of these natural areas, possibly due to development or land conversion. Water Bodies saw a slight decline of 6.56 km^2 or 0.09%, while Wetlands increased by 101.99 km^2 (1.46%), indicating growth in wetland areas, potentially from conservation efforts or natural processes.

6. Conclusions and Discussion

The aim of this research was the classification of the LULC in the Greater Dublin Area with respect to the five most predominant classes, namely Agricultural Areas, Artificial Surfaces, Forest and Seminatural Areas, Water Bodies, and Wetlands and change detection between the years 2018 and 2024 using ResNET50 deep learning architecture. Results from this research indicated that transfer learning is a reliable approach that can be used for LULC classification using satellite images. The proposed methodology was able to classify the images with an overall accuracy of 92.48 and an average classwise accuracy of 92.40%. Additionally, the pseudo-labelling technique demonstrates enhanced performance and reliability in the model's predictions and could be a favourable approach for LULC and change detection which involves working with unlabelled testing datasets. Lastly, the change detection from this research reflects ongoing shifts in land use and land cover, with notable expansions in urban and wetland areas and reductions in forested regions.

This work can be improved by optimizing the model and supplementing the model training by providing high quantity and quality data having higher inter and intra-class variability. This work can also be improved by comparing the prediction results with ground truth data is crucial for evaluating and validating the model's performance and exploring other techniques for semi-supervised and unsupervised classification, such as advanced algorithms or hybrid approaches. Furthermore, this study can be expanded to include more diverse geographical areas and temporal scales. This could offer a more comprehensive understanding of land cover dynamics and contribute to the development of more robust and adaptable classification systems.

REFERENCES

Alshari, E. A., & Gawali, B. W. (2021). 'Development of classification system for LULC using remote sensing and GIS', *Global Transitions Proceedings*, **2(1)**, pp. 8–17, doi: 10.1016/j.gltp.2021.01.002 [Online: Google Scholar] [Accessed 7 June 2024]

Brennan, M., Tubridy, M. & Associates. (2017). *National Landcover and Habitat Mapping Programme for Ireland*. Published by An Chomhairle Oidhreachta / the Heritage Council 2017. Available at: https://www.heritagecouncil.ie/content/files/Landcover%20and%20Habitat_Mapping.pdf [Accessed 7 June 2024]

Central Statistic Office (2022). *Census of Population 2022 Profile 1 - Population Distribution and Movements*. Available at: https://www.cso.ie/en/releasesandpublications/ep/p-cpp1/censusofpopulation2022profile1-populationdistributionandmovements/populationdistribution/ [Accessed 7 June 2024]

Chugtai, A. H., Abbasi, H., & Karas, I. R. (2021). 'A review on change detection method and accuracy assessment for land use land cover', *Remote Sensing Applications: Society and Environment*, **22**, doi: doi.org/10.1016/j.rsase.2021.100482 [Online: Google Scholar] [Accessed 14 April 2024]

Coppin, P., Jonckheere, I., Nackaerts, K., & Muys, B. (2004). 'Review ArticleDigital change detection methods in ecosystem monitoring: a review', *International Journal of Remote Sensing*, **25(9)**, pp.1565–1596, doi: 10.1080/0143116031000101675 [Online: Google Scholar] [Accessed7 June 2024]

Dastour, H. & Hassan, Q. K. (2023). ' A Comparison of Deep Transfer Learning Methods for Land Use and Land Cover Classification', *Sustainability*, **15(10)**, doi: 10.3390/su15107854 [Online: Google Scholar] [Accessed7 June 2024]

Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; et al. (2011). ' Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services'. *Remote Sensing of Environment*, 120, pp. 25-36, doi: 10.1016/j.rse.2011.11.026 [Online: Google Scholar] [Accessed 8 August 2024]

Environment Protection Agency (2020a). *Land use and land cover*. Available at: https://www.epa.ie/our-services/monitoring--assessment/assessment/irelands-environment/land--soil/current-trends-land-and-soil/#:~:text=Land%20cover%20describes%20what%20is,to%20from%20a%20human%20perspective [Accessed 7 June 2024]

Environment Protection Agency (2023b). *Ireland's Environment: Land and Soil*. Available at: https://www.epa.ie/our-services/monitoring--assessment/assessment/irelands-environment/land--soil/#:~:text=The%20most%20recent%20assessment%20(CORINE,%25)%20and%20forestry%20(9.5%25) [Accessed 7 June 2024]

Environment Protection Agency (2023c). *Land use and land cover*. Available at: https://www.epa.ie/our-services/monitoring--assessment/assessment/irelands-environment/land--soil/current-trends-land-and-soil/ [Accessed 7 June 2024]

Esfandeh, S., Danehkar, A., Salmanmahiny, A., Sadeghi, S. M. M., & Marcu, M. V. (2021). 'Climate Change Risk of Urban Growth and Land Use/Land Cover Conversion: An In-Depth Review of the Recent Research in Iran', *Sustainability*, **14(1)**, pp.338, doi: 10.3390/su14010338 [Online: Google Scholar] [Accessed 7 June 2024]

Garcia, M. C., Gutierrez, J. G., & Riquelme, J.G. (2019). 'A Framework for Evaluating Land Use and Land Cover Classification Using Convolutional Neural Networks', *Remote Sensing for Land Cover/Land Use Mapping at Local and Regional Scales*, **11(3)**, pp. 274, doi: 10.3390/rs11030274 [Online: Google Scholar] [Accessed 7 June 2024]

Gislason, P. O., Benediktsson, J. A., & Sveinsson, J. R. (2006). Random Forests for land cover classification', *Pattern Recognition Letters*, **27(4)**, pp. 294-300, doi: 10.1016/j.patrec.2005.08.011 [Online: Google Scholar] [Accessed 7 July 2024]

Hamidi, S., & Ewing, R. (2014). 'A longitudinal study of changes in urban sprawl between 2000 and 2010 in the United States', *Landscape and Urban Planning,* **128(1)**, pp. 72–82, doi: 10.1016/j.landurbplan.2014.04.021 [Online: Google Scholar] [Accessed 7 June 2024]

He, K.; Zhang, X.; Ren, S.; Sun, J. (2016) 'Deep residual learning for image recognition'. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA*. pp. 770–778. Available: https://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html [Online: Google Scholar] [Accessed 8 August 2024]

Hossen, H., Ibrahim, M. G., Mahmod, W. E., Negm, A., Nadaoka, K., & Saavedra, O. (2018). 'Forecasting future changes in Manzala Lake surface area by considering variations in land use and land cover using remote sensing approach', *Arabian Journal of Geosciences,* **11(5)** [Online: Google Scholar] [Accessed 14 April 2024]

Jagannathan, J., Divya, C. (2021). 'Deep learning for the prediction and classification of land use and land cover changes using deep convolutional neural network', *Ecological Informatics,* **65,** doi: 10.1016/j.ecoinf.2021. 101412 [Online: Google Scholar] [Accessed 14 April 2024]

Khurana, Y., & Gupta, S. (2022). 'LCLU Classification using SVM, MLC and ANN of multispectral imagery from Sentinel-2', doi: 10.17577/IJERTV11IS100056 [Online: Google Scholar] [Accessed 14 April 2024]

Kundu. N. (2023), ' Exploring ResNet50: An In-Depth Look at the Model Architecture and Code Implementation', Medium.com. Available at: https://medium.com/@nitishkundu1993/exploring-resnet50-an-in-depth-look-at-the-model-architecture-and-code-implementation-d8d8fa67e46f [Accessed 3 August 2024]

Kussul, N., Lavreniuk, M., Skakun, S., & Shelestov, A. (2017). 'Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data', *IEEE Geoscience and Remote Sensing Letters*, **14(5)**, pp. 778–782, doi: 10.1109/LGRS.2017.2681128 [Online: Google Scholar] [Accessed 15 April 2024]

Lin, C., & Doyog, N. D. (2023). 'Challenges of Retrieving LULC Information in Rural-Forest Mosaic Landscapes Using Random Forest Technique', *Forests,* **14(4)**, doi: 10.3390/f14040816 [Online: Google Scholar] [Accessed 7 June 2024]

Mahamunkar, G. S. and Netak, L. D. (2022). Comparison of Various Deep CNN Models for Land Use and Land Cover Classification. In: Kim, JH., Singh, M., Khan, J., Tiwary, U.S., Sur, M., Singh, D. (eds) Intelligent Human Computer Interaction. IHCI 2021. Lecture Notes in Computer Science, vol 13184. Springer, Cham, pp 449-510, doi: 10.1007/978-3-030-98404-5_46 [Online: Google Scholar] [Accessed 15 April 2024]

Mochizuki, S., & Murakami, T. (2012). 'Accuracy Comparison Of Land Cover Mapping Using The ObjectOriented Image Classification With Machine Learning Algorithms'. In The 33rd Asian Conference on Remote Sensing. Retrieved: https://pdfs.semanticscholar.org/e161/8c75b6ac4271da2bf8a071c51b674a99f8e1.pdf [Online] [Accessed 7 July 2024]

Naushad, R., Kaur, T., & Ghaderpour, E. (2021). 'Deep transfer learning for land use and land cover classification: A comparative study', *Sensors*, **21(23)**, doi: 10.3390/s21238083 [Online: Google Scholar] [Accessed 15 April 2024]

Noi, P. T., & Kappas, M. (2019). 'Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery', *Remote Sensing for Land Cover/Land Use Mapping at Local and Regional Scales*, **11(3)**, doi: 10.3390/rs11030274 [Online: Google Scholar] [Accessed 14 April 2024]

Phiri, D., & Morgenroth, J. (2017). ' Developments in Landsat Land Cover Classification Methods: A Review', *Remote Sensing,* **9(9)**, doi: 10.3390/rs9090967 [Online: Google Scholar] [Accessed 8 August 2024]

Qingmu, S., Liao, L., & Chen, K. (2021). 'The Impact of Land Use Change on Disaster Risk from the Perspective of Efficiency', *Sustainability*, **13(6)**, doi: 10.3390/su13063151 [Online: Google Scholar] [Accessed 7 June 2024]

Rodriguez-Galiano, V., Ghimire, B., Rogan, J., Chica, M., & Rigol-Sanchez, Juan. (2012). 'An assessment of the effectiveness of a random forest classifier for land-cover classification', *ISPRS Journal of Photogrammetry and Remote Sensing*, **67(1)**, pp. 93–104, doi: 10.1016/j.isprsjprs.2011.11.002 [Online: Google Scholar] [Accessed 7 June 2024]

Rousset, G., Despinoy, M., Schindler, K., & Mangeas, M. (2021). 'Assessment of Deep Learning Techniques for Land Use Land Cover Classification in Southern New Caledonia', *Remote Sensing,* **13(12)**, doi: 10.3390/rs13122257 [Online: Google Scholar] [Accessed 15 April 2024]

Samanta, S., & Paul, S. K. (2016). 'Geospatial analysis of shoreline and land use/land cover changes through remote sensing and GIS techniques', *Modelling Earth Systems and Environment,* **2(3)** [Online: Google Scholar] [Accessed 14 April 2024]

Sampath, V. K., &Radhakrishnan, N. (2022). 'A comparative study of LULC classifiers for analysing the cover management factor and support practice factor in RUSLE model' *Earth Science Informatics,* **16(10)**, doi: 10.1007/s12145-022-00911-7 [Online: Google Scholar] [Accessed 14 April 2024]

Sefrin, O., Riese, F.M., & Keller, S. (2020). 'Deep Learning for Land Cover Change Detection'. *Remote Sensing,* **13(1)**, pp. 78, doi: 10.3390/rs13010078 [Online: Google Scholar] [Accessed 11 June 2024]

Shabbir, A., Ali, N., Ahmed, J., Zafar, B., Rasheed, A., Sajid, M., Ahmed, A. & Dar, S. H. (2021). ' Satellite and Scene Image Classification Based on Transfer Learning and Fine Tuning of ResNet50', *Mathematical Problems in Engineering*, **2021.** doi: 10.1155/2021/5843816 [Online: Google Scholar] [Accessed 8 August 2024]

Sharma, N., Jain, V., & Mishra, A. (2018). ' An Analysis Of Convolutional Neural Networks For Image Classification'. *Procedia Computer Science,* **132**, pp. 377-384, doi: 10.1016/j.procs.2018.05.198 [Online: Google Scholar] [Accessed 8 August 2024]

Shao, Y., & Lunetta, R. S. (2012). 'Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points.' *ISPRS Journal of Photogrammetry and Remote Sensing,* **70**, pp. 78–87. doi: 10.1016/J.ISPRSJPRS.2012.04.001[Online: Google Scholar] [Accessed 7 July 2024]

Shorten, C. & Khoshgoftaar, T. M. (2019). 'A survey on Image Data Augmentation for Deep Learning', *Journal of Big Data*, **6(60),** doi: 10.1186/s40537-019-0197-0 [Online: Google Scholar] [Accessed 3 August 2024]

Srivastava, P. K., Han, D., Rico-Ramirez, M. A., Bray, M., & Islam, T. (2012). 'Selection of classification techniques for land use/land cover change investigation', *Advances in Space Research,* **50(9**), pp. 1250-1265, doi: 10.1016/j.asr.2012.06.032 [Online: Google Scholar] [Accessed 14 April 2024]

Tripathi, D. K., & Kumar, M. (2017). 'Remote sensing based analysis of land use/land cover dynamics in Takula block, Almora district (Uttarakhand)', *Journal of Human Ecology,* **38(3)**, pp. 207-212, doi: 10.1080/09709274.2012.11906489 [Online: Google Scholar] [Accessed 14 April 2024]

Tsang, S. H. (2022), ' Review — Pseudo-Label: The Simple and Efficient Semi-Supervised Learning Method for Deep Neural Networks', Medium.com. Available at: https://sh-tsang.medium.com/review-pseudo-label-the-simple-and-efficient-semi-supervised-learning-method-for-deep-neural-aa11b424ac29 [Accessed 10 August 2024]

Wang, J., Bretz, M., Dewan, M. A. A., & Delavar, M. A. (2022). 'Machine learning in modelling land-use and land cover-change (LULCC): Current status, challenges and prospects', *Science of The Total Environment,* **822(4)**, doi: 10.1016/j.scitotenv.2022.153559 [Online: Google Scholar] [Accessed 7 June 2024]

Yakimovich, A., Beaugnon, A., Huang, Y., Ozkirmli, E. (2021). ' Labels in a haystack: Approaches beyond supervised learning in biomedical applications', *Patterns,* **2(12),** doi: 10.1016/j.patter.2021.100383 [Online: Google Scholar] [Accessed 9 August 2024]

Yan, J., Wang, L., Mu, L., & Huang, L. (2020). 'Knowledge discovery from remote sensing images: A review', *WIREs data mining and knowledge discovery*, **10(5)**, doi: 10.1002/widm.1371 [Online: Google Scholar] [Accessed 7 June 2024]

Yao, Y., Yan, X., Luo, P., Liang, Y., Ren, S., Hu, Y., Han, J., & Guan, Q. (2021). 'Classifying land-use patterns by integrating time-series electricity data and high-spatial resolution remote sensing imagery', *International Journal of Applied Earth Observation and Geoinformation,* **106**, doi: 10.1016/j.jag.2021.102664 [Online: Google Scholar] [Accessed 7 June 2024]