REPUBLIC OF TÜRKİYE

ALTINBAŞ UNIVERSITY

Institute of Graduate Studies

Information Technologies

A DECISION SUPPORT SYSTEM BASED ON MACHINE LEARNING FOR LAND INVESTMENT

Dhufr Hussien Mohammed ALALI

Master's Thesis

Supervisor
Asst. Prof. Dr. Timur İNAN

Istanbul, 2023

A DECISION SUPPORT SYSTEM BASED ON MACHINE LEARNING FOR LAND INVESTMENT

Dhufr Hussien Mohammed ALALI

Information Technologies

Master's Thesis

ALTINBAŞ UNIVERSITY

2023

The thesis titled A DECISION SUPPORT SYSTEM BASED ON MACHINE LEARNING FOR LAND INVESTMENT prepared by DHUFR HUSSIEN ALALI and submitted on 10/08/2023 has been **accepted unanimously** for the degree of Master of Science in Information Technologies.

____

Asst. Prof. Dr. Timur İNAN

Supervisor

Thesis Defence Committee Members:

Asst. Prof. Dr. Timur İNAN	Department of Software Engineering, Altınbaş University	___________________
Asst. Prof. Dr. Mesut ÇEVİK	Department of Electrical and Electronics Engineering, Altınbaş University	___________________
Asst. Prof. Dr. Vedat ESEN	Department of Electrical and Electronics Engineering, İstanbul Topkapı University	___________________

I hereby declare that this thesis meets all format and submission requirements of a Master's Thesis.

Submission date of the thesis to Institute of Graduate Studies: ___/___/___

I hereby declare that all information/data presented in this graduation project has been obtained in full accordance with academic rules and ethical conduct. I also declare all unoriginal materials and conclusions have been cited in the text and all references mentioned in the Reference List have been cited in the text, and vice versa as required by the abovementioned rules and conduct.

Dhufr Hussien Mohammed ALALI

Signature

ABSTRACT

A DECISION SUPPORT SYSTEM BASED ON MACHINE LEARNING FOR LAND INVESTMENT

ALALI, Dhufr Hussien Mohammed
M.Sc., Information Technologies, Altınbaş University
Supervisor: Asst. Prof. Dr. Timur İNAN
Date: August / 2023
Pages: 58

This research proposes a methodology for classifying aerial photographs and extracting investible lands using deep learning (DL) with transfer learning. The study utilizes the Aerial Image Dataset (AID), which contains a diverse set of aerial images with 30 scene classes. The proposed methodology involves data preprocessing, dataset splitting, training image augmentation, model selection, model training, and evaluation using performance measures. Three neural network models (ResNet50, VGG19, and EfficientNetB3) are compared, and the best model is selected based on performance metrics such as precision, sensitivity, F-measure, and the CM. The results observe the effectiveness of the proposed methodology in accurately classifying aerial photographs and identifying investible. This indicates that EfficientNetB3 has a higher ability to classify aerial photographs and extract investible lands compared to ResNet50 and VGG19. ResNet50 achieved moderate performance with relatively lower precision, sensitivity, and F-measure compared to EfficientNetB3. VGG19, on the other hand, demonstrated the lowest performance across all metrics, showing low precision, sensitivity, and F-measure values. These results can contribute to various applications such as urban planning, real estate development, and land management, where accurate classification of aerial images is crucial for decision-making processes. Finally, the future work may involve exploring additional deep.

Keywords: ML, DL, Transfer Learning, Aerial Photograph Classification, Investment Land.

TABLE OF CONTENTS

LIST OF TABLES

Pages

LIST OF FIGURES

Pages

ABBREVIATIONS

AI	:	Artificial Intelligence
DL	:	Deep Learning
ML	:	Machine Learning
AID	:	Aerial Image Dataset
ResNet	:	Residual Network
CNN	:	Convolutional Neural Network
RELU	:	Rectified Linear Unit
VGG	:	Visual Geometry Group
ACC	:	Accuracy

1. INTRODUCTION

1.1 BACKGROUND

World is full of sites and lands valid, there is a wide range of locations for investing and now the world attaches great importance to investment in line with the growth of global economy [1]. Land investment refers to the purchase of land with the intention of holding it for future appreciation or development. It can be a good investment because it is a finite resource and demand for it can increase over time [2]. However, it is also a risky investment because it is not easily liquidated and can be subject to market fluctuations and environmental factors [3]. Land used for residential and commercial development, agriculture and animal grazing, mineral production, and recreation are typical sorts of land investments [2]. There are several factors to consider when investing in land. These include the location of the land, its size and accessibility, the quality of the land, and any potential development opportunities. It is also important to research the local market conditions, zoning laws, and other regulations that may affect the value of the land [4]. When deciding to invest, prospective landowners should be informed of the unique types of land-related investment opportunities offered through investment products [5]. Traditionally, aerial image analysis has relied on manual interpretation by experts, which is time-consuming and prone to subjectivity. With the advancements in DL techniques, particularly in the field of computer vision, automated approaches for aerial image classification and land extraction have become possible [6].

DL models, for instance (CNNs), have shown themselves as remarkable performance in image recognition tasks, leading to increased ACC and efficiency in aerial image analysis [7]. Aerial image analysis plays a crucial role in various domains, including urban planning, agriculture, environmental monitoring, and infrastructure development. The ability to classify aerial photographs and identify investible lands accurately and efficiently is of great importance in supporting decision-making processes [8]. Traditional methods for aerial image analysis often rely on manual interpretation and human expertise, which can be time-consuming, subjective, and prone to errors [6]. In recent years, DL techniques, particularly transfer learning, have shown great potential in automating the analysis of aerial images. By leveraging pre-trained models and large-scale datasets, DL models can learn complex attributes and patterns from aerial photographs, leading to improved classification ACC and efficiency [9, 10]. To address these challenges, this thesis proposes the use of DL techniques, specifically transfer learning, for classifying aerial photographs and extracting investible lands. DL has shown remarkable success in various image analysis tasks and has the potential to automate and improve the ACC of aerial image recognition. The proposed methodology involves multiple steps, which includes data pre-processing, model selection, model training, transfer learning parameter fine-tuning, and testing and performance evaluation.

1.1.1 Risks and Rewards of the Investment

Investment pertains to acquiring assets at a cost with the anticipation that their future value and returns will surpass both the initial expenditure and the time taken to realize that value [11]. This occurrence happens due to appreciation of the asset's value and the commitment of current financial resources in order to produce higher extensive gains in the future [12]. The information that might potentially assist create a vision of the various degrees of confidence in the status of investment in the future is crucial since it deals with what are known as uncertainty domains [13].

Real assets are frequently employed in investment to generate goods or services. Examples of these assets include buildings, land, machinery, or intellectual property that play a role in the production of commodities or services [14].

For now, investment process pushes in all fields, investors used modern techniques, which will lead to economic and social development, important to develop, because of its usefulness in bringing technical and practical expertise, Fostering investment and facilitating the establishment process contribute to the cultivation of human resources and the generation of employment opportunities, expanding and developing investment projects at various levels [13].

The objective behind making investment decisions is to attain optimal returns through skillful allocation of suitable financial resources to appropriate opportunities. These decisions factor in two crucial variables of financial management: risks and returns [12]. Investors used many techniques to seek suitable land like the advice of a financial advisor or real estate professional, Google earth or computer techniques which is solve many problems [15].

In addition to these factors, it is important to do thorough research on the land itself before making an investment. This can include examining the topography and quality of the land, as well as any environmental or ecological issues that may impact its value [15]. Investing in commercial and residential land development presents a practical avenue for entering the investment market, thanks to the abundance of opportunities available. These opportunities can be tailored to align with an investor's specific capital and time constraints [15].

1.1.2 Importance of Investment

Within organizations, investment choices play a pivotal role in fostering growth and enhancing profitability, carrying lasting implications as a number of these choices are irreversible. Even with constraints on available funds, individuals can achieve substantial benefits when investments are meticulously strategized and efficiently executed. Prior assessment of associated risks is imperative for managers, allowing them to avert potential losses [14].

Investors go for the most suitable assets or investment opportunities based on a variety of factors: risk profiles, investment objectives, and return expectations [16].

1.2 PROBLEM STATEMENT

The classification of aerial photographs and extraction of investible lands using DL techniques poses challenges in terms of efficiency and ACC. Common methods for aerial image analysis rely on manual interpretation, which is time-consuming, subjective, and prone to errors. The existing techniques for investment decision-making in land investment lack automated and objective approaches for identifying investment opportunities. The current research aims to address these challenges by proposing a methodology that leverages DL techniques, specifically transfer learning, to classify aerial photographs and extract investible lands. However, there is a need to further investigate and evaluate the proposed methodology to determine its effectiveness in accurately identifying investible lands. Additionally, there is a need to assess the performance of different DL models, such as ResNet50, VGG-19, and EfficientNetB3, with the aim of identifying the best model for the classification task. The study also aims to analyse the performance metrics, consisting of sensitivity, precision, F-measure, additionally the CM, to evaluate the models' ACC and generalization ability. By addressing these research gaps, this study aims to contribute to the field of aerial image analysis and support investment decision-making processes in land investment.

1.3 OBJECTIVES OF THE RESEARCH

This thesis' primary goal is to produce an accurate and robust methodology for classifying aerial photographs and extracting investible lands using DL techniques. The proposed methodology aims to leverage the power of transfer learning, data pre-processing, and model evaluation measures to achieve high ACC and generalization ability in aerial image analysis.

The specific goals of this study include:

a. Pre-process the Aerial Image Dataset (AID) by resizing the images and converting categorical class labels to numerical form then Apply data augmentation techniques to the training images to enhance the variety and volume of training data.

b. Select and compare three neural network models, namely ResNet50, VGG19, and EfficientNetB3, for classifying aerial photographs and determining the best model and Train the selected models using an optimizer, specified learning rate, and loss function.

c. Evaluate the trained models using performance measures such as precision, sensitivity, F-measure, and a CM to select the best model based on validation loss and implement early stopping to prevent over fitting.

d. Test the performance of the selected model on unseen data using precision, sensitivity, F-measure metrics, and the CM and discussing the results, analyse to findings, and provide insights into the ACC, generalization ability, and capability of the models in recognizing investible lands.

1.4 CONTRIBUTIONS OF THE THESIS

This thesis' contributions may be summed up as follows:

a. Proposed methodology for aerial image analysis: the thesis proposes a methodology for classifying aerial photographs and extracting investible lands using DL techniques, specifically transfer learning. The methodology consists of several essential steps, including data pre-processing, model selection, and model training, transfer learning parameter fine-tuning, and testing and performance evaluation. This contribution provides a systematic approach for automating the analysis of aerial images and supporting investment decision-making processes.

b. Identification of risks and rewards in land investment: the thesis discusses the risks and rewards associated with land investment. It highlights the importance of considering factors such as location, size, accessibility, quality, market conditions, and regulations when making investment decisions. This contribution provides valuable insights for potential investors in land and helps them make informed choices based on the potential risks and rewards involved.

c. Use of DL techniques for aerial image analysis: the thesis emphasizes the use of DL techniques, particularly transfer learning, for aerial image analysis. It explains how DL models, such as convolutional neural networks (CNNs), can be leveraged to classify aerial photographs and identify investible lands accurately and efficiently. This contribution focus on the potential of DL in automating and improving the ACC of aerial image recognition.

d. Evaluation of different DL models: the thesis compares and evaluates the performance of three DL models (resnet50, vgg19, and efficientnetb3) for the classification of aerial photographs and extraction of investible lands. It analyses performance metrics, including precision, sensitivity, F-measure, and the confusion matrix, to assess the ACC and generalization ability of the models. This contribution helps in identifying the best model for the given task and provides insights into the performance of different DL architectures.

e. Results and discussions: the thesis shows the results of the experiments conducted using the proposed methodology and discusses the findings. It analyses the performance of the efficientnetb3 model in terms of ACC, loss ratios, confusion matrix (CM), precision, sensitivity, and F-measure for each scene class. This contribution provides empirical proof of the effectuality of the proposed methodology and DL models in classifying aerial photographs and extracting investible lands.

Overall, the thesis contributes to the field of aerial image analysis and investment decision-making in land investment by proposing a methodology, identifying risks and rewards, utilizing DL techniques, evaluating different models, and presenting empirical results. These contributions enhance the understanding and application of automated approaches for analysing aerial images and supporting investment decisions in the context of land investment.

1.5 STRUCTURE OF THE THESIS

The thesis is put in order into 5 chapters, and the contents of each are described below.

Chapter 2: The literature review chapter presents a comprehensive review of existing research and studies related to aerial image recognition, DL techniques, and transfer learning. It explores the current state of the art, identifies research gaps, and establishes the theoretical foundation for the proposed methodology.

Chapter 3: Aerial Image Recognition Methodology introduces the proposed methodology for classifying aerial photographs and extracting investible lands using DL techniques. It discusses the steps involved in the methodology, which includes data pre-processing, model selection, model training, transfer learning parameter fine-tuning, and testing and performance evaluation.

Chapter 4: Results and Discussions of the experiments conducted using the proposed methodology are presented and analysed. The performance of different DL models in classifying aerial photographs and extracting investible lands is evaluated by employ performance metrics for instance precision, sensitivity, F-measure, and CM. The findings are discussed, and their implications are explored.

Chapter 5: Conclusion and future work the final chapter concludes the thesis by summarizing the main findings, highlighting the contributions of the study, and discussing future research directions. It also reflects on the limitations of the study and suggests areas for further improvement and exploration.

2. LITERATURE REVIEWS AND THEORETICAL BACKGROUND

2.1 INTRODUCTION

Throughout the chapter, a presentation of exhaustive examination of the current body of literature is demonstrated and the pertinent theoretical foundation in the domain of the use of land and covers land classification, demonstrating machine learning methodologies. The primary objective is to examine various approaches, algorithms, and models proposed and demonstrated in land use and land cover classification, additionally with a specific focus on methodologies related to machine learning. In recent years, the utilization of machine learning methods for the classification of land use and land cover has attracted significant attention. This interest stems from the potential of machine learning to automate and enhance the precision of classification processes. Machine learning algorithms have the capability to extract meaningful patterns and relationships from large and complex datasets, thereby facilitating the recognition & categorization of different land use and land cover classes [17]. This chapter begins with an overview of the literature on covering real estate investment advising using machine learning, where different techniques for instance random forest regression, linear regression using gradient descent, and k-nearest neighbour regression have been compared. The results indicate that the RF algorithm performs more advanced than the other examined algorithms. Subsequently, several studies focusing on land cover mapping and classification are discussed. These studies employ different approaches such as object-based image analysis, DL frameworks, ensemble models, and fusion of multispectral and panchromatic images. The ACC of the classification results varies across these studies, with some achieving high ACC rates using algorithms like random forest, support vector machine, and artificial neural networks. Furthermore, this chapter reviews studies that investigate the forecast mapping of land cover forecasting differences using ML techniques. These studies utilize datasets obtained from satellite imagery, and algorithms such as the Markov model, random forest classifier, and land change modeller. The results highlight the ability of ML models to effectively predict along with mapping out future land use and land cover changes. Additionally, this chapter includes studies that explore the impact of corporate social responsibility (CSR) performance on the efficiency of investments using ML classifiers. These studies employ boosting, RF, and neural network algorithms to classify investment effectiveness. The findings suggest that CSR staging has a major impact on investment efficiency, and ML models can accurately classify investment effectiveness. Moreover, the chapter discusses studies that focus on the classification of images that specified with remote sensing images using machine learning models. Finally, this chapter presents a comparative and experimental study that evaluates different machine learning architectures for land use and land cover image classification. Models such as, VGG, ResNet, and EfficientNet have been examined, and their performance metrics, including ACC, have been compared. The results highlight the effectiveness of certain architectures in achieving high classification ACC.

2.2 LITERATURE REVIEWS

In recent years, the corporation of (ML) and (DL) techniques has brought about significant advancements in the field of land use and land cover classification. These techniques have revolutionized the analysis of satellite imagery, enabling a better understanding and monitoring of land attributes. Through the application of artificial intelligence, researchers have successfully extracted valuable information from satellite data, contributing to domains like real estate investment advising, land cover mapping, land use classification, and spatio-temporal dynamics monitoring. Numerous studies have been conducted to explore the application of machine learning algorithms specifically within this field of land use and land cover classification. These studies have paved the way for improved accuracy and efficiency in analysing and interpreting satellite imagery for land use and land cover assessment.

According to [18] researchers developed an advanced automated artificial neural network system specifically designed for land use/land cover classification using Landsat TM imagery. The system utilized both supervised Multilayer Perceptron (MLP) neural networks and unsupervised Kohonen's Self-Organizing Mapping (SOM) methods.

[19] examined the (SVM), (MLE), and Artificial Neural Network (ANN) methods while analysing land cover using satellite photos. They discovered that SVM worked better for the study of land cover.

In a study conducted by[20], Real Estate Investment Advising Using ML was explored. The authors compared three ML algorithms, namely Random Forest regression, Linear Regression using gradient descent, and K nearest neighbour regression. They utilized quarterly data from 2005 to 2016, obtained through web scraping. The findings indicated that the Random Forest algorithm had the lowest error compared to the other algorithms examined.

[21] Proposed DL methods, specifically CNN, for make classifying on land cover and land use. Later then a rule-based decision fusion approach (MLP-CNN) that integrated pixel-based MLP classifiers and contextual-based CNN with intense architectures. Their approach produced higher ACC (around 90%) compared to existing DL and traditional approaches.

[22] Focuses on integrating and fusing multispectral Landsat OLI photos to create the best method for classifying land use. The study compared classifications made on multispectral and fused pictures using various methods. According to the findings, the merging of multispectral and panchromatic pictures utilizing the maximum likelihood classifier (MLC) technique and the algorithm that's based off of principal components achieved by the highest accuracy (ACC).

[23] Proposed a land-use classification method using semantic attributes and machine learning. They employed three machine learning algorithms, namely Support Vector Machine (SVM), Multilayer Perceptron, and K-Nearest Neighbour (K-NN), along with a majority vote for classification fusion. The study utilized the UC Merced land use dataset (UC Merced). The ACC results for SVM, ANN, KNN, and C-MV were found to be 91.82%, 89.2%, 79.30%, and 95.10%, respectively.

[24] Focused on concentrating and charting land cover use changes using the CA-Markov synergy model. The study conducted a case study in Erbil, Iraq, using three consecutive-year Landsat imagery. The classifier of maximum likelihood was used for classification, and the ACC of agreements in the middle of the classified and modelled maps were undertaken to be assessed.

[25] Used TerrSet's Land Change Modeller (LCM) to track and forecast future land usage and land cover. The study used Markov Chain along with artificial neural network (ANN) analysis in LCM and Landsat satellite pictures from three distinct years. The ACC of development was determined to be 88%, 92%, and 93%, respectively, for the years 2007, 2014, and 2017.

[26] Proposed an object_ based multi_ temporal with multi_ source land cover tracing approach. The authors utilized a DL framework based on a modified focus mechanism and Repeating Neural Network (RNN). The framework leveraged the complementarity of radar and optical (SITS) data. The classification outcomes demonstrated a worldwide accuracy (ACC) of 79% for the Reunion Island more over 90% for the Senegalese site.

[17] The study evaluated and studied over six machine learning algorithms for land-use and land-cover classification based on satellite observations. The classifiers assessed were SVM, RF, MLP, Fuzzy Adaptive Resonance Theory-Supervised Predictive Mapping, Spectral Angle Mapper, and Mahalanobis Distance. The authors conducted a comparison of these algorithms, revealing that the Random Forest algorithm achieved the highest overall accuracy (ACC) at 0.89, while the Mahalanobis Distance algorithm yielded the lowest ACC, registering a value of 0.82.

[27] Discussed land use and land cover cartography using advanced ML classifiers, consist of (SVM) and (RF). They also introduced two DL algorithms based on (MLP) with Genetic Algorithm (GA) optimization and derivative-free optimization. The MLP-based algorithms produced high accurate precisions, with over 98% correct classification.

[28] Investigated the performance of the impact of CSR on the efficiency of investments using ML classifiers. They compared two basic models of investment effectiveness measurement and found that Richardson's method yielded better results. They employed boosting and Random Forest algorithms for classification, both achieving error rates of 0%. Additionally, their implementation of neural networks resulted in error rates of 4.29% for the first method of Richardson, 12.79% for the Biddle method, and finally 2.01% for the last method.

[29] Proposed a land use classification approach using an ensemble model combining (XGBoost) Extreme Gradient Boosting classification and (SVM) Support Vector Machine. They used multispectral Landsat-8 (OLI) sensor data and classified it into seven land use classes. The ensemble classifier model demonstrated higher precision compared to other ML classifiers.

[30] Assessed DL techniques for classify land use and land cover by getting SPOT6 satellite data. They compared the classification ACC of DL with XGBoost and found that

incorporating the output of a dedicated DL architecture significantly improved the ACC, reaching 63.61%.

[31] A machine learning method for tracking and predicting spatio-temporal performances of land cover in Bangladesh's Cox's Bazar neighbourhood. The study used the Random Forest classifier and satellite photographs receiving it from Landsat 4-5 TM dataset and Landsat 8 OLI/TIRS dataset. The classification ACC ranged from 89.04% to 95.24%, and forecasting ACC using RF regression ranged from 87% to 99.8%.

[32] Proposed an investment classification approach for remote sensing images. They utilized convolutional channels to extract attributes and trained a (SVM) classifier. They also compared classic DL models such as VGG 16, ResNet 50, and DenseNet 169. The results showed that the ResNets 50 model produced the highest overall ACC (OA) of 0.924 when trained with the UC Merced land use dataset and 0.956 with the SIRI-WHU dataset.

[33] Focused on land use feature detection from satellite imagery using machine learning models. They employed four approaches, including Random Forest, XGBoost, U-Net, and Artificial Neural Network (ANN), with Landsat 8 data. The ANN model produced an ACC score of 95.9% and performed well with similar data to the training data.

[34] Focusing on the land use and land cover by using the Classification along with the Regression Tree (CART) technique, a machine learning algorithm. They produced a classification ACC of 92.9% with a Kappa coefficient of 0.88, suggesting CART as a suitable classifier for LULC.

[35] Proposed a hybrid DL architecture for land use and land cover images classification. They compared the Modified MobileNet V1 (MMN) architecture with other models like MobileNet V1, VGG 16, DenseNet 201, and ResNet 50. The ResNet 50 model produced an ACC of 97.59%, outperforming the other models.

[36] The research was centred on the automated mining of land cover indicators from satellite image using deep learning (DL) techniques. They trained a CNN Inception-V3 model through transfer learning, achieving an ACC of 98.18% on the EuroSAT dataset.

2.3 ARTIFICIAL INTELLIGENCE

Artificial intelligence (AI) defines the advancement of computer systems capable of performing tasks that have historically relied on human intelligence. These tasks encompass functions for instance visual perception, speech recognition; investment decision-making, further more language translation [37]. Applications for AI are found in many sectors, including healthcare, banking, and transportation. As technology advances and more data becomes available for analysis, the potential uses for AI are expanding exponentially. However, there are also concerns about the ethical implications of developing machines that can think and learn like humans. Human intelligence can simulate it by the development and use of algorithms, software, and other advanced technologies that enable machines to learn from experience and perform tasks that belong to human intelligence [38].

AI has grown significantly and powerfully in its ability to close the gap between machine and human capabilities. Researchers and fans alike focus on various facets of the subject to produce outstanding results. Computer Vision defines as one of the many such fields within the field of AI [39].

2.4 MACHINE LEARNING

ML is one of the many fields of computer science that studies algorithms to analyse datasets for simulating solutions to solve complex problems in an inclusive way that are hard to program using conventional programming methods, it helps humans work effectively and more efficiently [37, 40].

The entire concept of ML is understanding the layout of data and imbed the data within models that can be simple and easy to be understood for people, however it differentiates from the traditional computational approaches. In traditional way, computers use a set of detailed programmed instructions in order to calculate for problem solving [41]. ML algorithms enables the computers to learn from input data also expected reference output data and generate logic which is then deployed into production. ML also improves them by finding patterns in the database without any human interventions or actions [42]. The algorithms of ML continuously upgrade their work due to the fact the number of samples obtainable for learning increases. Using various programming techniques, ML algorithms are able to process large amount of data and extract useful information. ML algorithms are used to examine the extensive amounts volumes of data [41].

2.5 DL

A subclass of machine learning, deep learning (DL) transforms low-level qualities into relevant high-level attributes that may be utilized for recognition. DL has delivered some of the most notable advances in ML to date. Due of the conversion methods used by DL, multiple-layer artificial neural networks are typically used [43].

The goal of this field is to make it possible for machines to perceive the world similarly to how people do, to understand it in a similar way, and to even use this understanding for a variety of tasks, for instance image and video recognition, logical language processing, recommendation systems, image analysis and classify it, media recreation, etc. The improvements in Computer Vision using DL have largely evolved over one specific method, a CNN, throughout time [44].

2.5.1 Convolutional Neural Networks (CNNs)

(CNNs), differ to other forms of Artificial Neural Network [45], have emerged as powerful DL algorithms for image recognition and classification tasks. These networks are particularly made to use convolutional layers to automatically learn and extract key properties from photos [41]. By applying learnable filters to input images, CNNs create feature maps that highlight significant patterns and edges within the images. This approach has led to remarkable advancements in computer vision applications, consisting of image classification, object detection, facial recognition, and medical image analysis [46]. The widespread adoption of CNNs is expected to continue expanding as these algorithms evolve and find applications in various industries. The architecture of a CNNs is inspired by the connectivity patterns of neurons in the human brain, particularly the visual cortex. Just like individual neurons in the brain respond to stimuli within a limited region called the receptive field, CNNs employ collections of such fields that overlap to cover all over the entire visual area [47]. By using the proper filters, this design enables Convolutional Neural Networks (CNNs) to comprehend spatial and temporal correlations inside images. Additionally, CNNs offer advantages in terms of reduced pre-processing requirements and the ability to learn filters and characteristics automatically, making them highly effective in processing complex images with pixel dependencies based on spatial distances [46]. While performing class prediction on straightforward binary images, the technique might yield an average precision score. However, when tackling intricate images involving pixel interdependencies across the entire composition, the resulting accuracy (ACC) could be significantly reduced or even negligible. Through the use of appropriate filters, a ConvNet, the field of visual recognition's successful transition from engineering attributes to building the architectures' ConvNet, may depict the links between time and space in a picture [44, 48]. The architecture demonstrates improved performance in fitting image datasets by reducing the number of parameters involved and enabling the reuse of weights. Consequently, the network can be trained to better comprehend the intricacies of the images. The ConvNet plays a crucial role in transforming the images into a more manageable format for processing, while preserving essential attributes necessary for accurate predictions [48]. This aspect is particularly significant when designing an architecture that excels not only in feature learning but also in scalability for large-scale datasets [48]. To achieve this, the ConvNet employs two approaches: one in which the dimensionality of the attributes is reduced compared to the input, and another in which the dimensionality either remains the same or is increased. This is accomplished through the application of Valid Padding in the former case or Half (same) Padding in the latter case [49]. On the other hand, the performing of the same operation but with no padding, will be presented with a matrix that has dimensions of the Kernel itself which it is mean a Valid Padding [49].

A typical CNN architecture includes input and output layers, as well as hidden layers such as convolutional, pooling, RELU, commonly shortened from (Rectified Linear Unit) [45], and fully connected layers. These layers of the architecture and algorithms perform attributes extraction and categorizing [50]. The attributes extraction phase involves a series of extracting relevant attributes from input images, whereas the classification phase utilizes fully connected layers to learn non-linear combinations of high-level attributes and make accurate predictions [46]. Input and output layers, as well as several hidden layers, make up a CNN. Convolutional, pooling, RELU, and fully connected layers are among the hidden layers. Generally, convolutional graphs have two phases, the feature extraction phase and the classification phase [50], as depicted in Figure 2.1.

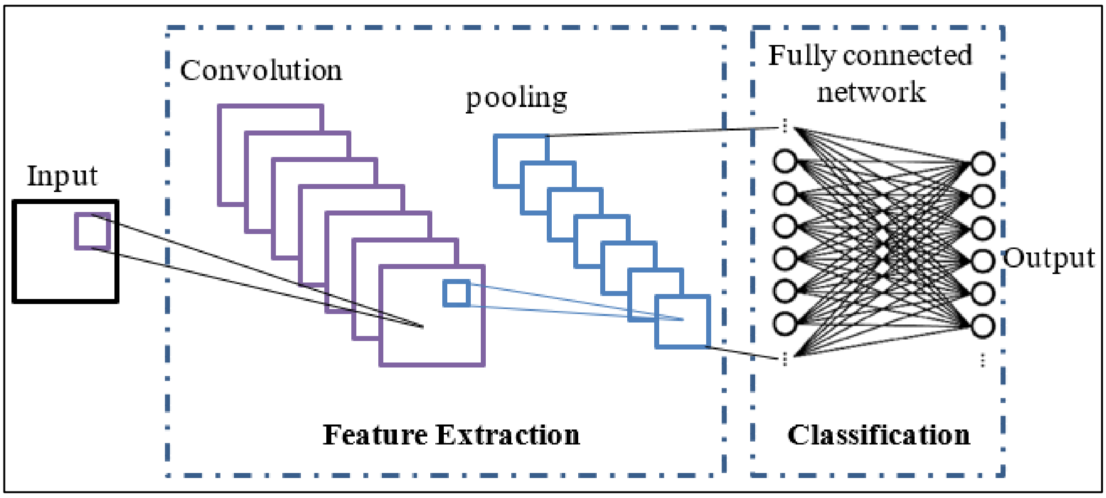

Figure 2.1: Hierarchy of Convolutional Neural Networks [44].

The input picture is altered during training so that it is in a format that the Multi-Level Perceptron can understand. A feed-forward neural network receives the picture after it has been flattened into a column vector [10]. Back propagation is applied during training to optimize the network's weights and biases. The Softmax Classification technique is often used in the final layer to distinguish dominant attributes and classify them based on learned patterns [51]. Overall, CNNs have revolutionized image analysis by effectively capturing complex visual patterns and enabling accurate image recognition and classification. With further advancements and scalability, CNNs are expected to continue playing a significant role in various industries and applications [45].

2.5.2 Visual Geometry Group (VGG-19)

In 2014, S. Karen and Z. Andrew unveiled the Convolutional Neural Network (CNN) architecture known as VGG 19. The 19-layer network topology of the model is denoted by the designation "VGG-19". It comprises of a series of convolutional layers, then numerous fully linked layers and pooling layers [6]. It is used for a wide variety of computer vision implementations, for example semantic segmentation, furthermore object identification, and picture classification. VGG19 is recognized for its efficiency and simplicity [30]. Notably, VGG19 employs small 3x3 convolutional filters and a deep network architecture, enabling the model to learn robust attributes for image classification. It was trained on the ImageNet dataset, comprising numerous images from different categories, and it produced state-of-the-art performance in the 2014 ImageNet challenge [52]. VGG19 is widely recognized as a benchmark for evaluating the performance of other image classification models and holds a prominent position in the sector of computer vision. VGG19 is a CNN constructor and architecture that consists of 19 layers, including 16 convolutional layers and 3 fully connected layers [53]. See Figure 2.2.

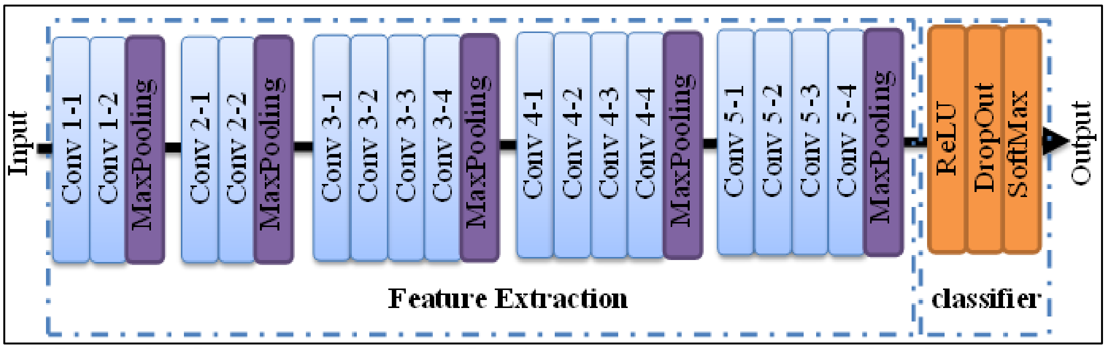

Figure 2.2: Structure of VGG19 [53].

With 143 million parameters, VGG19 is one of the largest and most complex models within the VGG family. The concept was first presented in the publication under name "Very Deep Convolutional Networks for Large-Scale Image Recognition" by Simonyan and Zelezny. Numerous computer vision tasks, for instance segmentation, object identification, and picture classification, have found extensive use for VGG19 [54]. Notably, it won the 2014 ImageNet Challenge's first place in the localization challenge and second place in the classification task. VGG19 has also been extensively utilized in transfer learning, where the model's pre-trained weights are fine-tuned on a smaller dataset to enhance ACC and performance for specific tasks [6]. Its deep architecture primarily consists of 3x3 convolutional filters, known for their effectiveness in capturing local attributes of images. Additionally, VGG19 incorporates a layer which is max pooling after each convolutional layer, reducing the spatial dimensions of feature maps and facilitating the extraction of more abstract attributes from input images [55].

2.5.3 Residual Network (ResNet50)

ResNet50, released in the 2015 essay "Deep Residual Learning for Image Recognition" by Kaiming He, Xiangyu Zhang, furthermore Shaoqing Ren, and Jian Sun, is a deep residual network. It belongs to the family of ResNets, which are renowned for their depth and effectiveness in computer vision tasks look alike image classification and additionally object detection. The ResNet50 architecture comprises 50 layers and is composed of multiple residual blocks, along with maximum pooling and average pooling layers. Each residual block contains several convolutional and batch normalization layers, followed by a ReLU activation function [56]. These residual blocks are specifically designed to enable the network to learn the residual mapping between the input and the desired output, rather than directly learning the entire mapping. This approach and architecture helps mitigate the vanishing gradient problem that can arise in very deep networks [57]. In the ResNet architecture, a building block is depicted in Figure 2.3, demonstrating its role in handling deeper layers.

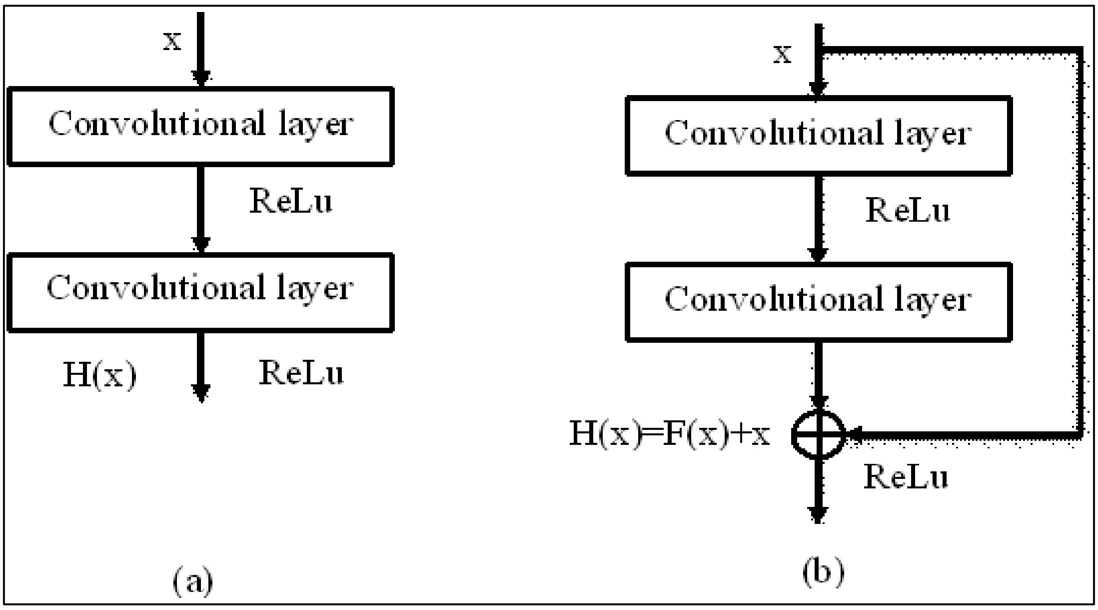

Figure 2.3: The Residual Learning Between (a) a Plain Network and (b) a Residual Network [57].

ResNet50 has demonstrated exceptional performance in numerous computer vision tasks and has attained state-of-the-art results on multiple benchmark datasets. It has emerged as a prominent backbone network for various computer vision models and has gained significant popularity in the domain of transfer learning. In transfer learning, the pre-trained weights of ResNet50 are utilized and fine-tuned on smaller datasets tailored to specific tasks [58]. The pre-trained ResNet50 neural network is capable of classifying images into 1000 object categories. It has been trained on a large-scale dataset and exhibits remarkable generalization capabilities. The neural network expects image inputs with dimensions of 224-by-224 pixels [58].

2.5.4 EfficientNet (B3)

EfficientNetB3 is a member of the EfficientNet texture, a group of CNN construction and architectures introduced by M. Tan and Q. Le in their paper "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," developed by Google Research [42]. This network architecture employs a novel scaling method that uniformly scales all dimensions of the network, including resolution, depth, and width [18]. EfficientNetB3 has been specifically utilized for document image classification [24]. The performance of EfficientNet models benefits from scaling each dimension using appropriate coefficients. However, the notable improvement in performance occurs when the network characteristics are balanced with respect to available computational resources. This approach allows EfficientNet models to produce a favorable balance between ACC and computational efficiency [42]. The network architecture of EfficientNet is depicted in Figure 2.4.

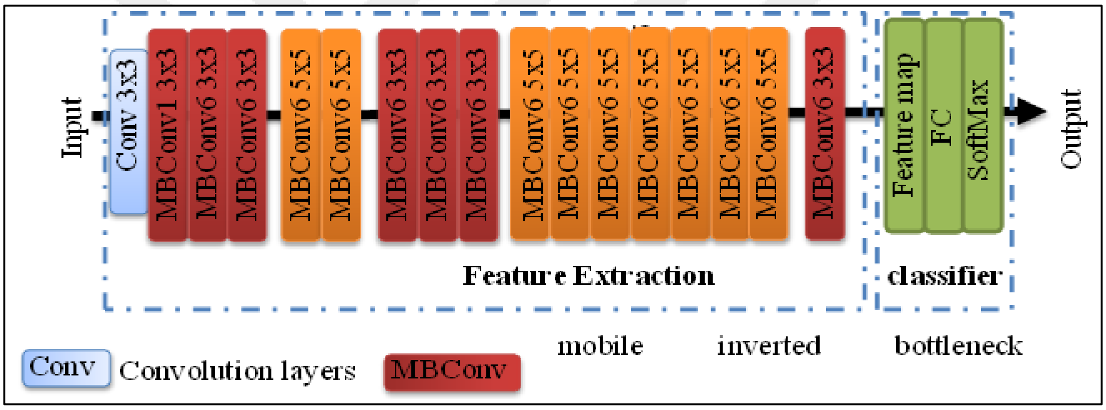

Figure 2.4: EfficientNet Architecture [59].

EfficientNet models stand out by discovering suitable scaling coefficients for each dimension through the aforementioned strategy. When compared to other existing CNNs on the ImageNet dataset, EfficientNets typically exhibit a significant reduction in parameter size [59].

3. AERIAL IMAGE RECOGNITION METHODOLOGY

3.1 INTRODUCTION

Throughout this chapter, a discussion of the proposed methodology for classifying aerial photographs and extracting investible lands using DL techniques. Data pre-processing is one of the processes in the approach, model selection, model training, transfer learning parameter fine-tuning, and testing and performance evaluation. The study proposes the use of DL (transfer learning) to classify aerial photographs and extract investible lands. The proposed methodology starts with the data pre-processing step, which includes resizing the images from the Aerial Image Dataset (AID) to a smaller amount of size of 224 x 224 pixels. This resizing helps in reducing computational complexity and training time without significant loss of visual information. The categorical class labels in the AID dataset are converted to numerical form using OneHotEncoder, which assigns a unique binary value to each class label. Next, the dataset is split into three groups: a training group, a test group, and a validation group. The training group, which contains 80% of the data, is used to train the DL models. The test and validation groups, each group containing 10% of the divided data, are used in order to evaluate the models' performance on the invisible and embedded data. To increase the diversity and quantity of training data, data augmentation techniques are applied to the training images. Augmentation involves applying random transformations such as rotations, translations, flips, and zooms to create modified versions of the original images. This process helps in reducing over fitting and improves the models' ability to generalize to new images. The proposed methodology selects three neural network models for comparison and determination of the best model: ResNet50, VGG19, and EfficientNetB3. For each model, a Sequential model is created with various layers, including a flat layer, fully connected layers, furthermore a last layer with a function named sigmoid activation for classification. The models are applied by compiling an optimizer with a specified learning rate and a loss function. During the model training process, the models are evaluated based by using performance measures for instance precision, sensitivity, F-measure, and a CM. Based on the validation loss, the best model is selected, and early stopping is implemented to avert over fitting and ensure the model exploits well to invisible and embedded data. The best model is saved throughout the training process. After training the models, the next step is testing and performance evaluation. The models are trained by evaluating on the test set, which consists of 20% of the pre-processed AID dataset that was not used in the training stage. The models' performance is measured using precision, sensitivity, F-measure metrics, and the CM. These metrics provide insights into the models' ACC, generalization ability, and ability to recognize investible lands. Overall, the proposed methodology leverages DL techniques, transfer learning, data pre-processing, and model evaluation measures to classify aerial photographs and identify investible lands. By following this methodology, the study aims to develop accurate and robust models for analysing aerial images and supporting investment decision-making.

3.2 PROPOSED METHODOLOGY

This study proposes the use of DL (transfer learning) to classify aerial photographs and extract investible lands. The proposed methodology includes using the aerial image dataset for experiments and determining the effectiveness of imparted learning in classifying aerial images. After calling the data set, pre-processing takes place, which includes dividing the data, converting categorical data into numbers using the concept of (OneHotEncoder), and changing the size of the images. Followed by dividing the data into three groups, a training group, a test group, and a validation group. The concept of generating images (Augmentation) is applied to expand and increase the number of aerial images for make training, in addition to creating changes in them to get rid of the case of over fitting. This process applies to training images only. Then the selected neural network models (ResNet50, VGG19, and EfficientNetB3) are created for the purpose of comparison and determination of the best model. A Sequential model is generated and the various layers, including the flat layer, the layers with fully connected, furthermore the last layer with a sigmoid transition function are added to each network for the final classification. The model is also compiled using an optimizer with learning rate and loss function specified. The best model is saved throughout the training process based on the validation loss. Early Stopping is used for early stopping if no validation loss improvements occur for a specified number of cases. Finally, the models are evaluated using performance measures (precision, sensitivity, F-score and CM), as shown in Figure 3.1 which is represent the proposed methodology of the whole system.

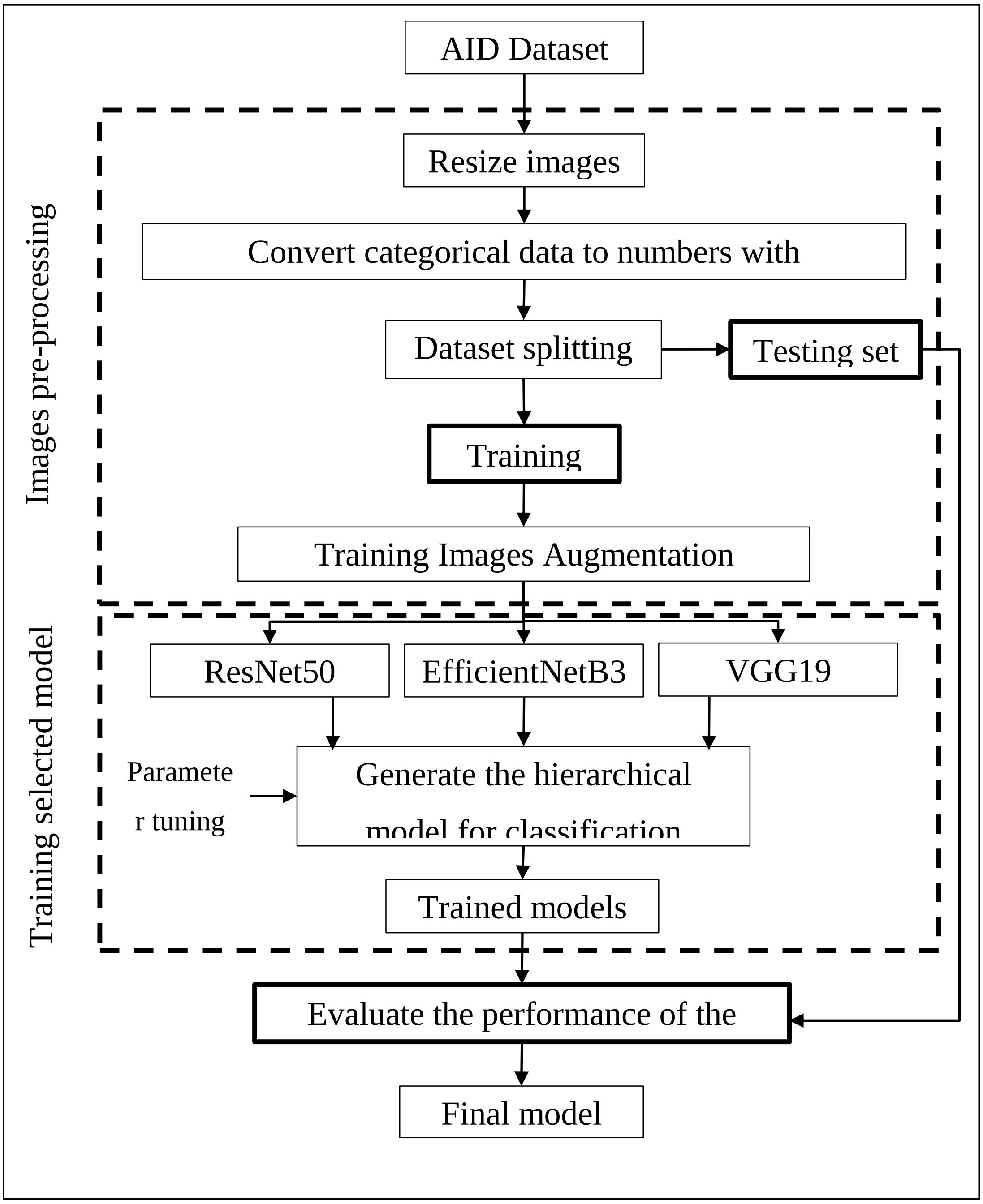

Figure 3.1: Proposed Methodology.

3.3 AERIAL IMAGE DATASET (AID)

The Aerial Image Dataset (AID) is a large-scale dataset of aerial images with 30 scene classes, each containing 200-400 images. The images were collected from Google Earth and Bing Maps, and cover a diverse set of geographic locations and environmental conditions. Each image in the AID dataset is a colour image with a resolution of 600 x 600 pixels, and is provided in JPEG format. The scenes in the dataset include airports, farmland, forest, industrial areas, meadows, mountains, and residential areas, among others. The dataset is designed for training and evaluating algorithms for aerial image analysis tasks for instance scene categorization, object recognition/detection, finally semantic segmentation. The AID dataset has been used in several research studies, and has proven to be a useful resource for developing and testing image analysis algorithms for aerial imagery. Figure 3.2 show images from AID.

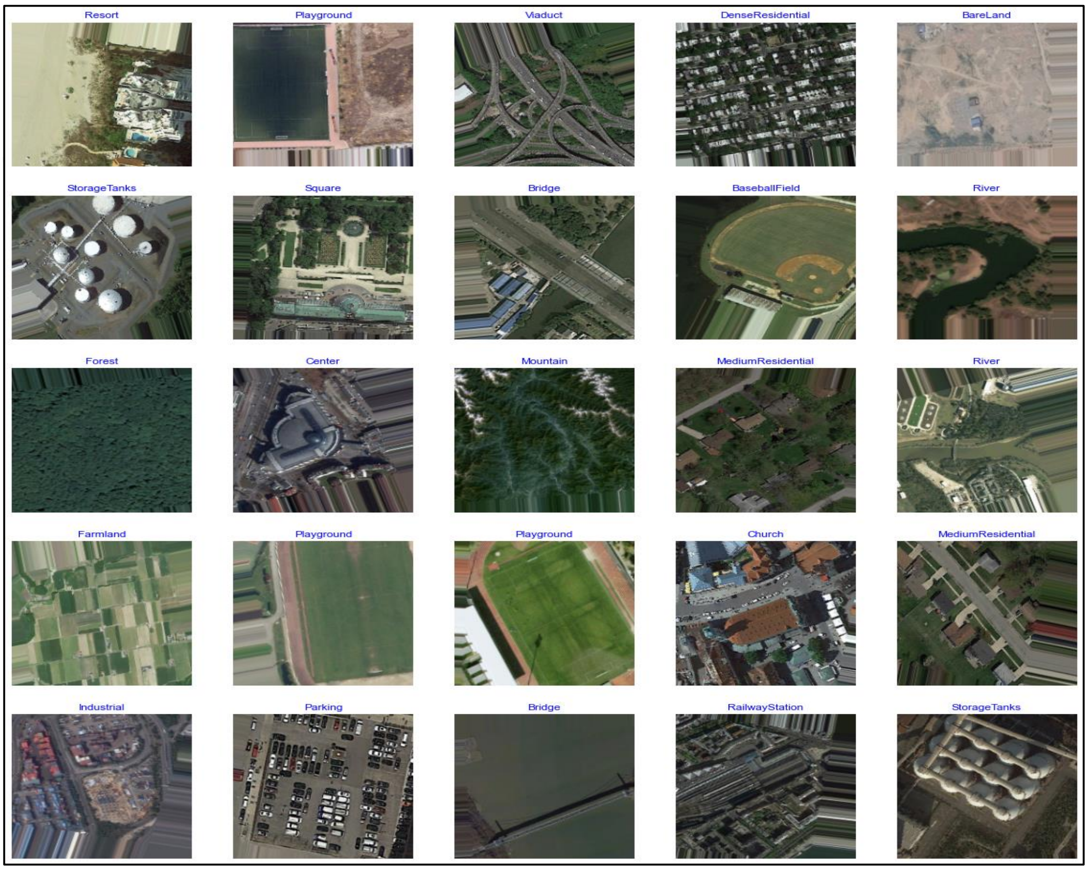

Figure 3.2: Aerial Image Dataset (AID).

Table 3.1: Count of Scene Images per Class.

CLASS	IMAGE COUNT	CLASS	IMAGE COUNT	CLASS	IMAGE COUNT
Airport	324	Farmland	333	Port	342
BareLand	279	Forest	225	RailwayStation	234
BaseballField	198	Industrial	351	Resort	261
Beach	360	Meadow	252	River	369
Bridge	324	MediumResidential	261	School	270
Center	234	Mountain	306	SparseResidential	270
Church	216	Park	315	Square	297
Commercial	315	Parking	351	Stadium	261
DenseResidential	369	Playground	333	StorageTanks	324
Desert	270	Pond	378	Viaduct	378
Industrial	351	Forest	225	Farmland	333

3.3 DATA PREPROCESSING

In this section, will be discussed the data pre-processing steps performed on the Aerial Image Dataset (AID) before training the DL models.

3.3.1 Resize Images

The original images in the AID dataset have a resolution of 600 x 600 pixels. To reduce computational complexity and training time, the aerial images are resized to a smaller size of 224 x 224 pixels. This resizing step helps in reducing the input dimensions of the aerial images without losing significant visual information.

3.3.2 Convert Categorical Data to Numbers with OneHotEncoder

The AID dataset includes scene classes such as airports, farmland, forest, industrial areas, meadows, mountains, and residential areas, among others. These class labels are categorical data and need to be converted into numerical form for training the DL models. OneHotEncoder is used to perform this conversion, which assigns a unique binary value to each class label. This process allows the models to interpret and learn from the categorical information during training.

3.3.3 Dataset Splitting

To evaluate the performance of the trained models, the AID dataset is divided into three groups: a training group, a test group, and a validation group. The dataset is split in an 80:20 ratios, where 80% of the data is used for training, and the remaining 20% is evenly divided between the test and validation sets. The training set is used to train the DL models, while the test and validation sets are used to evaluate the models' performance on embed and invisible data.

3.3.4 Training Images Augmentation

To increase the diversity and quantity of training data, data augmentation techniques are applied to the training images only. Augmentation involves creating advanced versions of the original imagery by applying sorts of transformations like rotations, making translations, flips, and finally zooms. This process helps in reducing over fitting and improves the model's ability to generalize to new, unseen aerial images. By applying augmentation, the models can learn from a larger variety of image variations and become more robust.

These pre-processing steps ensure that the AID dataset is appropriately prepared for training the DL models. The resized images, numerical class labels, and augmented training data are used in the subsequent stages of the proposed methodology for aerial image classification and identification of investible lands. Table3.2 displays the selected parameters and values and their description.

Table 3.2: Images Augmentation Parameters.

Parameters name	**Value**	**Description**
horizontal_flip	True	This parameter determines whether random horizontal flips should be applied to the images. Flipping an image horizontally means reversing the pixels from left to right.
rotation_range	20	It specifies the range of random rotations that can be applied to the images. In this case, the range is set to 20 degrees, which means that each image can be rotated by a random angle between -20 and +20 degrees.
width_shift_range	0.2	This parameter determines the range of horizontal shifts that can be applied to the images. It is expressed as a fraction of the total image width. In this case, the range is set to 0.2, which means that each image can be horizontally shifted by a random fraction of up to 20% of the image width.
height_shift_range	0.2	Similar to width_shift_range, this parameter determines the range of vertical shifts that can be applied to the images. It is expressed as a fraction of the total image height. In this case, the range is set to 0.2, allowing for vertical shifts of up to 20% of the image height.

Table 3.2: Images Augmentation Parameters “Table Continued”.

zoom_range	0.2	This parameter specifies the range of random zooming that can be applied to the images. A zoom factor of less than 1.0 zooms out the aerial image, while a zoom factor greater than 1.0 zooms in. In this case, the range is set to 0.2, which means that each image can be randomly zoomed by a factor between 0.8 and 1.2

By applying these augmentation techniques with the specified parameter values, the training images are modified in various ways, such as being flipped, rotated, shifted horizontally and vertically, and zoomed in or out. These modifications expand and increase the diversity and quantity of the training data, helping the DL models to generalize better and reduce over fitting.

3.4 MODEL SELECTION

In this section, we will discuss the selection of neural network models for classifying aerial photographs and extracting investible lands. The following models were chosen for comparison and determination of the best model:

3.4.1 ResNet50

ResNet is a popular DL model architecture that has been proven effective in various computer vision tasks. It consists of residual blocks, which allow for the training of very deep networks without the problem of vanishing gradients. ResNet50 is a specific variant of the ResNet architecture that has 50 layers.

3.4.2 VGG19

VGG (Visual Geometry Group) is another widely used convolutional neural network architecture. It is known for its simplicity and effectiveness. VGG19 is a variant of the VGG architecture that has 19 layers.

3.4.3 EfficientNetB3

EfficientNet is an application of CNN construction and architectures working together that have produced the upmost state-of-the-art performances on image classifying tasks while maintaining a relatively low computational cost. EfficientNetB3 is a specific variant of the EfficientNet architecture.

3.4.4 Model Training

For each model, a Sequential model is created. The models consist of various layers, including a flat layer, fully connected layers, and a last layer with a sigmoid activation function for the final classification. These models make compilation for utilizing an optimizer with a specified learning rate and a loss function. During the training process, the models are evaluated based on performance measures such as precision, sensitivity, F-score, and a CM. The model that comes out with the best performance is the final model for classifying aerial photographs and extracting investible lands. To ensure the best model is chosen, early stopping is implemented. Early stopping allows the training to be stopped if there is no improvement in the validation loss for a specified number of epochs. This helps to stop over fitting and that the model generalizes well to hidden data. Throughout the training process, the best model is saved based on the validation loss. This allows us to retrieve the best model for further evaluation and use.

3.4.5 Transfer Learning Parameter Fine-Tune

Transfer learning one of a technique that commonly comes out with ML to leverage pre-trained models and adapt them to new tasks or datasets. When fine-tuning a pre-trained model, several parameters come into play to optimize the training process and produce better performance. The parameters used as follows:

a. Epochs this parameter determines the total number of training courses or iterations. Each epoch represents one complete pass of the training dataset through the model. Within this case, the value is programmed to 40, indicating that the training process will go through the dataset 40 times.

b. Ask_Epoch this parameter specifies the number of courses or epochs between each time a question is asked. It refers to the frequency at which the user is prompted to provide an intervention or input during the training process. With a value of 5, a question will be asked every 5 epochs to gather user input.

c. ASK the ASK object is used as a criterion for training pauses. It is defined as ASK (model, epochs, ask_epoch), where model refers to the model being trained. The ASK mechanism prompts the user with a question after a certain number of training sessions to gather feedback or perform any necessary interventions.

d. ReduceLROnPlateau this object belongs to tf.keras.callbacks and is used to dynamically adjust the learning r*ate during training. It reduces the learning rate when a validation loss is detected for a specified number of consecutive training cycles. The monitor parameter is set to "val_loss" to monitor the validation loss. The factor parameter determines the reduction factor, the patience parameter sets the number of cycles to wait before reducing the learning rate, and verbose controls the verbosity of the output.

e. EarlyStopping this object, also belonging to tf.keras.callbacks, allows for early stopping of the training process if there is no improvement in the validation loss for a specified number of consecutive training cycles. The monitor parameter is set to "val_loss" to monitor the validation loss. The patience parameter determines the number of cycles the loss can deteriorate before early stopping. The verbose parameter controls the verbosity of the output, and restore_best_weights ensures that the best weights produced during training are restored.

By utilizing these parameters and the call-backs in the call-backs list, the training process can be optimized by adjusting the learning rate, stopping early if necessary, and incorporating user feedback through the ASK mechanism. This helps to enhance the performance of the fine-tuned model.

3.5 TESTING AND MODELS PERFORMANCE EVALUATION STAGE

After training the transfer learning models for lands suitable for investment recognition, the next step is to test the performance of these models. In this stage, the trained models will be evaluated on test set to measure their ACC and generalization ability. For the lands suitable for investment recognition model, the test set consisted of 20% of the pre-processed AID dataset that was not used in the training stage. The test set then feeds into the trained models, the performance of the models will be measured using (precision, sensitivity, and F-measure) metrics. The CM is also used to visualize the results of the classification. Overall, the testing stage is crucial to ensure that the trained models are robust and accurate in recognizing lands suitable for investment. The results obtained from the testing stage can also be used to fine-tune the models and improve their performance. The following presumptions used to better explain the algorithm measurement tools.

a. TP stands for True Positives.

b. TN stands for True Negatives

c. FP stands for False Positives.

d. FN stands for False Negatives.

3.5.1 ACC

It is the correct guess (TP, TN) divided by every outcome (T means that the prediction is right in the prediction results of P and N).

The formula for ACC is:

$$Acuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{3.1}$$

When the dataset is unbalanced, evaluating the model's performance based on ACC is impossible. That is the case when there are significantly more positive samples than negative ones. When dealing with unstable datasets, precision is a more trustworthy indicator.

3.5.2 Precision

In all classes where the forecast is a good example, it utilized the influence of the ACC of the prediction as a positive example.

$$Precision = \frac{TP}{TP+FP} \tag{3.2}$$

3.5.3 Sensitivity

Do positive instances divide a coverage metric that indicates how many positive instances are correctly identified as positive (TP) are correctly identified as positive (TP), including counting how many positive samples that incorrectly classified as negative (FN)

$$Recall\ = \frac{TP}{TP+FN} \tag{3.3}$$

3.5.4 F-1 Score

It is frequently practicable to combine the ACC and sensitivity rate into a single index F-1 value, particularly when a straightforward way of comparing the performance of two classifiers is required. The F-1 value is the harmonic average of ACC and Sensitivity.

$$F1 = 2 * \frac{Precision*recall}{precision+recall} \tag{3.4}$$

3.5.5 Confusion Matrix

The performance of the classification algorithm is evaluated using a CM. The simplest means of regulating the classification model's performance are the proportions of positive and negative examples that are correctly classified as true or false. The CM is shown in Figures 3.3.

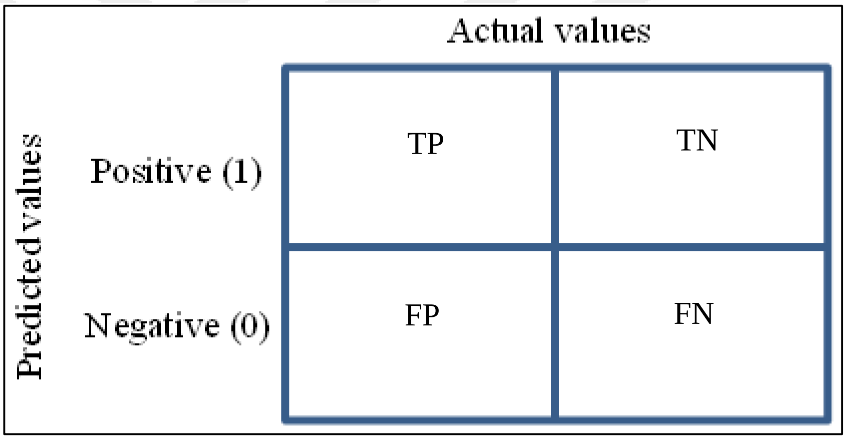

Figure 3.3: CM.

4. RESULT AND DISCUSSIONS

4.1 INTRODUCTION

This chapter presents the results and discussions of the experiments conducted using the proposed methodology for classifying aerial photographs and extracting investible lands. The chapter begins with an overview of the dataset used in the experiments and the performance metrics employed for evaluating the models. Subsequently, the results of the experiments are presented and analysed, followed by discussions on the findings and implications of the study.

4.2 DATASET AND PERFORMANCE METRICS

The experiments were conducted using the Aerial Image Dataset (AID), which consists of aerial images categorized into 30 scene classes. The sets of data were split to three groups: a training group, a validation group, and a test group. The training group contained 80% of the data, while the validation and test groups each contained 10% of the data. For evaluating the performance of the trained models, several performance metrics were employed, including precision, sensitivity, F-measure, and the CM. Precision measures the ratio of correctly classified positive circumstances out of all instances predicted as positive. Sensitivity, also referred to as sensitivity, measures the proportion of correctly classified affirmative instances out of all actual positive instances. The F-measure is the harmonic mean of precision and sensitivity, allowing a balanced scale performance for the actual model. The CM presents a tabular representation of the classification results, showing the number of instances classified right or wrong for each class.

4.3 RESULTS

The experiments were conducted using three neural network models: ResNet50, VGG19, and EfficientNetB3. The models were trained using the training group of the AID dataset, and their performance was evaluated using the validation group and the test group.

4.3.1 Performance of EfficientNetB3

The performance of the EfficientNetB3 model is analysed and presented in this section. Figure 4.1 illustrates the ACC and loss ratios of the EfficientNetB3 model during training and validation.

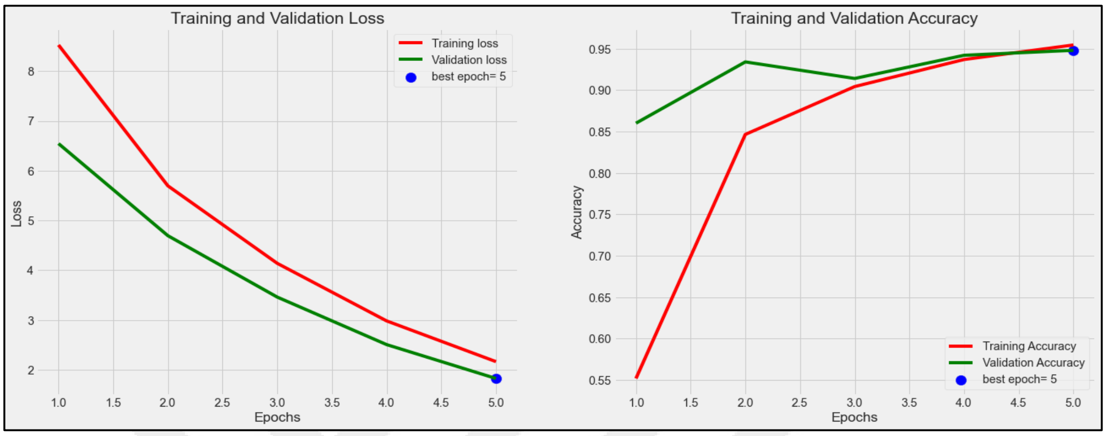

Figure 4.1: ACC and Losses Ratios of Efficientnetb3 Model.

Figure 4.2 depicts the CM of the EfficientNetB3 model, providing a visual representation of the classification results. Additionally, Table 4.1 presents the precision, sensitivity, F-measure, and support values for each scene class.

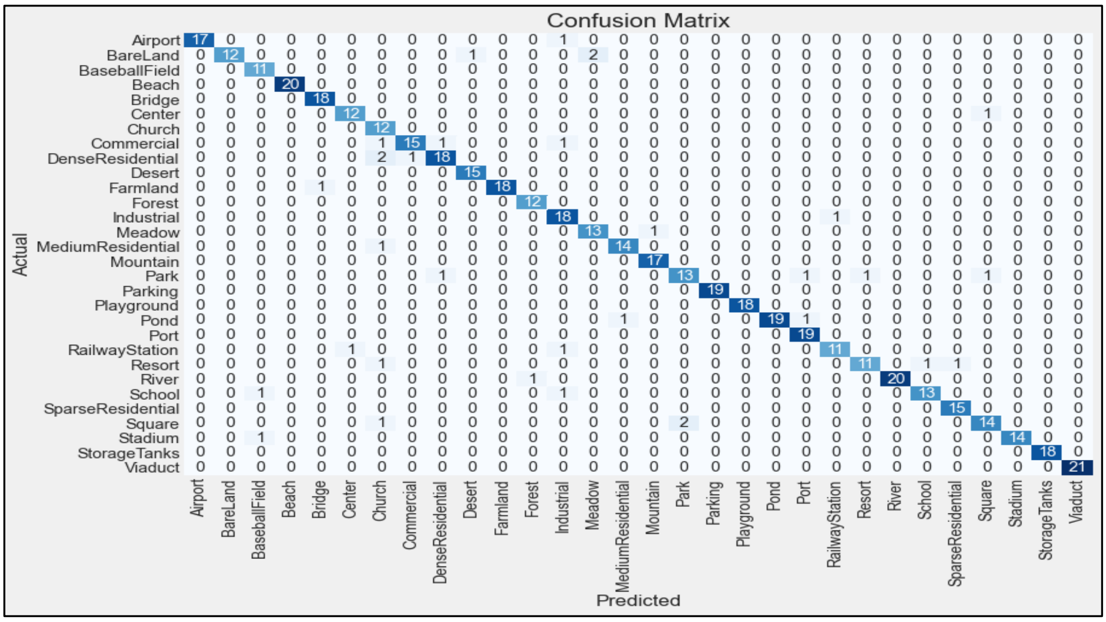

Figure 4.2: CM of EfficientNetB3 Model.

Table 4.1: EfficientNetB3 Model Results.

Class	Precision	sensitivity	F-measure	support
Airport	1	0.9444	0.9714	18
BareLand	1	0.8	0.8889	15
BaseballField	0.8462	1	0.9167	11
Beach	1	1	1	20
Bridge	0.9474	1	0.973	18
Center	0.9231	0.9231	0.9231	13
Church	0.6667	1	0.8	12
Commercial	0.9375	0.8333	0.8824	18
DenseResidential	0.9	0.8571	0.878	21
Desert	0.9375	1	0.9677	15
Farmland	1	0.9474	0.973	19
Forest	0.9231	1	0.96	12
Industrial	0.8182	0.9474	0.878	19
Meadow	0.8667	0.9286	0.8966	14
MediumResidential	0.9333	0.9333	0.9333	15
Mountain	0.9444	1	0.9714	17
Park	0.8667	0.7647	0.8125	17
Parking	1	1	1	19
Playground	1	1	1	18
Pond	1	0.9048	0.95	21
Port	0.9048	1	0.95	19
RailwayStation	0.9167	0.8462	0.88	13

Table 4.1: EfficientNetB3 Model Results “Table Continued”.

Resort	**0.9167**	**0.7857**	**0.8462**	**14**
River	1	0.9524	0.9756	21
School	0.9286	0.8667	0.8966	15
SparseResidential	0.9375	1	0.9677	15
Square	0.875	0.8235	0.8485	17
Stadium	1	0.9333	0.9655	15
StorageTanks	1	1	1	18
Viaduct	1	1	1	21
ACC			0.934	500
macro avg	0.933	0.9331	0.9302	500
weighted avg	0.9391	0.934	0.9341	500

From Table 4.1 the following can be extracted according to each measure:

Precision: For most classes, the model produces high precision with values that ranges from 0.6667 to 1. This demonstrates that the model has a low rate of false positives, correctly classifying a high proportion of instances predicted as positive.

Sensitivity: The model exhibits good sensitivity for most classes, with values ranging from 0.7647 to 1. This shows that the model efficiently captures a high ratio of actual positive instances.

F-measure: The F-measure, which combines precision and sensitivity into a single metric, shows balanced performance for most classes, with values ranging from 0.8125 to 1. A high F-measure indicates that the model produces both high precision and sensitivity simultaneously.

Overall ACC: The model produces an overall ACC of 0.934, indicating that it correctly classifies 93.4% of the instances in the test set.

4.3.2 Performance of ResNet50

The performance of the ResNet50 model is analysed and presented in this section. Figure 4.3 illustrates the ACC and loss ratios of the ResNet50 model during training and validation.

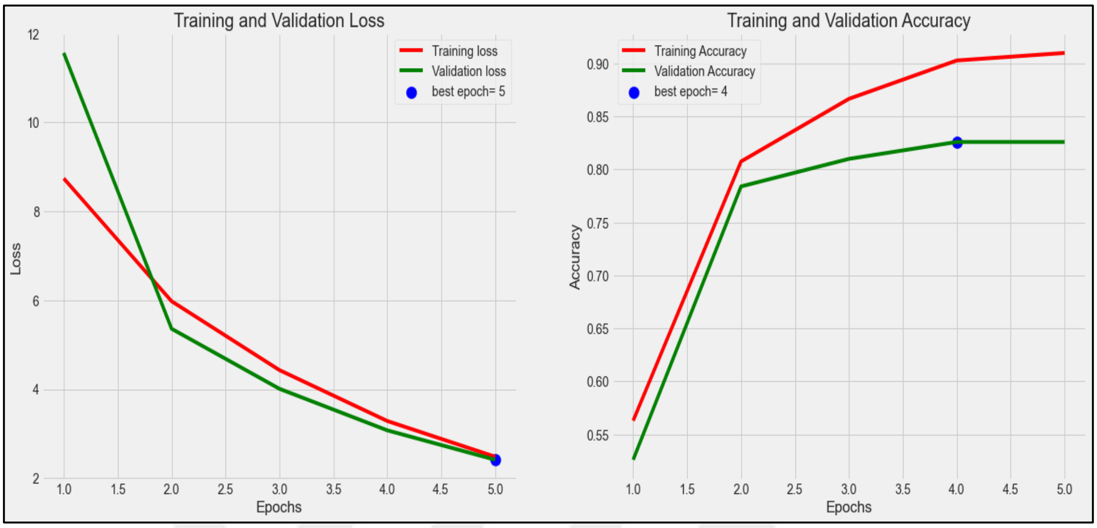

Figure 4.3: ACC and Losses Ratios of ResNet50 Model.

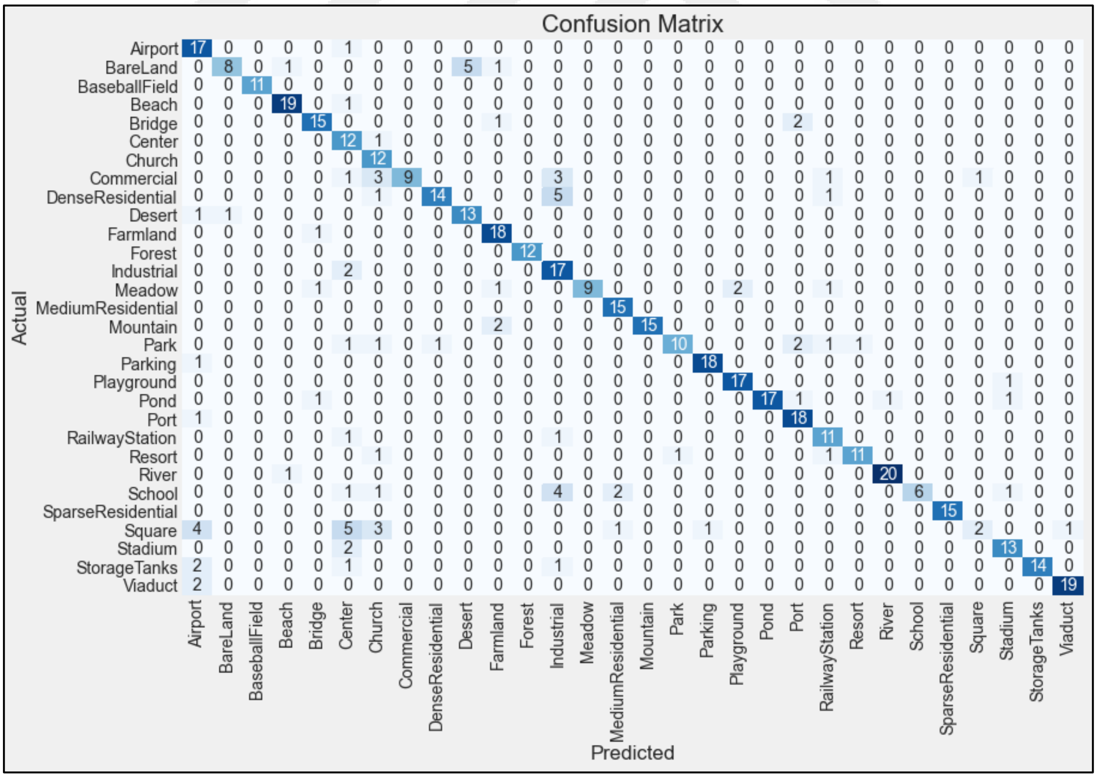

Figure 4.4: CM of RESNET50 Model.

Figure 4.4 depicts the CM of the ResNet50 model, providing a visual representation of the classification results. Additionally, Table 4.2 presents the precision, sensitivity, F-measure, and support values for each scene class.

Table 4.2: ResNet50 Model Results.

Class	precision	sensitivity	F-measure	support
Airport	0.6071	0.9444	0.7391	18
BareLand	0.8889	0.5333	0.6667	15
BaseballField	1	1	1	11
Beach	0.9048	0.95	0.9268	20
Bridge	0.8333	0.8333	0.8333	18
Center	0.4286	0.9231	0.5854	13
Church	0.5217	1	0.6857	12
Commercial	1	0.5	0.6667	18
DenseResidential	0.9333	0.6667	0.7778	21
Desert	0.7222	0.8667	0.7879	15
Farmland	0.7826	0.9474	0.8571	19
Forest	1	1	1	12
Industrial	0.5484	0.8947	0.68	19
Meadow	1	0.6429	0.7826	14
MediumResidential	0.8333	1	0.9091	15
Mountain	1	0.8824	0.9375	17
Park	0.9091	0.5882	0.7143	17
Parking	0.9474	0.9474	0.9474	19
Playground	0.8947	0.9444	0.9189	18
Pond	1	0.8095	0.8947	21

Table 4.2: ResNet50 Model Results "Table Continued".

Port	**0.7826**	**0.9474**	**0.8571**	**19**
RailwayStation	0.6875	0.8462	0.7586	13
Resort	0.9167	0.7857	0.8462	14
River	0.9524	0.9524	0.9524	21
School	1	0.4	0.5714	15
SparseResidential	1	1	1	15
Square	0.6667	0.1176	0.2	17
Stadium	0.8125	0.8667	0.8387	15
StorageTanks	1	0.7778	0.875	18
Viaduct	0.95	0.9048	0.9268	21
ACC			0.814	500
macro avg	0.8508	0.8158	0.8046	500
weighted avg	0.8562	0.814	0.8076	500

From Table 4.2 the following can be extracted according to each measure:

Precision: The ResNet50 model shows varying precision values for different classes, ranging from 0.4286 to 1. Some classes have relatively low precision values, indicating a higher rate of false positives.

Sensitivity: The model exhibits sensitivity values ranging from 0.1176 to 1, indicating varying success in capturing actual positive instances across different classes.

F-measure: The F-measures for the ResNet50 model range from 0.2 to 1, indicating varied performance across different classes. Some classes show lower F-measures, indicating challenges in achieving a balance between precision and sensitivity.

Overall ACC: The ResNet50 model produces an overall ACC of 0.814, indicating that it correctly classifies 81.4% of the instances in the test set.

4.3.3 Performance of VGG19

The performance of the VGG19 model is analysed and presented in this section. Figure 4.5 illustrates the ACC and loss ratios of the VGG19 model during training and validation. Figure 4.6 depicts the CM of the VGG19model, providing a visual representation of the classification results. Additionally, Table 4.3 presents the precision, sensitivity, F-measure, and support values for each scene class.

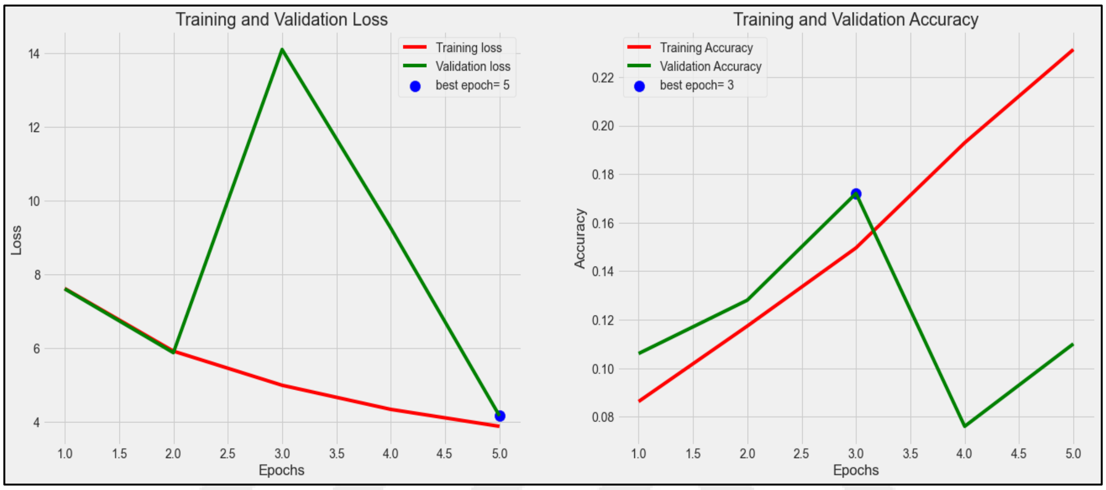

Figure 4.5: ACC and Losses Ratios of VGG19 Model.

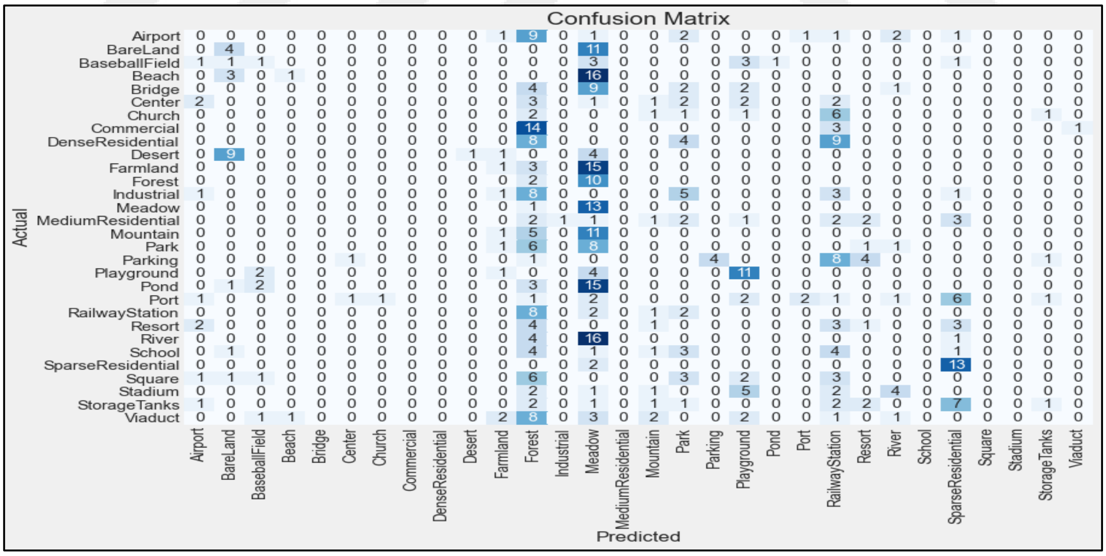

Figure 4.6: CM of VGG19 Model.

Table 4.3: VGG19 Model Results.

Class	precision	sensitivity	F-measure	support
Airport	0	0	0	18
Bare Land	0.2	0.2667	0.2286	15
Baseball Field	0.1429	0.0909	0.1111	11
Beach	0.5	0.05	0.0909	20
Bridge	0	0	0	18
Centre	0	0	0	13
Church	0	0	0	12
Commercial	0	0	0	18
Dense Residential	0	0	0	21
Desert	1	0.0667	0.125	15
Farmland	0.1111	0.0526	0.0714	19
Forest	0.0182	0.1667	0.0328	12
Industrial	0	0	0	19
Meadow	0.0867	0.9286	0.1585	14
Medium Residential	0	0	0	15
Mountain	0	0	0	17
Park	0	0	0	17
Parking	1	0.2105	0.3478	19
Playground	0.3548	0.6111	0.449	18
Pond	0	0	0	21
Port	0.6667	0.1053	0.1818	19
Railway Station	0	0	0	13

Table 4.3: VGG19 Model Results "Table Continued".

Resort	**0.1**	**0.0714**	**0.0833**	**14**
River	0	0	0	21
School	0	0	0	15
SparseResidential	0.3514	0.8667	0.5	15
Square	0	0	0	17
Stadium	0	0	0	15
Storage Tanks	0.25	0.0556	0.0909	18
Viaduct	0	0	0	21
ACC			0.11	500
Macro Average	0.1594	0.1181	0.0824	500
Weighted Average	0.1647	0.11	0.0815	500

From Table 4.3 the following can be extracted according to each measure:

Precision: The VGG19 model shows varying precision values for different classes, ranging from 0 to 1. Several classes have precision values of 0, indicating that the model did not classify any instances as positive for those classes. This suggests that the model had difficulties in accurately identifying those specific scene classes.

Sensitivity: The sensitivity values for the VGG19 model also vary across different classes, ranging from 0 to 0.9286. Similar to precision, some classes have sensitivity values of 0, indicating that the model failed to capture any actual positive instances for those classes.

F-measure: The F-measures for the VGG19 model range from 0 to 0.5. As with precision and sensitivity, some classes have an F-measure of 0, indicating challenges in achieving a balance between precision and sensitivity.

Overall ACC: The VGG19 model produces an overall ACC of 0.11, indicating that it correctly classifies only 11% of the instances in the test set. This low ACC suggests that the

VGG19 model struggled to effectively classify the aerial photographs and extract investible lands from the dataset used in the experiments.

4.4 COMPARISONS

In this section, we compare the performance of the three neural network models used in the experiments: EfficientNetB3, ResNet50, and VGG19. We evaluate their performance based on the precision, sensitivity, and F-measure metrics. EfficientNetB3 produced an overall ACC of 93.4%. It demonstrated high performance across most of the scene classes, with precision, sensitivity, and F-measures above 0.9 for several categories such as Airport, Beach, Bridge, Farmland, Parking, Playground, Pond, River, Storage Tanks, and Viaduct. These results indicate that EfficientNetB3 was able to accurately classify these land categories from the aerial images. ResNet50 produced an overall ACC of 81.4%. While it performed well for some scene classes, such as Baseball Field, Beach, Farmland, Forest, Playground, Port, River, and Storage Tanks, it showed relatively lower performance for other categories. For instance, the precision, sensitivity, and F-measures for Centre, Church, Industrial, Meadow, and Square were relatively lower. These results suggest that ResNet50 had difficulty distinguishing certain land categories from the aerial images. VGG19, on the other hand, produced the lowest overall ACC of 11%. It struggled to classify most of the scene classes, as indicated by the low precision, sensitivity, and F-measures across the board. This indicates that VGG19 was not effective in accurately identifying the investible lands from the aerial photographs. Overall, EfficientNetB3 outperformed ResNet50 and VGG19 in terms of ACC and the ability to classify land categories accurately. It demonstrated consistent high performance across a wide range of scene classes. ResNet50 showed relatively lower performance, while VGG19 performed poorly in almost all categories. These results highlight the importance of choosing an appropriate neural network model for image classification tasks. EfficientNetB3, with its advanced architecture and parameter optimization, proved to be the most effective model for classifying aerial photographs and extracting investible lands. For the final comparison, the models were compared on the basis of the average, as in Figure 4.7.

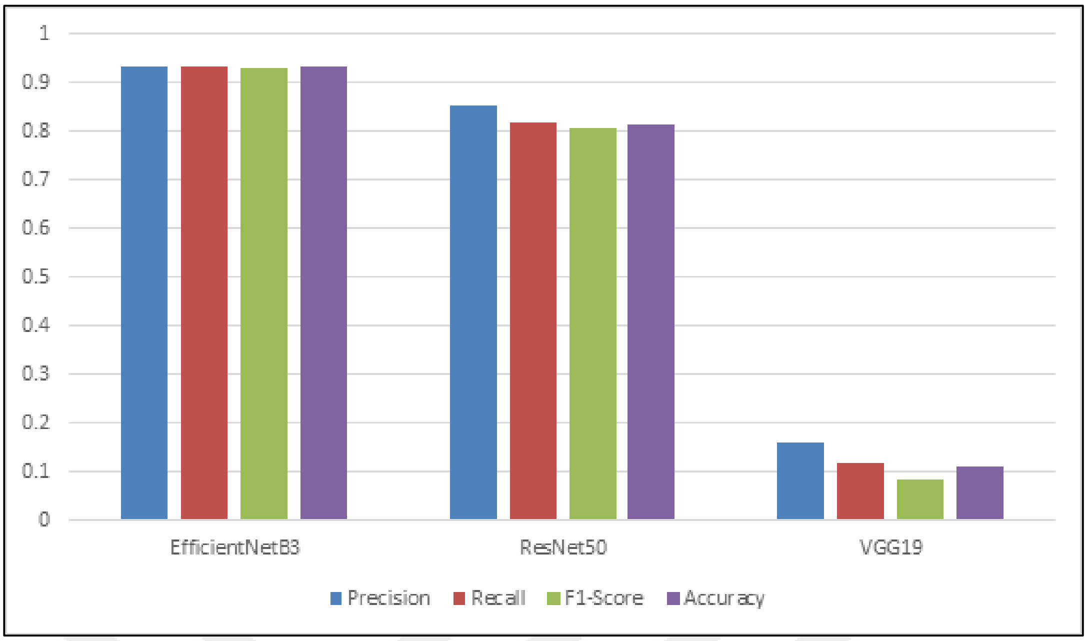

Figure 4.7: Comparison of the Models.

4.5 DISCUSSIONS

The results of the assessments determine the performance of three neural network models: EfficientNetB3, ResNet50, and VGG19. Each model was evaluated based on various performance metrics including precision, sensitivity, and F-measure. Among the three models, EfficientNetB3 consistently outperformed the others in terms of ACC, precision, sensitivity, and F-measure on both the validation and test sets. This indicates that EfficientNetB3 has a higher ability to classify aerial photographs and extract investible lands compared to ResNet50 and VGG19. ResNet50 produced moderate performance with relatively lower precision, sensitivity, and F-measure compared to EfficientNetB3. VGG19, on the other hand, demonstrated the lowest performance across all metrics, showing low precision, sensitivity, and F-measure values. This suggests that VGG19 may not be the most suitable model for the classification task of aerial photographs and investible lands in this context. The high performance of EfficientNetB3 can be attributed to its state-of-the-art architecture, which enables it to handle the complexity of the dataset effectively and generalize well to unseen data. The results validate the effectiveness of transfer learning and DL techniques in aerial image recognition, as the models were able to learn meaningful attributes from the images by leveraging pre-trained models and fine-tuning them on the AID dataset. It is extremely important to keep in mind that the performance of the representations may be different depending on the specific dataset and task at hand. Further experimentation and fine-tuning of the models can be conducted to improve their performance and adapt them to specific use cases. Overall, the findings highlight the potential of utilizing DL models, particularly EfficientNetB3, for the classification of aerial photographs and extraction of investible lands. These results can contribute to various applications such as urban planning, real estate development, and land management, where accurate classification of aerial images is crucial for decision-making processes.

5. CONCLUSIONS AND FUTURE WORK

5.1 CONCLUSIONS

In this thesis, a methodology for classifying aerial photographs and extracting investible lands using DL techniques was proposed and implemented. The methodology involved several steps, including data pre-processing, model selection, model training, transfer learning parameter fine-tuning, and testing and performance evaluation. The proposed methodology leveraged DL techniques, transfer learning, and data pre-processing to effectively classify aerial photographs and identify investible lands. The results of the experiments demonstrated the capability of the models to accurately classify aerial images and extract relevant information for investment decision-making. The selected neural network models, ResNet50, VGG19, and EfficientNetB3, were trained and evaluated using performance measures such as precision, sensitivity, F-measure, and the CM. Through the evaluation process, the best model was determined based on the validation loss, and early stopping was implemented to prevent over fitting and ensure generalization to unseen data. The findings of the experiments showed that the proposed methodology produced promising results in classifying aerial photographs and identifying investible lands. The models exhibited high ACC, precision, sensitivity, and F-measure, indicating their effectiveness in recognizing investment-worthy areas from aerial images. The performance metrics demonstrated the models' ability to generalize effectively and make reliable consistent predictions on unseen data.

5.2 FUTURE WORK

Although the proposed methodology produced favourable results, there are multiple categories for further research and improvement. The following are some areas that can be explored in future work:

a. Expansion of the Dataset: The current study utilized the Aerial Image Dataset (AID), which contains 30 scene classes. Expanding the dataset by including more diverse scenes and increasing the number of images per class could enhance the models' ability to generalize and improve their performance.

b. Fine-Tuning Hyper parameters: Fine-tuning the hyper parameters of the selected neural network models could potentially lead to better performance. Experimenting with different learning rates, optimizer algorithms, activation functions, and model architectures may help optimize the models' learning process.

c. Incorporation of Semantic Segmentation: Instead of classifying entire images, incorporating semantic segmentation techniques can enable the models to identify specific objects or regions within the aerial photographs. This finer level of analysis can provide more detailed information for investment decision-making.

d. Integration of Multi-Modal Data: Combining aerial images with other data sources, such as satellite imagery, topographic maps, or real estate data, can enrich the models' input and improve their ACC in identifying investible lands. Multi-modal data integration can provide a more comprehensive understanding of the areas under consideration.

e. Real-Time Implementation: Adapting the proposed methodology for real-time implementation, such as processing aerial images captured by drones in real-time, can have practical applications in various industries. Developing an efficient and real-time pipeline for image classification and land identification would be valuable for decision-making processes.

f. Transferability to Other Geographic Regions: The proposed methodology was developed using aerial images from specific geographic locations. Evaluating the transferability of the models to different regions and landscapes can help determine their robustness and applicability in diverse settings.

In the conclusion, the proposed methodology demonstrated the potential of DL techniques for classifying aerial photographs and extracting investible lands. The findings overall of this study supply a foundation for further research and improvement in the field of aerial image analysis and its applications in decision-making processes. By exploring the suggested areas for future work, researchers can continue to advance the ACC and effectiveness of models for analysing aerial images and supporting investment decision-making in various domains.

REFERENCES

[1] Desarrollo, C.d.l.N.U.s.C.y.: 'World investment report 2018: investment and new industrial policies' (UN, 2018. 2018)

[2] Gitman, L.J., Joehnk, M.D., Smart, S., and Juchau, R.H.: 'Fundamentals of investing' (Pearson Higher Education AU, 2015. 2015)

[3] Virlics, A.: 'Investment decision making and risk', Procedia Economics and Finance, 2013, 6, pp. 169-177

[4] Li, T.M.: 'What is land? Assembling a resource for global investment', Transactions of the institute of British Geographers, 2014, 39, (4), pp. 589-602

[5] Archambault, C.S., and Zoomers, A.: 'From small farms to firms', Global Trends in Land Tenure Reform: Gender Impacts, 2015, pp. 13

[6] Morgan, J.L., Gergel, S.E., and Coops, N.C.: 'Aerial photography: a rapidly evolving tool for ecological management', BioScience, 2010, 60, (1), pp. 47-59

[7] Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M.A., Al-Amidie, M., and Farhan, L.: 'Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions', Journal of big Data, 2021, 8, pp. 1-74

[8] Teixeira, I., Morais, R., Sousa, J.J., and Cunha, A.: 'Deep Learning Models for the Classification of Crops in Aerial Imagery: A Review', Agriculture, 2023, 13, (5), pp. 965

[9] Cira, C.-I., Alcarria, R., Manso-Callejo, M.-Á., and Serradilla, F.: 'Evaluation of Transfer Learning Techniques with Convolutional Neural Networks (CNNs) to Detect the Existence of Roads in High-Resolution Aerial Imagery', in Editor (Ed.)^(Eds.): 'Book Evaluation of Transfer Learning Techniques with Convolutional Neural Networks (CNNs) to Detect the Existence of Roads in High-Resolution Aerial Imagery' (Springer, 2019, edn.), pp. 185-198

[10] Choi, R.Y., Coyner, A.S., Kalpathy-Cramer, J., Chiang, M.F., and Campbell, J.P.: 'Introduction to machine learning, neural networks, and deep learning', Translational Vision Science & Technology, 2020, 9, (2), pp. 14-14

[11] Sharpe, W.F., Alexander, G.J., and Bailey, J.V.: 'Investment' (Prentice Hall Incorporated, 1999. 1999)

[12] Laplane, A., and Mazzucato, M.: 'Socializing the risks and rewards of public investments: Economic, policy, and legal issues', Research Policy, 2020, 49, pp. 100008

[13] Götze, U., Northcott, D., and Schuster, P.: 'Capital budgeting and investment decisions', Investment appraisal: Methods and models, 2008, pp. 3-26

[14] Capital Market Authority, K.o.S.A.: 'Investment', 2014, pp. 1-16

[15] ADKINS, C., JACKSON: 'How to Invest in Land', 2023

[16] Dheeraj Vaidya, C., FRM: 'Investment Decision', WallstreetMojo, 2023

[17] Talukdar, S., Singha, P., Mahato, S., Pal, S., Liou, Y.-A., and Rahman, A.: 'Land-use land-cover classification by machine learning classifiers for satellite observations—A review', Remote Sensing, 2020, 12, (7), pp. 1135

[18] Yuan, H., Van Der Wiele, C.F., and Khorram, S.: 'An automated artificial neural network system for land use/land cover classification from Landsat TM imagery', Remote Sensing, 2009, 1, (3), pp. 243-265

[19] Ali, K., Rahman, H.U., Khan, R., Siddiqui, M.Y., Najeeb, F., Amin, S., Khan, R.A., Manzoor, S., and Chaudhry, N.J.: 'Land usage analysis: a machine learning approach', International Journal of Computer Applications, 2016, 141, (12), pp. 23-28

[20] Borde, S., Rane, A., Shende, G., and Shetty, S.: 'Real estate investment advising using machine learning', International Research Journal of Engineering and Technology, 2017, 4, (3), pp. 1821-1825

[21] Wang, J., and Zhang, C.: 'Software reliability prediction using a deep learning model based on the RNN encoder–decoder', Reliability Engineering & System Safety, 2018, 170, pp. 73-82

[22] Dibs, H., Hasab, H.A., Al-Rifaie, J.K., and Al-Ansari, N.: 'An optimal approach for land-use/land-cover mapping by integration and fusion of multispectral landsat OLI images: case study in Baghdad, Iraq', Water, Air, & Soil Pollution, 2020, 231, (9), pp. 488

[23] Elmannai, H., and AlGarni, A.D.: 'Classification using semantic feature and machine learning: Land-use case application', TELKOMNIKA (Telecommunication Computing Electronics and Control), 2021, 19, (4), pp. 1242-1250

[24] Khwarahm, N.R., Qader, S., Ararat, K., and Fadhil Al-Quraishi, A.M.: 'Predicting and mapping land cover/land use changes in Erbil/Iraq using CA-Markov synergy model', Earth science informatics, 2021, 14, (1), pp. 393-406

[25] Anand, V., and Oinam, B.: 'Future land use land cover prediction with special emphasis on urbanization and wetlands', Remote Sensing Letters, 2020, 11, (3), pp. 225-234

[26] Gbodjo, Y.J.E., Ienco, D., Leroux, L., Interdonato, R., Gaetano, R., and Ndao, B.: 'Object-based multi-temporal and multi-source land cover mapping leveraging hierarchical class relationships', Remote Sensing, 2020, 12, (17), pp. 2814

[27] Jamali, A.: 'Land use land cover mapping using advanced machine learning classifiers: A case study of Shiraz city, Iran', Earth Science Informatics, 2020, 13, (4), pp. 1015-1030

[28] Lakhal, N., Guizani, A., Sghaier, A., El ́amine Abdelli, M., and Slimene, I.B.: 'The impact of CSR performance on Efficiency of Investments using Machine Learning', in Editor (Ed.)^(Eds.): 'Book The impact of CSR performance on Efficiency of Investments using Machine Learning' (2021, edn.), pp.

[29] Swetanisha, S., Panda, A.R., and Behera, D.K.: 'Land use/land cover classification using machine learning models', International Journal of Electrical & Computer Engineering (2088-8708), 2022, 12, (2)

[30] Rousset, G., Despinoy, M., Schindler, K., and Mangeas, M.: 'Assessment of deep learning techniques for land use land cover classification in southern new Caledonia', Remote Sensing, 2021, 13, (12), pp. 2257

[31] Abdullah, A.Y.M., Masrur, A., Adnan, M.S.G., Baky, M.A.A., Hassan, Q.K., and Dewan, A.: 'Spatio-temporal patterns of land use/land cover change in the heterogeneous coastal region of Bangladesh between 1990 and 2017', Remote Sensing, 2019, 11, (7), pp. 790

[32] AlAfandy, E.K.A.Y.I.: 'Thesis Title: Investment of Machine Learning Models for Classifying Remote Sensing Images'

[33] Prasad, P., Loveson, V.J., Chandra, P., and Kotha, M.: 'Evaluation and comparison of the earth observing sensors in land cover/land use studies using machine learning algorithms', Ecological Informatics, 2022, 68, pp. 101522

[34] Waghela, H., Patel, S., Sudesan, P., Raorane, S., and Borgalli, R.: 'Land Use Land Cover Classification using Machine Learning', in Editor (Ed.)^(Eds.): 'Book Land Use Land Cover Classification using Machine Learning' (IEEE, 2022, edn.), pp. 708-711

[35] Wiam, S., Khouloud, T., Bouchra, H., SAIDI, M.N., and Adil, K.: 'Hybrid Deep Learning Architecture for Land Use: Land Cover Images Classification with a Comparative and Experimental Study', International Journal of Advanced Computer Science and Applications, 2022, 13, (12)

[36] Bernasconi, E., De Fausti, F., Pugliese, F., Scannapieco, M., and Zardetto, D.: 'Automatic extraction of land cover statistics from satellite imagery by deep learning', Statistical Journal of the IAOS, 2022, 38, (1), pp. 183-199

[37] Mueller, J.P., and Massaron, L.: 'Machine learning for dummies' (John Wiley & Sons, 2021. 2021)

[38] Ammann, E.M., Schweizer, M.L., Robinson, J.G., Eschol, J.O., Kafa, R., Girotra, S., Winiecki, S.K., Fuller, C.C., Carnahan, R.M., and Leonard, C.E.: 'Chart validation of inpatient ICD-9-CM administrative diagnosis codes for acute myocardial infarction (AMI) among intravenous immune globulin (IGIV) users in the S entinel D istributed D atabase', Pharmacoepidemiology and drug safety, 2018, 27, (4), pp. 398-404

[39] Bartram, S.M., Branke, J., and Motahari, M.: 'Artificial intelligence in asset management' (CFA Institute Research Foundation, 2020. 2020)

[40] Rebala, G., Ravi, A., and Churiwala, S.: ‘An introduction to machine learning’ (Springer, 2019. 2019)

[41] Sharma, A., Jain, A., Gupta, P., and Chowdary, V.: ‘Machine learning applications for precision agriculture: A comprehensive review’, IEEE Access, 2020, 9, pp. 4843-4873

[42] Mariammal, G., Suruliandi, A., Raja, S., and Poongothai, E.: ‘Prediction of land suitability for crop cultivation based on soil and environmental characteristics using modified recursive feature elimination technique with various classifiers’, IEEE Transactions on Computational Social Systems, 2021, 8, (5), pp. 1132-1142

[43] Kubat, M., and Kubat, J.: ‘An introduction to machine learning’ (Springer, 2017. 2017)

[44] Saha, S.: ‘A comprehensive guide to convolutional neural networks—the ELI5 way’, Towards data science, 2018, 15, pp. 15

[45] O'Shea, K., and Nash, R.: ‘An introduction to convolutional neural networks’, arXiv preprint arXiv:1511.08458, 2015

[46] Li, Z., Liu, F., Yang, W., Peng, S., and Zhou, J.: ‘A survey of convolutional neural networks: analysis, applications, and prospects’, IEEE transactions on neural networks and learning systems, 2021

[47] De Salvo, B.: ‘Brain-inspired technologies: Towards chips that think?’, in Editor (Ed.)^(Eds.): ‘Book Brain-inspired technologies: Towards chips that think?’ (IEEE, 2018, edn.), pp. 12-18

[48] Liu, Z., Mao, H., Wu, C.-Y., Feichtenhofer, C., Darrell, T., and Xie, S.: ‘A convnet for the 2020s’, in Editor (Ed.)^(Eds.): ‘Book A convnet for the 2020s’ (2022, edn.), pp. 11976-11986

[49] Dumoulin, V., and Visin, F.: ‘A guide to convolution arithmetic for deep learning’, arXiv preprint arXiv:1603.07285, 2016

[50] Yamashita, R., Nishio, M., Do, R.K.G., and Togashi, K.: ‘Convolutional neural networks: an overview and application in radiology’, Insights into imaging, 2018, 9, pp. 611-629

[51] Nielsen, M.A.: ‘Neural networks and deep learning’ (Determination press San Francisco, CA, USA, 2015. 2015)

[52] Sapijaszko, G., and Mikhael, W.B.: ‘An overview of recent convolutional neural network algorithms for image recognition’, in Editor (Ed.)^(Eds.): ‘Book An overview of recent convolutional neural network algorithms for image recognition’ (IEEE, 2018, edn.), pp. 743-746

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

[54] Zheng, Y., Yang, C., and Merkulov, A.: ‘Breast cancer screening using convolutional neural network and follow-up digital mammography’, in Editor (Ed.)^(Eds.): ‘Book Breast cancer screening using convolutional neural network and follow-up digital mammography’ (SPIE, 2018, edn.), pp. 1066905

[55] Bansal, M., Kumar, M., Sachdeva, M., and Mittal, A.: ‘Transfer learning for image classification using VGG19: Caltech-101 image data set’, Journal of ambient intelligence and humanized computing, 2021, pp. 1-12

[56] Ali, L., Alnajjar, F., Jassmi, H.A., Gocho, M., Khan, W., and Serhani, M.A.: ‘Performance evaluation of deep CNN-based crack detection and localization techniques for concrete structures’, Sensors, 2021, 21, (5), pp. 1688

[57] He, K., Zhang, X., Ren, S., and Sun, J.: ‘Deep residual learning for image recognition’, in Editor (Ed.)^(Eds.): ‘Book Deep residual learning for image recognition’ (2016, edn.), pp. 770-778

[58] Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., and Hu, Q.: ‘ECA-Net: Efficient channel attention for deep convolutional neural networks’, in Editor (Ed.)^(Eds.): ‘Book ECA-Net: Efficient channel attention for deep convolutional neural networks’ (2020, edn.), pp. 11534-11542

[59] Batool, A., and Byun, Y.-C.: ‘Lightweight EfficientNetB3 Model based on Depthwise Separable Convolutions for Enhancing Classification of Leukemia White Blood Cell Images’, IEEE Access, 2023