Machine Learning Approaches for Building Inventory Characterisation

A thesis submitted in partial fulfilment of the requirements for the Degree of Master of Science in Geospatial Technologies

by

Sunil Tamang

Supervised by:

Dr. Bakhtiar Feizizadeh
Institute of Geoinformatics
University of Münster
Münster, Germany

Co-supervised by:

Dr. Marco Painho
NOVA Information Management School
Universidade Nova de Lisboa
Lisboa, Portugal

Dr. Christian Geiss
German Remote Sensing Data Center
German Aerospace Center (DLR)
Oberpfaffenhofen, Germany

20 February 2024

Funded by the European Union

Declaration

I hereby confirm that the thesis titled "**Machine Learning Approaches for Building Inventory Characterisation**" is entirely my original work, completed with the guidance of my supervisors. All sources used, including books, journals, handouts, unpublished manuscripts, and various internet resources, have been accurately cited. This thesis has not been approved for any degree and is not currently being submitted for any other academic qualification.

Sunil Tamang
Münster, Germany
20 February 2024

Copyright Statement

Acknowledgements

I want to express my sincere appreciation to my supervisors: Dr. Bhaktiar Fezizadeh from Institute of Geoinformatics (IFGI), University of Münster; Prof. Dr. Marco Painho from NOVA University of Lison; and Prof. Dr. Christian Geiss from German Aerospace Center (DLR) for guiding and supporting me throughout my thesis research. Special thanks to Anne Schauss from Heidelberg Institute for Geoinformation Technology (HeiGIT) for connecting me to Dr. Geiss, who proposed thesis topic and provided the data. I am equally thankful to Dr. Maciej Adamiak from HeiGIT for continued technical consultation and encouragement.

Acknowledgements are also due to Prof. Joaquin Huerta from University Jaume (UJI), and Dr. Chrisopher Brox at IFGI, and Prof. Dr. Marco Painho for their outstanding coordination of Erasmus Mundus joint master`s degree program. My heartfelt thanks to Erasmus Mundus program for providing a fully funded scholarship, which played important role in enriching my academic and personal experience.

I extend my gratitude to all professors from UJI and IFGI involved in the joint master program. A special note of gratitude goes to Prof. Sven Casteleyn for his unparalleled support and inspiration in learning programming skills. I am also thankful to Gloria from UJI for consistent assistance in navigating unexpected bureaucratic processes.

Thank you to all friends from Castellon cohort and Lisbon cohort for many beautiful experiences and mutual support that propelled us forward together. Lastly, my deepest thanks to my family for their unwavering belief in my pursuits and encouragement to continue learning and exploring new avenues.

Abstract

Accurate building inventories and relevant information are essential for sustainable urban governance, thereby contributing to achieve the Sustainable Development Goal 11, which focuses on sustainable cities and communities. Acquiring building information manually in complex-built environment with densely populated buildings is impractical and may not guarantee the high thematic and spatial detail necessary for real world applications.

Recent advancements in remote sensing technology, coupled with the availability of high-resolution satellite and drone-based imagery through open access channels and as machine learning (ML) and deep learning (DL) continue to advance, showcasing their capability to recognise complex patterns, new opportunities for interpreting various surface features on the Earth are made possible. A growing body of literature emphasises the use of both traditional ML and DL for building footprint extraction, yet available and accessible literature reveal limited application of DL in urban building characterisation.

This study involves exploring and implementing Random Forest (RF) as machine learning model and dense neural network (DNN) as deep learning model in the context of multi-class building characterisation encompassing six classes. A total of 35 geometric and distribution features calculated using VHR imagery and OpenStreetMap data are used to train the model. The experiments show that overall accuracy of RF (79.9%) is higher than that of the DNN (71.9%). Upon closer examination and comparison of diagonal elements, representing the number of correctly classified samples for each class, it is found that the DNN outperforms RF in correctly classifying more instances for four classes. Further the recall rate using DNN is greater for four classes- ‘Building block in closed construction’, ‘Detached building block’, ‘Free standing individual building’, and ‘Garage’, in comparison to that of RF. Implementation of the DNN and comparison with traditional machine learning algorithm- RF provide additional scientific contribution, especially in situation where there is limited use of deep learning algorithms in the building characterisation.

Keywords: Building Footprints, Characterisation, Random Forest, Dense Neural Network

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ADT	Aggregated Decision Tree
AI	Artificial Intelligence
CNN	Convolutional Neural Network
DL	Deep Learning
DNN	Dense Neural Network
DT	Decision Tree
GBDT	Gradient Boosted Decision Tree
GLCM	Grayscale Co-occurrence Matrix
kNN	k-Nearest Neighbours
LDA	Linear Discriminant Analysis
LiDAR	Light Detection and Ranging
MABI	Multi-Angular Built-up Index
ML	Machine Learning
MLP	Multi-Layer Perceptron
MRF	Markov Random Field
NB	Naïve Bayes
NC	Nearest Centroid
OBIA	Object Based Image Analysis
OSM	OpenStreetMap
PGIS	Participatory Geographic Information Systems
POI	Point of Interests
RF	Random Forest
SAR	Synthetic Radar Aperture
SVM	Support Vector Machine
UN	United Nations
VHR	Very High Resolution
WM	Word Mining

Chapter 1
Introduction

1.1 Background

Over 55% of people worldwide live in cities, and the United Nations predicts that by 2050, about 70% of the global population will reside in urban areas (UN, 2019). As cities rapidly grow to accommodate this influx of population growth, they expand in three ways: more buildings (horizontal expansion), filling spaces between existing structures, and constructing taller buildings (Lall et al., 2021). This highlights the importance of buildings as primary living spaces of cities and integral components of urban morphology (Oliveira, 2016).

Accurate building inventories and relevant information are essential for various practical applications such as disaster management (Cerri et al., 2021; Geiss et al., 2015), climate and energy modelling (Anand & Deb, 2023; Masson et al., 2020), urban planning, and transport management (Grippa et al., 2018; Scorza & Fortunato, 2021). They also play significant role in sustainable urban governance (Wurm et al., 2010), thereby contributing to achieve the Sustainable Development Goal 11, which focuses on sustainable cities and communities (UN, 2023). Traditionally, gathering building information involves labour-intensive and time-consuming processes such as field surveys and manually digitizing aerial images (Geiss et al., 2017; Lu et al., 2014). However, these methods may not guarantee the high thematic and spatial detail necessary for real world applications (Hall, 2003). In addition, acquiring building information manually in complex-built environment with densely populated buildings is impractical (Zhou & Chang, 2021). Hence, there is a need for research into automated methods for identifying and characterising buildings.

"Citizen scientists" (Goodchild, 2007) and participatory geographic information systems (PGIS) have empowered the general public to actively contribute to the creation of data and participate in collaborative decision-making processes using spatial information (Balram & Dragicevic, 2006). OpenStreetMap (OSM), a well-known platform for PGIS established in 2014, has achieved approximately 80% global data coverage (Barrington-Leigh & Millard-Ball, 2017; Hecht et al., 2015). Nonetheless, concerns have arisen regarding the accuracy and quality of OSM data due to infrequent updates to the database (W. Chen et al., 2020). In the context of buildings, 82% of building polygons lack further information (Cerri et al., 2021). There are noticeable gaps in attribute information within OSM, specifically in building features such as occupancy, structural type, roofing materials, number of floors, and the age of houses. For example, the OSM consists of floor area information for only about 5% of building polygons in Germany (Wurm et al., 2021).

Nevertheless, these attributes are crucial for developing practical models for real world applications. For instance, building features such as footprint area, floor area, and building use type are crucial for urban energy models (Anand & Deb, 2023; Wurm et al., 2021).

Recent advancements in remote sensing technology, coupled with the availability of high-resolution satellite and drone-based imagery through open access channels, offer new opportunities for interpreting various surface features on the Earth (Guo et al., 2018; Yuan et al., 2021). Multispectral optical imagery (W. Li et al., 2019), synthetic aperture radar (SAR) data (Pasquali et al., 2019), light detection and ranging (LiDAR) data (J. Wang & Shan, 2009) and drone-derived imagery (Wouters et al., 2021) have been leveraged to extract building footprints and compute additional features essential for characterising buildings (Bandam et al., 2022). Further, as machine learning (ML) and deep learning (DL) continue to advance, showcasing their capability to recognise complex patterns, there is increasing interest in research that employs artificial intelligence to automatically delineate building footprints and define their characteristics (Lu et al., 2014; Maxwell et al., 2018; Zhou & Chang, 2021).

1.2 Research Gap and Scientific Significance

Building characterisation involves two processes- outlining the boundary of building footprints and classifying these footprints into one or more classes based on attribute information. However, there are challenges in obtaining attribute information. Firstly, the OSM attribute information on building function or typology is incomplete (Cerri et al., 2021; Wurm et al., 2021). Secondly, governmental geodatabase, while accessible (may be challenging in developing country context), might lack consistent and required building information (W. Chen et al., 2020). Thirdly, buildings of different classes might have similar geometric and spectral characteristics, making solely remote sensing-based building classification quite challenging (Lu et al., 2014). This complexity necessitates the computation of additional features to numerically encapsulate building characteristics (Labetski et al., 2023) and identify the most significant features that contribute to data-driven building characterisation.

A growing body of literature emphasises the use of both traditional ML and DL for building footprint extraction. Notably, conventional machine learning models like Support Vector Machine (SVM), Decision Tree (DT), and Random Forest (RF), as well as deep learning models such as convolutional neural network (CNN), and U-Net algorithm, have been widely applied in building footprint extraction (Neupane et al., 2023; Shrestha & Vanneschi, 2018; Thottolil & Kumar, 2022; Virtriana et al., 2023; Wurm et al., 2021; Yuan et al., 2021). In the realm of building characterisation, traditional ML algorithms such as RF and SVM have been used (Geiss et al., 2015; Hecht et al., 2015), but there have been limited efforts in application of deep learning

models (Pelizari et al., 2021; Taoufiq et al., 2020). In addition, it is important to acknowledge that performance for ML and DL algorithms is specific to a domain, making it challenging to identify a single best algorithm across various datasets and applications (Ben-David et al., 2010; Sebastiani, 2002). Hence, further research and experimentation are necessary to find potentially more effective algorithm for building characterisation (Zhou & Chang, 2021).

To address above-mentioned gaps and limitations, the proposed research involves exploring and implementing Random Forest (refer to section 3.1.1) as machine learning model and dense neural network (refer to section 3.1.2) as deep learning model in the context of multi-class building characterisation encompassing six classes. Dataset comprises 152 features (refer to section 3.2.1) consisting of both 2D and 3D geometric indices, and spectral indices, computed using a very high-resolution satellite imagery. The indices are available in three spatial levels- individual buildings, aggregated buildings, and building blocks and are integrated into building polygons from the OSM dataset for city of Cologne, Germany. Subsequently, 215660 building data samples are used to train and test the models and their performances are compared. This research contributes to automated classification in the built environment. Given that this research is initial effort to implement both machine learning and deep learning for building characterisation, the results obtained are potentially significant for benchmarking model performance in similar future studies.

1.3 Objectives and Research Questions

The primary aim is the automated characterisation of building inventories using high-resolution multispectral satellite imagery and OpenStreetMap data. The following objectives and corresponding research questions are designed to achieve this goal:

Objective 1: To design and implement the Random Forest machine learning algorithm for building inventories characterisation.

Research questions:

a. How are the most significant features distributed across three spatial level- individual building, aggregated building, and building block- in characterising building inventories?

b. What is the performance of Random Forest in characterising building across six classes?

Objective 2: To design and implement the Dense Neural Network, deep learning algorithm, for building inventories characterisation.

Research questions:

a. How does Dense Neural Network perform in characterising buildings across six classes?

b. What difference exist in the model performance compared to Random Forest in characterising buildings across six classes?

1.4 Thesis structure

This thesis comprises of five chapters. Following the introductory chapter, **Chapter 2** reviews relevant literature to assess "state of the art" on application of artificial intelligence (both ML and DL) in building type characterisation. **Chapter 3** describes the study area, data, and provides conceptual background on Random Forest and Dense Neural Network. It further explains experimental set up for data preparation and preprocessing and for implementing RF and DNN. **Chapters 4** describes the results on implementation of models and their performance. It further discusses comparative analysis of RF and DNN model and the implications of the results on building characterisation. Finally, **Chapter 5** summarises the limitation of the study and provides recommendation for future work and summarises key findings and contributions of this study.

Chapter 2
Literature Review

2.1 Application of remote sensing in building detection

The initial phase of characterising building type involves the extraction of polygons or the delineation of building boundaries through remote sensing techniques (Lu et al., 2014). The data sources can be satellite and drone based optical imagery, SAR, and LiDAR data. Although SAR and LiDAR sensor data are used for building extraction, the very high-resolution (VHR) imagery from satellites such as WorldView, GeoEye, QuickBird and IKONOS are often utilised in building footprint extraction. The complex interpretation of SAR imagery and the substantial cost associated with LiDAR data collection have encouraged frequent use of optical imagery (J. Li et al., 2022; Wurm et al., 2010). The preference of optical imagery is driven by the benefits it offers. VHR optical imagery excels in capturing remote sensing data across the visible and infrared spectrum. With spatial resolution ranging from 0.05m to 4m, it presents combination of high geometric, spectral, and temporal resolution (Maktav et al., 2005), providing an effective means to identify buildings.

In line with studies exploring the application of VHR optical imagery in building extraction, J. Li et al., (2022) reviewed 417 peer-reviewed articles published between 2000 and 2021. The review subsequently categorised building detection methods into three groups: physical rule-based detection methods (24%); image segmentation methods (42%); and artificial intelligence (49.5%), covering both traditional ML and DL approaches.

Physical rule-based building footprint extraction relies on the principle that each building feature signifier consists of a unique set of characteristics. These characteristics facilitate binary characterization of an object when it satisfies the defined threshold for being identified as a building (Beykaei et al., 2014; J. Li et al., 2022). Consequently, a profound understanding of multiple building characteristic- geometric, spectral, textural, contextual, and vertical- plays a pivotal role in building detection (J. Li et al., 2022). In pure physical rule-based methods, techniques such as Harris corner detector (Harris & Stephens, 1988), and building indices like multi-angular built-up index (MABI), and built-up presence index (PanTex) are used to building footprint extraction. MABI (Liu et al., 2019), representing vertical building characteristics, may not perform optimally in low rise buildings. PanTex (Pesaresi et al., 2008), which represents textural characteristics using Grayscale co-occurrence matrix (GLCM) from panchromatic satellite data, may encounter challenges in differentiating land use types such as paved and concrete road from building structure with similar textures (J. Li et al., 2022; Shen et al., 2019).

In addressing such complexities, Object Based Image Analysis (OBIA) comes as viable alternative, providing better delineation of building boundaries.

Object Based Image Analysis (OBIA) swiftly emerged in the early 2000s in the field of remote sensing and geospatial analysis. This method involves grouping pixels with similar spectral and spatial attributes to construct distinct objects. The popularity of OBIA has risen due to its capability to capture complex features and patterns, proving especially advantageous in tasks such as categorizing land cover, detecting objects, and identifying changes in the landscape (Blaschke, 2010; Feizizadeh et al., 2021; Ma et al., 2019). OBIA encompasses several key steps. It commences with image segmentation, where pixels sharing similar spatial and spectral resolutions are clustered together to create coherent objects. Subsequently, feature extraction proceeds, wherein the outcomes of segmentation are employed to generate objects with diverse index such as texture, size, and shape. Classification follows, employing a set of predefined rules to assign categorical labels to objects. Spatial analysis is then employed to overlay the results with other geospatial layers, enhancing further interpretation. Belgiu et al., (2014) used OBIA approach to delineate building footprints from an airborne laser scanning data for their study in Germany. Gao et al., (2022) conducted a comparative study of the OBIA approach against manually annotated labels for identifying refugee dwelling footprints. Their findings revealed that OBIA outperformed manual labelling, achieving an Intersection over Union (IoU) greater than 0.5. Similarly, Prathiba et al., (2020) implemented the OBIA technique to extract building footprints from high-resolution satellite imagery. Both studies, however, faced under-segmentation and over-segmentation, negatively impacting the segmentation quality. Complex building structures, roads, and trees were prone to misclassification. The robust implementation of OBIA demands prior knowledge and familiarity with the study area, necessitating frequent parameter optimization by experts. Consequently, automation in OBIA is comparatively limited compared to pixel-based methodologies integrated into machine learning approaches.

Building detection based on artificial intelligence can be broadly divided into two groups: traditional machine learning, and deep learning approach. The machine learning mainly emphasise on feature extraction, feature engineering and the implementation of classifying algorithms (J. Li et al., 2022). Traditional machine learning methods like Random Forest, K-Nearest Neighbour, Support Vector Machine are widely used in building detection domain (Belgiu & Drăguţ, 2014; R. Chen et al., 2018; W. Chen et al., 2020; J. Li et al., 2022; Zhou & Chang, 2021). Further, the breakthrough in deep learning in 2012 revolutionized computer vision, including image segmentation, opening new possibilities and advancements in the field (Hoeser et al., 2020). The integration of deep learning techniques, such as fully convolutional networks (FCNs), U-Net, and ResNet architectures, have further improved the accuracy and efficiency of

image segmentation in remote sensing, enabling precise and detailed analysis of aerial and satellite imagery (Parente et al., 2023; Srivastava et al., 2024; Z. Wang et al., 2022). This section does not aim to present a comprehensive literature review on the application of deep learning in building detection, instead it highlights capabilities and effectiveness of deep learning in this field, using the U-Net as an illustrative example. The main innovation of U-Net is its unique architecture which includes an encoder-decoder design along with skip connection, and integration of fully convolutional networks (FCN). W. Li et al., (2019) employed the U-Net architecture with a combination of strategies, including data augmentation and post-processing techniques, to extract building footprints from high-resolution multispectral satellite images. The study utilized the SpaceNet building dataset provided by the DeepGlobe Satellite Challenge of the IEEE Conference on Computer Vision and Pattern Recognition 2018. The researchers explored the potential of GIS map datasets from OpenStreetMap, GoogleMaps, and MapWorld in conjunction with WorldView-3 satellite datasets in four cities. By applying the U-Net architecture and integrating various datasets, the proposed method achieved an impressive F1 score of 0.704. Furthermore, Pasquali et al., (2019) developed a structurally optimized U-Net architecture and implemented on a dataset provided to previous SpaceNet competition winners. They applied an ensemble of various U-Net architectures and distance transforms to enhance building boundary detection accuracy.

2.2 Machine learning approaches in building characterisation

2.2.1 Building types

The exploration of building type characterisation in literature is mainly divided into two key dimensions- functional and morphological characteristics. For functional building characterisation, researchers commonly employ at least three classes (Belgiu et al., 2014; Huang et al., 2017), with some extending to more than five classes (Du et al., 2015; Geiss et al., 2015; Zhou & Chang, 2021). Table 1 provides an overview of related research investigating characterisation of building types based on functional classes. Lu et al., (2014) classified building types into Single Family House, Multiple Family House, Non-residential Building. Similarly, Belgiu et al., (2014) characterised buildings into three classes, specifically referring to Industrial and Factory Building for what Lu et al., (2014) named as non-residential building. In addition, Geiss et al., (2015) and Zhou & Chang, (2021) approached building type characterisation based on structural function of the building. For example, Geiss et al., (2015) classified building type into seven classes- Confined Masonry, Reinforced Concrete Low, Reinforced Concrete High, Unreinforced Masonry, Timber Frame Residential, Timber Frame Non-Residential, Steel Frame.

Such classification plays important role in hazard risk assessments, particularly in urban areas located in active seismic zone.

In context of an architectural or morphological based building characterisation, buildings were classified into 5 to 11 classes. Table 2 provides overview of relevant research focused on characterising building types based on morphological classes. Goel et al., (2012) classified buildings into five European architectural classes. Wurm et al., (2015), taking Munich and Berlin as their study area, characterised building types into five classes- Perimeter Block Development, Block Development, Terraced Houses, Semi/Detached Houses, and Halls. In consistent with the nature of classes identified by Wurm et al., (2015), Henn et al., (2012) expanded morphological classes into seven classes. Castagno & Atkins, (2018) took a distinctive approach by exploring morphological classes based solely on the roof as the main feature. Their study conducted in New York, USA and Witten, Germany, resulted in the classification of building into eight morphological classes.

2.2.2 Features used in building characterisation

For robust and reliable performance of ML and DL algorithms, numerous relevant input features are important and can be extracted using VHR satellite imagery, drone based multispectral data, and LiDAR data (Du et al., 2015; Fan et al., 2014; Geiss et al., 2015; Tooke et al., 2014). Du et al., (2015) used 309 input features distributed across four categories- spectral features, texture features, geometry features and distribution features. Geometry features are also referred to as shape features (Lu et al., 2014; Wurm et al., 2015) or building specific features (Henn et al., 2012). Geometric Features includes but are not limited to area, length, radius, rectangular fit, elliptic fit, asymmetry, border index, main direction, shape index, compactness, roundness, density. Similarly, spectral features commonly used in analysis include mean brightness, mean blue, mean red, mean green, mean near infrared, GLCM contrast, GLCM entropy, and GLCM homogeneity. Further, neighborhood or spatial features such as proximity, distance from centroid are also calculated as additional input features (Henn et al., 2012; Lu et al., 2014). While most features are two dimensional (2D), researchers, for example (Geiss et al., 2015; Wurm et al., 2015) have also integrated three dimensional (3D) features by integrating height information.

2.2.3 Machine Learning and Deep Learning algorithms

Based on the literature reviewed (refer to Table 1 and Table 2), it is evident that Random Forest (RF) and Support Vector Machine (SVM) are commonly used traditional machine learning algorithms for building characterisation (Du et al., 2015; Geiss et al., 2015; Huang et al., 2017; Zhou & Chang, 2021). Huang et al., (2017) implemented RF and SVM using integrated data from

LiDAR and VHR imagery to classify buildings into three functional classes and achieved F1 score between 75% to 100% among the classes. Notably, RF acquired overall accuracy (OA) of 80.2%, approximately 12% higher than SVM. Similarly, Geiss et al., (2015) also used RF and SVM, where SVM outperformed RF with OA of 93% compared to RF`s OA of 82.5%. However, the authors observed that RF performed better when implemented with a limited set of relevant input features, as opposed to using all features. Lu et al., (2014) used DT, ADT, RF, SVM to classify building types using LiDAR data. The overall accuracy for urban context was 67.5%, 73.6%, 74.2% and 76.1% respectively. Zhou & Chang, (2021) employed twelve traditional machine learning algorithms- Adaboost, Bernoulli NB, DT, Gaussian NB, Gaussian Process (GP), Gradient Boosting Decision Trees (GBDT), k-Nearest Neighbors (kNN), Multi-Layer Perceptron (MLP), Nearest Centroid (NC), Quadratic Discriminant Analysis (QDA), RF, and SVM, to classify building types. Among these, DT, GP, GBDT, and MLP achieved an overall macro F1 greater than 85%, with GBDT performing better than RF and SVM.

Regarding deep learning in building characterisation, various variant models of Convolutional Neural Network (CNN) architecture have been implemented. Taoufiq et al., (2020) executed HierarchyNet, achieving overall accuracy of 87.34%. Hoffmann et al., (2019) implemented VGG16 using various geometric features obtained from OSM, aerial and street images to classify building types into four classes. The classification precision across all classes ranged from 68% to 76%. Yan et al., (2019) implemented graphical convolutional neural network (GCNN) architecture for building pattern classification, where vector data representing building types are modelled as graphs and twelve morphological indices are calculated as input features for building characterisation. It is important to note that there is limited literature available or accessible on the application of deep learning algorithms in building type characterisation.

Table 1. Overview of related work (functional building type classification)

Authors	**Study Area**	**Data**	**Functional Building Class**	**Input Features**	**Classification Algorithm**	**Model Performance**
Yang et al., (2022)	Futian district, China	OSM, POI	*4 classes:* Industrial, commercial, residential, other	Morphological Feature vector, socioeconomic feature vector	Ensemble stacking, where RF is base classifier and LR is meta classifier	Precision across classes ranged from 33.33% to 91.06%. Residential class achieved the highest precision. Industrial class with the lowest precision.
Zhou & Chang, (2021)	Beijing, China	POI	*5 classes:* Brick masonry, concrete frame, concrete frame-sheer wall, concrete sheer wall, steel	Geometric Features (10), Shape Features (1), POI Features (3)	Adaboost, Bernoulli NB, DT, Gaussian NB, GP, GBDT, kNN, MLP, NC, QDA, RF, SVM	Overall macro F1 > 85% for DT, GP, GBDT, MLP
Taoufiq et al., (2020)	Worldwide	Photographs from Google Images, Wikimedia	*4 classes:* Religious, Residential, Commercial, Business. Four classes were further divided into 10 functional sub classes.	Not available	HierarchyNet (a variant of CNN)	Overall Accuracy of 87.34%
Hoffmann et al., (2019)	49 States from USA	Aerial and Steet View Images, OSM	*4 classes*: Commercial, Residential, Public, Industrial	Geometric features	VGG16 (a variant of CNN)	Classification precision ranges from 68%-76%
Huang et al., (2017)	Guangzhou, China	LiDAR and VHR Imagery	*3 classes:* Villa, Apartment, Low-rise	Height extracted LiDAR data, texture, geometric and spectral features from VHR Imagery	RF, SVM	F1 score for classes for using integrated data with super object ranged from 75% to 100% for RF. Overall accuracy for RF is 80.02% and 68.97% for SVM using same data.

Du et al., (2015)	Beijing, China	VHR Imagery	*7 classes:* Low-Story Shantytowns, Medium Story Apartment, High Rising apartment, Administrative Buildings, Commercial Building, Industrial Building, Auxiliary Buildings	Spectral Features (47), Texture Features (216), Geometry Features (44), Distribution Features (2)	Improved RF	Overall accuracy of 71.5% for original RF and 79.54% for Improved RF. User`s accuracy of classes in improved RF ranged from 26.8% to 97.02%
Geiss et al., (2015)	Padang, Indonesia	Multispectral IKONOS Image, LANDSAT data and nDSM	*7 classes:* Confined Masonry, Reinforced Concrete Low, Reinforced Concrete High, Unreinforced Masonry, Timber Frame Residential, Timber Frame Non-Residential, Steel Frame	IKONOS Shape Features (18), IKONOS Spectral First Order Features (24), IKONOS Spectral Second Order Features (12), IKONOS Spatial Features (8), nDSM 3D Features (6), LANDSAT Temporal Feature (1)	RF, SVM	Overall accuracy for SVM is 93% and RF is 82.5% for all input features.
Belgiu et al., (2014)	Biberach an der Riss, Germany	Airborne Laser Scanning Data	*3 classes:* Residential/Small Buildings; Apartment Buildings; and Industrial and Factory Building	Shape Features (11), Area, Mean Height, Slope	RF	F1 score for classes ranged between 51% to 97.7%. Residential/Small Building achieved the highest F1 score.
Lu et al., (2014)	Denver, USA	LiDAR	*3 classes*: Single Family House, Multiple Family House, Non-residential Building	Basic Statistic Features (8), Shape Features (12), Building Spatial Relationship Features (9), Building and other Landuse Feature Relationship (12)	DT, ADT, RF, SVM	Overall accuracy for algorithms ranged from 67.5% to 76.1% for urban building context. However, no class-based model performance result was presented.

Table 2. Overview of related work (architectural building type classification)

Authors	**Study Area**	**Data**	**Architectural Building Class**	**Input Features**	**Algorithm**	**Model Performance**
Castagno & Atkins, (2018)	Witten, Germany; New York, & Ann Arbor, USA	Satellite Imagery and LiDAR	*8 roof classes:* Unknown, Complex flat, Gabled, Half-hipped, Hipped, Pyramidal, Skillion	Not available	RF	Accuracy of 87.2% when combined satellite imagery and LiDAR data is used.
Hecht et al., (2015)	Dresden, Halle, Krefeld, Stolpen and rural areas in Saxony, Germany	Topographic Raster Maps, cadastral databases, digital landscape model, 3D model	*11 classes:* Detached, semi-detached, Terraced, Multi-Family Houses, High Rise Buildings, Traditional Row Houses, Industrial Row Houses, Block perimeter Development, Rural houses, Industrial/commercial, Special purpose	Not available	RF	User accuracy for classes in Dresden dataset ranged from 63.2% to 100%
Wurm et al., (2015)	Berlin and Munich, Germany	Cadastral map, Stereoscopic Airborne images	*5 classes*: Perimeter Block Development, Block Development, Terraced Houses, Semi/Detached Houses, Halls	Shape based Features (26)	LDA	Mean Kappa of 0.9434 for Berlin data and 0.9355 for Munich data
Goel et al., (2012)	Europe	Images	*5 classes:* Art Noveau, Baroque, Gothic, Renaissance, Romanesque	Shape, Geometric Blur, HOG, Spatial Pyramid, Dense SIFT	WM, SVM, WM + SVM	Accuracy for WM, SVM and both combined were 32.64%, 46.22% and 48.25% respectively.
Henn et al., (2012)	Bonn, Germany	3D City Model	*7 classes:* Detached, Semi-Detached, Terraced, Wilhelminian Style, Villa, Apartment, Huge Dimensioned Building	Geometric Features (10), Neighbourhood Features (6), Infrastructure features (9)	SVM	Recall of classes ranged from 54.84% to 98.30%. Villa consisted of the lowest recall whereas detached building has maximum recall.

Chapter 3
Methods

3.1 Methods

Random Forest algorithm (refer to section 3.2.1) and Dense Neural Network (refer to section 3.2.2) are designed and implemented to achieve the goal of the proposed study.

3.1.1 Random Forest

Random Forest, introduced by Breiman is an ensemble classifier which consists of many decision tree (DT) classifiers (Breiman, 2001). A final classification is determined by considering the prediction of all the decision trees, and a class that receives the most votes is selected. This approach is often referred to as the 'wisdom of the crowd'. This collective decision based on maximum voting overcomes the limitation of relying on a single decision tree for classification. In the RF classifier, each tree is trained using randomly selected subset of training data and a subset of variables with a replacement bagging approach (Breiman, 2001; Pal, 2005). Therefore, some samples and variables may be sampled for several times and others may not be selected at all. This randomness and diversity among trees, achieved by combining fewer samples and variables (In-bag-data), introduce a higher bias but reduce variance, mitigating the risk of overfitting (Géron, 2022). Around one third of samples not used in the training, known as out-of-bag (OOB) samples can be employed for independent model evaluation without requiring separate validation set (Tooke et al., 2014). OOB error is calculated by comparing an aggregated ensemble prediction to true labels of the OOB samples, which helps to understand performance of RF in generalising to new and unseen data.

RF uses Gini Index for selection of variables, where the Index calculates the probability of randomly chosen element being incorrectly classified based on the distribution of classes at a particular node (Breiman, 2001; Pal, 2005). The lower the Gini Index indicates less impurity, and a better attribute for splitting. Gini Index can be calculated using formula below,

$$Gini\ Index = 1 - \sum_{i=1}^{n} (P_i)^2$$

where P_i is the probability of an element being classified into class *i*.

In RF model, users can define two parameters: a set of variables to generate tree and number of trees in the forest. Studies suggest that model performance is less sensitive to number of trees than to a set of variables chosen for best split in generating a tree. An advantage of RF is its ability

to interpret variable or feature importance which helps in understanding relative influence of a variable in overall prediction or classification model of random forest. Therefore, defining an appropriate number of variables and decision trees helps enhance model performance by addressing overfitting and underfitting (Belgiu & Drăguţ, 2014; Pal, 2005; Tooke et al., 2014). However, RF is difficult to interpret and visualise in comparison to DT (Breiman, 2001; Géron, 2022).

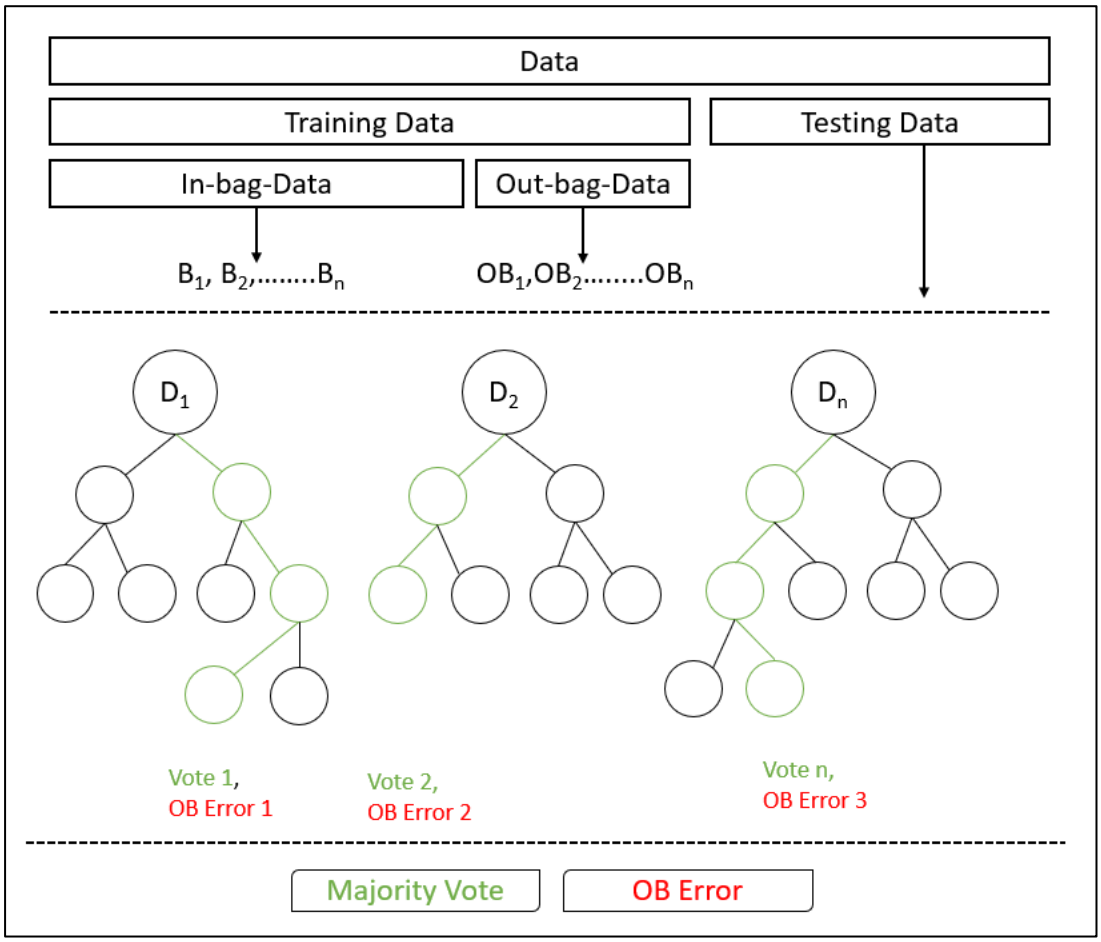


Figure 1. Random Forest Classifier

Figure 1 shows conceptual diagram for Random Forest Classifier algorithm, where data is divided into training and testing sets. All trees are grown in the forest using sub samples (for example, B_1, B_2….B_n) also known as 'in-bag' data from training set. Corresponding sub-samples (for example: OB_1, OB_2, …..OB_n) are called 'out-of-bag' data which are used for calculating out-of-bag error. When sample is introduced for classification, trees in the forest generate prediction and final classification obtained using majority voting among the trees.

3.1.2 Dense Neural Network

Dense Neural Network (DNN) represents the most straightforward artificial neural network architecture, where each neuron is connected to every neuron in both the preceding and succeeding layers (Agatonovic-Kustrin & Beresford, 2000; Géron, 2022). Neuron or also known

as perceptron is an essential component of neural networks. The primary difference between a neuron (Threshold Logic Unit) and a perceptron lies in their input and output handling: the former utilises binary numbers, while the latter deals with numerical values (Géron, 2022). Each neuron or perceptron multiple inputs, applies weights to them, sums them and passes result through an activation function to produce an output as shown in Figure 2. Step, sigmoid, hyperbolic tangent (tanh), and rectified linear unit (ReLU) are common activation functions.

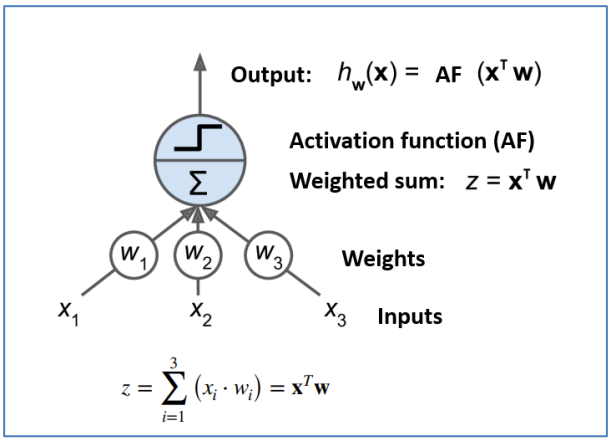


Figure 2. An artificial neuron or perceptron, adapted from (Géron, 2022, p.282)

A DNN comprises of an input layer, one or more hidden layers, and output layer as shown in Figure 3. Networks with only one hidden layer is called shallow network whereas networks with more than one hidden layer is referred to as deep networks.

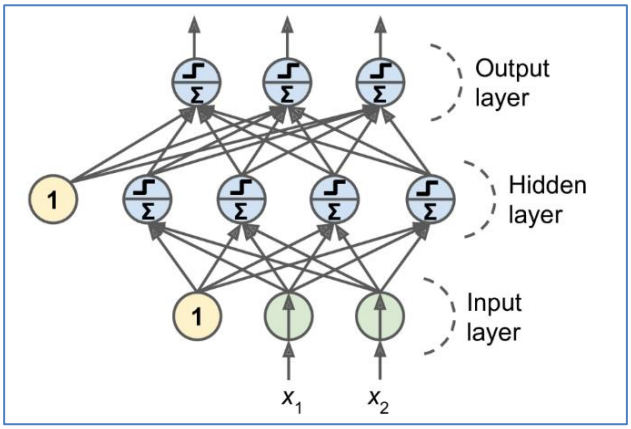


Figure 3. Dense Neural Network (Géron, 2022, p.286)

The DNN proposed in this study is a feedforward neural networks (FFNN), where information flows in one direction only- from input layer through hidden layer(s) to output layer (Abiodun et al., 2018). Application of FFNN includes but not limited to time series prediction and classification. Examples of FFNN algorithms Single layer perceptron, multilayer perceptron, and radial basis function network. Nazari & Yan, (2021) and Turhan et al., (2014) implemented DNN to study building energy requirements based on building shape and other features.

3.2 Experimental Setup

This section explains different approaches and methods applied in implementing both Random Forest and Dense Neural Network. Figure 4 shows the systematic workflow of the study.

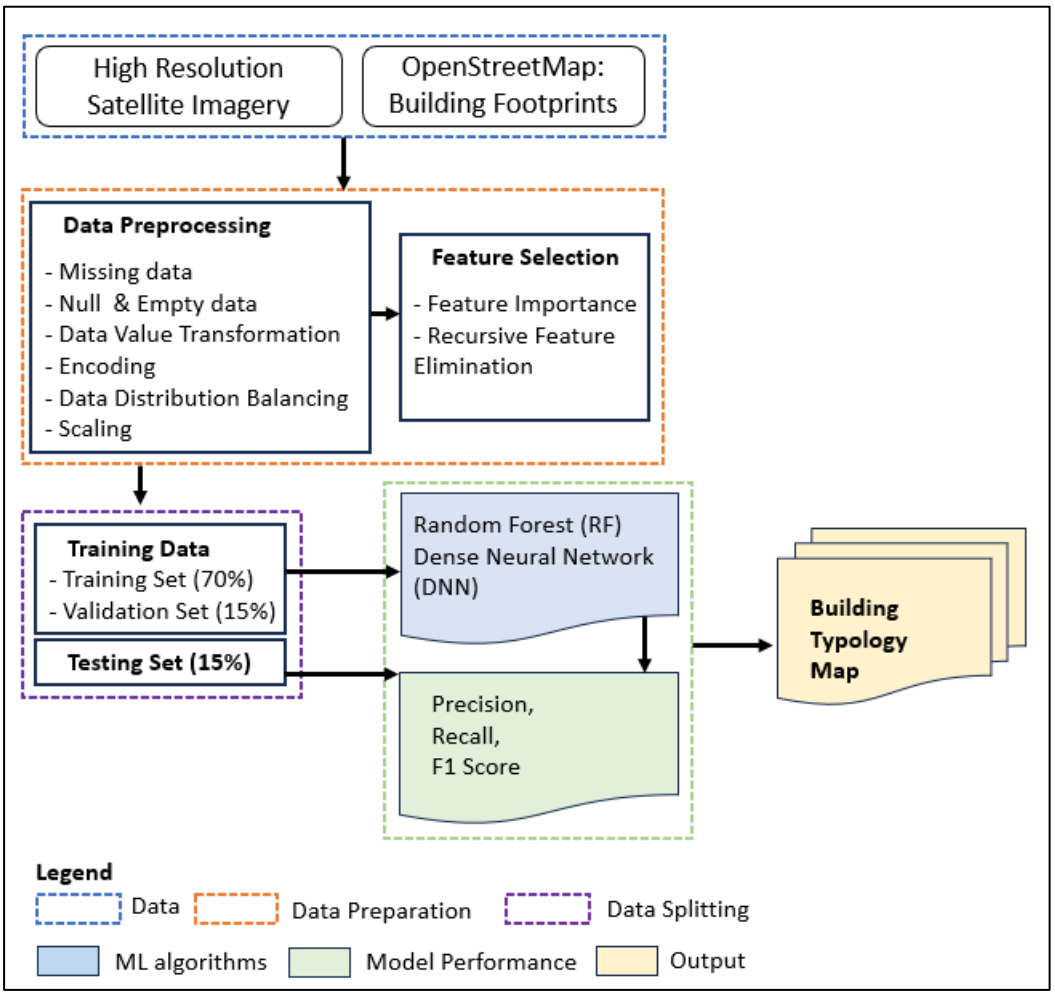


Figure 4. Experimental set up for the proposed study

3.2.1 Study Area and Data

A city of Cologne, Germany is selected as the primary study area, primarily for availability of data and as representation diverse urban buildings context. A dataset by Geiss et al., (2017) is used for this study. It consists of consists OSM building footprints and has integrated around 152 features calculated using eCognition software and ArcGIS software in three spatial levels- individual building; aggregated buildings; and building blocks. Spectral indices were computed using four bands multi-spectral and VHR satellite imagery obtained from WorldView II. Table 3 provides the complete list of features from the dataset.

Table 3. List of features from a dataset by Geiss et al., (2017, pp.3-5)

Feature Category	Features	Number of Features		
		Individual Building	Aggregated Building	Building Block
Geometry-Extent 2D	Area, length, width, length/width, length of main line, width of main line, length of main line/width of main line, perimeter	8	8	X
	Inverted floor area ratio	1	X	X
	Area of BB, average building area based on IB within BB, sum of building areas within BB, sum of bushes/trees areas within BB, sum of meadow areas within BB, sum of impervious and bare soil areas within BB, average length/width based on IB within BB, average length/width based on AB within BB, average length of main line/width of main line based on IB within BB, average length of main line/width of main line based on AB within BB	X	X	10
Geometry-Shape-2D	Number of building vertices, shape index, fractal dimension, density, border index, compactness, rectangular fit, radius of largest enclosed ellipse, radius of smallest enclosing ellipse, elliptic fit, roundness, normalized perimeter index, normalized proximity index, normalized spin index, areal asymmetry	15	15	X
Geometry-Extent 3D	Building height based on maximum height, building height based on median height	2	X	X
	Building height based on maximum height within AB, area weighted building height based on maximum heights of IB within AB, building height based on median height within AB, area weighted building height based on median heights of IB within AB	X	4	X
	Highest IB based on maximum height within BB, area-weighted building height based on maximum heights of IB within BB, median of building height based on median heights of IB within BB, area-weighted building height based on median heights of IB within BB,	X	X	4
	Building volume(s) based on maximum height(s), building volume(s) based on median height(s)	2	2	2
Geometry-Shape-3D	Shape index 3D based on maximum height, shape index 3D based on median height, asymmetry 3D based on maximum height, asymmetry 3D based on median height	4	4	X
	Shape index 3D based on area weighted maximum heights of IB, shape index 3D based on area-weighted median heights of IB, asymmetry 3D based on area weighted maximum heights of IB, asymmetry 3D based on area weighted median heights of IB	X	4	X

Spatial Context	Orientation of building according to major axis of minimum bounding rectangle, distance to nearest building, distance to city center (Cologne Cathedral)	3	3	X
	Standard deviation of orientations of IB, Standard deviation of orientations of AB, median of distances to nearest building based on IB, median of distances to nearest building based on AB, number of IB within BB, number of AB within BB, median of inverted floor area ratio based on IB, building aggregation measure, normalized building aggregation measure, vegetation fraction, impervious surface area, urban density, vegetation volume to built-up volume, urban vegetation index, site occupancy index	X	X	15
Spectral- First Order	Mean(blue), mean(green), mean(red), mean(NIR), StDev(blue), StDev(green), StDev(red), StDev(NIR), mean(blue) / mean(green), mean(blue) / mean(red), mean(blue) / mean(NIR), mean(green) / mean(red), mean(green) / mean(NIR), mean(red) / mean(NIR), mean(green) / [mean(green)+ mean(red)+ mean(NIR)], mean(red) / [mean(green)+ mean(red)+ mean(NIR)], mean(NIR) / [mean(green)+ mean(red)+ mean(NIR)], brightness, normalized differenced vegetation index, soil-adjusted vegetation index	20	20	X
Spectral- Second Order	$GLCM_{inv}$ (Angular 2nd moment), $GLCM_{inv}$ (entropy), $GLCM_{inv}$ (homogeneity)	3	3	X
Total Features		**58**	**63**	**31**

3.2.2 Data Preprocessing

In this step various activities of data cleaning and preparation for training RF and DNN was carried out as below,

***Data exploration*:** For all the input variables, null value, empty and zero in data sample were explored. Where necessary, new variable was created, for example, target variable for classification consisted of string data sample, therefore, new corresponding variable was created to compute numeric values.

Data Value Transformation: A target variable with string categorical data was converted into numerical values using scikit- learn`s 'LabelEncoder[1]'. RF algorithm requires numerical values for

[1] https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.LabelEncoder.html

the input and target variables because multiple decision trees are trained on different subsets of data and their predictions are combined through voting for classification.

For DNN, data for target variable is converted using one-hot encoding which transforms categorical numerical encoding into a binary matrix where each column represents a unique class, and each row represents a sample. In multiclass classification, where 'Categorical Crossentropy Loss' is commonly used as loss function, one-hot encoding provides efficient way to represent true distribution (Morales-Molina et al., 2021). One-hot encoding[2] can be implemented using 'to_categorical' function from the 'tensorflow.keras.utils' library.

***Balancing Data Distribution*:** As shown in Figure 5, data distribution among target variable classes were imbalanced with a class with minimal data sample of 3.15% and another class with maximum of 39.73% and other classes in between.

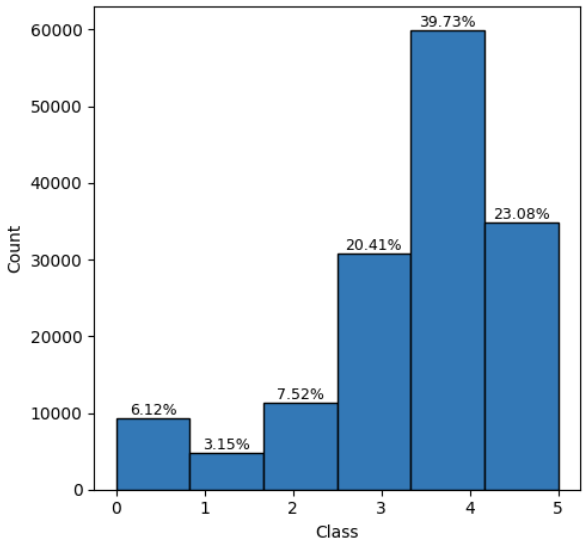


Figure 5. Distribution of data across six classes

Such imbalance in dataset can lead to problem always generalising majority class over the minority class (Maxwell et al., 2018; Morales-Molina et al., 2021). To address this issue, for both RF and DNN using scikit-learn and TensorFlow library, 'class_weight' parameter can be defined as 'balanced', which automatically adjusts the weight based on inverse of class frequencies. Mathematically, class weight can be expressed as below,

$$W_c = \frac{N}{Kx\ n_c}$$

[2] https://www.tensorflow.org/api_docs/python/tf/keras/utils/to_categorical

Where, W_c is a class weight; N is total number of samples in training set; K is the total number of classes; and N_c is number of samples belonging to sample c.

***Scaling*:** Numerical value of input variables or features were standardized using scikit-learn`s 'StandardScaler' function to ensure equal weights during learning process and therefore, each feature was transformed to have a mean of 0 and a standard deviation of 1.

3.2.3 Feature Selection

While not all the input variables or feature contribute equally to multiclass classification, most significant features were selected. Given sparsely distributed nature of the data across six classes, and 152 variables, 'ExtraTreesClassifier[3]' form scikit-learn was employed to select most significant features. This ensemble learning method grows multiple decision trees and combines their prediction to improve accuracy and control overfitting by balancing high against low variance (Géron, 2022).

In addition, Recursive Feature Elimination[4] (RFE) based feature selection technique was also experimented using 'RandomForestClassifier' algorithm. RFE process involves fitting the model (in this case Random Forest), ranking the features based on their importance, eliminating the least important feature, and repeating the process until the desired number of features is reached.

3.2.4 Implementation of RF and DNN

Google Colab Pro was used for training both Random Forest (RF) and Dense Neural Network (DNN). This cloud-based platform provided System RAM ranging from 35 GB to 51 GB, disk storage capabilities up to 225 GB, and a variable number of CPU cores, ranging from 4-40 depending on the availability.

Data Splitting: The dataset was divided into three sets: Training (70%), Validation (15%), and Testing (15%). When splitting data, a parameter '*stratify = y'* (target variable) was used to ensure that data will be split in a way that maintains the same distribution of classes as in the original dataset. This means the proportion of each class in target variable remains consistent across the training, validation, and testing sets. This consistency is valuable where data distribution among target variable classes is imbalanced.

[3] https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.ExtraTreesClassifier.html
[4] https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.RFECV.html

Random Forest (RF) was trained using RandomForestClassifier[5] from scikit-learn library. While training RF, following parameters as shown in Table 4 were defined.

Table 4. Parameter used in training Random Forest

Parameter	Value	Meaning
n_estimators	300	number of trees in the forest
max_depth	15	maximum depth of tree
class_weight	balanced	adjust weight inversely proportional to class frequencies
criterion	gini	Gini Impurity
bootstrap	TRUE	bootstrap samples with replacement

Optimal values n_estimators, and max_depth was explored by hyperparameters tuning using GridSearchCV[6]. During GridSearch, ten folds stratified cross validation was used and was evaluated against F1 score.

Dense Neural Network (DNN) was trained using Keras[7] for TensorFlow library. A DNN model used in this study consists of input layer, one hidden layer, and output layer (refer to Figure 6).

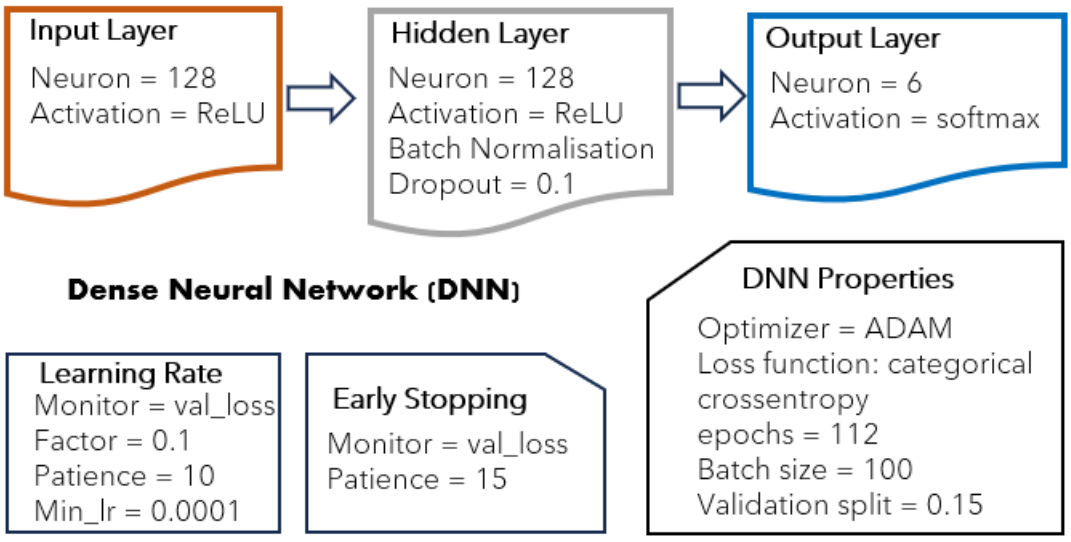


Figure 6. Dense Neural Network

Input layer consists of 128 units or neurons, with each neuron receiving input from input variables. ReLU, which introduced non-linearity by replacing negative input values with zero and leaving positive values unchanged is used as an activation function in both input layer and hidden layer (Shrestha & Vanneschi, 2018). And softmax activation is used in output layer for its suitability in multiclass classification. Categorical Crossentropy is defined as loss function, aligning with one-hot encoding of target variable, and addressing multiclass classification where

[5] https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html
[6] https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html
[7] https://www.tensorflow.org/api_docs/python/tf/keras/layers/Dense

each input belongs to one exact class among several classes. Training of DNN is facilitated by Adaptive Moment Estimation (Adam) optimisation algorithm.

To mitigate overfitting of the model, regularization technique known as Dropout at 20% rate is applied in hidden layer. This means that during training, around 20% of the neurons in the preceding layer will randomly drop out in each iteration. Also, the early stopping function is implemented by monitoring validation loss for at least fifteen consecutive iterations or epochs.

3.2.5 Model Evaluation

For better evaluation of models in multiclass classification, precision, recall and F1 score are considered appropriate evaluation metrics (Morales-Molina et al., 2021; Zhou & Chang, 2021). True Positives and True Negative are samples that are correctly classified, whereas False Negative and False Positives are samples that are incorrectly classified.

Precision: It is the percentage of predicted positives that were correctly classified.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

Recall: It is percentage of actual positives that were correctly classified.

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

F1 Score: It is weighted average of precision and recall which considers both false positives and false negative.

$$F1\ Score = 2\ X\ \frac{Precision\ \mathrm{X}\ Recall}{Precision + Recall}$$

Macro average of precision, recall, and F1 score treats all classes equally and compute the average without considering class size. Weighted average of above-mentioned metrics takes account number of sample in each class. Larger classes have more influences on the average, making it sensitive to class imbalance.

Chapter 4
Result and Discussion

This chapter presents the results of this study, consistent with objectives and associated research questions raised in section 1.3. Building footprints are classified into six classes and models` (RF and DNN) performance are evaluated and compared using evaluation metrics- precision, recall, and F1 score; confusion matrix and building footprint typology maps.

4.1 Most significant input features

Question 1.a: *How are the most significant features distributed across three spatial levels- individual building, aggregated building, and building block- in characterising building inventories?*

When implementing RFE with RF (refer to section 3.2.3) and evaluated with F1 score evaluation metric and five-folds cross validation (CV), F1 score largely increased up to five input variables. The F1 score continued to gradually rise until reaching a plateau when the number of input variables reached 35, as shown in Figure 7.

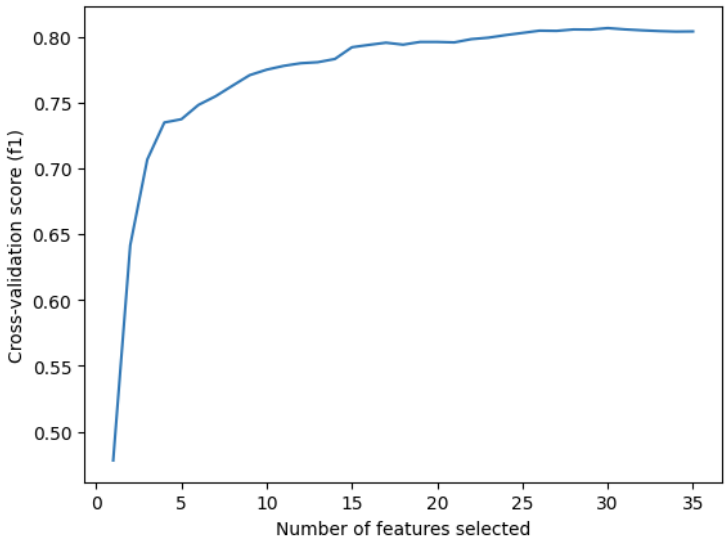


Figure 7. Recursive Feature Elimination (RFE) with RF and CV=5

Examining feature importance through 'ExtraTreeClassifier' and determining number of features necessary for a stable model performance using RFE, 35 most significant features, as shown in Table 5, were chosen for training Random Forest and Dense Neural Network. Among the selected 35 features, 54% originated from the IB category, 37% from the AB category, and 9% from the BB category. This result corresponds to addressing 'Question a' of the first objective in this study.

Table 5. Distribution of most significant features across three spatial levels

Spatial Level	Feature Rank	Feature	Acronyms
Individual Building (IB)	1	Inverted Floor Area Ratio	s_geom_iFA
	2	Building Height based on Median Height	s_geom_h_1
	8	Shape Length	Shape_Leng
	5	Proximity Index	s_circ_pro
	7	Building Perimeter	s_geom_per
	6	Building Volume based on Median Height	s_geom_vol
	12	Building Area	s_geom_are
	11	Shape Area	Shape_Area
	17	Length	s_geom_l_1
	9	Building Width	s_geom_wid
	15	Area Asymmetry	s_asym_2D
	25	Spin Index	s_circ_Spi
	27	Shape Index based on maximum height	s_cubi_SI_
	29	Building Height based on Maximum Height	s_geom_hgt
	18	Width of main line	s_geom_w_1
	33	Density 2D	s_squa_den
	35	Distance from city Center (Cologne Cathedral)	s_geom_dis
	28	Asymmetry 3D based on median height	s_asym_3D1
	25	Normalised Spin Index	s_circ_nSp
Aggregated Building (AB)	3	Width of Aggregated Building	m_geom_wid
	13	Area Asymmetry	m_asym_2D
	19	Area weighted building height based on maximum height of IB within AB	m_geom_h_1
	14	Width of main line	m_geom_w_1
	10	Volume based on median height	m_geom_v_1
	21	Volume based in maximum height	m_geom_vol
	16	Building height based on median height within AB	m_geom_h_2
	24	Aggregated Building Area	m_geom_are
	22	Building height based on maximum height within AB	m_geom_hgt
	32	Area-weighted Building Height Based on Median Height of IB within BB	m_geom_h_3
	26	Proximity Index	m_circ_pro
	30	Distance from city Center (Cologne Cathedral)	m_geom_dis
	34	Asymmetry 3D based on Median Height	m_geom_asym_3D 1
Building Block (BB)	4	Vegetation Volume to Built up Volume	b_rel_VV2B
	20	Impervious Surface Area	b_rel_ISA
	23	Urban Density	b_rel_UD

It is noteworthy that, among the 35 most significant input features, none originated from spectral attributes, despite the original dataset containing 46 spectral features. This result aligns with the finding of Du et al., (2015), who stated that spectral features demonstrate poor performance due to confusion among different urban buildings. All 35 selected input variables show geometric characteristics. Both studies by Du et al., (2015), and Fan et al., (2014) argued that texture and geometric feature such as area, width, length, and perimeter significantly contribute in urban building classification. In addition, Tooke et al., (2014) emphasised the importance of building height as an important attribute in building characterisation, and in line with this, the current study identified 13 input variables associated with building heights.

4.2 Building inventory characterisation using Random Forest

Question 1.b: *What is the performance of Random Forest in characterising building across six classes?*

RF classification results for multiclass classification, as detailed in section 3.2.4 and Table 4 for detail experimental set up, are presented in Table 6. The overall accuracy for testing set is 75.9%. Given that there are six classes for urban building characterisation, the overall accuracy may not reflect classification accuracy of each class. Evaluation metrics such as precision, recall, and F1 score provide better understanding of model performance in relation to specific classes (Morales-Molina et al., 2021; Zhou & Chang, 2021).

Table 6. Classification report for Random Forest Classifier

Building Class	**Training Set**			**Validation Set**			**Testing Set**		
	Precision	Recall	F1 score	Precision	Recall	F1 score	Precision	Recall	F1 score
Building block in closed construction	0.553	0.964	0.702	0.469	0.787	0.588	0.467	0.785	0.586
Detached building block	0.629	0.980	0.766	0.527	0.832	0.646	0.514	0.802	0.627
Free-standing individual building	0.795	0.920	0.853	0.666	0.723	0.693	0.665	0.717	0.690
Garage	0.835	0.972	0.899	0.786	0.905	0.841	0.789	0.908	0.844
Miscellaneous	0.965	0.706	0.816	0.836	0.630	0.719	0.836	0.632	0.720
Row Development	0.910	0.903	0.906	0.844	0.847	0.846	0.843	0.848	0.845
accuracy			0.846			0.759			0.759
macro average	0.781	0.908	0.824	0.688	0.787	0.722	0.686	0.782	0.719
weighted average	0.877	0.846	0.848	0.783	0.759	0.761	0.783	0.759	0.761

In Table 6, precision among six urban building classes ranges from 46.7% to 84.3%. While 'Building block in closed construction' class has the lowest precision, 'Row Development' class has the highest precision. Similarly, recall for the classes ranges from 63.2% to 90.8%. Three classes- 'Detached building block', 'Garage', and 'Row Development' achieved recall more than 80%, while five classes attained recall of 70%. Macro average of F1 score is 71.9%. Classification performance of the RF across six classes varies and a very possible reason is uneven distribution of data samples across the target classes. For example, 'Detached building' consists of only 3.15% of samples while the 'Miscellaneous' class comprises 39.73% of data samples.

The current RF, implemented with 'n_estimator = 300' and 'adepts = 15', exhibits better performance, as shown in Figure 8, compared to overfitted RF model with 'n_estimator = 100' and 'max_depth = None', as shown in Figure 9. Overall accuracy and F1 score are higher for Training Set than Validation Set. It is noted that that learning curve for both training and validation sets are converging, as shown in Figure 9, representing a better model performance. Evaluation metrics value for both Validation and Testing show similar characteristics. Optimal value for 'n_estimator', and 'max_depth' were identified by hyperparameter tuning using GridSearchCV.

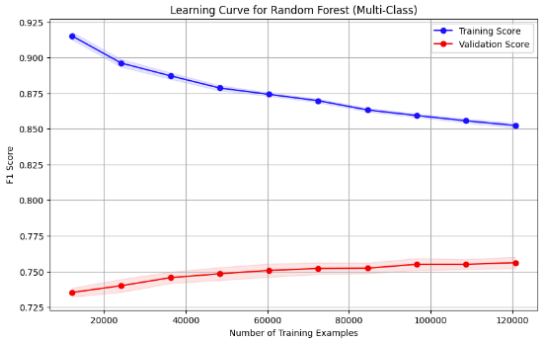


Figure 8. Proposed RF model

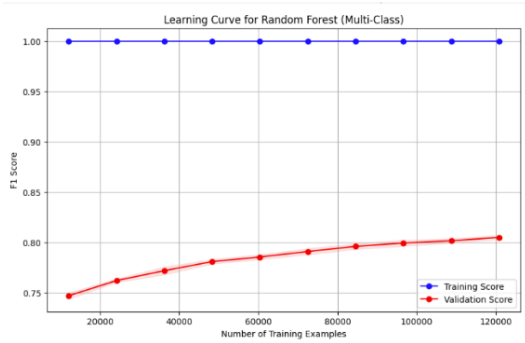


Figure 9. Overfitted RF model

Figure 10 shows the map where urban building inventories in city of Cologne, Germany are characterised using RF classification. In addition, Figure 11 show accurately characterised and wrongly classified urban building typology classes using RF algorithm.

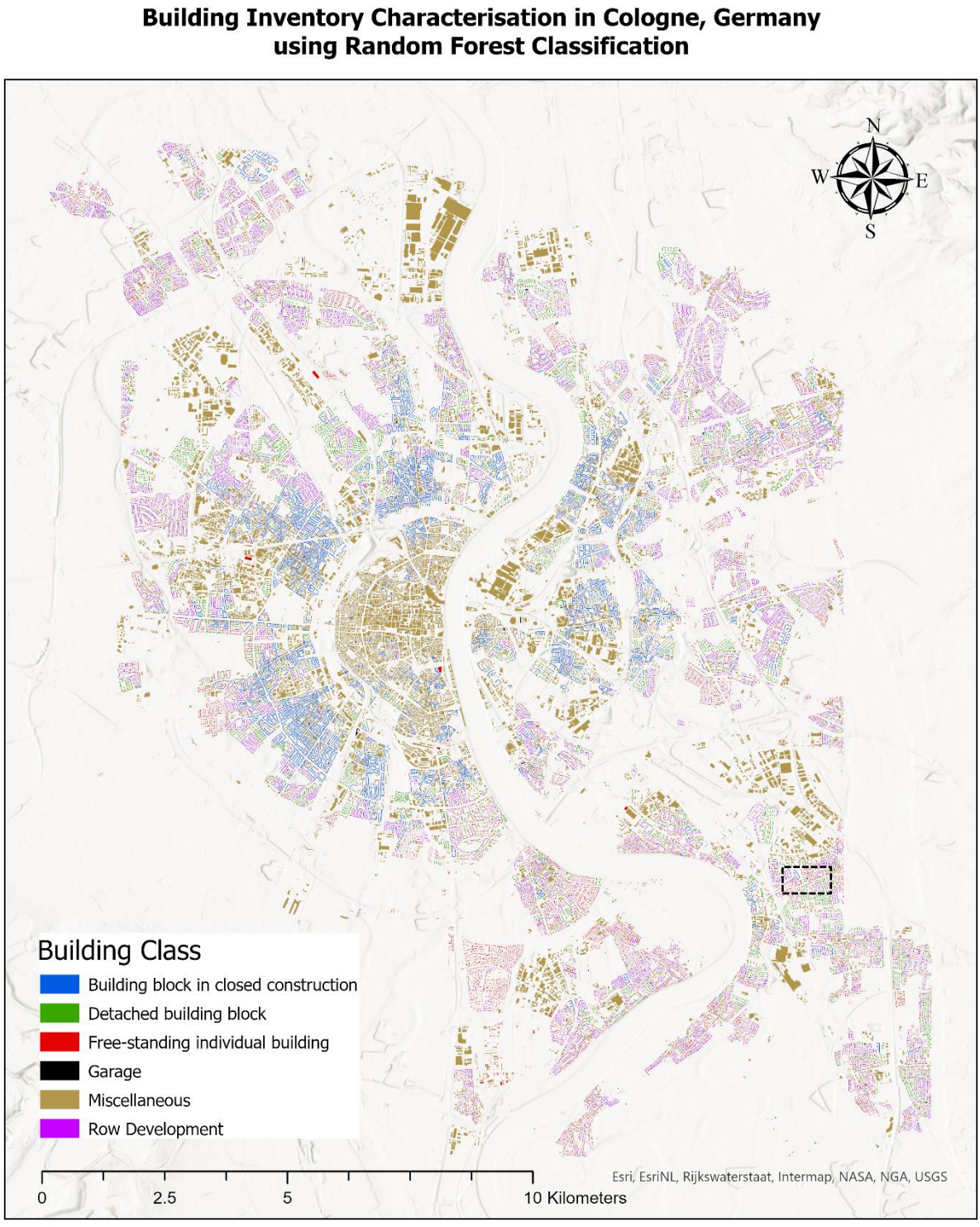


Figure 10. Urban Building Characterisation Map using Random Forest for Cologne, Germany

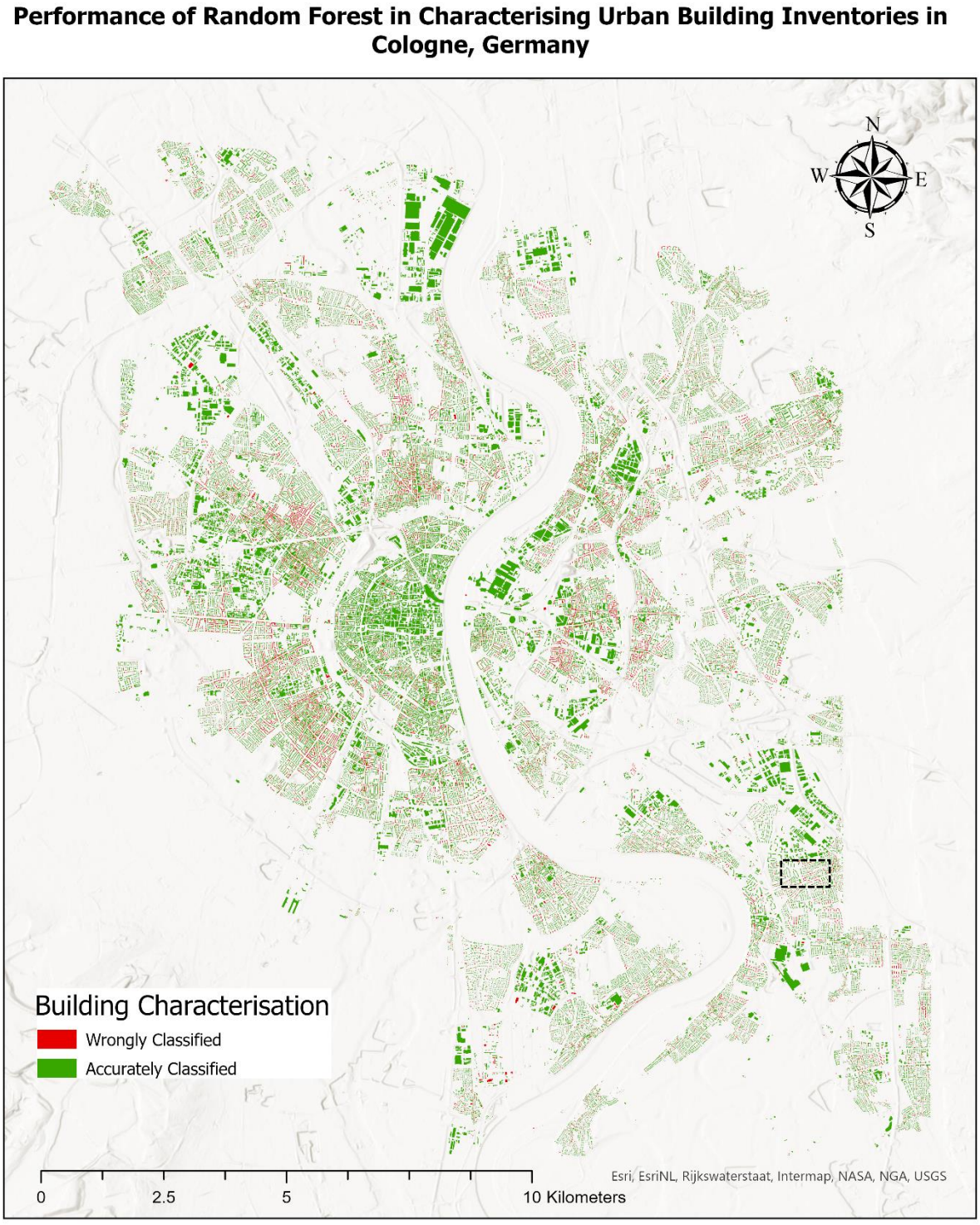


Figure 11. Map on Building Characterisation Performance using Random Forest for Cologne, Germany

4.3 Building inventory characterisation using Dense Neural Network

Question 2.a: *How does Dense Neural Network perform in characterising buildings across six classes?*

DNN classification results for multiclass classification, as detailed in section 3.2.4 and Figure 6 for detail experimental set up, are presented in Table 7. The overall accuracy for testing set is 71.8%.

Table 7. Classification Report for Dense Neural Network

Building Class	**Training Set**			**Validation Set**			**Testing Set**		
	Precision	Recall	F1 score	Precision	Recall	F1 score	Precision	Recall	F1 score
Building block in closed construction	0.416	0.870	0.563	0.393	0.816	0.531	0.389	0.816	0.527
Detached building block	0.488	0.951	0.645	0.426	0.833	0.564	0.434	0.853	0.575
Free-standing individual building	0.567	0.825	0.672	0.537	0.787	0.639	0.538	0.782	0.637
Garage	0.777	0.924	0.844	0.772	0.922	0.840	0.776	0.919	0.841
Miscellaneous	0.887	0.537	0.669	0.870	0.525	0.654	0.869	0.525	0.654
Row Development	0.862	0.816	0.838	0.846	0.798	0.821	0.850	0.806	0.828
accuracy			0.735			0.716			0.718
macro average	0.666	0.820	0.705	0.641	0.780	0.675	0.643	0.783	0.677
weighted average	0.794	0.735	0.737	0.776	0.716	0.719	0.778	0.718	0.721

In Table 7, precision among six urban building classes ranges from 38.9% to 86.9%. While 'Building block in closed construction' class has the lowest precision, 'Miscellaneous' class has the highest precision. Similarly, recall for the classes ranges from 52.5% to 91.9%. Four classes- 'building block in closed construction', 'Detached building block', 'Garage', and 'Row Development' achieved recall more than 80%. Macro average of F1 score is 67.7%. Showing similar results as the RF, classification performance of the DNN fluctuates across six classes, reflecting impact of imbalanced data distribution across target classes. For example, DNN when implemented in absence of regularisation technique such as Dropout and without defining 'class_weight' showed classification accuracy biased towards 'Miscellaneous' and 'Row Development' classes consisting large data sample distribution.

Figure 12 shows the map where urban building inventories in city of Cologne, Germany are characterised using DNN. In addition, Figure 13 shows a map illustrating building typology that were accurately classified as well as those that were incorrectly classified using DNN.

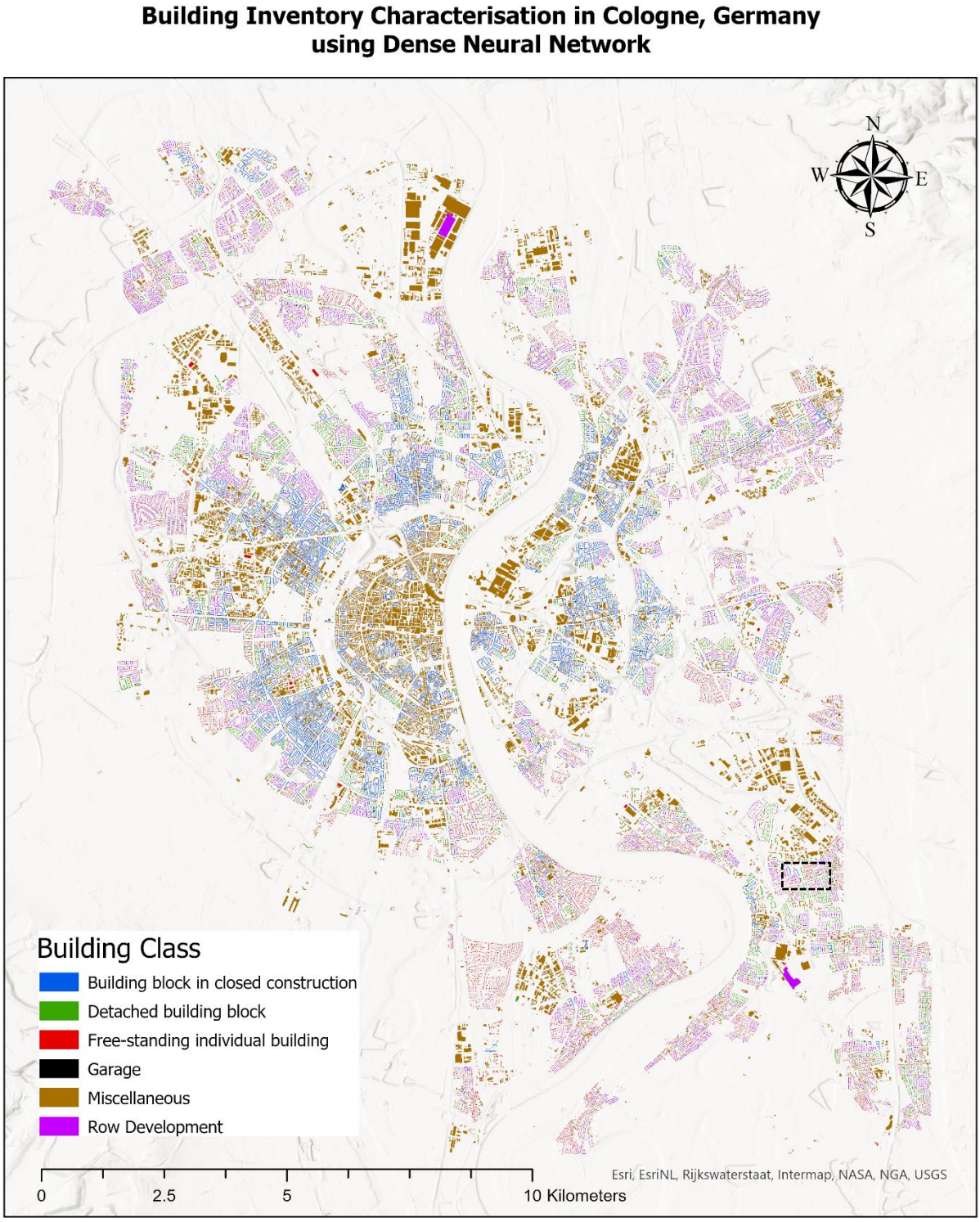


Figure 12. Urban Building Characterisation Map using Dense Neural Network for Cologne, Germany

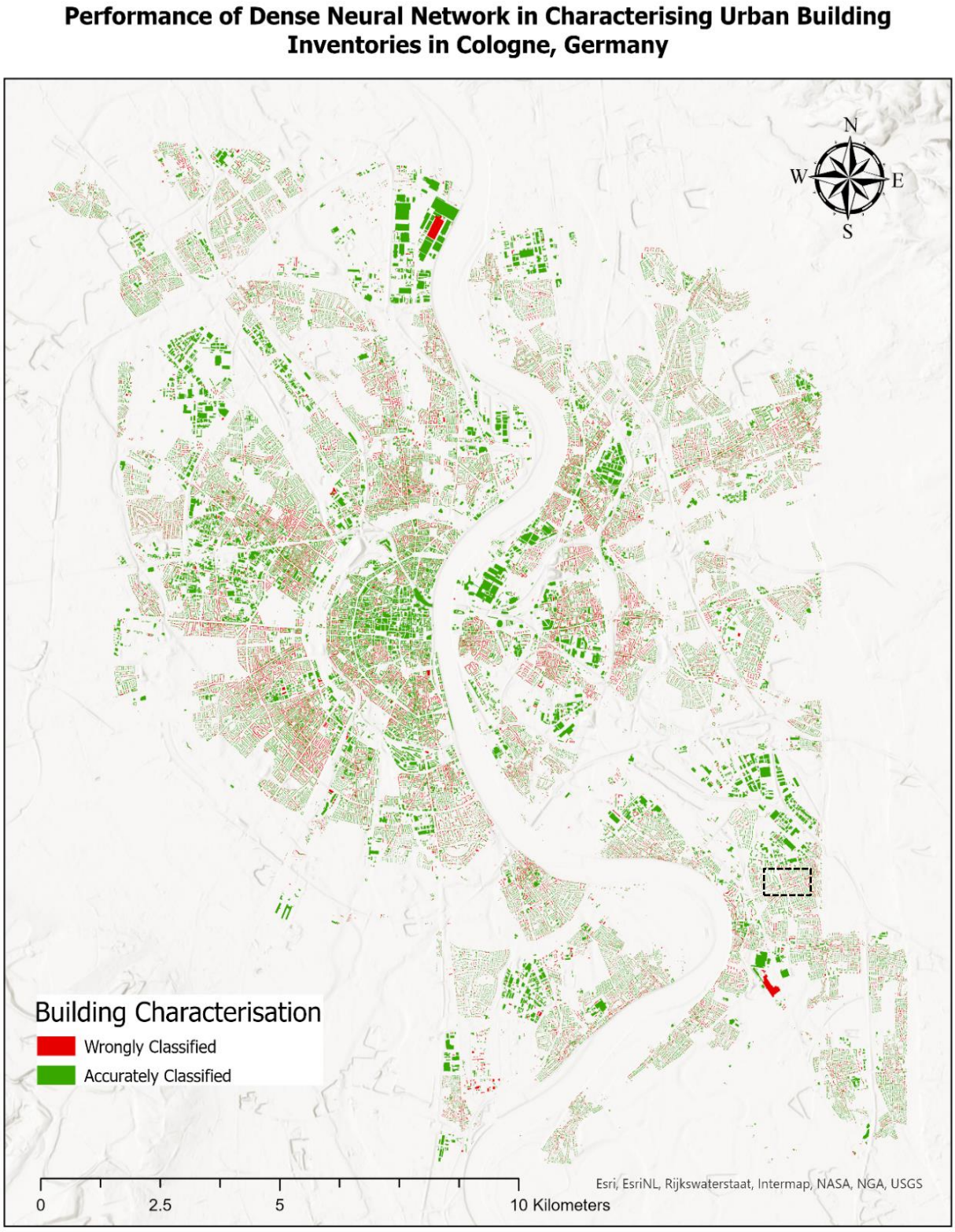


Figure 13. Map on Building Characterisation Performance using Dense Neural Network for Cologne, Germany

4.4 Comparing model performance of RF and DNN

Question 2.b: *What difference exist in the model performance compared to Random Forest in characterising buildings across six classes?*

Based on findings presented in Table 6 and Table 7, the Random Forest model shows a higher overall accuracy (79.9%) compared to Dense Neural Network (71.9%). Upon closer analysis of performance among target classes in this multiclass classification, it was found that recall rate using DNN is greater for four classes- 'Building block in closed construction', 'Detached building block', 'Free standing individual building', and 'Garage', in comparison to that of RF. However, in term of precision rate, RF model performs better over DNN model.

Figure 14 and Figure 15 shows the confusion matrix of Testing set for RF and DNN respectively. When diagonal elements, representing the number of correctly classified samples for each class are observed, it is found that the DNN outperforms RF in correctly classifying more instances for four classes. 'Miscellaneous' class which consists of the largest data samples of around 40% is better classified by the RF model. However, the nature of this class is somewhat ambiguous, requiring further study for a better understanding. For 'Row Development' class, RF performs slightly better than DNN. Therefore, observing the confusion matrix, DNN model performs better in characterising building typology that are clearly defined and understood.

Figure 16 consists of a set of maps developed using randomly selected a small area within the study area of Cologne, Germany (refer to black rectangle in Figure 8-11). These maps aim to provide a sample of visual representation and interpretation in comparing urban building characterisation using two machine learning models- RF and DNN. While other parts of wider study area exhibit similar results, a set of maps presented in Figure 16 is not intended to be referred to as a general pattern. Given large number of building inventories, totalling 2,15,660 samples initiate challenges to present all building for visual interpretation in a single A4 size map and therefore, the very small subset presented in Figure 16 serves as a sample for visual interpretation.

In Figure 16, map A presents actual building inventory types or classes. Maps- B1, and B2 show automatic building inventory characterisation using RF and DNN respectively. Lastly, maps- C1, and C2 represent correctly and incorrectly classified building typology using RF and DNN respectively.

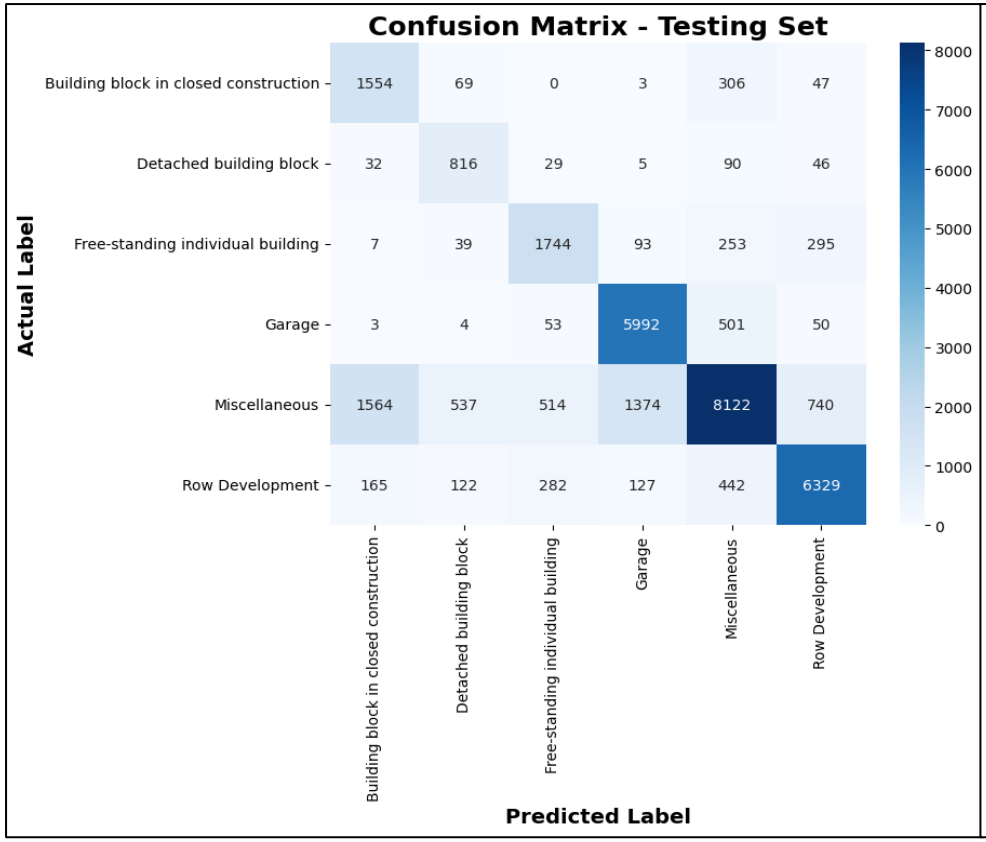


Figure 14. Confusion matrix for Random Forest Classifier

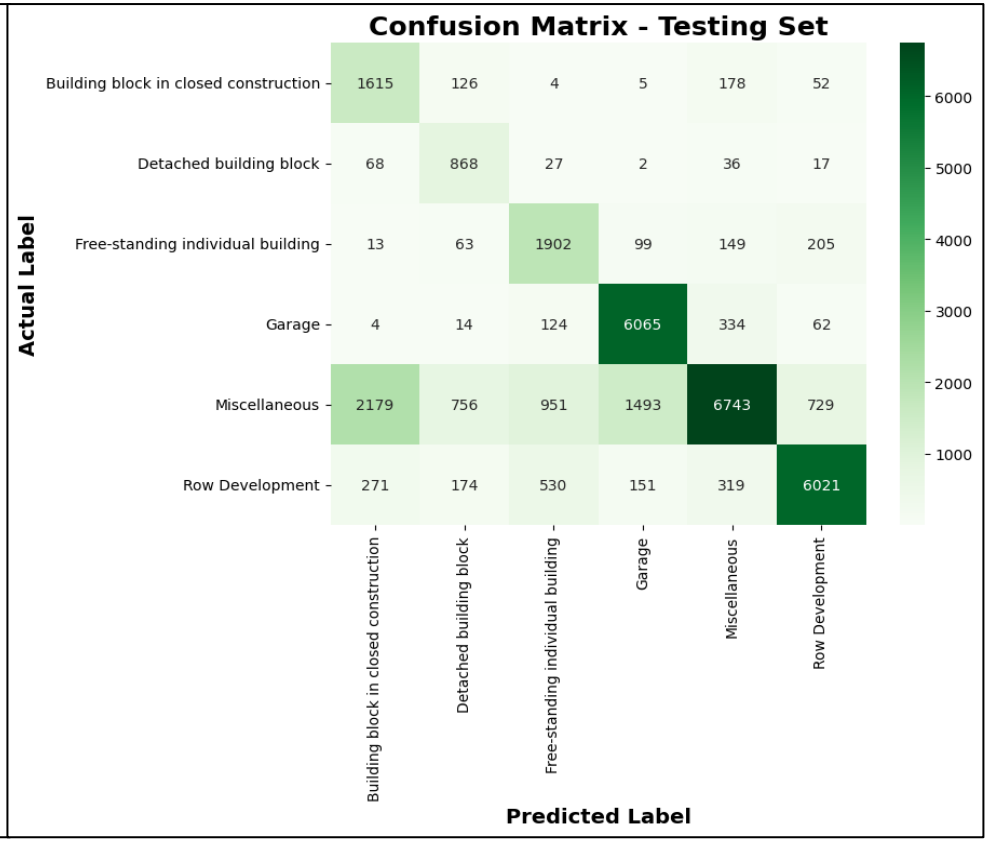


Figure 15. Confusion matrix for Dense Neural Network

A. *Actual Building Typology*

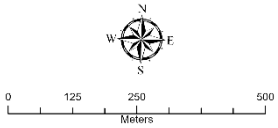
B1. *Building Characterisation using Random Forest*

C1. *Characterisation Performance of Random Forest*

For A, B1, and B2

Building Class

- Building block in closed construction
- Detached building block
- Free-standing individual building
- Garage
- Miscellaneous
- Row Development



For C1, and C2

Building Characterisation

- Wrongly Classified
- Accurately Classified

B2. *Building Characterisation using Dense Neural Network*

C2. *Characterisation Performance of Dense Neural Network*

Figure 16. Map based comparison on model performance of RF and DNN

Chapter 5
Conclusion

5.1 Limitation and Future Work

Some limitations of the study are: First, while the use of OSM building footprints provide reliable and convenient GIS data, thus avoiding additional work on image segmentation, certain difference in location may arise between OSM data and VHR image. This difference can impact classification performance. Both models incorporate input features previously computed using VHR imagery and OSM. The study assumes that all features calculation was carried out correctly. Second, the 'Miscellaneous' class consisting of the largest portion of data samples, is vague. Therefore, it requires further study to define or classify into sub-urban building typologies. In doing so, it may contribute to a more balanced data distribution, addressing the current issue of imbalanced data. Third, the Dense Neural Network implemented in this study, is a simplistic and feed forward neural network. Exploring other deep learning algorithms, such as Graph Convolutional Neural Network (Yan et al., 2019; Yang et al., 2022), by integrating VHR imagery as input image features and other vector data, and experimenting with post-classification methodologies, can potentially enhance the performance of urban building characterisation.

5.2 Conclusion

This study carried out automatic characterisation of urban building. From methodological perspective, RF machine learning model and DNN, a simple and feed forward neural network were experimented in imbalanced dataset. 'class_weight = balanced' parameter from sci-kit learn library, which automatically adjusts the weight based on inverse of class frequencies was used to address the issue of imbalanced data distribution. OSM building footprint data and 35 geometric features calculated using the VHR imagery were used. 19 features from individual building level, 13 features from aggregated building level, and 3 features from building block spatial level significantly contributed to models in urban building characterisation.

The experiments show that overall accuracy of RF (79.9%) is higher than that of the DNN (71.9%). However, upon examining and comparing diagonal elements, representing the number of correctly classified samples for each class in the confusion matrix, it is evident that the DNN outperforms RF in correctly classifying more instances for four classes- 'Building block in closed construction', 'Detached building block', 'Free standing individual building', and 'Garage'. Both models in this study uses geometric and distribution features calculated in three spatial levels, using VHR imagery and OSM data to classify urban building. Given that similar input features are

used for characterising urban building classes, these methods can be applied for other study areas.

Also, it is important to note that implementation of the DNN and comparison with traditional machine learning algorithm- RF provide additional scientific contribution, especially in situation where there is limited use of deep learning algorithms in the building characterisation.

Reference

Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A., & Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. *Heliyon*, *4*(11).

Agatonovic-Kustrin, S., & Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *Journal of Pharmaceutical and Biomedical Analysis*, *22*(5), 717–727.

Anand, A., & Deb, C. (2023). The Potential of Remote Sensing and GIS in Urban Building Energy Modelling. *Energy and Built Environment*.

Balram, S., & Dragicevic, S. (2006). Collaborative geographic information systems: Origins, boundaries, and structures. In *Collaborative geographic information systems* (pp. 1–23). Igi Global.

Bandam, A., Busari, E., Syranidou, C., Linssen, J., & Stolten, D. (2022). Classification of building types in Germany: A data-driven modeling approach. *Data*, *7*(4), 45.

Barrington-Leigh, C., & Millard-Ball, A. (2017). The world's user-generated road map is more than 80% complete. *PloS One*, *12*(8), e0180698.

Belgiu, M., & Drăguţ, L. (2014). Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high resolution imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, *96*, 67–75.

Belgiu, M., Tomljenovic, I., Lampoltshammer, T. J., Blaschke, T., & Höfle, B. (2014). Ontology-based classification of building types detected from airborne laser scanning data. *Remote Sensing*, *6*(2), 1347–1366.

Ben-David, S., Blitzer, J., Crammer, K., Kulesza, A., Pereira, F., & Vaughan, J. W. (2010). A theory of learning from different domains. *Machine Learning*, *79*, 151–175.

Beykaei, S. A., Zhong, M., Shiravi, S., & Zhang, Y. (2014). A hierarchical rule-based land use extraction system using geographic and remotely sensed data: A case study for residential uses. *Transportation Research Part C: Emerging Technologies*, *47*, 155–167.

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, *65*(1), 2–16.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32.

Castagno, J., & Atkins, E. (2018). Roof shape classification from LiDAR and satellite image data fusion using supervised learning. *Sensors*, *18*(11), 3960.

Cerri, M., Steinhausen, M., Kreibich, H., & Schröter, K. (2021). Are OpenStreetMap building data useful for flood vulnerability modelling? *Natural Hazards and Earth System Sciences*, *21*(2), 643–662.

Chen, R., Li, X., & Li, J. (2018). Object-based features for house detection from RGB high-resolution images. *Remote Sensing*, *10*(3), 451.

Chen, W., Zhou, Y., Wu, Q., Chen, G., Huang, X., & Yu, B. (2020). Urban building type mapping using geospatial data: A case study of beijing, china. *Remote Sensing*, *12*(17), 2805.

Du, S., Zhang, F., & Zhang, X. (2015). Semantic classification of urban buildings combining VHR image and GIS data: An improved random forest approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, *105*, 107–119.

Fan, H., Zipf, A., & Fu, Q. (2014). Estimation of building types on OpenStreetMap based on urban morphology analysis. *Connecting a Digital Europe through Location and Place*, 19–35.

Feizizadeh, B., Garajeh, M. K., Blaschke, T., & Lakes, T. (2021). An object based image analysis applied for volcanic and glacial landforms mapping in Sahand Mountain, Iran. *Catena*, *198*, 105073.

Gao, Y., Lang, S., Tiede, D., Gella, G. W., & Wendt, L. (2022). Comparing OBIA-Generated Labels and Manually Annotated Labels for Semantic Segmentation in Extracting Refugee-Dwelling Footprints. *Applied Sciences*, *12*(21), 11226.

Geiss, C., Pelizari, P. A., Marconcini, M., Sengara, W., Edwards, M., Lakes, T., & Taubenböck, H. (2015). Estimation of seismic building structural types using multi-sensor remote sensing and machine learning techniques. *ISPRS Journal of Photogrammetry and Remote Sensing*, *104*, 175–188.

Geiss, C., Thoma, M., Pittore, M., Wieland, M., Dech, S. W., & Taubenböck, H. (2017). Multitask active learning for characterization of built environments with multisensor earth observation data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(12), 5583–5597.

Géron, A. (2022). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media, Inc.

Goel, A., Juneja, M., & Jawahar, C. (2012). *Are buildings only instances? Exploration in architectural style categories*. 1–8.

Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal*, *69*, 211–221.

Grippa, T., Georganos, S., Zarougui, S., Bognounou, P., Diboulo, E., Forget, Y., Lennert, M., Vanhuysse, S., Mboga, N., & Wolff, E. (2018). Mapping urban land use at street block level using openstreetmap, remote sensing data, and spatial metrics. *ISPRS International Journal of Geo-Information*, *7*(7), 246.

Guo, Y., Liu, Y., Georgiou, T., & Lew, M. S. (2018). A review of semantic segmentation using deep neural networks. *International Journal of Multimedia Information Retrieval*, *7*, 87–93.

Hall, R. J. (2003). The roles of aerial photographs in forestry remote sensing image analysis. *Remote Sensing of Forest Environments: Concepts and Case Studies*, 47–75.

Harris, C., & Stephens, M. (1988). *A combined corner and edge detector*. *15*(50), 10–5244.

Hecht, R., Meinel, G., & Buchroithner, M. (2015). Automatic identification of building types based on topographic databases–a comparison of different data sources. *International Journal of Cartography*, *1*(1), 18–31.

Henn, A., Römer, C., Gröger, G., & Plümer, L. (2012). Automatic classification of building types in 3D city models: Using SVMs for semantic enrichment of low resolution building data. *GeoInformatica*, *16*, 281–306.

Hoeser, T., Bachofer, F., & Kuenzer, C. (2020). Object detection and image segmentation with deep learning on Earth observation data: A review—Part II: Applications. *Remote Sensing*, *12*(18), 3053.

Hoffmann, E. J., Wang, Y., Werner, M., Kang, J., & Zhu, X. X. (2019). Model fusion for building type classification from aerial and street view images. *Remote Sensing*, *11*(11), 1259.

Huang, Y., Zhuo, L., Tao, H., Shi, Q., & Liu, K. (2017). A novel building type classification scheme based on integrated LiDAR and high-resolution images. *Remote Sensing*, *9*(7), 679.

Labetski, A., Vitalis, S., Biljecki, F., Arroyo Ohori, K., & Stoter, J. (2023). 3D building metrics for urban morphology. *International Journal of Geographical Information Science*, *37*(1), 36–67.

Lall, S., Lebrand, M., Park, H., Sturm, D., & Venables, A. (2021). *Pancakes to Pyramids: City Form to Promote Sustainable Growth*. World Bank. https://documents1.worldbank.org/curated/en/554671622446381555/pdf/City-Form-to-Promote-Sustainable-Growth.pdf

Li, J., Huang, X., Tu, L., Zhang, T., & Wang, L. (2022). A review of building detection from very high resolution optical remote sensing images. *GIScience & Remote Sensing*, *59*(1), 1199–1225.

Li, W., He, C., Fang, J., Zheng, J., Fu, H., & Yu, L. (2019). Semantic segmentation-based building footprint extraction using very high-resolution satellite images and multi-source GIS data. *Remote Sensing*, *11*(4), 403.

Liu, C., Huang, X., Zhu, Z., Chen, H., Tang, X., & Gong, J. (2019). Automatic extraction of built-up area from ZY3 multi-view satellite imagery: Analysis of 45 global cities. *Remote Sensing of Environment*, *226*, 51–73.

Lu, Z., Im, J., Rhee, J., & Hodgson, M. (2014). Building type classification using spatial and landscape attributes derived from LiDAR remote sensing data. *Landscape and Urban Planning*, *130*, 134–148.

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *152*, 166–177.

Maktav, D., Erbek, F., & Jürgens, C. (2005). Remote sensing of urban areas. *International Journal of Remote Sensing*, *26*(4), 655–659.

Masson, V., Heldens, W., Bocher, E., Bonhomme, M., Bucher, B., Burmeister, C., de Munck, C., Esch, T., Hidalgo, J., & Kanani-Sühring, F. (2020). City-descriptive input data for urban climate models: Model requirements, data sources and challenges. *Urban Climate*, *31*, 100536.

Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, *39*(9), 2784–2817.

Morales-Molina, C. D., Hernandez-Suarez, A., Sanchez-Perez, G., Toscano-Medina, L. K., Perez-Meana, H., Olivares-Mercado, J., Portillo-Portillo, J., Sanchez, V., & Garcia-Villalba, L. J. (2021). A dense neural network approach for detecting clone id attacks on the rpl protocol of the iot. *Sensors*, *21*(9), 3173.

Nazari, F., & Yan, W. (2021). Convolutional versus dense neural networks: Comparing the two neural networks performance in predicting building operational energy use based on the building shape. *arXiv Preprint arXiv:2108.12929*. https://arxiv.org/abs/2108.12929

Neupane, B., Aryal, J., & Rajabifard, A. (2023). A novel dual skip connection mechanism in U-Nets for building footprint extraction. *arXiv Preprint arXiv:2303.09064*.

Oliveira, V. (2016). The Elements of Urban Form. In *Urban Morphology: An Introduction to the Study of the Physical Form of Cities* (pp. 7–30). Springer.

Pal, M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, *26*(1), 217–222.

Parente, J., Rodrigues, E., Rangel, B., & Martins, J. P. (2023). Integration of convolutional and adversarial networks into building design: A review. *Journal of Building Engineering*, 107155.

Pasquali, G., Iannelli, G. C., & Dell'Acqua, F. (2019). Building footprint extraction from multispectral, spaceborne earth observation datasets using a structurally optimized U-Net convolutional neural network. *Remote Sensing*, *11*(23), 2803.

Pelizari, P. A., Geiß, C., Aguirre, P., Santa María, H., Peña, Y. M., & Taubenböck, H. (2021). Automated building characterization for seismic risk assessment using street-level imagery and deep learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, *180*, 370–386.

Pesaresi, M., Gerhardinger, A., & Kayitakire, F. (2008). A robust built-up area presence index by anisotropic rotation-invariant textural measure. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *1*(3), 180–192.

Prathiba, A., Rastogi, K., Jain, G. V., & Govind Kumar, V. (2020). *Building footprint extraction from very-high-resolution satellite image using object-based image analysis (OBIA) technique*. 517–529.

Scorza, F., & Fortunato, G. (2021). Cyclable cities: Building feasible scenario through urban space morphology assessment. *Journal of Urban Planning and Development*, *147*(4), 05021039.

Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM Computing Surveys (CSUR)*, *34*(1), 1–47.

Shen, Y., Ai, T., & Li, C. (2019). A simplification of urban buildings to preserve geometric properties using superpixel segmentation. *International Journal of Applied Earth Observation and Geoinformation*, *79*, 162–174.

Shrestha, S., & Vanneschi, L. (2018). Improved fully convolutional network with conditional random fields for building extraction. *Remote Sensing*, *10*(7), 1135.

Srivastava, V., Avudaiammal, R., & V George, S. (2024). Investigations on extraction of buildings from RS imagery using deep learning models. *International Journal of Remote Sensing*, *45*(1), 68–100.

Taoufiq, S., Nagy, B., & Benedek, C. (2020). Hierarchynet: Hierarchical CNN-based urban building classification. *Remote Sensing*, *12*(22), 3794.

Thottolil, R., & Kumar, U. (2022). *Automatic Building Footprint Extraction using Random Forest Algorithm from High Resolution Google Earth Images: A Feature-Based Approach*. 1–6.

Tooke, T. R., Coops, N. C., & Webster, J. (2014). Predicting building ages from LiDAR data with random forests for building energy modeling. *Energy and Buildings*, *68*, 603–610.

Turhan, C., Kazanasmaz, T., Uygun, I. E., Ekmen, K. E., & Akkurt, G. G. (2014). Comparative study of a building energy performance software (KEP-IYTE-ESS) and ANN-based building heat load estimation. *Energy and Buildings*, *85*, 115–125.

UN. (2019). *World Urbanization Prospects 2018: Highlights*. United Nations (UN). https://population.un.org/wup/Publications/Files/WUP2018-Highlights.pdf

UN. (2023). The Sustainable Development Goals Report 2023: Special Edition. *United Nations*. https://unstats.un.org/sdgs/report/2023/

Virtriana, R., Harto, A. B., Atmaja, F. W., Meilano, I., Fauzan, K. N., Anggraini, T. S., Ihsan, K. T. N., Mustika, F. C., & Suminar, W. (2023). Machine learning remote sensing using the random forest classifier to detect the building damage caused by the Anak Krakatau Volcano tsunami. *Geomatics, Natural Hazards and Risk*, *14*(1), 28–51.

Wang, J., & Shan, J. (2009). *Segmentation of LiDAR point clouds for building extraction*. 9–13.

Wang, Z., Xu, N., Wang, B., Liu, Y., & Zhang, S. (2022). Urban building extraction from high-resolution remote sensing imagery based on multi-scale recurrent conditional generative adversarial network. *GIScience & Remote Sensing*, *59*(1), 861–884.

Wouters, L., Couasnon, A., De Ruiter, M. C., Van Den Homberg, M. J., Teklesadik, A., & De Moel, H. (2021). Improving flood damage assessments in data-scarce areas by retrieval of building characteristics through UAV image segmentation and machine learning–a case study of the 2019 floods in southern Malawi. *Natural Hazards and Earth System Sciences*, *21*(10), 3199–3218.

Wurm, M., Droin, A., Stark, T., Geiß, C., Sulzer, W., & Taubenböck, H. (2021). Deep learning-based generation of building stock data from remote sensing for urban heat demand modeling. *ISPRS International Journal of Geo-Information*, *10*(1), 23.

Wurm, M., Schmitt, A., & Taubenböck, H. (2015). Building types' classification using shape-based features and linear discriminant functions. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(5), 1901–1912.

Wurm, M., Taubenböck, H., & Dech, S. (2010). *Quantification of urban structure on building block level utilizing multisensoral remote sensing data*. *7831*, 123–134.

Yan, X., Ai, T., Yang, M., & Yin, H. (2019). A graph convolutional neural network for classification of building patterns using spatial vector data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *150*, 259–273.

Yang, M., Kong, B., Dang, R., & Yan, X. (2022). Classifying urban functional regions by integrating buildings and points-of-interest using a stacking ensemble method. *International Journal of Applied Earth Observation and Geoinformation*, *108*, 102753.

Yuan, X., Shi, J., & Gu, L. (2021). A review of deep learning methods for semantic segmentation of remote sensing imagery. *Expert Systems with Applications*, *169*, 114417.

Zhou, P., & Chang, Y. (2021). Automated classification of building structures for urban built environment identification using machine learning. *Journal of Building Engineering*, *43*, 103008.