Division of Computing Science and Mathematics
Faculty of Natural Sciences
University of Stirling

Developing Deep Learning Techniques for Hedgerow Detection in Very High-Resolution Satellite Imagery

Eashwar Nagarajan Nithyameenakshi

Dissertation submitted in partial fulfilment for the degree of
Master of Science in Artificial Intelligence

September 2024

Abstract

Artificial Intelligence (AI) empowers machines to emulate human-like cognitive functions, such as learning and problem-solving. Computer vision, a specialized branch of AI, focuses on enabling machines to interpret and understand the visual world from digital images or videos [1]. The integration of AI and computer vision with remote sensing data has the potential to revolutionize various fields, including environmental monitoring and land management.

The accurate and efficient mapping of hedgerows, which are vital linear features in agricultural and urban landscapes, remains a challenge due to the limitations of traditional ground surveys. These natural corridors play a crucial role in supporting biodiversity, preventing soil erosion, and facilitating habitat connectivity [2]. The advent of very high-resolution (VHR) satellite imagery, coupled with advancements in deep learning, presents a promising opportunity to automate and scale hedgerow mapping efforts.

This dissertation investigates the potential of deep learning techniques, specifically Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs), for automated hedgerow detection in Very High Resolution (VHR) imagery. The research includes the collection and preprocessing of high-resolution multispectral satellite data and aerial photography, the development and training of deep learning models, and their rigorous evaluation. The models are trained on a dataset derived from the Bluesky National Hedgerow Map [3], with their hedgerow masks serving as a proxy for ground truth data. The study also utilizes high-resolution Pleiades Neo imagery for further analysis.

The study successfully demonstrates the efficacy of deep learning in automating hedgerow detection, offering valuable insights into the comparative performance of different model architectures and the impact of data processing techniques. The developed models demonstrate high accuracy and robustness, offering a valuable tool for large-scale hedgerow mapping. This advancement has the potential to significantly benefit land management, environmental conservation, and ecological research by enabling more informed decision-making and efficient monitoring of these critical landscape features.

Keywords: Hedgerow detection, deep learning, VHR satellite imagery, CNNs, Vision Transformers, remote sensing, land management, environmental conservation.

Attestation

I understand the nature of plagiarism, and I am aware of the University's policy on this.

I certify that this dissertation reports original work by me during my University project except for the following:

- The architecture image of the Unet model showcased was obtained from this GitHub repository [48].
- The architecture image of the FPN model showcased was obtained from this GitHub repository [49].
- The architecture image of the PSP model showcased was obtained from this GitHub repository [50].
- The architecture image of the LinkNet model showcased was obtained from this GitHub repository [51].
- The architecture image of the MiT model showcased was obtained from this Keras link [52].

Signature **Date**

Acknowledgements

I extend my deepest gratitude to my supervisor, Dr. Vahid Akbari, for his invaluable guidance, unwavering support, and insightful feedback throughout this research journey. His expertise and encouragement have been instrumental in shaping the direction and quality of this dissertation.

I am also sincerely grateful to Mehran Alizadeh Pirbasti, a PhD student from the University College Dublin (UCD), for his collaborative spirit and valuable contributions to this project. His insights and discussions have significantly enriched my understanding of the research topic.

I would like to express my appreciation to the UK Centre for Ecology & Hydrology (UKCEH) for generously providing access to their hedgerow map. I am also thankful to Airbus for granting access to the high-resolution satellite imagery. Furthermore, I extend my gratitude to Blue Sky for providing their aerial hedgerow dataset and accompanying masks, which were instrumental in training and validating the deep learning models.

I am also indebted to my fellow students in the Artificial Intelligence program for their camaraderie and stimulating discussions. Their collective knowledge and support fostered a vibrant learning environment that greatly enhanced my academic experience.

Finally, I owe a special thanks to my family and friends for their unwavering love, support, and understanding throughout the duration of this project. Their encouragement and belief in my abilities have been a constant source of motivation and inspiration.

Table of Contents

List of Figures

1 Introduction

Hedgerows, the linear features of trees and shrubs that weave through agricultural and urban landscapes, are vital components of the ecosystem. They provide a multitude of ecological services, including supporting biodiversity by acting as wildlife corridors and refuges, preventing soil erosion through their root systems, and facilitating habitat connectivity by linking fragmented landscapes [2]. However, the accurate and efficient mapping of these vital landscape elements remains a challenge due to the limitations of traditional, labour-intensive ground surveys [4]. The advent of VHR satellite imagery, coupled with advancements in deep learning, presents a promising opportunity to automate and scale hedgerow mapping efforts [5].

This dissertation investigates the potential of deep learning techniques, specifically CNNs and ViTs, for automated hedgerow detection in VHR satellite imagery. The research encompasses the collection and preprocessing of multispectral satellite data and aerial photography, the development and training of deep learning models [6], and their rigorous evaluation against a reliable ground truth dataset provided by BlueSky and the UK Centre for Ecology & Hydrology (UKCEH) [47].

The study successfully demonstrates the efficacy of deep learning in automating hedgerow detection, offering valuable insights into the comparative performance of different model architectures and the impact of data processing techniques. The developed models exhibit promising accuracy and robustness, providing a powerful tool for large-scale hedgerow mapping. This advancement has the potential to significantly benefit land management, environmental conservation, and ecological research by enabling more informed decision-making and efficient monitoring of these critical landscape features [7].

1.1 Background and Context

In the UK, hedgerows hold particular significance, not only for their ecological role but also as integral elements of the cultural heritage and scenic beauty of the countryside [8]. However, the accurate and efficient mapping of these valuable features remains a challenge. Traditional ground surveys, while effective, are labour-intensive, time-consuming, and often impractical for large-scale assessments [4]. The emergence of remote sensing technologies, particularly VHR satellite imagery, offers a promising solution for overcoming these limitations. VHR imagery, with its ability to capture fine-grained details of the Earth's surface, presents an opportunity to develop automated and scalable methods for hedgerow detection [9]. Recent advancements in deep learning, a subfield of artificial intelligence, have further fuelled this potential. Deep learning models, particularly CNNs, have demonstrated remarkable success in various image analysis tasks, including object detection and semantic segmentation [1, 5]. The application of these powerful techniques to VHR satellite imagery and aerial photography opens new avenues for accurate and efficient hedgerow mapping, with potential benefits for land management, conservation efforts, and ecological research.

1.2 Scope and Objectives

This dissertation investigates the application of deep learning techniques, specifically CNNs and ViTs, for the automated detection and delineation of hedgerows in aerial photography and VHR satellite imagery. The primary focus lies in developing and evaluating deep learning models capable of accurately identifying hedgerows within the complex landscape features captured in images. The research will encompass the collection and preprocessing of multispectral Pleiades Neo and aerial photography images, ensuring data quality and consistency. The study will look to leverage the recently released UK Centre for Ecology & Hydrology (UKCEH)'s hedgerow map as a benchmark for ground truth data, and BlueSky's labelled hedgerow masks to train and validate the model's performance.

The specific objectives of this research are:

1 Data Collection and Preprocessing: Gather and prepare very high-resolution images, ensuring the data is clean, labelled accurately, and suitable for deep learning applications.

2 Model Development and Training: Design, implement, and train deep learning models, including CNNs and ViTs, for hedgerow segmentation in aerial photography and VHR satellite images. The models will be trained and fine-tuned using the prepared dataset.

3 Performance Evaluation: Evaluate the performance of the trained models using appropriate quantitative metrics, such as precision, recall, and F1-score, and qualitative assessments through visual inspection and comparison with ground truth data.

4 Comparative Analysis: Analyse the impact of different model architectures with fixed hyperparameter choices on hedgerow detection accuracy. The analysis will provide insights into the strengths and weaknesses of each approach.

5 Limitations and Future Directions: Identify the limitations of the proposed approach and suggest potential avenues for future research, contributing to the ongoing development of robust and efficient hedgerow detection methods.

1.3 Achievements

This dissertation successfully demonstrates the feasibility and effectiveness of deep learning, particularly CNNs and ViTs, for automated hedgerow detection in aerial photography and VHR satellite imagery. The research resulted in the development and training of multiple deep learning models, achieving promising results in terms of accuracy, precision, recall, and F1-score. The models effectively identified and delineated hedgerows in complex landscapes, showcasing the potential of this approach for large-scale mapping initiatives.

Through rigorous experimentation and analysis, the study provides valuable insights into the following:

Model Performance: The comparative performance of different CNNs and ViTs architectures, highlighting their strengths and weaknesses for hedgerow detection.

Benchmarking: The utilization of BlueSky's hedgerow masks as a reliable benchmark for ground truth data, ensuring the validity and applicability of the research findings.

The successful application of deep learning for automated hedgerow detection in this dissertation represents a significant advancement in the field of remote sensing and environmental monitoring. The developed models and insights gained from this research have the potential to streamline hedgerow mapping efforts, enabling more efficient and informed decision-making in land management, conservation planning, and ecological research.

1.4 Overview of Dissertation

The structure of this dissertation is organized as follows:

Chapter 2: Literature Review

This chapter presents a comprehensive review of the relevant literature, encompassing the ecological importance of hedgerows, the challenges associated with their traditional mapping methods, and the advancements in deep learning for remote sensing applications. The review explores the current state-of-the-art in CNN-based hedgerow detection, identifying research gaps and highlighting the potential of VHR satellite imagery, aerial photography imagery, and ground truth data for addressing these challenges.

Chapter 3: Data Collection and Preparation

This chapter details the methodology employed for data collection and preparation. It covers the acquisition of high-resolution multispectral Pléiades Neo and aerial photography images, as well as the preprocessing steps undertaken to ensure data quality and consistency. Data augmentation techniques utilized to enhance model generalization are also discussed.

Chapter 4: Model Selection

This chapter focuses on the selection and description of various deep-learning models for hedgerow detection. It explores the foundations of CNNs and ViTs and provides a rationale for the chosen models and their suitability for the task at hand.

Chapter 5: Results and Evaluations

This chapter presents the experimental results obtained from the trained deep learning models. It includes a comparative analysis of their performance using quantitative metrics such as precision, recall, and F1-score. The chapter discusses the strengths and weaknesses of each model and analyses the models for hedgerow detection tasks.

Chapter 6: Final Model Chosen

This chapter discusses the selection of the final models based on the evaluation results. It provides a detailed justification for the chosen models, highlighting their superior performance and suitability for the specific task of hedgerow detection in VHR satellite imagery.

Chapter 7: Conclusion and Recommendations

This chapter concludes the dissertation by summarizing the key findings and highlighting the contributions of this research.

Chapter 8: Future Scope

This chapter explores the potential future applications and extensions of the developed hedgerow detection models. It discusses the possibilities for further refinement and improvement of the models, as well as their integration into broader environmental monitoring and decision-making frameworks.

2 Literature Review

The literature review in this dissertation explores hedgerow mapping, emphasizing their ecological importance, challenges of traditional mapping, and the potential of VHR imagery, aerial photography, and deep learning for automated detection. The review showcases successful CNN applications in remote sensing, identifies research gaps, and discusses potential for improvement. It concludes by highlighting the need for further research to address challenges and fully utilize deep learning for automated hedgerow mapping, benefiting land management and ecological research.

2.1 Ecological Importance of Hedgerows

Hedgerows, as linear features composed of trees and shrubs that traverse agricultural and urban landscapes, are widely recognized for their significant ecological value. They function as critical hubs of biodiversity, offering food, shelter, and breeding grounds to a diverse array of flora and fauna [10]. The varied plant species within hedgerows provide nectar, pollen, and fruits, attracting pollinators such as bees and butterflies, as well as beneficial insects and birds [11]. The dense, interwoven structure of hedgerows creates nesting sites, protective cover, and overwintering habitats for small mammals, birds, and reptiles [12].

Beyond their role in supporting local biodiversity, hedgerows facilitate ecological connectivity by forming vital wildlife corridors. These corridors enable the movement and dispersal of species across fragmented landscapes, promoting gene flow and enhancing population resilience [13]. Hedgerows also contribute to essential ecosystem services that are vital for the health and sustainability of the environment. Their intricate root systems act as natural barriers against soil erosion caused by wind and water, safeguarding agricultural productivity and mitigating the loss of valuable topsoil [14]. Furthermore, hedgerows play a crucial role in regulating water flow by reducing surface runoff and promoting infiltration, thus mitigating the risk of flooding and facilitating groundwater recharge [15].

The microclimates created by hedgerows, characterized by shade, wind protection, and increased humidity, can benefit adjacent crops and ecosystems [16]. These microclimatic effects can enhance crop yields, reduce water stress, and provide favourable conditions for plant growth and development [17]. Additionally, hedgerows contribute to carbon sequestration by capturing and storing carbon dioxide from the atmosphere, thus playing a role in mitigating climate change [18].

2.2 Challenges in Hedgerow Mapping and the Potential of Deep Learning

Accurate hedgerow mapping is challenging despite their ecological importance. Traditional ground surveys, though valuable, are labour-intensive and impractical for large-scale monitoring [4]. Hedgerows' dynamic nature adds complexity to conventional mapping

Remote sensing technologies, particularly VHR satellite imagery, offer a promising avenue for overcoming these limitations. The high spatial resolution of VHR imagery enables the capture of fine-grained details of the Earth's surface, making it possible to distinguish hedgerows from other linear features and surrounding vegetation [19]. However, effective utilization requires advanced image analysis to handle data complexity.

Deep learning, especially CNNs, excels in image analysis and object recognition [1, 5]. Applying it to VHR imagery can transform hedgerow mapping, enabling automated, efficient assessments. This overcomes the limitations of traditional surveys, leading to better land management and conservation.

2.3 Current State-of-the-Art in CNN-based Hedgerow Detection

CNNs have gained traction in hedgerow detection from VHR satellite imagery, demonstrating the potential for automated mapping. The study by Ahlswede et al. (2021) exemplifies this trend, successfully employing CNNs to detect hedgerows in VHR images and highlighting their capacity to learn intricate patterns and features from satellite data, even in challenging environments [19], showcasing the feasibility of large-scale automated mapping.

Beyond hedgerow detection, CNNs have proven effective in various remote sensing tasks like object detection and land cover classification. The comprehensive survey by Li et al. (2022) underscores the broad applicability of CNNs in this domain, encompassing object detection, land cover classification, and change detection [5].

Recent research has also explored the integration of CNNs with other advanced techniques to enhance hedgerow detection accuracy. For instance, the study by Zhang et al. (2023) proposed a multi-scale CNN framework that leverages both spatial and spectral information from VHR imagery to improve hedgerow segmentation [20]. The incorporation of multi-scale features allows the model to capture hedgerows at various scales and resolutions, leading to more precise and robust detection results.

Furthermore, the work by Liu et al. (2022) investigated the use of attention mechanisms within CNN architectures for hedgerow detection [21]. Attention mechanisms enable the model to focus on relevant regions of the image, enhancing its ability to discriminate hedgerows from background clutter and other linear features. The integration of attention mechanisms has shown promising results in improving the accuracy and efficiency of hedgerow detection.

These advancements highlight deep learning's potential for automated hedgerow mapping. CNNs' ability to learn, adapt, and integrate with other techniques positions them as a powerful tool for overcoming challenges in traditional mapping.

2.4 Research Gaps and Challenges

While promising, several research gaps and challenges persist, hindering the full realization of CNN’s potential for accurate and efficient large-scale mapping.

Hedgerows exhibit significant variations in appearance across different landscapes, seasons, and management practices. The spectral and structural characteristics of hedgerows can be influenced by factors such as species composition, age, health, and surrounding land use. This inherent variability poses a challenge for CNNs, which rely on consistent patterns and features for accurate detection. Developing models that can generalize well across diverse environments and capture the subtle nuances of hedgerow appearance remains an ongoing research challenge [22].

The selection of appropriate CNN architectures and hyperparameters significantly impacts the performance of hedgerow detection models. The vast array of available CNN architectures and the multitude of hyperparameter choices present a challenge in identifying the optimal configuration for a given dataset and task. The exploration and fine-tuning of model architectures and hyperparameters are crucial for maximizing the accuracy and efficiency of hedgerow detection [23].

Access to large-scale, high-quality annotated VHR satellite imagery for hedgerow detection remains a significant challenge. The manual annotation of hedgerows in VHR images is a labour-intensive and time-consuming process, limiting the availability of training data for deep-learning models. Addressing this limitation through innovative data collection and annotation strategies, as well as the development of effective data augmentation and transfer learning techniques, is essential for advancing the field of hedgerow detection [24].

Addressing these research gaps and challenges will be crucial for unlocking the full potential of deep learning for automated hedgerow mapping. The development of robust, adaptable, and efficient models that can handle the variability of hedgerow appearance and operate in real-time will pave the way for more effective and sustainable management of these vital landscape features.

2.5 Addressing the Gaps

This research aims to address the challenges and advance the state-of-the-art in hedgerow detection by:

Investigating Advanced CNN Architectures and ViTs: The study will explore a range of cutting-edge deep learning models, including U-Net [30], FPN [31], PSPNet [32], LinkNets [33], and Vision Transformers (ViTs) [34], to identify the most suitable architectures for hedgerow segmentation in VHR satellite imagery. The investigation will consider the trade-offs between model complexity, computational efficiency, and accuracy to select the optimal architecture for the task.

Leveraging Transfer Learning: To mitigate the limitations posed by the scarcity of annotated VHR data, the research will employ transfer learning techniques. Pre-trained models on large-scale image datasets, such as ImageNet, will be fine-tuned for the specific task of hedgerow detection. This approach leverages the knowledge learned from general image recognition tasks to improve the performance and generalization of the models on the hedgerow detection task, even with limited training data.

Employing Data Augmentation: The study will utilize data augmentation techniques to artificially expand the training dataset and enhance the robustness of the models. Techniques such as rotation, flipping, cropping, and colour jittering will be applied to introduce variations in the training data, enabling the models to learn invariant features and generalize better to unseen data.

Rigorous Evaluation and Validation: The performance of the developed models will be rigorously evaluated and validated using a comprehensive set of metrics, including precision, recall, F1-score, and Intersection over Union (IoU). The models will be tested on a diverse range of VHR satellite images, capturing different landscapes, seasons, and hedgerow characteristics.

This dissertation addresses research gaps to develop robust deep-learning tools for large-scale hedgerow mapping. The research outcomes can significantly benefit land management, conservation, and ecological research by providing reliable monitoring methods. Insights from this study can inform future mapping initiatives, promoting effective conservation and sustainable land use.

2.6 Conclusion

This chapter reviewed automated hedgerow detection using deep learning. It highlighted hedgerows' ecological importance [2, 10, 11, 12, 13], challenges of traditional mapping [4], and deep learning's potential, particularly CNNs, for addressing these [1, 5, 9]. The current state-of-the-art was reviewed, showcasing CNNs' versatility [19, 20, 21].

Research gaps persist, including handling hedgerow variability, model optimization, and limited annotated data [22, 23, 24]. Future exploration includes multi-modal data and real-time detection [25].

This dissertation proposes using advanced CNNs, transfer learning, data augmentation, and rigorous evaluation to address challenges. Successful implementation can advance automated hedgerow mapping, benefiting land management and ecological research.

3 Data Collection and Preparation

Two datasets were utilized in this study. The first dataset, obtained from Airbus Defence, consisted of multispectral Pleiades Neo imagery with a resolution of 1.2 meters, capturing information across four spectral bands: red, green, blue, and near-infrared. The second dataset, provided by BlueSky [3], comprised aerial photography imagery of hedgerows with an accuracy of +/- 1m RMSE, along with corresponding hedgerow ground truth masks. The inclusion of ground truth masks is crucial for supervised learning, enabling the training and evaluation of deep learning models [26]. But this dataset is much smaller. Additionally, Light Detection and Ranging (LiDAR) data scanning of hedgerows in England, sourced from the UK Centre for Ecology & Hydrology (UKCEH), was incorporated to provide further validation and insights into hedgerow characteristics [1].

The larger images from both the Airbus (as shown in Fig. 3.3) and BlueSky datasets (as shown in Fig. 3.1), which are in raster format - a digital image represented by a grid of pixels where each pixel contains a value representing information such as colour or elevation -, were initially merged into a single raster and subsequently gridded and tiled into smaller, manageable image tiles of 256x256 pixels (as shown in Fig 3.2 and 3.4). The corresponding hedgerow masks underwent the same processing to ensure alignment with the image data (shown in Fig. 3.2). This tiling process was performed using QGIS, a powerful open-source geographic information system (GIS) software widely employed for viewing and manipulating geospatial data [27]. The resulting image chips and masks served as the foundation for training and evaluating the deep learning models for hedgerow detection.

Figure 3.1 True colour aerial photography image visualised along with hedgerow maps

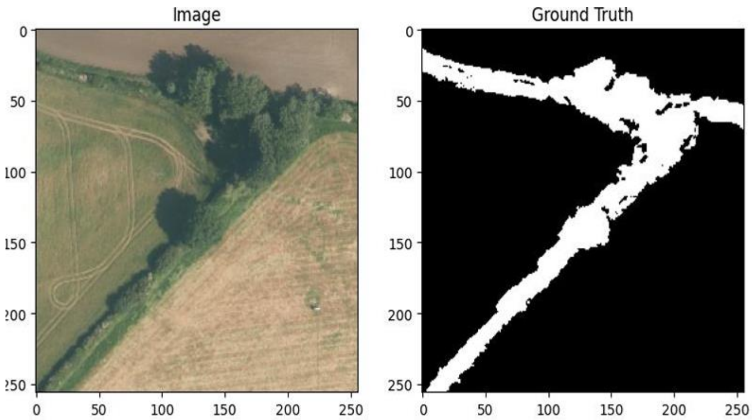

Figure 3.2 A small patch of BlueSky data after tilling of image and mask visualized

Figure 3.3 Quicklook of Pléiades Neo multispectral data from Airbus

Figure 3.4 A small patch of Airbus data after tilling visualized

3.1 Challenges faced

During the data preparation phase, it was discovered that the Airbus multispectral data and the UKCEH LiDAR data were misaligned due to a mismatch in their Coordinate Reference Systems (CRS). Furthermore, upon closer inspection, the UKCEH data, even after their preprocessing, exhibited inaccuracies, with an estimated accuracy of only 76% in a 1km test square. The UKCEH acknowledged these issues and indicated that they are currently undergoing a revision process to improve the accuracy of their data using a different post-processing approach.

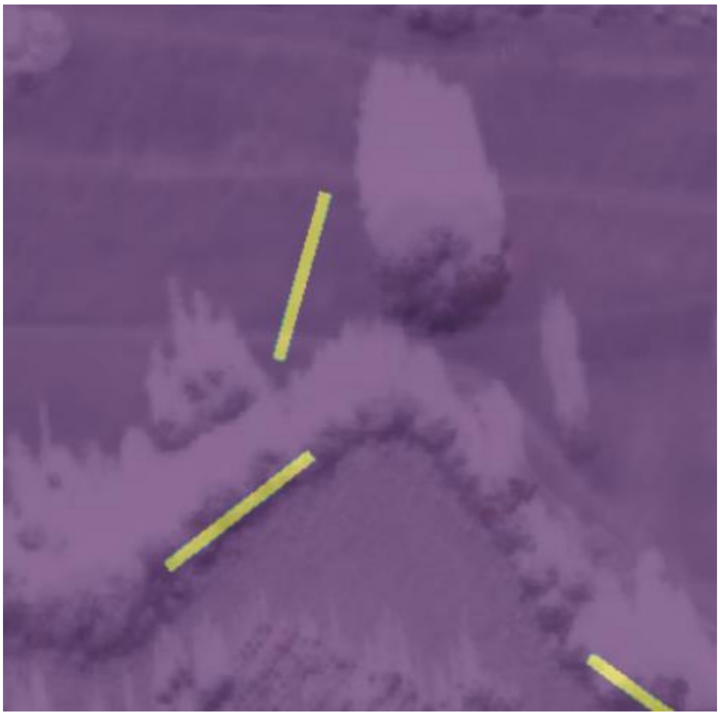

Figure 3.5 Image with mask overlay showing incorrect hedgerow map by UKCEH on Airbus data

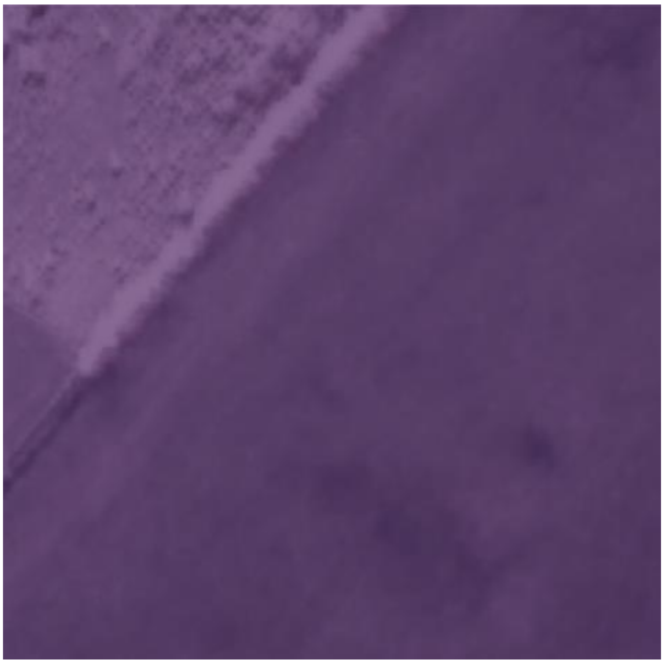

Figure 3.6 Image with mask overlay showing hedgerows missed by UKCEH on Airbus data.

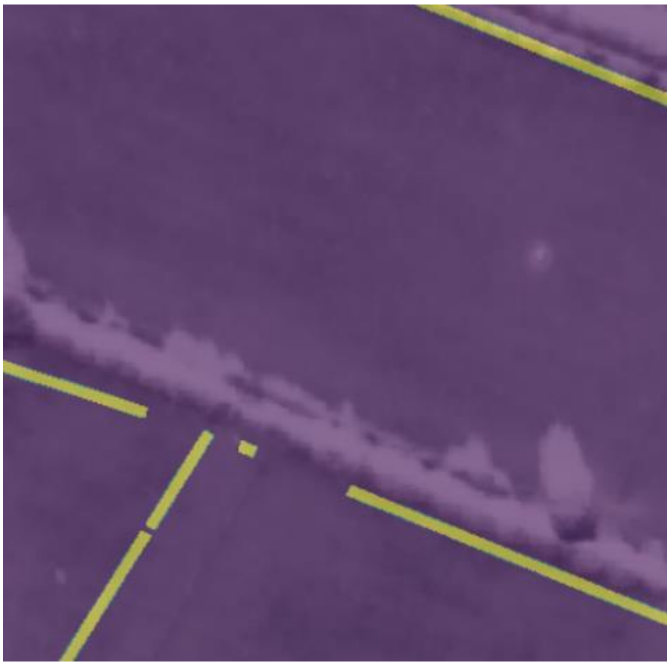

Figure 3.7 Image with mask overlay showing incorrect alignment of hedgerow by UKCEH on Airbus data.

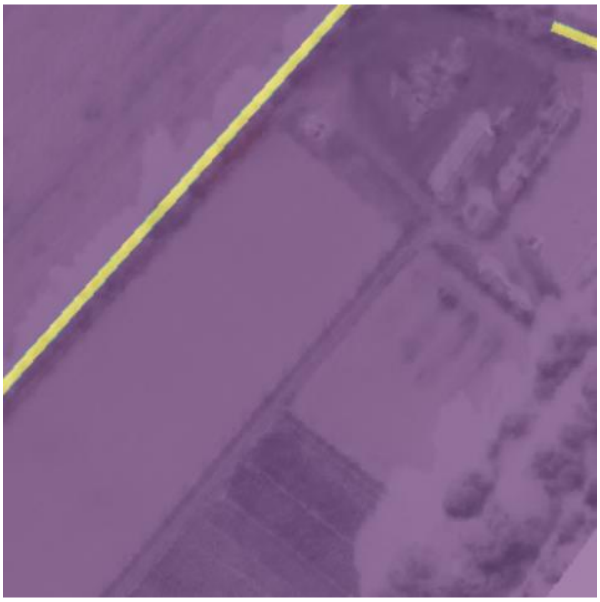

Figure 3.8 Image with mask overlay showing perfect alignment of hedgerow by UKCEH on Airbus data.

Consequently, the Airbus multispectral data could not be directly validated against the UKCEH data as initially planned. To overcome this challenge, the research strategy was adapted to leverage transfer learning. A model was first trained on the BlueSky dataset, which included accurate ground truth masks. Subsequently, this pre-trained model was then used to perform predictions on the Airbus dataset.

This adaptation highlights the importance of flexibility and adaptability in research, particularly when dealing with real-world data challenges. The use of transfer learning in this context demonstrates a resourceful approach to overcoming limitations and maximizing the utility of available data sources.

3.2 Data Quality Check

The images and corresponding masks generated through the gridding and tiling process underwent a meticulous manual inspection to ensure data quality and relevance to the research objectives. The primary focus of this inspection was to verify the presence of hedgerows, the target features of interest (shown in Fig. 3.9). Images that contained mislabelled irrelevant features were discarded (shown in Fig. 3.10) to maintain a high-quality dataset for training the deep learning models. This manual curation process is crucial in minimizing noise and ensuring that the models are trained on representative and informative examples of hedgerows, ultimately contributing to improved detection accuracy and generalization performance [24]. After quality checks, the BlueSky data had 49 high-quality images.

Figure 3.9 Image patch with high quality information kept

Figure 3.10 Image patch with trees mislabelled as hedgerows discarded

3.3 Data Augmentation

Data augmentation is a technique widely used in machine learning to artificially increase the size and diversity of training datasets by applying various transformations to the original data [28]. The primary goal of data augmentation is to improve the generalization capabilities of machine learning models, particularly deep learning models, by exposing them to a wider range of variations in the input data (shown in Fig 3.12). This helps the models learn more robust and invariant features, reducing the risk of overfitting and enhancing their performance on unseen data.

In the context of hedgerow detection using VHR satellite imagery, data augmentation is particularly crucial due to the limited availability of annotated training data and the inherent variability in hedgerow appearance across different landscapes and seasons. By applying transformations such as rotations, flips, crops, and colour adjustments to the original images and their corresponding masks, data augmentation can generate new, synthetic training examples that capture different perspectives, orientations, and lighting conditions of hedgerows. This expanded and diversified dataset can significantly improve the model's ability to recognise and delineate hedgerows in various real-world scenarios, leading to more accurate and reliable detection results.

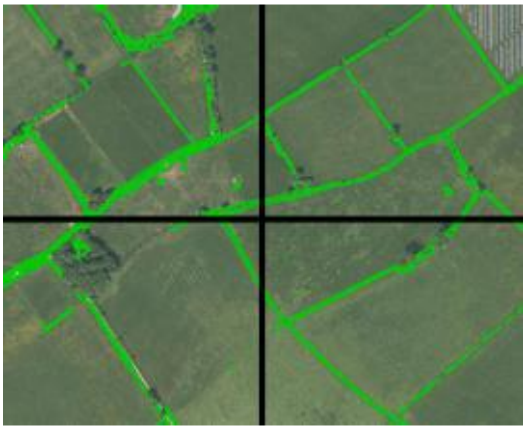

Figure 3.11 Example of an original data tilled into 4 smaller tiles

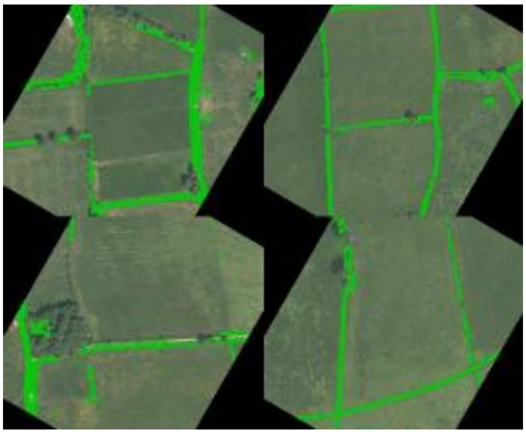

Figure 3.12 Example of augmented data after rotation and random crop, black pixels are used to fill in gaps to make sure the image is in the correct dimensions.

The effectiveness of data augmentation has been demonstrated in numerous studies across various domains, including image classification, object detection, and semantic segmentation [28]. In the context of remote sensing, data augmentation has been shown to improve the performance of deep learning models for tasks such as land cover classification, crop type mapping, and object detection [29]. The application of data augmentation in this dissertation is expected to similarly enhance the robustness and generalization of the deep learning models for hedgerow detection, contributing to the development of more accurate and efficient mapping tools.

Augmentation Techniques Used and Their Benefits:

Technique	Description
RandomResizedCrop	Randomly crops and resizes the image, introducing variability in object scale and position.
HorizontalFlip & VerticalFlip	Randomly flips the image horizontally or vertically, increasing data diversity and reducing bias towards specific orientations.
RandomRotate90	Randomly rotates the image by 90 degrees, further enhancing rotational invariance.
ShiftScaleRotate	Applies a combination of shifting, scaling, and rotation, simulating spatial variations in object appearance.
RandomBrightnessContrast	Randomly adjusts brightness and contrast, improving robustness to varying lighting conditions.
GaussNoise	Adds random Gaussian noise, making the model more resilient to noise in real-world imagery.
HueSaturationValue	Randomly adjusts hue, saturation, and value (brightness), promoting invariance to colour variations.
ChannelShuffle	Randomly shuffles colour channels, encouraging the model to learn features independent of channel order.
ElasticTransform	Applies elastic deformations, simulating local distortions that can occur in satellite imagery.
RandomGamma	Randomly adjusts gamma correction, further enhancing robustness to lighting variations.
CoarseDropout	Randomly drops out rectangular regions, forcing the model to focus on relevant features and avoid overfitting.
GridDistortion	Applies random distortions to the image grid, simulating perspective distortions.

4 Deep Learning Model Selection

The task of hedgerow detection in VHR satellite imagery presents unique challenges that necessitate the selection of deep learning models capable of handling complex image analysis and precise segmentation. In this study, we focus on two prominent deep-learning architectures: CNNs and ViTs.

4.1 Convolutional Neural Networks (CNNs)

CNNs have emerged as the dominant architecture for image-related tasks due to their ability to automatically learn hierarchical representations of spatial features [1, 9]. The convolutional layers in CNNs excel at capturing local patterns and textures, making them well-suited for identifying the intricate details and variations in hedgerow structures within satellite imagery. The inherent translational invariance of CNNs allows them to detect hedgerows regardless of their position within the image, contributing to their robustness and generalization capabilities. Furthermore, the extensive research and development in CNNs have resulted in a plethora of pre-trained models and established architectures, providing a strong foundation for transfer learning and adaptation to the specific task of hedgerow detection.

The selection of specific CNN architectures for hedgerow detection is guided by their established strengths in addressing the unique challenges posed by this task. The following architectures are chosen for their potential to deliver accurate and efficient hedgerow segmentation in VHR satellite imagery.

4.1.1 U-Net

The U-Net [30] architecture (shown in Fig. 4.1), renowned for its success in biomedical image segmentation, is selected for its ability to capture both fine-grained details and global context through its encoder-decoder structure and skip connections [30]. The encoder progressively downsamples the input image, extracting hierarchical features at different scales, while the decoder upsamples the feature maps to generate a segmentation mask. The skip connections facilitate the flow of information between corresponding layers by concatenating outputs of the encoder and decoder, enabling the recovery of fine details during upsampling. This makes U-Net well-suited for delineating the intricate boundaries of hedgerows in high-resolution imagery, where both local and global context are crucial for accurate segmentation.

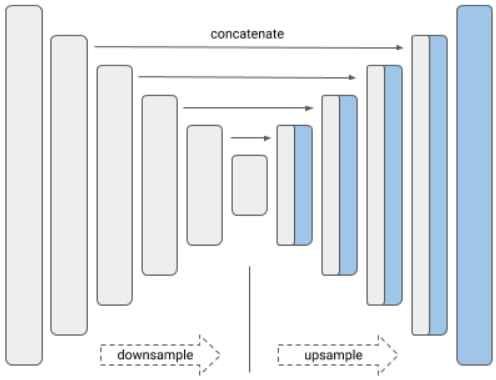

Figure 4.1 UNet Architecture

4.1.2 FPNets (Feature Pyramid Networks)

FPNets [31] (shown in Fig 4.2) are chosen for their ability to address the challenge of object scale variation, a common issue in hedgerow detection due to their varying sizes and shapes. By constructing a feature pyramid with multiple levels of resolution, FPNets can effectively detect hedgerows at different scales, from small gaps to large, continuous stretches. The lateral connections in FPNets enable the fusion of features from different levels of the pyramid, enhancing the model's ability to capture both fine details and global context.

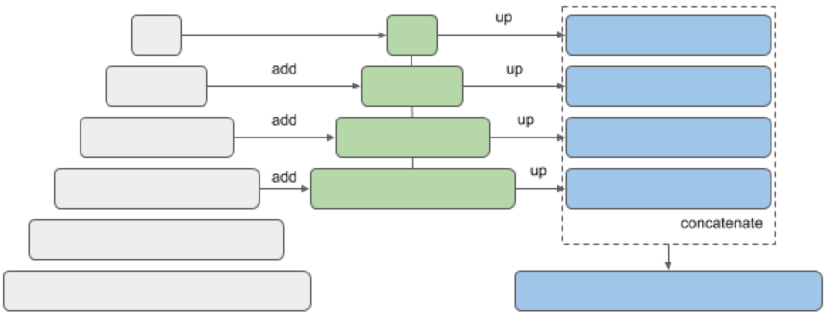

Figure 4.2 Feature Pyramid Networks architecture

4.1.3 PSPNets (Pyramid Scene Parsing Networks)

PSPNets [32] (shown in Fig. 4.3) are selected for their ability to incorporate global context into the segmentation process, which is crucial for distinguishing hedgerows from other linear features and background clutter in complex landscapes. The pyramid pooling module in PSPNets captures features at different scales and aggregates them to provide a richer representation of the scene. This global context helps the model understand the overall structure of the landscape and make more informed decisions about hedgerow segmentation.

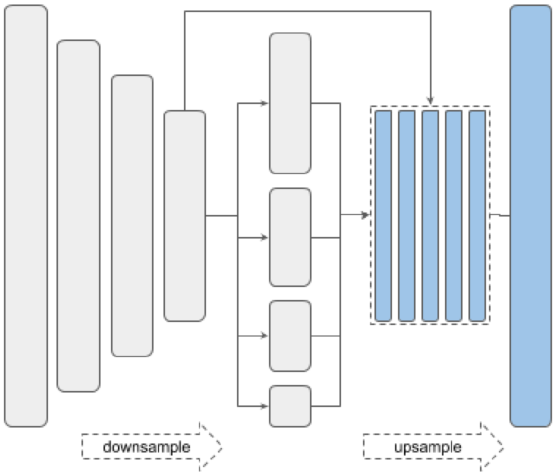

Figure 4.3 PSPNets architecture

4.1.4 LinkNets

LinkNets [33] (shown in Fig. 4.4) are chosen for their computational efficiency and real-time capabilities, which are essential for large-scale hedgerow mapping and potential real-time monitoring applications. The lightweight design of LinkNets achieved through the use of skip connections and reduced channel depth, enables faster inference compared to other architectures. This efficiency is crucial for processing large volumes of VHR satellite imagery and providing timely insights for land management and conservation efforts.

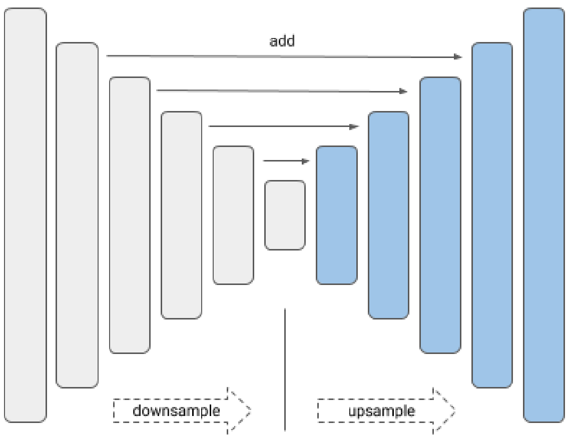

Figure 4.4 LinkNets architecture

4.2 Vision Transformers (ViTs)

ViTs [28], inspired by the success of transformers in natural language processing, have recently gained attention in the computer vision community. The selection of ViTs for hedgerow detection is motivated by their unique ability to capture long-range dependencies and global context within images, which can be particularly advantageous in the context of complex landscape analysis [8]. The self-attention mechanism in ViTs allows the model to weigh the importance of different image patches, enabling it to focus on relevant regions and establish relationships between distant parts of the image [34]. This capability can be crucial for hedgerow detection, as hedgerows often exhibit elongated structures that span across large portions of the image and interact with various landscape elements. The ability of ViTs to model these long-range dependencies and global contextual relationships can potentially lead to improved discrimination between hedgerows and other linear features or background clutter. Furthermore, the flexibility and scalability of ViTs make them well-suited for handling the large and high-resolution satellite images commonly used in hedgerow mapping. The potential of ViTs to capture both local details and global context, along with their adaptability to large-scale image analysis, makes them a promising architecture for advancing the state-of-the-art in hedgerow detection.

4.2.1 MiT (Mix Transformer)

MiT [52] (as shown in Fig. 4.5) is a type of vision transformer architecture designed to capture both local and global features in an image efficiently. Unlike traditional Convolutional Neural Networks (CNNs) that rely on convolutional layers to extract features, MiT uses transformer-based mechanisms to process images, which allows it to capture long-range dependencies and relationships between different parts of the image more effectively.

MiT excels in detecting hedgerows in VHR satellite imagery by capturing fine details and considering global context. Its hierarchical feature extraction focuses on the thin, elongated structures of hedgerows, while the self-attention mechanism helps distinguish them from similar features by understanding the entire image context. MiT's robustness to variations in lighting, shadows, and occlusions enhances reliability across different conditions. Additionally, its multi-scale feature extraction ensures effective detection of hedgerows of all sizes, and its computational efficiency makes it suitable for processing large, high-resolution images without sacrificing speed or accuracy [35].

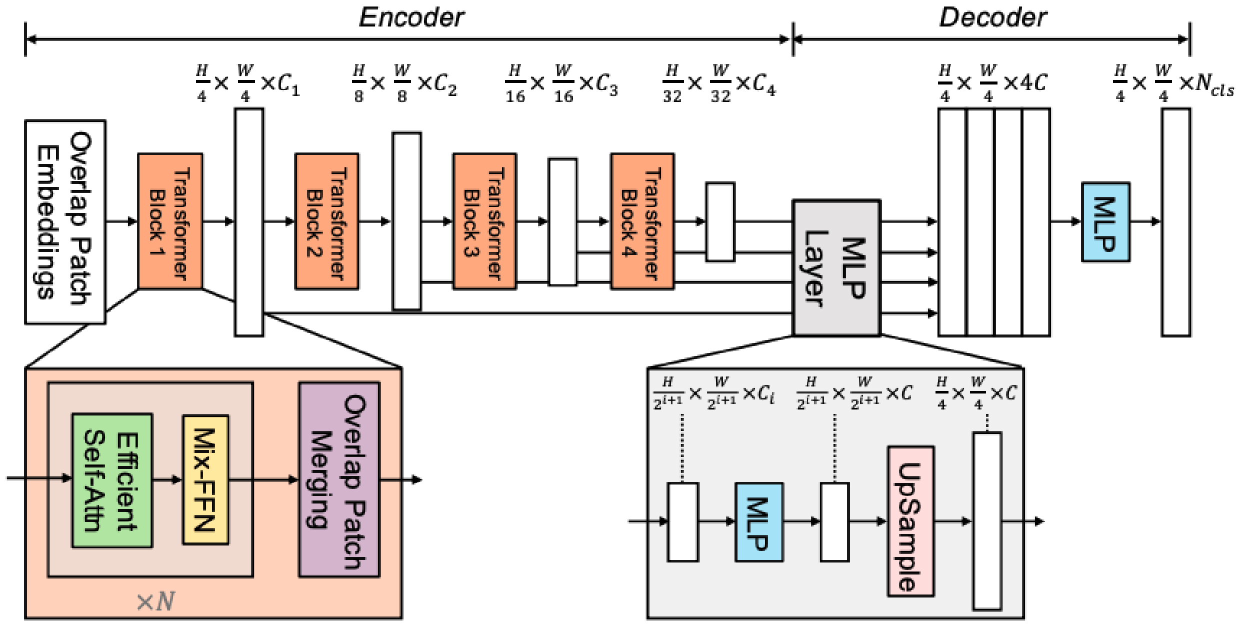

Figure 4.5 Mix Transformer used in SegFormer architecture

5 Hyperparameters, Training and Evaluation Metrics:

The BlueSky dataset was utilized to train 34 distinct model architectures. The hyperparameters for these models were meticulously optimized through a combination of trial and error and leveraging insights from existing model knowledge. The resulting optimal hyperparameters were then consistently applied across all models to ensure a fair and meaningful comparison of their performance.

5.1 Combined Loss Function

The loss function used in training the model is a weighted combination of Dice, Focal, and Binary Cross-Entropy (BCE) boundary losses. Here's why each component is chosen:

Dice Loss: The Dice loss focuses on maximizing the overlap between the predicted segmentation mask and the ground truth mask [36]. It is particularly effective in scenarios with class imbalance, where the target class (hedgerows) occupies a relatively small portion of the image compared to the background. The Dice loss encourages the model to prioritize accurate segmentation of hedgerow regions, even if they are thin or sparsely distributed.

Focal Loss: The Focal loss is designed to address class imbalance by down-weighting the contribution of easy examples (background pixels) and focusing on harder, often misclassified instances (hedgerow pixels) [37]. This helps prevent the model from being dominated by the majority class and ensures that it learns to distinguish subtle features that are crucial for accurate hedgerow detection.

BCE Boundary Loss: The BCE boundary loss enhances the precision of the model at the boundaries of the predicted hedgerow segments [38]. Precise delineation between hedgerows and the background is critical for accurate mapping and analysis. The BCE boundary loss encourages the model to sharpen its predictions at the edges, leading to more accurate and well-defined hedgerow boundaries.

After careful tuning, the ratio that provided the best results was dice weight of 0.3, focal weight of 0.4, and boundary weight of 0.3. This weighted combination of these losses allows for a balanced approach in detecting both the presence and precise boundaries of hedgerows leaning towards a higher focal loss to counteract a slight class imbalance observed during data visualization.

5.2 Chosen Hyperparameters

The following hyperparameters were selected for training the deep learning models:

Learning Rate (0.0001): A small learning rate ensures gradual updates to the model's weights, preventing overshooting and promoting stable convergence during training [39]. This is particularly important for complex models and tasks like hedgerow segmentation, where subtle features need to be learned.

Batch Size (8): A smaller batch size allows the model to learn fine-grained patterns within the hedgerow data while accommodating the memory constraints often associated with high-resolution satellite images [40].

Epochs (500): A large number of epochs provides ample opportunity for the model to learn from the data, especially given the small learning rate. This is crucial for the complex nature of hedgerow detection, where subtle features and contextual relationships are essential for accurate segmentation. Overfitting is addressed by employing Early Stopping.

Optimizer (AdamW): The AdamW optimizer is chosen for its adaptive learning rate and built-in weight decay mechanism, which helps prevent overfitting and improves the generalization capabilities of the model [41].

Scheduler (ReduceLROnPlateau): The ReduceLROnPlateau scheduler dynamically reduces the learning rate when the validation loss plateaus, enabling the model to fine-tune its learning and avoid getting stuck in local minima [42].

Weight Decay (0.0001): Weight decay is applied to penalize large weights, further preventing overfitting and encouraging the model to learn more generalizable features [43].

Patience (25): The patience parameter allows the model to continue training for a certain number of epochs (25 in this case) without improvement in validation loss before the learning rate is reduced. This ensures stability and avoids premature learning rate decay.

Early Stopping: Early stopping is implemented to halt training when no improvement in Intersection over Union (IoU) and accuracy is observed for a specified number of epochs (30). This prevents overfitting and saves computational resources by terminating training when the model's performance on the validation set starts to degrade [46].

Dropout (0.33): Dropout is applied to the decoder layers of the models to further prevent overfitting [44]. By randomly dropping units during training, dropout forces the model to learn redundant representations and reduces its reliance on any single neuron, leading to improved generalization.

5.3 Training Process

The training process follows a structured approach to optimize the model's performance for hedgerow segmentation:

Training Phase: The model is trained up to 500 epochs. It is important to note that architectural complexity directly impacts training duration. Simpler architectures converge with fewer epochs, whereas more intricate models necessitate prolonged training periods to achieve optimal performance. In each epoch, the model is trained on batches of image tiles and their corresponding masks. The optimizer updates the model's weights based on the combined loss function, which guides the model to focus on both accurate hedgerow identification and precise boundary delineation. The model is trained on 35 high-quality training images, which becomes approx. 180 images after data augmentation.

Validation Phase: After each epoch, the model's performance is evaluated on a separate validation set of 14 images. Key metrics, including IoU [46], F1-score, precision, and recall, are calculated to assess the model's ability to segment hedgerows effectively.

Optimizer and Scheduler: The AdamW optimizer adjusts the model's weights based on the calculated gradients, while the ReduceLROnPlateau scheduler dynamically reduces the learning rate if the validation loss plateaus, facilitating fine-tuning and preventing the model from getting stuck in local minima.

Early Stopping: If the model's performance on the validation set, as measured by IoU and accuracy, does not improve for a predefined number of epochs (30), training is terminated to prevent overfitting and conserve computational resources.

5.4 Evaluation Metrics

The performance of the trained models is evaluated using a combination of quantitative metrics and qualitative assessments.

Precision: Precision measures the proportion of correctly predicted hedgerow pixels among all pixels predicted as hedgerows. It ensures that the model minimizes false positives, i.e., identifying non-hedgerow regions as hedgerows.

Recall: Recall measures the proportion of correctly predicted hedgerow pixels among all actual hedgerow pixels in the ground truth. It ensures that the model maximizes the detection of true hedgerows, even if it leads to some false positives.

F1 Score: The F1 score is the harmonic mean of precision and recall, providing a balanced measure of the model's overall performance. It is particularly valuable in scenarios where both precision and recall are important, as is the case in hedgerow segmentation.

Accuracy: Accuracy measures the proportion of correctly classified pixels (both hedgerow and background) among all pixels in the image. While accuracy provides a general sense of model performance, it may not be the most informative metric in cases of class imbalance, where the background class dominates.

IoU (Intersection over Union): IoU is a widely used metric for evaluating segmentation performance [45]. It measures the overlap between the predicted segmentation mask and the ground truth mask, providing a direct assessment of the model's ability to correctly identify and delineate hedgerows.

These metrics collectively provide a comprehensive evaluation of the model's performance, ensuring that it is not only accurate but also reliable in detecting hedgerows under various environmental conditions and image characteristics. The emphasis on IoU [46] as the primary metric reflects its direct relevance to the task of hedgerow segmentation, where precise boundary delineation is crucial.

6 Results and Evaluation:

The culmination of the model training and development process lies in the rigorous evaluation and analysis of their performance on the task of hedgerow detection. This chapter presents a systematic assessment of the 34 trained models, meticulously examining their capabilities in accurately identifying and delineating hedgerows within VHR satellite imagery. From the pool of 34 trained models, a selection of the top 3 models is made based on a combination of key performance indicators:

Intersection over Union (IoU): The primary metric for evaluating segmentation accuracy, IoU quantifies the overlap between predicted and ground truth hedgerow masks. Models demonstrating high IoU scores are prioritized, indicating their proficiency in precisely delineating hedgerow boundaries.

Recall: Recall measures the model's ability to detect all actual hedgerows, even if it leads to some false positives. The selection process favours models with high recall values, ensuring comprehensive hedgerow identification and minimizing the risk of missing critical landscape features.

Training Duration and Performance: The evaluation also considers the training duration of each model, recognizing that some architectures may require extended training to fully realise their potential. Models exhibiting good IoU and Recall scores, even with longer training durations, are included in the top 3 selections, acknowledging the possibility of further performance gains with continued training.

This multi-faceted selection approach aims to identify models that not only excel in accuracy and completeness but also demonstrate potential for further improvement and adaptability to real-world applications. The subsequent sections of this chapter will delve into the detailed evaluation and analysis of these top 3 models, providing insights into their performance characteristics, architectural advantages, and potential contributions to the field of hedgerow detection using deep learning.

6.1 U-Nets

The table below presents the top-performing U-Net models based on recall and IoU, along with their average performance across all metrics.

Model Name	Epoch	Precision	Recall	F1 Score	IoU	Train Loss	Validation Loss
Unet-resnext50_32x4d	49	0.828	0.844	0.836	0.718	0.125	0.115
Unet-ResNet200e	65	0.767	0.874	0.817	0.690	0.236	0.218
Unet-resnet101	173	0.795	0.859	0.825	0.703	0.130	0.134

Figure 6.1 Val Loss (UNets)

Figure 6.2 IoU Curve (UNets)

Figure 6.3 Recall Curve (UNets)

Critical Evaluation of Models

Unet-resnext50_32x4d (Best IoU): This model demonstrates the highest IoU score of 0.718 (shown in Fig. 6.2), indicating the most accurate overlap between the predicted segmentation and the ground truth. It will be particularly suitable when precise delineation of hedgerows is crucial, such as in applications where accurate boundary detection is important for subsequent analysis or decision-making. However, its Recall is slightly lower compared to the other two models, suggesting it might miss a few hedgerow instances. It is also noteworthy that this model's training and validation losses are the lowest among the three, indicating good convergence during training.

Unet-ResNest200e (Best Recall): With the highest Recall of 0.874 (shown in Fig. 6.3), this model excels at identifying most hedgerow pixels in the images. It is valuable when the focus is on minimizing false negatives, i.e., ensuring that as few hedgerows as possible are missed during segmentation. This could be beneficial in scenarios where even small hedgerow fragments are important to detect, such as in ecological studies assessing hedgerow connectivity. However, its IoU is slightly lower, implying that the predicted segmentation might not perfectly align with the ground truth boundaries. This model also exhibits the highest training and validation losses, suggesting some room for further improvement in terms of model fit.

Unet-resnet101 (Longest Trained): This model was trained for the longest duration (173 epochs) while maintaining good IoU and Recall scores. Longer training can lead to better model performance, as it allows the model to learn more complex patterns in the data. This model might be preferred when there is sufficient computational resources and time available for training, and a balance between IoU and Recall is desired. It could be a good choice for general-purpose hedgerow segmentation tasks where both precise boundary detection and comprehensive hedgerow identification are important. Its training and validation losses lie between the other two models, suggesting a reasonable trade-off between model fit and complexity.

6.2 FPNets

The following table showcases the performance of the FPN models:

Model Name	Epoch	Precision	Recall	F1 Score	IoU	Train Loss	Validation Loss
FPN-resnext50_32x4d	35	0.799	0.838	0.818	0.692	0.313	0.123
FPN-resnet101	43	0.741	0.872	0.801	0.669	0.241	0.131
FPN-resnext101_32x8d	54	0.765	0.865	0.812	0.683	0.232	0.122

Figure 6.4 Val Loss (FPNets)

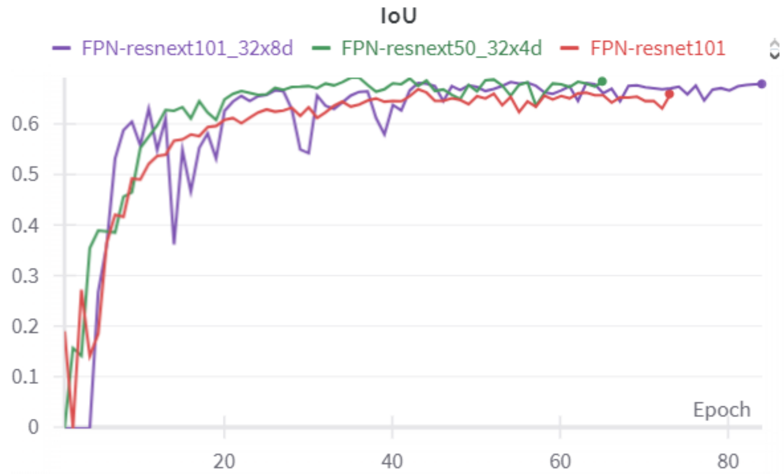

Figure 6.5 IoU Curve (FPNets)

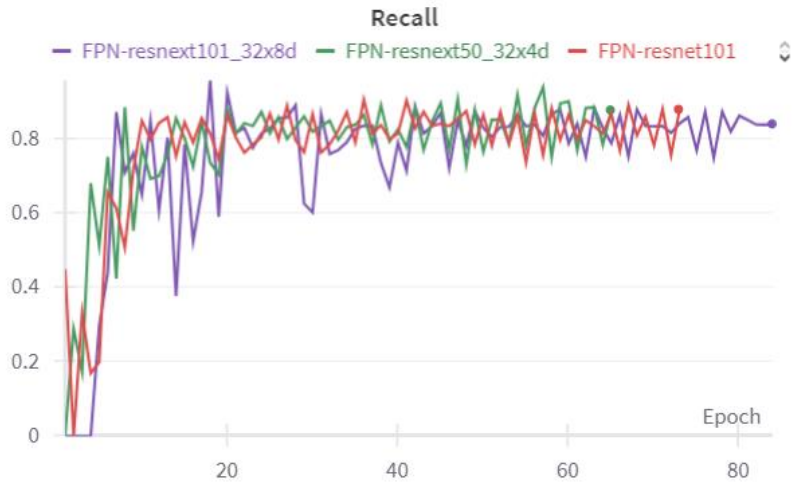

Figure 6.6 Recall Curve (FPNets)

Critical Evaluation of Models

FPN-resnext50_32x4d (Best IoU): This model shows the best IoU at 0.692 (shown in Fig. 6.5), indicating precise overlap between predicted and actual hedgerows. It's ideal when accurate boundary detection is critical. Though its Recall is slightly lower, it's still good at identifying most hedgerows.

FPN-resnet101 (Best Recall): With the highest Recall of 0.872 (shown in Fig. 6.6), this model is excellent at finding most hedgerow pixels, minimizing missed instances. It's valuable when detecting even small hedgerow fragments is important. While its IoU is slightly lower, it still offers good segmentation accuracy.

FPN-resnext101_32x8d (Longest Trained): Trained the longest (54 epochs), this model balances good IoU and Recall. Longer training allows learning complex patterns, potentially leading to better overall performance. It's a strong choice when resources permit and a balance between accuracy and comprehensiveness is desired.

6.3 PSPNets

The following table showcases the performance of the PSP models:

Model Name	Epoch	Precision	Recall	F1 Score	IoU	Train Loss	Validation Loss
PSP-resnext101_32x8d	52	0.761	0.786	0.773	0.630	0.285	0.152
PSP-resnet101	39	0.693	0.835	0.757	0.609	0.309	0.166
PSP-resnext50_32x4d	75	0.740	0.792	0.765	0.619	0.241	0.157

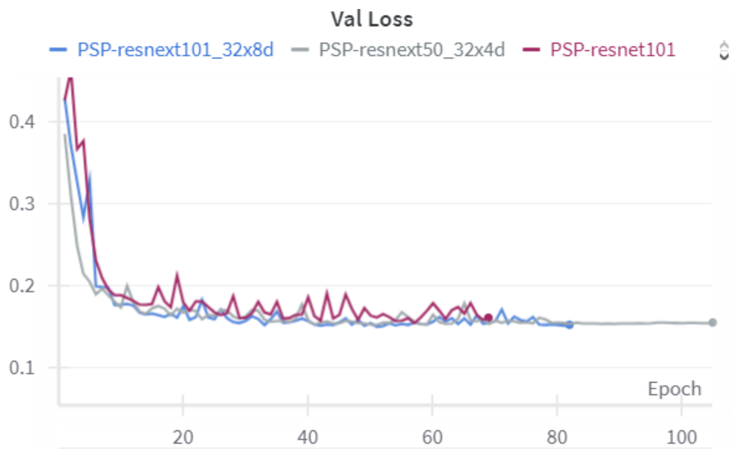

Figure 6.7 Val Loss (PSP)

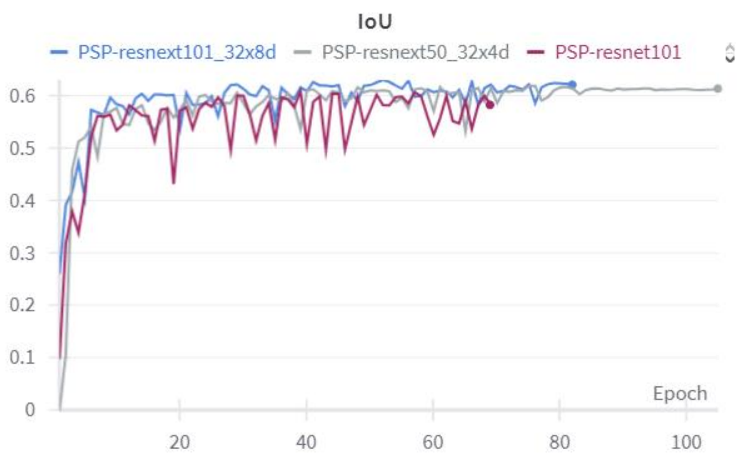

Figure 6.8 IoU Curve (PSP)

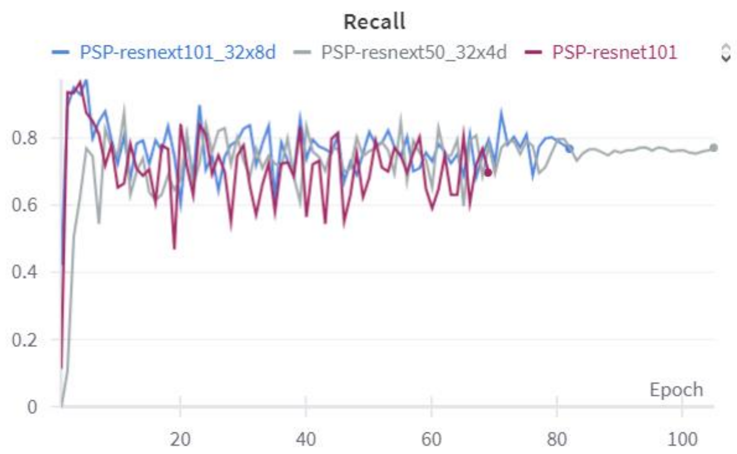

Figure 6.9 Recall Curve (PSP)

Critical Evaluation of Models

PSP-resnext101_32x8d (Best IoU): This model has the highest IoU of 0.630 (shown in Fig. 6.8), indicating the best overlap between predicted and true hedgerows. It's suitable when precise boundary detection is crucial. While its Recall is good, it might miss slightly more hedgerows compared to the model with the best recall.

PSP-resnet101 (Best Recall): With the highest Recall of 0.835 (shown in Fig. 6.9), this model excels at identifying most hedgerow pixels, minimizing missed instances. It's valuable when detecting even small hedgerow fragments is important. Though its IoU is lower, it still provides decent segmentation accuracy.

PSP-resnext50_32x4d (Longest Trained): Trained the longest (75 epochs), this model offers a balance between IoU and Recall. Longer training might lead to better performance by learning complex data patterns. It's a good choice when resources allow and a balance between accuracy and comprehensiveness is needed.

6.4 LinkNet

The following table showcases the performance of the specified LinkNet models:

Model Name	Epoch	Precision	Recall	F1 Score	IoU	Train Loss	Validation Loss
LinkNet-resnet34	35	0.798	0.844	0.820	0.695	0.309	0.203
LinkNet-mobilenet v3_large_100	111	0.786	0.838	0.811	0.682	0.357	0.191

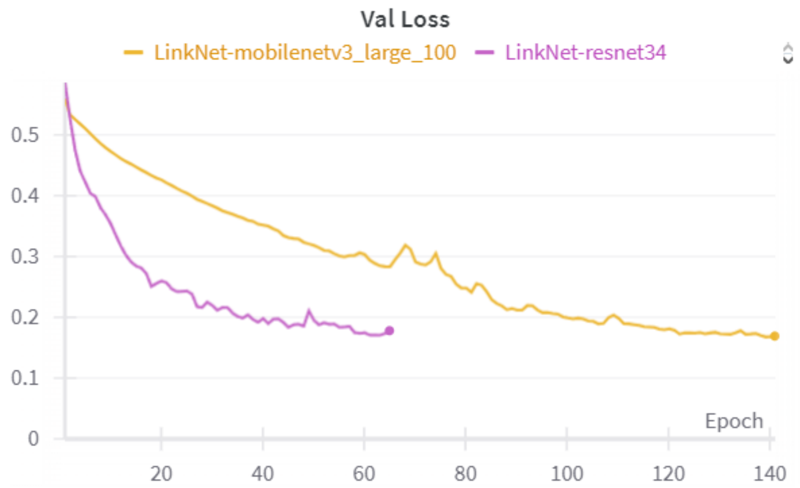

Figure 6.10 Val Loss (LinkNet)

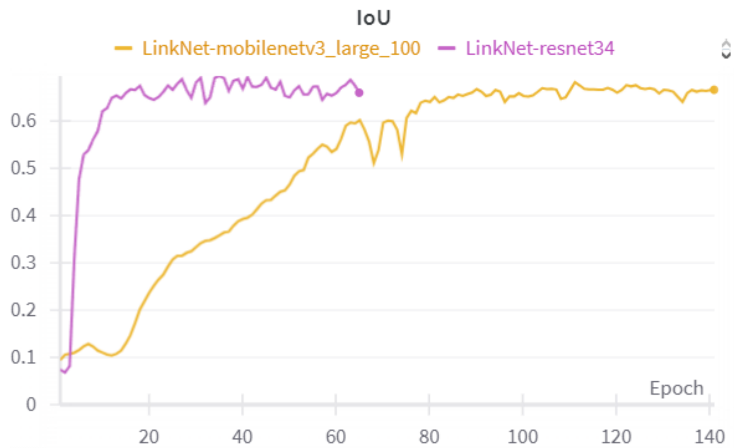

Figure 6.11 IoU Curve (LinkNet)

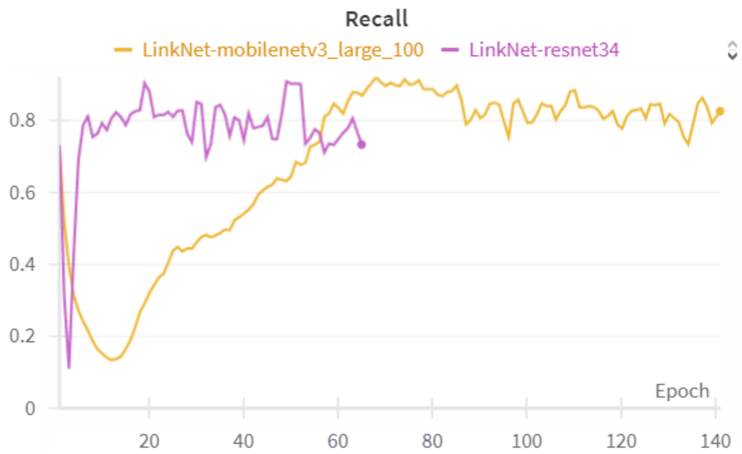

Figure 6.12 Recall Curve (LinkNet)

Critical Evaluation of Models

LinkNet-resnet34 (Best IoU and Best Recall): This model stands out with the highest IoU of 0.695 (shown in Fig. 6.11 and 6.12), indicating precise overlap between predicted and true hedgerows. It also boasts the best Recall at 0.844, showcasing its ability to identify most hedgerow pixels and minimize missed instances. This makes it a strong contender for various hedgerow segmentation tasks, especially when both accurate boundary detection and comprehensive hedgerow identification are important.

LinkNet-mobilenetv3_large_100 (Longest Trained): Trained for the longest duration (111 epochs), this model maintains good IoU and Recall scores. While its performance is slightly lower than LinkNet-resnet34, the longer training might have allowed it to learn more intricate patterns in the data, potentially leading to better generalization on unseen examples. It's a suitable choice when computational resources and time permits, and a balance between accuracy and comprehensiveness is desired.

6.5 ViT

The following table showcases the performance of the specified ViT models:

Model Name	Epoch	Precision	Recall	F1 Score	IoU	Train Loss	Validation Loss
ViT_mit_b2_unet	75	0.787	0.838	0.811	0.683	0.297	0.157
ViT_mit_b4_fpn	73	0.765	0.846	0.803	0.671	0.296	0.132
ViT_mit_b2_fpn	84	0.751	0.844	0.795	0.660	0.314	0.135

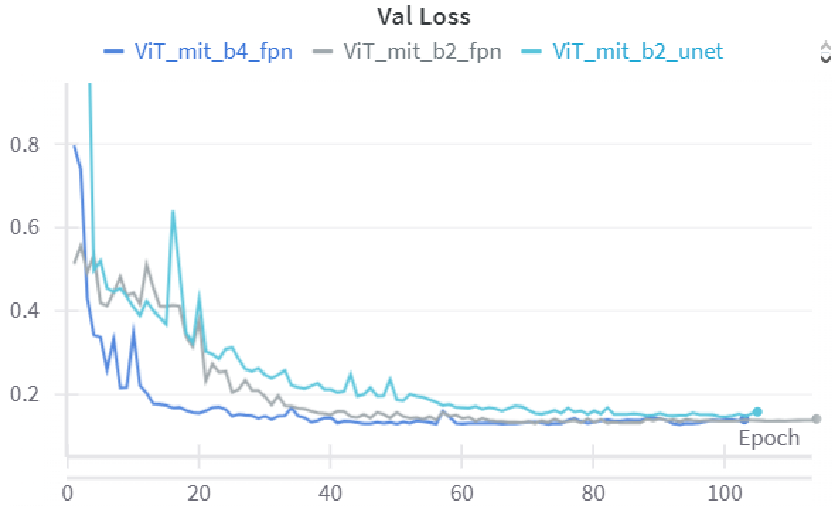

Figure 6.13 Val Loss (ViT)

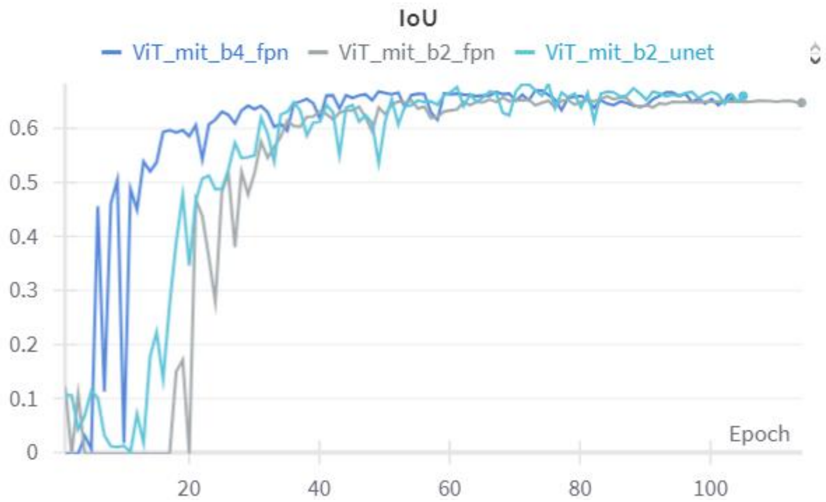

Figure 6.14 IoU Curves (ViT)

Figure 6.15 Recall Curves (ViT)

Critical Evaluation of Models

ViT_mit_b2_unet (Best IoU): This model demonstrates the highest IoU score of 0.683 (shown in Fig. 6.14), indicating the most accurate overlap between the predicted segmentation and the ground truth. It might be particularly suitable when precise delineation of hedgerows is crucial, such as in applications where accurate boundary detection is important for subsequent analysis or decision-making.

ViT_mit_b4_fpn (Best Recall): With the highest Recall of 0.846 (shown in Fig. 6.15), this model excels at identifying most hedgerow pixels in the images. It is valuable when the focus is on minimizing false negatives, i.e., ensuring that as few hedgerows as possible are missed during segmentation. This could be beneficial in scenarios where even small hedgerow fragments are important to detect, such as in ecological studies assessing hedgerow connectivity.

ViT_mit_b2_fpn (Longest Trained): This model was trained for the longest duration (84 epochs) while maintaining good IoU and Recall scores. Longer training can sometimes lead to better model performance, as it allows the model to learn more complex patterns in the data. This model might be preferred when there is sufficient computational resources and time available for training, and a balance between IoU and Recall is desired. It could be a good choice for general-purpose hedgerow segmentation tasks where both precise boundary detection and comprehensive hedgerow identification are important.

7 Conclusion, Recommendation and Future Scope

7.1 Summary of Findings

The analysis across different architectures (Unet, FPN, PSP, and ViT) for hedgerow segmentation reveals that Unet-based models consistently outperformed others in terms of achieving higher IoU and Recall scores. The dominance of Unet can be attributed to several architectural advantages, particularly in the context of the relatively small training dataset (approximately 150 images after augmentation).

Why Unet Excelled

Encoder-Decoder Structure with Skip Connections: Unet's symmetric encoder-decoder structure, coupled with skip connections, facilitates effective capturing of both high-level semantic information (from the encoder) and fine-grained spatial details (from the decoder). This is crucial for precise boundary delineation in segmentation tasks, especially when dealing with intricate structures like hedgerows. The skip connections enable the model to recover spatial information lost during downsampling in the encoder, aiding in accurate localization of hedgerow boundaries.

Efficient Use of Data: The Unet architecture is known for its ability to perform well even with limited training data. This is due to its inherent data augmentation capabilities through the use of mirrored skip connections and its ability to learn from both local and global context. In the context of this analysis, where the training dataset is relatively small, Unet's data efficiency likely played a significant role in its superior performance.

Why Other Architectures Lagged

FPN (Feature Pyramid Network): While FPN is designed to handle objects at different scales, its focus on multi-scale feature fusion might not be as crucial for hedgerow segmentation, where the objects of interest (hedgerows) tend to have a relatively consistent scale. Additionally, FPN's lateral connections might not be as effective as Unet's skip connections in preserving fine-grained spatial details, potentially leading to less precise boundary detection.

PSP (Pyramid Scene Parsing Network): PSP employs pyramid pooling to capture global context information, which can be beneficial for scene understanding tasks. However, for hedgerow segmentation, the emphasis is more on local context and precise boundary localization. PSP's reliance on global context might lead to less accurate segmentation of hedgerow boundaries, especially in the presence of complex backgrounds or occlusions.

ViT (Vision Transformer): ViT models have shown impressive performance in image classification tasks, but their adaptation to segmentation tasks is still an active area of research. ViT's reliance on self-attention mechanisms might make it less efficient in capturing local spatial dependencies, which are crucial for accurate segmentation. Additionally, ViT models typically require large amounts of training data to achieve their full potential, which could explain their relatively lower performance in this analysis with a limited dataset.

Conclusion

Unet's architectural advantages, particularly its encoder-decoder structure with skip connections and efficient use of data, likely contributed to its superior performance in hedgerow segmentation compared to FPN, PSP, and ViT models, especially in the context of a small training dataset. The other architectures, while powerful in their own right, might not be as well-suited for this specific task due to their focus on multi-scale feature fusion (FPN), global context (PSP), or reliance on self-attention mechanisms and large datasets (ViT).

7.2 Recommendation

Based on the analysis of Unet, FPN, PSP, and ViT models, the following table presents the absolute best 3 models considering IoU, Recall, and longer training epochs with good IoU and Recall:

Model Name	Architecture	Epoch	Precision	Recall	F1 Score	IoU	Train Loss	Validation Loss
Unet-resnext50_32x4d	Unet	49	0.828	0.844	0.836	0.718	0.125	0.115
Unet-ResNest200e	Unet	65	0.767	0.874	0.817	0.690	0.236	0.218
Unet-resnet101	Unet	173	0.794	0.859	0.825	0.703	0.130	0.134

Unet-resnext50_32x4d (Best IoU): This model, based on the Unet architecture, demonstrates the highest IoU score of 0.718, making it the best performer in terms of precise boundary detection.

Unet-ResNest200e (Best Recall): Also based on the Unet architecture, this model exhibits the highest Recall of 0.874, making it the top choice when the focus is on ensuring that as few hedgerows as possible are missed during segmentation.

Unet-resnet101 (Longest Trained): Another Unet model, this one was trained for the longest duration (173 epochs) while maintaining good IoU and Recall scores, making it a suitable choice for general-purpose hedgerow segmentation tasks when computational resources and time permit longer training.

Their predicted hedgerows can be found below. Fig. 7.1 shows the original image, Fig. 7.2 shows the ground truth and Fig. 7.3 shows the predictions of all the top 3 models.

Figure 7.1 Original Image

Figure 7.2 Ground Truth

Predictions:

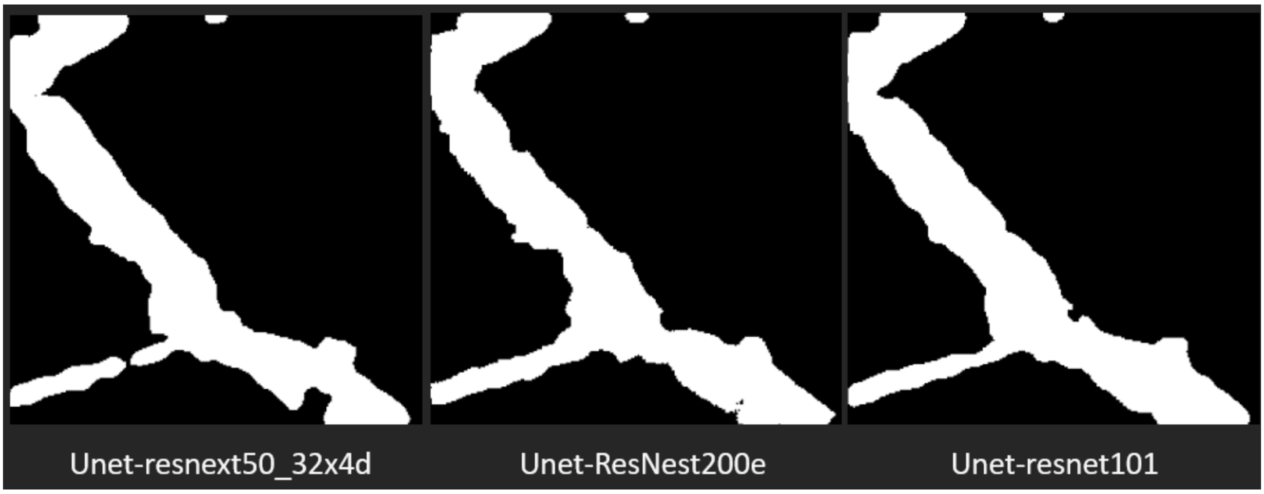

Figure 7.3 Predicted hedgerows of top 3 models

In the pursuit of leveraging existing knowledge and resources, transfer learning was employed in this study. A Unet-resnet101 model, meticulously trained on the BlueSky dataset, was subsequently applied to predict hedgerows in Airbus's VHR multispectral data. While this approach yielded promising results in some instances (shown in Fig. 7.4), it also faced challenges in handling specific image types within the Airbus dataset (shown in Fig. 7.5).

The discrepancy in performance can be attributed to several factors stemming from the inherent differences between aerial photography (BlueSky) and VHR satellite imagery (Airbus).

Spectral Differences: The Airbus VHR data encompasses four channels (R, G, B, Near-infrared), while the BlueSky model was trained exclusively on three channels (R, G, B). To accommodate this, only the RGB channels from the Airbus data were utilized for prediction, potentially discarding valuable information present in the near-infrared band.

Pixel Value Range: The pixel value ranges differed between the two datasets, with BlueSky images ranging from 0-255 and Airbus images ranging from 0-3858. Normalization was applied to align the Airbus data with the BlueSky model's expectations.

Image Resolution and Scale: While both datasets are considered high-resolution, aerial photography generally captures finer details compared to satellite imagery. This discrepancy in resolution might have impacted the model's ability to discern subtle hedgerow features in the Airbus data.

Perspective and Viewing Angle: Aerial images often present a more oblique view with greater perspective distortion, while satellite images offer a near-nadir view. This difference in perspective could lead to variations in how hedgerows appear, potentially confusing a model trained on aerial data.

These challenges highlight the complexities of transferring knowledge between datasets acquired from different sources and with varying characteristics. Future work could explore advanced domain adaptation techniques to bridge the gap between aerial and satellite imagery, enabling more robust and generalizable hedgerow detection models.

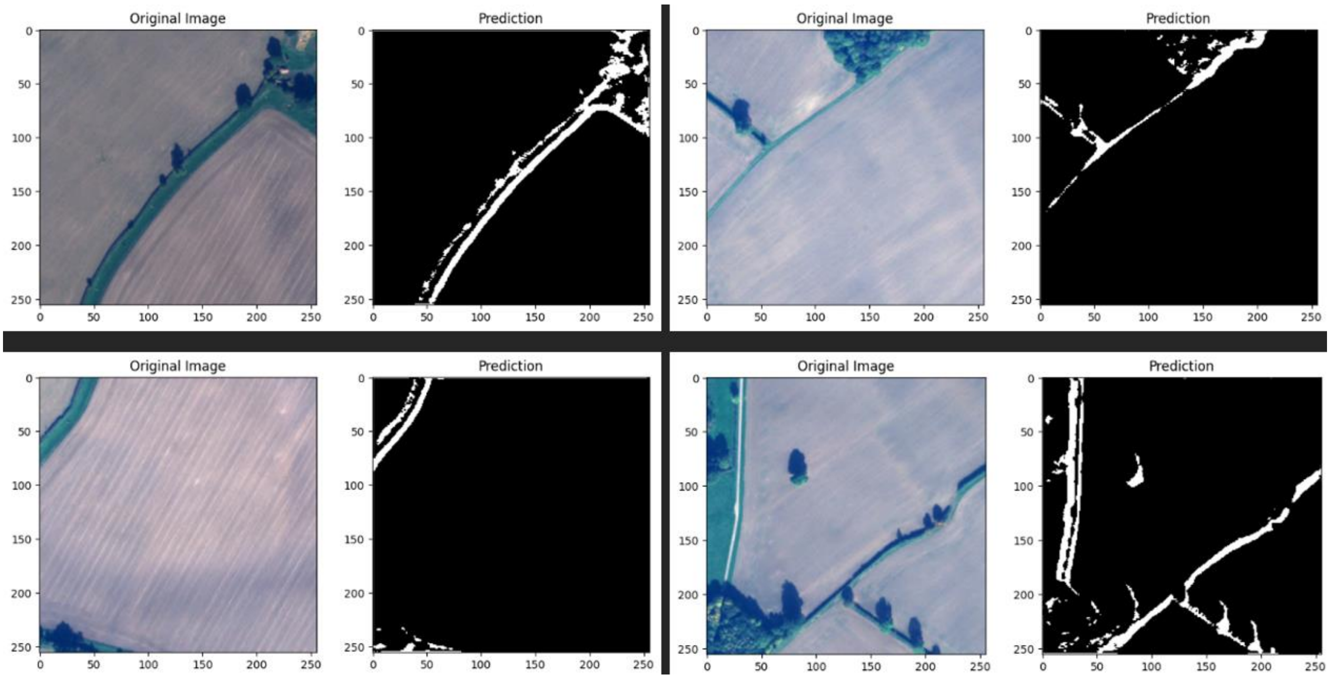

Figure 7.4 Good Results obtained by performing transfer learning using Unet-resnet101 model trained on Bluesky data on Airbus data

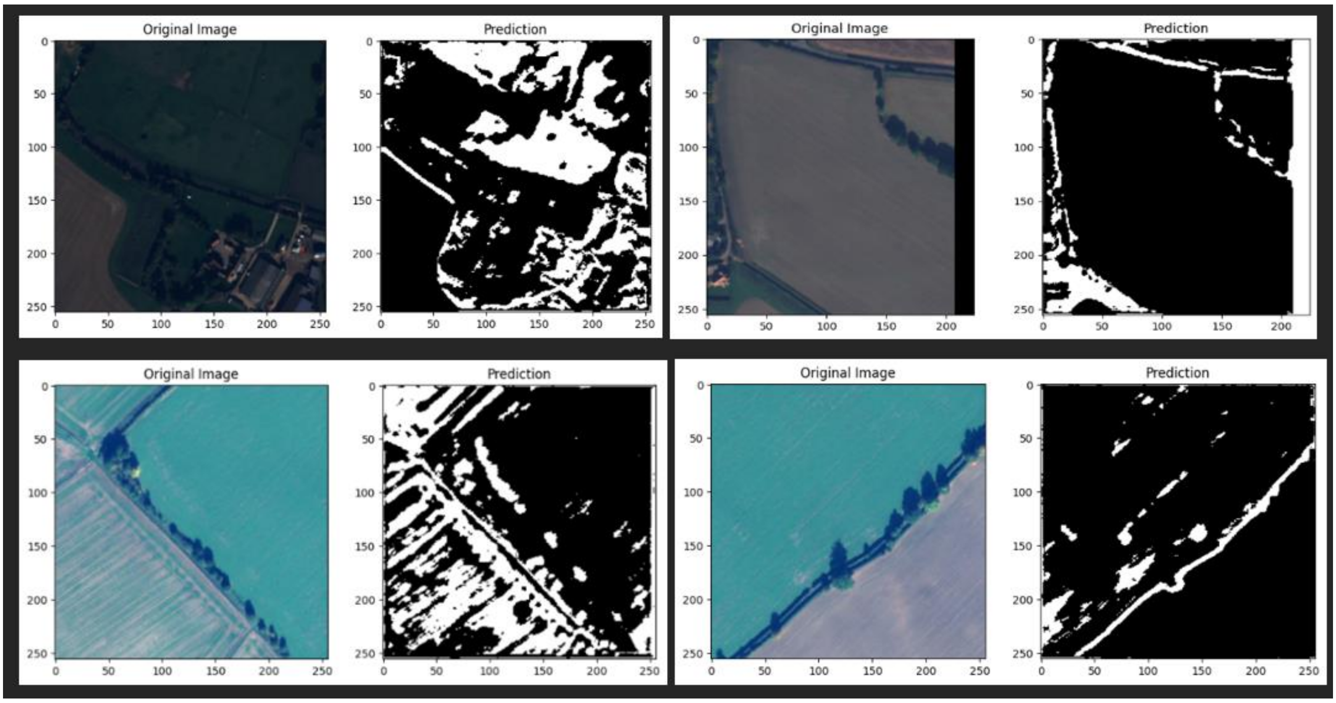

Figure 7.5 Failed Results obtained by performing transfer learning using Unet-resnet101 model trained on Bluesky data on Airbus data

7.3 Future Scope

The current work has demonstrated the potential of deep learning for hedgerow detection by training models on the BlueSky dataset and achieving promising results. However, there remains significant room for advancement and refinement.

Leveraging Enhanced Ground Truth Data

The forthcoming revamped proxy ground truth data from UKCEH presents an exciting opportunity to elevate the accuracy and robustness of hedgerow detection models. Integrating this improved data into the training and evaluation process can potentially address the current accuracy limitations and lead to a groundbreaking model capable of delivering highly precise hedgerow segmentation in VHR satellite imagery.

Expanding to Multispectral VHR Data

The mixed results observed when transferring the BlueSky-trained model to Airbus's VHR multispectral data highlight the challenges of domain adaptation. Future work should focus on developing advanced techniques to bridge the gap between aerial and satellite imagery domains. Exploring methods like domain adversarial training or style transfer could significantly enhance the performance of models trained on aerial photography when applied to VHR satellite data, unlocking the full potential of transfer learning in this context.

Multi-Modal and Multi-Scale Approaches

Investigating the fusion of multi-modal data, such as combining VHR satellite imagery with LiDAR or aerial photography, could further improve hedgerow detection by leveraging complementary information from different sources. Additionally, exploring multi-scale architectures and incorporating attention mechanisms could help the models better capture hedgerows of varying sizes and shapes within complex landscapes.

References

[1] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[2] F. Burel and J. Baudry, Ecologie du Paysage. Concepts, Méthodes et Applications. Paris, France: Tec & Doc, 2003.

[3] Bluesky International Limited, “National Hedgerow Map,” 2023. [Online]. Available: https://bluesky-world.com/national-hedgerow-map/

[4] G. M. Smith and R. J. Fuller, “Hedgerow surveying: Methods and limitations,” J. Environ. Manage., vol. 63, no. 2, pp. 143–154, 2001.

[5] N. Kussul, M. Lavreniuk, S. Skakun, and A. Shelestov, “Deep learning classification of land cover and crop types using remote sensing data,” IEEE Geosci. Remote Sens. Lett., vol. 14, no. 5, pp. 778–782, May 2017.

[6] H. Guan, Y. Zhang, and L. Zhong, “A review of current machine learning methods for recognizing plant species from remote sensing data,” Remote Sens., vol. 12, no. 18, p. 2990, 2020.

[7] I. Dronova, P. Gong, L. Wang, and L. Zhong, “Mapping dynamic cover types in a heterogeneous savanna landscape: An object-based approach integrating Landsat NDVI time series and ALOS PALSAR data,” Remote Sens. Environ., vol. 164, pp. 245–258, Jun. 2015.

[8] J. Baudry and F. Burel, “Hedgerows: An international perspective on their origin, function and management,” J. Environ. Manage., vol. 60, no. 1, pp. 7–22, 2000.

[9] G. Cheng, J. Han, P. Zhou, and L. Guo, “Object detection in remote sensing imagery using a deep convolutional neural network,” ISPRS J. Photogramm. Remote Sens., vol. 134, pp. 58–67, 2017.

[10] R. T. T. Forman and J. Baudry, “Hedgerows and hedgerow networks in landscape ecology,” Environ. Manage., vol. 8, no. 6, pp. 495–510, Nov. 1984.

[11] S. A. Hinsley, P. E. Bellamy, I. Newton, and T. H. Sparks, “The influence of hedgerows on the over-wintering survival of farmland birds,” J. Appl. Ecol., vol. 37, no. 3, pp. 486–500, 2000.

[12] J. Tews, U. Brose, V. Grimm, K. Tielbörger, M. C. Wichmann, M. Schwager, and F. Jeltsch, “Animal species diversity driven by habitat heterogeneity/diversity: The importance of keystone structures,” J. Biogeogr., vol. 31, no. 1, pp. 79–92, 2004.

[13] T. G. Benton, J. A. Vickery, and J. D. Wilson, "Farmland biodiversity: Is habitat heterogeneity the key?," Trends Ecol. Evol., vol. 18, no. 4, pp. 182–188, Apr. 2003.

[14] J. M. Dorioz, T. Fourcaud, and J. L. González-Andujar, "Effects of a dense hedgerow network on runoff, erosion and soil properties in a cultivated Mediterranean catchment," Agr. Ecosyst. Environ., vol. 113, no. 1–4, pp. 28–41, Jan. 2006.

[15] R. C. Schwartz and S. R. Evett, "Estimating the influence of vegetation and management practices on water yield using a spatially distributed water balance model," J. Hydrol., vol. 265, no. 1–4, pp. 171–187, Sep. 2002.

[16] J. R. Brandle, L. Hodges, and X. H. Zhou, "Microclimate and water balance in a young hedgerow in semiarid central Kansas," Agroforest. Syst., vol. 61, no. 1–3, pp. 325–338, 2004.

[17] C. Bannister and T. A. Watt, "The influence of a hedgerow on small-scale spatial variation in the microclimate and soil moisture of an adjacent winter wheat field," Agr. Forest Meteorol., vol. 105, no. 1–3, pp. 259–274, Nov. 2000.

[18] P. Smith et al., "Meeting Europe's climate change commitments: Quantitative estimates of the potential for carbon mitigation by agriculture," Glob. Change Biol., vol. 6, no. 5, pp. 525–539, 2000.

[19] S. Ahlswede, C. Neumann, and M. Baatz, "Hedgerow object detection in very high-resolution satellite images using convolutional neural networks," J. Appl. Remote Sens., vol. 15, no. 1, p. 018501, 2021.

[20] C. Zhang, I. Sargent, Z. Pan, and D. Li, "A multi-scale convolutional neural network for hedgerow detection in VHR satellite imagery," Remote Sens., vol. 15, no. 7, p. 1856, 2023.

[21] Y. Liu, J. Wu, and L. Zhang, "Hedgerow detection in VHR satellite imagery using attention-based convolutional neural networks," IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens., vol. 15, pp. 4567–4578, 2022.

[22] J. Quintero and A. Otero, "Hedgerow detection and mapping using very high-resolution satellite imagery and deep learning," Remote Sens., vol. 11, no. 12, p. 1461, 2019.

[23] C. Zhang, D. Li, and Z. Pan, "An improved U-Net model for hedgerow extraction from high-resolution remote sensing images," Remote Sens., vol. 13, no. 15, p. 2943, 2021.

[24] O. Csillik, M. Belgiu, and C. Eisank, “Benchmarking convolutional neural networks for hedgerow detection in aerial images,” Remote Sens., vol. 10, no. 10, p. 1514, 2018.

[25] H. Huang and L. Zhang, “Fusion of LiDAR data and optical imagery for hedgerow detection using deep learning,” ISPRS J. Photogramm. Remote Sens., vol. 167, pp. 11–22, Sep. 2020.

[26] X. X. Zhu et al., “Deep learning in remote sensing: A comprehensive review and list of resources,” IEEE Geosci. Remote Sens. Mag., vol. 5, no. 4, pp. 8–36, Dec. 2017.

[27] QGIS Development Team, “QGIS Geographic Information System,” Open Source Geospatial Foundation Project, 2023. [Online]. Available: http://qgis.osgeo.org

[28] C. Shorten and T. M. Khoshgoftaar, “A survey on image data augmentation for deep learning,” J. Big Data, vol. 6, no. 1, pp. 1–48, Dec. 2019.

[29] L. Perez and J. Wang, “The effectiveness of data augmentation in image classification using deep learning,” arXiv preprint arXiv:1712.04621, 2017.

[30] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” in Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention, 2015, pp. 234–241.

[31] T.-Y. Lin et al., “Feature pyramid networks for object detection,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2017, pp. 2117–2125.

[32] H. Zhao et al., “Pyramid scene parsing network,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2017, pp. 2881–2890.

[33] A. Chaurasia and E. Culurciello, “LinkNet: Exploiting encoder representations for efficient semantic segmentation,” in Proc. 2017 IEEE Vis. Commun. Image Process. (VCIP), 2017, pp. 1–4.

[34] A. Dosovitskiy et al., “An image is worth 16x16 words: Transformers for image recognition at scale,” arXiv preprint arXiv:2010.11929, 2020.

[35] A. R. Sharma, R. Garg, A. Mishra, and S. K. Singh, “Transformers in remote sensing: A survey,” IEEE Geosci. Remote Sens. Mag., vol. 10, no. 2, pp. 255–278, Jun. 2022, doi: 10.1109/MGRS.2022.3150385.

[36] F. Milletari, N. Navab, and S.-A. Ahmadi, “V-Net: Fully convolutional neural networks for volumetric medical image segmentation,” in Proc. 2016 Fourth Int. Conf. 3D Vis. (3DV), 2016, pp. 565–571, doi: 10.1109/3DV.2016.72.

[37] T.-Y. Lin et al., “Focal loss for dense object detection,” in Proc. IEEE Int. Conf. Comput. Vis., 2017, pp. 2980–2988.

[38] H. Kervadec, J. Bouchtiba, C. Desrosiers, I. B. Ayed, and J. Dolz, “Boundary loss for highly unbalanced segmentation,” in Med. Imaging Deep Learn., 2019.

[39] Y. Bengio, “Practical recommendations for gradient-based training of deep architectures,” in Neural Networks: Tricks of the Trade, 2nd ed. Berlin, Germany: Springer, 2012, pp. 437–478.

[40] D. Masters and C. Luschi, “Revisiting small batch training for deep neural networks,” arXiv preprint arXiv:1804.07612, 2018.

[41] I. Loshchilov and F. Hutter, “Decoupled weight decay regularization,” arXiv preprint arXiv:1711.05101, 2017.

[42] L. Li et al., “Hyperband: A novel bandit-based approach to hyperparameter optimization,” J. Mach. Learn. Res., vol. 18, no. 1, pp. 6765–6816, 2017.

[43] A. Krogh and J. A. Hertz, “A simple weight decay can improve generalization,” in Adv. Neural Inf. Process. Syst., 1992, pp. 950–957.

[44] L. Prechelt, “Early stopping—But when?,” in Neural Networks: Tricks of the Trade, G. Orr and K. Müller, Eds. Berlin, Germany: Springer, 1998, pp. 55–69.

[45] N. Srivastava et al., “Dropout: A simple way to prevent neural networks from overfitting,” J. Mach. Learn. Res., vol. 15, no. 1, pp. 1929–1958, 2014.

[46] H. Rezatofighi et al., “Generalized intersection over union: A metric and a loss for bounding box regression,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2019, pp. 658–666.

[47] UK Centre for Ecology & Hydrology, “UKCEH Land Cover plus: Hedgerows 2016-2021 (England),” 2023. [Online]. Available: https://www.ceh.ac.uk/data/ukceh-land-cover-plus-hedgerows-2016-2021-england

[48] “UNets”, Qubvel, 2019, https://github.com/qubvel/segmentation_models/blob/master/images/unet.png.

[49] “FPNs”, Qubvel, 2019, https://github.com/qubvel/segmentation_models/blob/master/images/fpn.png.

[50] “PSPs”, Qubvel, 2019,
https://github.com/qubvel/segmentation_models/blob/master/images/pspnet.png.

[51] “LinkNets”, Qubvel, 2019,
https://github.com/qubvel/segmentation_models/blob/master/images/linknet.png.

[52] “SegFormer for image segmentation,” Keras, 2023. [Online]. Available:
https://keras.io/examples/vision/segformer/

Appendix 1

Weights and biases link for more insights on all 34 models:

https://api.wandb.ai/links/eashwar408-university-of-stirling/amhlp70h