Hybrid Image Reconstruction Algorithm for Low-Resolution Images

低解像度画像用のハイブリッド画像再構成アルゴリズム

山梨大学大学院

医工農学総合教育部

博士課程学位論文

２０２４年３月

NG SUIT MUN

SUMMARY

This dissertation explores the advanced field of Image Super-Resolution (SR), a method that enhances Low-Resolution (LR) images into High-Resolution (HR) versions by filling in missing high-frequency components. SR is divided into two significant types: Multiple Image Super-Resolution (MISR) and Single Image Super-Resolution (SISR). MISR combines multiple LR images to create a single detailed and noise-minimized HR image, but it requires exact alignment, which can be time-consuming. SISR, in contrast, creates an HR image from a single LR image, offering broader adaptability but with the challenge of deriving detailed information from limited data.

The techniques used in image SR fall into three primary categories: interpolation-based, reconstruction-based, and learning-based. Interpolation-based methods like nearest neighbour, bilinear, and bicubic interpolation are simple but often result in blurred HR images. Reconstruction-based methods model the degradation of LR images and reverse-solve this model to produce HR images. These methods typically outperform interpolation-based methods but necessitate detailed knowledge about image degradation. The third category, learning-based methods, applies machine learning to predict HR images from LR ones. This category includes various approaches such as dictionary learning-based SR, deep learning-based SR, Support Vector Machine (SVM)-based SR, and random-forest based SR. This dissertation primarily concentrates on learning-based SISR, specifically employing the dictionary-learning based SR with sparse representation algorithms.

This approach demonstrates the efficacy of using dictionary learning with appropriate sparse representation algorithms in producing images with improved performance, even in noisy environments. A custom modification has also been made to the dictionary learning-based method to enhance computational efficiency. Although dictionary learning methods such as k-SVD are adept at enhancing contrast in grayscale images, they are less effective with colour images, as training typically focuses on grayscale patches. Deep learning-based SR techniques, while producing high-quality colour HR images, face difficulties when processing unseen data and varying image types. Addressing this, the dissertation proposes a hybrid SR technique that capitalizes on the strengths of both approaches: the contrast enhancement of dictionary learning for grayscale images and the colour fidelity of deep learning methods. This research emphasizes the value of SR for both colour and grayscale images and seeks to broaden the applicability of SR techniques in diverse imaging fields, including medical imaging. A significant application of the proposed SR technique discussed in this dissertation is its use in the agricultural sector for detecting and classifying thrips pests. The proposed approach seeks to enhance image resolution and accuracy, facilitating more effective identification

and categorization of these pests. An end-to-end automatic counting system for farmers is proposed, incorporating the developed SR technique. This system aims to improve the precision of pest detection and classification, thereby streamlining and enhancing the efficiency of the monitoring process in agricultural practices.

Firstly, the dissertation first conducts a comparative analysis to assess the effectiveness of combining dictionary learning with sparse representation algorithms in producing enhanced denoised images. The performance of different algorithms integrating dictionary learning with various sparse representation methods, including the Douglas-Rachford algorithm, Soft Thresholding, and Orthogonal Matching Pursuit (OMP), is evaluated. The effectiveness of these algorithms is determined based on average Peak-Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Metric (SSIM) values, compared with images enhanced solely through sparse regularization methods like Gradient Descent, Newton Method, and Alternating Direction Method of Multipliers (ADMM). The findings demonstrate that the integration of dictionary learning with Douglas-Rachford based algorithm yields the highest average PSNR and SSIM values, indicating its superior efficacy in enhancing LR images.

Then, this dissertation introduces a SISR image reconstruction scheme, named KSVD_DR, which employs the Douglas-Rachford algorithm and k-Singular Value Decomposition (k-SVD) technique for improved computational efficiency. This system operates in two phases: training and testing. The training phase involves computing LR and HR dictionaries using KSVD_DR, while the testing phase uses these dictionaries to reconstruct HR images from LR image patches. The final HR image is assembled through a patch averaging operation. Performance evaluations indicate that the Douglas-Rachford algorithm leads to faster computation times in comparison to other methods including OMP, OMP, Group OMP (GOMP), Group MP (GMP), and Block MP (BMP). This also highlighting the effectiveness of KSVD_DR in the SISR image reconstruction process, especially in enhancing grayscale image resolution.

After that, a novel hybrid technique is developed, merging dictionary learning-based methods with transformer-based deep learning for post-SR image enhancement. This hybrid approach employs the KSVD_DR method for Pathway 1 and the Image Processing Transformer (IPT) for Pathway 2, ensuring high-quality image resolution for various applications, including both general colour and medical grayscale images. Comparative evaluations against leading dictionary learning and deep learning-based methods showcase the hybrid method's superiority in visual quality and quantitative performance metrics. Although the images produced by the proposed hybrid method narrowly trails by a marginal 0.2 in PSNR when compared to the latest state-of-the-art of deep learning-based SR, however, it demonstrates a better performance on grayscale images, even after finetuning the IPT model with grayscale images. Hence, the results highlighted the potential of the proposed hybrid method in SR task.

Furthermore, this dissertation extends the application of the KSVD_DR method to the agricultural

sector, particularly for the efficient detection and classification of thrips pests. By enhancing image resolution and accuracy, this method enables more effective identification and categorization of these pests. The validation results from the Yolov8 detection model to discern the impact of SR processing on the accuracy of thrips detection were examined. For detection and counting purpose, the validation results comparing the accuracy of thrips detection with and without the application of SR processing, using the metric mAP50 (mean Average Precision at 50% Intersection Over Union) was conducted. The accuracy for the 'Without SR' scenario is listed as 0.665, at the 50% IoU threshold. In contrast, the 'With SR' scenario demonstrates a significant improvement in accuracy, with the KSVD_DR method achieving 89.7% (0.897) and the hybrid method reaching 87.7% (0.877) in detecting thrips, underscoring the effectiveness of SR enhancement. The dissertation proposes an end-to-end automatic counting system for farmers, integrating the developed SR technique to streamline the monitoring process in agricultural practices. After that, for the classification of thrip, the validation results were evaluated by comparing the Top-1 accuracy of thrips classification with and without the application of SR processing. For the 'Without SR' scenario, the Top-1 accuracy is recorded as 0.585, indicating that the model correctly classifies thrips with 58.5% accuracy. Conversely, in the 'With SR' scenario, there is a notable improvement in accuracy: the KSVD_DR method achieves 65.3% (0.653) accuracy, while the hybrid method slightly trails at 64.2% (0.642), demonstrating the model's enhanced capability in classifying thrips with SR enhancement. These findings underscore the influence of SR processing on the model's ability to classify thrips more accurately.

However, this dissertation also acknowledges that its final images slightly lag deep-learning-based SR methods in terms of PSNR values for colour images. However, it excels in processing grayscale images, even surpassing the IPT model after fine-tuning. Future optimization efforts will focus on improving colour image processing. In terms of agricultural pest detection and classification, the current model is specifically tailored for blue traps, which limits its broader application. Future work includes expanding the dataset to encompass yellow traps to enhance the model's versatility and improving the classification model to address data imbalance issues. These enhancements aim to provide more accurate and comprehensive pest management solutions in agriculture.

In conclusion, this dissertation not only makes significant strides in advancing SR techniques but also demonstrates their practical application in agriculture field.

Keywords: *Single-Image Super Resolution*, *Dictionary Learning-based SR*, *Douglas-Rachford*, *Hybrid Method*

CONTENTS

1.   INTRODUCTION

1.1 Background

Image Super-Resolution (SR) is a technique aimed to refine a Low-Resolution (LR) image into its High-Resolution (HR) equivalent by extrapolating the omitted high-frequency components (Nasrollahi & Moeslund, 2014; Wang et al., 2015; Yue et al., 2016). An LR image has a few pixels, creating a blurry or pixelated appearance. It has a lower spatial resolution and less detail compared to HR images. LR images are often produced by digital cameras or other imaging devices with limited capabilities or when images are compressed for storage or transmission. It is essential to understand that an image's resolution must differ from its physical size to grasp the concept of SR imaging(Tian & Ma, 2011). In other words, image SR techniques are mainly used to produce images with more apparent content from its input LR image.

Although technological advancements have improved imaging devices, their high cost and substantial size often hinder the acquisition of HR images in various scenarios. In such instances, SR imaging offers a solution by compensating for the limitation of these imaging systems. It achieves this by enhancing and combining multiple images of the same scene to produce a single image with superior resolution, effectively addressing the challenges posed by less-than-ideal image capturing conditions (Tian & Ma, 2011). According to Milanfar explains that image resolution is about the amount of detail in an image; higher resolution means more pixels and clearer details (Milanfar, 2017).

Figure 1.1 shows the basic block diagram for SR imaging techniques, while Figure 1.2 shows a simple illustration of the block diagram. In another words, the progress in image processing has made SR imaging key for better visual clarify and extracting more information for recognition purposes. As a result, researchers in image processing are focusing on developing SR image reconstruction algorithms capable of producing high-quality SR images from a series of LR images of the same scene.
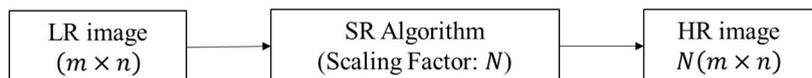

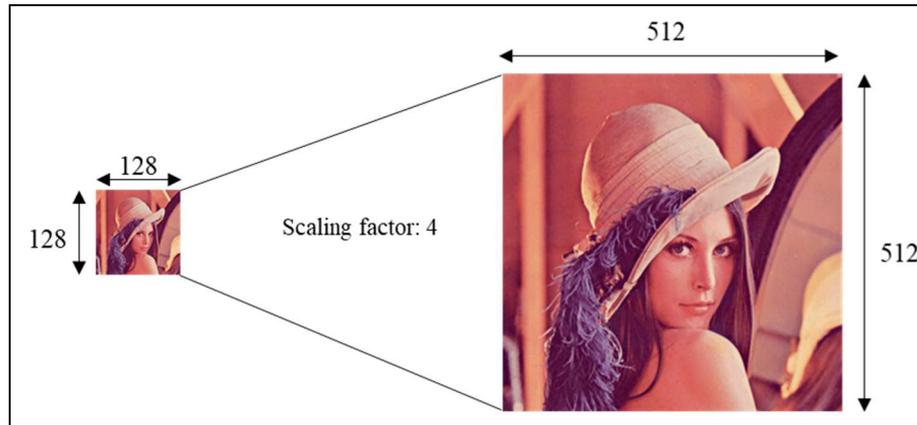**Figure 1.1: Basic block diagram for SR imaging techniques.**

Figure 1.2: Simple example of SR imaging.

The data-type-based categorization primarily includes Multiple Image Super-Resolution (MISR) and Single Image Super-Resolution (SISR). MISR is a technique that seeks to augment the resolution of a collection of LR images. It accomplishes this by merging multiple LR images to form a single HR image which is rich in details and minimal in noise or distortions. One of the complexities of MISR techniques is the need for a precise registration process. This procedure necessitates the alignment of multiple LR images with subpixel precision, which can be time-consuming. On the other hand, SISR is a process which uses only a single LR image to generate the HR counterpart. Since the SISR approach works on single LR image, and hence more suitable for a wide range of applications. However, SISR still presents challenges as it requires prediction or inference of more details from limited information. In essence, SISR aims to recreate high-quality images from their low-resolution counterparts, making it a difficult endeavour.

When it comes to the techniques applied in image SR, it can be classified into three main categories: interpolation-based, reconstruction-based, and learning-based methods. First, interpolation-based SR such as the nearest neighbour, bilinear and bicubic approaches are the simplest form of SR techniques, which involve the estimation of the high-frequency details by using nearby pixel values. However, significant high-frequency information is always missing in the target HR images, resulting in blurry images. Besides, reconstruction-based methods involve modelling the image degradation process and then inversely solving this model to reconstruct the HR image. Although these methods often offer better performance than interpolation-based methods, these methods need precise knowledge about the degradation of the LR image, which is not always available. Therefore, the reconstruction-based methods are unable to reconstruct effectively in some cases. After that, learning-based SR utilizes machine learning techniques to predict HR images from LR counterparts. For example, dictionary learning-based SR, deep learning-based SR, Support Vector Machine (SVM) based SR and random-forest based SR. In this dissertation, the focus is on the learning based SISR. A hybrid approach that merges the strengths of dictionary learning-based and transformer-

based deep learning methods for image enhancement after the SR process is developed. Besides, a custom modification has also been done in the dictionary learning-based method to improve the computation efficiency.

SR imaging has a significant impact across various sectors. In surveillance, it sharpens the CCTV footage, essential for security and criminal investigations. Additionally, in consumer electronics, such as smartphones and cameras, SR technology improves photo and video quality, allowing for clearer images. These diverse applications highlight SR's versatility and its ability to improve visual details in multiple areas. While in the medical field, SR imaging is revolutionizing diagnostics. By enhancing the details and clarity of medical images, SR imaging is crucial for precise diagnosis and effective treatment planning. SR techniques are highly advantageous in medical imaging, applicable not just to colour images but also to grayscale images. This includes the images from Computer Tomography (CT) scans, Magnetic Resonance Imaging (MRI), and X-ray Imaging, where SR can significantly enhance the image quality. Therefore, this dissertation not only focuses on improving the image quality of the general colour images, but also involves evaluations of the grayscale images, to make sure the general use of SR techniques proposed. Figure 1.3 and Figure 1.4 present sample images from the test datasets used in dissertation, illustrating examples of both colour and grayscale images that analysed.

Figure 1.3: Colour images from test dataset (Set14 (Zeyde et al., 2012), Set5 (Bevilacqua et al., 2012), Urban100 (Huang et al., 2015) and B100 (Martin et al., 2001)).

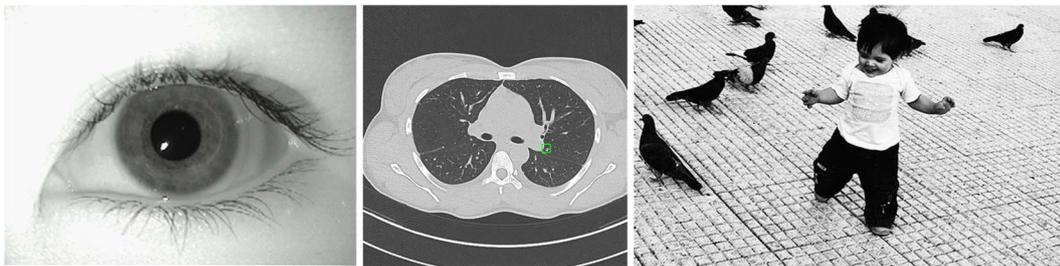**Figure 1.4: Grayscale images from test dataset (MMU Iris Dataset (*MultiMedia University Iris dataset for Biometric Attandance System*, 2020), NIH CT Scan (Yan et al., 2018), and MIRFLICKR (Huiskes & Lew, 2008)).**

Then, SR techniques can also be useful in the field of smart agriculture, especially for early detection and management of small pests, such as the thrips insects, which are difficult to identify with

LR images. In recent years, the damage caused by thrips species, particularly to citrus crops has been increasing, highlighting the need for effective control measures. Thrips pests damage the fruit of citrus plants, with their activity spanning from May, during young fruit phase, to November, the coloration period. Adult insects originating from surrounding fields fly in and damage the fruit surface, leading to deterioration in appearance and a decrease in marketable yield as shown in Figure 1.5. Therefore, it is crucial to predict and estimate the occurrence of thrips species in advance. In this dissertation, the application of the SR technique proposed will be applied for the detection and classification of thrips pests. This approach aims to enhance image resolution and accuracy, facilitating more effective identification and categorization of these agricultural pests.

ネーブル果実 レモン果実

Figure 1.5: Damage on the citrus and lemon fruit surface caused by thrips pets (アザミウマ類発生予察について, 2022).

To analyse the presence of thrips pests in fruit orchards, farmers implement a methodical approach by placing the coloured traps as shown in Figure 1.6 within their farms. Yellow traps are used extensively in citrus orchards that suffer from *Yellow Tea Thrips* widespread impact. Conversely, in lemon orchard frequently afflicted by *Flower Thrips*, blue traps are preferred and will be introduced in the planting season, leveraging the pest's known attraction to the blue colour traps. Following the placement of these traps, farmers conduct a manual count of the thrips that are captured. This approach facilitates the tracking and measurement of pest populations, but it is time intensive. Therefore, this dissertation proposes the development of an end-to-end automatic counting system for farmers. This system will incorporate the proposed SR technique to enhance the precision of pest detection, aiming to streamline and improve the efficiency of the monitoring process.
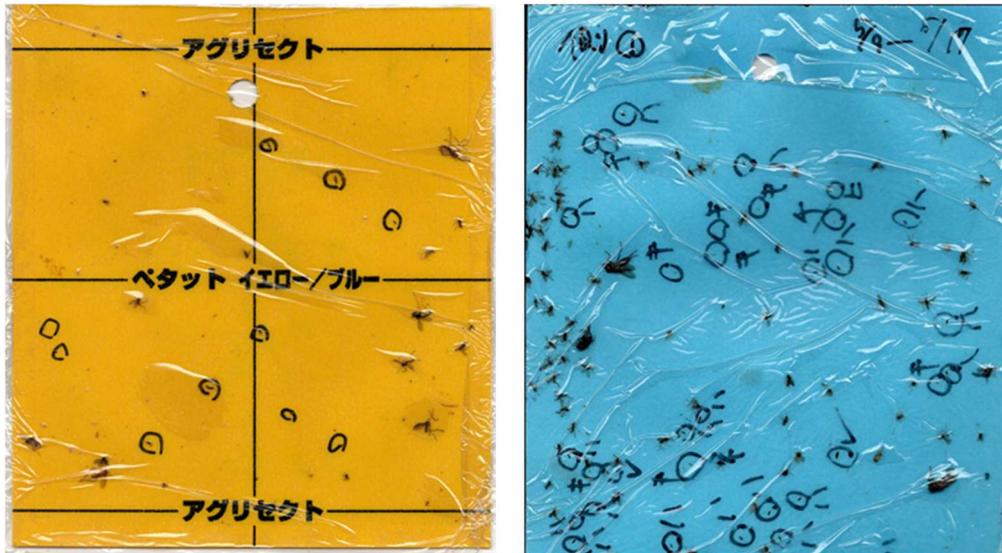
Figure 1.6: Examples of yellow and blue traps.

In addition, classifying various types of thrips is also part of the analysis process, as it helps in understanding the specific pest dynamics and in tailoring pest management strategies accordingly. As depicted in Figure 1.7, there are five distinct types of thrips that affect fruit orchards each with unique characteristics based on the size and colors.
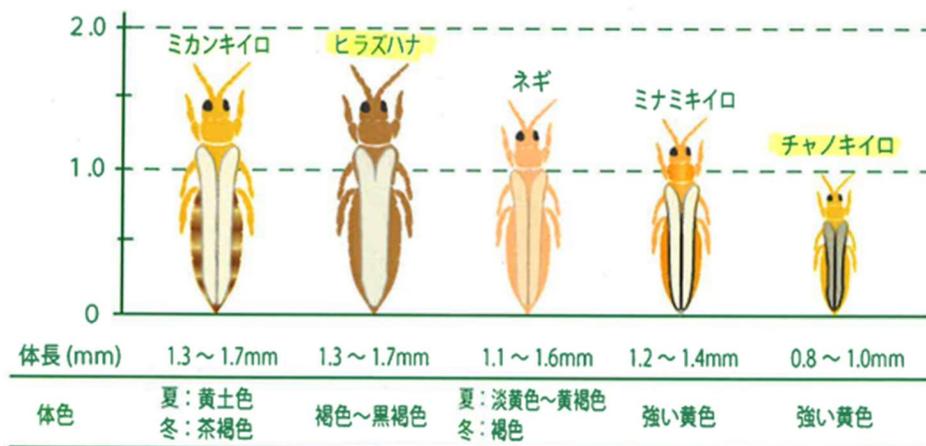
Figure 1.7: Classification of Thrips based on their size and colours (アザミウマ 5 種, 2022).

Recognizing and distinguishing between these types is important because it determines the most effective control measures. The farmers provided the ground truth for this classification, marking various classes of thrips on the blue trap as shown in Figure 1.8. As illustrated in Figure 1.8 and Table 1.1, six distinct types of thrips including *Flower Thrips*, *Honeysuckle Thrips*, *Western Flower Thrips*, *Yellow Tea Thrips*, *Onion Thrips* and others have been identified and labeled by the farmers. In this

dissertation, a classification framework is introduced with the implementation of proposed SR technique to assist farmers in identifying different types of thrips.
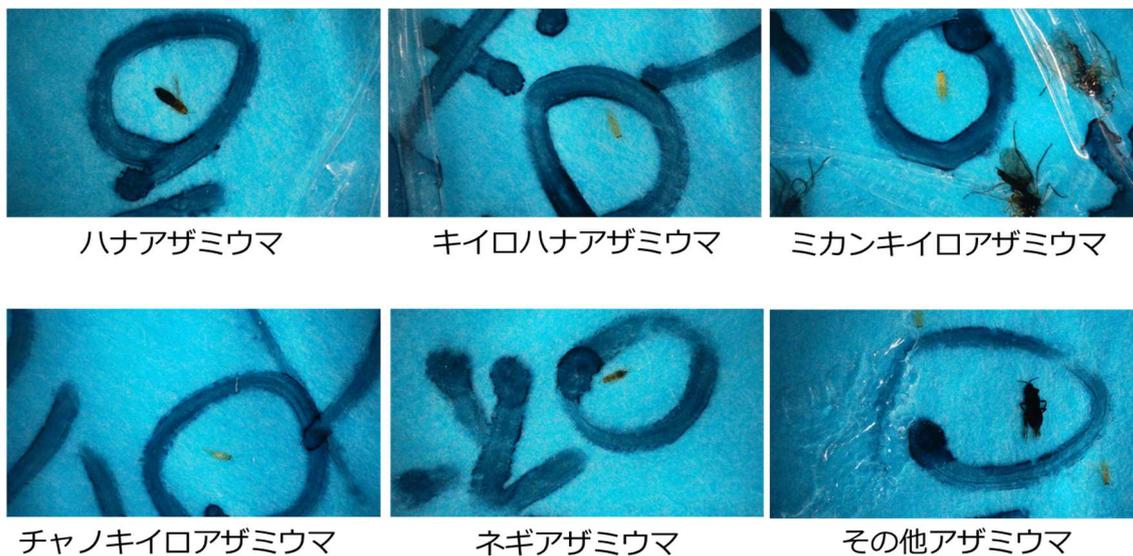
Figure 1.8: Different classes of Thrips captured using Jiusion HD USB Microscope from blue trap.

Table 1.1: Classes of Thrips.

Class	Labels by Farmers	Types of Thrips (Japanese)	Types of Thrips
1	ハ	ハナアザミウマ	*Flower Thrips*
2	キ	キイロハナアザミウマ	*Honeysuckle Thrips*
3	ミ	ミカンキイロアザミウマ	*Western Flower Thrips*
4	チ	チャノキイロアザミウマ	*Yellow Tea Thrips*
5	ネ	ネギアザミウマ	*Onion Thrips*
6	その他	その他アザミウマ	Other Thrips

1.2 Contributions

To summarize, this dissertation introduces four main contributions:

1. An analysis comparing the effectiveness of combining dictionary learning with sparse representation algorithms to produce enhanced denoised images.

2. A SISR image reconstruction scheme designed for LR images, utilizing the Douglas-Rachford algorithm and the k-Singular Value Decomposition (k-SVD) technique, referred to as KSVD_DR, for improved computational efficiency.

3. A hybrid image reconstruction algorithm for enhancing LR images, applicable to both general colour and medical grayscale images.

4. The application of SR for the detection and classification of thrips in agriculture.

The first contribution of this dissertation is a comparative analysis focused on the efficacy of combining dictionary learning steps with sparse representation algorithms to produce better denoised images. This analysis involved evaluating the performance of three different algorithms, each integrating dictionary learning with distinct sparse representation methods: Douglas-Rachford (Algorithm 1), Soft Thresholding (Algorithm 2), and Orthogonal Matching Pursuit (OMP) (Algorithm 3). The effectiveness of these algorithms was measured using the average Peak-Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Metric (SSIM) values of the denoised images produced. These values were then compared to those obtained from images enhanced solely through sparse regularization methods such as Gradient Descent, Newton Method, and Alternating Direction Method of Multipliers (ADMM). The findings revealed that the denoised images generated by Algorithm 1, which combines dictionary learning with the Douglas-Rachford method, achieved the highest average PSNR and SSIM values. This outcome demonstrates the superior capability of Algorithm 1 in enhancing LR images, thereby validating the effectiveness of incorporating dictionary learning processes with sparse representation methods in image SR tasks.

The second key contribution of this dissertation is the development of a SISR system based on dictionary learning with a sparse representation method. This system employs the Douglas-Rachford algorithm and k-Singular Value Decomposition (k-SVD) technique, named KSVD_DR, to optimize computational efficiency. The SISR system functions in two phases: training and testing. In the training phase, LR and HR dictionaries were computed using KSVD_DR. During the testing phase, these dictionaries were used to reconstruct HR images from LR image patches. The process involves computing sparse representation coefficients of the LR patches using the Douglas-Rachford algorithm and then reconstructing the HR patches using these coefficients and the HR dictionary. The final HR image was assembled through a patch averaging operation. Performance evaluations of the k-SVD algorithm, particularly against methods like Orthogonal Matching Pursuit (OMP), Group OMP (GOMP), and Block MP (BMP), demonstrate that the Douglas-Rachford algorithm leads to faster computation times. This result underscores the effectiveness of the proposed SISR image

reconstruction process, showcasing the method's capability to enhance image resolution, particularly in grayscale, beyond that achieved by state-of-the-art SR methods such as Image Processing Transformer (IPT).

The third contribution involves developing a novel hybrid technique that combines dictionary learning-based methods with transformer-based deep learning for post-SR image enhancement. The KSVD_DR method produces images for Pathway 1 of the hybrid system, while Pathway 2 uses the IPT by Chen *et al.*, chosen for its flexibility and ability to focus on image areas potentially overlooked by KSVD_DR. This approach ensures the final HR images are suitable for diverse applications, from general colour imagery to the specific requirements of grayscale medical imaging. Comparative evaluations against leading dictionary learning and deep learning-based methods, conducted on both colour and grayscale images, demonstrate that this hybrid method excels in visual quality and quantitative performance metrics.

For the last contribution, this research applies the modified dictionary-based SR method, KSVD_DR to real-world agricultural challenges due to its advantages include not needing large image datasets and having shorter training periods, due to its simpler structural design. It specifically addresses the need for efficient detection and classification of thrips in farming. This application proves the practical utility of the research, offering a viable solution to the critical issue of pest management in agriculture.

In conclusion, this dissertation demonstrates advancements in the field of image SISR through its four key contributions. It establishes the effectiveness of combining dictionary learning with sparse representation algorithms in improving image denoising, as evidenced by the superior performance of the combination of dictionary learning with Douglas-Rachford algorithm. The development of the dictionary learning based SISR system, named KSVD_DR marks a significant stride in optimizing computational efficiency without compromising image quality. The novel hybrid technique brings together the strength of dictionary learning-based methods with transformer-based deep learning paves the way for versatile image enhancement, catering to both colour and grayscale images. Lastly, the practical application of the KSVD_DR method in agricultural pest detection and classification underlines the real-world impact of this research, addressing critical challenges in pest management with a technologically advanced solution. Collectively, these contributions not only push the boundaries of image processing technology but also demonstrate the potential for such innovations to solve practical problems, particularly in the agricultural sector.

1.3 Structure of the Dissertation

This dissertation is structured into four chapters, each focusing on a distinct element of the study. The first chapter introduces the research, providing a background and outlining the key contributions. Chapter 2 delves into a comprehensive review of existing studies on SR techniques and analyses the HR image outputs. Chapter 3 details the dissertation's first major contribution, which involves analysing the effectiveness of integrating dictionary learning methods with sparse representation algorithms to enhance image denoising. Chapter 4 then discusses the second contribution, presenting a SISR system developed using the Douglas-Rachford algorithm and k-SVD, referred to as KSVD_DR. This chapter focuses on the computational efficiency of KSVD_DR compared to other methods. Additionally, Chapter 5 highlights the third contribution, which is the development of a hybrid image reconstruction method for LR images. This chapter will present the methodology adopted, including the selection of datasets, the experimental setup, and the algorithm used to produce HR images. The results which include a detailed analysis and discussion of the HR images produced will also be written in this chapter. After that, Chapter 6 presents the methodology adopted for the final contribution of this research. The results of the detection and classification which contributes to the analysis of pests for farmers will also be presented in this chapter. Lastly, Chapter 7 concludes this dissertation, discussing the limitations, and introduces future work.

2.   RELATED WORKS

2.1 Super-Resolution (SR) Techniques

Image Super-Resolution (SR) is a technique that involves creating one or multiple High-Resolution (HR) images from their respective Low-Resolution (LR) observations. Over the years, researchers have extensively studied SR technologies due to their practical applications in various fields, including medical image processing (Jiang et al., 2018; Malczewski, 2020; Pham et al., 2019), satellite and aerial imaging, infrared imaging, text image enhancement, video enhancement and others.

In the pursuit of producing HR images, researchers can opt for either hardware enhancements or software methods. One example of a hardware enhancement is the technique described by Ji *et al*., the process proposed in their paper involves using adaptive optics and pupil segmentation to improve image resolution in biological tissues (Ji et al., 2010). On the other hand, software methods can be a more cost-effective choice that does not require hardware upgrades. Examples of software methods for producing HR images include interpolation-based methods, Fourier-based methods, maximum likelihood estimation methods, and iterative restoration methods. However, these software methods can still be computationally expensive or slow, and the choice between hardware and software methods depends on the specific imaging needs and constraints.

Although hardware enhancements can produce HR images with faster computational time by reducing the need for extensive software processing, they also have practical limitations (Pandey & Ghanekar, 2018). Increasing the chip size to improve image resolution is expensive and can generate high capacitance, reducing the charge transfer rate. Conversely, decreasing the sensor size can limit the amount of light available to each sensor, resulting in shot noise and making the hardware more susceptible to diffraction phenomena. Due to these issues, software techniques such as SR methods have become a popular alternative to obtaining HR images from one or multiple LR input images. These SR methods can be computationally intensive but offer a more cost-effective and practical solution for producing HR images in various applications.

In 2014, researchers published a comprehensive survey that reviewed most of the papers on proposed SR methods up to 2012 and provided an overview of the advantages and disadvantages of these models (Nasrollahi & Moeslund, 2014). Similarly, Tai and Ma reviewed SR image and video reconstruction methods, highlighting future challenges in SR research, such as multi-view SR imaging and temporal SR video (Tian & Ma, 2011). Yue *et al.* also described commonly used regularised SR techniques and summarised current applications of SR models (Yue et al., 2016). These reviews offer insight into SR methods and highlight the importance of HR imaging in many modern applications

that require accurate results.

2.1.1 Image Degradation Model

The image degradation model mathematically represents how an image undergoes degradation during imaging. The image degradation model is critical in implementing an image SR reconstruction algorithm. It illustrates obtaining a degraded observed image from a real-life imaging application. In SR models, the goal is to reconstruct an HR image from one or more LR images. This process can be challenging because the LR images contain incomplete information about the original scene. The image degradation model provides essential information about the nature of this incompleteness, allowing the SR model to effectively fill in the missing details and produce a more accurate HR image. By incorporating the image degradation model into SR models, researchers can tailor the algorithm to account for the specific type and level of degradation in the LR images, resulting in a more effective and accurate SR process. This results in a more effective and precise SR process. Additionally, the image degradation model can be used to evaluate the performance of SR algorithms by comparing the reconstructed HR image to the ground truth HR image. Therefore, one of the primary tasks in developing SR algorithms is formulating an exact image degradation model.

Figure 2.1 illustrates a basic image degradation model researchers can use to address image super-resolution (SR) problems in current imaging systems. As shown in the figure, the main idea of SR algorithms is to reverse the degradation process and reconstruct a high-resolution (HR) image. Single Image Super-Resolution (SISR) and Multiple Image Super-Resolution (MISR) use the same image degradation model, the only difference being that MISR obtains more than one observed LR image.
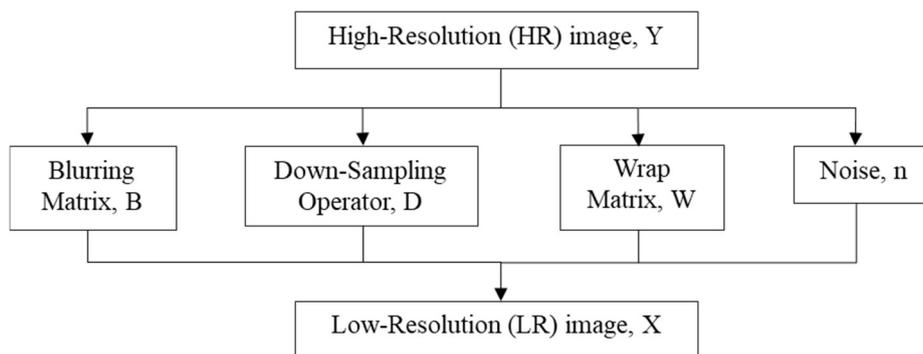
Figure 2.1: Basic image degradation model.

Figure 2.2 depicts the images generated at each step of the degradation factors from an original HR image using the standard test image, "*cameramen*." Generally, the observed images obtained from the existing imaging applications undergo degradation through blurring, warping, down-sampling, and adding white Gaussian noise.
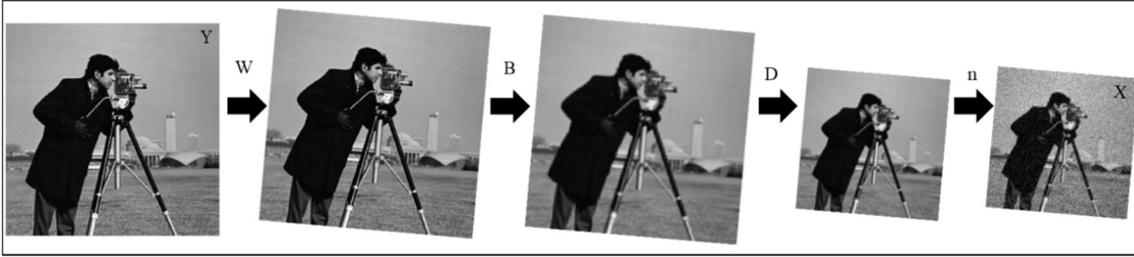
Figure 2.2: Illustration of the process in the imaging degradation model.

The process of generating a practical LR image, Y, from an HR image, X, can be mathematically defined using Equation (2.1), where B represents the blurring matrix, D is the down-sampling operator, W is the wrap matrix, and n represents the white Gaussian noise. However, to reduce the complexity of solving SR problems, researchers usually assume a noise-free process and define the blurring matrix, B as the identity matrix.

$$Y = BDWX + n \qquad (2.1)$$

2.1.2   Taxonomy of Super-Resolution Techniques

Solving SR problems is a highly ill-posed task, as the details captured in LR observations are often insufficient to uniquely reconstruct the desired HR image, as noted by Kien Nguyen *et al.* (Nguyen et al., 2018). In 1996, M. Elad and A. Feuer tackled this problem by proposing a hybrid reconstruction algorithm that combined the benefits of the Maximum Likelihood Estimator (ML) and the Projection Onto Convex Sets (POCS) techniques (Elad & Feuer, 1996). Their work has inspired other researchers to investigate new SR models and designs for addressing LR imaging problems.

Huang *et al.* proposed a self-similarity-based image SR algorithm that uses transformed self-exemplars, but their algorithm had drawbacks in recovering consistent structures and had slow computation time (Huang et al., 2015). Cai *et al.* developed an SR model trained on standard datasets to generalise to real-world images to address these issues. They constructed a real-world SR (RealSR) dataset with the real-world dataset to increase the accuracy of SR algorithms, mainly when applied to real-world LR images (Cai et al., 2019). Developing a real-world dataset is essential for improving the accuracy of proposed SR algorithms, mainly when applied to real-world LR images.

Chatterjee *et al.* identified three main tasks in solving SR problems; alias-free up-sampling of images to increase maximum spatial frequency, removing degradations such as blurring and noise from captured images, and registering and fusing multiple samples (Chatterjee et al., 2009). Figure 2.3 categorises SR algorithms based on factors such as the operational domain, number of input images, and reconstruction methods. The highlighted portion indicates that the algorithm proposed in this research focuses on the learning-based approach in SR methods.
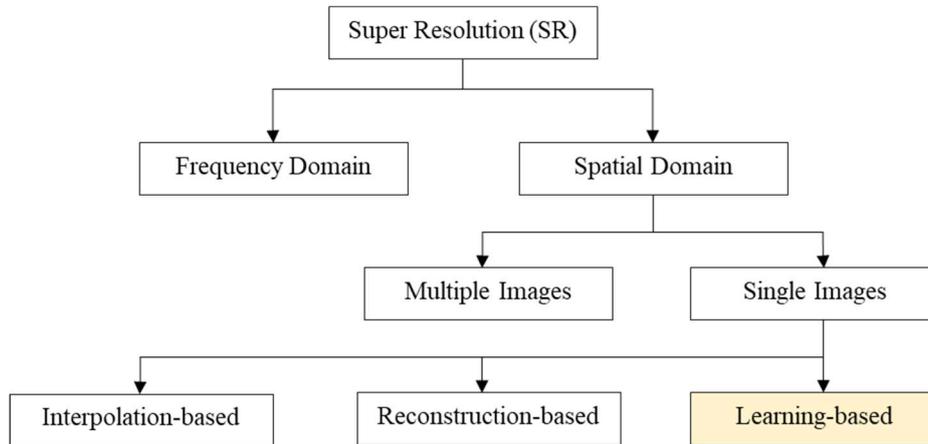
Figure 2.3: Categorisation of SR approaches.

Figure 2.3 in this taxonomy illustrates the division of SR techniques based on their operational domain, the frequency and spatial domain. In the spatial domain-based, the SR algorithms operate on the pixel intensities of an image. In contrast, in the frequency domain, the algorithms operate on the Fourier transform of the picture. Frequency domain-based SR methods can achieve good results for specific images, such as images with repetitive patterns or textures. Still, they have limitations when handling complex image structures and spatially variant distortions. Additionally, spatial domain-based SR methods can incorporate prior knowledge about the underlying image structure, such as edges and textures, leading to more accurate and visually pleasing SR results.

As shown in Figure 2.4, approaches in the spatial field perform SR on the pixel intensity values of an image. In contrast, frequency-domain SR approaches capitalise on the aliasing present in LR images. SR algorithms initially emerged from processing techniques in the frequency domain due to their low computational complexity compared to the spatial domain (Nasrollahi & Moeslund, 2014). However, most SR algorithms have since been implemented in the spatial domain-based because frequency domain-based SR methods have limited flexibility in the observation model, making them insufficient for dealing with real-world application (Tian & Ma, 2011).

Figure 2.4: Difference between spatial and frequency domain of an image.

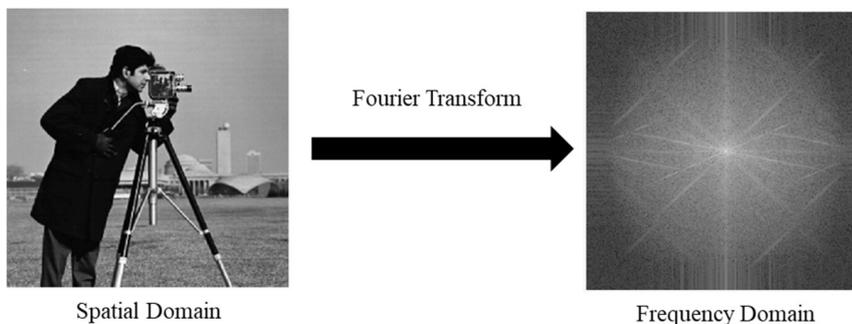SR techniques in the spatial domain can be classified based on the number of input images into two types: Single Image Super-Resolution (SISR) and Multiple Image Super-Resolution (MISR). In MISR, multiple input images are used, which can be captured simultaneously using the same camera from other cameras or angles (Park et al., 2003). The main task in MISR is to estimate the controlled or uncontrolled motions in the imaging system, which are used for merging the sub-pixel shifting of LR images on an HR grid for SR reconstruction. However, the motion estimation process is volatile and complex in MISR, as objects in practical applications presented in the same frame could differ in motion and direction. Therefore, SISR is used to overcome this limitation of MISR. After that, it is also important to note that MISR and SISR are not mutually exclusive and can be combined in some cases. For instance, a MISR approach can be used to align multiple input LR images, and then a SISR approach can be applied to each aligned image individually. Recent advancements in deep learning-based SR methods have shown that both MISR and SISR can benefit from deep convolutional neural networks.

SR reconstruction algorithms for SISR typically use one LR input image to produce the HR image. Based on the SR reconstruction process, SISR can be classified into three approaches: interpolation-based, reconstruction-based, and learning-based, as discussed in this research while implementing the SR reconstruction algorithm (Hui et al., 2020). Interpolation-based SISR approaches use neighbourhood pixels to define missing pixels and produce an HR image from an LR input image. The basic interpolation based SISR approaches include a nearest neighbour, bilinear, and bicubic interpolation methods (Lehmann et al., 1999; Thévenaz et al., 2000). The bicubic interpolation produces sharper results than the bilinear interpolation and is the most commonly used (Pandey & Ghanekar, 2018). However, the interpolation-based approaches are faster and less accurate.

Reconstruction-based SISR approaches can be applied to overcome the limitations of interpolation-based methods. These approaches recover the lost high-frequency components by finding them in multiple LR images, producing accurate results when rich and complementary details are present. Reconstruction-based methods face performance degradation, particularly when the upscale factor increases. Recently, learning-based SISR approaches, such as sparse coding, neighbour embedding, regression, and deep network-based methods, have been commonly used by SR models. These approaches study the relationship between the LR and HR feature spaces through training and find the lost high-frequency components from the training dataset (Ayas & Ekinci, 2020; Chang et al., 2004; Timofte et al., 2013; Yang et al., 2008; Yang et al., 2010; Zeyde et al., 2012).

In this dissertation, a learning-based SR image reconstruction algorithm is proposed, aiming to enhance SR performance. Specifically, dictionary learning, and sparse representation techniques are employed in one of our pathways in our design architecture. Dictionary learning facilitates the acquisition of an overcomplete set of basic functions from the LR and HR training image pairs. At the same time, sparse representation exploits the sparsity of image signals to reconstruct the HR image

from the LR input. The proposed algorithm is then evaluated by comparing its results with those obtained using the commonly employed SISR method, bicubic interpolation.

Overall, using learning based SISR approaches, including dictionary learning and sparse representation techniques, offers promising prospects for improving the quality of SR reconstructions. The references provided can provide more detailed information on the specific aspects and methodologies of the learning based SISR algorithms using dictionary learning and sparse representation (Ayas & Ekinci, 2020).

2.1.3   Dictionary Learning and Sparse Representation Algorithms

The implementation of image SR methods has drawn the attention of researchers to dictionary learning and sparse representation algorithms. Generally, dictionary learning, also known as sparse coding, is a technique in machine learning where the goal is to find a set of 'atoms' or 'dictionary elements' that can be linearly combined to represent a given set of data (Aharon et al., 2006). Furthermore, sparse representation is a technique where data is represented as a linear combination of basic functions (Zhang et al., 2015). These basis functions are only a few from a large set, and hence the representation is 'sparse'.

In this case, Zhang *et al.* conducted a comprehensive study and update review, exploring the motivations, theories, and applications of different sparse representation methods (Zhang et al., 2015). The authors concluded that developing an effective sparse representation method remains a significant challenge. Furthermore, they emphasised the importance of designing an efficient dictionary, which is expected to enhance the performance of the proposed methods in future applications.

In 2006, M. Elad and M. Aharon introduced a straightforward approach for image denoising, which relies on sparse and redundant representation using dictionaries trained with the k-Singular Value Decomposition (k-SVD) (Elad & Aharon, 2006). The details of the k-SVD algorithm are also explained in detail by Aharon *et al.* (Aharon et al., 2006). Building upon this research, Rubinstein *et al.* proposed a new algorithm called Sparse k-SVD (S-KSVD) for dictionary training (Rubinstein et al., 2009). These studies demonstrated the effectiveness of the k-SVD algorithm in training dictionaries for developing SR methods based on sparse representation. In this research, the k-SVD algorithm is an essential reference during the training phase design. Figure 2.5 illustrates the k-SVD algorithm, which consists of three main steps: dictionary initialisation, sparse coding methods, and dictionary update. The dictionary is initially initialised during training, followed by applying sparse coding techniques. The dictionary is then updated to generate the final dictionary.
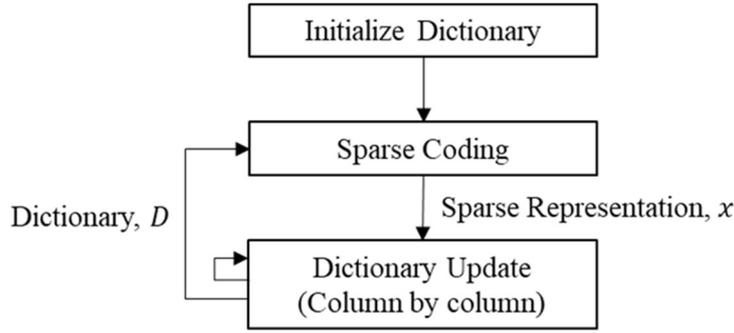
Figure 2.5: General algorithm outline for training the dictionary using k-SVD.

Sparse reconstruction optimisation problems can also be solved by using the regularisation methods, and the greedy algorithms. The regularisation methods include the Gradient Descent method, Newton method, and Alternating Direction Method of Multipliers (ADMM). In practice, the sparse coding step often employs greedy algorithms such as Orthogonal Matching Pursuit (OMP). Theoretically, the main idea of the greedy strategy is to determine the position according to the relationship between the atom and probe sample and then evaluate the amplitude value using the least square. The greedy algorithm used most frequently by the researchers in the process for the k-SVD algorithm is OMP (Ayas & Ekinci, 2020; Elad & Aharon, 2006; Timofte et al., 2013; Zeyde et al., 2012). The OMP algorithm, as described in (Pati et al., 1993; Tropp & Gilbert, 2007), is known as the improved version of the MP algorithm. The MP algorithm is mainly applied to find the best matching atom from the dictionary for constructing the sparse approximation in each iteration, computing the signal representation residual, and lastly, choosing the best matching atom until reached the stopping criterion of the iteration (Zhang et al., 2015).

On the other hand, the OMP algorithm will guarantee the orthogonal direction of projection in each iteration by using the orthogonalization process (Tropp & Gilbert, 2007). Angshul Majumdar also proposed the group and block sparse signal reconstruction via MP, such as the Block Matching Pursuit (BMP), Group Matching Pursuit (GMP), and Group Orthogonal Matching Pursuit (GOMP) (Majumdar, 2009). Among these methods, GOMP specifically employs a selection criterion based on the highest average correlation within each group. Then, Li and Qi put forth the non-local Douglas-Rachford (NLDR) algorithm for solving the CS image recovery problem (Li & Qi, 2015). Their research demonstrates the utility of the Douglas-Rachford algorithm for sparse coding in the context of image denoising. These studies collectively contribute to the advancement of SR techniques by harnessing the power of dictionary learning, sparse representation, and innovative algorithmic frameworks to enhance image resolution and quality.

In this dissertation, the emphasis is on assessing the effectiveness of combining dictionary learning with sparse representation algorithms. This involves comparing the k-SVD method integrated with Douglas Rachford, soft thresholding, and OMP against sparse regularization methods such as the

Gradient Descent method, Newton method, and ADMM. The approach of combining dictionary learning with sparse representation is contrasted with the implementation of sparse regularization methods that do not incorporate dictionary learning. The results show that the method combining k-SVD with Douglas Rachford offers the most advantageous outcomes for enhancing LR images with noise.

Furthermore, although typically a greedy algorithm replaces the sparse coding method, it is suggested to use the Douglas-Rachford algorithm as the sparse coding method in k-SVD algorithm, for its enhanced noise management and effectiveness during training. The proposed KSVD_DR will undergo comparison with these methods in terms of computational time. This comparison aims to demonstrate that the effectiveness of and efficiency of the KSVD_DR approach, highlighting its advantages in processing speed without compromising image quality.

Many researchers have explored the implementation of SR reconstruction schemes based on dictionary learning with sparse representation algorithms to generate HR images. Yang *et al.* proposed a sparse coding based SR method (ScSR) with a global dictionary learned from both LR and HR image patches (Yang et al., 2010). Zeyde *et al.* proposed a simplified SISR process using sparse representation, improving computational complexity and algorithm architecture by employing a different training process to generate trained dictionary pairs. Their approach also enabled running without a training set by boot-strapping the up-sampling task from observed LR images (Zeyde et al., 2012). Yang *et al.* introduced an image SISR technique through sparse representation, using a jointly trained coupled dictionary from HR and LR image patches. Their method handles patch-based SR with noisy inputs for generic and face images (Yang et al., 2010). Additionally, Timofte *et al.* presented the Anchored Neighbourhood Regression (ANR) method, which fixes the Neighbourhood Embedding (NE) of an LR image patch to the nearest atom in the dictionary and precomputes the corresponding embedding matrix (Timofte et al., 2013). These advancements in dictionary learning and sparse representation algorithms have improved the quality and performance of SR reconstruction for HR image generation. Ayas and Ekinci also proposed to use a novel multi-scale directional feature descriptor approach to improve the quality of HR image. While current methods for dictionary learning, especially methods that use the k-SVD with different pursuit method in the training phase, is still limited by lengthy training times (Ayas & Ekinci, 2020). Despite advancements in dictionary learning methods such as k-SVD combined with various pursuit algorithms, they have limitations in processing colour information, as they are typically trained on grayscale patches. Therefore, a hybrid method is proposed to improve the limitation of dictionary learning-based methods.

2.1.4   Deep Learning Algorithms

In recent years, deep learning has emerged as a powerful paradigm for tackling various computer vision tasks, including SISR. Deep learning models, particularly CNNs, have demonstrated

remarkable performance improvements in generating HR images from LR inputs. By leveraging their ability to learn complex mappings between LR and HR image spaces, deep learning based SISR algorithms have pushed the boundaries of possible image quality and detail restoration. This subsection explores the application of deep learning techniques in SISR, examining the key components, network architectures, and training strategies that have driven advancements in the field. Notable achievements and practical approaches are highlighted, providing insight into the profound impact of deep learning on pushing the boundaries of SISR research.

Pham *et al.* proposed a deep three-dimensional CNN approach for multiscale brain Medical Resonance Imaging (MRI) SR. Their method enables the reconstruction of LR images using an HR reference image from the same patient (Pham et al., 2019). Similarly, Jiang *et al.* developed a lightweight CNN network structure for SR applications using a multi-view information fusion strategy (Jiang et al., 2020). These authors' work highlights the significant advancements made in the field of SR imaging through the utilisation of CNN-based models. One challenge that CNN-based SR methods face is the large model size and significant graphic memory consumption. Hui *et al.* proposed a lightweight image SR approach called Feature Enhancement Residual Network (FERN) to address this issue to achieve accurate SR results (Hui et al., 2020). Additionally, most existing CNN-based SISR methods assume that an LR image is down-sampled from an HR image. However, in cases where the actual degradation does not adhere to this assumption, the performance of these models may suffer. To tackle this problem, Zhang *et al.* proposed an SR network capable of handling multiple degradations, leading to improved practicality in the image SR process (K. Zhang et al., 2018).

Kin *et al*. presented an accurate SR method using deep convolutional networks named Very Deep SR (VDSR) (Kim et al., 2016), significantly enhancing image enhancement. The researchers focused on developing deep learning algorithms for producing high-quality images from LR inputs. Then, Lim *et al*. proposed an Enhanced Deep Residual Network for SISR (EDSR) (Lim et al., 2017). EDSR is a deep learning model that employs advanced residual learning techniques to improve the quality of unsampled images. Zhang *et al*. also proposed an SR method in 2018, utilising very deep Residual Channel Attention Networks (RCAN) (Y. Zhang et al., 2018). This model integrates attention mechanisms in deep residual networks to enhance the model's focus on relevant image features. Dai *et al.* developed the Second-order Attention Network (SAN) for SISR in 2019 (Dai et al., 2019). SAN employs a second-order attention mechanism to enhance the model's ability to focus on essential image features. He *et al.* offered a novel approach to image SR in 2019. They designed an ODE-inspired network to enhance the quality of single images (OISR-RK3) (He et al., 2019). They designed an ODE-inspired network to enhance the quality of single images, and this methodology leverages the principles of ordinary differential equations to optimise network architecture. Zhang *et al.* advanced the field of image restoration in 2019 by proposing Residual Non-Local Attention Networks (RNAN) (Zhang et al., 2019). This approach enhances image quality by integrating non-local attention

mechanisms into deep residual networks.

After that, SRCNN which stands for Super-Resolution Convolutional Neural Network, is an innovative algorithm that utilizes a deep convolutional neural network to enhance the resolution of images (Vb, 2020). It employs a layered architecture to learn how to upscale LR images into HR counterparts effectively and provides a reference for researchers in the field of implementing SR techniques using CNN. Niu *et al.* introduced a SISR method using a Holistic Attention Network (HAN) (Niu et al., 2020). HAN integrates holistic attention mechanisms into deep learning models, enabling superior image enhancement. Zhou *et al.* also proposed the cross-scale Internal Graph Neural Network (IGNN) for image SR in 2020 (Zhou et al., 2020). This model leverages graph neural networks to exploit the internal correlations of an image across different scales, leading to improved SR results. Then, Zhang *et al.* introduced the Residual Dense Network (RDN) (Zhang et al., 2020). RDN is a deep learning-based model designed to enhance the quality of images by employing densely connected convolutional layers, resulting in improved performance in image SR tasks. In this dissertation, all these methods, representing significant advancements in SR techniques, will be considered as the state-of-the-art benchmark for comparison.

In the field of SR, various methods have been developed to address the challenge of out-of-distribution scaling factors. MetaSR pioneered the concept of an arbitrary-scale meta-upscale module for SR task (Hu et al., 2019). Following this, several other arbitrary-scale methods were introduced to enhance performance. One such method is the Local Image Function (LIIF), which approaches images as continuous functions rather than discrete 2D pixel arrays, effectively creating a single, efficient model for SR tasks (Y. Chen et al., 2021). Building on this, the Local Texture Estimator (LTE) was developed to analyse image textures in the Fourier space (Lee & Jin, 2022). Wei and Zhang further extended this field with the introduction of the SR Neural Operator (SRNO) for continuous SR (Wei & Zhang, 2023). In their work, Wei and Zhang used the RDN as a baseline, evaluating various models including RDN-MetaSR, RND-LIIF, RDN-LTE, and RDN-SRNO. Since this dissertation does not focus on out-of-distribution scales, it will only compare the in-distribution results from their paper. Additionally, a Dual-Domain Learning (DDL) approach was proposed, utilizing the Enhanced Deep Super-Resolution (EDSR) as a base model, and extending it to DDL-EDSR. This DDL approach, especially when combined with Bayesian methods, allows for the quantification of spectral uncertainty in the frequency domain (Liu et al., 2023).

The transformer architecture, initially introduced for natural language processing, has recently gained significant attention in image processing, including SISR. With its ability to capture long-range dependencies and model complex relationships, the transformer has shown promising results in various computer vision tasks. In the context of SISR, the transformer's capability to learn global image representations and handle non-local relationships makes it an intriguing choice for enhancing image resolution. First, Chen *et al.* proposed a pre-trained Image Processing Transformer (IPT) that

can effectively handle various image processing tasks, including denoising, super-resolution, and JPEG artefact reduction (H. Chen et al., 2021). By leveraging the power of transformer-based architectures, the IPT achieves competitive performance on these tasks without requiring task-specific fine-tuning. The trained IPT model demonstrates the potential of transformers in image processing and provides a valuable tool for improving image quality and enhancing various visual applications. The transformer framework in IPT is based on the novel approach for object detection using transformer-based architectures presented by Carion *et al.* (Carion et al., 2020). The authors propose an end-to-end framework that replaces conventional object detection components with transformer models, such as region proposal networks and non-maximum suppression. This work highlights the potential of transformers in revolutionising object detection pipelines, which inspired the use of the transformer model in other tasks in computer vision, such as the SR. Although the deep learning-based SR techniques show superior results in producing HR images, however, as compared with the dictionary learning-based SR methods, deep learning-based SR methods may not generalize well to unseen data or different kinds of images. Hence, a hybrid approach, merging the strengths of both dictionary learning-based and deep learning-based super-resolution (SR) techniques is proposed in this dissertation.

2.1.5 Performance Analysis Measurement

An image quality assessment is crucial for evaluating the output HR image generated by the proposed SR reconstruction scheme. This assessment is essential for analysing the performance of the proposed work and determining the level of image quality achieved. Goyal *et al.* showed an analytical relation and comparison of pixel difference-based measurement metrics and Human Vision System based (HVS) metrics on the same baboon image (Goyal et al., 2015). The measurements like the Structural Similarity Index Metric (SSIM) and Peak Signal Square Error (PSNR) values are the image quality metric that is simple to be computed in the field of image processing. Therefore, PSNR, and SSIM values were used to analyse the final HR images produced by the SR reconstruction methods proposed in this dissertation.

2.1.6 Summary

The initial contribution of this dissertation lies in a comparative study focusing on the integration of dictionary learning with sparse representation algorithms for enhanced image denoising. This study encompasses an analysis of three distinct algorithms, each blending dictionary learning with a different sparse representation technique: Douglas-Rachford, Soft Thresholding, and OMP. The evaluation of these algorithms' performance was based on their average PSNR and SSIM in the produced denoised images. These metrics were then contrasted with the results from images solely enhanced through sparse regularization methods like Gradient Descent, Newton Method, and ADMM. It was observed

that the combination between dictionary learning method with Douglas-Rachford based algorithm led in achieving the highest average PSNR and SSIM values, indicating its superior efficacy in LR image enhancement and underscoring the value of integrating dictionary learning in SR tasks.

This dissertation then proposed a SISR framework, leveraging the power of dictionary learning combined with sparse representation. This system, named KSVD_DR, incorporates the Douglas-Rachford algorithm and k-SVD technique, aiming to enhance computational efficiency. Comparisons of k-SVD with methods like OMP, GOMP and BMP as shown in Table 2.1 indicated that can overpower other methods in terms of computation time. A checkmark indicates that the KSVD_DR method demonstrates superior performance in computational time compared to the respective algorithm.

Table 2.1: Comparison evaluation of computation time (seconds) with the proposed KSVD_DR.

Proposed Work (KSVD_DR)	Algorithms			
	k-SVD+OMP	k-SVD+GOMP	k-SVD+GMP	k-SVD+BMP
Computation Time (seconds)	✅	✅	✅	✅

After that, this dissertation introduces an innovative hybrid model that synergizes the capabilities of dictionary learning-based and deep learning-based SR methods. In this hybrid structure, Pathway 1 utilizes the KSVD_DR algorithm, which is built upon dictionary learning and sparse representation strategies, selecting the Douglas-Rachford algorithm for its proficiency in noise management and training efficiency. Concurrently, Pathway 2 integrates the Image Processing Transformer (IPT), a cutting-edge deep-learning-based SR technique. IPT's primary strength is in its ability to comprehend and interpret broader context information within images, contrasting with the patch-by-patch processing typical of dictionary learning-based methods. Despite their advantages, deep learning-based SR techniques often struggle with generalizing across unseen data or varying image types. The dissertation conducts an extensive comparative analysis of the HR images produced by this hybrid approach against existing state-of-the-art dictionary learning-based and deep-learning-based SR methods. The assessment of super-resolution methods covered in this research includes a variety of dictionary learning techniques such as NE, ScSR, ANR, Zedge, Ayas, and Ekinci, detailed in Table 2.2, where a checkmark indicates that the image produced using proposed hybrid method demonstrates better performance in terms of PSNR and SSIM values compared to the respective algorithm

Table 2.2: Comparison evaluation of proposed work with dictionary learning-based SR in terms of PSNR and SSIM values.

Datasets	Scale	Bicubic	Dictionary Learning-based SR				
			NE	ScSR	ANR	Zedge	Ayas & Ekinci
Set14	×2	✅	✅	✅	✅	✅	✅
	×3	✅	✅	✅	✅	✅	✅
	×4	✅	✅	✅	✅	✅	✅
Set5	×2	✅	✅	✅	✅	✅	✅
	×3	✅	✅	✅	✅	✅	✅
	×4	✅	✅	✅	✅	✅	✅
Urban100	×2	✅	NA	NA	NA	NA	NA
	×3	✅	NA	NA	NA	NA	NA
	×4	✅	NA	NA	NA	NA	NA
B100	×2	✅	NA	NA	NA	NA	NA
	×3	✅	NA	NA	NA	NA	NA
	×4	✅	NA	NA	NA	NA	NA

Additionally, a range of deep learning models were reviewed, including VDSR, EDSR, RCAN, SAN, OISR-RK3, RNAN, SRCNN, HAN, IGNN, RDN, and several others, as documented in Table 2.3, where a checkmark indicates the proposed hybrid method outperforms others, an equal sign indicates similar performance, and a cross indicates inferior performance. The results indicate that the newly proposed hybrid approach outperforms the established dictionary learning-based SR methods and is surpassed only by the IPT model among the deep learning-based SR methods.

Table 2.3: Comparison evaluation of proposed work with deep-learning SR in terms of PSNR values.

Datasets	Scale	Deep Learning-based SR					
		SRCNN	VDSR	EDSR	RCAN	RDN	OISK-RK3
Set14	×2	✅	✅	✅	✅	✅	✅
	×3	✅	✅	✅	✅	✅	✅
	×4	✅	✅	✅	✅	✅	✅
Set5	×2	✅	✅	✅	✅	✅	✅
	×3	✅	✅	✅	✅	✅	✅
	×4	✅	✅	✅	✅	❌	✅

Urban100	×2	-	✓	✓	✓	✓	✓
	×3	-	✓	✓	✓	✓	✓
	×4	-	✓	✓	✓	✓	✓
B100	×2	-	✓	✓	✓	✓	✓
	×3	-	✓	✓	✓	✓	✓
	×4	-	✓	✓	✓	✓	✓

Datasets	Scale	Deep Learning-based SR					
		RNAN	SAN	HAN	IGNN	IPT	RDN-MetaSR
Set14	×2	✓	✓	✓	✓	✗	✓
	×3	✓	✓	✓	✓	✗	✓
	×4	✓	✓	✓	✓	✗	✓
Set5	×2	✓	✗	✓	✓	✓	✓
	×3	✓	✗	✗	✓	✓	✓
	×4	✓	✓	✗	✓	✓	✓
Urban100	×2	✓	✓	✓	✓	✗	✓
	×3	✓	✓	✓	✓	✗	✓
	×4	✓	✓	✓	✓	✗	✓
B100	×2	✓	✓	✓	✓	✗	✓
	×3	✓	✓	✓	✓	✗	✓
	×4	✓	✓	✓	✓	✗	✓

Datasets	Scale	Deep Learning-based SR			
		RDN-LIIF	RDN-LTE	RDN-SRNO	DDL-EDSR
Set14	×2	✓	=	✓	✓
	×3	✓	=	✓	✓
	×4	✓	✓	✗	✓
Set5	×2	✓	✓	✗	✓
	×3	✓	✓	✗	✓
	×4	✓	✓	✗	✓
Urban100	×2	✓	=	✓	✓
	×3	✓	=	✓	✓
	×4	✓	=	✓	✓
B100	×2	✓	=	✓	✓
	×3	✓	=	✓	✓
	×4	✓	=	✓	✓

Table 2.4 extends the comparative evaluation to include PSNR values for various grayscale images when contrasted with the IPT model. The proposed method surpasses IPT in grayscale image enhancement, particularly for upscale cases ×2 and ×3 using NIH CT Scan images, except for certain scenarios. This hybrid method combines the advantages of various algorithms to produce superior image quality. It demonstrates its strength within the domain of dictionary learning-based SR methods. Relative to deep-learning-based SR methods, the proposed approach is competitive, trailing the top-performing method by a close margin in PSNR values and outperforming IPT in grayscale image tests.

Table 2.4: Comparison evaluation of PSNR values with IPT.

Grayscale Datasets	Scale	IPT
MMU Iris	×2	✅
	×3	✅
	×4	✅
MIRFLICKR	×2	✅
	×3	✅
	×4	✅
NIH CT Scan	×2	❌
	×3	❌
	×4	✅

2.2 Detection and Classification of Thrips

Recent studies have highlighted the critical importance of accurately identifying and managing thrips, pests known for damaging crops and spreading plant diseases. The research by Hulagappa *et al.* emphasizes the need to understand the potential invasion risks of thrips under changing climate conditions highlighting the complexities faced in pest management due to environmental changes (Hulagappa et al., 2022). The challenge is further exacerbated by the difficulty in detecting thrips, given their small size. Addressing this identification challenge, the review by Mehle and Trdan compares traditional and modern methods, suggesting a combination approach for the most accurate identification of thrips (Mehle & Trdan, 2012). While new techniques in morphometry, molecular biology, and biochemistry offer efficient alternatives, traditional morphological methods remain vital for initial identification and subsequent validation.

Building on these methodologies, Lin *et al.* introduced a YOLO-based method for the automatic detection and counting of small yellow thrips (SYT) on lotus leaves, effectively tackling the challenges

posed by the leaf's curved structure (Lin et al., 2021). This approach, employing VDSR and Deep Plug-and-play SR (DPSR) networks, led to a significant improvement in detection rates. In a similar vein, Amarathunga *et al.* demonstrated the effectiveness of the Vision Transform (ViT) architecture for fine-grained classification of microscopic insect pests, achieving high accuracy in distinguishing morphologically similar species (Amarathunga et al., 2022). Talking about object detection models, this study specifically focuses on the YOLOv8 model, a continuation of the "You Only Look Once" object detection AI model series. Originating from Joseph Redmon's pioneering work in object detection, the YOLO framework processes an entire image in a single instance, simultaneously estimating object classes and locations (Redmon et al., 2016). This allows for fast and accurate real-time object detection. Over the years, several iterations, including YOLOv2 and YOLOv3, as well as other improved versions like YOLOX and YOLOv5, have been developed, each enhancing accuracy and speed (Redmon & Farhadi, 2017). In this dissertation, the latest iteration, YOLOv8, announced by Ultralytics in 2023, is utilized. YOLOv8 stands out for its combination of high-speed processing and high accuracy, making it an ideal choice for the efficient and effective detection and classification of thrip in agricultural settings.

Hence, this dissertation builds upon these advancements by implementing a novel dictionary learning-based SR method, KSVD_DR, alongside the advanced object detection model YOLOv8, for the automated detection and classification of thrips in agricultural setting. The proposed system, featuring the KSVD_DR method, is designed to overcome the challenges posed by existing approaches. Its advantages include not needing large image datasets and having shorter training periods, due to its simpler structural design. This simplicity also makes it easier to integrate into existing systems, while still being capable of delivering reliable and effective results. The objective is to not only emulate the success of deep learning methods, as demonstrated by Lin *et al.*, but also to expand on Amarathunga *et al.* work by classifying six types of thrip. This approach promises to enhance the process of thrips counting, offering a more efficient and less labor-intensive alternative to manual methods. The anticipated improvements are expected to be evident both with and without the use of SR, underscoring the potential of this methodology to significantly impact pest management practices.

3. PERFORMANCE ANALYSIS ON DICTIONARY LEARNING AND SPARSE REPRESENTATION ALGORITHM

The discussion regarding the effectiveness of applying dictionary learning steps with sparse representation algorithms was conducted as shown in Figure 3.1. To establish a baseline, sparse regularization methods like the Gradient Descent (GD) method, Newton method, and Alternating Direction Method of Multipliers (ADMM), employing either isotropic or anisotropic total variation (TV), were executed without the dictionary learning process. For comparative analysis, three distinct algorithms, referred to as Algorithm 1, 2 and 3 which meld dictionaries learning with various sparse coding methods, including Douglas Rachford algorithm, soft thresholding algorithm, and OMP algorithm were evaluated.

For all the experiments described in the dissertation, laptop equipped with an 11th Gen Intel(R) Core (TM) i7-1165G7 processor at 2.80 GHz and 16.0 GB of RAM (15.8 GB usable) was used. The system runs on a 64-bit OS with an x64-based processor, ensuring efficient performance and software compatibility. All computations, including algorithm execution and data processing, were carried out using MATLAB software, due to its capability to handle demanding computational tasks.
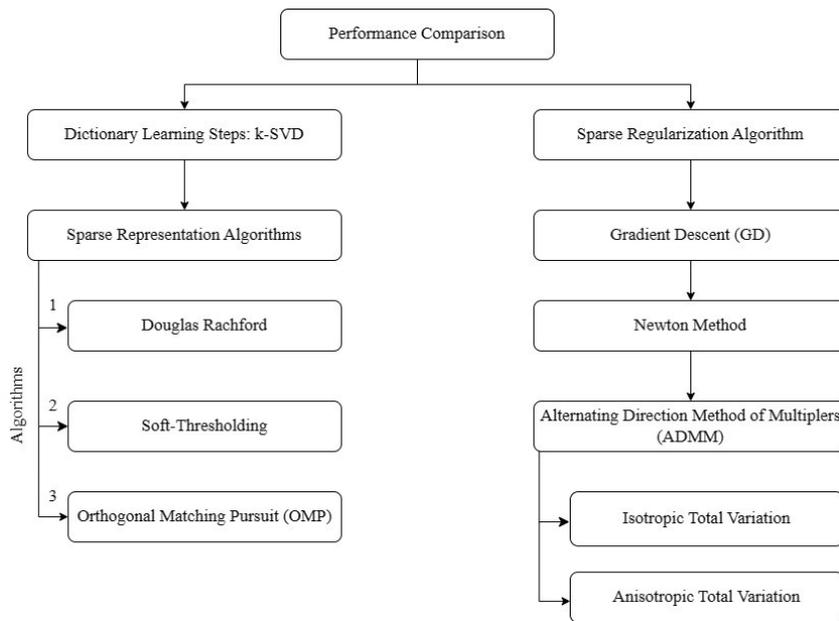

Figure 3.1: Performance comparison between dictionary learning with sparse representation algorithms and sparse regularisation algorithms.

First, the LR image reconstruction is done by adding the Gaussian noise with $\delta = 0.06$ on the image as shown in figure below. After that, the final LR image was used to produce a large number m of image patches and stored in a matrix Y. The width of the image patches, w was set to be 10 and there were 12,000 image patches stored in the matrix Y. These 12,000 image patches were processed to remove the mean and only image patches with the largest energy were kept. In this case, 4,000 image patches from the matrix Y were kept and used for dictionary training later.

Figure 3.2: HR image (left) and LR image constructed (right).

Then, an initial dictionary, $D_0 \in \mathbb{R}^{100 \times 200}$ was produced. Figure 3.3 shows an example of the initial dictionary produced for LR image constructed.
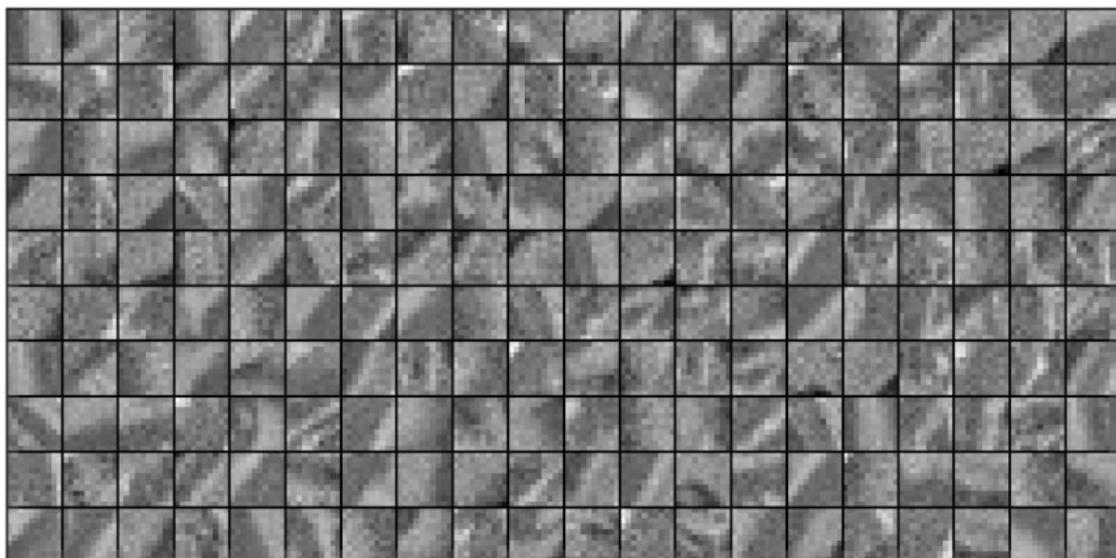

Figure 3.3: Initial dictionary, $D_0 \in \mathbb{R}^{100x200}$.

The initialization of the overcomplete dictionary set the stage for the creation of the learned dictionary through the k-SVD algorithm. Within this training process, a variety of sparse representation algorithms were employed, including the Douglas-Rachford, soft thresholding, and OMP methods. These were instrumental in refining the k-SVD algorithm's output. Figure 3.4, Figure 3.5 and Figure 3.6 show the corresponding final dictionary produced by using Algorithm 1, 2 and 3 respectively.
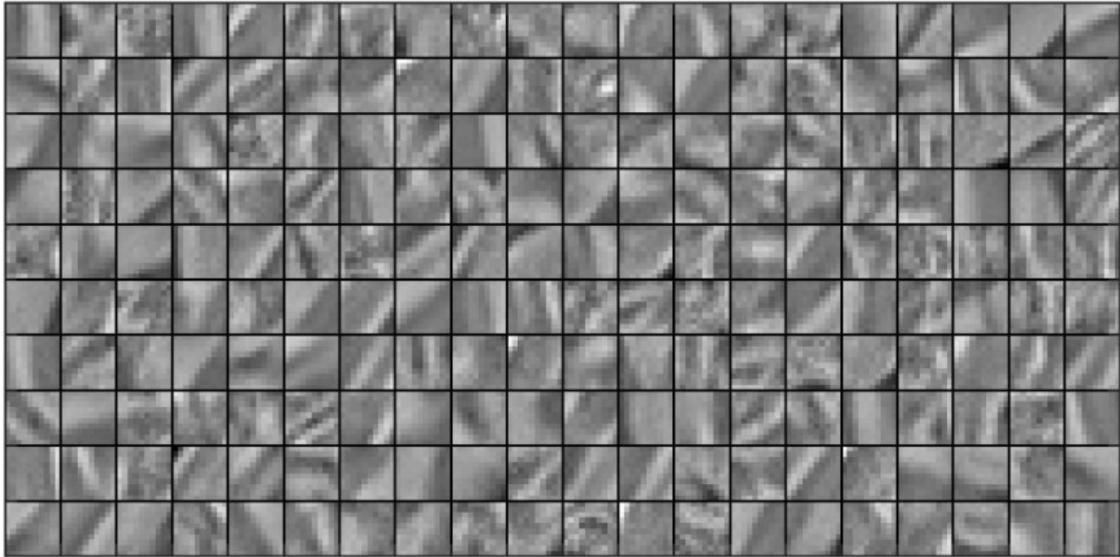

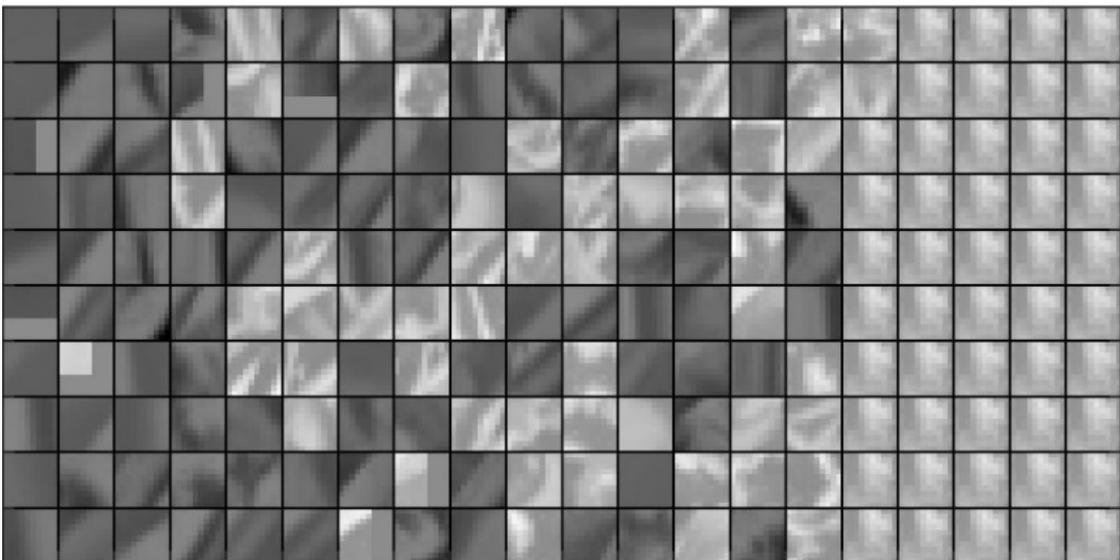**Figure 3.4: Final trained dictionary, $D \in \mathbb{R}^{100x200}$ using Algorithm 1.**

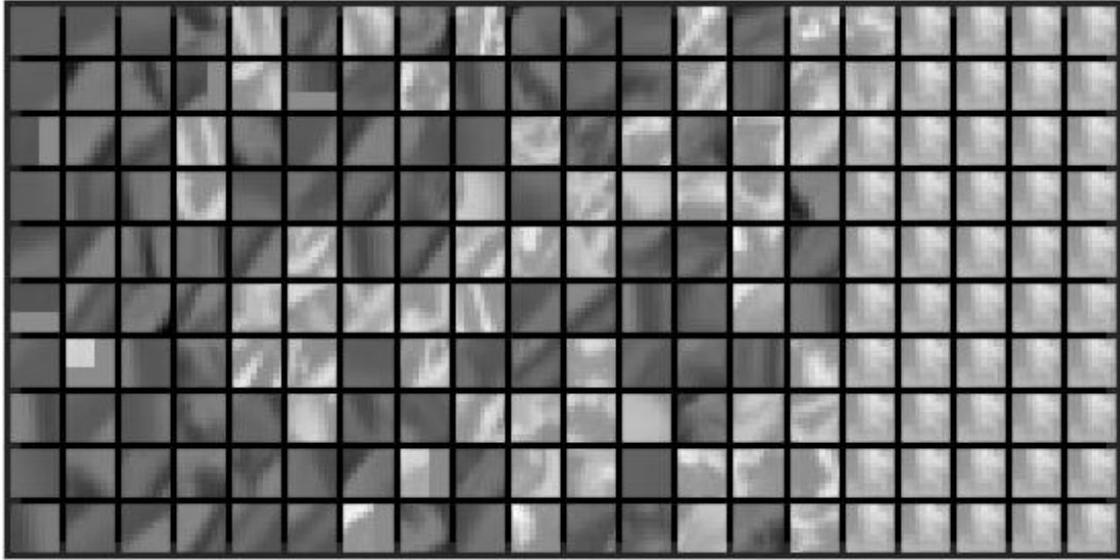**Figure 3.5: Final trained dictionary, $D \in \mathbb{R}^{100x200}$ using Algorithm 2.**

Figure 3.6: Final trained dictionary, $D \in \mathbb{R}^{100x200}$ using Algorithm 3.

After that, different sparse coding methods in Algorithms 1, 2, and 3 were used to conduct the image denoising process. In this case, the overlap factor was set as 2 and 15,876 image patches were extracted from the LR image. These image patches were fitted into the sparse coding methods in the respective algorithms to produce the denoised image patches. These denoised image patches were then averaged and the HR image was successfully developed.

However, this dissertation also aimed to discuss the effectiveness of implementing the dictionary learning process with sparse representation algorithms. Hence, all the LR images used in this dissertation were also processed using the sparse regularisation methods: GD method, Newton method, and ADMM by using isotropic TV or anisotropic TV to produce HR images without undergoing the dictionary learning process.

3.1 Results and Discussion

Figure 3.7 presents the denoised outcomes obtained from Algorithms 1, 2, and various sparse regularization methods. The analysis demonstrated that Algorithm 1 yields the highest-quality HR image, effectively reconstructing the image without the blurring effect that diminishes the results of Algorithm 2. Algorithm 3, while somewhat effective, could not eliminate all the Gaussian noise as efficiently as Algorithm 1. Conversely, the HR images resulting from the sparse regularization methods exhibited noticeable blurring effects.

Figure 3.7: LR Image, HR image produced by Algorithm 1, 2, 3 and sparse regularisation methods (Gradient Descent (GD) method, Newton Method, ADMM by isotropic TV and anisotropic TV).

To quantify the performance of the algorithms, PSNR and SSIM values were calculated. For these calculations, the original HR image serves as the reference. PSNR, which qualifies the ratio between the maximum possible power of a signal and the power of corrupting noise, is expressed in decibels (dB) and reflects image fidelity. The computation of the PSNR value follows the formula presented in Equation (3.1), where x is the targeted image, y is the reference image, MPP is the Maximum Possible Pixel in an image and MSE is the Mean Square Error of an image. The larger the PSNR value, the greater the image quality obtained by the output image.

$$PSNR\ (x, y) = 10 \log_{10}[\frac{MPP^2}{MSE\ (x, y)}] \qquad (3.1)$$

Next, the SSIM value of an image was computed based on the comparison on luminance, contrast and structure in image x and y as shown by the Equation (3.2), where the $l(x, y)$, $c(x, y)$ and $s(x, y)$ is defined as the luminance comparison function, contrast comparison function and structure comparison function respectively. The closer the SSIM value to 1, the greater the image quality of the denoised image produced as the SSIM value only lies between 0 and 1.

$$SSIM = l(x, y) \cdot c(x, y) \cdot s(x, y) \qquad (3.2)$$

As depicted in Figure 3.8, Algorithm 1 attained the highest average PSNR value at 31.817 dB for the standard test images. Among the sparse regularization methods, the GD method recorded the highest average PSNR value of 29.730 dB for the same set of images.
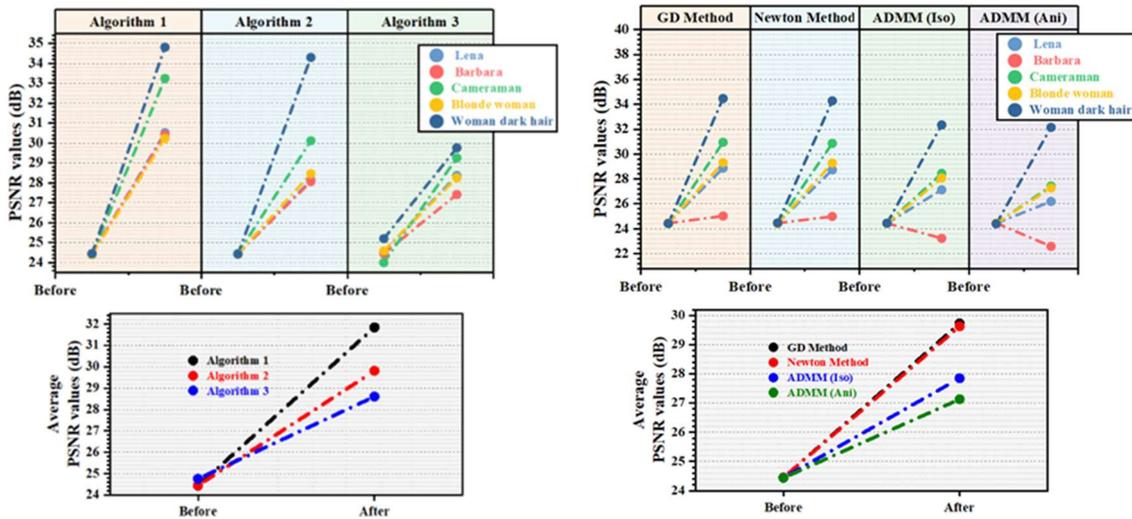

Figure 3.8: PSNR values of Algorithm 1, 2, 3 and sparse regularisation methods (Gradient Descent (GD) method, Newton Method, ADMM by isotropic TV and anisotropic TV).

Figure 3.9 presents the SSIM metrics for Algorithms 1, 2, and 3, alongside the sparse regularization methods. Consistent with the PSNR outcomes, Algorithm 1, which combines KSVD for dictionary learning and the Douglas-Rachford algorithm for sparse representation, secured the highest average SSIM score of 0.8764 on standard test images. The GD method emerged as the top performer among the sparse regularization techniques, with an average SSIM of 0.8420.
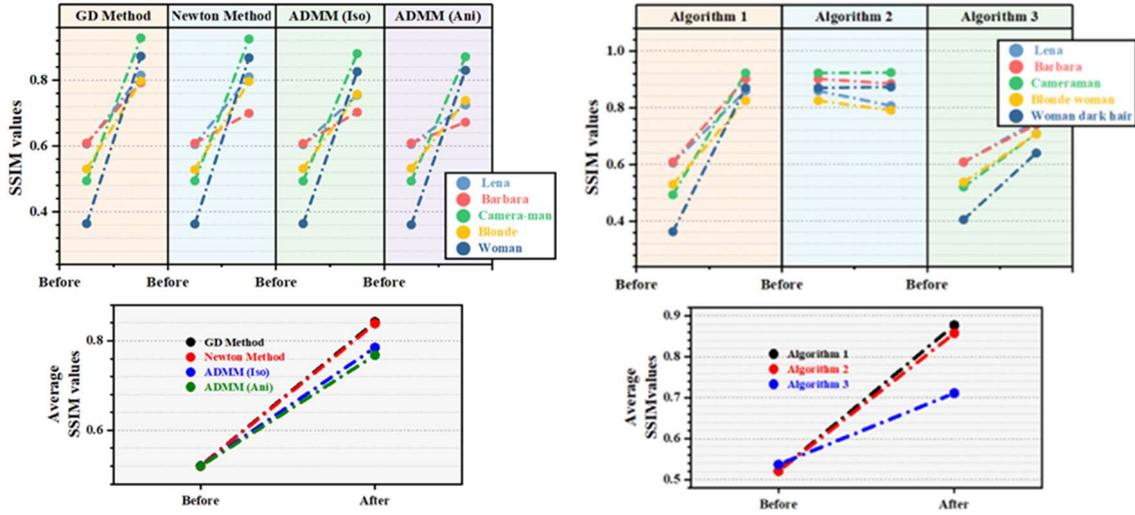
Figure 3.9: SSIM values of Algorithm 1, 2, 3 and sparse regularisation methods (Gradient Descent (GD) method, Newton Method, ADMM by isotropic TV and anisotropic TV).

3.2 Summary

Overall, the PSNR and SSIM results indicate that the HR images reconstructed by Algorithm 1 outperformed those from Algorithms 2 and 3, as well as the sparse regularization methods executed without dictionary learning. In summary, Algorithm 1's integration of dictionary learning with sparse representation methods demonstrated superior effectiveness in enhancing the quality of LR images, as evidenced by its leading PSNR and SSIM values. Hence, a dictionary learning-based SR algorithm that utilizes the Douglas-Rachford algorithm in conjunction with the k-SVD method was employed in the subsequent parts of this dissertation.

4. SINGLE-IMAGE SUPER-RESOLUTION (SISR) USING DICTIONARY LEARNING AND SPARSE REPRESENTATION ALGORITHM

4.1 Methodology

Figure 4.1 shows the SISR algorithm proposed based on the dictionary learning and sparse representation algorithm, by incorporating the Douglas-Rachford algorithm with k-SVD technique, named KSVD_DR.
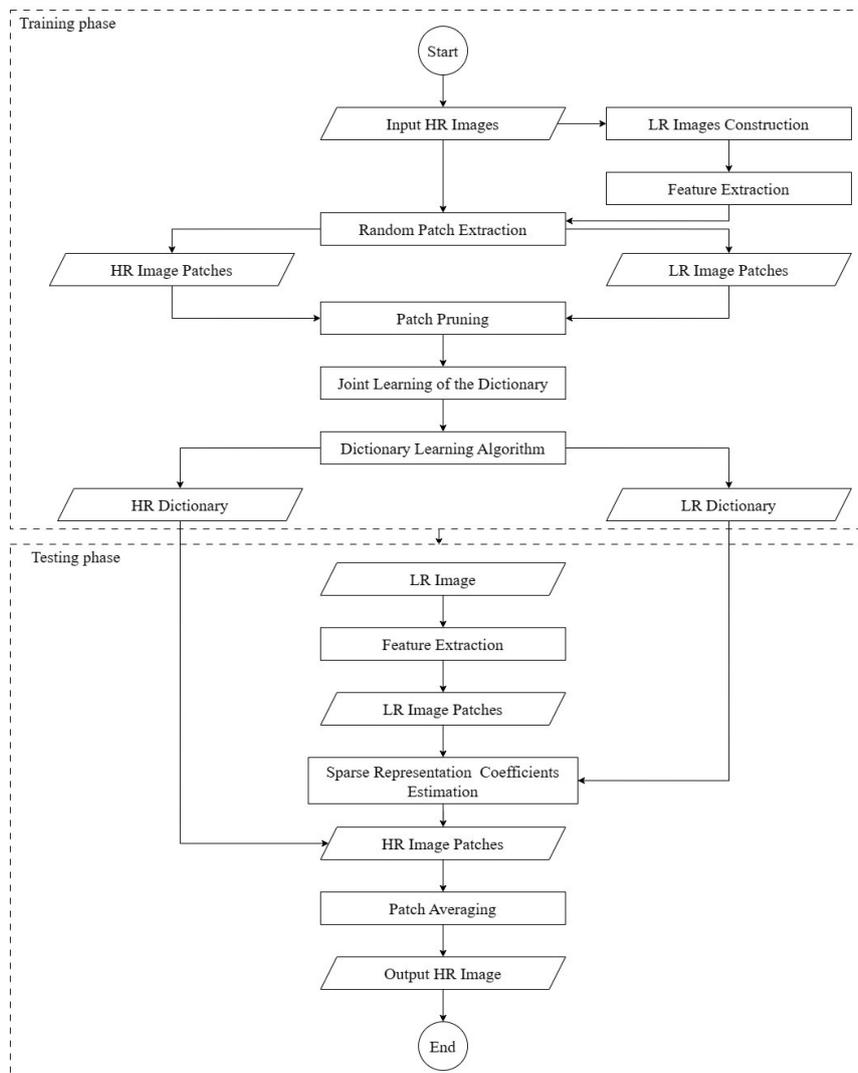

Figure 4.1: System architecture of the KSVD_DR.

33

The training stage was started by collecting different coloured HR images, including various types like plants, buildings cars, and fruit, from publicly available database (Gonzalez, 2009) as shown in Figure 4.2. In this dissertation, 69 coloured images which includes were used for training the dictionary. These HR images were firstly converted to grayscale and then downgraded to LR images by performing the down-sampling operator with factor of 2 as shown in Figure 4.3. Different down-sampling factors such as the $\times 2$, $\times 3$, and $\times 4$ were also consider in forming the LR images for analysis purpose in this dissertation.

Figure 4.2: Different types of training images, (A) flower (176×197 pixels), (B) building (351×436 pixels), (C) car (163×195 pixels) and (D) fruit (307×308 pixels).

Figure 4.3: Down sampled LR images in grayscale, (A) flower (88×99 pixels), (B) building (176×218 pixels), (C) car (82×98 pixels) and (D) fruit (154×154 pixels).

Then, these LR images were resized to match the pixel size of the original HR images using an interpolation operation. Then, features were extracted from the interpolated LR images by applying the feature extraction filters, f_1, f_2, f_3 and f_4. The filters f_1 and f_2 are defined as $f_1: [-1, 0, 1]$, $f_2: [1, 0, -2, 0, 1]$ respectively. Filters f_3 and f_4 are the transpose of f_1 and f_2 as shown in Equation (4.1) and (4.2).

$$f_3 = f_1^T \tag{4.1}$$
$$f_4 = f_2^T \tag{4.2}$$

In this dissertation, four distinct filters were applied to LR image patches, extracting four separate feature vectors. These vectors were then concatenated to form a comprehensive feature representation for each LR image patch. Consequently, for every LR image in the training set, four gradient maps

were produced by the direct application of these filters. Subsequently, at each location, LR image patches were extracted from these gradient maps and combined to create the final feature vector. Figure 4.4 shows a simple illustration of the process of feature extraction for LR image. By doing this, it brings benefits in improving the compatibility among adjacent patches in the final HR image as the feature representation for each LR image patch obtained was encoded with its neighbouring information.
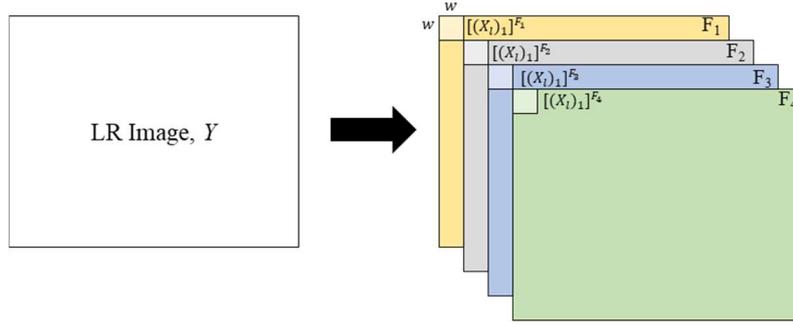

Figure 4.4: Simple illustration of feature extraction process for LR image.

Next, a large number m of the HR and LR image patches which will be stored in matrix X_h and X_l respectively were extracted with size $n = w \times w$. The width of the image patches, w was set as 3 and the value of total number of image patches, m was set to be 100,000. Since a total number of 69 images were used in this research for training purpose, therefore, value of m cannot be equally divided between these training images. The dimension of both the LR and HR image patches were explained by the Equation (4.3) and (4.4).

$$X_h \in \mathbb{R}^{9 \times 99966} \tag{4.3}$$

$$X_l \in \mathbb{R}^{36 \times 99966} \tag{4.4}$$

Image patches from the LR and HR images were then taken from the same areas in the respective images. The mean of each HR image patches was first removed to produce dictionary with zero mean and unit norm in the following steps. To make the operation simpler, a process called patch pruning was used, which reduces the number of image patches for the dictionary learning algorithm. The LR and HR image patches were pre-normalized before conducting the patch pruning process. In the patch pruning process, the values obtained from calculating the variance of elements in each column of LR image patches were then used to compare with the threshold value. In this case, the threshold value was set to be 10 in this research and the number of values which were larger than the threshold value that will be kept for the dictionary process. For example, in the case for w is equal to 3 with an upscale factor of 2, the number of patches extracted from both LR and HR images were reduced from 99966 patches to 65999 patches as denoted by Equation (4.5) and (4.6) as shown below.

$$X_h \in \mathbb{R}^{9 \times 65999} \tag{4.5}$$

$$X_l \in \mathbb{R}^{36 \times 65999} \tag{4.6}$$

By using the same example of w was set to be 3 with upscale value of 2, the matrix dimension of X_c obtained by combining the LR and HR image patches for joint learning of dictionary was defined as Equation (4.7).

$$X_c \in \mathbb{R}^{45 \times 65999} \tag{4.7}$$

In this case, the LR and HR dictionaries were trained together using a dictionary training method. In this dissertation, the k-SVD algorithm was employed for its proficiency in finding an optimal dictionary for sparse signal representation. The k-SVD algorithm begins with dictionary initialization. Following this, the dictionary was held constant, the sparsest representation of the input data was determined using sparse coding method. After this, while maintain a fixed sparse representation, the dictionary was updated to fit the data more accurately. This process of finding the sparsest representation and updating the dictionary continues until the dictionary reaches a stable solution that cannot be improved further.

In this dissertation, although greedy algorithms were commonly used for sparse coding, the Douglas-Rachford algorithm was favoured due to its enhanced noise management and training effectiveness. This preference was informed by performance analysis results, which previously compared k-SVD with various sparse representation and regularization methods. The advantages of the Douglas-Rachford algorithm in this context were particularly highlighted. The computational efficiency of the k-SVD algorithm was evaluated in combination with several methods, including Orthogonal Matching Pursuit (OMP), Douglas Rachford, Group OMP (GOMP), Group MP (GMP), and Block MP (BMP). These comparisons were conducted over 20 iterations to assess the efficacy of k-SVD when paired with each method. The laptop specifications remain consistent with those outlined in the beginning of Chapter 3.

After the process of dictionary learning by using k-SVD algorithm with Douglas-Rachford (KSVD_DR) as sparse coding method, both the LR and HR learned dictionaries were obtained as shown in Equations (4.8) and (4.9). The dictionary size, p used in this research is set as 512.

$$D_h \in \mathbb{R}^{9 \times 512} \tag{4.8}$$

$$D_l \in \mathbb{R}^{36 \times 512} \tag{4.9}$$

For the testing phase, the input LR image was generated from the corresponding HR image through down-sampling process. Following this, features were derived from the LR image using the same feature filter employed during the training phase. LR image patches were then extracted from this filtered LR image. These patches, alongside the LR dictionary produced in the previous phase, were used to calculate the sparse coefficients via the same method as used in KSVD_DR which is the Douglas Rachford algorithm. Consequently, HR image patches were computed using the sparse representation coefficients and the HR dictionary obtained from the training phase. After the patch averaging operation, the final HR was produced. The evaluation will also be done in this stage to

observe the effect it has on the HR images.

4.2 Results and Discussion

4.2.1   Evaluation on Computational Time

Table 4.1 presents the computational time, in seconds, for 20 iterations across various combinations of the k-SVD algorithm paired with various methods, such as OMP, Douglas Rachford (DR), GOMP, GMP and BMP. Of all these combinations, the k-SVD with DR emerged as the most time-efficient in the training phase of the dictionary learning algorithm. The superiority of the use of k-SVD with DR can be attributed to the inherent qualities of the Douglas Rachford algorithm. For instance, while OMP relies on a threshold value which can sometimes limit its efficiency, the Douglas Rachford approach offers more robustness in handling such parameters, making it a commendable choice for combination with k-SVD.

Table 4.1: The computation time of 20 iterations.

Algorithms	Computation Time (seconds)
k-SVD + OMP	885.7643
k-SVD + DR	**797.3006**
k-SVD + GOMP	863.1495
k-SVD + GMP	4787.7249
k-SVD + BMP	1858.7987

4.2.2   Evaluation on Grayscale Images

The evaluation of grayscale images in this dissertation was focused on assessing the PSNR values, positioning them against the benchmark set by the Image Processing Transformer (IPT), a leading SR method. A subset of grayscale images from the MIRFLICKR dataset, which originates from a diverse collection on Flickr, was selected for testing. Despite the KSVD_DR algorithm being trained on a relatively small dataset and requiring a shorter training duration, it demonstrated superior PSNR values compared to IPT as shown in Table 4.2. The IPT model, by contrast, is pretrained on the extensive ImageNet dataset and typically necessitates a longer training period. The findings indicate that our KSVD_DR algorithm not only achieves higher-resolution images in grayscale but also enhances computational efficiency during training, outperforming IPT in these key areas.

Table 4.2: Evaluation on grayscale images in terms of PSNR values.

Grayscale Dataset	Scale	PSNR (dB)	
		IPT	KSVD_SR
MIRFLICKR	×2	24.80486	**24.82395**
	×3	23.89823	**23.92559**
	×4	22.13902	**22.16586**

4.3 Summary

As a summary, this dissertation provided the SISR algorithm proposed based on the dictionary learning and sparse representation algorithm, by incorporating the Douglas-Rachford algorithm with k-SVD technique, known as KSVD_DR. In this case, this dissertation benchmarks the efficiency of the k-SVD algorithm with various sparse coding methods, concluding that k-SVD paired with the Douglas Rachford algorithm yields the best computational efficiency in dictionary learning. In grayscale image evaluation, despite a smaller training dataset and less training time, the KSVD_DR algorithm surpasses the IPT in PSNR values, highlighting its ability to produce high-quality images with greater training efficiency.

5. HYBRID IMAGE RECONSTRUCTION ALGORITHM FOR LOW-RESOLUTION IMAGES

5.1 Methodology

Figure 5.1 illustrates the comprehensive flowchart of the proposed approach. The HR output images, generated by the SR algorithm from both pathways, will undergo the same hybrid procedure. Pathway 1 uses the SR algorithm called KSVD_DR, which is based on the dictionary learning and sparse representation method. Pathway 2 employs the SR algorithm named Image Processing Transformer (IPT) which is a state of the art deep-learning-based SR technique (H. Chen et al., 2021).

In this chapter, the IPT in Pathway 2 was executed on a server equipped with an NVIDIA GeForce RTX 4090 GPU, supporting CUDA Version 12.2 and featuring a total memory of 24,564 MiB. This setup was essential for handling the computationally intensive tasks of the IPT. All other computations and analyses were carried out using the laptop specifications detailed in Chapter 3
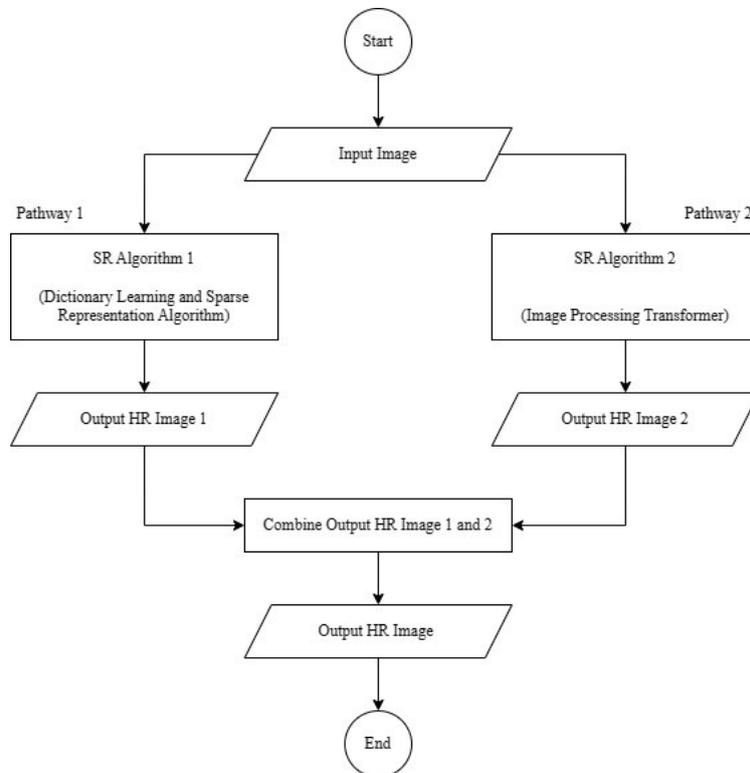

Figure 5.1: Hybrid image reconstruction algorithm.

The SR strategy presented in this dissertation introduces a dual-pathway approach to enhance image resolution. In Pathway 1 of the SR strategy outlined in this dissertation, the focus is on dictionary learning and sparse representation, utilizing the KSVD_DR method. As detailed in Chapter 4, while this approach is highly effective in enhancing contrast information in grayscale images, its primary limitation lies in processing colour information. This is due to the method being trained predominantly on grayscale patches, limiting its applicability to colour images.

On the other hand, Pathway 2 introduces the Image Processing Transformer (IPT) model as conceived by Chen et al., offering significant benefits (H. Chen et al., 2021), which, despite being trained on extensive datasets like ImageNet, encounters challenges when dealing with unseen data and may not perform as well on grayscale images. The IPT model excels in processing colour images, benefiting from its training on diverse and comprehensive colour datasets. However, its effectiveness is somewhat diminished when it comes to images that lack colour diversity or are outside its training scope.

To address these complementary strengths and limitations, this dissertation proposes a novel post-processing hybrid method. This approach combines the advantages of both Pathway 1 and Pathway 2, leveraging the contrast enhancement capabilities of KSVD_DR for grayscale images and the colour processing proficiency of the IPT model. The hybrid method is designed to overcome the individual limitations of each pathway, offering a more versatile and comprehensive solution for super-resolution tasks. By integrating the distinct benefits of dictionary learning and deep learning, the hybrid method provides a robust solution that enhances image quality across both grayscale and colour images, effectively dealing with the challenges posed by unseen data.

5.1.1   Hybrid Techniques

Figure 5.2 shows the flowchart of the hybrid technique proposed in this paper. First, the process was started with the input images, I_{SR} and I_{SR1}, representing the output HR images from both SR algorithm 1 and 2 respectively. Then, a predefined 2D Gaussian filter with standard deviation, σ of 1 and a size of 5×5, was created. Then, normalization process was applied to the predefined 2D Gaussian filter as shown in Equation (5.1) and (5.2).

$$g_1(i, j) = (G(i,j))^2 \tag{5.1}$$

$$g(i, j) = \frac{g_1(i, j)}{\sum g_1(:)} \tag{5.2}$$

where $g_1(i, j)$ is the results of squaring the $G(i,j)$, $G(i,j)$ is the Gaussian filter value at coordinate (i,j), $g(i, j)$ is the normalized filter value at coordinate (i,j) and $\sum g_1(:)$ is the sum of all elements in the Gaussian filter matrix.

After that, the pixel values of I_{SR1}, I_{SR2} were converted to double-precision floating-point format as shown in Equation (5.3) and (5.4).

$$I_{SR1_y}(x, y) = F(I_{SR1}(x,y)) \tag{5.3}$$

$$I_{SR2_y}(x, y) = F(I_{SR2}(x,y)) \tag{5.4}$$

Afterwards, the following process was repeated for each iteration, $i = 1, 2, ..., maxIter$. The maximum iteration, $maxIter$ is empirically set as 20 in the current application. Firstly, the $I_{SR1_y}(x, y)$, and $I_{SR2_y}(x, y)$ were down sampled to the size of $row\_l \times col\_l$ (depends on the scaling factor) by using bicubic interpolation. This down-sampling process is to reduce the computational complexity while reconstructing the image. After that, the difference between the output images from both pathways, $I_{diff}(x, y)$ was calculated in Equation (5.5), where the I_{SR1_s} and I_{SR1_s} down-sampled images from the previous process.

$$I_{diff}(x,y) = I_{SR1_s}(x,y) - I_{SR2_s}(x, y) \tag{5.5}$$

The $I_{diff}(x,y)$ was then up sampled again to the size of HR image, I_{HR} with the size of $row\_h \times col\_h$ by using the proposed KSVD_DR. The final image after this operation was expressed as $I_{diff_s}(x,y)$.

Finally, a convolutional operation was applied to $I_{diff_s}(x,y)$ using the Gaussian filter kernel, $g(i, j)$ which created at the beginning of the process. The convolution operation was computed using the '*convn*' function in MATLAB, which performs a multi-dimensional convolution operation. The parameter '*same*' was specified to preserve the spatial dimensions of the input, thereby ensuring that the convolved image retains its original size. Finally, the outcome of the convolutional operation was added to the $I_{SR2_y}(x, y)$, to produce the final output image, $I_{COM}(x, y)$.
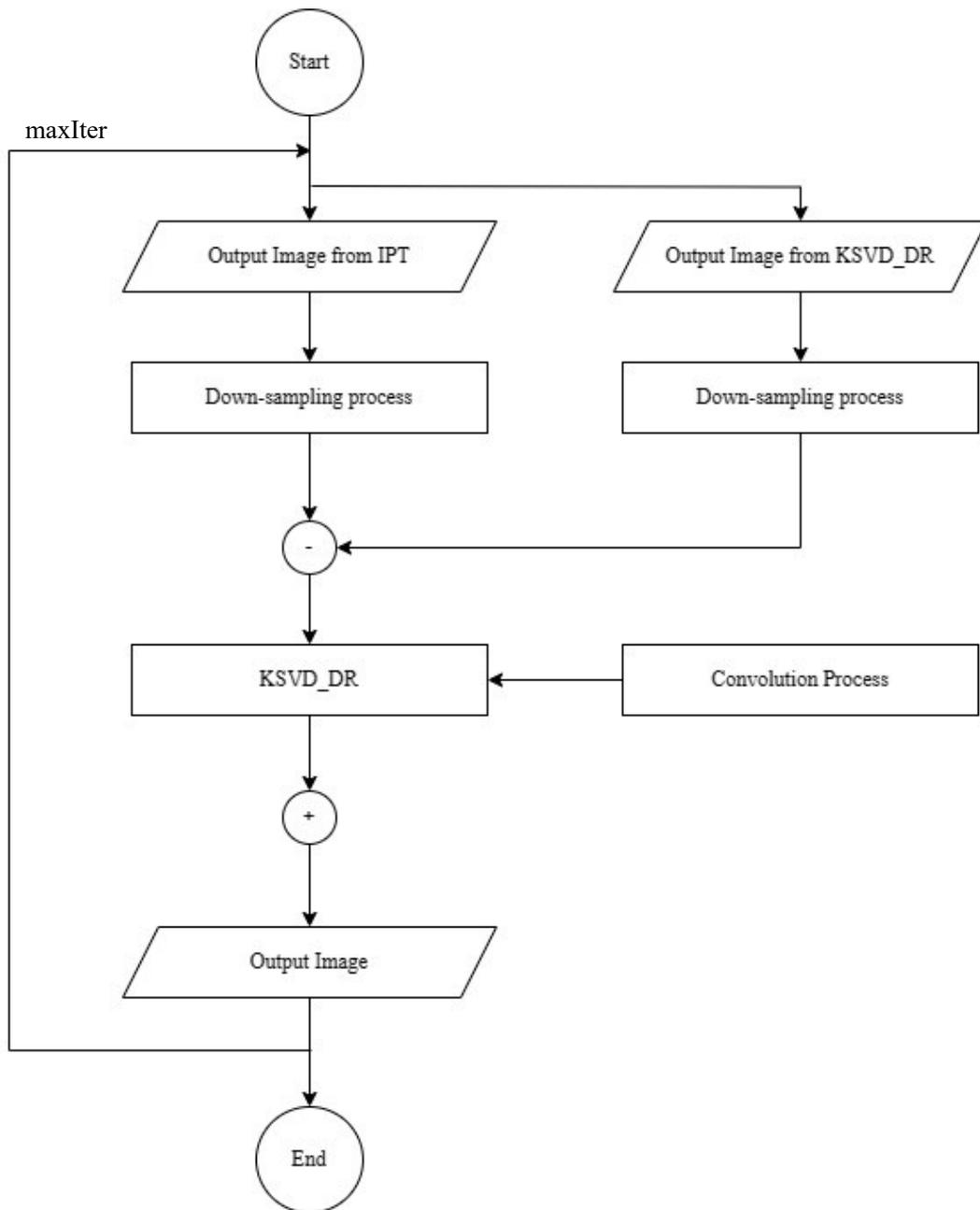
Figure 5.2: Flowchart of the hybrid image reconstruction algorithm.

5.2 Results and Discussion

5.2.1   Evaluation on General Datasets

The quality of the HR images generated by the proposed hybrid technique were also assessed using well-established image quality evaluation process indices: RSNR and SSIM. In this dissertation, the Set14, Set5, Urban100, and B100 datasets were used as the testing images. The Set14, Set5, Urban100, and B100 datasets are integral in the field of image SR, offering a diverse range of challenges for testing SR algorithms. The Set5 dataset is a compact collection used frequently as a benchmark, comprising five high-quality images that include both natural scenes and man-made objects. Due to its small size, Set5 is ideal for quick tests and demonstrations. Expanding upon this, the Set14 dataset includes 14 HR images featuring a wider variety of scenes such as landscapes, urban environments, and text. This dataset is particularly valuable for more comprehensive testing owing to its diverse image content.

The Urban100 dataset, with its focus on urban scenes, contains 100 HR images. It is characterized by its complexity, featuring detailed architectural structures and dense urban landscapes, making it a rigorous testbed for assessing the quality of SR methods on intricate and textured images. On the other hand, the B100 dataset, also known as BSD100 or Berkeley Segmentation Dataset 100, consists of 100 natural images. These images are sourced from a variety of natural scenes and provide a broad spectrum of textures and patterns. The B100 dataset is commonly utilized to evaluate SR algorithms in natural settings, where the preservation of fine details and textures is of utmost importance. Each dataset offers a unique set of images, enabling a thorough assessment of SR algorithms under various conditions, from simple scenarios to highly complex and detailed environments.

Table 5.1 and Table 5.2 offer a comparative analysis with dictionary learning-based SR techniques such as NE, ScSR (Yang et al., 2010), ANR (Timofte et al., 2013), Zedge (Zeyde et al., 2012), and Ayas & Ekinci (Ayas & Ekinci, 2020), focusing on PSNR values and SSIM metrices. The data presented in these tables clearly indicated that our proposed approach, consistently outperforms in terms of achieving higher PSNR and SSIM values.

Table 5.1: Comparison with dictionary learning-based SR in terms of PSNR values (dB).

Datasets	Scale	Bicubic	Dictionary Learning-based SR					Proposed Work
			NE	ScSR	ANR	Zedge	Ayas & Ekinci	
Set14	×2	30.09	31.76	-	31.80	31.76	31.84	**34.29**
	×3	27.42	28.60	28.31	28.65	28.67	28.79	**30.75**
	×4	25.88	26.81	-	26.85	26.90	26.99	**28.79**
Set5	×2	33.65	35.77	-	35.83	35.72	35.81	**38.13**

	×3	28.62	31.84	31.42	31.92	31.90	32.04	**34.67**
	×4	28.42	29.61	-	29.69	29.71	29.75	**32.50**
Urban100	×2	26.86	-	-	-	-	-	**33.75**
	×3	24.45	-	-	-	-	-	**29.58**
	×4	23.13	-	-	-	-	-	**27.27**
B100	×2	29.55	-	-	-	-	-	**32.47**
	×3	27.20	-	-	-	-	-	**29.41**
	×4	25.97	-	-	-	-	-	**27.84**

Table 5.2: Comparison with dictionary learning-based SR in terms of SSIM values.

Datasets	Scale	Bicubic	Dictionary Learning-based SR					Proposed Work
			NE	ScSR	ANR	Zedge	Ayas & Ekinci	
Set14	×2	0.9059	0.8959	-	0.8971	0.8944	0.8967	**0.9457**
	×3	0.8378	0.8000	0.7957	0.8015	0.7998	0.8039	**0.8895**
	×4	0.7871	0.7202	-	0.7223	0.7218	0.7251	**0.8459**
Set5	×2	0.9654	0.9449	-	0.9458	0.9442	0.9453	**0.9807**
	×3	0.9199	0.8846	0.8816	0.8857	0.8854	0.8882	**0.9655**
	×4	0.9077	0.8210	-	0.8228	0.8235	0.8262	**0.9521**
Urban100	×2	0.8816	-	-	-	-	-	**0.9584**
	×3	0.8011	-	-	-	-	-	**0.9096**
	×4	0.7408	-	-	-	-	-	**0.8640**
B100	×2	0.9060	-	-	-	-	-	**0.9424**
	×3	0.8433	-	-	-	-	-	**0.8890**
	×4	0.8001	-	-	-	-	-	**0.8485**

On the other hand, Figure 5.3, Figure 5.4 and Figure 5.5 show the qualitative results for the ×3 upscale scenario, utilizing images sourced from Set14 and Set5 databases. These figures compare our results with the state-of-the-art dictionary learning methods, such as the NE, ScSR (Yang et al., 2010), ANR (Timofte et al., 2013), Zedge (Zeyde et al., 2012), and Ayas & Ekinci (Ayas & Ekinci, 2020). These figures demonstrate the higher image quality achieved by our proposed hybrid approach.
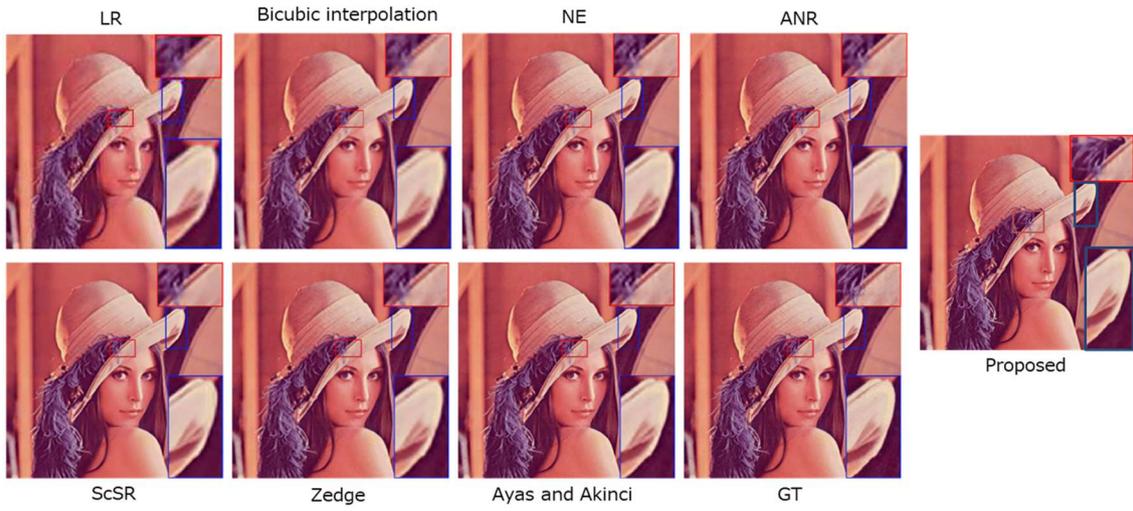
Figure 5.3: Qualitative results on the upscale of $\times 3$ using *Lenna* image (Set14).
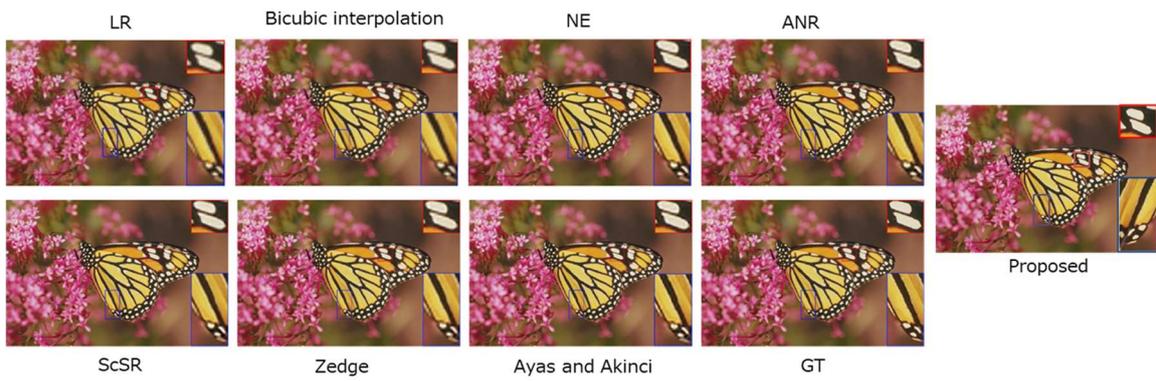

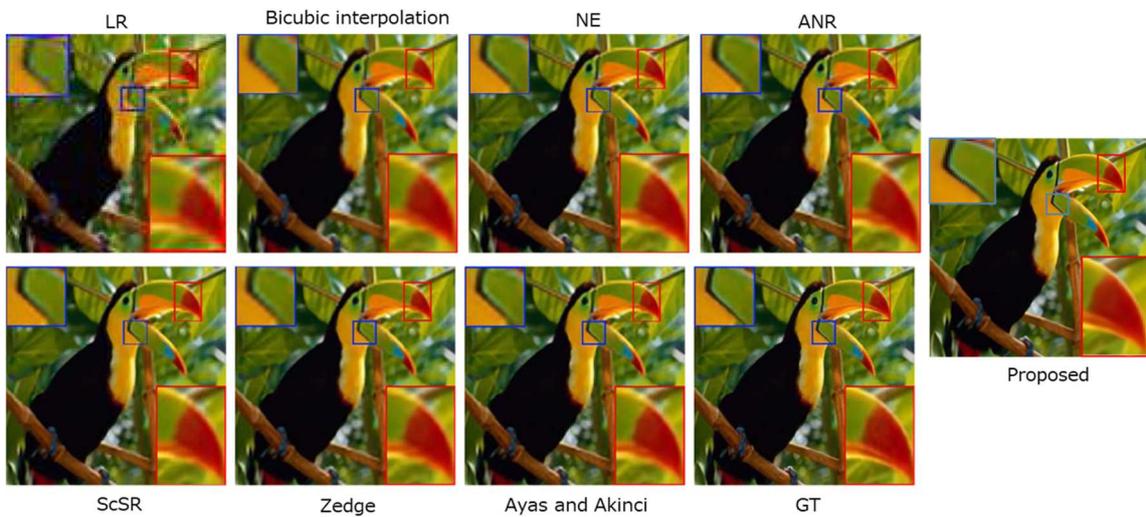**Figure 5.4 Qualitative results on the upscale of $\times 3$ using *Monarch* image (Set14).**

Figure 5.5: Qualitative results on the upscale of $\times 3$ using *Bird* image (Set5).

Table 5.3 provides a comparison with deep learning-based SR methods such as SRCNN (Vb, 2020), VDSR (Kim et al., 2016), EDSR (Lim et al., 2017), RCAN (Y. Zhang et al., 2018), RDN (Zhang et al., 2020), OISK-RK3 (He et al., 2019), RNAN (Zhang et al., 2019), SAN (Dai et al., 2019), HAN (Niu et al., 2020), IGNN (Zhou et al., 2020), IPT (H. Chen et al., 2021), RDN-MetaSR (Hu et al., 2019), RDN-LIIF (Y. Chen et al., 2021), RDN-LTE (Lee & Jin, 2022), RDN-SRNO (Wei & Zhang, 2023), DDL-RDSR (Liu et al., 2023), specifically focusing on PSNR values. For clarity, the top-performing method is highlighted, and the second-best is underlined. Upon examination, it's evident that our proposed approach frequently attains the second-best rank in many scenarios. Notably, the gap between our proposed method and the top-ranked technique is often a marginal difference of approximately 0.2dB in terms of PSNR values, emphasizing the competitive performance of our approach to the state-of-the-art deep learning methods. In this case, it can be observed that the IPT method consistently ranks as the highest too. Considering these results, additional test was conducted on other datasets to validate and further explore these outcomes.

Table 5.3: Comparison with deep-learning SR in terms of PSNR values (dB).

Datasets	Scale	Deep Learning-based SR					
		SRCNN	VDSR	EDSR	RCAN	RDN	OISK-RK3
Set14	×2	32.45	33.05	33.92	34.12	34.01	33.94
	×3	29.30	29.78	30.52	30.65	30.57	30.57
	×4	27.50	28.02	28.80	28.87	28.92	28.81
Set5	×2	36.66	37.53	38.11	38.27	38.24	38.21
	×3	32.75	33.67	34.65	34.74	34.71	34.72
	×4	30.49	31.35	32.46	32.63	<u>32.64</u>	32.47
Urban100	×2	-	31.90	32.32	32.41	32.34	32.36
	×3	-	27.14	28.80	29.09	28.80	28.95
	×4	-	25.18	26.64	26.82	26.79	26.61
B100	×2	-	31.90	32.32	32.41	32.34	32.36
	×3	-	28.83	29.25	29.32	29.26	29.29
	×4	-	27.29	27.71	27.77	27.78	27.72
Datasets	Scale	Deep Learning-based SR					
		RNAN	SAN	HAN	IGNN	IPT	RDN-MetaSR
Set14	×2	33.87	34.07	34.16	34.06	**34.49**	33.98
	×3	30.52	30.59	30.67	30.66	**30.84**	30.54

Datasets	Scale						
	×4	28.86	28.83	28.90	28.85	<u>28.84</u>	28.78
Set5	×2	38.17	<u>38.31</u>	38.27	38.24	38.28	38.22
	×3	34.66	<u>34.75</u>	<u>34.75</u>	34.72	34.73	34.63
	×4	32.53	32.49	<u>32.64</u>	32.57	32.53	32.38
Urban100	×2	32.32	32.42	32.41	32.41	**33.93**	32.92
	×3	28.75	28.93	29.10	29.03	**29.65**	28.82
	×4	26.79	26.61	26.85	26.84	**27.30**	26.55
B100	×2	32.32	32.42	32.41	32.41	**32.52**	32.33
	×3	29.26	29.33	29.32	29.31	**29.42**	29.26
	×4	27.75	27.72	27.80	27.77	**27.84**	27.71

Datasets	Scale	Deep Learning-based SR				Proposed Work	
		RDN-LIIF	RDN-LTE	RDN-SRNO	DDL-EDSR		
Set14	×2	33.97	<u>34.29</u>	34.27	33.75	<u>34.29</u>	
	×3	30.53	<u>30.75</u>	30.71	30.45	<u>30.75</u>	
	×4	28.80	28.79	**28.97**	28.70	28.79	
Set5	×2	38.17	38.13	**38.32**	38.05	38.13	
	×3	34.68	34.67	**34.84**	34.50	34.67	
	×4	32.50	32.50	**32.69**	32.30	32.50	
Urban100	×2	32.87	<u>33.75</u>	33.33	32.41	<u>33.75</u>	
	×3	28.82	<u>29.58</u>	29.14	28.45	<u>29.58</u>	
	×4	26.68	<u>27.27</u>	26.98	26.34	<u>27.27</u>	
B100	×2	32.32	<u>32.47</u>	32.43	32.21	<u>32.47</u>	
	×3	29.26	<u>29.41</u>	29.37	29.16	<u>29.41</u>	
	×4	27.74	**27.84**	<u>27.83</u>	27.64	**27.84**	

5.2.2 Evaluation of Grayscale Datasets

After the initial tests, an extended evaluation was conducted to further probe the efficacy of the proposed hybrid method as shown in Table 5.4. For a more comprehensive understanding, grayscale datasets are utilized, comprising 200 images from the MMU iris dataset, 100 images sourced from the MIRFLICKR dataset, and 50 images from the NIH CT scan dataset. Generally, the MMU Iris, MIRFLICKR, and NIH CT Scan datasets are valuable resources for grayscale image processing research. The MMU Iris Dataset, hailing from Multimedia University, consists of high-quality iris images, widely used in biometric research for iris recognition and authentication, featuring diverse iris patterns and textures. The MIRFLICKR dataset, sourced from Flickr, encompasses a broad range of images, including landscapes, people, and urban scenes. Some of the grayscale images from this dataset were selected for testing purposes in this dissertation. Lastly, the NIH CT Scan dataset is a compilation of computed tomography scan images used primarily in medical imaging research. It provides detailed internal body views, crucial for developing and testing medical imaging technologies. Together, these datasets offer a rich spectrum of grayscale imagery, from biometric patterns to diverse general imagery and intricate medical scans, making them indispensable in advancing grayscale image SR and analysis.

Hence, comparative analyses between the proposed hybrid methods and IPT, based on PSNR metrics, revealed that the hybrid method surpasses the performance of IPT in most of the evaluated cases. In the color image tests, our method was slightly behind IPT by 0.2. However, for grayscale images, our hybrid approach surpassed IPT, showing its effectiveness in different image settings.

Table 5.4: Comparison with IPT in terms of PSNR value using grayscale datasets.

Grayscale Datasets	Scale	PSNR (dB)	
		IPT	Proposed Work
MMU Iris	×2	35.89	**36.01**
	×3	31.31	**31.38**
	×4	30.84	**30.87**
MIRFLICKR	×2	24.48	**24.51**
	×3	23.21	**23.24**
	×4	21.34	**21.39**
NIH CT Scan	×2	**28.92**	28.90
	×3	**24.81**	24.81
	×4	26.92	**26.93**

5.2.3   PSNR Assessment of Multiple Upscaling Iterations on Grayscale Images

In a series of experiments, the PSNR assessment for the multiple upscaling iterations on grayscale images was examined. Firstly, images were down sampled twice, upscaled back to their original size using both the KSVD_DR and IPT. Figure 5.6 shows the results for the MMU iris dataset. To further investigate the effectiveness of the methods, the proposed hybrid approach was introduced. In this, the first upscaling utilized the integration of IPT, and the KSVD_DR. Once this intermediate HR output was achieved using the hybrid method, SR was performed again using IPT and so on. Notably, when applied to the iris dataset, the hybrid approach achieved higher PSNR scores, suggesting a better restoration of image details and noise reduction compared to the individual use of KSVD_DR or IPT. This indicates that the hybrid method successfully leverages the complementary strengths of dictionary learning for detail preservation in grayscale images and deep learning's capability to reconstruct broader image contexts. Thus, the hybrid approach not only enhances the visual fidelity of the iris images but also potentially improves the performance of applications relying on precise iris recognition.

For the MIRFLICKR dataset, represented in Figure 5.7, a similar experimental process was applied. The results indicated that while the IPT method initially produced HR images with lower PSNR values than those generated by the KSVD_DR method, the introduction of the hybrid approach improved IPT's PSNR performance. This suggests that while IPT is generally strong in handling color information, it may not initially excel with the grayscale data in this dataset. The hybrid method, by incorporating the KSVD_DR technique, compensates for this by boosting the detail and quality in the grayscale domain, leading to an overall enhancement in SR outcomes as reflected by the higher PSNR values. This experiment underscores the value of combining different SR techniques to improve the quality of the resulting images.
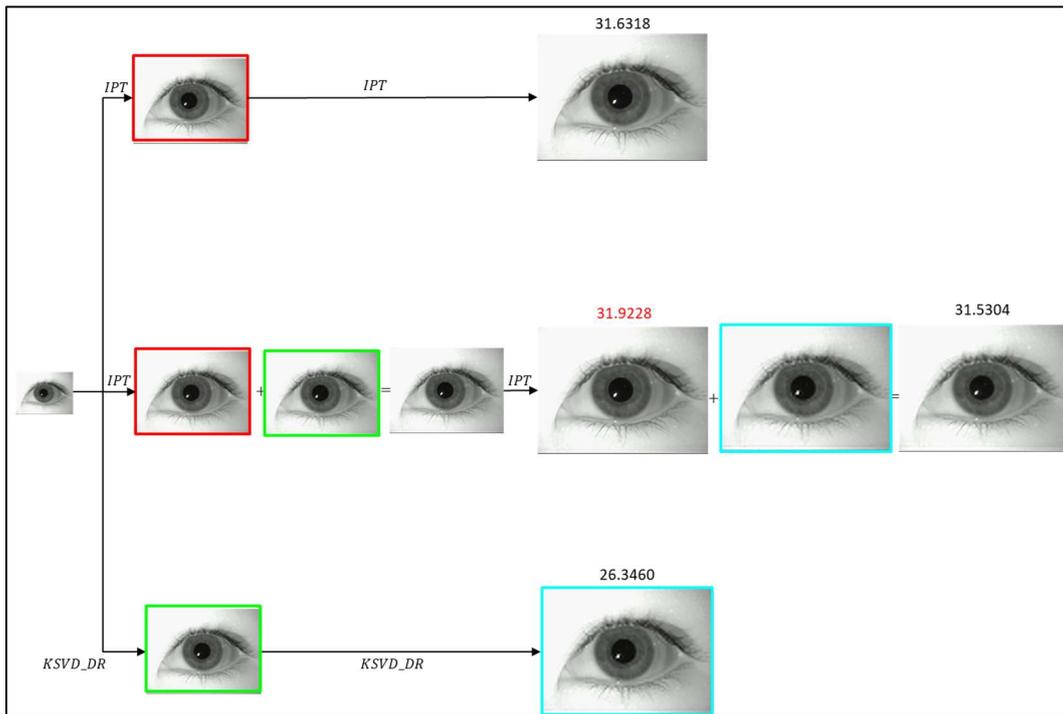
Figure 5.6: PSNR assessment of multiple upscaling iterations on MMU iris dataset.
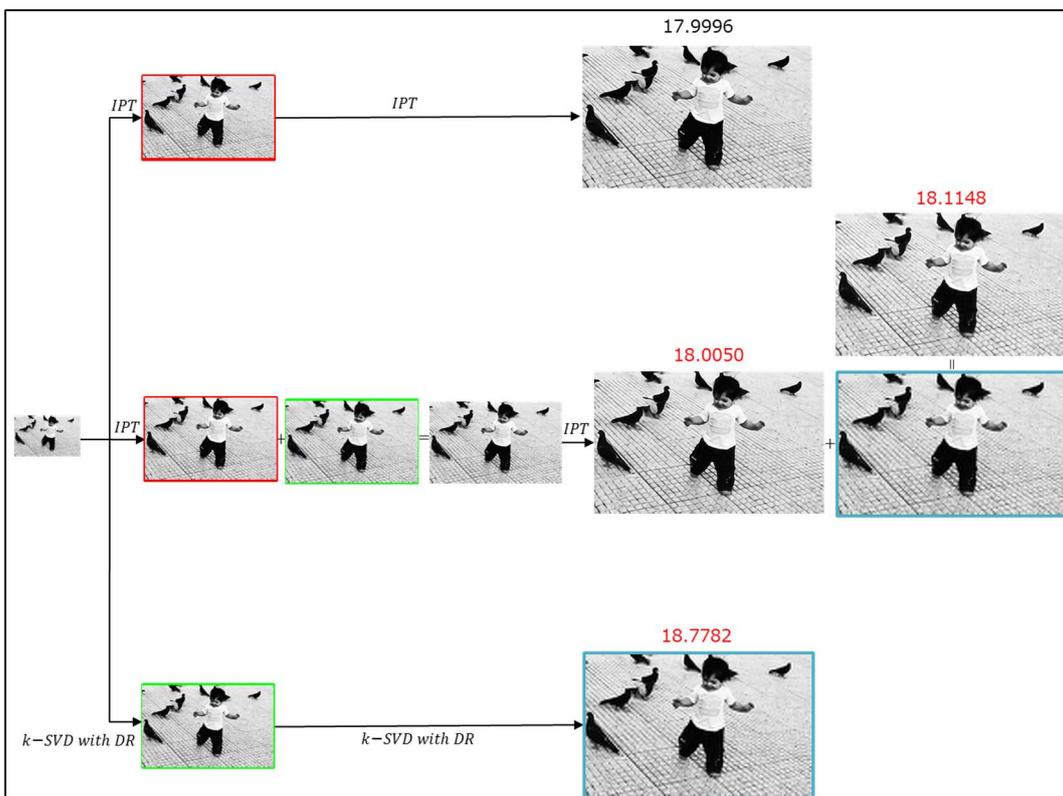

Figure 5.7: PSNR assessment of multiple upscaling iterations on MIRFLICKR database.

5.2.4   Finetuning of IPT on grayscale images

In this dissertation, the hybrid method proposed consists of two pathways: Pathway 1 utilizes the KSVD_DR algorithm, trained on a dataset of 69 images converted to grayscale, sourced from a publicly available dataset. On the other hand, Pathway 2 employs the IPT model, which is initially trained on the ImageNet dataset, predominantly composed of colour images. This distinction in training datasets is critical for understanding the performance outcomes of the study.

The findings of this study reveal that while the proposed hybrid method slightly trails the IPT model by a margin of 0.2 in terms of performance on general colour images, it exhibits better performance in the processing of grayscale images. To further validate these results, the IPT was fine-tuned using the same 69 grayscale images employed for KSVD_DR training for an upscale factor of $\times 2$. This process aimed to ensure a fair and direct comparison between the two methods under similar training conditions. Additionally, it's important to note that the IPT model has certain limitations regarding input image size, specifically requiring images of size larger than 64×64 pixels. This restriction poses a challenge in cases of upscale factors $\times 3$ and $\times 4$, where the LR image pairs of the dataset are smaller than the minimum size required. Table 5.5 shows the hyper-parameters for the finetuning process of IPT model done in this dissertation.

Table 5.5: Hyper-parameters for finetuning IPT model.

Hyper-parameters	Settings
Optimizer	ADAM
Epochs	50
Decay	0.0001
Momentum	0.9
Batch size	1
Learning rate	0.0004

Figure 5.8 illustrates the training process where the loss rapidly decreases and then stabilizes, with no indication of overfitting. This pattern reflects efficient fine-tuning approach where the model has effectively learned from the grayscale images. In the evaluation phase, a grayscale image from the MIRFLICKR database as shown in Figure 5.9 were utilized to assess the fine-tuned model's performance. The results, based on the PSNR and SSIM metrics, indicate that the proposed hybrid method, with a PSNR of 22.2552dB and SSIM of 0.7560, outperforms the fine-tuned IPT model, which achieved a PSNR of 22.0921dB and SSIM of 0.7451, even after the IPT was fine-tuned with grayscale images.
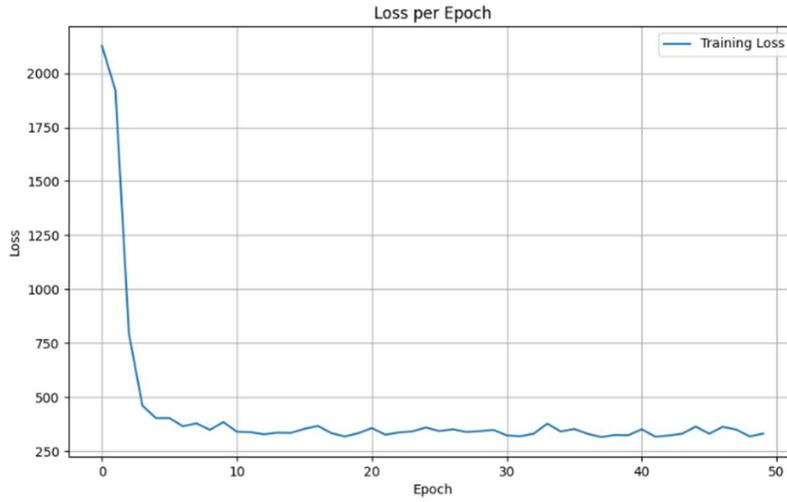
Figure 5.8: Training Loss per Epoch.

<table>
<tr><td>PSNR: 22.0921dB, SSIM: 0.7451</td><td>PSNR: **22.2552dB**, SSIM: **0.7560**</td></tr>
</table>

Figure 5.9: Output images computed from finetuned IPT model (left), and proposed hybrid method (right).

5.3 Summary

As a summary, the hybrid method successfully leverages the strengths of various algorithms, resulting in enhanced image quality and outperforming traditional dictionary learning-based SR techniques. A close comparison with deep-learning-based SR methods reveals that the proposed approach narrowly trails by a marginal 0.2 in PSNR values. However, it demonstrates superior performance on grayscale datasets, outshining the IPT model even after fine-tuning. Overall, the findings underscore the potential of the proposed hybrid method in SR tasks, showing promising results in both colour and grayscale image enhancement. Future research will focus on further optimizing this technique, particularly for improving its performance with colour images.

6. APPLICATION OF SUPER-RESOLUTION FOR THRIPS DETECTION AND CLASSIFICATION

6.1 Automatic Counting System of Thrips

Figure 6.1 shows the pipeline of the proposed technique for counting the number of thrips appeared in a single blue trap. The process started with an input image acquired from scanning an actual blue trap. The initial step involved preprocessing to isolate the region of interest (ROI). Following this, the dataset without SR was created by dividing the ROI image into two halves. Simultaneously, the dataset with SR was generated by enlarging the divided images by a factor of four using both the proposed SR method, KSVD_DR and hybrid method in this dissertation. These enhanced images were then segmented into patches. Both the datasets were uploaded into an annotation tool to mark the presence of thrips on the traps. Subsequently, the images were processed through the Yolov8 detection model for training and validation. The final stage produced output images with marked bounding boxes indicating the thrips, with the total count derived from the number of detected bounding boxes. All analyses were conducted on the server, utilizing the same specifications as detailed in Chapter 5.
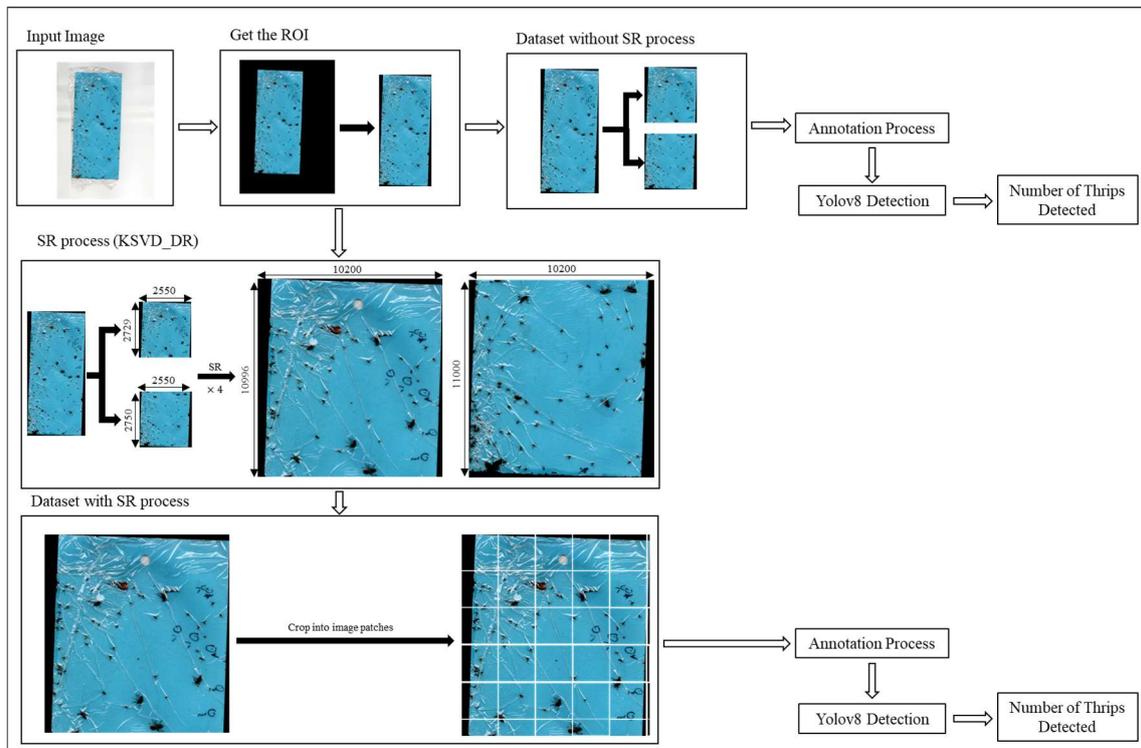

Figure 6.1: Proposed thrips detection pipeline.

6.1.1   Methodology

The input dataset comprises images of blue traps collected from orchards by farmers, amounting to a total of 34 scans. These blue traps, once retrieved, were scanned to produce HR digital images, each with an approximate resolution of 4000×7000 pixels and were stored in JPEG format. For illustrative purposes, the original image in Figure 6.1 features an example of these images, showcasing an original resolution of 4960×7015 pixels, which provides a high level of detail for further image processing and analysis tasks.

The initial step in processing involves the removal of the white background, as detailed in the preprocessing stage shown in Figure 6.2. This process started with loading the image and converting it to the HSV colour space, a method effective for colour segmentation. The specific lower and upper bounds were set to define the blue colour range within the HSV spectrum: lower blue at [90, 50, 50] and upper blue at [130, 255, 255], aimed at identifying varying shades of blue. A binary mask was then generated to pinpoint the blue pixels that fall within these predefined ranges. Contours were extracted from this mask, and the largest one, assumed to be the target blue insect plane, was identified. The bounding box for this contour is determined, giving the precise coordinates needed to accurately crop the relevant section from the original image. The cropped image, featuring the isolated blue insect plane, was then saved for further analysis.
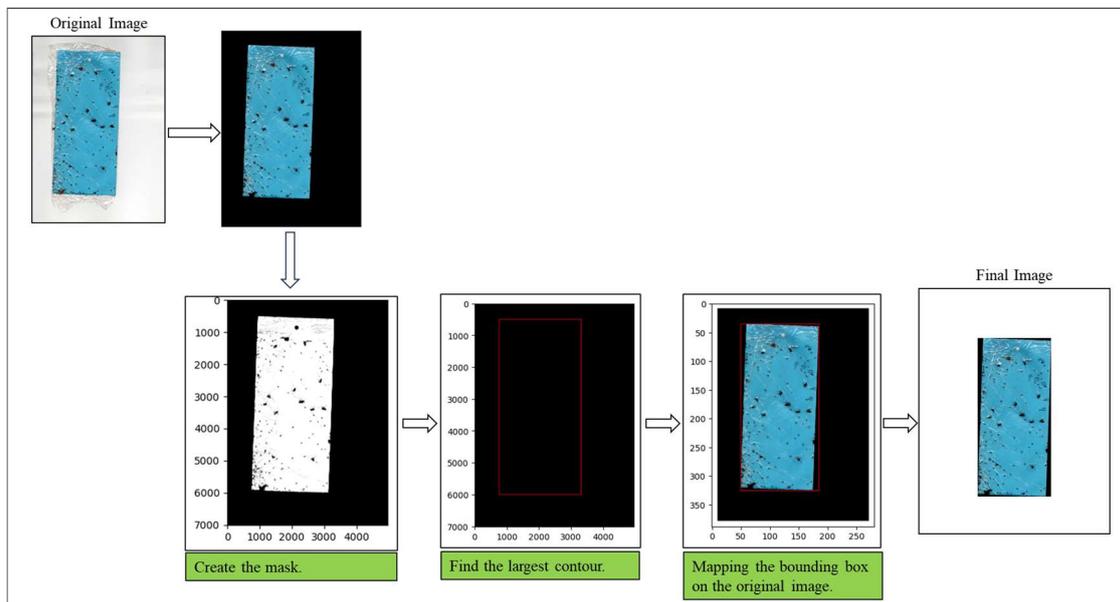

Figure 6.2: Pre-processing process on getting the ROI.

The images, once pre-processed, were further divided into top and bottom sections if their total size exceeds 11,100,000 pixels. This division results in a dataset referred to as the 'without SR process' dataset, where each cropped image is saved for subsequent use. Given the focus of this dissertation on

the effectiveness of the SR process, the KSVD_DR method was employed for its notable advantages, which include a reduced need for large image datasets and shorter training durations, attributed to its simpler structural design. An upscale factor of four was implemented throughout this study for the SR process. Using the same example provided in Figure 6.1, the original images with resolutions of 2550×2729 pixels and 2550×2750 pixels were transformed through the SR process. After applying the SR technique, the resolutions of these images were significantly increased to 10200×10996 pixels and 10200×11000 pixels, respectively. This substantial increase in resolution was a result of the upscale factor applied in the SR process, illustrating the effectiveness of the technique in enhancing the image size and detail.

Following the application of both the proposed SR method, KSVD_DR and hybrid method proposed in this dissertation, these enlarged images were then segmented into smaller patches. These segmented patches created after applying the super-resolution process was collectively referred to as the 'with SR' dataset. Table 6.1 shows the total number of images collected in datasets with and without applying SR process.

Table 6.1: Number of images in datasets with and without SR process.

Datasets	Number of images
Without SR	62
With SR	1938

Both the datasets were then prepared for annotation in the CVAT application, a tool used for labelling and annotating images (Sekachev et al., 2020). Annotation of thrips within the images is conducted based on the labels provided by the farmer on the physical blue traps, as depicted in Figure 6.3. Initially, the labelling process involves categorizing different classes of thrips, facilitating the use of these labels for classification tasks during the training process. However, since the primary objective of this study is to count the number of thrips present on the blue traps, the classification labels for various thrips classes are consolidated into a single class after exporting the data. This simplification allows for a focused approach towards quantifying thrips in the blue traps, aligning with the main goal of accurately determining their numbers.

Figure 6.3: Manually labelled thrips using the CVAT application (Sekachev et al., 2020), image in datasets without SR (left) and with SR (right).

The datasets, both with and without SR, were exported complete with labelled bounding boxes to facilitate the training and validation processes using the Yolov8 detection model. The distribution of images for training and validation in each dataset followed an 8:2 ratio, meaning 80% of the images were allocated for training while the remaining 20% were set aside for validation. This distribution is detailed in Table 6.2 and Table 6.3, which respectively present the breakdown for the datasets with and without the SR process.

Table 6.2: Number of images in dataset without SR process used for training and validation process.

	Number of images
Training	50
Validation	12

Table 6.3: Number of images in dataset with SR process used for training and validation process.

	Number of images
Training	1938
Validation	402

Finally, the Yolov8 model was subsequently executed, and its accuracy was assessed using the validation dataset. This evaluation focused on determining the effectiveness of the model in accurately detecting thrips within the images. Table 6.4 in this dissertation details the hyperparameters used during the training phase. This table serves as a reference point for understanding the specific configurations

and settings employed to train the Yolov8 model for this application in thrips detection.

Table 6.4: Hyper-parameters for training Yolov8 detection model.

Hyper-parameters	Settings
Pretrained Model	yolo8n.pt
Epochs	100
Image Size	640
Optimizer	Auto
Decay	0.0005
Momentum	0.937
Batch size	16
Learning rate	0.01

6.1.2   Results and Discussion

In this subsection, the validation results from the Yolov8 detection model to discern the impact of SR processing on the accuracy of thrips detection were examined. Table 6.5 presents the validation results comparing the accuracy of thrips detection with and without the application of SR processing, using the metric mAP50 (mean Average Precision at 50% Intersection Over Union). The accuracy for the 'Without SR' scenario is listed as 0.665, indicating that the model correctly identified and localized thrips in the images with a 66.5% accuracy at the 50% IoU threshold. In contrast, the 'With SR' scenario demonstrates superior performance over the 'Without SR' approach, with different methods yield varying levels of accuracy: the KSVD_DR method achieves the highest accuracy of 89.7% (0.897), while the hybrid method also shows improvement with an accuracy of 87.7% (0.877). This indicates that both the KSVR_DR and hybrid method outperform the 'Without SR' approach in detecting thrips, with KSVD_DR leading the way in terms of effectiveness with SR enhancement.

Table 6.5: Validation results of the case with and without SR.

	Without SR (mAP50)	With SR (mAP50)	
		KSVD_DR	**Hybrid Method**
Accuracy	0.665	**0.897**	0.877

Following the quantitative analysis of the Yolov8 detection model's performance, visual assessments are utilized to corroborate the numerical analysis of the Yolov8 detection model's performance. Figure 6.4, Figure 6.5 and Figure 6.6 showcase the model's predicted detections alongside the actual labelled data for batches without and with SR process, employing both KSVD_DR and

Hybrid Method, respectively. Figure 6.4 illustrates the model's predictions without SR, revealing areas where the model may have challenges in thrips localization and detection due to the original resolution of the images. Conversely, Figure 6.5 highlights the model's enhanced ability to detect thrips when SR is applied, as evidenced by the increased accuracy and the precision of the bounding boxes in relation to the labelled ground truth. Similarly, Figure 6.6 illustrates the effectiveness of the Hybrid Method in enhancing detection precision. These figures provide a visual confirmation of the numerical findings and underscore the benefits of SR in improving the model's detection capabilities for practical applications in thrips detection.

Finally, the validation results suggest that the implementation of SR techniques, particularly the KSVD_DR method due to its highest accuracy, can significantly enhance the accuracy of detecting small pests, specifically thrips, as explored in this dissertation. This enhancement is pivotal for farmers, as it aids in the more precise detection of thrips within fruit orchards, potentially improving pest management and crop health.
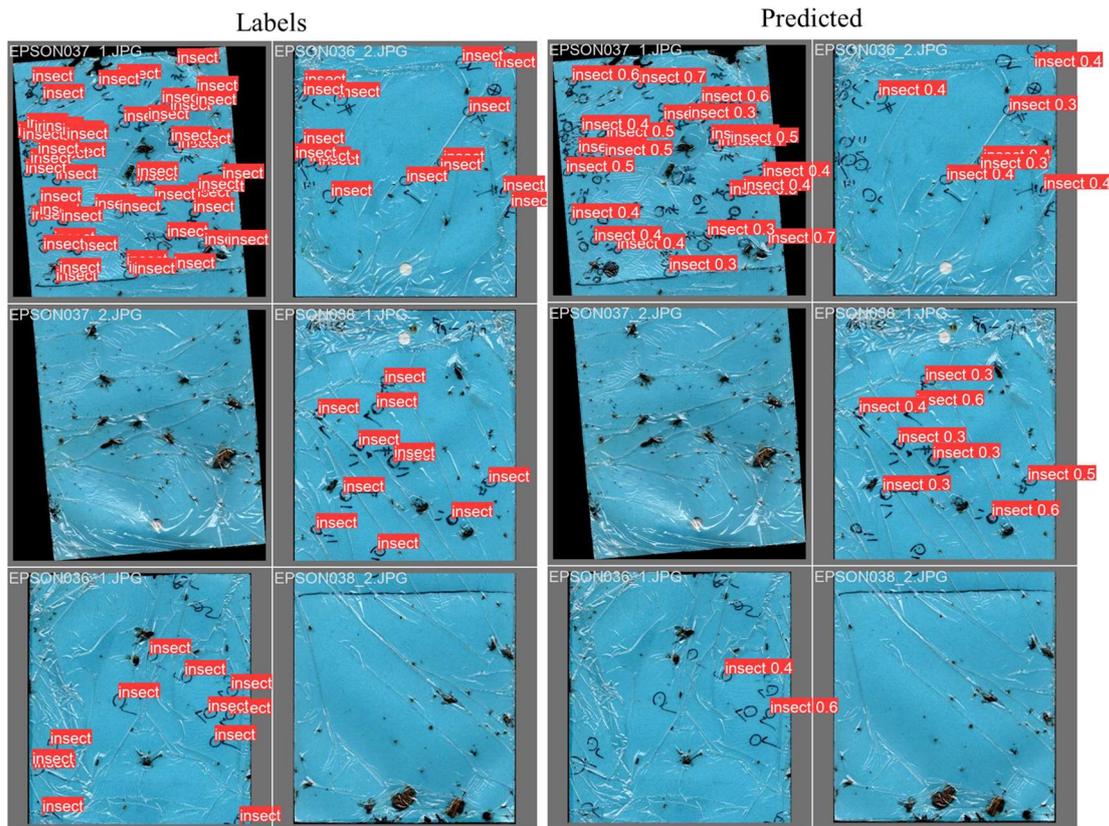
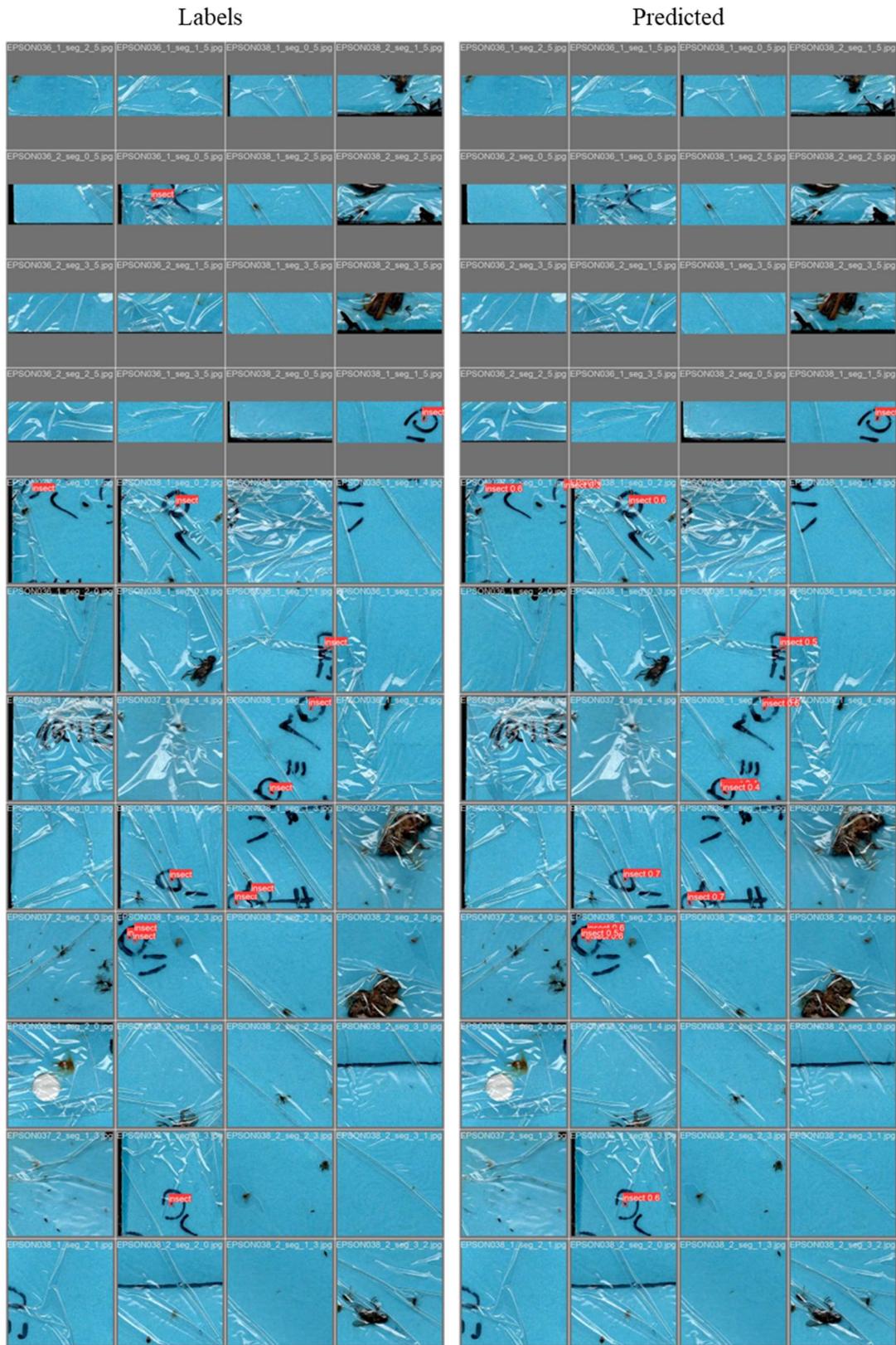**Figure 6.4: Predicted detection and the actual labels for batches (without SR).**

Labels   Predicted

Figure 6.5: Predicted detection and the actual labels for batches (with SR, KSVD_DR).
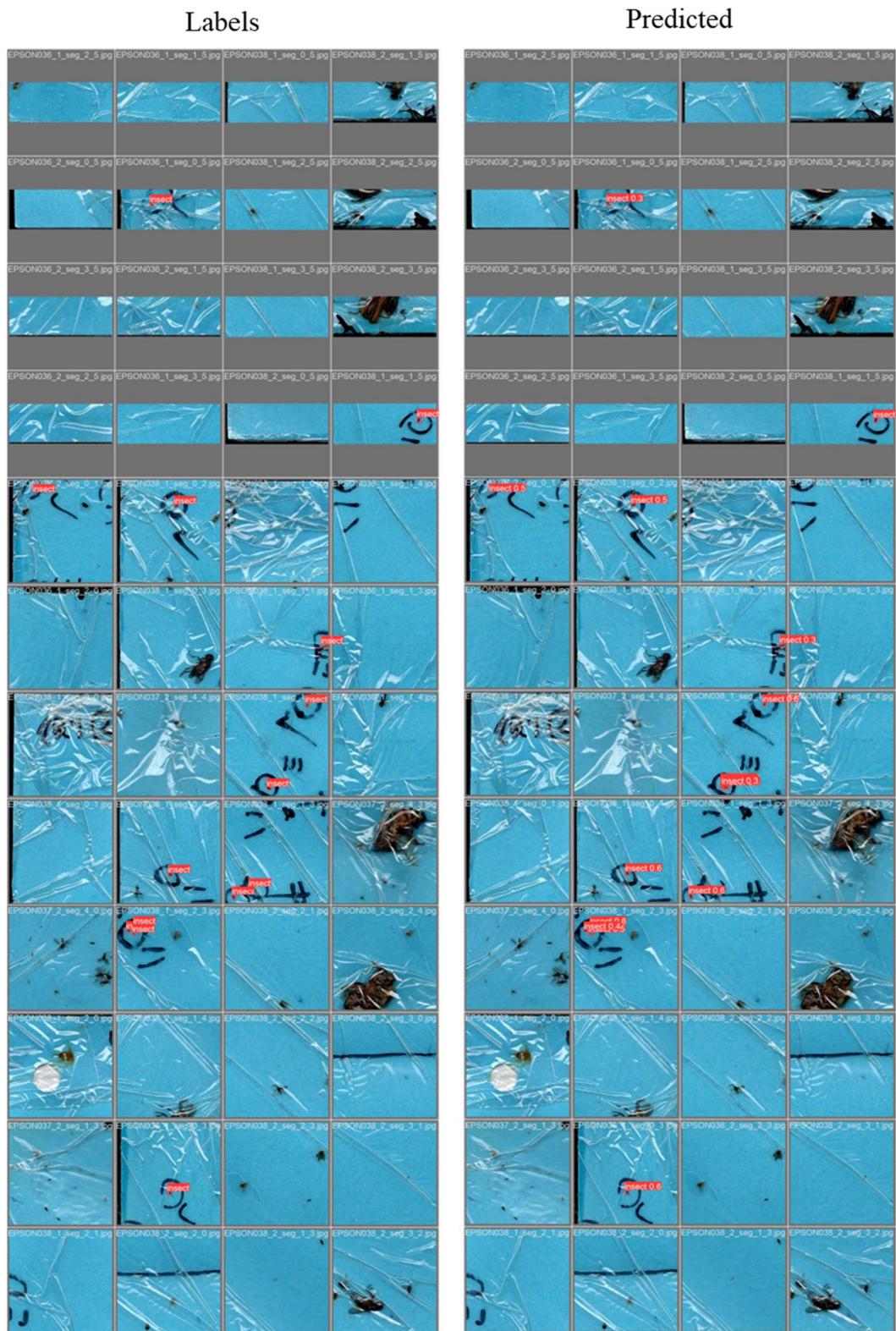
Figure 6.6: Predicted detection and the actual labels for batches (with SR, Hybrid Method).

6.1.3   Application of Automatic Counting System

Figure 6.7 depicted the user interface of a web application created using Streamlit, designed to assist farmers in detecting and counting thrips on blue trap images. The application boasted straightforward functionality: users could upload an image of a blue trap, which was then displayed on the screen. The system automatically initiated a counting process, detecting bounding boxes around the thrips present on the trap. The total count of detected thrips was prominently displayed at the bottom of the interface. The system's processing time was contingent upon the size of the image uploaded by the user. With an accuracy of 0.897 in the validation phase, specifically with images processed using the KSVD_DR method, the trained model embedded within the application provided farmers with a reliable tool to streamline the analysis and monitoring of thrip populations in their orchards. The KSVD_DR method, known for its practical applicability and shorter processing times, was chosen for this application to ensure efficient and effective monitoring.

Figure 6.7: User interface of automatic counting system for thrips detected on a single image.

6.2 Classification of Thrips

Figure 6.8 depicted the process employed for the classification of thrips found on blue traps. As with thrips detection, both datasets with and without SR process were loaded into an annotation tool to identify the presence of thrips on the traps. However, in this instance, the various types of thrips had already been labelled according to the classifications provided by the farmer, furnishing pre-existing information about the bounding boxes for different thrips classes. Consequently, the Yolov8 classification model was utilized, necessitating the extraction of different classes of thrips based on the labelled bounding boxes for training purposes. The concluding phase presented the aggregate count, which was derived from the quantity of detected thrips across each class. All analyses were conducted on the server, utilizing the same specifications as detailed in Chapter 5.
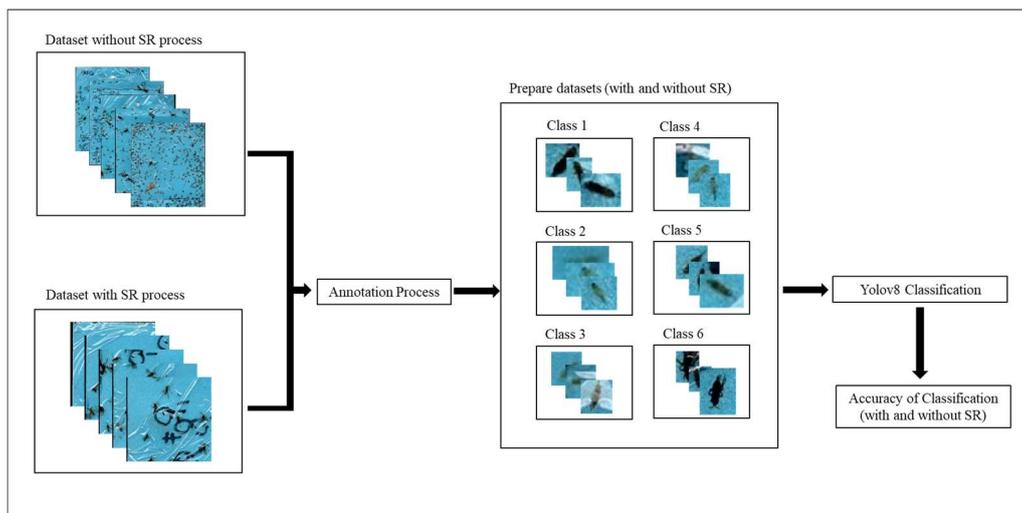
Figure 6.8: Proposed thrips classification pipeline.

6.2.1   Methodology

In the classification section, as the preprocessing has already been addressed in the detection section, the datasets with and without SR process were directly utilized. The thrips images for each class were extracted using the CVAT annotation tool based on the labelled bounding boxes. Table 6.6 lists the number of thrips per class. Blue traps, typically used to attract *Flower Thrips*, recorded the highest count. On the other end, *Yellow Tea Thrips* were the least in number, which was about ten times less than the *Flower Thrips*. While the quantity of thrips per class remained almost consistent in both the with and without SR datasets, the extracted image sizes differed; images from dataset without SR process were four times smaller than those from the dataset with SR process. This size difference

was a factor in the Yolov8 classification model's ability to distinguish between the various types of thrips.

Table 6.6: Number of thrips in each class in both datasets with and without SR process.

Types of Thrips	ID	Count	
		Dataset without SR	Dataset with SR
Flower Thrips	0	351	350
Honeysuckle Thrips	1	128	128
Western Flower Thrips	2	112	113
Yellow Tea Thrips	3	28	28
Onion Thrips	4	66	66
Other Thrips	5	54	54

The datasets, one enhanced with SR and the other without, underwent training and validation using the Yolov8 classification model. The specific distribution of images designated for training and validation is outlined in Table 6.7 and Table 6.8. The division between training and validation sets corresponds directly to their use in the detection phase; images used for validation in detection had their extracted bounding boxes serve as validation data in classification, and similarly for training data. This method ensures consistency in the evaluation of the model's performance across both detection and classification tasks.

Table 6.7: Distribution of each type of thrips in dataset without SR for training and validation.

	Types of Thrips	Number of images
Training	*Flower Thrips*	312
	Honeysuckle Thrips	118
	Western Flower Thrips	103
	Yellow Tea Thrips	23
	Onion Thrips	57
	Other Thrips	32
Validation	*Flower Thrips*	39
	Honeysuckle Thrips	10
	Western Flower Thrips	9
	Yellow Tea Thrips	5

	Onion Thrips	9
	Other Thrips	22

Table 6.8: Distribution of each type of thrips in dataset with SR for training and validation.

	Types of Thrips	Number of images
Training	*Flower Thrips*	311
	Honeysuckle Thrips	117
	Western Flower Thrips	101
	Yellow Tea Thrips	23
	Onion Thrips	52
	Other Thrips	32
Validation	*Flower Thrips*	41
	Honeysuckle Thrips	12
	Western Flower Thrips	8
	Yellow Tea Thrips	4
	Onion Thrips	8
	Other Thrips	22

The Yolov8 classification model was then implemented, with its performance evaluated using the validation dataset. The primary goal of this evaluation was to assess the model's accuracy in classifying thrips within images. To provide a comprehensive understanding of the model's training process, Table 6.9 in this dissertation enumerates the hyperparameters utilized.

Table 6.9: Hyper-parameters for training Yolov8 classification model.

Hyper-parameters	Settings
Pretrained Model	yolo8cls.pt
Epochs	100
Image Size	640
Optimizer	Auto
Decay	0.0005
Momentum	0.937
Batch size	16
Learning rate	0.01

6.2.2   Results and Discussion

In this subsection, the focus shifts to the classification aspect of thrips using the Yolov8 classification model, evaluating how SR processing influences the model's classification accuracy. The validation results, as outlined in Table 6.10, comparing the Top-1 accuracy of thrips classification with and without the application of SR processing. For the 'Without SR' scenario, the Top-1 accuracy is recorded as 0.585, indicating that the model correctly classifies thrips with 58.5% accuracy. Conversely, the 'With SR' scenario shows a discernible improvement: the KSVD_DR method enhances accuracy to 65.3% (0.653), while the hybrid method, although slightly less effective, still improves performance to 64.2% (0.642) compared to the baseline. These results highlight the beneficial impact of SR processing on the model's classification accuracy, with the KSVD_DR method emerging as the most effective for accurately detecting thrips, followed closely by the hybrid method.

Table 6.10: Validation results of the case with and without SR.

	Without SR	With SR	
		KSVD_DR	**Hybrid Method**
Accuracy	0.585	**0.653**	0.642

Figure 6.9, Figure 6.10 and Figure 6.11 display the validation results of predicted classes for a single batch and their corresponding actual labels in datasets without and with the application of the SR process, specifically KSVD_DR and hybrid method as discussed in this dissertation. The figures predominantly show predictions classifying thrips as *Flower Thrip*, labeled as 'class 0'. This skew in classification is attributed to the data imbalance within the datasets, where a majority of the images represent the *Flower Thrips* class. To address this issue and improve classification accuracy for other thrips classes, it is suggested to augment the dataset with additional data for underrepresented classes. This expansion of the dataset is expected to rectify the current imbalance and enhance the overall effectiveness of the classification model.
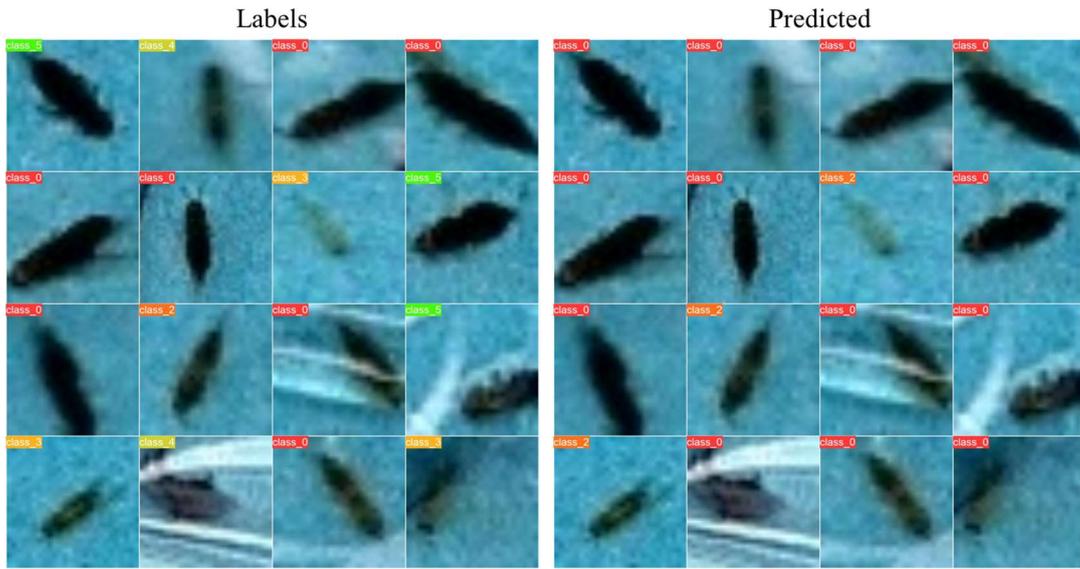
Figure 6.9: Predicted class and the actual labels for one batch (without SR).
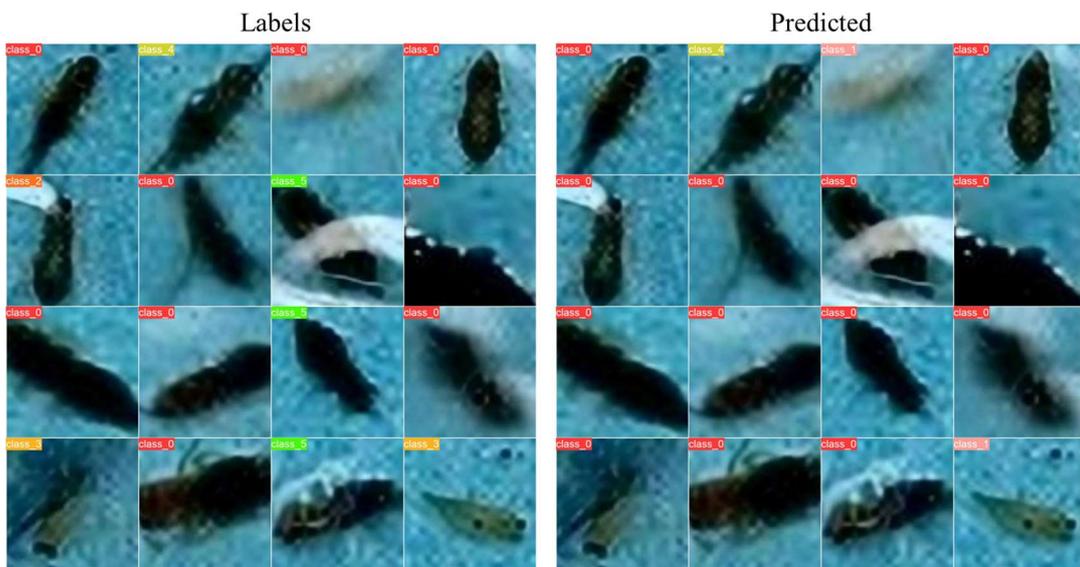

Figure 6.10: Predicted class and the actual labels for one batch (with SR, KSVD_DR).
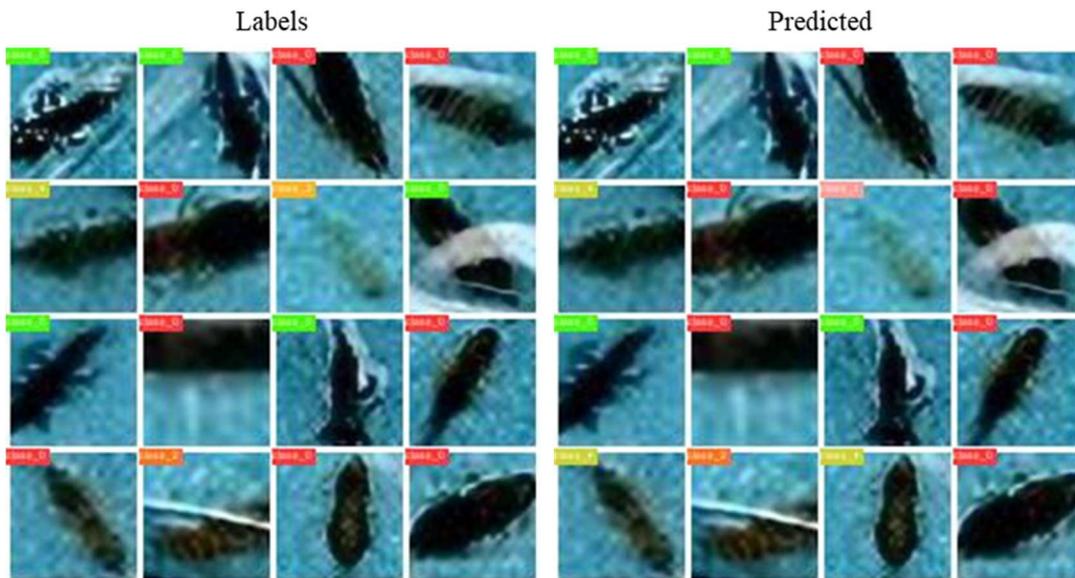
Figure 6.11: Predicted class and the actual labels for one batch (with SR, Hybrid Method).

6.2.3　Limitations and Future Works

For the automatic counting system of thrips which involved the detection of thrips, the current detection model is designed specifically for blue traps. This specialization presents a limitation as it restricts the model's application solely to these traps, potentially reducing its effectiveness in scenarios involving traps of other colours, or in diverse environmental conditions. To address this, future work will involve incorporating datasets for yellow traps. Expanding the model's training to include this additional variety will not only enhance its adaptability but also improve its utility across different agricultural settings, making it more versatile and broadly applicable.

On the other hand, a notable limitation in the thrips classification model is the issue of data imbalance. The existing dataset predominantly favours certain classes of thrips, leading to a potential bias in the model's classification accuracy. This imbalance could affect the model's overall performance, particularly in its ability to generalize across different thrip types. In future developments, there will be a focus on augmenting the dataset with a more balanced representation of each thrips class. By diversifying the dataset, the model's classification capabilities can be significantly enhanced, ensuring a more accurate and comprehensive understanding of thrips populations in various agricultural contexts.

7. CONCLUSION AND FUTURE WORK

Background: Image Super-Resolution (SR) is a pivotal technique designed to convert Low-Resolution (LR) images into High-Resolution (HR) counterparts by extrapolating omitted high-frequency components. SR encompasses two primary categories: Multiple Image Super-Resolution (MISR) and Single Image Super-Resolution (SISR). MISR involves merging several LR images into one HR image, rich in detail and low in noise or distortion. This process requires precise alignment of the LR images, a procedure that can be time-consuming due to the need for subpixel precision. SISR, on the other hand, generates an HR image from a single LR image and is more adaptable to various applications. However, SISR faces challenges in inferring detailed information from limited data, making it a complex task. The methodologies in image SR are classified into three main groups: interpolation-based, reconstruction-based, and learning-based methods. Interpolation-based SR methods, such as nearest neighbour, bilinear, and bicubic, are basic techniques that estimate high-frequency details from adjacent pixel values but often lead to blurred HR images. Reconstruction-based methods model the degradation process of LR images and reconstruct the HR images by inversely solving this model. While they generally surpass interpolation-based methods in performance, they require precise knowledge of the image degradation, which is not always available, limiting their effectiveness in certain scenarios. Learning-based SR, incorporating machine learning techniques to predict HR images from LR ones, represents a more advanced approach. This category includes dictionary learning-based SR, deep learning-based SR, Support Vector Machine (SVM)-based SR, and random-forest based SR. The focus of this dissertation is on learning-based SISR which involve in implementing the SISR image reconstruction algorithm using the dictionary-learning based SR with sparse representation algorithms, for showing that the efficiency of using dictionary learning algorithm with suitable sparse representation algorithms in producing image with enhanced performance in noise too. Additionally, a custom modification in the dictionary learning-based method was implemented to improve computational efficiency. While dictionary learning methods like k-SVD offer effective contrast enhancements in grayscale images, its limitation is in colour image processing. Conversely, deep learning-based SR methods can produce high-quality HR images but may struggle with unseen data or varying image types. This dissertation introduces a hybrid SR approach that combines the quick contrast enhancement of dictionary learning methods for grayscale images with the colour image processing strengths of deep learning techniques. This combination aims to provide a more comprehensive SR solution, improving image quality for both grayscale and colour images and ensuring robust performance even with new data types. The result is a versatile SR technique with broad applications in diverse imaging fields. Also, one of the significant applications of the proposed

SR technique is in the agricultural sector, particularly in the detection and classification of thrips pests. This approach aims to improve image resolution and accuracy, enabling more effective identification and categorization of these pests. This dissertation proposes an end-to-end automatic counting system for farmers, integrating the developed SR technique to enhance the precision of pest detection and classification, streamlining, and improving the efficiency of the monitoring process in agricultural practices.

Proposal: This dissertation focused on four contributions. Firstly, this dissertation has analysed the integration of dictionary learning with sparse representation algorithms to produce higher quality denoised images, establishing a new benchmark in image enhancement. Specifically, the study has shown that the combination of dictionary learning with the Douglas-Rachford method outperforms other algorithms in achieving superior PSNR and SSIM values for denoised images. Additionally, a SISR system, named KSVD_DR, has been developed, leveraging the Douglas-Rachford algorithm and k-Singular Value Decomposition (k-SVD) technique for increased computational efficiency. This proposed system is notable for its speed and accuracy, particularly in grayscale image enhancement, where it surpasses the capabilities of existing state-of-the-art SR methods like the IPT. A novel hybrid image reconstruction algorithm, which unites dictionary learning-based methods with transformer-based deep learning, constitutes another significant contribution. This innovative approach ensures high-quality image resolution for various applications, covering both general colour and grayscale images, and outshines traditional and deep learning-based SR methods in terms of visual quality and quantitative performance metrics. Lastly, the application of the KSVD_DR method to the agricultural sector underscores the practicality of the research. By efficiently detecting and classifying thrips, a common agricultural pest, the research offers a robust solution to pest management, demonstrating the viability of advanced image processing techniques in real-world scenarios. In conclusion, the proposal put forth in this dissertation not only advances the technological aspects of image processing but also provides tangible solutions to pressing agricultural challenges. By bridging the gap between theoretical research and practical application, the contributions made herein pave the way for further advancements in the detection and classification of pests, with the potential to significantly impact agricultural practices.

Limitation and Future Direction: As shown in the results part, the final image slightly lags deep-learning-based SR methods by a small margin in PSNR values in colour images, but it excels in processing grayscale images, even surpassing the IPT model after fine-tuning. These findings highlight the potential of the hybrid method in enhancing both colour and grayscale images, setting a direction for future optimization, especially in improving colour image processing. In the realm of agricultural pest detection and classification, the research faced certain limitations. The current detection model is tailored specifically for blue traps, limiting its broader application. Future work aims to expand the dataset to include yellow traps, enhancing the model's versatility in different agricultural settings.

Additionally, the classification model of thrips encountered challenges due to data imbalance, favouring certain thrips classes and potentially biasing classification accuracy. Future developments will concentrate on creating a more balanced dataset for each thrips class, ensuring more accurate and comprehensive pest management solutions in agriculture. This approach is expected to significantly improve the model's ability to accurately classify different types of thrips, aiding in effective pest control and management strategies.

ACKNOWLEDGMENT

REFERENCES

Aharon, M., Elad, M., & Bruckstein, A. (2006). K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on signal processing*, *54*(11), 4311-4322.

Amarathunga, D. C., Ratnayake, M. N., Grundy, J., & Dorin, A. (2022). Fine-grained image classification of microscopic insect pest species: Western Flower thrips and Plague thrips. *Computers and Electronics in Agriculture*, *203*, 107462.

Ayas, S., & Ekinci, M. (2020). Single image super resolution using dictionary learning and sparse coding with multi-scale and multi-directional Gabor feature representation. *Information Sciences*, *512*, 1264-1278.

Bevilacqua, M., Roumy, A., Guillemot, C., & Alberi-Morel, M. L. (2012). Low-complexity single-image super-resolution based on nonnegative neighbor embedding.

Cai, J., Zeng, H., Yong, H., Cao, Z., & Zhang, L. (2019). Toward real-world single image super-resolution: A new benchmark and a new model. Proceedings of the IEEE/CVF International Conference on Computer Vision,

Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020). End-to-end object detection with transformers. European conference on computer vision,

Chang, H., Yeung, D.-Y., & Xiong, Y. (2004). Super-resolution through neighbor embedding. Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.,

Chatterjee, P., Mukherjee, S., Chaudhuri, S., & Seetharaman, G. (2009). Application of Papoulis–Gerchberg method in image super-resolution and inpainting. *The computer journal*, *52*(1), 80-89.

Chen, H., Wang, Y., Guo, T., Xu, C., Deng, Y., Liu, Z., Ma, S., Xu, C., Xu, C., & Gao, W. (2021). Pre-trained image processing transformer. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition,

Chen, Y., Liu, S., & Wang, X. (2021). Learning continuous image representation with local implicit image function. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition,

Dai, T., Cai, J., Zhang, Y., Xia, S.-T., & Zhang, L. (2019). Second-order attention network for single image super-resolution. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition,

Elad, M., & Aharon, M. (2006). Image denoising via sparse and redundant representations over learned

dictionaries. *Ieee Transactions on Image Processing*, *15*(12), 3736-3745.

Elad, M., & Feuer, A. (1996). Super-resolution reconstruction of an image. Proceedings of 19th convention of electrical and electronics engineers in Israel,

Gonzalez, R. C. (2009). *Digital image processing*. Pearson education india.

Goyal, M., Lather, Y., & Lather, V. (2015). Analytical relation & comparison of PSNR and SSIM on babbon image and human eye perception using matlab. *International Journal of Advanced Research in Engineering and Applied Sciences*, *4*(5), 108-119.

He, X., Mo, Z., Wang, P., Liu, Y., Yang, M., & Cheng, J. (2019). Ode-inspired network design for single image super-resolution. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,

Hu, X., Mu, H., Zhang, X., Wang, Z., Tan, T., & Sun, J. (2019). Meta-SR: A magnification-arbitrary network for super-resolution. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition,

Huang, J.-B., Singh, A., & Ahuja, N. (2015). Single image super-resolution from transformed self-exemplars. Proceedings of the IEEE conference on computer vision and pattern recognition,

Hui, Z., Gao, X., & Wang, X. (2020). Lightweight image super-resolution with feature enhancement residual network. *Neurocomputing*, *404*, 50-60.

Huiskes, M. J., & Lew, M. S. (2008). The mir flickr retrieval evaluation. Proceedings of the 1st ACM international conference on Multimedia information retrieval,

Hulagappa, T., Baradevanal, G., Surpur, S., Raghavendra, D., Doddachowdappa, S., Shashank, P. R., Mallaiah, K. K., & Bedar, J. (2022). Diagnosis and potential invasion risk of Thrips parvispinus under current and future climate change scenarios. *PeerJ*, *10*, e13868.

Ji, N., Milkie, D. E., & Betzig, E. (2010). Adaptive optics via pupil segmentation for high-resolution imaging in biological tissues. *Nature methods*, *7*(2), 141-147.

Jiang, C., Zhang, Q., Fan, R., & Hu, Z. (2018). Super-resolution CT image reconstruction based on dictionary learning and sparse representation. *Scientific reports*, *8*(1), 8799.

Jiang, X., Wang, N., Xin, J., Yang, X., Yu, Y., & Gao, X. (2020). Image super-resolution via multi-view information fusion networks. *Neurocomputing*, *402*, 29-37.

Kim, J., Lee, J. K., & Lee, K. M. (2016). Accurate image super-resolution using very deep convolutional networks. Proceedings of the IEEE conference on computer vision and pattern recognition,

Lee, J., & Jin, K. H. (2022). Local texture estimator for implicit representation function. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition,

Lehmann, T. M., Gonner, C., & Spitzer, K. (1999). Survey: Interpolation methods in medical image processing. *IEEE transactions on medical imaging*, *18*(11), 1049-1075.

Li, S., & Qi, H. (2015). A Douglas–Rachford splitting approach to compressed sensing image recovery

using low-rank regularization. *Ieee Transactions on Image Processing*, *24*(11), 4240-4249.

Lim, B., Son, S., Kim, H., Nah, S., & Mu Lee, K. (2017). Enhanced deep residual networks for single image super-resolution. Proceedings of the IEEE conference on computer vision and pattern recognition workshops,

Lin, C.-H., Chen, P.-H., Lin, C., Chen, Y.-C., Huang, M.-J., & Liu, W.-M. (2021). Automatic detection and counting of small yellow thrips on lotus leaf back based on YOLO combined with VDSR and DPSR network. Thirteenth International Conference on Digital Image Processing (ICDIP 2021),

Liu, T., Cheng, J., & Tan, S. (2023). Spectral Bayesian Uncertainty for Image Super-Resolution. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,

Majumdar, A. (2009). *Greedy Algorithms promoting Group Sparsity V2.* . https://www.mathworks.com/matlabcentral/fileexchange/22835-greedy-algorithms-promoting-group-sparsity-v2

Malczewski, K. (2020). Super-Resolution with compressively sensed MR/PET signals at its input. *Informatics in Medicine Unlocked*, *18*, 100302.

Martin, D., Fowlkes, C., Tal, D., & Malik, J. (2001). A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001,

Mehle, N., & Trdan, S. (2012). Traditional and modern methods for the identification of thrips (Thysanoptera) species. *Journal of Pest Science*, *85*, 179-190.

Milanfar, P. (2017). *Super-resolution imaging*. CRC press.

MultiMedia University Iris dataset for Biometric Attandance System. (2020). MMU Iris Dataset. https://www.kaggle.com/datasets/naureenmohammad/mmu-iris-dataset

Nasrollahi, K., & Moeslund, T. B. (2014). Super-resolution: a comprehensive survey. *Machine vision and applications*, *25*, 1423-1468.

Nguyen, K., Fookes, C., Sridharan, S., Tistarelli, M., & Nixon, M. (2018). Super-resolution for biometrics: A comprehensive survey. *Pattern Recognition*, *78*, 23-42.

Niu, B., Wen, W., Ren, W., Zhang, X., Yang, L., Wang, S., Zhang, K., Cao, X., & Shen, H. (2020). Single image super-resolution via a holistic attention network. Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XII 16,

Pandey, G., & Ghanekar, U. (2018). A compendious study of super-resolution techniques by single image. *Optik*, *166*, 147-160.

Park, S. C., Park, M. K., & Kang, M. G. (2003). Super-resolution image reconstruction: a technical overview. *IEEE signal processing magazine*, *20*(3), 21-36.

Pati, Y. C., Rezaiifar, R., & Krishnaprasad, P. S. (1993). Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. Proceedings of 27th

Asilomar conference on signals, systems and computers,

Pham, C.-H., Tor-Díez, C., Meunier, H., Bednarek, N., Fablet, R., Passat, N., & Rousseau, F. (2019). Multiscale brain MRI super-resolution using deep 3D convolutional networks. *Computerized Medical Imaging and Graphics*, *77*, 101647.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. Proceedings of the IEEE conference on computer vision and pattern recognition,

Redmon, J., & Farhadi, A. (2017). YOLO9000: better, faster, stronger. Proceedings of the IEEE conference on computer vision and pattern recognition,

Rubinstein, R., Zibulevsky, M., & Elad, M. (2009). Double sparsity: Learning sparse dictionaries for sparse signal approximation. *IEEE Transactions on signal processing*, *58*(3), 1553-1564.

Sekachev, B., Manovich, N., Zhiltsov, M., Zhavoronkov, A., Kalinin, D., & Hoff, B. (2020). Computer vision annotation tool (CVAT). *Accessed: Jun*, *21*, 2021.

Thévenaz, P., Blu, T., & Unser, M. (2000). Image interpolation and resampling. *Handbook of medical imaging, processing and analysis*, *1*(1), 393-420.

Tian, J., & Ma, K.-K. (2011). A survey on super-resolution imaging. *Signal, Image and Video Processing*, *5*, 329-342.

Timofte, R., De Smet, V., & Van Gool, L. (2013). Anchored neighborhood regression for fast example-based super-resolution. Proceedings of the IEEE international conference on computer vision,

Tropp, J. A., & Gilbert, A. C. (2007). Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on information theory*, *53*(12), 4655-4666.

Vb, S. K. (2020). Perceptual image super resolution using deep learning and super resolution convolution neural networks (SRCNN). *Intelligent Systems and Computer Technology*, *37*(3).

Wang, Z., Liu, D., Yang, J., Han, W., & Huang, T. (2015). Deep networks for image super-resolution with sparse prior. Proceedings of the IEEE international conference on computer vision,

Wei, M., & Zhang, X. (2023). Super-Resolution Neural Operator. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,

Yan, K., Wang, X., Lu, L., & Summers, R. M. (2018). DeepLesion: automated mining of large-scale lesion annotations and universal lesion detection with deep learning. *Journal of medical imaging*, *5*(3), 036501-036501.

Yang, J., Wright, J., Huang, T., & Ma, Y. (2008). Image super-resolution as sparse representation of raw image patches. 2008 IEEE conference on computer vision and pattern recognition,

Yang, J., Wright, J., Huang, T. S., & Ma, Y. (2010). Image super-resolution via sparse representation. *Ieee Transactions on Image Processing*, *19*(11), 2861-2873.

Yue, L., Shen, H., Li, J., Yuan, Q., Zhang, H., & Zhang, L. (2016). Image super-resolution: The techniques, applications, and future. *Signal Processing*, *128*, 389-408.

Zeyde, R., Elad, M., & Protter, M. (2012). On single image scale-up using sparse-representations. Curves and Surfaces: 7th International Conference, Avignon, France, June 24-30, 2010, Revised Selected Papers 7,

Zhang, K., Zuo, W., & Zhang, L. (2018). Learning a single convolutional super-resolution network for multiple degradations. Proceedings of the IEEE conference on computer vision and pattern recognition,

Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B., & Fu, Y. (2018). Image super-resolution using very deep residual channel attention networks. Proceedings of the European conference on computer vision (ECCV),

Zhang, Y., Li, K., Li, K., Zhong, B., & Fu, Y. (2019). Residual non-local attention networks for image restoration. *arXiv preprint arXiv:1903.10082*.

Zhang, Y., Tian, Y., Kong, Y., Zhong, B., & Fu, Y. (2020). Residual dense network for image restoration. *IEEE transactions on pattern analysis and machine intelligence*, *43*(7), 2480-2495.

Zhang, Z., Xu, Y., Yang, J., Li, X., & Zhang, D. (2015). A survey of sparse representation: algorithms and applications. *Ieee Access*, *3*, 490-530.

Zhou, S., Zhang, J., Zuo, W., & Loy, C. C. (2020). Cross-scale internal graph neural network for image super-resolution. *Advances in neural information processing systems*, *33*, 3499-3509.

アザミウマ5種. (2022).

アザミウマ類発生予察について. (2022).