Compression and enhancement of medical images and video

A Dissertation

submitted to the designated
by the Assembly
of the Department of Computer Science and Engineering
Examination Committee

by

Stamatia-Christina Zerva

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

University of Ioannina
School of Engineering
Ioannina 2024

Advisory Committee:

- **Lisimachos Paul Kondi**, Professor, Department of Computer Science and Engineering, University of Ioannina
- **Christoforos Nikou**, Professor, Department of Computer Science and Engineering, University of Ioannina
- **Konstadinos Parsopoulos**, Professor, Department of Computer Science and Engineering, University of Ioannina

Examining Committee:

- **Lisimachos Paul Kondi**, Professor, Department of Computer Science and Engineering, University of Ioannina
- **Christoforos Nikou**, Professor, Department of Computer Science and Engineering, University of Ioannina
- **Konstadinos Parsopoulos**, Professor, Department of Computer Science and Engineering, University of Ioannina
- **Michail Vrigkas**, Assis. Professor, Department of Communication and Digital Media, University of Western Macedonia
- **Alexandros Tzallas**, Assoc. Professor, Department of Informatics and Telecommunications, University of Ioannina
- **Angeliki Katsenou**, Senior lecturer, School of Computer Science, University of Bristol, UK
- **Aggelos K. Katsaggelos**, Professor, Department of Electrical and Computer Engineering, Northwestern University, USA

Dedication

To everyone who believed in me, even when I doubted myself.

Acknowledgements

This PhD dissertation represents not only my work at the keyboard but also a milestone in more than a few years of collaborative work. Several people have contributed to this accomplishment, and I would like to acknowledge their support and encouragement.

First and foremost, I would like to express my deepest gratitude to my advisor, Professor Lisimachos Paul Kondi. Your guidance, encouragement, and invaluable feedback have been crucial to the completion of this dissertation. Your patience and understanding during challenging times have meant a great deal to me, and your expertise has profoundly shaped my academic journey.

I am also grateful to the other two members of my dissertation committee, Professor Christoforos Nikou, and Professor Konstadinos Parsopoulos. Your insightful comments and constructive criticism have been instrumental in refining my research and improving this work.

My heartfelt thanks go to my family. To my parents, Theofanis and Angeliki, your unconditional love and support have been my foundation. You have always believed in me and encouraged me to pursue my dreams. To my siblings, Cleo, Evangelia and Konstandinos, your encouragement and support have been invaluable.

A special note of appreciation goes to Associate Professor Alexandros Tzallas, whose altruistic support and guidance have been a beacon of light throughout my academic journey. Your selfless dedication and willingness to help without seeking personal gain have inspired me deeply and have been instrumental in my success.

Special thanks to my friends and my beloved ones. The stimulating discussions, collaborative work, support and shared experiences have made this journey enjoyable and intellectually rewarding.

I would like to acknowledge the financial support from University of Ioannina, which made this research possible. Your support has been fundamental in allowing me to pursue and complete my research.

To those who believed in me and provided their unwavering support, thank you from the bottom of my heart. Your faith in me has been a driving force in this accomplishment.

To those who doubted or did not believe in me, I also extend my gratitude. Your skepticism fueled my determination to prove myself and achieve my goals.

To everyone who has been part of this journey, directly or indirectly, thank you. This achievement is as much yours as it is mine.

Table of Contents

List of Figures

List of Tables

List of Algorithms

Glossary

ADMM - Alternating Direction Method of Multipliers: An optimization algorithm that decomposes complex problems into smaller subproblems, which are then easier to solve.

APGM - Accelerated Proximal Gradient Method: An optimization technique that improves the convergence speed of traditional proximal gradient methods. BM3D - Block-matching and 3D filtering: An image denoising technique that exploits the redundancy of the image structure.

CNN - Convolutional Neural Network: A deep learning algorithm which can take in an input image, assign importance to various aspects/objects in the image and differentiate one from the other.

CT - Computed Tomography: A diagnostic imaging test used to create detailed images of internal organs, bones, soft tissue, and blood vessels.

CWDR - Color Wavelet Difference Reduction: An extension of the Wavelet Difference Reduction method tailored to compress color images.

DCT - Discrete Cosine Transform: A technique used to convert a signal into elementary frequency components.

DICOM - Digital Imaging and Communications in Medicine: A standard for storing and transmitting medical imaging information.

DnCNN - Denoising Convolutional Neural Network: A deep learning model specifically designed for image denoising tasks.

DRCN - Deeply-Recursive Convolutional Network: A network structure for deep learning focusing on image super-resolution.

DWT - Discrete Wavelet Transform: A transformation that analyzes data at different scales or resolutions.

EDVR - Enhanced Deformable Convolutional Networks for Video Restoration: A method that uses deformable convolutions for high-performance video restoration, including super-resolution.

FFT - Fast Fourier Transform: An algorithm that computes the Discrete Fourier Transform (DFT) and its inverse, widely used for digital signal processing.

GANs - Generative Adversarial Networks: An approach in machine learning using two neural networks, contesting with each other to generate new, synthetic instances of data that can pass for real data.

HR - High Resolution: Refers to images or videos with a higher pixel count, resulting in greater detail and clarity.

JPEG - Joint Photographic Experts Group: A commonly used method of compression for digital images.

JPEG 2000 - An updated version of JPEG that provides better compression with minimal information loss.

LR - Low Resolution: Refers to images or videos with a lower pixel count, which affects the sharpness and detail.

LSTM - Long Short-Term Memory: A type of recurrent neural network capable of learning order dependence in sequence prediction problems.

MCPD - Mean Co-located Pixel Difference: Used for assessing the similarity in the spatial and temporal domain between slices in medical imaging.

MCSE - Multi-Dimensional Compression by Substring Enumeration: A generalization of compression algorithms for handling high-dimensional data.

MLP - Multi-Layer Perceptron: A type of deep neural network used in supervised learning.

MOS - Mean Opinion Score: A measure used in the domain of telecommunication and video to obtain the human users' view on the quality of a network. MRI - Magnetic Resonance Imaging: A medical imaging technique used in radiology to form pictures of the anatomy and the physiological processes of the body.

NIQE - Natural Image Quality Evaluator: A no-reference image quality assessment metric that evaluates the perceptual quality of images without a reference image.

PnP - Plug-and-Play: Refers to the capability of a system to automatically detect and configure hardware or software without user intervention.

PPP - Plug-and-Play Priors: A framework used in computational imaging that integrates physical models with learned models to solve inverse problems through iterative algorithms.

PSNR - Peak Signal-to-Noise Ratio: A measure used to assess the quality of reconstruction of lossy compression codecs.

RAISR - Rapid and Accurate Image Super Resolution: An algorithm that uses machine learning to rapidly and accurately upscale images.

RGB - Red, Green, Blue: A color model in which red, green, and blue light are added together in various ways to reproduce a broad array of colors.

RLC - Run-Length Coding: An image compression method that encodes consecutive data elements within a file.

SOF-VSR - Super-resolving Optical Flow for Video Super-Resolution: A method for enhancing the resolution of video sequences.

SPIHT - Set Partitioning in Hierarchical Trees: An algorithm used for image compression that exploits the inherent similarities across the sub-bands in a wavelet transform.

SR - Super Resolution: A class of techniques that enhance the resolution of an imaging system.

SSIM - Structural Similarity Index: A method for measuring the similarity between two images.

TCIA - The Cancer Imaging Archive: A service that de-identifies and hosts a large archive of medical images of cancer accessible for public download.

TV - Total Variation: A method used in image processing to reduce image noise without removing significant parts of the image content.

VDSR - Very Deep Super Resolution: A deep learning model for image super-resolution.

WDR - Wavelet Difference Reduction: A method for compressing images by exploiting spatial and temporal coherence between adjacent slices.

WDR-3D - 3D Wavelet Difference Reduction: A method for compressing 3D images by exploiting spatial and temporal coherence between adjacent slices.

YUV - A color encoding system typically used as part of a color image pipeline. It encodes a color image or video taking human perception into account, allowing reduced bandwidth for chrominance components, thereby typically enabling transmission errors or compression artifacts to be more efficiently masked by the human perception than using a direct RGB-representation.

Abstract

Stamatia-Christina Zerva, Ph.D., Department of Computer Science and Engineering, School of Engineering, University of Ioannina, Greece, 2024.
Compression and enhancement of medical images and video.
Advisor: Lisimachos Kondi, Professor.

This dissertation explores the advancement of medical imaging technologies through the development of innovative methodologies for medical image and video compression and super-resolution. The research focuses on improving the efficiency and quality of medical image compression by introducing an enhanced method based on Wavelet Difference Reduction (WDR). Additionally, it proposes novel approaches for video and MRI super-resolution utilizing Plug-and-Play Priors (PnP) integrated within the Alternating Direction Method of Multipliers (ADMM) framework. The results demonstrate significant improvements in Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) across various datasets, emphasizing the practical applicability and superior performance of the proposed methods in clinical settings. Future directions include further optimization for real-time applications, extending the methodologies to other medical imaging modalities, and integrating robust security measures to safeguard patient data.

Εκτεταμενη Περιληψη

Σταματία-Χριστίνα Ζέρβα, Δ.Δ., Τμήμα Μηχανικών Η/Υ και Πληροφορικής, Πολυτεχνική Σχολή, Πανεπιστήμιο Ιωαννίνων, 2024.
Συμπίεση και υπερ-ανάλυση ιατρικών εικόνων και video.
Επιβλέπων: Λυσίμαχος Κόντης, Καθηγητής.

Εισαγωγή Ο τομέας της ιατρικής απεικόνισης έχει γνωρίσει αξιοσημείωτες προόδους με την εμφάνιση εξελιγμένων τεχνικών συμπίεσης και βελτίωσης εικόνας. Αυτή η διατριβή στοχεύει να συμβάλει σε αυτόν τον τομέα αναπτύσσοντας αποτελεσματικές μεθοδολογίες για τη συμπίεση και την υπερ-ανάλυση ιατρικής εικόνας, αντιμετωπίζοντας την κρίσιμη ανάγκη για υψηλής ποιότητας απεικόνιση στην κλινική διάγνωση και τη φροντίδα των ασθενών.

Στόχοι Οι πρωταρχικοί στόχοι αυτής της διατριβής είναι: Να αναπτυχθεί μια νέα, αποτελεσματική μέθοδο για τη συμπίεση ιατρικής εικόνας. Να να προταθούν μια νέα, αποτελεσματική τεχνική για υπερ-ανάλυση βίντεο. Να εισαχθεί μια προηγμένη μεθόδος για την υπερ-ανάλυση ιατρικής εικόνας.

Δομή Η διατριβή είναι οργανωμένη σε επτά κεφάλαια, καθένα από τα οποία εστιάζει σε διαφορετικές πτυχές των προόδων της τεχνολογίας ιατρικής απεικόνισης. Στο πρώτο κεφάλαιο παρουσιάζονται οι στόχοι και η δομή της διατριβής. Το δεύτερο κεφάλαιο παρέχει ένα ολοκληρωμένο υπόβαθρο για τις τεχνολογίες ιατρικής απεικόνισης, συμπεριλαμβανομένης της εξέλιξης της συμπίεσης εικόνας και της υπερ-ανάλυσης. Τα κεφάλαια τρία έως πέντε παρουσιάζουν τις βασικές ερευνητικές συνεισφορές, περιγράφοντας λεπτομερώς τις νέες μεθοδολογίες για τη συμπίεση εικόνας και την υπερ-ανάλυση, και τις εφαρμογές τους σε ιατρικά πλαίσια. Το έκτο κεφάλαιο ολοκληρώνει τη διατριβή, συνοψίζοντας τα βασικά ευρήματα και προτείνοντας μελλοντικές ερευνητικές κατευθύνσεις.

Συμπίεση ιατρικής εικόνας Η συμπίεση ιατρικής εικόνας διαδραματίζει κρίσιμο ρόλο στην αποτελεσματική αποθήκευση, μετάδοση και διαχείριση μεγάλων συνόλων ιατρικών δεδομένων. Αυτή η έρευνα εισάγει μια βελτιωμένη μέθοδο συμπίεσης που βασίζεται στη Μείωση Διαφοράς Κυματιδίων (WDR). Η προτεινόμενη μέθοδος, που ονομάζεται 3D-WDR-MCPD, χρησιμοποιεί τη Μέση Διαφορά Pixel Co-Located (MCPD) για την επιλογή βέλτιστων τμημάτων για συμπίεση, διασφαλίζοντας υψηλές τιμές PSNR και SSIM διατηρώντας παράλληλα κρίσιμες διαγνωστικές πληροφορίες. Εκτενείς αξιολογήσεις σε δημόσια διαθέσιμα σύνολα δεδομένων, όπως το Cancer Image Archive (TCIA), καταδεικνύουν την ανώτερη απόδοση της μεθόδου, επιτυγχάνοντας υψηλούς λόγους συμπίεσης με ελάχιστη απώλεια οπτικής ποιότητας.

Υπερ-ανάλυση βίντεο Η υπερ-ανάλυση βίντεο είναι ζωτικής σημασίας για τη βελτίωση της ποιότητας των ιατρικών βίντεο, τα οποία είναι απαραίτητα σε διαδικασίες όπως η ενδοσκόπηση και η παρακολούθηση ασθενών. Αυτή η διατριβή προτείνει μια νέα μέθοδο για υπερ-ανάλυση βίντεο χρησιμοποιώντας Plug-and-Play Priors (PnP) εντός του πλαισίου ADMM. Η μέθοδος αξιοποιεί δίκτυα βαθιάς αποθορυβοποίησης, όπως το DnCNN, για να βελτιώσει τις λεπτομέρειες της εικόνας χωρίς σημαντικές υπολογιστικές επιβαρύνσεις. Η προσέγγιση επικυρώνεται μέσω εκτεταμένων πειραμάτων, παρουσιάζοντας ουσιαστικές βελτιώσεις στην ποιότητα του βίντεο σε διάφορα σύνολα δεδομένων.

Υπερ-ανάλυση μαγνητικής τομογραφίας Η μαγνητική τομογραφία (MRI) είναι ένα κρίσιμο εργαλείο στην ιατρική διαγνωστική, που απαιτεί εικόνες υψηλής ανάλυσης για ακριβή ανάλυση. Η διατριβή επεκτείνει το πλαίσιο PnP σε υπερ-ανάλυση μαγνητικής τομογραφίας, ενσωματώνοντας καινοτόμους denoisers στον αλγόριθμο ADMM. Αυτή η τεχνική βελτιώνει σημαντικά την ανάλυση και τη σαφήνεια των εικόνων MRI, παρέχοντας πλουσιότερο οπτικό πλαίσιο για κλινική διάγνωση.

Αποτελέσματα και συζήτηση Οι προτεινόμενες μεθοδολογίες καταδεικνύουν σημαντικές προόδους στην ποιότητα της ιατρικής εικόνας και βίντεο. Τα βασικά ευρήματα περιλαμβάνουν: Η μέθοδος 3D-WDR-MCPD επιτυγχάνει έως και 3,8 dB βελτίωση του PSNR, εξασφαλίζοντας υψηλή απόδοση συμπίεσης και διατήρηση του διαγνωστικού περιεχομένου. Η μέθοδος υπερ-ανάλυσης βίντεο που βασίζεται σε PnP ξεπερνά σταθερά τις σύγχρονες τεχνικές, παρέχοντας βίντεο ευκρινέστερα και υψηλότερης ποιότητας. Η προσέγγιση υπερ-ανάλυσης MRI ενισχύει τη σαφήνεια της εικόνας, υποστηρίζοντας καλύτερες κλινικές αξιολογήσεις. Αυτά τα αποτελέσματα υπογραμμίζουν τις πρακτικές επιπτώσεις των προτεινόμενων μεθόδων, υπογραμμίζοντας τις δυνατότητές τους να φέρουν επανάσταση στην τεχνολογία ιατρικής απεικόνισης, επιτρέποντας πιο αποτελεσματική διαχείριση δεδομένων και βελτιωμένη διαγνωστική ακρίβεια.

Μελλοντικές έρευνες Η μελλοντική έρευνα θα επικεντρωθεί σε πολλές πολλά υποσχόμενες οδούς, όπως: Περαιτέρω βελτιστοποίηση μοντέλων βαθιάς εκμάθησης για βελτιωμένη υπερ-ανάλυση εικόνας. Επέκταση των μεθόδων που αναπτύχθηκαν σε άλλες μεθόδους ιατρικής απεικόνισης, όπως αξονικές τομογραφίες και υπερηχογράφημα. Δίνοντας έμφαση στις δυνατότητες επεξεργασίας σε πραγματικό χρόνο για την υποστήριξη κλινικών πρακτικών και τηλεϊατρικής. Ενσωμάτωση ισχυρών μέτρων κρυπτογράφησης και απορρήτου για την προστασία των δεδομένων ιατρικών εικόνων.

Συμπέρασμα Αυτή η διατριβή παρουσιάζει πρωτοποριακές συνεισφορές στην τεχνολογία ιατρικής απεικόνισης, προσφέροντας καινοτόμες λύσεις για συμπίεση εικόνας και βίντεο και υπερ-ανάλυση. Αντιμετωπίζοντας κρίσιμες προκλήσεις και αξιοποιώντας τις δυνατότητες της βαθιάς μάθησης, η έρευνα θέτει ένα νέο σημείο αναφοράς για μελλοντικές εξελίξεις στον τομέα, υποσχόμενη να βελτιώσει σημαντικά τη φροντίδα των ασθενών και τη διαγνωστική ακρίβεια.

Chapter 1

Introduction

In the context of the ever-evolving field of digital image processing, which utilizes mathematical operations and signal processing techniques to analyze images and videos, significant technological advancements have been made in medical imaging and telemedicine.

Specifically, developments such as medical image compression, super-resolution of medical images, and video super-resolution play pivotal roles. These technologies are not only essential for improving the efficiency of medical data storage and transmission but also enhance the quality and resolution of medical images and videos. Consequently, they contribute significantly to facilitating better diagnosis, treatment planning, and research, leveraging the digital processing techniques that underpin modern image interpretation.

Medical Image Compression involves reducing the size of medical images without substantially degrading their quality. This process is essential for effective data storage and rapid transmission, especially in telemedicine applications where bandwidth may be limited. Techniques like JPEG, JPEG 2000, and advanced deep learning methods are commonly used for this purpose. The goal is to achieve high compression ratios while preserving the critical diagnostic features of the images [7].

Medical Image Super-Resolution refers to the process of enhancing the resolution of medical images. This is particularly important for improving the visibility of fine details that are crucial for accurate diagnosis and treatment planning. Techniques like interpolation, deep learning-based approaches, and generative adversarial networks (GANs) have been widely adopted for super-resolution. These methods aim to reconstruct high-resolution images from their low-resolution counterparts, making it possible to identify and analyze medical conditions with greater precision [8].

Video Super-Resolution in the medical field involves increasing the resolution of video sequences. This is vital in various medical applications, such as endoscopy, surgery recordings, and patient monitoring, where higher resolution can lead to better visibility and understanding of dynamic processes. Techniques for video super-resolution often rely on frame interpolation, motion estimation, and deep learning algorithms to enhance the detail and clarity of video frames. This not only improves the quality of medical videos but also supports better decision-making and documentation in clinical practices [9].

These technologies collectively contribute to the advancement of medical imaging by enabling more efficient data management and enhancing the quality and clarity of medical images and videos. Their ongoing development and application hold the promise of further improving patient care, diagnostic accuracy, and medical research.

1.1 Objectives

Image compression and image enhancement techniques are the most widely used techniques these days in the field of medical images. So, in this PhD dissertation we compare the performance quality of the different compression and enhancement techniques based on different performance metrics.

Based on the need of developing more efficient techniques of compression and super-resolution, the objectives of this dissertation are:

- To develop a new, efficient method of medical images compression.
- To develop new, efficient methods of video super-resolution.
- To develop a new, efficient method of medical images super-resolution.

1.2 Structure

This dissertation consists of six chapters, each focusing on different aspects of advancements in medical imaging technologies, specifically in the domains of image compression, medical image super-resolution, and video super-resolution.

Chapter one is the introduction of the dissertation, which results in the objectives and structure of the dissertation

Chapter two provides a comprehensive background on medical imaging technologies, including the evolution of image compression, the importance of image and video super-resolution in medical applications, and a review of existing technologies and methodologies. It sets the stage for understanding the current challenges and opportunities in the field.

Chapter three focuses on the crucial aspect of medical image compression, delving into the technical details of various compression techniques, from traditional methods like JPEG and JPEG 2000 to advanced deep learning approaches and suggests a wavelet-based MRI compression method, based on wavelet difference reduction.

Chapter four examines the technology's significance in medical contexts such as endoscopy, surgical procedures, and patient monitoring. It presents various techniques for improving video quality, including frame interpolation and motion estimation, and evaluates their effectiveness in clinical practice and finally suggests a video super-resolution method using plug-and-play priors and an MRI super-resolution method using plug-and-play priors and rigid transformation.

Chapter five extends on the method presented in Chapter 5 and presents a new technique for video super-resolution that incorporates an innovative denoiser within the ADMM algorithm.

The concluding Chapter six summarizes the key findings of the dissertation, emphasizing the contribution of the research to the field of medical imaging technology. It outlines the prospects for future advancements in image and video compression and super-resolution, highlighting areas for further investigation and development.

Chapter 2

Theoretical Background

2.1 Medical Image Compression

Medical image compression became a prevalent tool with a significant impact on diagnosing diseases in clinical practice [10]. The problem of compressing and transmitting an image in real-time, given the bandwidth of the communication channel, was of great importance, especially in a low-speed connection environment. This problem was not easy to solve because medical images typically contained a huge amount of important diagnostic information that needed to be losslessly [11] transmitted and stored [7]. Real-time constraints limited image compression applications for transmission purposes. On the other hand, image compression applications for storage purposes were less stringent since most algorithms were not executed in real time.

Many irreversible standards were used for compressing images [12]. One of the most popular for medical applications was the standard from the joint photographic experts group (JPEG) [13]. The vital feature of JPEG was that it enabled compression at various levels, thus allowing the user to choose the quality of the compressed image so that information losses were not visible to physicians. JPEG 2000 [14] was the successor of the JPEG standard that provided compression with no or very little information loss, so the image quality did not deteriorate but approximated the image quality without compression. Compared to the JPEG standard, the JPEG 2000 standard provided a typical compression gain of 20

A wide variety of image transform-based coding techniques existed, which among others included discrete cosine transform (DCT) [15] and discrete wavelet transform (DWT) [16]. The DCT method proposed by Nasir [17] transformed image pixels from the spatial domain into the frequency domain, allowing redundancy to be found. The Hungarian mathematician Alfréd Haar created the first DWT method. The main characteristic of this method was that the wavelets were discretely sampled, and it had a temporal resolution that allowed capturing both frequency and location information [18]. For the compression of volumetric medical datasets, it appeared that 3D wavelet-based encoders outperformed DCT-based solutions while providing the required functions such as quality scaling and resolution, random access and region coding [19]. Narmatha *et al.* [20] proposed a two-stream method for encoding and decoding medical images by dividing and merging different regions of the wavelet subbands. Amri *et al.* [21] combined into a single processing pipeline image reduction and expansion techniques of different lossless compression standards such as JPEG-LS and TIFF formats to compress medical images.

In recent years, much effort focused on volumetric medical image compression, where 3D medical images could be viewed as time sequences or volume tomographic slices of an object. Bruylants et al. [22] employed the wavelet transformation to allow support for volumetric image datasets. Ravichandran et al. [23] also used the wavelet transform to compress 3D medical images. Senapati et al. [24] proposed the 3D hierarchical listless block (3D-HLCK) algorithm, a modified 3D block coding algorithm containing a listless variant. Tang and Pearlman [25] created the 3D set partitioning embedded block method (3D-SPECK), which encoded 3D volumetric image data by utilizing the dependencies in each dimension. Chen et al. [26] developed an end-to-end learning-based framework for 3D volumetric image compression. The framework used the intra-slice and inter-slice information to predict the entropy coding distribution. Also, it utilized two novel gating mechanisms for better aggregation of the intra-slice and inter-slice features. Nagoor et al. [27] proposed a lossless compression algorithm that trained a neural network as a 3D data predictor for medical image volumes containing images with 65,536 levels of colors and tones.

Additionally, Zerva et al. [1] proposed an extension of the standard wavelet difference reduction (WDR) method using mean co-located pixel difference (MCPD) to select the optimal number of slices that exhibit the highest similarity in the spatial and temporal domain. The slices with large spatiotemporal coherence were encoded together as one volume in terms of higher PSNR and structural similarity index (SSIM). It was found that the perceptual quality of the medical image was remarkably high. The results indicated that the PSNR improvement over existing schemes might reach up to 3.8 dB and could guide the implementation of a mobile and web platform that can be used for compressing and transmitting medical images in real time.

Among reversible compression algorithms, Huffman coding was one of the oldest methods of compressing image data. Developed by Huffman [28], it was used to reduce coding redundancy without degrading the quality of the reconstructed image. Other reversible compression algorithms included arithmetic coding and lossless predictive coding. Arithmetic coding was a form of entropy encoding used in image compression and other data compression methods. Unlike traditional coding techniques, which assigned a distinct code or set of bits to each symbol in the data, arithmetic coding represented the entire message as a single number, a fraction n where $0.0 \leq n < 1.0$ [29]. Lossless predictive coding was a two-stage approach that utilized a lossless adaptive predictor followed by arithmetic coding [30].

Recent lossless approaches involved the multi-dimensional compression by substring enumeration (MCSE) by Dubé [31]. CSE was a compression algorithm for bit strings that was generalized to higher dimensions to handle all types of images. Makarichev et al. [32] modified the irreversible discrete atomic compression (DAC) algorithm by adding compressed data describing the difference between the original image and the compressed one inside the corresponding DAC file. The addition combined with the compressed image resulted in a reconstructed image without any distortions. Lee et al. [33] developed a high-throughput image-compression technique using the Golomb-Rice coding and its hardware architecture. Descampe et al. [34] proposed the JPEG XS compression algorithm for visually lossless, low-latency lightweight image coding. It was an international standard that achieved similar (slightly lower) compression ratios compared to the JPEG 2000 method. One advantage over JPEG 2000 was that it consumed significantly less power and required fewer logic resources in hardware implementations.

A significant number of approaches utilized artificial neural networks (ANNs) for specific tasks to increase the compression ratio. Min et al. [35] created a hybrid approach to compress three-dimensional (3D) medical images. The hybrid algorithm utilized the medical images' anatomical features to divide the medical data into specific areas. Then, a deep neural network created optimal predictors in each area. The predictors could be switched adaptively according to the area's characteristics being compressed. Finally, the residuals were compressed using an entropy coding scheme. Yang et al. [36] created an image compression-encryption algorithm with the help of a fractional-order memristive BPF chaotic circuit and a back-propagation (BP) trained neural network. The neural network compressed the image while the encryption process was done using a zigzag algorithm with a xor operation. Rhee et al. [37] created a lossless compression technique based on the multi-layer perceptron (MLP) neural network. The MLP outputted prediction errors and contexts which were introduced as input to adaptive arithmetic encoders. Zhu et al. [38] used a long short-term memory (LSTM) neural network for building a predictor, which was used in lossless compression.

In contrast, among the irreversible compression algorithms (created by the need to produce significantly lower bit rates), there were various approaches to lossy image compression, such as vector quantization, coding prediction [39], and transform coding. The general components of a lossy image compression technique involved the three stages (decomposition or transformation, quantization, and symbol encoder) shown in Fig. 2.1.

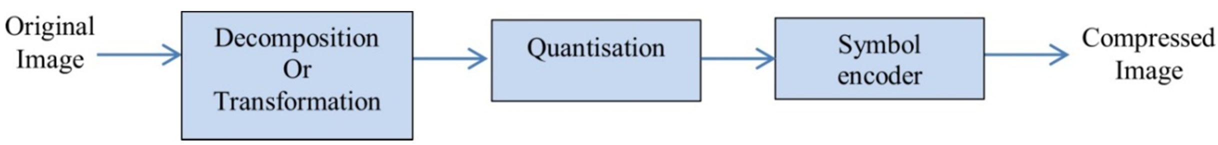


Figure 2.1: General elements of an irreversible compression technique. An irreversible compression technique is, in general, a three-stage process. The procedure begins by decomposing or transforming the image, followed by a quantization and symbol encoding process.

There were various transformations used in image compression. The Karhunen-Lo‘eve transform (KLT) [40, 41, 42] was an orthogonal linear transformation technique that removed pairwise statistical correlation amongst the transform coefficients.

The piecewise Fourier transform (PFT) [43] maintained image quality by compressing the images' bandwidth. The discrete Walsh-Hadamard transform (DWHT) [44, 45] was an orthogonal transformation type that broke down a signal into a series of Walsh functions (orthogonal and rectangular wave-forms). Finally, the wavelet transform algorithms [46], which prevailed in the compression of medical images, were used as frequency analysis and signal coding tools in complex non-stationary signals.

Recent irreversible compression algorithms included the work of Xu et al. [47], which improved the singular value decomposition (SVD) method using a singular vector sparse reconstruction strategy. Guo et al. [48] developed an image compression framework for computer vision applications in embedded systems. The framework made use of the trade-off between memory traffic and vision performance. Sadchenko et al. [49] created a compression algorithm based on the samples decimation method for medical images, which considered medical image peculiarities.

Some lossy approaches utilized artificial neural networks (ANNs) to increase the compression ratio. Dua et al. [50] used a convolutional neural network (CNN) for compressing hyperspectral images. The algorithm combined CNN's auto-encoder, convolution, and max-pooling layers to reduce the image's dimensions and produce a compressed image. The image could be restored with some loss of information by reversing the CNN's steps using the CNN's decoder and transpose convolution layer. Zhao et al. [51] utilized multiple description CNNs to compress images for transmission. Multiple description coding was used for signal transmission in unreliable and non-prioritized networks. Mishra et al. [52] proposed a two-stage auto-encoder-based framework for compressing and decompressing malaria red blood cell images. The above irreversible methods managed to achieve high compression ratios, but they were unsuitable for medical images since they could lose potentially valuable medical information.

2.2 Medical Image Compression Evaluation Methods

Several methods evaluated the clinical acceptance of the compression level [53]. The first was the numerical analysis of the pixel before and after compression [54]. This simple method was recommended for calculating the mean pixel error for the compressed image but had no correlation with radiologists' evaluations and therefore had no clinical significance. A second method used subjective observers to evaluate with a focus on visual acceptance and presumptive diagnostic value. Many approaches were proposed, including image scores from the least to the most compressed or subjective evaluations of the onset of a pathological process. None of this led to reliable and reproducible results. A third method was the objective measurement of diagnostic accuracy using blind method evaluation. This category of methods was the most reliable.

The relationship between "optically lossless" compression and "diagnostically lossless" is complex. There was evidence that despite the apparent visual degradation from compression, high performance equivalent to that of uncompressed images for certain details, body parts, and diagnostic methods could be achieved. This equivalent did not alter the ability of a radiologist to successfully interpret a poor-quality image (perhaps with less confidence). On the other hand, many physicians were reluctant to interpret compression-degraded images, so the "visually lossless" limit may have been the limiting factor despite the "diagnostically lossless" limit, assuming that the former implied less compression than the latter. Conversely, although it was often assumed that if there was no visual quality loss, there could be no diagnostic loss. The above claim had not been sufficiently investigated, and there was a possibility that the experimental way which defined the thresholds for visual perception without losses was insufficient to guarantee diagnostic performance. Challenging tasks, including low-contrast detection, must maintain high-frequency information, or they would be vulnerable to high compression rates, which were misinterpreted as false-positive findings [55].

Simple mathematical measurements that quantified the difference between the original and the decompressed image, such as PSNR and mean square error (MSE), were poorly correlated with visual or diagnostic performance, and more advanced measurements were developed. SSIM [56] was a method for measuring the similarity between two images. The SSIM index could be considered a measure of the quality of one of the images being compared, provided that the other image was considered of excellent quality. Another method based on mathematical models simulated human physiology. These software tools could help measure image similarity or differences and determine noticeable difference (JND), signal-to-noise ratio (SNR) ratios, or levels. Probability for detecting differences in the number of pixels. Here, the structural similarity method (SSIM) was an improvement over traditional methods such as

PSNR and MSE because it appeared to be more consistent with HVS performance [57].

The amount of "information" in an image is described as its "entropy", which could be estimated mathematically, with varying degrees of complexity. A simple measure was the zero-order entropy (sum of the environmentally independent probabilities of each pixel value). The degree to which an image could be compressed using reversible compression could also be used as a measure of entropy. An image's entropy determined its compression ratio before the difference was visually or diagnostically detectable. A significant factor in a medical image's entropy was the amount of rectangular pixel panel occupied by the body part (e.g., consider a small versus a large breast on a fixed-size mammography scanner). Also important was the amount of noise in any unstable background (non-static) or area that had been separated [58].

High entropy images should probably be processed with lower compression ratios to irreversible compression than those with more uniform content. A simple approach was to measure the file output size of a reversible image compression method (JPEG lossless or JPEG 2000), which should be larger for images with higher entropy. Other reliable methods, such as image compositional complexity (ICC), fractal dimension (FD), or region of interest (ROI), might have been more effective at computing and creating images more noise resistant [59].

2.3 Wavelet transform

The wavelet transform combined low-pass and high-pass filtering into a spectral signal decomposition and extremely fast implementation. Before considering the wavelet transformations of 2D images, it was useful first to consider the wavelet transformations of one-dimensional (1D) signals [60, 61]. Given a 1D signal $s_0[n]$, its one-level wavelet transform is the mapping $s_0[n] \to (s_1[2n]|d_1[2n])$ defined by the formulas (2.1) and (2.2).

$$s_1[2n] = \sum_{k=-M}^{M} a_k s_0[2n+k] \tag{2.1}$$

$$d_1[2n] = \sum_{k=-N}^{N} \beta_k s_0[2n + k + 1] \tag{2.2}$$

The signals $s_1[2n]$ and $d_1[2n]$ are respectively low-pass and high-pass filterings of $s_0[n]$. These filterings have also been down-sampled and are defined over the indices $\{2n\}$ rather than $\{n\}$. Viewed as sampled, signals are sampled at half the rate as s_0. The coefficients $\{a_k\}$ are the low-pass coefficients and the coefficients $\{\beta_k\}$ are the high-pass coefficients [60].

These coefficients have some basic properties which are shared by other wavelet systems. One important property is that they define an invertible transform. Perhaps just as importantly, the high-pass coefficients satisfy $\sum \beta_k = 0$ and $\sum k\beta_k = 0$. Consequently, if s_0 is linear (or approximately linear) over the indices $2n, 2n+1, 2n+2$, then $d_1[2n] = 0$ (or $d_1[2n] \approx 0$). When s_0 is obtained from samples of a piecewise smooth function, the high-pass filtering d1 will be essentially zero-valued (except near transitions between pieces of the piecewise smooth function). This provides the foundation for compression. When the transform $s_m \to (s_{m+1}|d_{m+1})$ is iterated on the low-pass outputs $s_1, s_2, \ldots$, then many levels of transformation will produce large numbers of zero values (or almost zero values) at high-pass outputs $d_2, d_3, \ldots$ Such high redundancy of zero values, in $d_1, d_2, d_3, \ldots$, allows for significant compression [60].

A wavelet transform for 1D signals can easily be generalized to 2D images by applying it separately to each dimension. The first level of a discrete particle transformation of a matrix $F = J \times K$, where J and K are both even, is obtained in a two-step manner. The first step can be seen in equation (2.3) and involves transforming each row of F with a 1D particle transformation by taking a matrix $\tilde{F}$.

$$F \to \begin{pmatrix} s_1^1|d_1^1 \\ s_2^1|d_2^1 \\ \vdots \\ s_j^1|d_j^1 \end{pmatrix} \tag{2.3}$$

The second step, shown in formula (2.4) transforms each column of $\tilde{F}$ by the same 1D transform where A^1, V^1, H^1, and D^1 are each $\frac{J}{2} \times \frac{K}{2}$ sub-matrices. Steps one and two are independent and may be performed in either order [60].

$$F \rightarrow \left(\frac{A_1^1|V_1^1}{H^1|D^1} \right) \tag{2.4}$$

The wavelet transform can be iterated on the row-low-pass/column-low-pass outputs (two-level transform). Doing this on A^1 produces submatrices A^2, V^2, H^2, and D^2. As with 1D signals, the second level sub-matrices are responses to the 2D image values having twice the range of pixels (twice the scale) as the first level sub-matrices [60].

2.4 Image and Video Super-Resolution

Super-resolution (SR) involved the generation of high-resolution images or videos from their low-resolution counterparts, presenting a complex challenge within the field of computer vision. Its applications spanned diverse domains, including medical imaging, surveillance, remote sensing, and multimedia. The enhancement of resolution and visual quality in low-resolution videos was a particularly demanding task within super-resolution, as it aimed to address issues like motion, subsampling, additive noise, and point spread function (PSF) blurring between frames in a low-resolution (LR) sequence [9].

Researchers, over time, introduced various techniques and algorithms to tackle the intricate problem of super-resolution. Within a low-resolution sequence, each frame captured only a fraction of the original high-resolution (HR) image's information due to inherent degradations. However, frames with subpixel motion offered unique partial information of the original HR image. Consequently, with sufficient LR frames containing distinct information, the HR image could be reconstructed through digital image or video processing [62].

The use of deep learning for image super-resolution attracted considerable interest for its efficiency in transforming low-resolution images into high-resolution counterparts. A range of deep learning frameworks and techniques were developed for this purpose. The field saw significant progress with the adoption of Convolutional Neural Networks (CNNs) in enhancing image resolution, as evidenced by various studies and models [63, 64, 65, 66, 67, 68, 69, 70] such as SRCNN by Dong et al. [71], which marked a groundbreaking development in single-image super-resolution. This was followed by the introduction of more complex models like VDSR and DRCN by Kim et al., which featured deeper networks [63] and recursive learning [64], respectively.

Unlike traditional methods, learning-based strategies aimed to directly learn the transformation from low-resolution to high-resolution images. Examples included approaches that substituted low-resolution patches with corresponding high-resolution matches from a predefined dictionary [72, 73, 74, 75], as suggested by Timofte et al. [76, 73], or self-example methods [77] that leveraged recurring patterns within the image to enhance resolution. In recent developments, deep neural networks were recognized for their ability to learn complex, hierarchical data representations, further advancing the capabilities of image super-resolution.

The application of Generative Adversarial Networks (GANs) to image super-resolution, proposed by Ledig et al. [65] with SRGAN, revolutionized the field. GANs enabled high-quality SR by generating more visually pleasing details, albeit with the risk of introducing non-realistic textures. Deep learning models combined with regularization techniques were explored in works like Zhang et al. [78], where a CNN was used with a sparsity-promoting regularizer to achieve superior super-resolution results. Recent research shifted toward joint image restoration and super-resolution, where the low-resolution image was restored before being super-resolved. An example was the work of Zhang et al. [79], which combined the power of CNNs with an image restoration framework.

Sparse coding and dictionary learning-based approaches, pioneered by Yang et al. [80], made significant contributions to image super-resolution. These techniques modeled the relationship between low and high-resolution patches by learning overcomplete dictionaries and optimizing for sparsity.

Video super-resolution followed image super-resolution. Tsai and Huang [81] used the Fourier Transform's shifting property and the aliasing connection between the continuous and discrete Fourier transforms. In the spatial domain, Stark and Oskoui [82] introduced the projection onto convex sets (POCS) method, which aligned convex constraint sets that reflected the desired image attributes with the high-resolution image domain. This technique was adapted for dynamic motion blur through methods such as block matching and phase correlation [83, 84].

In video super-resolution, the accurate estimation of motion assumed a pivotal role. This process, essential for enhancing the resolution of low-resolution videos, involved aligning and consolidating information from multiple frames to generate a high-resolution output [85].

The deterioration of images in video super-resolution commonly involved the representation of a linear blur, motion, subsampling, and Gaussian noise. This was typically conceptualized through an observation model, assuming the acquisition of multiple low-resolution (LR) images through a specific process [86]. According to this model, the LR input images were obtained from the high-resolution (HR) original scene through operations such as warping, blurring, and downsampling. It was assumed that the HR image remained constant during the acquisition of several LR images [87].

2.4.1 Plug and Play Priors

Plug-and-Play Priors (PPP) stood out as a widely adopted framework that integrated physical and learned models to address computational imaging challenges. It was a robust framework that merged conventional optimization techniques with modern denoising methods and priors to efficiently tackle inverse problems [88]. Initially introduced by Venkatakrishnan et al. in 2013 [5], PPP garnered significant attention across various domains of computer vision and image processing. This literature review delved into key contributions that shaped the development and application of PPP.

The original PPP framework proposed by Venkatakrishnan et al. [5] showcased its efficacy in solving inverse problems, such as image denoising and deblurring. Their work demonstrated that by alternately applying denoising and data fidelity steps, PPP achieved state-of-the-art results. The denoising step employed robust algorithms like Non-Local Means (NLM) or Block-matching and 3D filtering (BM3D) [89] to eliminate noise and enhance image quality. The data fidelity step ensured consistency between the denoised image and the observed measurements. Despite the original formulation relying on ADMM [90], PPP proved equally effective when combined with other proximal algorithms like primal-dual splitting (PDS) [91] and fast iterative shrinkage/thresholding algorithm (FISTA) [92].

To further enhance denoising capabilities within PPP, Zhang et al. [78] introduced a deep denoising network named DnCNN. Integrating DnCNN into the PPP framework demonstrated its effectiveness in tasks such as image super-resolution and inpainting. The utilization of deep neural networks within PPP provided a more flexible and potent denoising tool, surpassing traditional handcrafted denoisers in performance.

Ghassab and Bouguila [93] explored the utilization of a Student-t mixture model as a promising tool for the reconstruction of video super-resolution. The Student-t mixture model, renowned for its heavy tail, was deemed robust and well-suited for the prior of video frame patches, offering a mixture model with a rich log-likelihood for information retrieval. Edge-preserving filtering was implemented to address potential data uncertainties and preserve areas with abrupt lighting changes in video frames. The Plug-and-Play Priors (PPP) structure was subsequently employed to integrate the Student-t mixture prior model and edge-preserving filtering into the super-resolution algorithm. Empirical evaluations conducted on various video frame sets demonstrated the effectiveness of the proposed algorithm. Comparisons with eight other state-of-the-art super-resolution methods affirmed that the proposed framework generally outperformed others across different super-resolution scales, even in the absence of leveraging motion estimation to exploit frame correlations.

PnP-ADMM is widely recognized for its efficiency and fast empirical convergence within the realm of frequently employed operators in computational imaging. However, it demands the computation of the proximal map, in contrast to PnP-FISTA, which solely requires the computation of the gradient ∇g. While the gradient is theoretically less complex than the proximal map, numerous applications enable the efficient computation or approximation of the proximal map. General techniques such as conjugate gradient or specialized methods, particularly when the forward model incorporated a spatial blurring operator computed through fast Fourier transform (FFT), could be employed for this purpose [94].

The incorporation of an extra state variable, employed as an initiation for the proximal minimization problem, streamlined this procedure. An iterative solver, commencing from this initialization, performed a series of steps to estimate the minimization effectively. This state variable also converged with the outer loop, resulting in decreased computational requirements through partial updates while maintaining the accuracy of the final solution [95].

In a research paper by Brifman et al. [96], scientists introduced a straightforward and robust super-resolution framework applicable to individual images and easily adaptable to videos. The foundation of the framework was rooted in the observation that the denoising of both images and videos could be effectively accomplished through various methods. By leveraging the Plug-and-Play-Prior framework and adopting the Regularization-by-Denoising (RED) approach, the researchers illustrated how denoisers could be harnessed to tackle both Single-Image Super-Resolution (SISR) and Video Super-Resolution (VSR) challenges using a unified formulation. Instead of incorporating motion estimation between frames, the VBM3D video denoiser was employed in this approach.

2.4.2 Regularization Methods for Video Super-Resolution

The role of regularization methods in video super-resolution was pivotal. Regularization methods introduced mathematical constraints into the super-resolution process, guiding it towards a solution that adhered to prior knowledge about the data. These constraints were vital for producing visually appealing results.

In video super-resolution, numerous regularization techniques were developed. A significant technique introduced by Tekalp et al. [97] enhanced the approach by adopting a least squares solution for equation systems and integrating a linear shift-invariant blur model. Kim et al. [98] further refined this method by implementing a weighted least squares algorithm to better manage noisy data, though these approaches required prior knowledge of global motion.

Stochastic methods formed another significant class of resolution improvement algorithms, notably including maximum likelihood and maximum a posteriori (MAP) strategies [99]. MAP strategies were, in fact, equivalent to regularization methods. The MAP approach, particularly, utilized an edge-preserving Huber–Markov random field as an image prior, offering a sophisticated solution for resolution enhancement while estimating registration parameters [100, 101, 102, 103, 104, 105, 106]. This method was supported by the use of Gibbs-Markov random fields with a focus on local interactions. The selection of an appropriate regularization parameter, which was pivotal for high-resolution image reconstruction, was adeptly addressed using the L-curve method to pinpoint the optimal "L-corner."

The accurate characterization of the point spread function (PSF) and the precise registration of subpixel movement were crucial for high-resolution image reconstruction. Nonetheless, accurately determining these parameters remained a challenge in practical scenarios. Lee and Kang [107] proposed a regularized adaptive high-resolution reconstruction technique that accounted for inaccuracies in subpixel registration, using Gaussian noise assumptions related to the magnitude of registration errors. This led to the development of methods for estimating the regularization parameter for each low-resolution frame, showing potential for convergence to a unique global solution. Additionally, a hierarchical Bayesian framework was employed [108] to tackle image restoration challenges in the face of partially known blurs, introducing iterative algorithms aimed at enhancing image restoration fidelity in complex scenarios [109, 110, 111].

Lucas et al. [112] proposed VSRResFeatGAN, which was a generator network optimized for the VSR problem, enhanced with two regularizers, a distance loss in featurespace and pixel-space.

An early approach was Total Variation (TV) regularization, proposed by Rudin et al. [113], which minimized the gradient magnitude of an image to promote piecewise smoothness. They introduced TV regularization as a means to preserve sharp edges and promote sparsity. Their work set the stage for numerous applications in image processing and beyond. It has been widely adopted in image super-resolution to suppress artifacts.

The Non-local Means (NLM) algorithm, introduced by Buades et al. [114], was incorporated into super-resolution to exploit non-local self-similarity within images, enabling better estimation of missing high-frequency details. Interestingly, the algorithm that we used for denoising in a plug-and-play context was, at least in terms of statistical inference, superior to the Non-local Means, as explained in [115]. This meant practically that Non-local Means was a sub-optimal case of the adopted model herein, since the effectiveness of the denoising task by the former was compromised for the sake of computational speed, see [115].

Dabov et al. [89] introduced the BM3D algorithm, which leveraged a 3D transform-domain collaborative filtering approach to effectively remove noise from images. This work had a profound impact on image denoising techniques.

Furthermore, Gu et al. [116] extended BM3D by incorporating weighted nuclear norm minimization. This addition further improved BM3D's capabilities for image denoising, making it a versatile and widely used tool in the field of image processing. BM3D was an image denoising algorithm, and it was not specifically a regularization algorithm. However, image denoising often involved regularization techniques to suppress noise and enhance image quality. BM3D employed collaborative filtering and 3D transform-domain methods to denoise images, which could be seen as a form of signal processing rather than traditional regularization.

The alternating direction method of multipliers (ADMM) was an algorithm that solved convex optimization problems by breaking them into smaller pieces, each of which were then easier to handle. The foundational work on ADMM by Boyd et al. [117] provided a comprehensive overview of the algorithm, including its theoretical underpinnings and practical applications in various domains, such as machine learning and distributed optimization. Another key development in the ADMM family was the Split Bregman method, introduced by Goldstein and Osher [118]. This method extended the principles of ADMM to efficiently solve L1-regularized problems, making it invaluable in image processing and sparse signal recovery tasks. ADMM became a powerful optimization tool, especially for solving problems with structured or separable objective functions.

The Accelerated Proximal Gradient Method (APGM) was another optimization algorithm primarily used for solving non-smooth convex optimization problems [90]. Similar to ADMM, it could be applied to problems with regularization terms as part of the objective function. APGM was particularly useful for solving problems with non-smooth components and could be used in regularization scenarios. APGM was widely employed for efficiently solving non-smooth convex optimization problems. The theoretical foundations of accelerated gradient methods, including APGM, were explored in Nesterov's work [119]. This research provided valuable insights into the convergence properties and efficiency of these algorithms, reinforcing their significance in optimization. Finally, Bayesian frameworks and variational methods have also been employed for image super-resolution [78].

2.5 MRI Super-Resolution

Medical imaging played a pivotal role in the diagnosis and monitoring of various neurological conditions. Magnetic Resonance Imaging (MRI) was a widely used modality for brain imaging due to its non-invasive nature and excellent soft tissue contrast. However, the spatial resolution of MRI images was limited by various factors, such as acquisition time, hardware constraints, and patient motion. This limitation could hinder the accurate diagnosis and evaluation of brain abnormalities [120].

Super-resolution techniques aimed to address this limitation by enhancing the resolution of low-resolution images, providing detailed images that were crucial for clinical assessments. The literature encompassed a variety of algorithms addressing image super-resolution. Basic interpolation methods such as nearest neighbor, bilinear, and bicubic techniques were commonly employed for upscaling low-resolution images. However, these methods often introduced artifacts [8]. Irani et al. introduced the iterative back-projection algorithm [121], a straightforward implementation used in MRI super-resolution. More advanced sparse-coding super-resolution algorithms, which involved finding a sparse representation, demonstrated super-resolution gains, particularly through techniques like dictionary learning [122] or exploitation of multiple modalities [123]. However, the application of dictionary learning in 3D MRI images for medical imaging faced challenges such as limited resolution improvement and slow execution speed, limiting its adoption [124].

Additionally, approaches based on a priori information, such as deterministic regularized methods [125, 126, 127, 128, 129, 130] and statistical regularized methods [131, 132, 133, 134], were employed in MRI SR. The field of deep learning emerged as a prominent area of research with significant implications across various domains, including MRI. In recent years, deep learning was extensively explored for imaging tasks, contributing to improvements in image quality through denoising, artifact reduction, system calibration, and sparse-data-based reconstruction [135, 136, 137, 138, 139, 140, 141].

In the context of super-resolution imaging, deep learning techniques witnessed successful applications in computed tomography (CT) [142], while early attempts at MRI super-resolution utilized methods like 3D convolutional neural networks [143], Very Deep Residual Neural Networks (VDSR) [144], generative adversarial networks (GANs) [145, 146, 147, 148, 149, 150, 151, 152, 153, 154], and densely connected networks [155, 156].

2.6 Discussion

Current compression and super-resolution techniques, while having achieved significant advancements, still face several challenges that limit their practical application. One major limitation is their dependency on extensive, high-quality training datasets. Models trained on limited types of data may not generalize well to different video content, resulting in poor performance when faced with data that deviate from training conditions. For instance, models trained primarily on urban landscapes might struggle with rural or indoor scenes.

Moreover, many state-of-the-art methods are computationally intensive, relying on deep neural networks that require substantial computational resources for both training and inference. This makes them less suitable for real-time applications or for use on devices with limited processing power, such as smartphones or embedded systems. Additionally, these methods often do not adequately handle real-world variables such as motion blur, varying lighting conditions, and compression artifacts, which are common in typical video streams.

This thesis tries to fill in these gaps by suggesting improved methods for the compression of medical images and the super-resolution of video and medical images.

Chapter 3

Medical image compression

3.1 Introduction

In this chapter, our objectives were: to (i) take into account the spatial and temporal coherence of adjacent slices in a volumetric medical image; and (ii) improve upon previous 3D-based compression methods in terms of PSNR and SSIM [157]. To accomplish these tasks, we propose an extension of the 3D wavelet difference reduction (3D-WDR) method [158] that employs the mean co-located pixel difference to estimate an optimal number of slices to be efficiently compressed in one volumetric object. Given the requirements for the best possible reconstructed images, the proposed method meet these objectives.

Our main contribution is the design of a volumetric medical image compression method that can be easily reproduced, is suitable for use in a variety of medical images such as MRI and CT scans, and achieves state-of-the-art compression results with high compression ratio and small information loss within an acceptable range. This performance originates from computing the spatiotemporal difference between adjacent slices in an image volume and compress as a single volume only those slices that exhibit the largest similarity on the pixel values. This allows us to progressively transmit a medical image through a communication channel and also allows for a gradual improvement of image quality. This work aspires to serve as a bar in the 3D medical image compression domain that future works may improve upon. We performed extensive evaluations in publicly available datasets and achieved high compression ratios in all of them while maintaining a high visual-quality, which ensures that compression of medical images that are used for diagnosis, is of critical importance as diagnostic data are preserved.

3.2 The 3D-WDR-MCPD Method

The method proposed in this section was published in 2020 [1]. The wavelet difference reduction (WDR) algorithm [158] follows the basic concepts of the set partitioning in hierarchical trees (SPIHT) algorithm [159] by incorporating extra features that aggregate the coefficients to an area of interest. By reducing the difference between the wavelet coefficients, it recognizes the important wavelet coefficients and improves their accuracy to achieve high compression ratios. During WDR encoding, the compressed output produced consists of the most important coefficients along with the series of bits that briefly describe the exact position of the significant values. It offers good perceptual quality and high compression rate while preserving the edges of the images. It is suitable for compressing medical images at a low bit rate per pixel.

In this method, the structure of the data tree used by SPIHT is avoided, and the principles of integrating and partitioning the encoded bit-level without loss are preserved (Fig. 3.1). Also, the WDR method implements run-length coding (RLC) that allows the encoder to achieve faster transmission of image details over networks.

The term "reduction of difference" refers to how the WDR algorithm encodes the positions of significant coefficient values of the wavelet transformation by using a difference-coding method. These positions are not directly encoded, but instead, the distances between the important coefficients are encoded. Thus, the WDR method encodes the path between two important coefficients. The importance of the WDR method lies in the fact that it increases the data transmission speed because the method employs the basic concepts of the run-length coding [160].

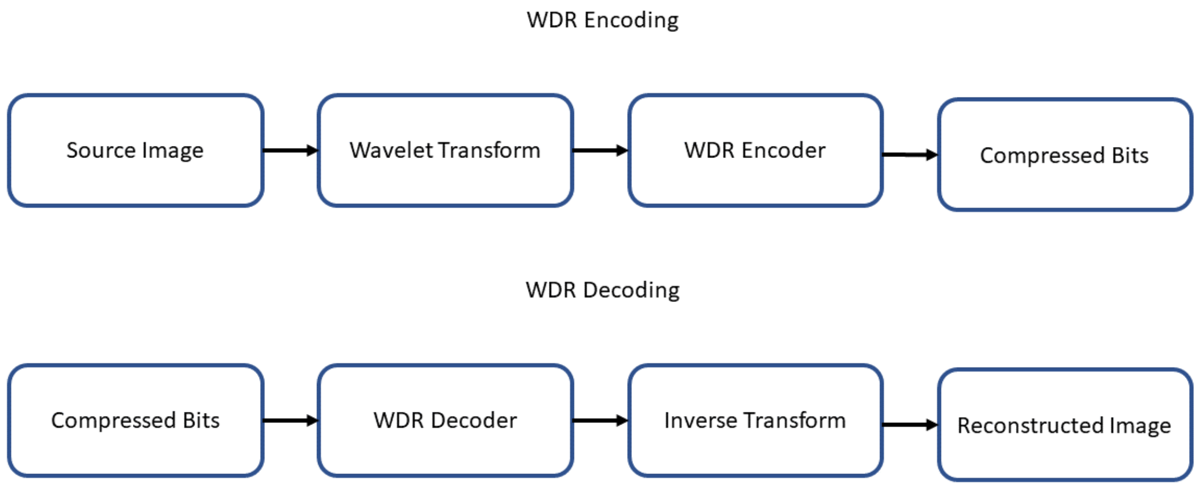


Figure 3.1: WDR block diagram. The encoding step of the WDR algorithm involves importing an uncompressed image that undergoes a wavelet transformation phase. Then, it is sent as input to the WDR encoder, which produces the compressed image's bits. The reconstructed image is produced by importing the compressed image bits to the WDR decoder. The decoded data undergo an inverse transform procedure, producing the final reconstructed image [1].

The WDR algorithm consists of five parts, as shown in Fig. 3.2. In the Initialization section, an initial threshold value of T_0 is selected so that all transform values are less than T_0 and at least one is greater than or equal to $T_0 = 2$. The purpose of the loop indicated in Fig. 3.2 is to encode significant transformation values by the bit-level encoding method. In relation to the quantity T_0, a binary expansion is calculated for each transformed value. The loop is the process by which these binary extensions are calculated. As the threshold is halved, the significance pass and refinement pass calculate the next bit.

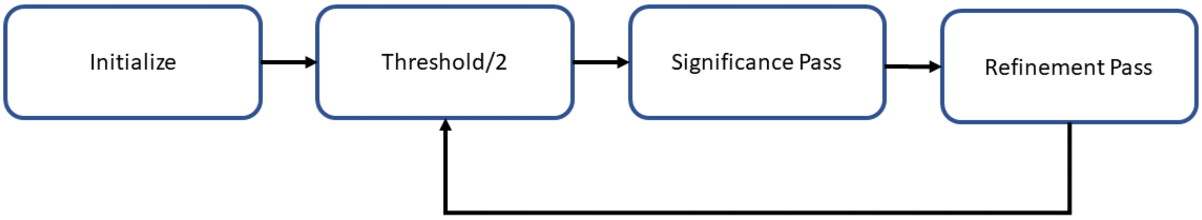


Figure 3.2: WDR compression diagram. The WDR algorithm initially calculates the image's DWT, then classifies the particle transformation coefficients from the largest scale to the finest scale and sets an initial T-threshold. The next step is the significance pass which involves two stages. The first stage is finding the relative to the T-threshold significant coefficients positions, while the second stage exports the significant coefficients. The next step is the refinement pass which gets the improvement values from all significant factors except those found in the classification step of the current iteration round. Finally, the loop divides the threshold by two and repeats the process from step 2 [1].

The general model of the WDR method is shown in the following distinct steps of the algorithm.

- Initialize: Calculate the DWT of the original image
- Threshold/2: Classify the particle transformation coefficients from the largest scale to the finest scale and set an initial T threshold.
- Significance pass: Find the significant coefficients' positions relative to the T-threshold and export these significant coefficients.
- Refinement pass: Get the improvement values of all significant factors, except those found in the classification step of this iteration round.
- Loop: Divide the threshold T by two and go to step 2.

Specifically, each step of the WDR algorithm can be seen in Algorithm 3.1.

The WDR method is equipped with a built-in encoding scheme that can achieve any compression ratio and is competitive with other image compression algorithms.

In embedded wavelet-based coding, the significance map forms a binary image; consequently, techniques that have been employed for the coding of bi-level images apply to significance-map coding. For example, run-length coding has a long history

Algorithm 3.1 The WDR method

Step 1: (Initialize). Choose an initial threshold T_0 so that all transform values satisfy $|x_m| < T_0$ and at least one transform value satisfies $|x_m| \geq \frac{T_0}{2}$

Step 2: (Update threshold). Let $T_k = \frac{T_{k-1}}{2}$

Step 3: (Significance pass). Perform the following procedure while scanning through insignificant values for higher thresholds

Initialize step counter $C = 0$
Let $C_{old} = 0$
repeat
  Get next insignificant index m
  Increment step-counter C by 1
  if $|x_m| \geq T_k$ **then**
    Output sign x_m and set $q_m = sgn(x_m) \cdot T_k$
    Move m to the end of significant indices sequence
    Let $n = C - C_{old}$
    if $n > 1$ **then**
      Output reduced binary expansion of n
    else if $|x_m| < T_k$ **then**
      Let q_m retain its initial value of 0
    end if
  end if
until end of insignificant indices
Output end-marker
The end-marker is a plus sign followed by the reduced binary expansion of $n = C + 1 - C_{old}$ and a final plus sign

Step 4: (Refinement pass). Scan through significant values found with higher threshold values T_j, for $j < k$ (if $k = 1$ skip this step). For each significant value x_m, do the following

if $|x_m| \in [|q_m|, |q_m| + T_k]$ **then**
  Output bit 0
else if $|x_m| \in [|q_m| + T_k, |q_m| + 2T_k]$ **then**
  Output bit 1
  Replace value of q_m by $q_m + sgn(q_m) \cdot T_k$
end if

Step 5: (Loop). Repeat steps 2 through 4 (exiting at any point if bit budget is exceeded)

Algorithm 3.2 Proposed 3D-WDR-MCPD algorithm.

Input: Original uncompressed image.

Compute the average $MCPD$ values for each slice using eq. (3.1) .

Predefine a threshold $T_{MCPD} = 0.5$.

Select slices with the highest spatiotemporal coherence: $MCPD \leq T_{MCPD}$ and construct the 3D volume.

Employ Algorithm 3.1 to compress the optimal 3D volume.

Output: Compressed image.

of such binary-coding use. The wavelet difference reduction algorithm combines run-length coding of the significance map with an efficient lossless representation of run-length symbols to produce an embedded image coder.

WDR was originally developed as a 2D encoder but is straightforwardly extended to 3D [161]. Also, WDR can be extended to shape adaptive by "skipping" over flat regions and not coding any significant information for them or including them in the run-length. This 3D extension deploys the run-length scanning as a 3D raster scan of each subband of the 3D discrete wavelet transform, which is easily accomplished in either dyadic or packet DWT decompositions.

The proposed method is an extension of the 3D-WDR method. Specifically, we extended the 3D-WDR method by using the mean co-located pixel difference (MCPD) to estimate the optimal number of frames that can be encoded given the best peak signal-to-noise ratio. MCPD measures the temporal difference between slices on the pixel values. The MCPD between two slices of dimensions $N \times M$ i.e., slice **x** and slice **y** for each pixel is computed as:

$$MCPD = \frac{1}{NM} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} |x(i,j) - y(i,j)| \,, \tag{3.1}$$

where $x(i,j)$ and $y(i,j)$ correspond to the pixel value at position (i,j) of slices **x** and **y**, respectively.

In particular, we compute the $MCPD$ of each slice i with all the slices that follow, where $i = 1, \ldots K$, and K being the total number of slices, for each MRI volume, (e.g., slice 1 with slices $2, 3, 4, \ldots K$, slice 2 with slices $3, 4, 5, \ldots K$) and construct a volumetric image consisting of only the slices that exhibit an average $MCPD$ value ≤ 0.5 to keep those with similar spatial and temporal coherence. After extensive evaluation and cross-validation, the threshold of 0.5 has been found to be the optimal value for selecting the most similar slices in the spatial and temporal domain. In case of volumetric data such as MRI and CT exams, the term “temporal coherence” may refer to the similarity between different slices. Note that the slices that are not selected to be part of any volumetric image are compressed separately using the standard WDR method, using the same bit rate in bits per voxel as the rest of the slices. The main steps of the proposed method are summarized in Algorithm 3.2.

3.3 Results

Dataset: To evaluate our method, a widely-used publicly available dataset named the cancer image archive (TCIA) [162] was used. This is a collection of medical de-identified datasets related to a common disease such as lung cancer or brain tumor from 20 subjects with primary newly diagnosed glioblastoma. For each subject, two MRI exams of brain tumor are included in DICOM format containing 16-bit images with at least 20 slices per MRI exam. We used a 20-slices dataset of brain tumor.

Evaluation metrics: For the evaluation of the results, we computed the peak signal-to-noise ratio (PSNR) and the structural similarity index measure (SSIM) [157]. PSNR computes the peak signal-to-noise ratio between two images, in decibels (dB). This ratio is a quality measurement between the original and the compressed image. PSNR can take values up to infinity, the higher the PSNR, the better the quality of the compressed image. Since the MRI exams in the TCIA dataset contain 16-bit images, in this case, the PSNR is computed as:

$$PSNR = 10\log_{10}\left(\frac{(2^{16}-1)^2}{MSE}\right), \tag{3.2}$$

where $MSE = \frac{\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}(x(i,j)-\hat{x}(i,j))^2}{NM}$ with $x(i,j)$ and $\hat{x}(i,j)$ correspond to the pixel value at position (i,j) of the ground truth $\mathbf{x}$ (original uncompressed image) and the compressed image $\hat{\mathbf{x}}$ of dimensions $N \times M$, respectively. Note that, the term $2^{16}-1$ is the maximum pixel value in the input image data type. SSIM is a metric that represents a visual distortion between a reference image and the observed/compressed image. The SSIM is a function between two images $\mathbf{x}$ and $\hat{\mathbf{x}}$ and is computed between pairs of local square overlapping windows x and $\hat{x}$ of the two images, respectively:

$$SSIM(x,\hat{x}) = \frac{(2\mu_x\mu_{\hat{x}}+C_1)(2\sigma_{x\hat{x}}+C_2)}{(\mu_x^2+\mu_{\hat{x}}^2+C_1)(\sigma_x^2+\sigma_{\hat{x}}^2+C_2)}, \tag{3.3}$$

where μ_x and $\mu_{\hat{x}}$ denote the mean intensity of the ground truth and the compressed image, respectively, σ_x and $\sigma_{\hat{x}}$ are the standard deviations at patches x and $\hat{x}$ of the two images, $\sigma_{x\hat{x}}$ is the covariance of x and $\hat{x}$, and C_1 and C_2 are constants added to avoid instability. Values close to 1 indicate that the compressed image preserves high visual quality (i.e., identical patches). Finally, the mean SSIM index value is computed to evaluate the total image similarity.

First, we computed PSNR and SSIM values for the 20 consecutive slices to build the 3D model used for compression. Table 3.1 shows the results in terms of PSNR and SSIM for 20 slices using the 3D-WDR algorithm at bit rate one bit per voxel. As it can be seen, PSNR varies from 46.9228 to 51.3733 and SSIM varies from 0.6403 to 0.7841.

Table 3.2 depicts the average $MCPD$ values for all slices, while Figure 3.3 depicts the distribution of MCPD. It can be observed that $MCPD$ is less than 0.5 for the slices $2, 4, 6, 15, 18, 19, 20$ (the numbers in bold), in that sense a threshold of 0.5 is considered as a reasonable choice to form our 3D model.

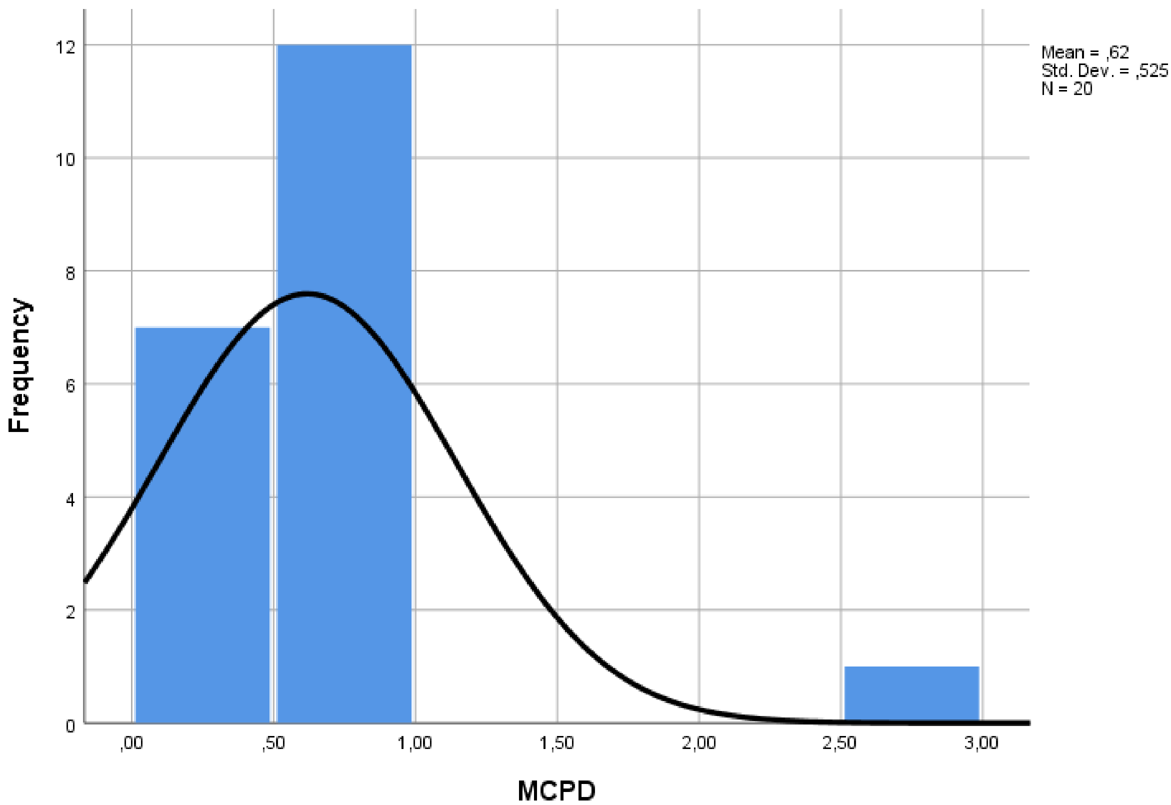


Figure 3.3: Distribution of MCPD values [1].

In Table 3.3, we report the results of the proposed 3D-WDR-MCPD method in

Table 3.1: Evaluation of standard 3D-WDR model in terms of PSNR and SSIM for the 20 slices of the MRI exam [1].

Slice index	**PSNR**	**SSIM**
1	48.0704	0.7484
2	50.8077	0.7263
3	49.2582	0.7670
4	48.7388	0.7124
5	51.3304	0.7490
6	48.9123	0.6838
7	48.0288	0.7515
8	49.3198	0.7133
9	49.4259	0.6935
10	46.9228	0.7815
11	49.7743	0.7841
12	48.7383	0.6992
13	51.0344	0.7106
14	47.5173	0.7577
15	49.2868	0.7049
16	48.3499	0.7481
17	51.3733	0.6467
18	49.0812	0.6800
19	49.3517	0.6917
20	48.0205	0.6403

terms of PSNR, SSIM, and PSNR improvement. Specifically, PSNR varies from 50.7433 to 52.2776 and SSIM varies from 0.6834 to 0.7810. Also, note that the PSNR improvement with respect to the standard 3D-WDR algorithm (Table 3.1) is remarkably high in terms of absolute dB difference. For instance, PSNR for slice 2, when the standard 3D-WDR method was employed (Table 3.1), is 50.8077 and SSIM is 0.7263, while when the proposed 3D-WDR-MCPD method is used PSNR and SSIM increase to 52.2776 and 0.7578, respectively. This corresponds to a PSNR improvement of approximately 1.5 dB, which also corresponds to visual improvement. The maximum

Table 3.2: Estimated MCPD values for the 20 slices of the MRI exam [1].

Slice index	**MCPD average**
1	0.6009
2	**0.4716**
3	0.5357
4	**0.4882**
5	0.5173
6	**0.4784**
7	0.5229
8	0.5806
9	0.5325
10	2.7478
11	0.6987
12	0.6331
13	0.6603
14	0.5302
15	**0.4501**
16	0.6162
17	0.6703
18	**0.3852**
19	**0.2561**
20	**0.0000**

PSNR improvement is achieved for slice 12 as it reaches 5.13 dB, which indicates that compressing similar slices (seven slices) in one volume per time is more efficient than directly compressing the 3D image as a whole.

Figures 3.4 and 3.5 depict qualitative results when the proposed 3D-WDR-MCPD method is employed for compressing slice 2 and slice 4, respectively. For the depicted MRI images the compression rate is achieved at one bit per voxel ($16:1$) compression ratio. Also, the visual perception of the compressed image is retained and is very close to the original uncompressed image. The main reason behind this is the smooth transition of images throughout the MRI slices.

Table 3.3: Evaluation of the proposed method in terms of PSNR, SSIM and PSNR improvement for the seven slice indexes originated from the computed MCPD values in Table 3.2 given the predefined threshold. Numbers in bold refer to the slices compressed as one volume with the proposed method, while the rest of the slices are compressed with WDR, one-by-one [1].

Slice index	**PSNR**	**SSIM**	**PSNR Improvement (in dB)**
1	48.4718	0.7448	0.4
2	**52.2776**	**0.7578**	1.47
3	50.9928	0.7415	1.73
4	**50.6824**	**0.7810**	1.94
5	51.3411	0.7541	0.01
6	**50.7814**	**0.7503**	1.87
7	50.2464	0.7282	2.22
8	51.2849	0.8430	1.97
9	51.5872	0.8050	2.16
10	47.3742	0.7639	0.45
11	50.0361	0.7534	0.26
12	53.8705	0.8507	5.13
13	51.5479	0.8445	0.51
14	52.3852	0.7261	4.87
15	**51.0053**	**0.7596**	1.72
16	51.0327	0.7264	2.68
17	52.0943	0.8316	0.72
18	**51.9189**	**0.7742**	2.84
19	**50.7433**	**0.7375**	1.39
20	**51.8223**	**0.6834**	3.80

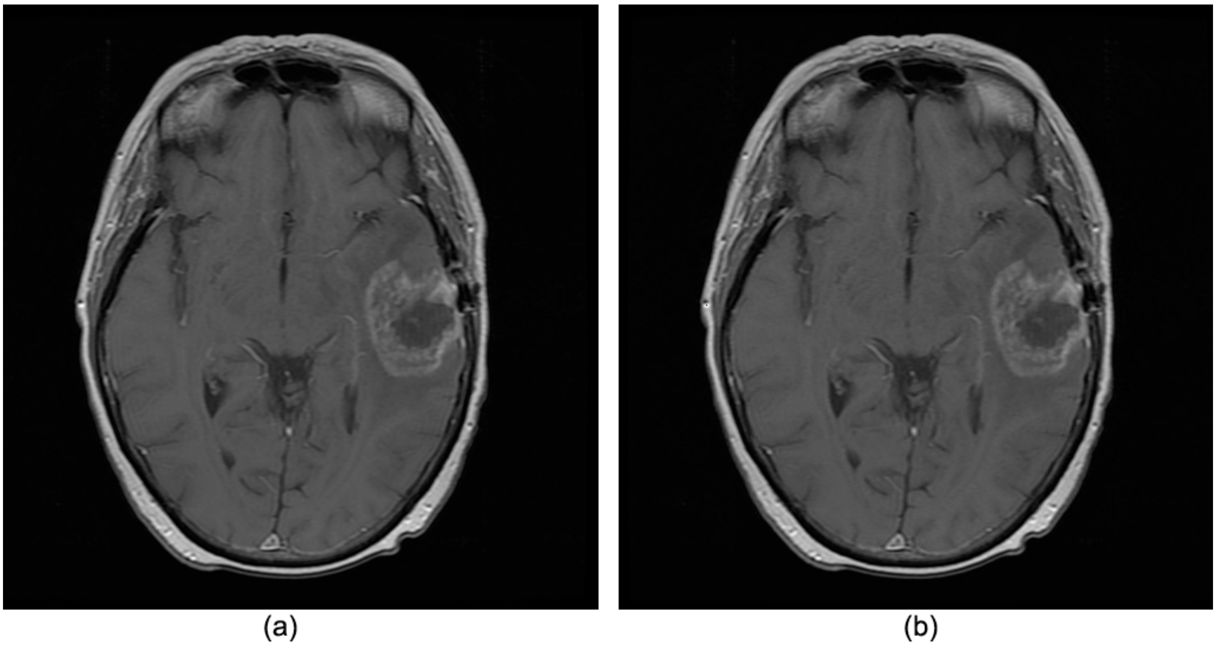


Figure 3.4: Slice 2: (a) Original uncompress slice and (b) compressed slice using the 3D-WDR-MCPD method for compression [1].

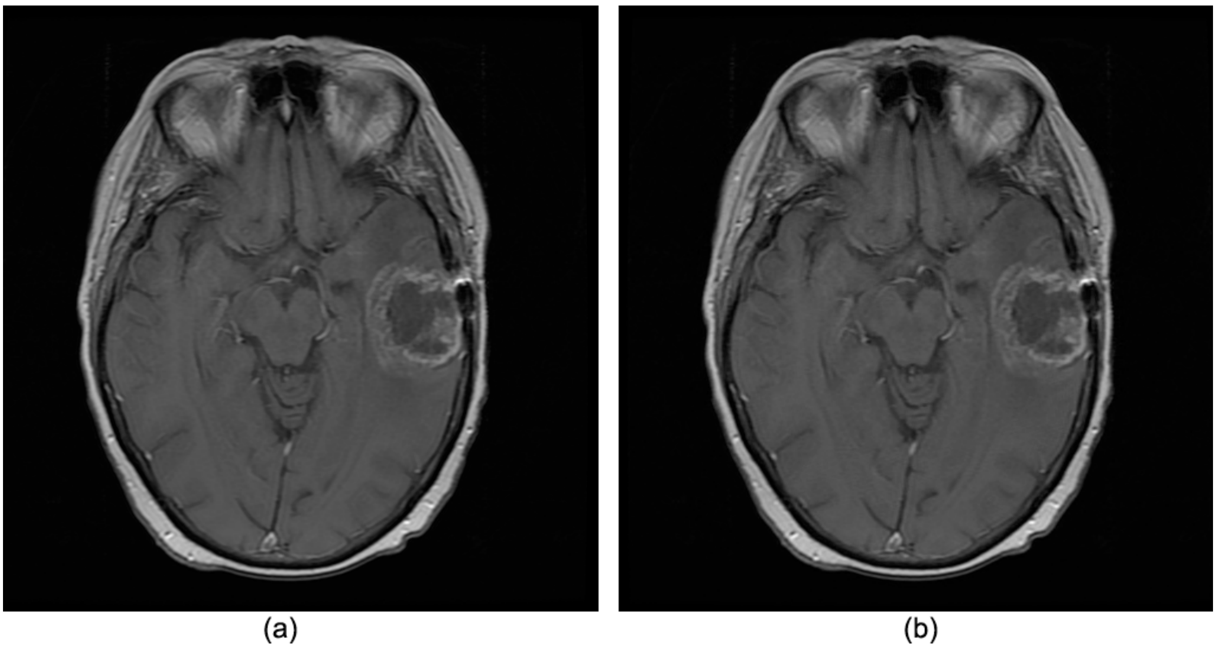


Figure 3.5: Slice 4: (a) Original uncompress slice and (b) compressed slice using the 3D-WDR-MCPD method for compression [1].

3.4 The CWDR Method

3.4.1 Introduction

According to our knowledge, the compression of microscopic images using set partitioning methods has not been tried yet. Color pictures display more information than grayscale images since color pictures display the same number of grayscale tones as in grayscale images plus a number of colors on every image, thus, improving contrast resolution. From a medical point of view, color images disclose important information, which can be critical for diagnostic purposes. Therefore, it motivated us to propose an extension of the original WDR method to effectively compress microscopic images, namely color wavelet difference reduction (CWDR).

Our main contribution is designing a medical image compression method that can be easily reproduced. The method proposed in this section was published in 2023 [2]. It is suitable for use in various color medical images of big sizes, such as microscopic images. It achieves state-of-the-art compression results with a high compression ratio and small information loss within an acceptable range. Extensive evaluations have been performed in a custom-created dataset containing image data extracted using the Hamamatsu NanoZoomer 210. The dataset contained 31 slides of colorectal cancer, and the proposed CWDR algorithm achieved high compression ratios in all images while maintaining high visual quality.

3.4.2 Method

The family of set partitioned methods was initially designed for grayscale image compression. To apply them to color images, we must first understand color space. The color image is usually in RGB format. The RGB color spaces are highly correlated, so transformation to a less correlated space is required for efficient lossy compression. The original RGB images were transformed using standard transformations to code the YC_bC_r color space such that the luminance channel Y is stored as one byte for each pixel. On the other hand, the two chrominance channels are stored as one byte for each block of, say, $n \times n \times n$ pixels, i.e., Cb and Cr are the blue component and red component related to the chroma component.

The proposed method, CWDR, is an extension of WDR for color images. This method compresses each color plane at the coding stage and generates three separate bitstreams of the same bitrate. Then, the generated bitstream of each color space would be serially concatenated. The proposed system structure flowchart is shown in Fig. 3.6 while its algorithmic structure can be seen in Algorithm 3.3. The operation starts by selecting suitable colors and image scales. The next step represents the application of variable filters of the wavelet transforms. Then, some quantization processes are performed to show the elements of a big set in terms of a smaller set to lower the number of bits necessary to indicate all possible values of mapping outputs to fewer bits.

In the RGB color model, a color image can be represented by the following intensity function.

$$I_{RGB} = (F_R, F_G, F_B) \tag{3.4}$$

where $F_R(x, y)$ is the intensity of the pixel (x, y) in the red channel, $F_G(x, y)$ is the intensity of pixel (x, y) in the green channel, and $F_B(x, y)$ is the intensity of pixel (x, y) in the blue channel. The intensity of each color channel is usually stored using eight bits, which indicates that the quantization level is 256. That is, a pixel in a color image requires total storage of 24 bits. A 24 bit memory can express $2^{24} = 16777216$ distinct colors. The number of colors should adequately meet the display effect of most images. Such images may be called true color images, where each pixel's information is kept using a 24-bit memory.

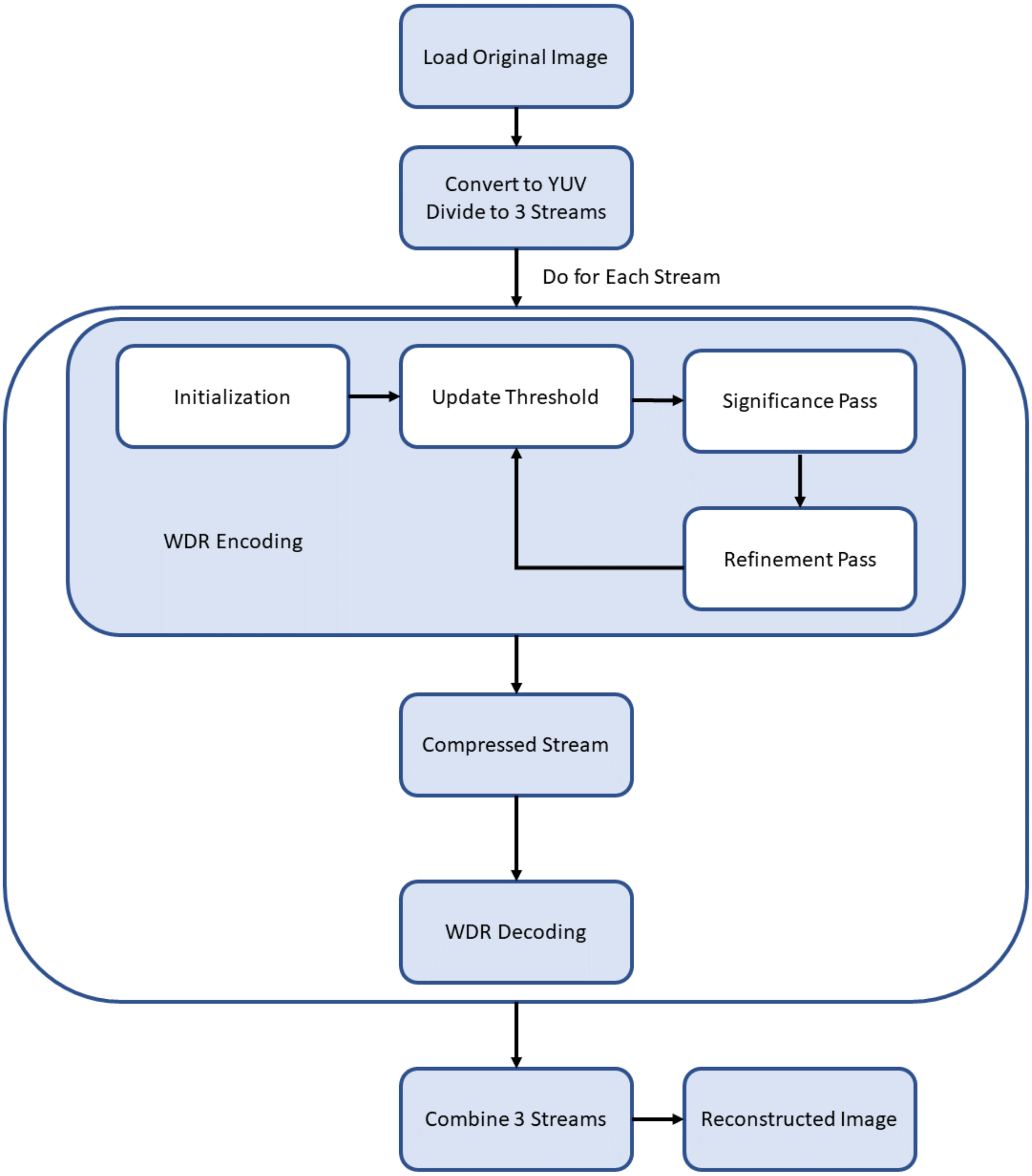


Figure 3.6: System Flowchart [2].

To split the RGB image into three streams, we separately save each channel to different variables as seen in (3.5).

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} I_{RGB}(F_R, 0, 0) \\ I_{RGB}(0, R_G, 0) \\ I_{RGB}(0, 0, R_B) \end{bmatrix} \tag{3.5}$$

After splitting the RGB image into three streams, it was converted to the YUV format using the following formula.

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.14713 & -0.28886 & 0.436 \\ 0.615 & -0.51499 & -0.10001 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{3.6}$$

Thus, the color wavelet difference reduction (CWDR) algorithm follows the basic concepts of WDR [163] by incorporating extra features that aggregate the coefficients to an area of interest. It is suitable for compressing medical images at a low bit rate per pixel.

Algorithm 3.3 The CWDR Algorithm

Input: Original uncompressed image.

1: Convert to YUV- Divide into three streams.
2: **for all** streams **do**
3: Calculate the DWT of the stream.
4: **while** (Predetermined number of bits is not reached) **do**
5: Sort the wavelet transform coefficients from the larger scale to the finer scale.
6: Set an initial threshold: $T_n = 2^N$, where $N = \log_2(\max_{(i,j)} \vee\gamma(i,j))$, with $n = 1$, where $\gamma(i,j)$ are the wavelet coefficients in the set of non-significant coefficients and N is the total number of bit planes.
7: Sorting pass: Find the positions of the significant coefficients concerning the threshold, and keep the coefficients that satisfy the condition: $\gamma(i,j) \geq T_n$.
8: Improvement process: Get the improvement rates of all significant coefficients, except those found in the sorting pass step of the current iteration.
9: Update threshold: $n = n + 1; T_{n-1} = T_n; T_n = \frac{T_n}{2}$
10: **end while**
11: **end for**
12: Combine three streams.

Output: Compressed image.

After the WDR process, we converted the YUV image back to RGB, using the following formula:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.164 & 0 & 1.596 \\ 1.164 & -0.392 & -0.813 \\ 1.164 & 2.017 & 0 \end{bmatrix} \begin{bmatrix} Y \\ U \\ V \end{bmatrix} \tag{3.7}$$

Finally, we combined the three streams back into one RGB image. To combine the three streams into an RGB image, we save them to one variable, a.k.a.:

$$RGB = I(R, G, B) \tag{3.8}$$

3.4.3 Results

The proposed method is evaluated in the demanding field of histopathological microscopy image analysis. The diagnosis and prognosis systems based on histological image analysis present significant growth during the last five years, utilizing whole slide scanning technologies, computational resources management, distributed systems, and multiple cores. According to the medical question, histological microscopy images are extracted using standard tissue preparation procedures. The employed dataset has been extracted using the Hamamatsu NanoZoomer 210, scanning 31 slides of colorectal cancer. The scanning system provides two optical magnification options (20x and 40x), which can scan 210 slides automatically. According to digitalization, each pixel of a Whole Slide Image (WSI) corresponds to a physical area of several tens of nm^2.

Specifically, in 40× magnification mode, the Hamamatsu NanoZoomer scanner extracts an image where the size of each pixel edge corresponds to 227 nm. The above image digitization procedure provides an appropriate resolution for most histological findings. In most cases, the extracted images are stored in compressed JPEG-based or uncompressed TIFF format. The resolution of a typical WSI in 40x magnification is about 100K x 100K pixels, whereas an uncompressed format could require hundreds of GBs of memory. Commonly, the challenge of compressing images focuses on image size minimization, along with the high performance of quality measures (Signal to Noise Ratio - SNR, Peak Signal to Noise Ratio - PSNR, Structural Similarity – SSIM). However, the most significant issue for histological microscopy images must be the image quality assessment of the medical regions of interest, such as cells and nuclei, cell degeneration and cancer, inflammation and fibrosis areas, and other histological lesions. For the above reasons, parts from each WSI after applying the compression procedure, were extracted for evaluation by a specialist.

Table 3.4, Fig. 3.7, and Fig. 3.8 show the results in terms of PSNR and SSIM for the 31 images of our dataset using the CWDR algorithm at a bit rate of 0.5 bit per pixel, which gives a compression ratio of 16. It can be seen that PSNR varies from 32.01 to 47.49. On the other hand, SSIM varies from 0.68 to 0.98, with 30 out of 31 images having an SSIM value of 0.92 or higher.

Table 3.4: Evaluation of CWDR model in terms of PSNR and SSIM for the 31 images [2].

	Image	PSNR	SSIM
	1768-18TH_1	41.1934	0.9473
	1870-18H_1	47.2838	0.9308
	1870-18H_2	45.1102	0.9344
	1870-18H_3	33.6743	0.9393
	1884-1921_1	36.2088	0.8996
	2529-20H_1	37.6522	0.9316
	2529-20TH_1	37.6522	0.9220
	3211-20AH_1	37.8373	0.9245
	3211-20B_1	36.3347	0.9200
	3469-18H_1	45.7910	0.9300
	4015-14F_1	45.0333	0.9268
	4015-14L_1	44.4740	0.9369
	4339-20N_1	38.9031	0.9311
	4339-20X_1	37.1588	0.9300
Bit Rate=0.5 bpp CR= 16	5820-14N_1	47.4880	0.9333
	5820-141I_1	45.1940	0.9288
	6448-19MX_1	35.8616	0.9300
	7438-16C_1	44.5228	0.9583
	7438-16D_1	43.8831	0.9600
	7870-18A_1	46.3653	0.9400
	7995-17DK_1	45.3092	0.9300
	8036-1BK_1	42.5244	0.9401
	8036-1BK_2	41.8953	0.9288
	8036-1BK_3	42.6988	0.9377
	8036-1BK_4	42.8339	0.9384
	8036-1BK_5	42.6134	0.9391
	8036-1BP_1	43.3838	0.9330
	8036-1BP_2	41.9381	0.9292
	8036-1BP_3	40.7767	0.9404
	8036-1BP_4	41.7617	0.9292
	8036-1BP_5	45.7987	0.9299

Regarding 8-bit images like ours, PSNR values over 30 indicate a very good image quality, and values over 40 indicate that the image quality is excellent (i.e., very close to the original image). It can be seen that the proposed method gives auspicious results since 23 out of 31 images have PSNR values over 40.

Additionally, SSIM values are high (which implies that the two images are very similar), further supporting the promising results of our method.

The Wilcoxon signed-rank test was used to compare the PSNR values of the proposed method with the respective values for DWT, JPEG 2000, HEIC (high-efficiency image format) and WEBP (web picture format) methods. High-Efficiency Image File Format is a container format for storing individual digital images and image sequences. The standard covers multimedia files that can also include other media streams, such as timed text, audio and video. WEBP is a modern image format that provides superior lossless and lossy compression for images on the web. The Wilcoxon signed-rank test is a non-parametric statistical hypothesis test used either to test the location of a population based on a sample of data or to compare the locations of two populations using two matched samples. The results obtained with those statistical tests are shown in Fig. 3.9 and indicated statistically significant differences between the CWDR and the other four methods.

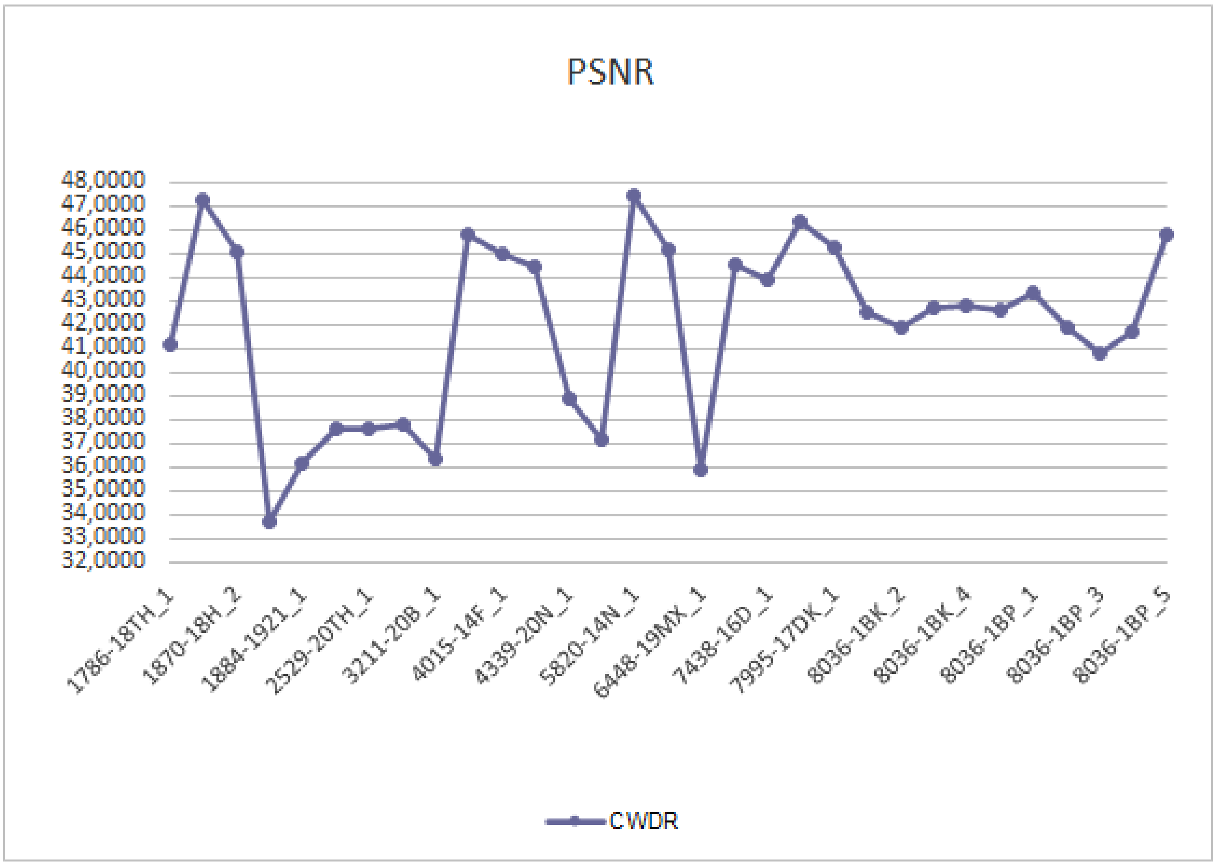


Figure 3.7: PSNR values of the 31 images. This figure depicts the PSNR values for all 31 images. It can be seen that the proposed CWDR method managed to achieve values over 30 in all cases, which indicates very good results. Additionally, 74.19% of the cases got an even higher PSNR value (> 40), indicating excellent image quality [2].

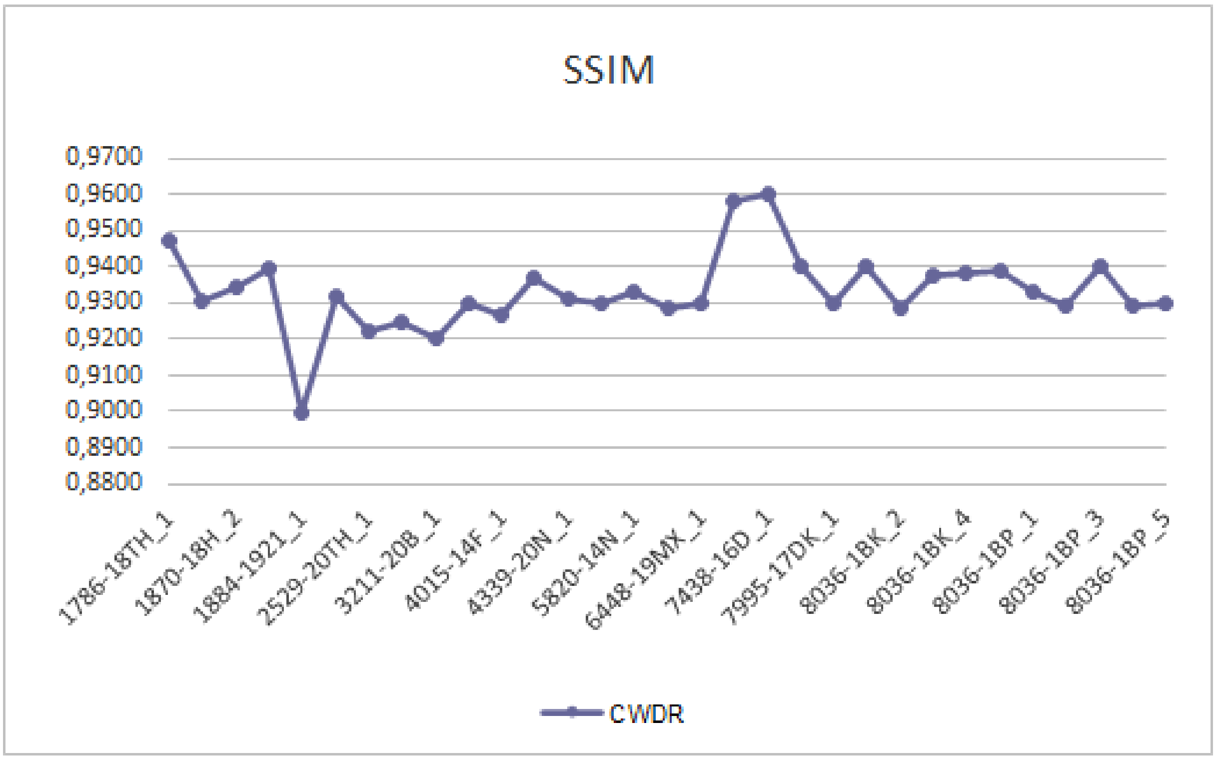


Figure 3.8: SSIM values of the 31 images. This figure depicts the SSIM values for all 31 images. It can be seen that the proposed CWDR method managed to achieve values very close to 1 in 30 out of 31 cases which indicates that the compressed images are completely the same as the original uncompressed ones [2].

We also compared the SSIM values of the proposed method with the respective values for Lossless Compressing Using DWT Technique [3], JPEG 2000, HEIC, and WEBP methods using the Wilcoxon signed-rank test. The results obtained with those statistical tests are shown in Fig. 3.10 and indicated statistically significant differences between the CWDR and the other four methods.

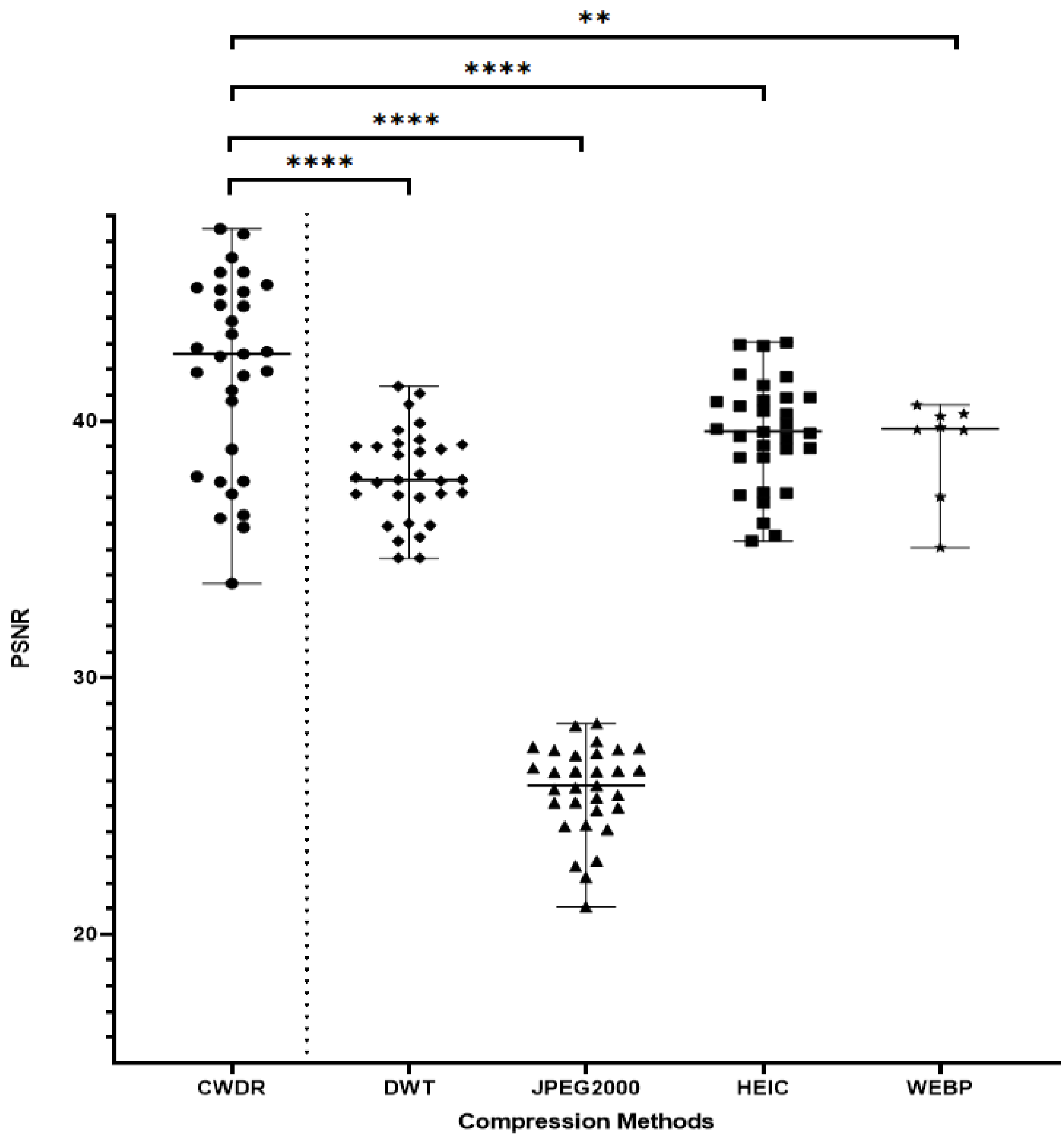


Figure 3.9: Scatter plot representation and the Wilcoxon signed-rank test results of the comparison for each of the two well-known compression methods (Lossless Compressing Using DWT Technique [3], JPEG 2000, HEIC and WEBP) with the CWDR method regarding PSNR values. Stars links join significantly different values; two stars (**) typically indicate that the result is very statistically significant, usually at the 0.01 level (p-value < 0.01), three stars (***) typically denote a significance level of 0.001, indicating that the p-value of the test result is less than 0.001, which means that if the null hypothesis tested were indeed true, there would be a one in 1,000 chance of observing results at least as extreme. Four stars (****) are less commonly used than one, two, or three asterisks in standard practice. If used, they might denote an extremely high level of significance, possibly at the 0.0001 level (p-value < 0.0001) [2].

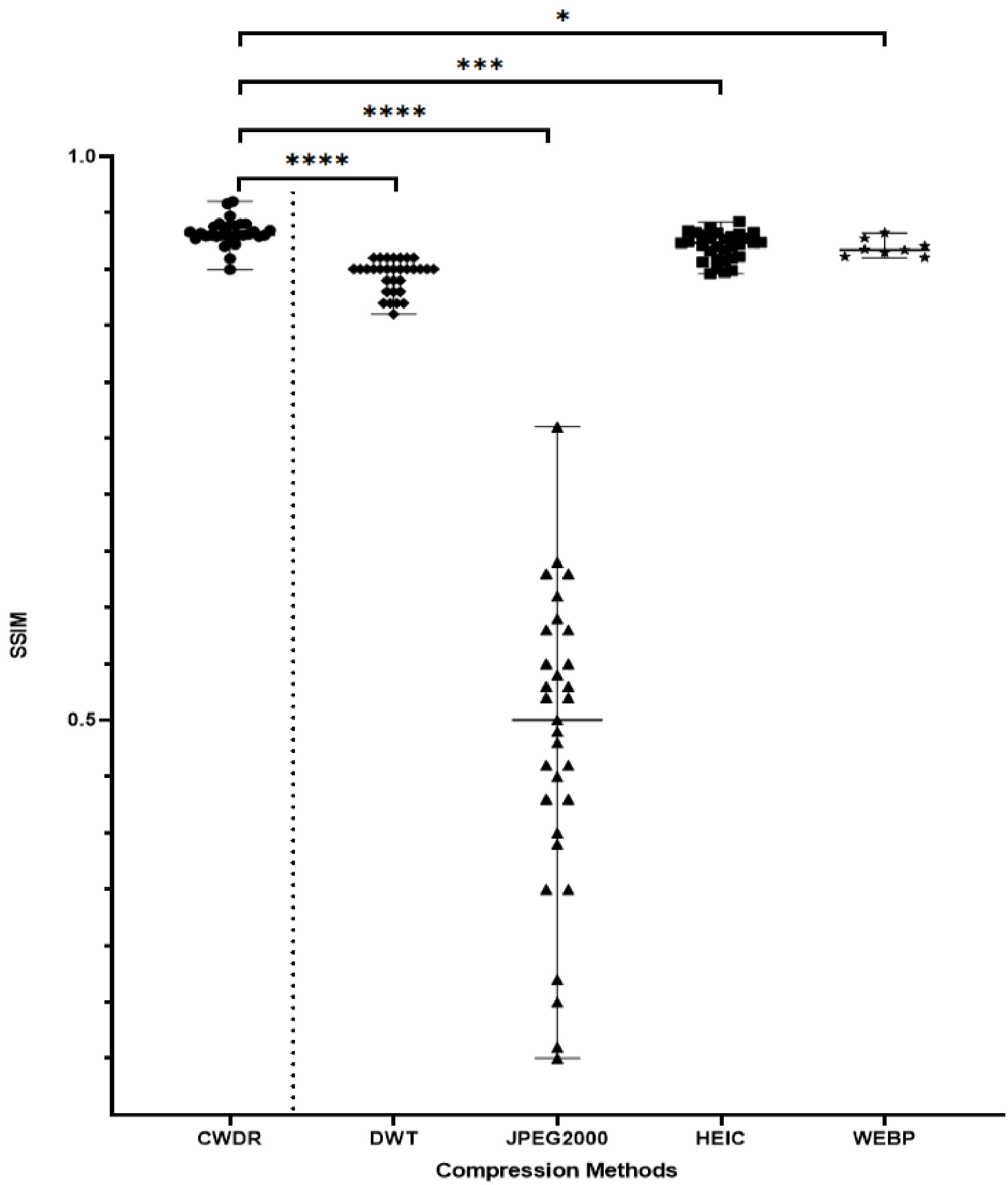
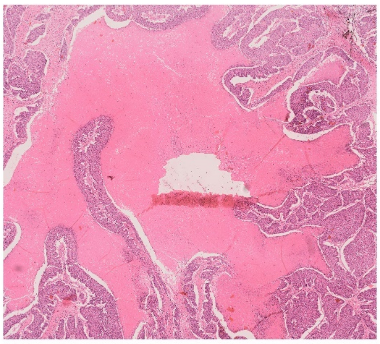
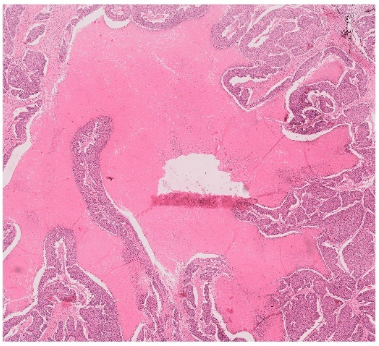


Figure 3.10: Scatter plot representation and the Wilcoxon signed-rank test results of the comparison for each of the two compression methods (Lossless Compressing Using DWT Technique [3], JPEG 2000, HEIC and WEBP) with the CWDR method regarding the SSIM values. Stars links join significantly different values; three stars (***) stand for $p < 0.001$ [2].

It should be noted that the images that WEBP method can compress are limited to 16383 pixels in height and width. Therefore, we could only compress 16 out of our 31 images. It can be seen in Fig. 3.11, and Fig. 3.12, that the decoded images maintain all the diagnostically important information. Thus, they can be considered as ”visually and diagnostically lossless.” Even in Fig. 3.12, which depicts the results of the image with the worst SSIM value (image 1870-18H_3), the result is visually and diagnostically lossless.



(a)



(b)

Figure 3.11: 7438-16D_1: (a) Original uncompressed image and (b) compressed image using the CWDR method for compression. It can be seen from both images that they are visually and diagnostically lossless [2].

3.4.4 Discussion

The proposed method is evaluated in the demanding field of histopathological microscopy image analysis. Utilizing the advantages of whole slide scanning technologies, computational resources management, distributed systems, and multiple cores, the diagnosis and prognosis systems based on histological image analysis presented significant growth during the last five years. Histological microscopy images are extracted after standard tissue preparation procedures. The employed dataset has been extracted using the Hamamatsu NanoZoomer 210, which scanned 31 slides of colorectal cancer. The scanning system provides two optical magnification options ($20\times$ and $40\times$), which can scan 210 slides automatically. According to digitization, each pixel of a whole slide image (WSI) corresponds to a physical area of several tens of nm^2. Specifically, in $40\times$ magnification mode, the Hamamatsu NanoZoomer scanner extracts an image where the size of each pixel edge corresponds to 227 nm. Such image digitization provides appropriate resolution for most histological findings. In most cases, the extracted images are stored in compressed JPEG-based or uncompressed TIFF format. The resolution of a typical WSI in $40x$ magnification is about 100K $\times$ 100K pixels, which uncompressed format could require hundreds of GBs of memory. Commonly, the challenge of compressing images focuses on image size minimization, along with a high performance of quality measures (SNR, PSNR, and SSIM). However, the most significant issue for histological microscopy images must be the image quality assessment of the medical regions of interest, such as cells and nuclei, cell degeneration and cancer, inflammation and fibrosis areas, and other histological lesions. Due to this, after applying the compression procedure, parts from each WSI have been extracted for evaluation by a specialist.

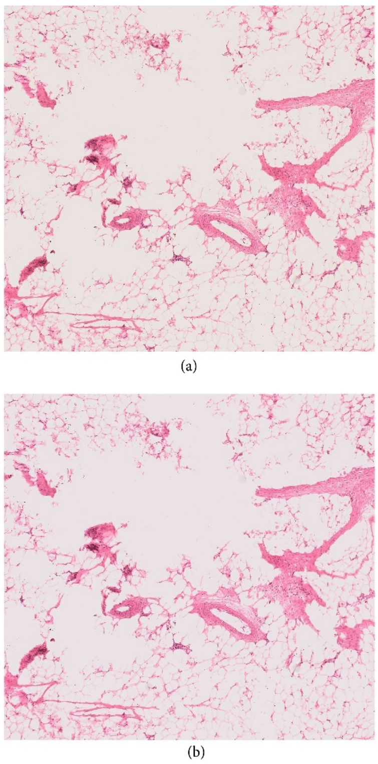


Figure 3.12: 1870-18H_3: (a) Original uncompressed image and (b) compressed image using the CWDR method for compression. It can be seen from both images that they are visually and diagnostically lossless [2].

The perceived quality of the compressed images was evaluated with a Mean Opinion Score (MOS) scale ranging from 1 to 5 (bad, poor, fair, good, and excellent). Even though its suitability may be debatable, the MOS scale provides a different method of gauging the quality and depicting how specialists evaluates it. Four qualified histopathologists evaluated the quality of the compressed images without blindly consulting each other. Specifically, the histopathologists asked to evaluate all the images regarding the diagnostic value, by accessing them one-by-one. The proposed compression method met high qualitative criteria obtaining similar image quality rating score, with a statistically significant association comparing to original uncompressed images (Fig. 3.13).

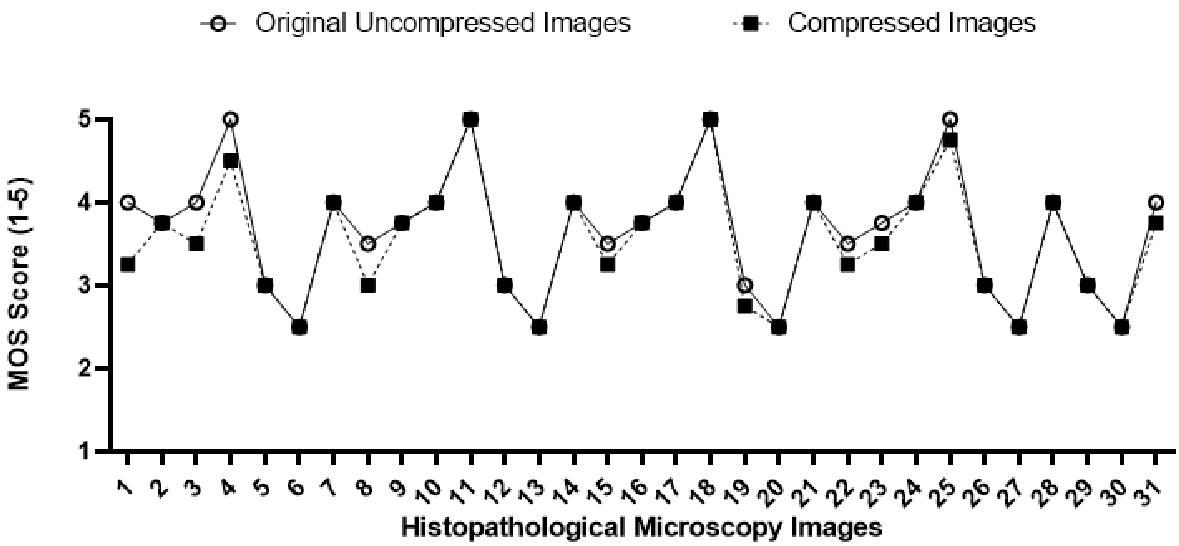


Figure 3.13: Spearman's rank-order correlations were to examine the association between the MOS score of original uncompressed images and compressed images using the proposed method. There were positive and significant associations between the two MOS scores, ($rs = 0.9539$, $N = 31$, $p < 0.001$), while Wilcoxon signed-rank test showed no significant differences between the two MS scores ($p > 0.05$) [2]

3.5 Summary

Medical image compression plays an important role in efficient storage, transmission, and management of these datasets. In this chapter, a 3D image compression model based on discrete wavelet transform was proposed, which is a clinically acceptable option for medical image compression. In particular, we extended the standard WDR method using $MCPD$ to select the optimal number of slices that exhibit the highest similarity in the spatial and temporal domain. The slices with large spatiotemporal coherence are then encoded together as one volume in terms of higher PSNR and SSIM. It is found that the perceptual quality of the medical image is remarkably high. The results indicate that the PSNR improvement over existing schemes may reach up to 3.8 dB and they may guide us through the implementation of a mobile and web platform that may be used for the compression and transmission of medical images in real-time.

The proposed 3D image compression method extends the original WDR model. This method is clinically acceptable according to the histopathological microscopy image analysis. It achieves state-of-the-art compression results with a high compression ratio and slight information loss within an acceptable range. Specifically, extensive evaluations have been performed in a custom-created dataset containing in 31 slides of colorectal cancer. The proposed CWDR algorithm achieved high compression ratios in all images while maintaining high visual quality. These results might guide us through implementing a mobile and web platform that may be used to compress and transmit medical images in real-time.

Chapter 4

Video Super-Resolution Using Plug-and-Play Priors

4.1 Introduction

The Plug-and-Play Priors (PPP) framework is recognized as one of the extensively used methodologies for addressing computational imaging challenges through the integration of physical and learned models. PPP takes advantage of high-fidelity physical sensor models and robust machine learning techniques for data pre-modeling, incorporating cutting-edge reconstruction algorithms. PPP algorithms follow a cycle of minimizing data fidelity terms to uphold data consistency and enforcing learned regularization through image denoising [164]. Recent achievements of PPP algorithms span applications in biomicroscopy, computed tomography, magnetic resonance imaging, and joint ptychotomography [5].

We introduced a PnP method for video super-resolution, using motion estimation, which has not been done yet and, afterwards, we extended it in order to use it for MRI Super-Resolution. The method proposed in this section was published in 2024 [4].

4.2 Method

The acquisition model we are assuming is:

$$\mathbf{y} = A\mathbf{x} + \varepsilon, \tag{4.1}$$

where:

- $\mathbf{y}$ is the full set of LR frames, described as $\mathbf{y} = [\ \mathbf{y}_1{}^T, \mathbf{y}_2{}^T, ..., \mathbf{y}_p{}^T]\ ^T$, where $\mathbf{y}_k, k = 1, 2, ..., p$ are the p LR images. Each observed LR image is of size $N_1 \times N_2$. Let the kth LR image be denoted in lexicographic notation as $\mathbf{y}_k = [\ y_{k,1}, y_{k,2}, ..., y_{k,M}]\ ^T$, for $k = 1, 2, ..., p$ and $M = N_1 N_2$.
- $\mathbf{x}$ is the desired HR image, of size $L_1 N_1 \times L_2 N_2$, written in lexicographical notation as the vector $\mathbf{x} = [\ x_1, x_2, ..., x_N]\ ^T$, where $N = L_1 N_1 L_2 N_2$ and L_1 and L_2 represent the up-sampling factors in the horizontal and vertical directions, respectively. $\mathbf{x}$ is the ideal un-degraded image that is sampled at or above the Nyquist rate from a continuous scene which is assumed to be band-limited.
- $\varepsilon = [\ \varepsilon_1, \varepsilon_2, ..., \varepsilon_p]\ ^T$, where ε_k is the noise vector for frame k and contains independent zero-mean Gaussian random variables.
- $A = [\ A_1, A_2, ..., A_p]\ ^T$ is the degradation matrix which performs the operations of blur, motion and subsampling.

Assuming that each LR image is corrupted by additive noise, we can then represent the observation model as [87]:

$$\mathbf{y}_k = A_k \mathbf{x} + \varepsilon_k \text{ for } 1 \leq k \leq p \tag{4.2}$$

where

$$A_k = SB_kM_k. \tag{4.3}$$

M_k is a warp matrix of size $L_1N_1L_2N_2 \times L_1N_1L_2N_2$, B_k represents a $L_1N_1L_2N_2 \times L_1N_1L_2N_2$ blur matrix, and S is a $N_1N_2 \times L_1N_1L_2N_2$ subsampling matrix. In our case $B_k = I$, since we assumed no added blur on video frames.

The goal is to find the estimate $\hat{\mathbf{x}}$ of the HR image $\mathbf{x}$ from the p LR images y_k by minimizing the cost function

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbb{R}^N} f(\mathbf{x}) \text{ with } f(\mathbf{x}) = g(\mathbf{x}) + h(\mathbf{x}), \tag{4.4}$$

where $g(\mathbf{x}) = \sum_{k=1}^{p} \frac{1}{2}\|A_k\mathbf{x} - \mathbf{y}_k\|_2^2$ is the "fidelity to the data" term, and $h(\mathbf{x})$ is the regularization term, which offers some prior knowledge about $\mathbf{x}$. In this work, we utilize the Plug-and-Play Prior methodology, where $h(\mathbf{x})$ is not explicitly defined. Instead, the ADMM algorithm is modified so that the proximal operator that depends on $h(\mathbf{x})$ is replaced by a denoising neural network [6].

We next outline the steps of the proposed algorithm.

1. The first step of our algorithm is to evaluate the term M_k from Eq. (4.3), by using optical flow motion estimation. The motion estimation method used is a popular optical flow method, called the Farneback algorithm, named after its creator, Gunnar Farneback. The Farneback algorithm generates an image pyramid, where each level has a lower resolution compared to the previous level. The Farneback method employs a dense approach, meaning it estimates the motion vector for every pixel in the image. The algorithm consists of the following steps [165]:

   (a) *Preprocessing*: The input frames are preprocessed to enhance their quality. Preprocessing steps include noise reduction (via Gaussian Blurring), image denoising (via Non-Local Means Denoising), and color space conversion (cnversion to Grayscale).

   (b) *Image pyramids*: The Farneback algorithm constructs a Gaussian pyramid for each frame. This involves creating a series of downsampled versions of the original image, forming a hierarchy of images with decreasing resolution. The pyramids enable capturing motion at multiple scales, improving the accuracy of the optical flow estimation.

(c) *Optical flow estimation*: For each level of the pyramid, the Farneback algorithm computes the optical flow using a combination of polynomial expansion and spatial filtering. It estimates the local flow vectors by calculating the phase difference between the polynomials corresponding to neighboring image patches.

(d) *Upsampling and refinement*: Once the optical flow is computed at the coarsest level of the pyramid, it is successively refined by upsampling the flow field and incorporating the local information from higher-resolution levels. This refinement process improves the accuracy of the flow estimation, particularly for small and fast-moving objects.

The result of the Farneback method is a dense optical flow field, where each pixel has an associated motion vector. These vectors represent the direction and magnitude of the motion of objects in the scene between consecutive frames.

We assume that one of the LR images, $\mathbf{y}_{mid}$ (typically the middle one), is produced from the HR image $\mathbf{x}$, by applying only downsampling, without motion shift. Thus, $M_{mid} = I$. Optical flow is calculated between $\mathbf{y}_{mid}$ and the rest of the LR images. Following that, we get M_k for the remaining $p-1$ images.

2. The second step of our algorithm is based on the PnP-ADMM method. Specifically, we run PnP-ADMM, following the steps described in Algorithm 4.1 until convergence, where $\mathbf{x}^0$ is the initial value of the HR image, obtained from $\mathbf{y}_{mid}$ multiplied by the pseudo-inverse of $\mathbf{A}_{mid}$, followed by denoising using DnCNN, while D is the image denoising operator (neural network) and g is defined as $g(\mathbf{x}) = \sum_{k=1}^{p} \frac{1}{2}\|A_k\mathbf{x} - \mathbf{y}_k\|_2^2$.

One important property of ADMM is that it does not explicitly require knowledge of $g(\mathbf{x})$ or their gradients, relying instead on the proximal operator, which is defined as:

Algorithm 4.1 PnP-ADMM [5]

$\mathbf{u}^0 = \mathbf{0}, \mathbf{x}^0,$ and $\gamma > 0$

for $k = 1, 2, ..., t$ **do**

  $\mathbf{z}^k \leftarrow prox_{\gamma g}(\mathbf{x}^{k-1} - \mathbf{u}^{k-1})$

  $\mathbf{x}^k \leftarrow D(\mathbf{z}^k + \mathbf{u}^{k-1})$

  $\mathbf{u}^k \leftarrow \mathbf{u}^{k-1} + (\mathbf{z}^k - \mathbf{x}^k)$

end for

return $\mathbf{x}^t$

$$prox_{\gamma g}(\mathbf{x}) := \arg\min_{\mathbf{x} \in \mathbb{R}^N} \{\frac{1}{2}\|\mathbf{x} - \mathbf{z}\|_2^2 \ + \gamma g(\mathbf{x})\}. \tag{4.5}$$

4.3 Proof of Convergence

The crucial conceptual observation lies in the fact that PnP algorithms incorporating black-box denoisers often struggle to address optimization problems. In other words, while the original ADMM algorithm effectively solves the optimization problem, the introduction of a black-box denoiser, denoted as D, disrupts this process by eliminating a corresponding function h for minimization. Specifically, the numerical assessment of widely employed denoisers, such as BM3D and DnCNN, demonstrates that their Jacobians lack symmetry, suggesting that these denoisers do not function as either gradient descent steps or proximal maps [166].

Nevertheless, it remains feasible to establish a criterion for the converged solution in PnP by employing a consensus equilibrium formulation, as proposed by [167].

$$\mathbf{x} = G(\mathbf{x} - \mathbf{u}) \text{ and } \mathbf{x} = D(\mathbf{x} + \mathbf{u}), \tag{4.6}$$

where $G := prox_g$ and $\mathbf{x}$, $\mathbf{u}$ are the converged values of PnP-ADMM.

Notably, within the consensus equilibrium expression in (4.6), $\mathbf{x}$ represents the final reconstruction and $\mathbf{u}$ can be construed as noise, eliminated by the denoiser in $\mathbf{x} = D(\mathbf{x} + \mathbf{u})$ on one hand and counterbalanced by the fidelity to the data effect in

$\mathbf{x} = G(\mathbf{x} - \mathbf{ux})$ on the other. To derive (4.6), it is important to recognize that the fixed points $\mathbf{z}$, $\mathbf{x}$, and $\mathbf{u}$ of the PnP-ADMM iteration satisfy

$$\mathbf{z} = G(\mathbf{x} - \mathbf{u}), \mathbf{x} = D(\mathbf{z} + \mathbf{u}), \mathbf{u} = \mathbf{u} + \mathbf{z} - \mathbf{x}. \tag{4.7}$$

From the last equation we conclude that $\mathbf{x} = \mathbf{z}$, which leads directly to (4.6). Also, the first-order optimality condition for the minimization problem $\mathbf{x} = G(\mathbf{x} - \mathbf{u}) = prox_{\gamma g}(\mathbf{x} - \mathbf{u})$ is $0 = \mathbf{x} - (\mathbf{x} - \mathbf{u}) + \gamma \nabla g(\mathbf{x})$, so $\mathbf{u} = -\gamma \nabla g(\mathbf{x})$.

The application of monotone operator theory, as outlined in [168], allows for the illustration of the convergence of PnP algorithms. In this approach, the initial phase involves identifying a fixed point for a high-dimensional operator that can be iteratively used to discover a solution, provided the appropriate assumptions are met. In the proof of PnP-ADMM convergence presented in [167], [169], the initial step is to establish a one-to-one correspondence between the fixed points of PnP-ADMM and those of the operator:

$$T := (2G - I)(2D - I). \tag{4.8}$$

After a linear coordinate transformation, Algorithm 1 is essentially identical to the Mann iterations of T, expressed as $\mathbf{v}^k \leftarrow \frac{1}{2}\mathbf{v}^{k-1} + \frac{1}{2}T(\mathbf{v}^{k-1})$ [167]. This results in linear convergence towards a unique fixed point when T functions as a contraction, a condition satisfied when g is strongly convex and $R := I - D$ serves as a suitably strong contraction [169]. Weaker conditions lead to sublinear convergence, reaching a potentially non-unique fixed point [170]. Additional notable theoretical findings on PnP-ADMM encompass its convergence for implicit proximal operators [88], applicability with bounded denoisers [171], and suitability for linearized Gaussian mixture model (GMM) denoisers [94]. Even CNN-based denoisers can be trained to meet these contractive, non-expansive, or Lipschitz conditions through the implementation of spectral normalization techniques [169], [172]. Conversely, when g exhibits only mild convexity and the denoiser D is strongly non-expansive, the iteration converges sublinearly towards its fixed point [173].

4.4 Results

We implemented our PnP method in SCICO [174], which is an open source library for computational imaging that includes implementations of PnP algorithms.

We conducted extensive experiments on benchmark subsets “calendar” and “city”, from Vid4 dataset to evaluate the performance of our proposed method. The ”Calendar” and ”City” subsets from the Vid4 dataset are commonly used in the evaluation of video super-resolution (VSR) methods. These subsets provide a diverse range of challenges for super-resolution algorithms, including varying degrees of motion, texture, and image content. The primary use of these subsets in VSR experiments is to test the algorithm’s effectiveness in enhancing video quality by increasing the resolution of both temporal and spatial dimensions. The original videos in the Vid4 dataset are in standard definition, with a spatial resolution of 720x480 pixels. This lower resolution allows researchers to demonstrate how effectively their super-resolution algorithms can upscale video to higher resolutions, such as 1080p (1920x1080 pixels) or even 4K in some experimental setups. The temporal resolution, referring to the frame rate of the videos, is generally standard across datasets used for super-resolution unless explicitly modified for specific experiments. Standard frame rates are often 30 frames per second (fps), but specific details can vary based on how the dataset was created and processed. ”Calendar” typically includes scenarios with rich textures and gradual changes in scenes, such as calendar pages turning or detailed artistic images. This tests the algorithm’s ability to handle fine details and text, which are crucial for practical applications like digital archiving or art restoration. On the other hand, ”City” involves urban environments with moving objects (e.g., cars, people), varying lighting conditions, and complex geometries. This provides a dynamic scene to test the algorithm’s robustness in handling motion and reconstructing high-frequency details in a natural setting. Both subsets provide unique challenges in terms of motion estimation and handling artifacts like aliasing or motion blur, which are critical aspects of video super-resolution. Algorithms tested on these subsets need to effectively interpolate spatial details while maintaining temporal consistency across frames. Due to their common use in the VSR research community, results on these subsets can be easily compared against existing methods, providing a benchmark for evaluating the improvement brought by new algorithms.

Specifically, we used $p = 3$ frames, with the second in order being the zero-motion image, and we added Gaussian noise with a standard deviation of 0.02. The up-sampling factors in the horizontal and vertical directions were $L_1 = L_2 = 4$. For the denoising operator D, the DnCNN neural network [78] was used, as it was pre-trained by SCICO. Finally, we compared our results against other successful video super-resolution techniques in terms of both quantitative metrics, such as PSNR (Peak signal-to-noise ratio), and visual quality.

The results that were compared to ours were acquired by APGM (accelerated proximal gradient method) [90], BM3D (Block-matching and 3D filtering) [113], Total Variation [89], bicubic, SOF-VSR (Super-resolving Optical Flow for Video Super-Resolution) [175] and EDVR (Video Restoration with Enhanced Deformable Convolutional Networks) [176].

Table 4.1 show PSNR results for the two datasets for all the methods tested. It can be seen that average PSNR for our method is 22.86 dB for "Calendar" dataset and 25.74 dB for "City" dataset, while all the other methods have lower values. The highest PSNR values for Frame 17 of "Calendar" (Fig. 4.1) and Frame 14 of "City" (Fig. 4.3). It is obvious that SOF-VSR gave the best result for "City" dataset, however there was no significant difference of the proposed method with SOF-VSR (p=0.207 in Wilcoxon signed-rank test).

Table 4.1: PSNR statistics for the two datasets for all the methods [4].

	Calendar		City	
	Average	St. Dev	Average	St. Dev
PPP V1	**22.86**	0.38	**25.74**	0.28
APGM	20.58	0.29	23.91	0.27
BM3D	21.09	0.38	24.37	0.25
TV	21.66	0.42	25.13	0.20
bicubic	19.36	0.10	22.61	0.29
SOF-VSR	21.69	0.15	25.61	0.59

Apart from the numerical results, the visual proofs are also in favor of our method, since the super-resolved pictures are clearer than the pictures produced with the other methods. Examples of the results can be seen in Fig. 4.2 and Fig. 4.4, which are the results for the original Fig. 4.1 and Fig. 4.3. It should be noted that there is no image result for EDVR, since results were taken from [177].

Figure 4.1: Original “Calendar” Image [4].

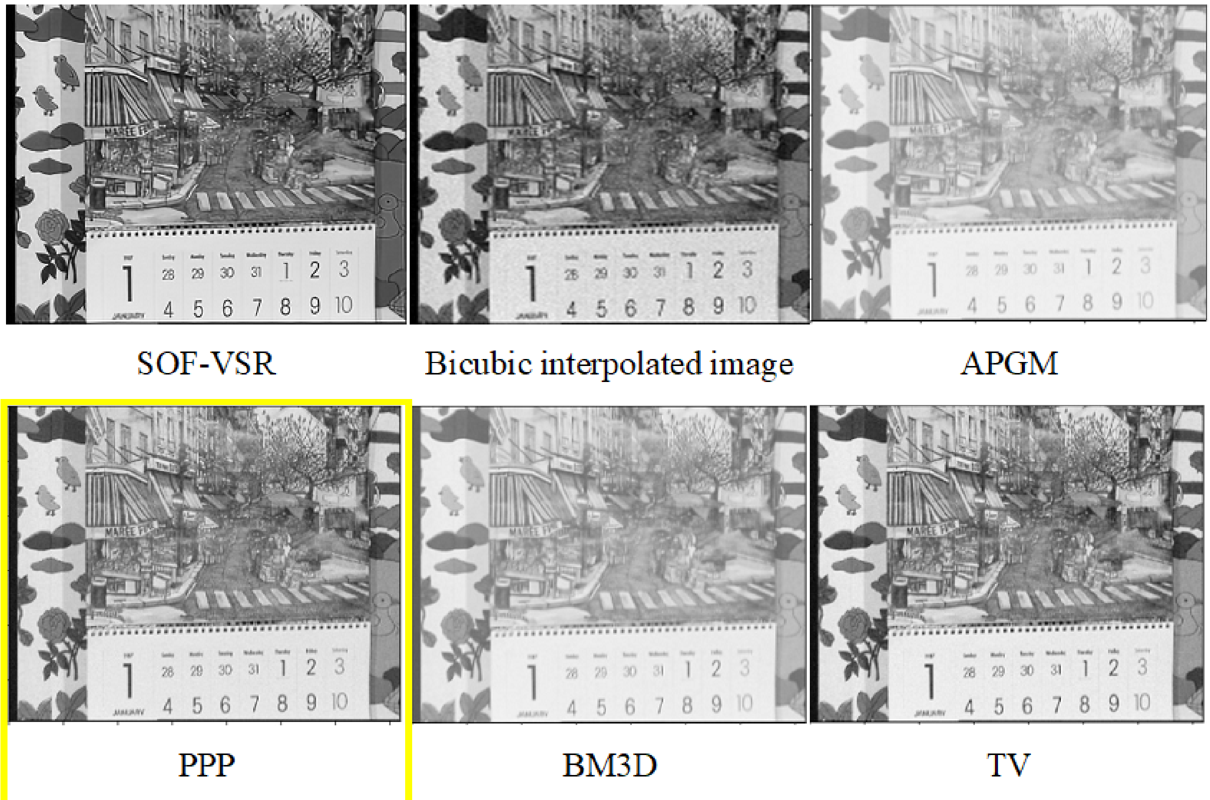


Figure 4.2: Result of “Calendar” Image [4].

Figure 4.3: Original “City” Image [4].

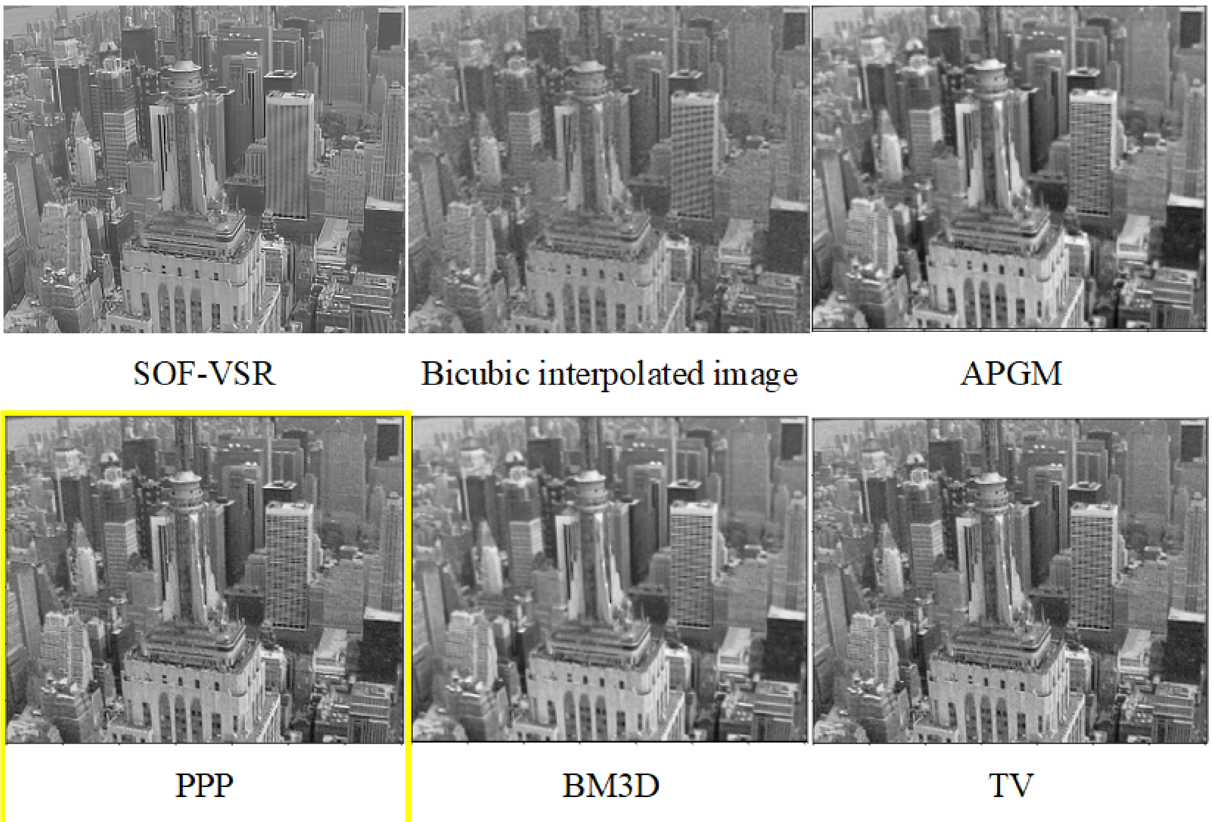


Figure 4.4: Result of “City” Image [4].

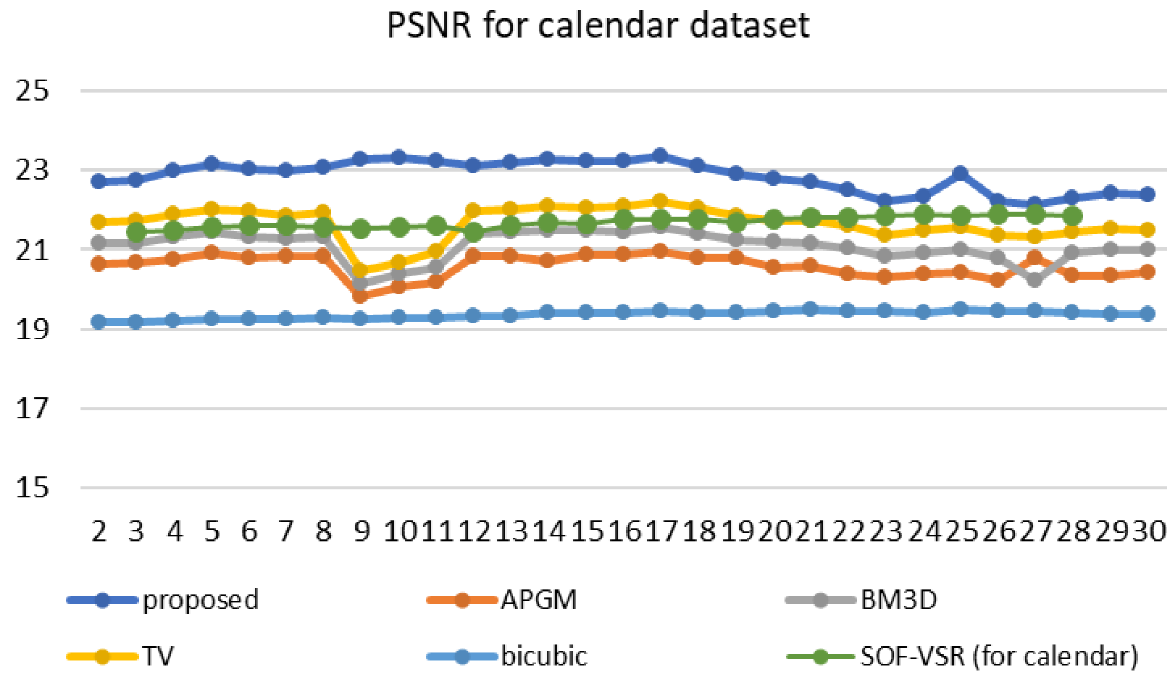


Figure 4.5: PSNR values of the 29 images of "Calendar" dataset for all the methods tested [4].

Fig. 4.5 and Fig. 4.6 show the results in terms of PSNR for the images of "Calendar" and "City" datasets accordingly, for all the methods tested. Frames 9, 10 and 11 from "Calendar" dataset show a much lower PSNR for APGM, BM3D, and TV, because these images have greater difference from the others and these methods are more motion-sensitive than ours. The results demonstrate the superior performance of our approach in terms of reconstruction accuracy and preservation of fine details and textures. It is worth mentioning that our method needs no training, since DnCNN is pre-trained. Finally, the runtime of our method per frame is 12 seconds, ran in Google Colab with T4 GPU.

The Wilcoxon signed-rank test was used to compare the PSNR values of the proposed method with the respective values for Pseudo-inverse, Bicubic, APGM, BM3D and TV methods. The results obtained with those statistical tests are shown in Fig. 4.7 and indicated statistically significant differences between the PPP V1 and the other five methods.

Our method's superior performance can be attributed to several factors: Effective Handling of Noise: The use of the DnCNN neural network for denoising effectively

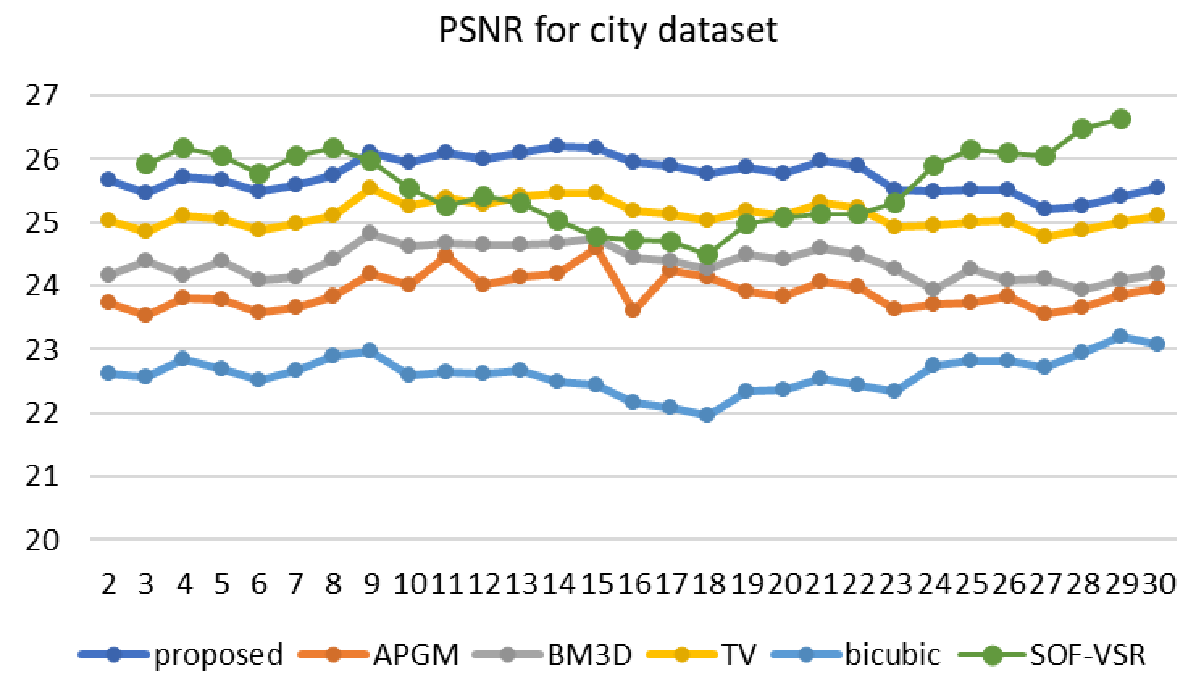


Figure 4.6: PSNR values of the 29 images of "City" dataset for all the methods tested [4].

removes noise while preserving details, which is crucial for achieving high PSNR values. Robust Motion Estimation: The PnP framework with an effective prior model and accurate optical flow estimation helps in better alignment and fusion of frames, leading to improved reconstruction quality. High-Frequency Detail Preservation: The method excels in preserving fine details and textures, as evidenced by the high PSNR values and visually superior images.

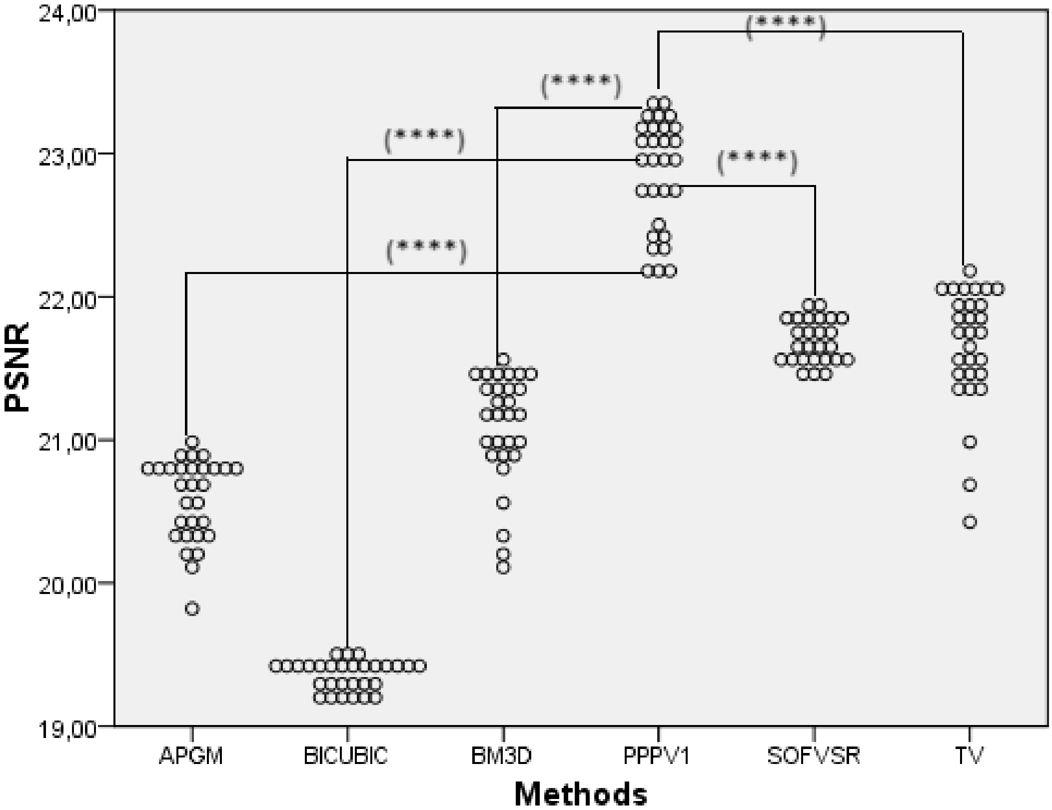


Figure 4.7: Scatter plot representation and the Wilcoxon signed-rank test results of the comparison for each of the five super-resolution methods (Pseudo-inverse, Bicubic, APGM, BM3D and TV) with the PPP V1 method regarding PSNR values. Four stars (****) denote an extremely high level of significance, possibly at the 0.0001 level (p-value < 0.0001), while in this case all results were p=0.000, indicating an ultimately significant correlation.

4.5 MRI Super-Resolution Using Plug-And-Play Priors and Rigid Transformation

The use of Plug-and-Play Priors in MRI super-resolution offers numerous advantages, including the integration of advanced denoising techniques, flexibility, improved handling of complex image features, enhanced convergence and stability, scalability to 3D data, incorporation of motion correction, and quantitative improvements in image quality. These benefits make PPP a powerful and effective framework for enhancing the resolution of MRI images, ultimately contributing to better diagnostic accuracy and patient care. Therefore, extending our, previously analysed, PPP method, we propose its use in MRI super-resolution. The step we changed was the way we evaluated the term M_k from the function (4.3), which now is being evaluated by using rigid registration. Rigid registration, also known as rigid body registration or rigid transformation, is a fundamental technique in medical image processing and computer vision. It is used to align two images by performing translations and rotations while preserving the shape and size of the structures within the images [178].

In a 2D plane, a rigid transformation can be represented using a 3×3 matrix, often referred to as the transformation matrix. For example, a 2D translation can be represented as [179]:

$$T = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

Rotation and reflection matrices can also be formulated similarly. The result of the rigid transformation is represented as an affine transformation matrix. This matrix captures the translation and rotation parameters applied to the original image [179].

To evaluate our method, a widely-used publicly available dataset named the cancer image archive (TCIA) [162] was used. Specifically, we conducted experiments using a dataset of LR brain MRI images and a corresponding HR reference dataset. Our method achieved notable improvements in image quality, as demonstrated by Figure 4.8 and Figure 4.9.

To quantitatively assess the performance of our method, we computed PSNR and compared the results both with other methods, which are APGM (accelerated proximal gradient method) [90], BM3D (Block-matching and 3D filtering) [113], Total Variation [89] and RAISR (Rapid and Accurate Image Super Resolution) [180] and with the pseudo-inverse and the denoised pseudo-inverse versions of the images. The results, presented in Table 4.2, clearly indicate the superiority of our approach in terms of image fidelity.

The Wilcoxon signed-rank test was used to compare the PSNR values of the proposed method with the respective values for Pseudo-inverse, Denoised Pseudoinverse, APGM, BM3D and TV methods. The results obtained with those statistical tests are shown in Fig. 4.10 and indicated statistically significant differences between the PPP V1 and the other five methods, since no per-slice data was available for RAISR and MIRNetv2.

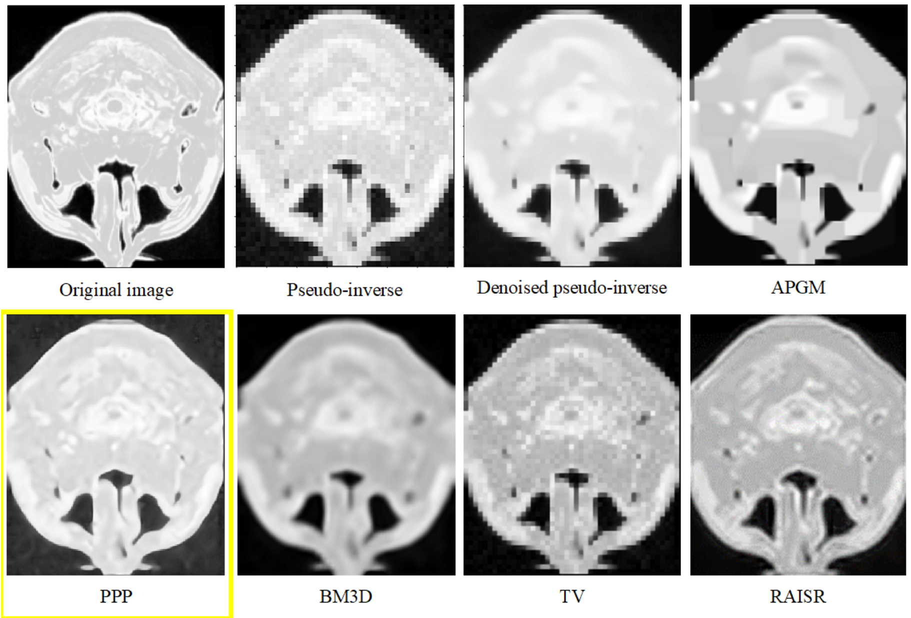


Figure 4.8: Result of image 001 from Dataset 1

Table 4.2: PSNR statistics for the two datasets of all the methods

	Dataset 1		**Dataset 2**	
	Average	**St.Dev**	**Average**	**St.Dev**
PPP	**26.59**	0.49	25.67	0.65
Pseudo-inverse	19.52	0.56	22.81	0.26
Denoised pseudo-inverse	20.36	0.51	23.73	0.28
APGM	19.91	0.34	23.78	0.22
BM3D	20.58	0.82	23.72	0.36
TV	22.48	0.44	23.50	0.29
RAISR	21.99	0.43	**25.77**	0.32

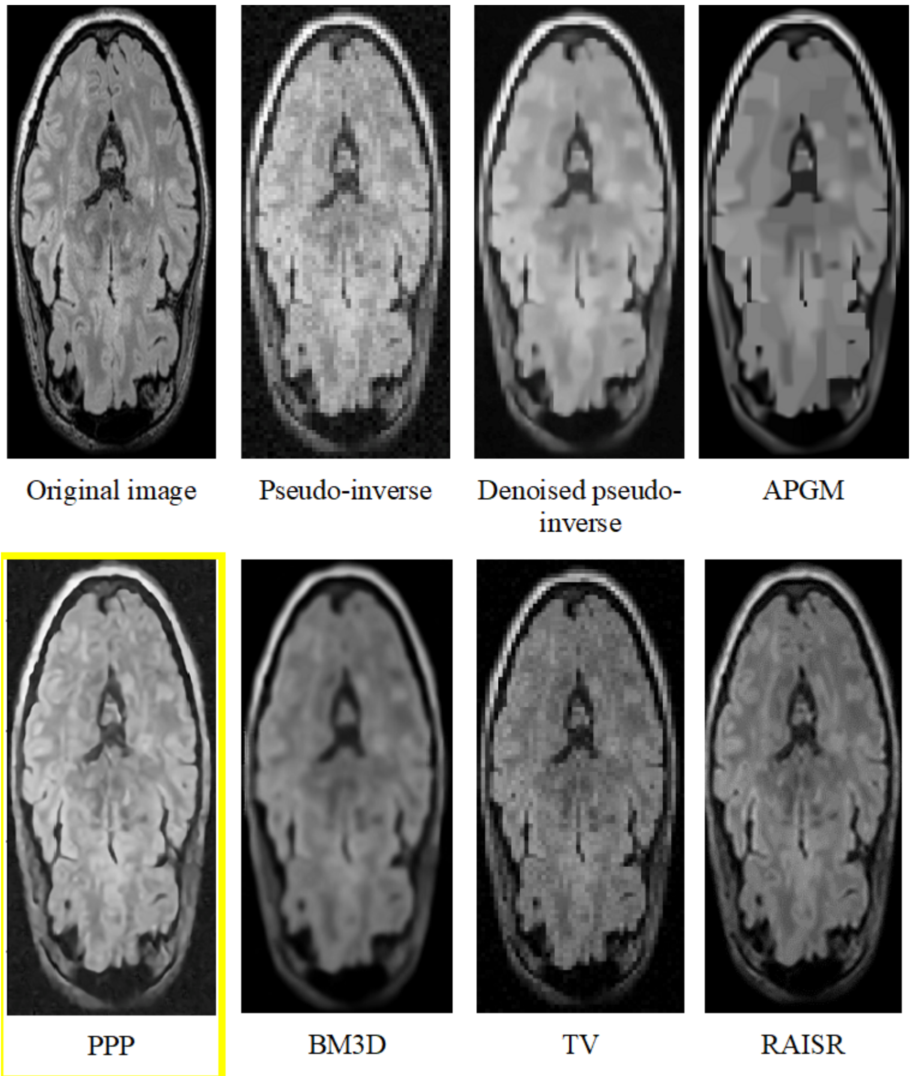


Figure 4.9: Result of image 261 from Dataset 2

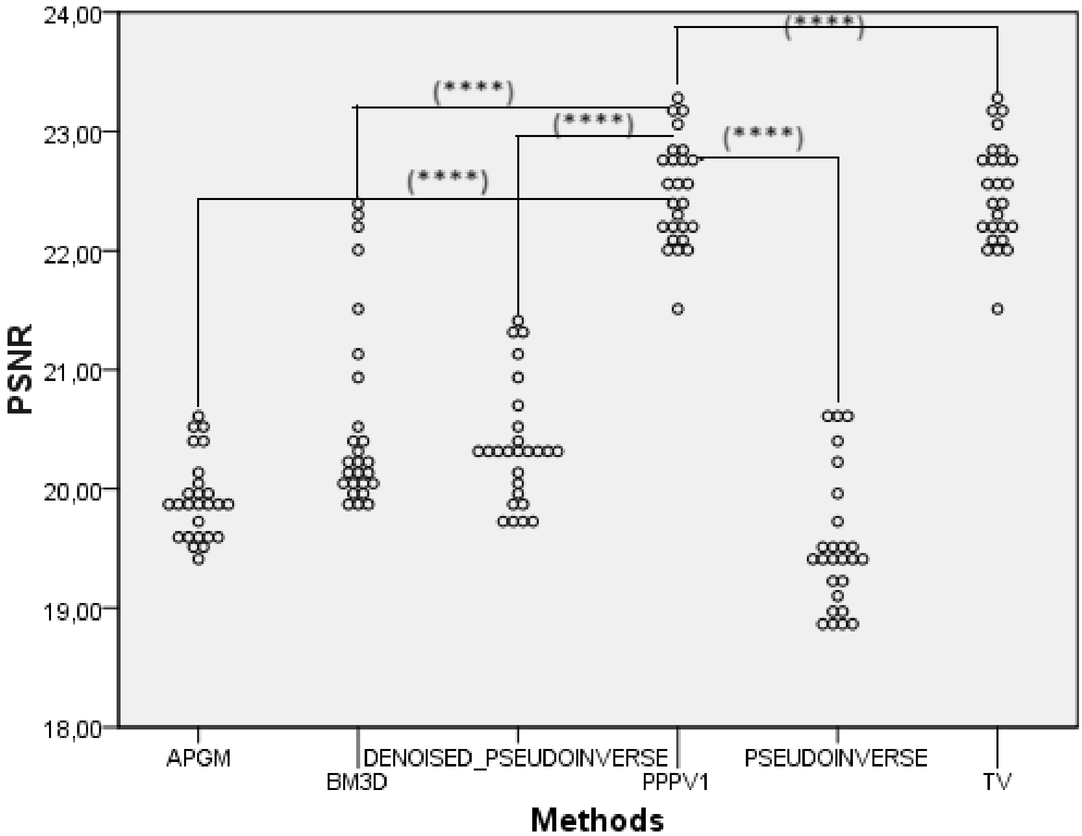


Figure 4.10: Scatter plot representation and the Wilcoxon signed-rank test results of the comparison for each of the five super-resolution methods (Pseudo-inverse, Denoised Pseudoinverse, APGM, BM3D and TV) with the PPP V1 method regarding PSNR values. Four stars (****) denote an extremely high level of significance, possibly at the 0.0001 level (p-value < 0.0001), while in this case all results were p=0.000, indicating an ultimately significant correlation.

4.6 Summary

PnP techniques have established themselves as a standard tool for computational imaging since their introduction in 2013. They have been utilized in a remarkable variety of applications that provide cutting-edge performance. They were arguably the first practical approach to integrating learned models with imaging physics to solve inverse imaging issues when they were first introduced. The ease with which they can be implemented was a major factor in their rapid popularity. Since then, alternative strategies have emerged that, in some cases, result in improved reconstruction performance; however, this is achieved at the expense of a potentially time-consuming and data-dependent application-specific training procedure.

In this chapter, we proposed a PnP method for video super-resolution (resolution enhancement) with motion estimation. The convergence property of the proposed algorithm is analyzed in detail. More importantly, experimental results show the validity of our algorithm and its superiority compared to other state-of-the-art methods. One important advantage of our method is that it needs no training, sonce DnCNN is pre-trained.

Regarding its use for brain MRI super-resolution, the experimental results demonstrate the superiority of our approach over existing techniques, underscoring its potential for clinical applications in neuroimaging.

Chapter 5

An Improved Regularization Method for Video Super-Resolution Using an Effective Prior

5.1 Introduction

As previously mentioned, video super-resolution stands as a significant challenge in the field of computer vision, drawing substantial interest for its wide-ranging applications in areas such as surveillance, entertainment, and healthcare imaging. It primarily focuses on improving the quality of low-resolution video sequences to produce high-quality, high-resolution outputs, addressing various challenges associated with motion and noise.

Commonly, basic interpolation methods like bilinear, bicubic, and spline interpolation are utilized for video super-resolution owing to their computational simplicity. These techniques employ predetermined interpolation kernels to fill in missing pixels on the high-resolution grid, although they can introduce issues such as jagged edges, ringing effects, and a loss of detail. More sophisticated interpolation methods, as cited in certain studies, consider the image's structure to somewhat reduce these issues but may still result in somewhat blurred images, especially with considerable upscaling [181, 182, 183, 184, 185].

Video super-resolution approaches [186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199] combine multiple images of the same scene to create a single high-resolution image, building on the premise that different frames offer unique details about the scene. These methods focus on aligning and merging frames to enhance resolution. Traditional multi-frame super-resolution techniques [186, 187, 191, 198] align frames with sub-pixel accuracy and reconstruct the high-resolution frame using a specific observation model. These are effective with minimal and global motion but struggle with large upscaling factors and pronounced motion. Learning-based approaches, in contrast, derive a direct correlation from low- to high-resolution frames, employing optical flow estimation for frame warping and learning multi-frame fusion from extensive databases [193, 194, 195]. Some approaches, for example, leverage deep learning for complex motion scenarios in multi-frame super-resolution [193]. The latest methods aim to learn both frame registration and fusion using deep neural networks, though the complexity of motion remains a hurdle, sometimes leading to the loss of critical image details [192, 196].

In this chapter a robust regularization method for video super-resolution is proposed, which utilizes an effective prior for denoising and takes into account the motion between consequent frames.

5.2 Our Method

The acquisition model we are assuming is:

$$\mathbf{y} = A\mathbf{x} + \varepsilon, \tag{5.1}$$

where:

- $\mathbf{y}$ is the full set of low resolution (LR) frames, described as $\mathbf{y} = [\,\mathbf{y}_1{}^T, \mathbf{y}_2{}^T, ..., \mathbf{y}_p{}^T]\,^T$, where $\mathbf{y}_k, k = 1, 2, ..., p$ are the p LR images. Each observed LR image is of size $N_1 \times N_2$. Let the kth LR image be denoted in lexicographic notation as $\mathbf{y}_k = [\; y_{k,1}, y_{k,2}, ..., y_{k,M}]\;^T$, for $k = 1, 2, ..., p$ and $M = N_1 N_2$.

- $\mathbf{x}$ is the desired high resolution (HR) image, of size $L_1 N_1 \times L_2 N_2$, written in lexicographical notation as the vector $\mathbf{x} = [\; x_1, x_2, ..., x_N]\;^T$, where $N = L_1 N_1 L_2 N_2$ and L_1 and L_2 represent the up-sampling factors in the horizontal and vertical directions, respectively.

- $\varepsilon = [\; \varepsilon_1, \varepsilon_2, ..., \varepsilon_p]\;^T$, where ε_k is the noise vector for frame k and contains independent zero-mean Gaussian random variables.

- $A = [\; A_1, A_2, ..., A_p]\;^T$ is the degradation matrix which performs the operations of blur, motion and subsampling.

Assuming that each LR image is corrupted by additive noise, we can then represent the observation model as [87]:

$$\mathbf{y}_k = A_k \mathbf{x} + \varepsilon_k \text{ for } 1 \leq k \leq p \tag{5.2}$$

where

$$A_k = S B_k M_k. \tag{5.3}$$

M_k is a warp matrix of size $L_1 N_1 L_2 N_2 \times L_1 N_1 L_2 N_2$, B_k represents a $L_1 N_1 L_2 N_2 \times L_1 N_1 L_2 N_2$ blur matrix, and S is a $N_1 N_2 \times L_1 N_1 L_2 N_2$ subsampling matrix. In our case $B_k = I$, since we assumed no added blur on video frames.

The goal is to find the estimate $\hat{\mathbf{x}}$ of the HR image $\mathbf{x}$ from the p LR images y_k by minimizing the cost function

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x} \in \mathbb{R}^N} f(\mathbf{x}) \text{ with } f(\mathbf{x}) = g(\mathbf{x}) + h(\mathbf{x}), \tag{5.4}$$

where $g(\mathbf{x}) = \sum_{k=1}^{p} \frac{1}{2} \|A_k \mathbf{x} - \mathbf{y}_k\|_2^2$ is the "fidelity to the data" term, and $h(\mathbf{x})$ is the regularization term, which offers some prior knowledge about $\mathbf{x}$. In this study, we adopt the Plug-and-Play Priors approach, in which the ADMM algorithm is modified so that the proximal the proximal operator related to $h(\mathbf{x})$ is replaced by a denoiser that solves the problem of Eq. (5.5. The denoiser used is based on the work by Chantas et al.[115].

The following outlines the algorithm we propose:

1. The initial phase involves computing the term M_k as indicated in Eq. (4.3), achieved through the application of optical flow motion estimation via the Farneback algorithm. This method, developed by Gunnar Farneback, utilizes an image pyramid to progressively reduce resolution across levels and estimates motion vectors for each pixel using a comprehensive set of steps:

    (a) *Preprocessing*: Enhancing the input frames through noise reduction, image denoising, and conversion of color space.

    (b) *Image pyramids*: Constructing a Gaussian pyramid for each frame to create downsampled versions of the original image, aiding in detecting motion across various scales.

    (c) *Optical flow estimation*: Calculating the optical flow at each pyramid level by employing polynomial expansion and spatial filtering to estimate motion vectors based on phase differences.

    (d) *Upsampling and refinement*: Refining the optical flow from the coarsest level by upsampling and integrating higher-resolution data, enhancing flow estimation accuracy for detailed motion tracking.

    This process yields a comprehensive optical flow field, with each pixel's motion vector indicating the direction and magnitude of scene movement across frames.

    We presuppose that one of the low-resolution (LR) images, $\mathbf{y}_{mid}$ (usually the central image), is derived from the high-resolution (HR) image $\mathbf{x}$ through downsampling alone, without motion, thus setting $M_{mid} = I$. Optical flow is then computed between $\mathbf{y}_{mid}$ and the other LR images to determine M_k for the remaining $p - 1$ images.

2. The subsequent phase is centered on employing the PnP-ADMM technique. We execute the PnP-ADMM, adhering to the procedure outlined in Algorithm 5.1 until reaching convergence. The initial HR image guess, $\mathbf{x}^0$, is generated from

$\mathbf{y}_{mid}$ using the pseudo-inverse of $\mathbf{A}_{mid}$ and then denoised via DnCNN. Here, D represents the denoising operator, introduced and discussed in Section 5.3, and g is formulated as $g(\mathbf{x}) = \sum_{k=1}^{p} \frac{1}{2}\|A_k\mathbf{x} - \mathbf{y}_k\|_2^2$.

Algorithm 5.1 PnP-ADMM [6]

1: $\mathbf{u}^0 = \mathbf{0}, \mathbf{x}^0,$ and $\gamma > 0$
2: **for** $k = 1, 2, ..., t$ **do**
3: $\quad \mathbf{z}^k \leftarrow prox_{\gamma g}(\mathbf{x}^{k-1} - \mathbf{u}^{k-1})$
4: $\quad \mathbf{x}^k \leftarrow D(\mathbf{z}^k + \mathbf{u}^{k-1})$
5: $\quad \mathbf{u}^k \leftarrow \mathbf{u}^{k-1} + (\mathbf{z}^k - \mathbf{x}^k)$
6: **end for**
7: **return** $\mathbf{x}^t$

We next explain the modification made to the standard ADMM algorithm to obtain PnP-ADMM. Line 4 or the standard ADMM is $\mathbf{x}^k \leftarrow prox_{\beta h}(\mathbf{z}^k + \mathbf{u}^{k-1})$. In the PnP-ADMM, the proximal operator is replaced by a denoiser that solves the problem

$$\mathbf{z} = \mathbf{x}_0 + \mathbf{w}, \text{ where } \mathbf{x}_0 \sim p, \mathbf{w} \sim N(0; \beta I). \tag{5.5}$$

It can be shown that the Maximum A Posteriori (MAP) estimator $\hat{\mathbf{x}_0}$ of $\mathbf{x}_0$ is the proximal operator:

$$\hat{\mathbf{x}_0} = prox_{\beta h}(\mathbf{z}) = \arg\min_{\mathbf{x} \in \mathbb{R}^N} \{\frac{1}{2}\|\mathbf{x} - \mathbf{z}\|_2^2 \ + \beta h(\mathbf{x})\}, \tag{5.6}$$

for $h(\mathbf{x}) = -\log(p(\mathbf{x}))$.

5.3 The denoising algorithm

In this section, we describe the algorithm we use to implement the denoising step of Eq. (5.6). The algorithm is a simplification of that proposed in ([115]), it is formulated in a probabilistic (Variational Bayes) context and utilizes an effective prior distribution, which we describe in short next.

5.3.1 The prior distribution

The prior distribution we employ for the denoising step was proposed in [115] for the single image Super-Resolution, and it is of the form:

$$p(\mathbf{x}) \propto \prod_{w\in\Omega}\left(\sum_{\delta\in\mathcal{D}}\left(1+\frac{\lambda}{\nu}\epsilon_{w,\delta}(\mathbf{x})\right)^{-\frac{\nu+1}{2}}\right), \tag{5.7}$$

where λ, ν are the real-positive distribution parameters and $\boldsymbol{\epsilon}_{w,\delta}$ is a similarity measure between two patches each of center pixel w and $w+\delta$. The above distribution is produced after integrating out the hidden variables of the prior in [115]. However, this form in never explicitly used (it is not necessary) in the optimization algorithm. We show it here in this form for simplicity of presentation. Indeed, $h(\mathbf{x})$ enables us to interpret the prior in a deterministic context, analogous with the penalty function imposed on the video frames, see equation 5.6.

We introduce a similarity measure between two image patches, denoted as $\mathcal{N}_w$ and $\mathcal{N}_{w'}$, where $\mathbf{x}(w)$ and $\mathbf{x}(w')$ represent the central pixel of the first and second patch, respectively.

The complete set of pixel coordinates is represented by $\Omega = \{1, \ldots, N\}$. Furthermore, we define δ as the integer displacement between the center pixels of the two patches, such that $w' = w+\delta$. For measuring similarity, we employ a weighted Euclidean norm, represented by $\epsilon_{w,\delta}$, to quantify the difference between $\mathcal{N}_w$ and $\mathcal{N}_{w'}$ (or $\mathcal{N}_{w+\delta}$) as follows:

$$\epsilon_{w,\delta} = \sum_{i\in\Omega}\mathbf{v}_\delta^2(i)\mathbf{g}_w(i), \tag{5.8}$$

where $\mathbf{v}_\delta$ is defined by: $\mathbf{v}_\delta = \mathbf{Q}_\delta\mathbf{x}$ and $\mathbf{v}_\delta^2$ indicates the vector obtained by squaring each element of $\mathbf{v}_\delta$. $\mathbf{Q}_\delta$ represents the difference operator, an $N\times N$ matrix, such that the i-th component of $\mathbf{Q}_\delta\mathbf{x}$ equals $\mathbf{x}(i)-\mathbf{x}(i')$ for all $i, i' \in \Omega$ with $i'-i=\delta$. The matrix $\mathbf{G}_w$ is an $N\times N$ diagonal matrix, where its diagonal elements corresponding to the pixels in $\mathcal{N}_w$ are the only non-zero values, specifically, $\mathbf{G}_w(i,i) = 0$ for all i not in $\mathcal{N}_w$. Lastly, we denote by $\mathbf{g}_w$ the $N\times 1$ vector with elements the weights of the weighted norm: the closer to the central pixel of the patches the larger the weight value.

The norm defined by (5.8) retains its value even if the summation (5.8) runs over only the subset $\mathcal{N}_w \subset \Omega$ instead of Ω, since $\mathbf{g}_w(i) = 0$ for $i \notin \mathcal{N}_w$. However, we use the full summation range over Ω for enabling fast computations with the Fast Fourier Transform, as explained next.

The distance between the patch $\mathcal{N}_{w=1}$ and an arbitrary patch $\mathcal{N}_{w'}$, $w' \in \Omega$, is $\delta = w - w' = 1 - w'$. Given that the image patches correspond to $\mathbf{g}_1$ and $\mathbf{g}_{w'}$, it is:

$$\mathbf{g}_{w'}(i) = \mathbf{g}_{w=1}(i-\delta) = \mathbf{g}_1(i+1-w), \;\; \forall i \in \Omega. \tag{5.9}$$

As we can see, each $\mathbf{g}_{w'}$, is a circularly shifted by w' version of $\mathbf{g}_1 \equiv \mathbf{g}$ (denoted simply by $\mathbf{g}$ from now on). The formula (5.8) for calculating $\epsilon_{w,\delta}$, expressed in terms of $\mathbf{g}$, is:

$$\epsilon_{w,\delta} = \sum_{i \in \Omega} \mathbf{v}_\delta^2(i)\mathbf{g}_w(i) = \sum_{i \in \Omega} \mathbf{v}_\delta^2(i)\mathbf{g}(i + 1 - w). \tag{5.10}$$

Clearly, the values of $\epsilon_{w,\delta}$ for all w's, are the result of the *correlation* (denoted by a star in line 4 of Algorithm 2) between $\mathbf{v}_\delta^2$ and $\mathbf{g}$, since the indices of $\mathbf{v}^2$ and $\mathbf{g}$ always differ by the constant $1 - w$. To calculate the correlation required for the super-resolution technique discussed in the following section, we use the Fast Fourier Transform (FFT). This approach decreases the computational complexity of the algorithm from $O(N^2)$, typical for correlation calculations, to $O(N \log N)$, which is the complexity for multiplication in the DFT (Discrete Fourier Transform) domain.

5.3.2 Denoising in PnP-ADMM

Next, we describe the algorithm we employ in the PnP-ADMM context of Algorithm 5.1, and specifically for the denoising step (line 4). The algorithm we employ, as a denoising sub-problem of the general super-resolution algorithm (Algorithm 5.1), is in essence a special case of the VBPS algorithm in [115], where there is no blurring nor decimation. Mathematically speaking, this means that the imaging operator $\mathbf{DH}$ is the $N \times N$ identity matrix $\mathbf{I}$, as shown in line 8 of Algorithm 2.

More specifically, the imaging model assumed for the denoising step is a simplified form of Eq. (2.1) in [115], because it is now $\mathbf{DH} = \mathbf{I}$ (i.e., no blur/decimation, hence it is just the identity matrix). Also, in this form, $\mathbf{z}^k + \mathbf{u}^{k-1}$ has the role of the "noisy image" and $\mathbf{x}^k$ is the uncorrupted one, meant to be estimated by the denoising algorithm.

In parallel with imaging model, we assume the imaging model, i.e., the prior distribution introduced above and given by Eq. (5.5). This is in essence the prior distribution for the uncorrupted image to be estimated via the denoising procedure. This means that the Algorithm 2 is the result of the adoption of both the imaging model mentioned above and the prior (5.5) for $\mathbf{x}$. Lastly, note that the denoising Algorithm 2 selects automatically, in the initialization step, the noise variance β, among other parameters.

Algorithm 5.2 Variational Bayes Patch Similarity Denoising

Input: Noisy image $\mathbf{z}^k + \mathbf{u}^{k-1}$.

Output: Denoised image $\mathbf{x}^k$.

Initialization:

Image initial estimate: Set $\alpha_{\text{new}} = \alpha/2$, where α is the regularization parameter obtained from [200]. Then, set $\mathbf{m}^{(0)} = \mathbf{x}_{\text{Stat}}$, where $\mathbf{x}_{\text{Stat}}$ is the super-resolved image obtained after setting $\alpha = \alpha_{\text{new}}$. *Parameter selection*: Set $t = 0$, and $\beta = N/\|\mathbf{x} - \mathbf{z}\|_2^2$, $\lambda = 10^3\alpha_{\text{new}}$, $\nu = 7$, $rmax = 280$, MAXITER $= 25$ and err$= 10^{-7}$.

while $\|\mathbf{m}^{(t)} - \mathbf{m}^{(t-1)}\|_2^2/N > err$ AND $t <$ MAXITER **do**

  for every δ in $\mathcal{D}$ **do**

    $\mathbf{v}_\delta \leftarrow \mathbf{Q}_\delta \mathbf{m}^{(t)}$ **for** every w in Ω **do**

    Calculate the expectations of the following model's random variables:

$$\langle a_{w,\delta} \rangle_{(t)} = \frac{1+\nu}{\lambda \hat{\epsilon}_{w,\delta} + \nu},$$

$$\langle z_{w,\delta} \rangle_{(t)} = \frac{e^{-\frac{\lambda}{2}\langle a_{w,\delta}\rangle_{(t)}\hat{e}_{w,\delta} - \frac{\nu}{2}\log\langle a_{w,\delta}\rangle_{(t)}}}{\sum_\delta e^{-\frac{\lambda}{2}\langle a_{w,\delta'}\rangle_{(t)}\hat{e}_{w,\delta'} - \frac{\nu}{2}\log\langle a_{w,\delta'}\rangle_{(t)}}},$$

    where $\hat{\epsilon}$ is the ϵ in (5.8), calculated with the image estimation provided in the previous iteration $t-1$,

    calculate $\mathbf{b}_\delta^{(t)}(w) = \langle a_{w,\delta}\rangle_{(t)} \langle z_{w,\delta}\rangle_{(t)}$, for all w and δ,

    set $\square_\delta^{(t)} = \text{diag}\{\mathbf{b}_\delta^{(t)} * \mathbf{g}\}$ (convolution),

    $t \leftarrow t+1$

    Obtain $\mathbf{m}^{(t)}$ by solving the linear system $\left(\beta\mathbf{I} + \lambda \sum_\delta \mathbf{Q}_\delta^T \Lambda_\delta^{(t)} \mathbf{Q}_\delta\right)\mathbf{m}^{(t)} = \beta\mathbf{y}$ with the Conjugated Gradients algorithm.

   end for

  end for

end while

T=t; $\mathbf{x}^k = \mathbf{m}^{(T)}$.

5.4 Results

We implemented our method in SCICO [174], which is an open source library for computational imaging that includes implementations of several algorithms.

Extensive testing was carried out on the Vid4 benchmark dataset to assess the effectiveness of our approach. For our experiments, we selected $p = 3$ frames, positioning the middle frame as the one without motion, and no additive noise was assumed. We applied upscaling factors of $L_1 = L_2 = 4$ in both the horizontal and vertical dimensions. Our method's performance was then benchmarked against other established video super-resolution methods using quantitative indicators like PSNR and subjecive assessment of visual quality.

The experimental results on the *Calendar* and *City* datasets demonstrate the superior performance of the proposed method in comparison to several state-of-the-art video super-resolution techniques, which are SOF-VSR [175], VSR-DUF [201], RBPN [202], DBPN [203], FRVSR [204], EDVR [176] and PPP V1 [4]. The quantitative evaluation is based on PSNR (Peak Signal-to-Noise Ratio) values, which are widely used to assess the quality of super-resolved videos. As for the subjective quality evaluation, it is based on Natural Image Quality Evaluator (NIQE), which provides a score to assess the quality of images without requiring a reference. This no-reference quality metric is valuable because it does not need prior knowledge of specific types of image distortions or perceived degradation. NIQE works independently of any manually degraded data, which potentially makes it more adaptable to unexpected quality issues in images. A lower NIQE score suggests higher perceptual quality of the image.

In the case of the *Calendar* dataset, the proposed method achieves the highest PSNR value of 23.04 dB, outperforming all the competing methods, as it can be seen in Table 5.1. Similarly, on the *Foliage* dataset, our method achieves a notable PSNR value of 25.82 dB, surpassing the performance of the existing methods. The datasets that our algorithm doesn't give the best PSNR are *City* and *Walk*, where only RBPN surpasses our method.

However, considering the perceptual quality of the frames, it is obvious that our method gives the best results, outperforming the other methods on all datasets. This outcome demonstrates the robustness and effectiveness of our method in enhancing the natural quality of super-resolved videos for this specific dataset.

Table 5.1: PSNR results.

	Calendar	City	Foliage	Walk	Average	St. Dev.
SOF-VSR	16.02	21.34	18.89	20.06	19.08	2.27
VSR-DUF	16.12	20.06	18.40	18.73	18.33	1.64
RBPN	22.65	**26.39**	24.90	**29.37**	**25.83**	2.82
DBPN	20.93	23.95	21.72	25.73	25.37	2.18
FRVSR	21.55	25.4	24.11	26.21	24.32	2.04
EDVR	21.70	25.51	24.93	24.01	24.39	1.64
PPP v1	19.47	24.27	20.43	24.45	22.16	2.58
Ours	**23.04**	25.64	**25.82**	27.92	25.61	2.00

Table 5.2: NIQE results.

	Calendar	City	Foliage	Walk	Average	St. Dev.
SOF-VSR	5.56	6.52	8.23	5.95	6.57	1.18
VSR-DUF	4.50	5.72	6.53	5.06	5.45	0.87
RBPN	4.36	5.17	7.14	5.18	5.46	1.18
DBPN	4.87	5.66	7.69	5.67	5.97	1.20
FRVSR	5.30	5.80	7.12	5.22	5.86	0.88
PPP v1	6.83	6.70	7.24	6.74	6.88	0.25
Ours	**4.34**	**4.50**	**4.63**	**3.74**	**4.30**	0.39

Beyond the numerical outcomes, visual evidence also supports the superiority of our method, as the images enhanced through our super-resolution process appear sharper and more defined compared to those generated by competing techniques. Example of the results can be seen in Fig. 5.1, which show example image of the *Calendar* dataset.

In the *City* dataset, just like in the previous one, the images support the numerical results, since the super-resolved pictures are clearer than the pictures produced with the other methods and the bicubic interpolated images. Example of the results can be seen in Fig. 5.2.

It should be noted that there is no image and NIQE result for EDVR, since we were unable to run the code provided so we included only the PSNR values referred to [177].

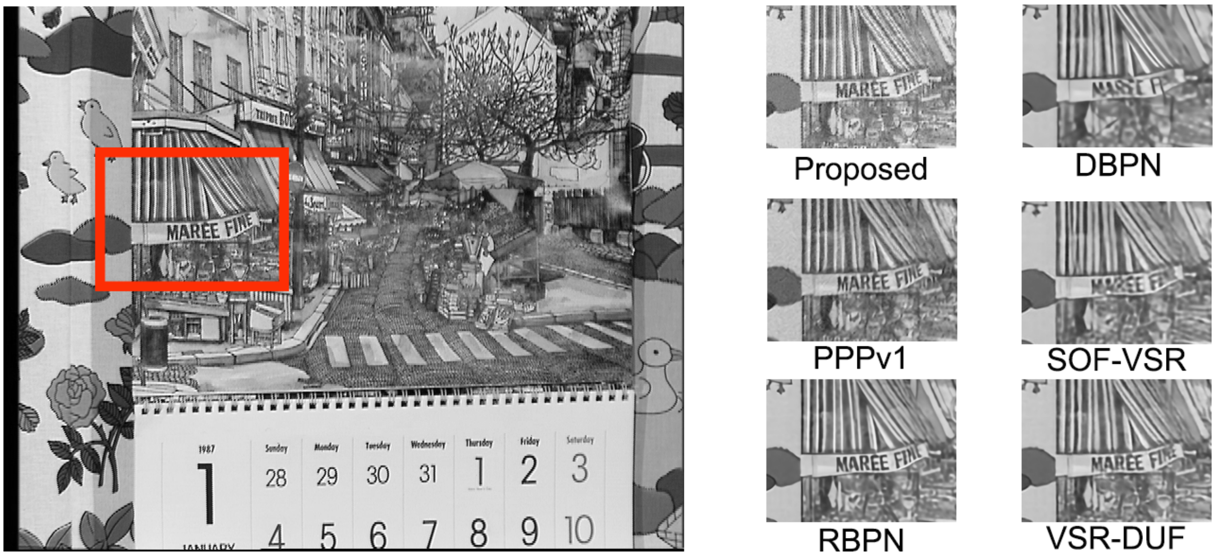


Figure 5.1: Original Image and Super-resolved Image 20, from Dataset Calendar, with all methods

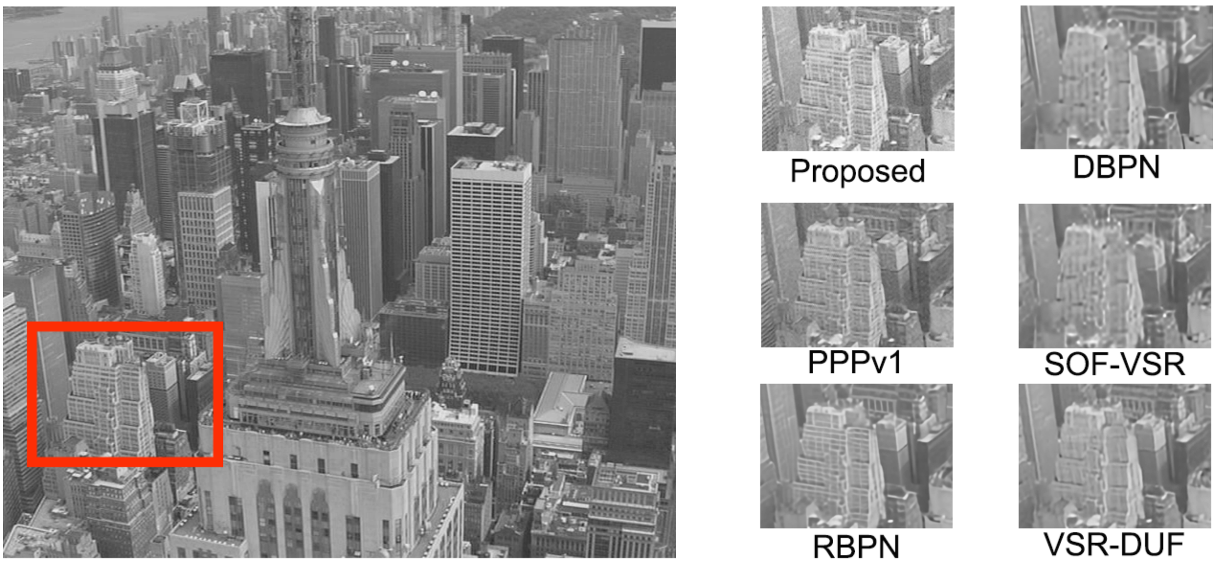


Figure 5.2: Original Image and Super-resolved Image 20, from dataset City, with all methods

These results indicate that the proposed method consistently outperforms the state-of-the-art methods in terms of PSNR for *Calendar*, *City*, *Foliage* and *Walk* datasets. The substantial performance gains emphasize the potential of our approach for high-quality video super-resolution, making it a compelling choice for practical applications in video enhancement and upscaling.

5.5 Discussion

The results of our experiments clearly demonstrate the effectiveness of the proposed method in the context of video super-resolution when compared to several established methods. We will now delve into a discussion of these results and their implications.

5.5.1 Interpreting the PSNR Gains

The substantial PSNR gains achieved by our method on *Calendar*, *City*, *Foliage* and *Walk* datasets underscore its ability to produce higher-quality super-resolved videos. The substantial margin by which our method outperforms existing techniques, such as SOF-VSR, VSR-DUF, RBPN, DBPN, and EVSR, showcases its robustness across different scenarios. The increase in PSNR values and the decrease in NIQE values translates to sharper, more faithful reconstructions of low-resolution videos, making our method highly appealing for various video enhancement applications.

5.5.2 Applicability Across Diverse Datasets

Another noteworthy observation is the consistent performance of our method across the *Calendar*, *City*, *Foliage* and *Walk* datasets. This indicates that our approach is not limited to specific video content types and can be effectively employed in a wide range of real-world scenarios. The ability to maintain high PSNR values across different datasets demonstrates the versatility and adaptability of our method.

5.5.3 Practical Implications

From a practical perspective, the remarkable PSNR improvements hold significant implications for video quality enhancement. Whether it's enhancing low-resolution surveillance footage in urban environments (*City* dataset) or improving the clarity of complex, high-motion scenes (*Calendar* dataset), our method showcases its potential to make a substantial difference in various real-world applications.

5.6 Using the Improved Regularization Method for MRI Super-Resolution

Following the steps we took in Chapter four, we, once again, modified our enhanced PPP method suggested in this Chapter, to use it in MRI super-resolution. As referred in sub-section 4.5, the step we changed was the way we evaluated the term M_k from the function (4.3), which now is being evaluated by using rigid registration.

To evaluate our method, the widely-used publicly available dataset named the cancer image archive (TCIA) [162] was, once again, used,in order to compare our results to the previously proposed method. Specifically, we conducted experiments using a dataset of LR brain MRI images and a corresponding HR reference dataset. Our method with the effective prior achieved notable improvements in image quality, as demonstrated by Figure 5.3 and Figure 5.4.

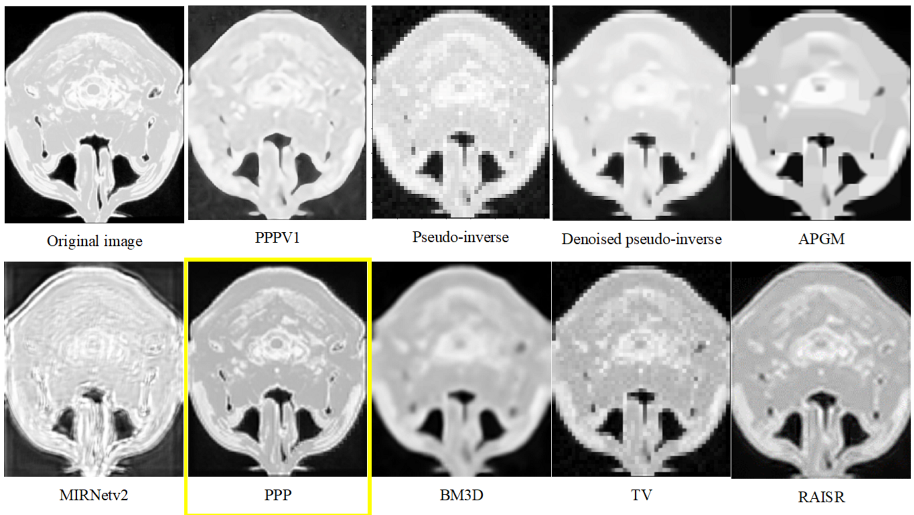


Figure 5.3: Result of image 001 from Dataset 1

To objectively evaluate the effectiveness of our improved technique, we calculated the PSNR and conducted comparisons with both alternative approaches and enhanced versions of our own method. Specifically, we compared against PPPV1, APGM (accelerated proximal gradient method) [90], BM3D (Block-matching and 3D filtering)

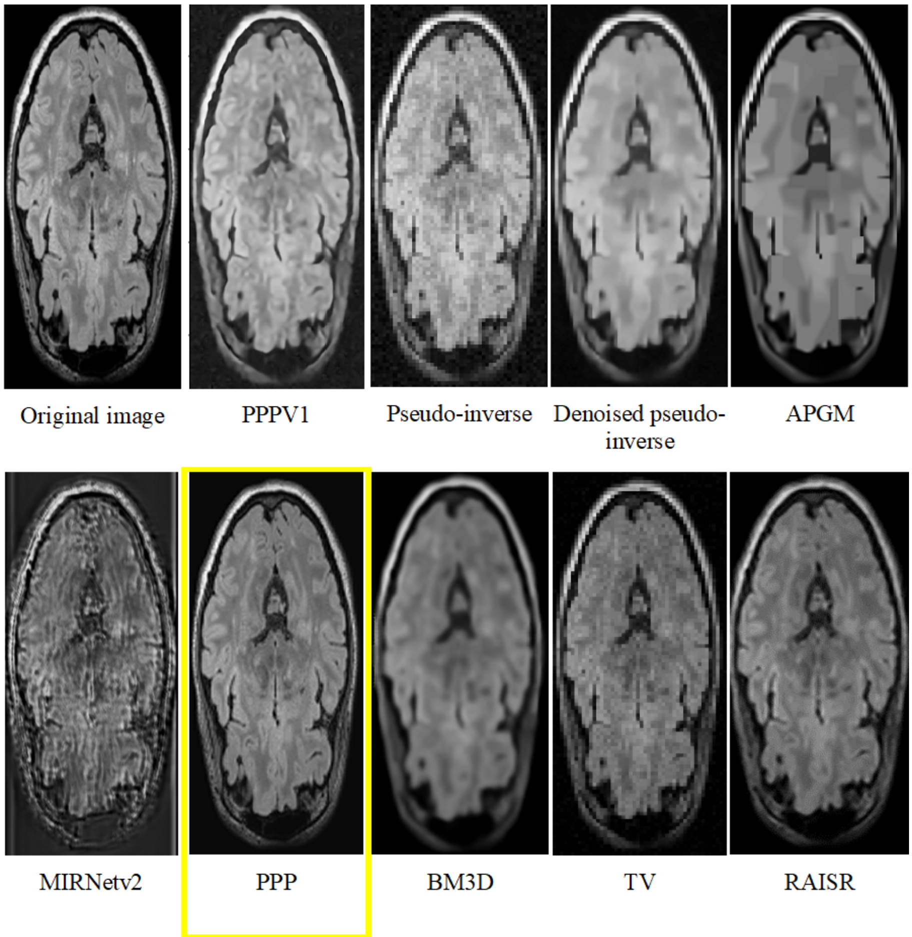


Figure 5.4: Result of image 261 from Dataset 2

[113], Total Variation [89], and RAISR (Rapid and Accurate Image Super Resolution) [180], as well as with the pseudo-inverse and the denoised pseudo-inverse images. The outcomes, detailed in Table 5.3, unequivocally demonstrate that our method surpasses others in delivering higher image quality.

The Wilcoxon signed-rank test was used to compare the PSNR values of the proposed method with the respective values for PPP V1, Pseudo-inverse, Denoised Pseudoinverse, APGM, BM3D and TV methods. The results obtained with those statistical tests are shown in Fig. 5.5 and indicated statistically significant differences between the PPP and the other six methods, since no per-slice data was available for RAISR and MIRNetv2.

Considering the perceptual quality of the frames, it is obvious from Table 5.4 that our method gives the best results, outperforming the other methods on all datasets. This outcome demonstrates the robustness and effectiveness of our method in en-

Table 5.3: PSNR statistics for the two datasets of all the methods

	Dataset 1		Dataset 2	
	Average	**St.Dev**	**Average**	**St.Dev**
PPPV1	22.49	0.44	25.26	0.25
PPP	**26.59**	0.49	25.67	0.65
Pseudo-inverse	19.52	0.56	22.81	0.26
Denoised pseudo-inverse	20.36	0.51	23.73	0.28
APGM	19.91	0.34	23.78	0.22
BM3D	20.58	0.82	23.72	0.36
TV	22.48	0.44	23.50	0.29
RAISR	21.99	0.43	**25.77**	0.32
MIRNetv2	14.05	0.27	14.26	0.18

hancing the natural quality of super-resolved videos for this specific dataset.

Table 5.4: NIQE statistics for the two datasets of all the methods

	Dataset 1		Dataset 2	
	Average	**St.Dev**	**Average**	**St.Dev**
PPPV1	6.14	0.15	6.66	0.17
PPP	**5.82**	0.15	**6.39**	0.16
Pseudo-inverse	14.13	0.36	14.08	0.35
Denoised pseudo-inverse	14.13	0.36	14.08	0.35
APGM	13.86	0.35	13.03	0.33
BM3D	10.66	0.27	11.92	0.30
TV	12.22	0.31	12.82	0.32
RAISR	5.87	0.15	9.61	0.24
MIRNetv2	7.18	0.18	7.95	0.20

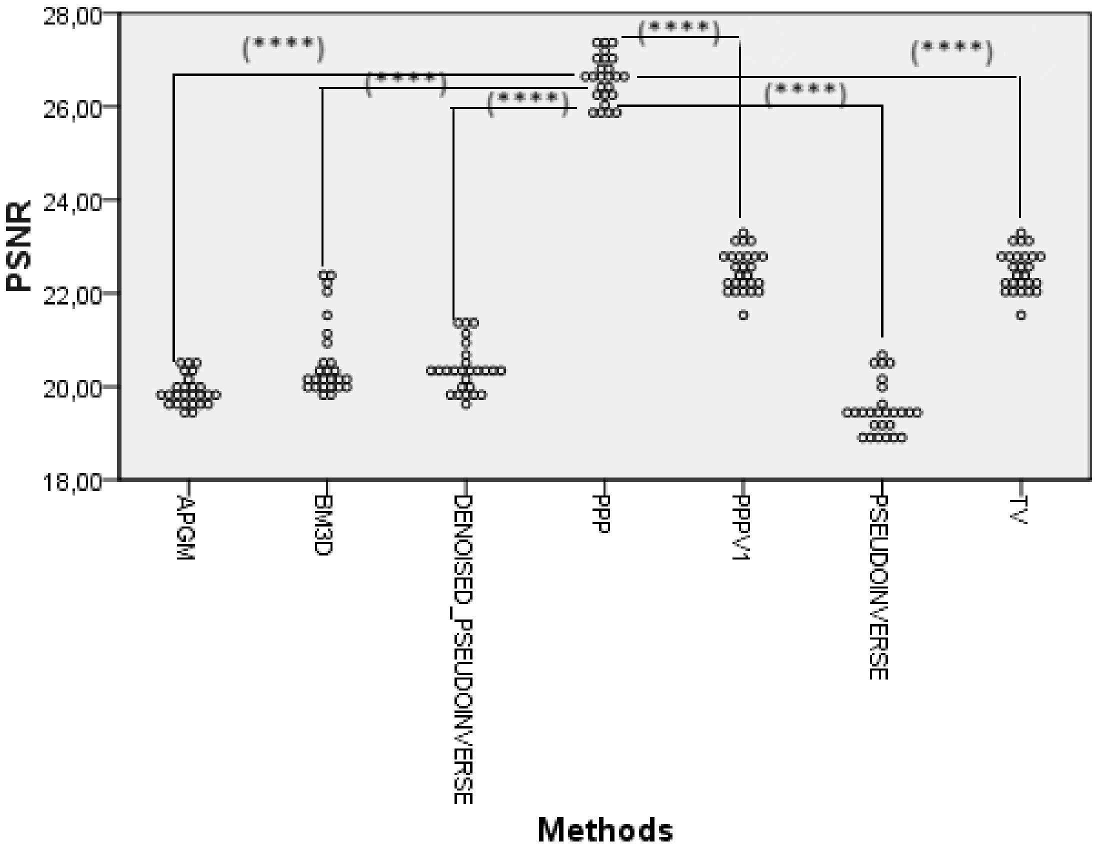


Figure 5.5: Scatter plot representation and the Wilcoxon signed-rank test results of the comparison for each of the six super-resolution methods (PPP V1, Pseudo-inverse, Denoised Pseudoinverse, APGM, BM3D and TV) with the PPP method regarding PSNR values. Four stars (****) are less commonly used than one, two, or three asterisks in standard practice. If used, they might denote an extremely high level of significance, possibly at the 0.0001 level (p-value < 0.0001), while in this case all results were p=0.000, indicating an ultimately significant correlation.

5.7 Summary

In summary, the results presented in this study highlight the superior performance of the proposed method in the field of video super-resolution. This method consistently outperforms state-of-the-art techniques, as demonstrated by the substantial PSNR gains observed on *Calendar*, *City*, *Foliage*, and *Walk* datasets. The following key takeaways can be drawn:

- Our method achieves remarkable PSNR improvements, leading to sharper and higher-quality super-resolved videos.
- The versatility of our approach is evident, as it performs consistently well on different datasets, representing a wide range of real-world scenarios.
- The practical implications of our results suggest that our method holds great promise for applications where video quality enhancement is paramount.
- Computational efficiency is another significant advantage of our method. Unlike Deep Neural Network-based methods, our approach does not rely on neural networks and requires no training, making it faster and less resource-intensive.

Regarding its use for brain MRI super-resolution, the experimental results demonstrate the superiority of our approach over existing techniques, underscoring its potential for clinical applications in neuroimaging.

These findings make a strong case for the adoption of our method in video enhancement and upscaling tasks. We believe that the approach we suggest has the potential to contribute significantly to the field of video super-resolution and benefit a wide range of applications. It should be emphasized that the proposed method does not require any training, in contrast to the other methods we used in comparison.

Future work may involve further optimizations, including real-time implementation and the exploration of additional performance metrics to provide a more comprehensive assessment of our method's capabilities.

Chapter 6

Conclusion and Future Directions

This PhD thesis has embarked on an in-depth exploration of novel methodologies in the compression and super-resolution of medical images and videos, marking significant strides in the advancement of medical imaging technology. With a keen focus on pioneering algorithms and leveraging the potential of deep learning, the research encapsulates a series of innovative contributions to the field, establishing a solid foundation for future explorations and enhancements.

6.1 Summary of Key Findings

Throughout this research, the thesis has unfolded novel techniques for medical image compression and super-resolution, with Chapters 4, 5, and 6 introducing ground-breaking methods that stand at the forefront of this domain. The thesis showcased an improved medical image compression method predicated on Wavelet Difference Reduction (WDR) and the efficacious application of Plug-and-Play Priors (PnP) for advancing video and MRI super-resolution.

Chapter 4 delineated an innovative compression method, highlighting the utilization of the mean co-located pixel difference (MCPD) to determine the optimal encoding frames. This methodology not only achieved exemplary compression ratios but also meticulously preserved the critical diagnostic content within the medical images, hence fostering an efficient storage and transmission process.

In Chapter 5, the thesis navigated through the intricacies of video super-resolution, employing Plug-and-Play Priors to substantially uplift the resolution of low-quality videos. This chapter underlined the method's adeptness in enhancing image details without the customary computational burdens, thereby offering a robust solution for the improvement of medical imaging quality.

Chapter 6 further extended the discourse into advanced regularization methods for video super-resolution, revealing a novel technique that integrates an innovative denoiser within the ADMM algorithm. This approach underscored the potential to significantly refine the resolution and clarity of medical videos, thereby providing a richer visual context for clinical analysis and diagnostics.

6.2 Contribution to Medical Imaging Technology

The methodologies developed in this thesis have contributed to medical imaging technology by:

- Enhancing Data Efficiency: The new compression method introduced offers a sophisticated approach to managing the vast data volumes associated with medical imaging, thereby enabling more effective storage and transmission solutions.
- Elevating Image Quality: The super-resolution techniques presented have substantially elevated the quality and resolution of medical images and videos, facilitating more precise and detailed clinical evaluations.
- Innovating with Deep Learning: By incorporating deep learning into super-resolution processes, the research has opened new avenues for automating and improving image enhancement tasks, setting a new benchmark for future endeavors in the field.

6.3 Future Directions

The findings and methodologies introduced in this thesis pave the way for several promising avenues of future research, including:

- Exploration of Advanced Deep Learning Models: Investigating and developing cutting-edge deep learning models to further enhance the efficacy of image super-resolution.

- Extension to Various Medical Imaging Modalities: Adapting the developed methods to encompass a broader spectrum of imaging modalities, thereby amplifying their utility and application in medical diagnostics.

- Emphasis on Real-time Processing: Focusing on algorithms and technological solutions that facilitate real-time image processing could revolutionize clinical practices and telemedicine.

- Integration of Security Measures: As advancements continue, integrating robust encryption and privacy measures will be crucial in safeguarding the integrity and confidentiality of medical images.

6.4 Closing Remarks

The contributions of this thesis to medical imaging technology are not just evolutionary but revolutionary, offering a glimpse into the future of healthcare diagnostics. By addressing the pivotal challenges and harnessing the opportunities within medical imaging, this research not only propels the field forward but also sets the stage for transformative breakthroughs that promise to enhance patient care and outcomes significantly. The integration of artificial intelligence and computational methods heralds a new era in medical imaging, one that is bound to unfold new dimensions of diagnostic accuracy and efficiency.

Bibliography

[1] M. C. Zerva, M. Vrigkas, L. P. Kondi, and C. Nikou, “Improving 3d medical image compression efficiency using spatiotemporal coherence,” *Electronic Imaging*, vol. 2020, no. 10, pp. 63–1, 2020.

[2] M. C. H. Zerva, V. Christou, N. Giannakeas, A. T. Tzallas, and L. P. Kondi, “An improved medical image compression method based on wavelet difference reduction,” *IEEE Access*, vol. 11, pp. 18 026–18 037, 2023.

[3] F. Grillo and J. Balbás, “Lossless compressing using dwt tecnique,” *MATLAB Central File Exchange*, 2020.

[4] M. C. Zerva and L. P. Kondi, “Video super-resolution using plug-and-play priors,” *IEEE Access*, vol. 12, pp. 11 963–11 971, 2024.

[5] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, “Plug-and-play priors for model-based reconstruction,” in *Proc. IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2013, pp. 945–948.

[6] U. Kamilov, C. Bouman, G. Buzzard, and B. Wohlberg, “Plug-and-play methods for integrating physical and learned models in computational imaging: Theory, algorithms, and applications,” *IEEE Signal Processing Magazine*, vol. 40, no. 1, pp. 85–97, Jan. 2023.

[7] A. J. Hussain, A. Al-Fayadh, and N. Radi, “Image compression techniques: A survey in lossless and lossy algorithms,” *Neurocomputing*, vol. 300, pp. 44–69, 2018.

[8] W.-C. Siu and K.-W. Hung, “Review of image interpolation and super-resolution,” in *Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC), 2012 Asia-Pacific*. IEEE, 2012, pp. 1–10.

[9] S. Borman and R. Stevenson, "Spatial resolution enhancement of low-resolution image sequences - a comprehensive review with directions for future research," Lab. for Image and Sugnal Analysis, Univ. Notre Dame, Tech. Rep., 1998.

[10] R. Duszak and L. Harvey, "Medical imaging: Is the growth boom over? the neiman report no 1," *Neiman Health Policy Institute, Reston, VA. Cite*, 2012.

[11] L. F. Lucas, N. M. Rodrigues, L. A. da Silva Cruz, and S. M. de Faria, "Lossless compression of medical images using 3-d predictors," *IEEE transactions on medical imaging*, vol. 36, no. 11, pp. 2250–2260, 2017.

[12] A. AbuBaker, M. Eshtay, and M. AkhoZahia, "Comparison study of different lossy compression techniques applied on digital mammogram images," *Volume*, vol. 7, pp. 149–155, 2016.

[13] G. K. Wallace, "The JPEG still picture compression standard," *IEEE transactions on consumer electronics*, vol. 38, no. 1, pp. xviii–xxxiv, 1992.

[14] M. Rabbani and R. Joshi, "An overview of the JPEG 2000 still image compression standard," *Signal processing: Image communication*, vol. 17, no. 1, pp. 3–48, 2002.

[15] W. B. Pennebaker and J. L. Mitchell, *JPEG: Still Image Data Compression Standard*. Science Business Media: Springer, 1992.

[16] R. Leung and D. Taubman, "Transform and embedded coding techniques for maximum efficiency and random accessibility in 3-d scalable compression," *IEEE Trans. Image Process*, vol. 14, no. 10, pp. 1632–1646, October 2005.

[17] N. Ahmed, "How i came up with the discrete cosine transform," *Digital Signal Processing*, vol. 1, no. 1, pp. 4–5, 1991.

[18] V. Britanak, P. C. Yip, and K. R. Rao, *Discrete cosine and sine transforms: general properties, fast algorithms and integer approximations*. Elsevier, 2010.

[19] P. Schelkens, A. Munteanu, J. Barbarien, M. Galca, X. Giro-Nieto, and J. Cornelis, "Wavelet coding of volumetric medical datasets," in *IEEE Trans.* vol. 22, no. 3: Med. Imaging, 2003, pp. 441–458.

[20] C. Narmatha, P. Manimegalai, and S. Manimurugan, "A lossless compression scheme for grayscale medical images using a p2-bit short technique," *Journal of Medical Imaging and Health Informatics*, vol. 7, no. 6, pp. 1196–1204, October 2017.

[21] H. Amri, M. G. A. Khalfallah, N. NebhaniJean, J. C. Lapayre, and M. S. Bouhlel, "Medical image compression approach based on image resizing, digital watermarking and lossless compression," *Journal of Signal Processing Systems*, vol. 87, p. 203, 2017.

[22] T. Bruylants, A. Munteanu, and P. Schelkens, "Wavelet based volumetric medical image compression," *Signal processing: Image communication*, vol. 31, pp. 112–133, 2015.

[23] D. Ravichandran, M. G. Ahamad, and M. A. Dhivakar, "Performance analysis of three-dimensional medical image compression based on discrete wavelet transform," in *2016 22nd International Conference on Virtual System & Multimedia (VSMM)*. IEEE, 2016, pp. 1–8.

[24] R. K. Senapati, P. Prasad, G. Swain, and T. Shankar, "Volumetric medical image compression using 3d listless embedded block partitioning," *SpringerPlus*, vol. 5, no. 1, pp. 1–16, 2016.

[25] X. Tang and W. A. Pearlman, "Three-dimensional wavelet-based compression of hyperspectral images," in *Hyperspectral data compression*. Springer, 2006, pp. 273–308.

[26] Z. Chen, S. Gu, G. Lu, and D. Xu, "Exploiting intra-slice and inter-slice redundancy for learning-based lossless volumetric image compression," *IEEE Transactions on Image Processing*, 2022.

[27] O. H. Nagoor, J. Whittle, J. Deng, B. Mora, and M. W. Jones, "Lossless compression for volumetric medical images using deep neural network with local sampling," in *2020 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2020, pp. 2815–2819.

[28] D. A. Huffman, "A method for the construction of minimum-redundancy codes," *Proceedings of the IRE*, vol. 40, no. 9, pp. 1098–1101, 1952.

[29] G. G. Langdon, "An introduction to arithmetic coding," *IBM Journal of Research and Development*, vol. 28, no. 2, pp. 135–149, 1984.

[30] J. McCoy, N. Magotra, and S. Stearns, "Lossless predictive coding," in *Proceedings of 1994 37th Midwest Symposium on Circuits and Systems*, vol. 2. IEEE, 1994, pp. 927–930.

[31] D. Dubé, "Lossless compression of grayscale and colour images using multidimensional cse," in *2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA)*. IEEE, 2019, pp. 222–227.

[32] V. Makarichev, V. Lukin, and I. Brysina, "Lossless discrete atomic compression of full color digital images," in *2021 IEEE 16th International Conference on the Experience of Designing and Application of CAD Systems (CADSM)*. IEEE, 2021, pp. 43–46.

[33] J. Lee, J. Yun, J. Lee, I. Hwang, D. Hong, Y. Kim, C. G. Kim, and W.-C. Park, "An effective algorithm and architecture for the high-throughput lossless compression of high-resolution images," *IEEE Access*, vol. 7, pp. 138 803–138 815, 2019.

[34] A. Descampe, T. Richter, T. Ebrahimi, S. Foessel, J. Keinert, T. Bruylants, P. Pellegrin, C. Buysschaert, and G. Rouvroy, "JPEG XS—Anew standard for visually lossless low-latency lightweight image coding," *Proceedings of the IEEE*, vol. 109, no. 9, pp. 1559–1577, 2021.

[35] Q. Min, X. Wang, B. Huang, and Z. Zhou, "Lossless medical image compression based on anatomical information and deep neural networks," *Biomedical Signal Processing and Control*, vol. 74, p. 103499, 2022.

[36] F. Yang, J. Mou, K. Sun, and R. Chu, "Lossless image compression-encryption algorithm based on bp neural network and chaotic system," *Multimedia Tools and Applications*, vol. 79, no. 27, pp. 19 963–19 992, 2020.

[37] H. Rhee, Y. I. Jang, S. Kim, and N. I. Cho, "Lossless image compression by joint prediction of pixel and context using duplex neural networks," *IEEE Access*, vol. 9, pp. 86 632–86 645, 2021.

[38] C. Zhu, H. Zhang, and Y. Tang, “Lossless image compression algorithm based on long short-term memory neural network,” in *2020 5th International Conference on Computational Intelligence and Applications (ICCIA)*, 2020, pp. 82–88.

[39] A. Al-Fayadh, A. J. Hussain, P. Lisboa, and D. Al-Jumeily, “Hybrid classified vector quantisation and its application to image compression,” in *EUROCON 2007-The International Conference on" Computer as a Tool"*. IEEE, 2007, pp. 157–162.

[40] R. Dony *et al.*, “Karhunen-loeve transform,” *The transform and data compression handbook*, vol. 1, no. 1-34, p. 29, 2001.

[41] K. Karhunen, “Under lineare methoden in der wahr scheinlichkeitsrechnung,” *Annales Academiae Scientiarun Fennicae Series A1: Mathematia Physica*, vol. 47, 1947.

[42] M. Loeve, “Functions aleatoires du second ordre,” *Processus stochastique et mouvement Brownien*, pp. 366–420, 1948.

[43] G. Anderson and T. Huang, “Piecewise fourier transformation for picture bandwidth compression,” *IEEE Transactions on Communication Technology*, vol. 19, no. 2, pp. 133–140, 1971.

[44] K. G. Beauchamp, *Applications of Walsh and related functions, with an introduction to sequency theory*. Academic press, 1984, vol. 2.

[45] T. Beer, “Walsh transforms,” *American Journal of Physics*, vol. 49, no. 5, pp. 466–472, 1981.

[46] P. S. Addison, “Wavelet transforms and the ecg: a review,” *Physiological measurement*, vol. 26, no. 5, p. R155, 2005.

[47] S. Xu, J. Zhang, L. Bo, H. Li, H. Zhang, Z. Zhong, and D. Yuan, “Singular vector sparse reconstruction for image compression,” *Computers & Electrical Engineering*, vol. 91, p. 107069, 2021.

[48] L. Guo, D. Zhou, J. Zhou, S. Kimura, and S. Goto, “Lossy compression for embedded computer vision systems,” *IEEE Access*, vol. 6, pp. 39 385–39 397, 2018.

[49] A. Sadchenko, O. Kushnirenko, and O. Plachinda, "Fast lossy compression algorithm for medical images," in *2016 International Conference on Electronics and Information Technology (EIT)*, 2016, pp. 1–4.

[50] Y. Dua, R. S. Singh, K. Parwani, S. Lunagariya, and V. Kumar, "Convolution neural network based lossy compression of hyperspectral images," *Signal Processing: Image Communication*, vol. 95, p. 116255, 2021.

[51] L. Zhao, H. Bai, A. Wang, and Y. Zhao, "Multiple description convolutional neural networks for image compression," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 8, pp. 2494–2508, 2019.

[52] D. Mishra, S. K. Singh, and R. K. Singh, "Lossy medical image compression using residual learning-based dual autoencoder model," in *2020 IEEE 7th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, 2020, pp. 1–5.

[53] R. M. Slone, D. H. Foos, B. R. Whiting, E. Muka, D. A. Rubin, T. K. Pilgram, K. S. Kohm, S. S. Young, P. Ho, and D. D. Hendrickson, "Assessment of visually lossless irreversible image compression: comparison of three methods by using an image-comparison workstation," *Radiology*, vol. 215, no. 2, pp. 543–553, 2000.

[54] B. J. Erickson, "Irreversible compression of medical images," *Journal of Digital Imaging*, vol. 15, no. 1, pp. 5–14, 2002.

[55] S. Krivenko, V. Lukin, O. Krylova, L. Kryvenko, and K. Egiazarian, "A fast method of visually lossless compression of dental images," *Applied Sciences*, vol. 11, no. 1, p. 135, 2021.

[56] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[57] N. Ponomarenko, V. Lukin, J. Astola, and K. Egiazarian, "Analysis of hvs-metrics' properties using color image database tid2013," in *International Conference on Advanced Concepts for Intelligent Vision Systems*. Springer, 2015, pp. 613–624.

[58] K. J. Balakrishnan and N. A. Touba, "Relationship between entropy and test data compression," *IEEE Transactions on computer-aided design of integrated circuits and systems*, vol. 26, no. 2, pp. 386–395, 2007.

[59] A. Zemliachenko, V. Lukin, N. Ponomarenko, K. Egiazarian, and J. Astola, "Still image/video frame lossy compression providing a desired visual quality," *Multidimensional Systems and Signal Processing*, vol. 27, no. 3, pp. 697–718, 2016.

[60] H. L. Resnikoff, O. Raymond Jr *et al.*, *Wavelet analysis: the scalable structure of information*. Springer Science & Business Media, 2002.

[61] T. Nguyen and J. Walker, "Wavelet-based image compression," 2001.

[62] L. Yue, H. Shen, J. Li, Q. Yuan, H. Zhang, and L. Zhang, "Image super-resolution: The techniques, applications, and future," *Signal Processing*, vol. 128, pp. 389–408, 2016.

[63] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proc. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 1646–1654.

[64] ——, "Deeply-recursive convolutional network for image super-resolution," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1637–1645, 2015.

[65] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 105–114.

[66] W. Shi, J. Caballero, F. Huszar, J. Totz, A. Aitken, R. Bishop, D. Rueckert, and Z. Wang, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, June 27-30 2016, pp. 1874–1883.

[67] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep CNN denoiser prior for image restoration," in *Proceedings of the 2017 IEEE Conference on Computer*

Vision and Pattern Recognition, Honolulu, HI, USA, July 21-26 2017, pp. 2808–2817.

[68] S. Donn, L. Meeus, H. Luong, B. Goossens, and W. Philips, "Exploiting reflectional and rotational invariance in single image superresolution," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Honolulu, HI, USA, July 21-26 2017, pp. 1043–1049.

[69] B. Lim, S. Son, H. Kim, S. Nah, and K. Lee, "Enhanced deep residual networks for single image super-resolution," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Honolulu, HI, USA, July 21-26 2017, pp. 1132–1140.

[70] Y. Fan, H. Shi, J. Yu, D. Liu, W. Han, H. Yu, Z. Wang, X. Wang, and T. Huang, "Balanced two-stage residual networks for image super-resolution," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Honolulu, HI, USA, July 21-26 2017, pp. 1157–1164.

[71] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295–307, 2016.

[72] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Transactions on Image Processing*, vol. 19, pp. 2861–2873, 2010.

[73] R. Timofte, V. Smet, and L. V. Gool, "A+: Adjusted anchored neighborhood regression for fast super-resolution," in *Lecture Notes in Computer Science*, vol. 9006, 2014, pp. 111–126.

[74] C. Dong, C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, pp. 295–307, 2015.

[75] W. Freeman, T. Jones, and E. Pasztor, "Example-based super-resolution," *IEEE Computer Graphics and Applications*, vol. 22, pp. 56–65, 2002.

[76] R. Timofte, V. De, and L. V. Gool, "Anchored neighborhood regression for fast example-based super-resolution," in *Proceedings of the IEEE International*

Conference on Computer Vision, Sydney, Australia, December 1-8 2013, pp. 1920–1927.

[77] G. Freedman and R. Fattal, "Image and video upscaling from local self-examples," *ACM Transactions on Graphics*, vol. 30, pp. 474–484, 2011.

[78] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Residual learning of deep CNN for image denoising," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, 2017.

[79] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual dense network for image super-resolution," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[80] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2861–2873, 2010.

[81] R. Y. Tsai and T. S. Huang, "Multiframe image restoration and registration," in *Proc. Advances in Computer Vision and Image Processing*, 1984, pp. 317–339.

[82] H. Stark and P. Oskoui, "High-resolution image recovery from image-plane arrays, using convex projections." *Journal of the Optical Society of America. A, Optics and image science*, vol. 6 11, pp. 1715–1726, 1989.

[83] A. Patti, M. Sezan, and A. Tekalp, "Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time," *IEEE Transactions on Image Processing*, vol. 6, no. 8, pp. 1064–1076, 1997.

[84] P. E. Eren, M. Sezan, and A. Tekalp, "Robust, object-based high-resolution image reconstruction from low-resolution video," *IEEE Transactions on Image Processing*, vol. 6, no. 10, p. 1446–1451, 1997. [Online]. Available: https://hdl.handle.net/11511/100397

[85] P. Torr and A. Zisserman, "Feature based methods for structure and motion estimation," in *Vision Algorithms: Theory and Practice*, 03 2000, pp. 278–294.

[86] H. He and L. Kondi, "An image super-resolution algorithm for different error levels per frame," *IEEE Transactions on Image Processing*, vol. 15, no. 3, pp. 592–603, 2006.

[87] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 21–36, 2003.

[88] S. Sreehari, S. V. Venkatakrishnan, B. Wohlberg, G. T. Buzzard, L. F. Drummy, J. P. Simmons, and C. A. Bouman, "Plug-and-play priors for bright field electron tomography and sparse interpolation," *IEEE Transactions on Computational Imaging*, vol. 2, no. 4, pp. 408–423, 2016.

[89] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image restoration by sparse 3d transform-domain collaborative filtering," in *Proc. Image Processing: Algorithms and Systems VI*, ser. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, J. T. Astola, K. O. Egiazarian, and E. R. Dougherty, Eds., vol. 6812, feb 2008, p. 681207.

[90] U. S. Kamilov, H. Mansour, and H. B Wohlberg, "A plug-and-play priors approach for solving nonlinear imaging inverse problems," *IEEE Signal Process. Lett*, vol. 24, no. 12, p. 1872–1876, 2017.

[91] S. Ono, "Primal-dual plug-and-play image restoration," *IEEE Signal Processing Letters*, vol. 24, p. 1108–1112, 2017.

[92] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[93] V. K. Ghassab and N. Bouguila, "Plug-and-play video super-resolution using edge-preserving filtering," *Computer Vision and Image Understanding*, vol. 216, p. 103359, 2022.

[94] A. M. Teodoro, J. Bioucas-Dias, and M. Figueiredo, "A convergent image fusion algorithm using scene-adapted gaussian-mixture-based denoising," *IEEE Transaction on Image Processing*, vol. 28, no. 1, p. 451–463, 2019.

[95] V. Sridhar, X. Wang, G. T. Buzzard, and C. C. Bouman, "Distributed iterative CT reconstruction using multi-agent consensus equilibrium," *IEEE Transactions on Computational Imaging*, vol. 6, p. 1153–1166, 2020.

[96] A. Brifman, Y. Romano, and M. Elad, "Unified single-image and video super-resolution via denoising algorithms," *IEEE Transactions on Image Processing*, vol. 28, no. 12, pp. 6063–6076, Dec. 2019.

[97] A. Tekalp, M. Ozkan, and M. Sezan, "High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration," in *Proc. ICASSP-92: 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, 1992, pp. 169–172.

[98] S. Kim, N. Bose, and H. Valenzuela, "Recursive reconstruction of high resolution image from noisy undersampled multiframes," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 6, pp. 1013–1027, 1990.

[99] B. C. Tom, N. P. Galatsanos, and A. K. Katsaggelos, *Super-Resolution Imaging*. Boston, MA: Springer US, 2000, ch. Reconstruction of a High Resolution Image from Multiple Low Resolution Images, pp. 73–105.

[100] R. Schultz and R. Stevenson, "A Bayesian approach to image expansion for improved definition," *IEEE Transactions on Image Processing*, vol. 3, no. 3, pp. 233–242, 1994.

[101] ——, "Extraction of high-resolution frames from video sequences," *IEEE Transactions on Image Processing*, vol. 5, no. 6, pp. 996–1011, 1996.

[102] ——, "Improved definition video frame enhancement," in *Proc. 1995 International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, 1995, pp. 2169–2172.

[103] R. Hardie, K. Barnard, and E. Armstrong, "Joint MAP registration and high-resolution image estimation using a sequence of undersampled images," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1621–1633, 1997.

[104] L. Kondi, D. Scribner, and J. Schuler, "A comparison of digital image resolution enhancement techniques," in *Proc. SPIE AeroSense Conference*, vol. 4719, 2002, pp. 220–229.

[105] H. He and L. P. Kondi, "Resolution enhancement of video sequences with simultaneous estimation of the regularization parameters," in *Proc. Image and*

Video Communications and Processing 2003, ser. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, B. Vasudev, T. R. Hsing, A. G. Tescher, and T. Ebrahimi, Eds., vol. 5022, May 2003, pp. 1123–1133.

[106] H. He and L. Kondi, "MAP based resolution enhancement of video sequences using a huber-markov random field image prior model," in *Proc. 2003 International Conference on Image Processing*, vol. 2, 2003, pp. II–933.

[107] E. S. Lee and M. G. Kang, "Regularized adaptive high-resolution image reconstruction considering inaccurate subpixel registration," *IEEE Transactions on Image Processing*, vol. 12, no. 7, p. 826—837, 2003.

[108] N. P. Galatsanos, V. Z. Mesarović, R. Molina, and A. K. Katsaggelos, "Hierarchical bayesian image restoration from partially known blurs," *IEEE Transactions on Image Processing*, vol. 9, no. 10, p. 1784—1797, 2000.

[109] A. Zomet, A. Rav-Acha, and S. Peleg, "Robust super-resolution," in *Proc. of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, vol. 1, 2001, pp. 645–650.

[110] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Robust shift and add approach to super-resolution," *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 5203, Oct. 2003.

[111] S. Farsiu, M. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE Transactions on Image Processing*, vol. 13, no. 10, pp. 1327–1344, 2004.

[112] A. Lucas, S. López-Tapia, R. Molina, and A. K. Katsaggelos, "Generative adversarial networks and perceptual losses for video super-resolution," *IEEE Transactions on Image Processing*, vol. 28, no. 7, pp. 3312–3327, 2019.

[113] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, pp. 259–268, 1992.

[114] B. Coll and J.-M. Morel, "A review of image denoising algorithms, with a new one," *SIAM Journal on Multiscale Modeling and Simulation*, vol. 4, 01 2005.

[115] G. Chantas, S. Nikolopoulos, and I. Kompatsiaris, “Heavy-tailed self-similarity modeling for single image super resolution,” *IEEE Transactions on Image Processing*, vol. 30, pp. 838 – 852, 11 2020.

[116] S. Gu, L. Zhang, W. Zuo, and X. Feng, “Weighted nuclear norm minimization with application to image denoising,” 06 2014, pp. 2862–2869.

[117] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Foundations and Trends in Machine Learning*, vol. 3, pp. 1–122, 01 2011.

[118] T. Goldstein and S. Osher, “The split bregman method for l1-regularized problems,” *SIAM J. Imaging Sciences*, vol. 2, pp. 323–343, 01 2009.

[119] N. Yurii, “Gradient methods for minimizing composite functions,” *Mathematical Programming*, vol. 140, pp. 125 – 161, 2012.

[120] F. Shi, J. Cheng, L. Wang, P.-T. Yap, and D. Shen, “Lrtv: Mr image super-resolution with low-rank and total variation regularizations,” *IEEE Transactions on Medical Imaging*, vol. 34, no. 12, pp. 2459–2466, 2015.

[121] M. Irani and S. Peleg, “Super resolution from image sequences,” in *Pattern Recognition, 1990. Proceedings., 10th International Conference on*, vol. 2. IEEE, 1990, pp. 115–120.

[122] Y. Wang, J. Qiao, J. Li, P. Fu, S. Chu, and J. Roddick, “Sparse representation-based mri super-resolution reconstruction,” *Measurement*, vol. 47, pp. 946–953, 2014.

[123] E. Tsiligianni, M. Zerva, I. Marivani, N. Deligiannis, and L. Kondi, “Interpretable deep learning for multimodal super-resolution of medical images,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*. Springer International Publishing, 2021, pp. 421–429.

[124] A. Chaudhari, Z. Fang, F. Kogan, J. Wood, K. Stevens, E. Gibbons, J. Lee, G. Gold, and B. Hargreaves, “Super-resolution musculoskeletal mri using deep learning,” *Magnetic Resonance in Medicine*, 2018.

[125] A. Ben-Ezra, H. Greenspan, and Y. Rubner, "Regularized super-resolution of brain mri," in *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*. IEEE, 2009, pp. 254–257.

[126] S. H. Joshi, A. Marquina, S. J. Osher, I. Dinov, J. D. Van Horn, and A. W. Toga, "Mri resolution enhancement using total variation regularization," in *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*. IEEE, 2009, pp. 161–164.

[127] S. Nedjati-Gilani, D. C. Alexander, and G. J. M. Parker, "Regularized super-resolution for diffusion MRI," in *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on*. IEEE, 2008, pp. 875–878.

[128] R. Peeters, P. Kornprobst, M. Nikolova, S. Sunaert, T. Vieville, G. Malandain, R. Deriche, O. Faugeras, M. Ng, and P. Van Hecke, "The use of super-resolution techniques to reduce slice thickness in functional mri," *International Journal of Imaging Systems Technology*, vol. 14, no. 3, pp. 131–138, 2004.

[129] F. Rousseau, "A non-local approach for image super-resolution using intermodality priors," *Medical Image Analysis*, vol. 14, no. 4, pp. 594–605, 2010.

[130] F. Rousseau, K. Kim, and C. Studholme, "A groupwise super-resolution approach: application to brain MRI," in *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on*. IEEE, 2010, pp. 860–863.

[131] Y. Bai, X. Han, and J. Prince, "Super-resolution reconstruction of MR brain images," in *Proc. of 38th Annual Conference on Information Sciences and Systems (CISS04)*, 2004, pp. 1358–1363.

[132] A. Gholipour, J. Estroff, and S. Warfield, "Robust super-resolution volume reconstruction from slice acquisitions: application to fetal brain MRI," *IEEE Transactions on Medical Imaging*, vol. 29, no. 10, pp. 1739–1758, 2010.

[133] S. Rahman and S. Wesarg, "Upsampling of cardiac mr images: Comparison of averaging and super-resolution for the combination of multiple views," in *Information Technology and Applications in Biomedicine (ITAB), 2010 10th IEEE International Conference on*. IEEE, 2010, pp. 1–4.

[134] ——, “Combining short-axis and long-axis cardiac mr images by applying a super-resolution reconstruction algorithm,” in *Medical Imaging 2010: Image Processing*, vol. 7623. International Society for Optics and Photonics, 2010, p. 76230I.

[135] E. Cha, E. Kim, and J. Ye, “k-space deep learning for parallel mri: Application to time-resolved mr angiography,” 2018, available: https://arxiv.org/abs/1806.00806.

[136] H. Chen, Y. Zhang, M. Kalra, F. Lin, Y. Chen, P. Liao, J. Zhou, and G. Wang, “Low-dose ct with a residual encoder-decoder convolutional neural network,” *IEEE Transactions on Medical Imaging*, vol. 36, no. 12, pp. 2524–2535, 2017.

[137] L. Gjesteby, Q. Yang, Y. Xi, B. E. H. Claus, Y. Jin, B. De Man, and G. Wang, “Reducing metal streak artifacts in ct images via deep learning: Pilot results,” in *The 14th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 2017, pp. 611–614.

[138] D. Lee, J. Yoo, and J. C. Ye, “Deep residual learning for compressed sensing MRI,” in *2017 IEEE 14th International Symposium on Biomedical Imaging ISBI 2017*, 2017, pp. 15–18.

[139] H. Shan, Y. Zhang, Q. Yang, U. Kruger, M. K. Kalra, L. Sun, W. Cong, and G. Wang, “3-D convolutional encoder-decoder network for low-dose ct via transfer learning from a 2-D trained network,” *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1522–1534, 2018.

[140] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M. K. Kalra, Y. Zhang, L. Sun, and G. Wang, “Low-dose CT image denoising using a generative adversarial network with wasserstein distance and perceptual loss,” *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1348–1357, 2018.

[141] C. You, Q. Yang, H. Shan, L. Gjesteby, G. Li, S. Ju, Z. Zhang, Z. Zhao, Y. Zhang, W. Cong, and G. Wang, “Structurally-sensitive multi-scale deep neural network for low-dose CT denoising,” *IEEE Access*, vol. 6, pp. 41 839–41 855, 2018.

[142] C. You, W. Cong, M. W. Vannier, P. K. Saha, E. A. Hoffman, G. Wang, G. Li, Y. Zhang, X. Zhang, H. Shan, M. Li, S. Ju, Z. Zhao, and Z. Zhang, “CT super-

resolution gan constrained by the identical, residual, and cycle learning ensemble (GAN-CIRCLE)," *IEEE Transactions on Medical Imaging*, no. 1, p. 188–203, Jan 2020.

[143] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, p. 295–307, Feb. 2016. [Online]. Available: https://doi.org/10.1109/TPAMI.2015.2439281

[144] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proc. IEEE conference on computer vision and pattern recognition*, 2016, pp. 1646–1654.

[145] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and A. Bengio, "Generative adversarial networks," 2014.

[146] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of gans for improved quality, stability, and variation," 2018.

[147] C. Donahue, J. McAuley, and M. Puckette, "Adversarial audio synthesis," in *International Conference on Learning Representations*, 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:52890982

[148] B. M. Li, T. Amvrosiadis, N. Rochefort, and A. Onken, "Synthesising realistic calcium traces of neuronal populations using GAN," 2023.

[149] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 2242–2251.

[150] A. Bansal, S. Ma, D. Ramanan, and Y. Sheikh, "Recycle-gan: Unsupervised video retargeting," in *ECCV*, 2018.

[151] D. Pathak, P. Krähenbühl, J. Donahue, T. Darrell, and A. Efros, "Context encoders: Feature learning by inpainting," in *CVPR*, 2016.

[152] J. Wu, C. Zhang, T. Xue, B. Freeman, and J. Tenenbaum, "Learning a probabilistic latent space of object shapes via 3D generative-adversarial modeling," in *Advances in Neural Information Processing Systems*, D. Lee,

M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett, Eds., vol. 29. Curran Associates, Inc., 2016. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2016/file/44f683a84163b3523afe57c2e008bc8c-Paper.pdf

[153] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE conference on computer vision and pattern recognition*, 2017, pp. 4681–4690.

[154] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. C. Loy, "Esrgan: Enhanced super-resolution generative adversarial networks," in *The European Conference on Computer Vision Workshops (ECCVW)*, September 2018.

[155] C. Yuhua, S. Feng, A. Christodoulou, Z. Zhou, Y. Xie, and D. Li, "Efficient and accurate MRI super-resolution using a generative adversarial network and 3D multi-level densely connected network," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2018, pp. 91–99.

[156] Y. Chen, Y. Xie, Z. Zhou, F. Shi, A. G. Christodoulou, and D. Li, "Brain MRI super resolution using 3D deep densely connected neural networks," in *Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium on*. IEEE, 2018, pp. 739–742.

[157] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[158] J. Tian and R. O. W. Jr., "A lossy image codec based on index coding." *Proc IEEE Computer Society Data Compression Conference*, 1996.

[159] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Transactions on Circuits Systems and Video Technology*, vol. 6, no. 3, p. 243–250, June 1996.

[160] T. Sultana, "Image compression technique wdr and ezw for different wavelet codes." *J. Multim. Process. Technol.*, vol. 6, no. 2, pp. 43–52, 2015.

[161] J. E. Fowler, "Qccpack: An open-source software library for quantization, compression, and coding," *Proc. IEEE Computer Society Data Compression Conference*, vol. 554, 2000.

[162] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle, L. Tarbox, and F. Prior, "The cancer imaging archive (TCIA): Maintaining and operating a public information repository," *Journal of Digital Imaging*, vol. 26, no. 6, pp. 1045–1057, 2013.

[163] X. Wu and N. Memon, "Context-based, adaptive, lossless image coding," *IEEE transactions on Communications*, vol. 45, no. 4, pp. 437–444, 1997.

[164] A. Brifman, Y. Romano, and M. Elad, "Turning a denoiser into a super-resolver using plug and play priors," in *2016 IEEE International Conference on Image Processing, ICIP 2016 - Proceedings*, ser. Proc. International Conference on Image Processing, ICIP, aug 2016, pp. 1404–1408.

[165] G. Farnebäck, "Two-frame motion estimation based on polynomial expansion," in *Proc. Scandinavian Conference on Image analysis*, vol. 2749, 06 2003, pp. 363–370.

[166] E. T. Reehorst and P. Schniter, "Regularization by denoising: Clarifications and new interpretations," *IEEE Transactions on Computational Imaging*, vol. 5, no. 1, pp. 52–67, 2019.

[167] G. T. Buzzard, S. H. Chan, S. Sreehari, and C. A. Bouman, "Plug-and-play unplugged: Optimization free reconstruction using consensus equilibrium," *SIAM Journal on Imaging Science*, vol. 11, no. 3, p. 2001–2020, 2018.

[168] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, 2017.

[169] E. Ryu, J. Liu, S. Wang, X. Chen, Z. Wang, and W. Yin, "Plug-and-play methods provably converge with properly trained denoisers," in *Proc. 36th Int. Conf. Machine Learning (ICML)*, vol. 97, 2019, pp. 5546–5557.

[170] Y. Sun, Z. Wu, X. Xu, B. Wohlberg, and U. S. Kamilov, "Scalable plug-and-play ADMM with convergence guarantees," *IEEE Transactions on Computational Imagimg*, vol. 2, p. 849–863, 2021.

[171] S. H. Chan, X. Wang, and O. A. Elgendy, "Exploiting intra-slice and inter-slice redundancy for learning-based lossless volumetric image compression," *IEEE Transactions on Computational Imagimg*, vol. 2, no. 1, p. 84–98, 2017.

[172] Y. Sun, J. Liu, X. Xu, B. Wohlberg, and U. S. Kamilov, "Block coordinate regularization by denoising," in *Proc. Advances in Neural Information Processing Systems*, vol. 33, 2019, pp. 382—392.

[173] Y. Sun, B. Wohlberg, and U. S. Kamilov, "An online plug-and-play algorithm for regularized image reconstruction," *IEEE Trans. Comp. Imag.*, vol. 5, no. 3, pp. 395–408, 2019.

[174] T. Balke, F. R. Davis, C. Garcia-Cardona, M. McCann, L. Pfister, and B. Wohlberg, "Scientific computational imaging code (SCICO)," *Journal of Open Source Software*, vol. 7, no. 78, Oct. 2022.

[175] L. Wang, Y. Guo, L. Liu, Z. Lin, X. Deng, and W. An, "Deep video super-resolution using HR optical flow estimation," *IEEE Transactions on Image Processing*, vol. 29, pp. 4323–4336, 2020.

[176] X. Wang, K. Yu, K. C. Chan, C. Dong, and C. C. Loy, "BasicSR," https://github.com/xinntao/BasicSR, 2020.

[177] Y. Li, P. Jin, F. Yang, C. Liu, M.-H. Yang, and P. Milanfar, "COMISR: Compression-informed video super-resolution," in *Proc. 2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021, pp. 2523–2532.

[178] K. Anjyo and H. Ochiai, *Rigid Transformation*. Cham: Springer International Publishing, 2017, pp. 5–21. [Online]. Available: https://doi.org/10.1007/978-3-031-02592-1_2

[179] R. Szeliski, *Computer Vision: Algorithms and Applications*, 1st ed. Berlin, Heidelberg: Springer-Verlag, 2010.

[180] S. He and B. Jalali, "Brain MRI image super resolution using phase stretch transform and transfer learning," *arXiv preprint arXiv:1807.11643*, 2018.

[181] X. Zhang and X. Wu, "Image interpolation by adaptive 2-D autoregressive modeling and soft-decision estimation," *IEEE Transactions on Image Processing*, vol. 17, pp. 887–896, 2008.

[182] X. Li and M. Orchard, "New edge-directed interpolation," *IEEE Transactions on Image Processing*, vol. 10, pp. 1521–1527, 2000.

[183] S. Dai, H. Mei, X. Wei, W. Ying, and Y. Gong, "Soft edge smoothness prior for alpha channel super resolution," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Minneapolis, MN, USA, June 17-22 2007, pp. 1–8.

[184] J. Xie, R. Feris, and M. Sun, "Edge-guided single depth image super resolution," *IEEE Transactions on Image Processing*, vol. 25, pp. 428–438, 2016.

[185] W. Dong, L. Zhang, R. Lukac, and G. Shi, "Sparse representation based image interpolation with nonlocal autoregressive modeling," *IEEE Transactions on Image Processing*, vol. 22, pp. 1382–1394, 2013.

[186] S. Farsiu, M. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE Transactions on Image Processing*, vol. 13, no. 10, pp. 1327–1344, 2004.

[187] M. Protter, M. Elad, H. Takeda, and P. Milanfar, "Generalizing the nonlocal-means to super-resolution reconstruction," *IEEE Transactions on Image Processing*, vol. 18, no. 1, pp. 36–51, 2009.

[188] H. Takeda, P. Milanfar, M. Protter, and M. Elad, "Super-resolution without explicit subpixel motion estimation," *IEEE Transactions on Image Processing*, vol. 18, no. 9, pp. 1958–1975, 2009.

[189] C. Liu and D. Sun, "On Bayesian adaptive video super resolution," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 2, pp. 346–360, 2014.

[190] T. Kohler, X. Huang, F. Schebesch, A. Aichert, A. Maier, and J. Hornegger, "Robust multi-frame super-resolution employing iteratively re-weighted minimization," *IEEE Transactions on Computational Imaging*, vol. 2, pp. 42–58, 2016.

[191] Q. Yuan, L. Zhang, H. Shen, and P. Li, "Adaptive multiple-frame image super-resolution based on U-curve," *IEEE Transactions on Image Processing*, vol. 19, no. 12, pp. 3157–3170, 2010.

[192] Y. Huang, W. Wang, and L. Wang, "Video super-resolution via bidirectional recurrent convolutional networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, pp. 1015–1028, 2018.

[193] R. Liao, T. Xin, R. Li, Z. Ma, and J. Jia, "Video super-resolution via deep draft-ensemble learning," in *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, Chile, December 7-13 2015.

[194] A. Kappeler, S. Yoo, Q. Dai, and A. Katsaggelos, "Video super-resolution with convolutional neural networks," *IEEE Transactions on Computational Imaging*, vol. 2, pp. 109–122, 2016.

[195] L. Ding, Z. Wang, Y. Fan, X. Liu, Z. Wang, S. Chang, X. Wang, and T. Huang, "Learning temporal dynamics for video super-resolution: A deep learning approach," *IEEE Transactions on Image Processing*, vol. 27, pp. 3432–3445, 2018.

[196] D. Li and Z. Wang, "Video super-resolution via motion compensation and deep residual learning," *IEEE Transactions on Computational Imaging*, vol. 3, pp. 749–762, 2017.

[197] Q. Dai, S. Yoo, A. Kappeler, and A. Katsaggelos, "Sparse representation based multiple frame video super-resolution," *IEEE Transactions on Image Processing*, vol. 26, pp. 765–781, 2017.

[198] R. Borsoi, G. Costa, and J. Bermudez, "A new adaptive video super-resolution algorithm with improved robustness to innovations," *IEEE Transactions on Image Processing*, vol. 28, pp. 673–686, 2018.

[199] X. Liu, L. Chen, W. Wang, and J. Zhao, "Robust multi-frame super-resolution based on spatially weighted half-quadratic estimation and adaptive btv regularization," *IEEE Transactions on Image Processing*, vol. 27, pp. 4971–4986, 2018.

[200] G. K. Chantas, N. P. Galatsanos, and N. A. Woods, "Super-resolution based on fast registration and maximum a posteriori reconstruction," *IEEE Transactions on Image Processing*, vol. 16, no. 7, pp. 1821–1830, 2007.

[201] M. Haris, G. Shakhnarovich, and N. Ukita, "Recurrent back-projection network for video super-resolution," in *2019 CVPR*, 06 2019, pp. 3892–3901.

[202] ——, "Deep back-projection networks for super-resolution," *CoRR*, vol. abs/1803.02735, 2018. [Online]. Available: http://arxiv.org/abs/1803.02735

[203] Y. Jo, S. W. Oh, J. Kang, and S. J. Kim, "Deep video super-resolution network using dynamic upsampling filters without explicit motion compensation," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 3224–3232.

[204] M. S. M. Sajjadi, R. Vemulapalli, and M. Brown, "Frame-Recurrent Video Super-Resolution," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

Author's Publications

M. C. Zerva, M. Vrigkas, L. P. Kondi, and C. Nikou, "Improving 3d medical image compression efficiency using spatiotemporal coherence," *Electronic Imaging*, vol. 2020, no. 10, pp. 63–1, 2020.

E. Tsiligianni, M. Zerva, I. Marivani, N. Deligiannis, and L. Kondi, "Interpretable deep learning for multimodal super-resolution of medical images," *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*. Springer International Publishing, 2021, pp. 421–429.

M. C. H. Zerva, V. Christou, N. Giannakeas, A. T. Tzallas, and L. P. Kondi, "An Improved Medical Image Compression Method Based on Wavelet Difference Reduction," *IEEE Access*, vol. 11, pp. 18026-18037, 2023, doi: 10.1109/ACCESS.2023.3246948.

M. Ch. Zerva and L. P. Kondi, "Video Super-Resolution Using Plug-and-Play Priors," *IEEE Access*, vol. 12, pp. 11963-11971, 2024, doi: 10.1109/ACCESS.2024.3355195.

Short Biography

Matina Zerva was born in Greece in 1989. She is a researcher and Ph.D student in the department of Computer Science and Engineering of the University of Ioannina, Greece. She received her M.Sc. in Applied Mathematics and Computer Science from the department of Mathematics and her B.Sc. in Computer Science from department of Computer Science and Engineering of the same institution. She works as a computer teacher in a private computer school she owns. Her main areas of interest are medical images compression and enhancement.