Improving the diagnostic quality of fetal MRI: Evaluation of an AI-based post-processing approach for ultrafast image enhancement

Master Thesis

For attainment of the academic degree of
Master of Science in Engineering (MSc)

in the Master Program Digital Healthcare
at St. Pölten University of Applied Sciences

by
Rebecca Mittermaier, BSc.
dh231809

First advisor: FH Prof. Andreas Jakl, MSc
Second advisor: Univ.-Prof.in Dr.in Daniela Prayer

St. Pölten, 10.09.2025

Declaration of Honour

I hereby declare that,

- I have written the work at hand on my own without help from others and I have used no other resources and tools than the ones acknowledged.
- I have complied with the Standards of Good Scientific Practice in accordance with the St. Pölten UAS' Guidelines for Scientific Work when writing this work.
- I have neither published nor submitted the work at hand to another higher education institution for assessment or in any other form as examination work.

Regarding the use of generative artificial intelligence tools such as chatbots, image generators, programming applications, paraphrasing and translation tools, I declare that,

☐ no generative artificial intelligence tools were used in the course of this work.

☐ I have used generative artificial intelligence tools to proof-read this work.

☒ I have used generative artificial intelligence tools to create parts of the content of this work. I certify that I have cited the original source of any generated content. The generative artificial intelligence tools that I used are acknowledged at the respective positions in the text.

Having read and understood the St. Pölten UAS' Guidelines for Scientific Work, I am aware of the consequences of a dishonest declaration.

Preface

My fascination with fetal MRI stems not only from its technical complexities, but also from its profound impact on prenatal diagnostics. It was my daily work as a radiologic technologist that inspired me to deepen my understanding of advanced imaging technologies, leading me to pursue a master's degree in digital healthcare. There, I was able to undertake research in the emerging field of artificial intelligence, a technology that is set to transform our discipline.

The idea for this thesis emerged from my desire to combine my two primary interests, fetal MRI and AI, in a project that integrates technological innovation with tangible clinical relevance. I firmly believe that it is crucial for radiological technologists to engage with such advancements not just as users, but also as critical evaluators. This project has given me the invaluable opportunity to investigate how AI-driven post-processing can enhance diagnostic imaging, as well as consider how such innovations can be responsibly integrated into clinical practice.

I would like to express my heartfelt gratitude to my two thesis supervisors, FH Prof. Andreas Jakl and Univ.-Prof.in Dr.in Daniela Prayer, whose expertise and guidance were invaluable throughout this project. I would also like to thank Univ. Prof. Dr. Gregor Kasprian, MBA, Director of Neuroradiology and MSK Radiology at General Hospital Vienna, for granting me access to the imaging data and for welcoming me so warmly into his highly specialized neuroradiology, fetal imaging and DINLAB team. Their collective support and insightful feedback were instrumental in shaping this thesis.

Finally, I would like to extend a special thank you to my tenderly partner, our families and my closest friends. Their unwavering patience, motivation and belief in me provided me with the foundation I needed to get through the most challenging phases of this journey.

Vienna, 10. September 2025

Rebecca Mittermaier

Abstract

Background: Fetal Magnetic Resonance Imaging (MRI) is crucial tool in prenatal diagnostics, yet it is often challenged by long acquisition times and motion artifacts, which can compromise image quality. Artificial intelligence (AI)-based post-processing techniques, particularly super-resolution, offer a promising solution to enhance image quality without extending scan duration. This thesis evaluates the clinical and technical efficacy of an AI-based super-resolution algorithm for T2-weighted fetal MRI.

Methods: This retrospective, single-center study analyzed 30 fetal MRI datasets acquired on a 3T Philips scanner. Three T2-weighted turbo spin-echo (TSE) sequences were compared for each case: a conventional accelerated sequence (T2 TSE cs), the same sequence enhanced by the AI algorithm (T2 AI SR), and a standard high-resolution sequence (T2 HR cs). Image quality was assessed using both quantitative metrics (Signal-to-Noise Ratio [SNR], Contrast-to-Noise ratio [CNR], Edge-Rise Distance [ERD]) and a qualitative evaluation performed by two expert pediatric neuroradiologists using a blinded, randomized setup with a 4-point Likert scale.

Results: The AI-enhanced sequence (T2 AI SR) was qualitatively rated as significantly superior to both the conventional and high-resolution sequences in perceived image sharpness, tissue contrast, overall image quality, and diagnostic confidence (all < 0.001). This superior diagnostic quality was achieved in 38% less scan time compared to the high-resolution reference. Quantitative metrics, however, did not consistently align with this subjective preference, revealing no significant improvements in SNR and CNR and showing a paradoxical increase in ERD for the AI sequence. This suggests that traditional metrics may not adequately capture the perceptual benefits of AI-driven reconstruction.

Conclusion: AI-based super-resolution post-processing can substantially improve the perceived diagnostic quality and clinical usability of ultrafast fetal MRI sequences while significantly reducing the need for high-resolution images with substantially longer acquisition times. The findings strongly support the integration of this technology to enhance diagnostic confidence and workflow efficiency in the time-sensitive and motion-prone environment of fetal imaging. However, the study also highlights the critical importance of expert perceptual evaluation, as traditional quantitative metrics may not fully reflect the diagnostic utility of AI-generated images.

Kurzfassung

Hintergrund: Die fetale Magnetresonanztherapie (MRT) ist ein entscheidendes Instrument in der pränatalen Diagnostik, wird jedoch oft durch lange Aufnahmezeiten und Bewegungsartefakte beeinträchtigt. Künstliche Intelligenz (KI)-basierte Nachverarbeitungsverfahren, insbesondere Super-Resolution Algorithmen, bieten eine vielversprechende Möglichkeit, die Bildqualität zu verbessern, ohne die Scanzeit zu verlängern. Diese Masterarbeit evaluiert die klinische und technische Wirksamkeit eines solchen KI-basierten Super-resolution Algorithmus für T2-gewichtete fetale MRT-Sequenzen.

Methodik: in dieser retrospektiven Single-Center Studie wurden 30 fetale MRT-Datensätze analysiert, welche an einem 3T Philips Scanner akquiriert wurden. Pro Fall wurden drei T2-gewichtete Turbo-Spin-Echo (TSE) Sequenzen verglichen: eine konventionell beschleunigte Sequenz (T2 TSE cs), dieselbe Sequenz nach Anwendung des KI-Algorithmus (T2 AI SR) und eine Standard-hochauflösende Sequenz (T2 HR cs). Die Bildqualität wurde sowohl anhand quantitativer Metriken (Signal-Rausch-Verhältnis [SNR], Kontrast-Rausch-Verhältnis [CNR], Edge-Rise Distance [ERD]) als auch durch eine qualitative Bewertung durch zwei erfahrene pädiatrische Neuroradiolog*Innen beurteilt. Die qualitative Analyse erfolgte verblindet, randomisiert und mittels einer 4-Punkte-Likert Skala.

Ergebnisse: Die KI-verbesserte Sequenz (T2 AI SR) wurde qualitativ als signifikant überlegen gegenüber der konventionellen und der hochauflösenden Sequenz in den Kategorien Bildschärfe, Gewebekontrast, Gesamtbildqualität und diagnostische Sicherheit bewertet (alle < 0.001). Diese überlegende diagnostische Qualität wurde in 38% kürzerer Scanzeit im Vergleich zur hochauflösenden Referenzsequenz erreicht. Die quantitativen Metriken spiegelten diese subjektive Präferenz jedoch nicht konsistent wider: Es zeigten sich keine signifikanten Verbesserungen bei SNR und CNR sowie ein paradoxer Anstieg der ERD-Wertes für die KI-Sequenz. Dies deutet darauf hin, dass traditionelle Metriken die perzeptuellen Vorteile der KI-rekonstruktion nur unzureichend erfassen.

Schlussfolgerung: Die KI-basierte Super-Resolution-Nachverarbeitung kann die wahrgenommene diagnostische Qualität und klinische Nützlichkeit von ultraschnellen fetalen MRT-Sequenzen erheblich verbessern und gleichzeitig den Bedarf an hochauflösenden Bildern mit wesentlich längeren Aufnahmezeiten deutlich reduzieren. Die Ergebnisse unterstützen nachdrücklich die Integration dieser Technologie zur Steigerung der diagnostischen Sicherheit und der

Workflow-Effizienz in der zeitkritischen und bewegungsanfälligen fetalen Bildgebung. Die Studie unterstreicht jedoch auch die entscheidende Bedeutung der perzeptuellen Expertenbewertung, da traditionelle quantitative Metriken den diagnostischen Nutzen von KI-generierten Bildern möglicherweise nicht vollständig abbilden.

Abbreviations

MRI	Magnetic Resonance Imaging
US	Ultrasound
CNS	Central Nervous System
AI	Artificial Intelligence
SNR	Signal-to-Noise Ratio
CNR	Contrast-to-Noise Ratio
ERD	Edge-to-Rise Distance
CNS	Central-Nervous System
CCA	Corpus Callosum Agenesis
T	Tesla
SAR	Specific Absorption Rate
T2w	T2-weighted
TSE	Turbo Spin Echo
HASTE	Half-Fourier Acquisition Single-shot Turbo spin Echo
SSFSE	Single-Shot Fast Spin Echo
DWI	Diffusion Weighted Imaging
EPI	Echo Planar Imaging (EPI)
bSSFP	balanced Steady-State Free Precision
FFT	Fast Fourier Transformation
CS	Compressed Sense
CNN	Convolutional Neural Network
GAN	Generative Adversarial Network
ISTA	Iterative Shrinkage Threshold Algorithm

SRGAN	Super Resolution Generative Adversarial Networks
NGSA	Next Generation Scan Acceleration
GDPR	Data Protection Regulation
HR	High Resolution
PACS	Picture Archiving and Communication System
ROI	Regions of Interest
CSF	Cerebro Spinal Fluid
IQR	Interquartile Range
GPU	Graphic Processing Unit
Con	Conventional

Table of Content

1 Introduction

Fetal Magnetic Resonance Imaging (MRI) has emerged as a valuable complementary imaging tool in prenatal diagnostics, particularly when ultrasound (US) alone is insufficient for a comprehensive assessment [1]. Unlike US, fetal MRI offers superior soft tissue contrast and a broader field of view, enabling detailed visualization of fetal anomaly even when structural anomalies are partially obscured by ossification or unfavorable positioning. It plays a crucial role in the assessment of complex congenital abnormalities, particularly those involving the central nervous system (CNS), such as corpus callosum agenesis (CCA) and other brain malformations [2], [3], [4].

Despite its strength, fetal MRI is not without challenges. The need to acquire diagnostically usable images of a moving fetus imposes technical limitations that are absent in adulting imaging. Fetal MRI cannot utilize respiratory gating or external motion tracking, which increases the risk of motion artifacts [3]. Prolonged acquisition times, susceptibility to motion related image degradation, and dependence on the cooperation of the pregnant patient can reduce image quality and limit its broader use in routine clinical workflows. Additionally, fetal MRI interpretation remains highly specialized and time-consuming, requiring experienced radiologists and often manual post-processing [3]. Even with advances in high-field MRI systems such as 3 Tesla, where higher signal-to-noise ratio enable improved resolution, new challenges emerge – namely, increased scan duration, heightened sensitivity to motion, and risks of maternal discomfort due to increased energy deposition and heating [5]

A systematic review has shown that fetal MRI provides added diagnostic value in approximately 22.5% of cases where ultrasound was inconclusive [4], underscoring its importance in high-risk pregnancies and in diagnostic clarification of CNS anomalies. Yet, the technical barriers mentioned above limit its accessibility and reliability, especially in time-sensitive clinical environments. Even with improvements in acquisition protocols, fetal MRI remains resource-intensive and requires post processing workflows that are often slow and manual.

In this context, artificial intelligence (AI)-based post processing emerges as a promising solution to overcome the trade-offs between image quality and acquisition speed. Recent research has focused on the application of convolutional neural networks to enhance image quality retrospectively, after acquisition, through techniques such as denoising, artifact correction, and super resolution. These approaches aim to reconstruct high resolution images from low resolution or rapidly acquired scans without requiring longer scan times or additional patient exposure [3], [4]. AI-based post processing could also introduce potential workflow improvements by reducing the need for repeated scans, increasing diagnostic confidence, and standardize image quality across varying acquisition parameters. Moreover, deep learning models are increasingly capable of automating key aspects of image evaluation, including anomaly detection, segmentation, and quantitative analysis, thereby possibly supporting radiologists in decision-making and reducing inter-observer variability [3].

By leveraging AI-enhanced imaging techniques, fetal MRI has the potential to contribute meaningfully to both clinical and operational improvements in prenatal diagnostics. It could improve diagnostic confidence, particularly in cases requiring detailed morphological assessment of the fetal brain. In addition to image-level improvements, AI-based models could also support postnatal outcome prediction by integrating multi-parametric MRI data with genetic and clinical information [4]. These potential benefits underline the need for rigorous validation. For AI-enhanced fetal MRI to become a standard clinical tool, it must be assessed for its technical robustness, clinical feasibility, patient acceptability, and its impact on diagnostic decision-making [3].

This study evaluates the clinical feasibility and diagnostic benefits of AI-enhanced fetal MRI. The central aim is to investigate whether AI-based post processing can improve image quality and usability in fetal MRI scans acquired under time or resolution constraints. The research specifically focuses on a convolutional neural network method designed to apply super resolution and artifact reduction retrospectively to standard fetal MRI sequences. A retrospective dataset is analyzed using both quantitative and qualitative assessment methods. Quantitative image quality is evaluated using metrics such as Signal-to-Noise Ratio (SNR), Contrast-to-Noise-Ratio (CNR), and Edge-to-Rise Distance (ERD), while qualitative analysis involves structured expert evaluation of diagnostic usability. Statistical methods, including paired t-test are used to compare AI-enhanced sequences with their conventional counterparts.

The following research questions guide this investigation:

1. To what extent can AI-based post processing improve image quality of ultrafast acquired fetal MRI images?

2. How does AI-enhanced low-resolution MRI compare to standard high-resolution sequences in diagnostic value?
3. Can AI-based postprocessing facilitate shorter acquisition times while maintaining or improving image quality?
4. How do expert radiologists evaluate the diagnostic usability of AI-enhanced fetal MRI images?

This thesis is structured as follows: Chapter 1 introduces the problem statement, motivation and research objectives. Chapter 2 provides an overview of fetal MRI, including its technical foundations, clinical challenges, and AI-driven image enhancement techniques. Chapter 3 details the methodology, including data acquisition, implementation of AI-based post processing tools, and the evaluation framework for both quantitative and qualitative analyses. Chapter 4 presents the experimental results, followed by a critical interpretation of findings in Chapter 5, placing them in the context of current literature. Finally, Chapter 6 concludes the study, summarizing key insights, limitations, and directions for future research.

2 Background and Related Work

2.1 Fetal MRI – Technical Principles

Fetal MRI has established itself as an important supplemental tool in prenatal diagnostics, particularly when conventional ultrasound reaches its diagnostic limits. As a non-invasive and radiation-free modality, fetal MRI enables detailed anatomical assessment by providing multiplanar views and a high soft tissue contrast. It is particularly useful for clarifying complex or uncertain findings, such as suspected anomalies of midline brain structure or cortical development [6], [7].

MR imaging is based on the magnetic properties of hydrogen protons within body tissues. When exposed to a strong magnetic field and radiofrequency pulses, these protons emit signals that are captured and transformed into detailed images via mathematical operations such as the Fourier transformation. In the context of fetal imaging, MRI is conventionally conducted without the application of contrast agents. The utilization of rapid sequence protocols is paramount in this setting, as they serve to minimize the degradation that is often associated with motion-related artefacts [6].

A significant factor to be taken into consideration when performing fetal MRI is the selection of the appropriate field strength. The utilization of both 1.5 Tesla (T) and 3T scanners is a common clinical practice. Whilst 3T scanners offer a higher signal-to-noise ratio, better spatial resolution, and potentially shorter scan times, they also increase sensitivity to artefacts and can pose challenges due to specific absorption rate (SAR) limitations and maternal discomfort [5]. Therefore, both field strengths are utilized in accordance with the clinical indication and the institutional setup.

The most common sequence applied in fetal MRI is the T2-weighted (T2w) single-shot turbo spin echo (TSE), also known as HASTE (Half-Fourier Acquisition Single-shot Turbo spin Echo) or SSFSE (Single-Shot Fast Spin Echo). They provide excellent contrast between cerebrospinal fluid and brain parenchyma, facilitating the visualization of sulcation, ventricles, midline structures, and early cortical development [7].

Whilst T1-weighted sequences are utilized less frequently, they offer essential diagnostic information within specific clinical context. They are particularly useful for identifying fat-containing lesions, hemorrhages, and calcifications, and thus contribute to the evaluation of conditions such as teratomas, bowel obstructions, and intracranial bleeding [8].

The employment of additional sequences may be justified to support differential diagnoses or to clarify complex findings. These include diffusion-weighted imaging (DWI) to assess white matter integrity, echo planar imaging (EPI) for vascular anomalies, and balanced steady-state free precision (bSSFP) sequences for detailed thoracoabdominal evaluation [1], [4], [7].

Contemporary fetal MRI protocols typically involve the acquisition of multiplanar images, encompassing the axial, sagittal and coronal planes, through the fetal brain and body. These are adjusted in real time according to the fetus position, with slice thickness typically ranging from 2 to 4 mm depending on field strength, desired resolution, and the need for expediency [5], [7].

Whilst these imaging strategies facilitate comprehensive anatomical assessment, fetal MRI remains subject to challenges pertaining to motion, scan duration, and the consistency of image quality. These issues are addressed in the following section.

2.2 Challenges of fetal MRI

Despite its many advantages, fetal magnetic resonance imaging is associated with several significant limitations that can affect image quality, diagnostic reliability, and clinical efficiency. These challenges stem primarily from the dynamic nature of fetal anatomy and technical constraints of MRI acquisition under prenatal conditions.

One of the most prominent challenges is motion, particularly spontaneous fetal movement. Unlike adult patients, fetuses cannot be instructed to remain still, and sedation is not ethically or clinically justifiable in most cases. Techniques such as respiratory gating or external motion tracking are not applicable in the fetal context [3]. As a result, even rapid sequences are susceptible to motion related blurring, which can limit the visibility of small anatomical structures or pathology.

The time required to acquire high-quality, high-resolution MRI scans of the fetus is relatively long, especially when multiple planes and sequences are necessary. Longer acquisition times not only increase the risk of motion artifacts but can also cause maternal discomfort, particularly in late pregnancy, leading to further image degradation due to involuntary movements [5].

Another challenge is the limited standardization of fetal MRI protocols across institutions. Variability in sequence parameters, field strength, and post processing techniques can make it difficult to compare studies or establish universal diagnostic thresholds. Additionally, the interpretation of these MRI requires specialized expertise, often limited to tertiary care centers, which restricts broader access to this modality [1], [3].

To mitigate motion effects, ultrafast sequences such as single shot T2-weighted imaging are frequently used. However, these come as the cost of reduced spatial resolution and signal-to-noise ratio, especially in brain structures. This trade-off between speed and detail often limits the detection of subtle anomalies or the precise delineation of cortical and white matter development [4], [7].

While 3T scanners offer higher signal strength and better resolution compared to 1.5T, they also produce more pronounced susceptibility artifacts, particularly in fluid-filled structure such as amniotic fluid. Moreover, increased SAR and the associated maternal heating may lead to discomfort, although current evidence suggests no fetal harm [5]. Nonetheless, these issues can necessitate repetition of sequences, prolong the exam and increase patient burden.

2.3 Image Optimization and Acceleration

One of the principal challenges in fetal MRI lies in minimizing scan duration while maintaining high diagnostic image quality. To mitigate the inherent limitations of fetal MRI, including motion artifacts, prolonged acquisition times, and limited spatial resolution, both conventional and advanced image reconstruction strategies have been employed. These range from zero-filling, wavelet-based denoising, and compressed sensing to state-of-the-art, AI-driven approaches such as deep learning-based reconstruction algorithms. The latter could be increasingly incorporated into clinical workflows through commercially available solutions like Philips SmartSpeed and Siemens DeepResolve. Collectively, these methods seek to accelerate data acquisition and improve image fidelity, thereby could be enhancing the robustness and clinical applicability of fetal MRI.

2.3.1 Zero-Filling and Wavelet-Based Denoising

MRI data acquired in the frequency domain, known as k-space, represents the spatial frequency content of the object in question rather than its visual appearance. The relationship between the raw data in k-space and the final anatomical image is fundamental to MRI reconstruction (shown in Figure 1).

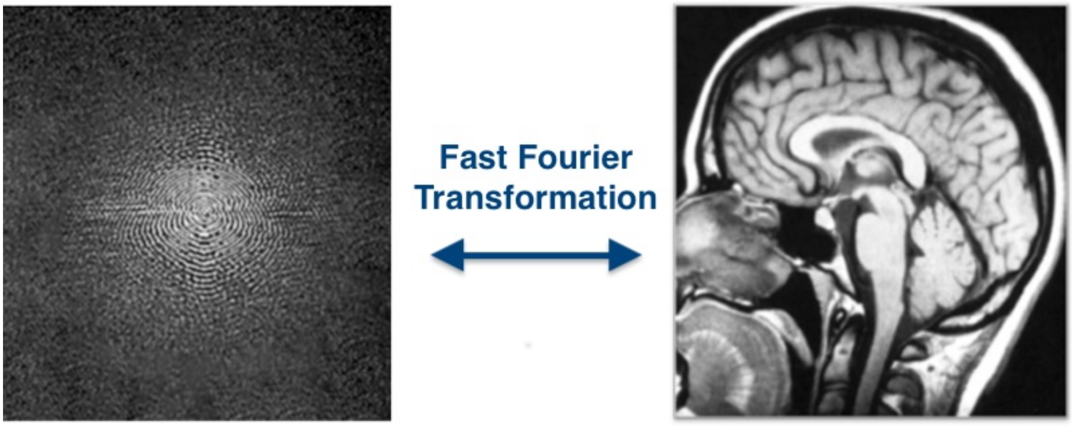

Figure 1: The relationship between k-space and the final MR image (source: https://mriquestions.com/what-is-k-space.html, Courtesy of Allen D. Ellster, MRIquestions.com)

The central region k-space contains low-frequency signals that define image contrast and general shape, while the outer regions encode high-frequency information responsible for sharp edges and fine anatomical detail. To reconstruct a viewable image, the acquired data must be mathematically transformed into the spatial domain using a Fast Fourier Transform (FFT). As the acquisition of high-frequency k-space data is both time consuming and susceptible to noise and motion, methods such as zero-filling are employed to enhance apparent resolution without extending scan time [9], [10].

Zero filling involves the padding of the outer regions of the acquired k-space matrix with zeros prior to the Fast Fourier Transformation. Even though this technique does not contribute to the acquisition of new physical or anatomical information, it has been demonstrated to enhance the density of the image matrix. This, in turn has been shown to assist in the reduction of partial volume artifacts and pixelation. This results in smoother contour transitions and improved edge definition, especially in regions with steep signal gradients (shown in Figure 2). Furthermore, this technique has been demonstrated to enhance the signal-to-artifact ratio and emulate the effects of sinc-based interpolation, while preserving the integrity of original signal. [10], [11].

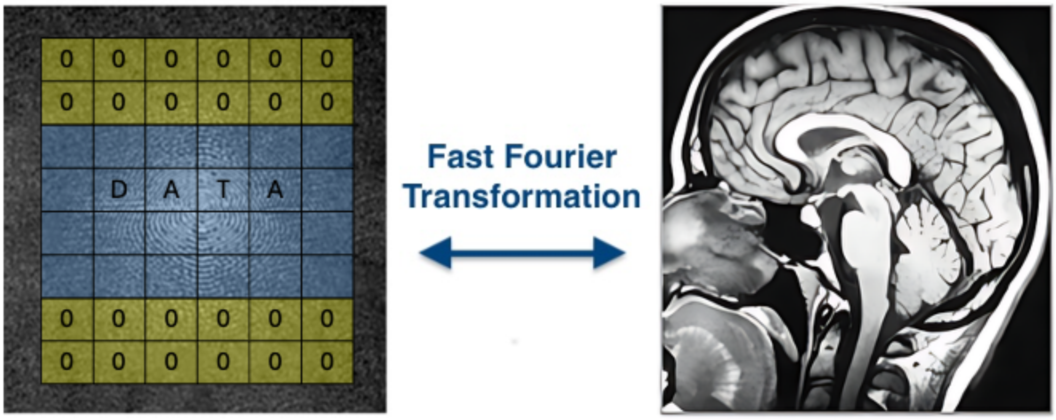

Figure 2: Schematic illustration of the zero-filling process (source: own illustration, adapted by https://mriquestions.com/what-is-k-space.html)

It is important to distinguish zero-filling from interpolation techniques, which estimate missing voxel values in the spatial domain based on surrounding intensity patterns. While interpolation improves visual impressions by smoothing, it does not enhance frequency resolution or impact the actual data acquired. Zero-filling, by contrast, operates direct in the k-space and consequently exerts an influence on the outcome of the reconstruction process itself [11].

Beyond zero-filling, wavelet-based denoising and encoding techniques have emerged as powerful tools to enhance MRI quality, particularly in conditions of undersampling, low SNR, or motion. The wavelet transform is a process which decomposes the MR images into multiple frequency bands, allowing selective suppression of noise, while preserving relevant structures [12], [13]. In addition, wavelet-based encoding has shown distinct advantages in comparison to conventional phase encoding methods. Firstly, wavelet encoding is less prone to Gibbs ringing artifacts, which often result from Fourier undersampling near sharp transitions, Secondly, wavelet methods demonstrated robustness enhanced against motions artifacts. And thirdly, it is conceivable that shorter imaging times can be attained by meticulously arranging wavelet-encoding selective excitations [12].

2.3.2 Compressed Sensing

While zero-filling and wavelet-based denoising enhance image quality during reconstruction phase, they do not address the fundamental limitation of conventional MRI: namely, the time intensive process of acquiring fully sampled k-space data. Compressed Sensing (CS) has been demonstrated to affect a paradigm shift in the realm of medical imaging, with the potential to reduce scan time by acquiring a reduced number of data points during acquisition, while still enabling high-fidelity image reconstruction [14]. This approach renders CS particularly valuable in time sensitive settings such as fetal MRI.

The theory of compressed sensing is predicted on three fundamental principles, as illustrated schematically in Figure 3 (Philips).

1. Sparse undersampling of k-space:

Instead of acquiring data in uniform manner across k-space, CS employs a pseudo random or variable density sampling pattern. This strategy preserves dense sampling at the center of k-space – where most of the signal energy resides – and sparsely samples the periphery. Such incoherent undersampling generates noise-like artifacts rather than structured aliasing, which are easier to eliminate during reconstruction. [15]

2. Transform Domain Sparsity:

A fundamental assumption in the field of Compressed Sense is that MR images are sparse or compressible in specific transform domains, such as wavelet or Fourier. In these domains, only a small number of coefficients carry meaningful signal, while the rest are negligible. This property facilitates precise reconstruction from a limited number of measurements. By leveraging this sparsity, CS significantly reduces the data acquisition burden without compromising diagnostic content. [16]

3. Iterative Image Reconstruction:

The reconstruction process in CS is not a one-step process as is the case in traditional inverse Fourier transforms. The proposed methodology involves a nonlinear, iterative optimization algorithm that alternates between the image domain and a sparsifying transform domain. The following simplified steps are typically taken:

- The initial image is generated from the undersampled k-space using a Fourier transform, and this image appears to be noisy and prone to artifacts.
- A sparsifying transform (e.g. wavelet) is applied to concentrate the relevant signal into a small number of coefficients.
- The process of denoising involves the suppression of low intensity, noise-like components. This can be achieved through various techniques such as thresholding.

- The signal that has been cleaned up is then transformed back into image space, and consistency with the original k-space data is enforced.
- The updated image is then compared with the previous version, and the process is repeated until convergence is achieved.

Through this cyclical refinement process, the algorithm is able to converge on an image that is consistent with the measured data and devoid of incoherent artifacts [15].

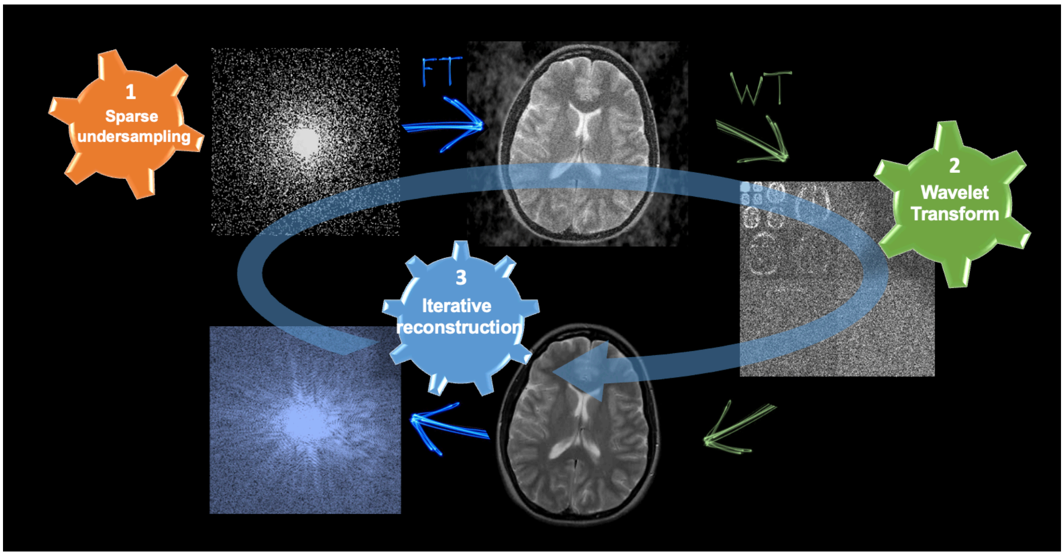

Figure 3: Schematic representation of the Compressed Sensing pipeline (adapted from Philips Healthcare). (1) Sparse, undersampling of k-space. (2) Transform domain sparsity (e.g. wavelet transform). (3) iterative Image Reconstruction.

Compressed sensing has been proven to enhance acquisition efficiency through mathematical optimization [14]. However, recent advances in artificial intelligence, particularly convolutional neural networks, offer even greater potential for real-time image enhancement and reconstruction [17]. The subsequent section will introduce AI-driven techniques, with a focus on a vendor-integrated solution such as Philips SmartSpeed.

2.3.3 Deep Learning-Based Acceleration Philips SmartSpeed

Recent advances in deep learning have resulted in the development of AI-based reconstruction techniques that overcome the limitations of zero-filling and compressed sensing. These models have the capacity to learn complex mappings between undersampled k-space data and fully reconstructed images using training datasets, thereby improving image quality and acquisition speed in challenging conditions [17]. For instance, Pezzotti et al. (2020) demonstrated that the Adaptive-

CS-Net significantly improved SNR and perceptual image quality in undersampled MRI data, while Beljaards et al. (2024) showed improved robustness under motion using AI-driven artifact detection and reconstruction selection [18], [19].

The Philips SmartSpeed innovation is based on these principles, integrating deep learning into the MR reconstruction pipeline. The program is composed of two core modules. The Adaptive-CS-Net is utilized for denoising, and the Precise Image Net for image refinement [20], [21]. The following subsection explains the function of each module and their role in the reconstruction process.

The SmartSpeed Architecture: a dual AI engine for MR reconstruction:

1. **Adaptive-CS-Net:**

A key component of the SmartSpeed framework is the Adaptive-CS-Net, a convolutional neural network (CNN) that is used to reconstruct undersampled MR data at an early stage in the reconstruction pipeline. Unlike traditional zero-filling or handcrafted compressed sensing algorithms, this model is trained to predict and fill in missing k-space information using frequency-based priors that have been learned. The CNN operates on raw, multichannel k-space data prior to coil combination, allowing for better preservation of fine structural detail and reducing noise typically resulting from acceleration-induced undersampling [18], [22].

The conceptual basis of the Adaptive-CS-Net is inspired by the Iterative Shrinkage-Thresholding Algorithm (ISTA), a classical optimization method used to solve sparse reconstruction problems. In ISTA, each iteration alternates between applying a sparsity-promoting transform (e.g. wavelets), suppressing small or noisy coefficients through thresholding and enforcing consistency with the measured k-space data. Adaptive-CS-Net generalizes this idea by replacing the fixed mathematical operators with learnable components, namely CNN modules trained on large-scale datasets. This enables the network to learn to suppress noise more effectively and restore fine anatomical features without relying on manually defined thresholds or fixed transforms [17], [18].

As Jurka et al. (2024) emphasize, this architecture enables Adaptive-CS-Net to reduce noise in accelerated sequences and mitigate the blurring effects typically observed in compressed sensing reconstructions. In the context of medical imaging, this model was trained using 740,000 MR images from various anatomical regions and contrast types at 1.5T and 3T field strengths [20], [22].

2. **Precise Image Net:**

As part of the SmartSpeed reconstruction framework, Precise Image Net is applied after the initial CNN-based k-space reconstruction performed by Adaptive-CS-Net. While the first stage focuses on reconstructing undersampled raw data in the frequency domain, Precise Image Net operates entirely in the image domain. It aims to enhance image quality further by addressing residual artifacts and improving spatial resolution.

Jurka et al. (2024) describe Precise Image Net in their technical implementation as a deep learning model trained to replace the traditional zero-filling step, which is used to increase the matrix size to improve display resolution. Rather than zero-padding the outer regions of the k-space – an approach that visually sharpens images but adds no real structural information – Precise Image Net generates high-resolution data. The authors also highlight that the network was trained using synthetically cropped k-space data to induce Gibbs ringing artifacts, enabling the model to learn to identify them and suppress them. In their evaluation, this resulted in images with higher SNR, greater sharpness, and a notably incidence of ringing artifacts [22].

The architecture of this network is inspired by the structure of Super Resolution Generative Adversarial Networks (SRGANs), which were introduced by Ledig et al. (2017). These models use a generator – discriminator pair: the generator network learns to produce photorealistic, high-resolution images, and the discriminator network learns to distinguish these from genuine, high-resolution references. This adversarial training framework enables the model to optimize for both pixel accuracy and perceptual quality by incorporating high-level image features via a Visual Geometry Group based perceptual loss function [23].

In the context of MRI, Precise Image Net serves multiple functions. It replaces zero-filling, enabling a more informed upsampling of image resolution. It reduces Gibbs ringing by training on artifacts generated via controlled k-space cropping and it enhances visual detail and texture via GAN-based super-resolution and perceptual loss optimization [22], [23].

Alongside Adaptive-CS-Net, Precise Image Net forms a two stage, AI driven reconstruction pipeline (systematically shown in Figure 4, adapted by Philips Healthcare) that replaces the traditional compressed sensing and filtered back projection steps with data-driven, clinically tuned modules.

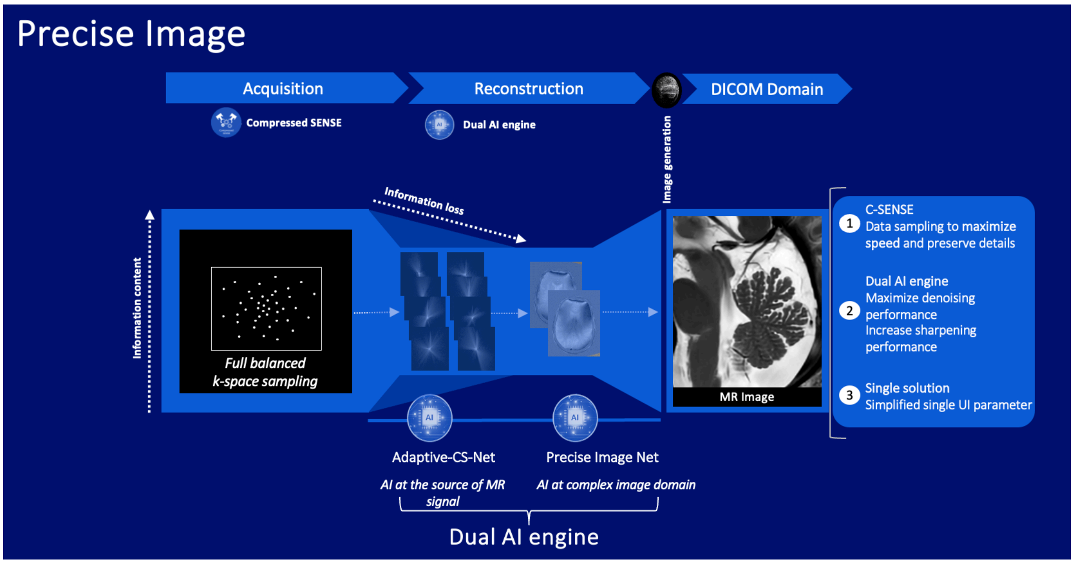

Figure 4: Philips SmartSpeed Dual AI Engine Reconstruction Pipeline, Illustration from Philips

As illustrated in Figure 4, the reconstruction process follows a sequential architecture. The individual steps can be described as follows:

1. **Data Acquisition**: Raw MRI data are acquired using a pseudo-randomly undersampling pattern (CS) in k-space. This allows a reduction in scan time while preserving the central frequency components essential for image contrast.
2. **Adaptive-CS-Net:** In the early reconstruction stage, the CNN-based Adaptive-CS-Net is applied directly to the undersampled multichannel k-space data. This model reconstructs missing data by leveraging learned anatomical priors, reducing noise and preserving structure prior to coil combination.
3. **Fourier transformation:** The now completed k-space data are converted into image space via inverse Fourier transform, resulting in a first-stage spatial image.
4. **Precise Image Net:** The second AI module acts on this intermediate image, performing denoising, ringing artifact suppression, and super-resolution enhancement. By replacing traditional zero-filling, this model increases matrix size and sharpness while maintaining data consistency.
5. **Final Output:** The processed image exhibits improved SNR, reduced artifacts, and enhanced spatial detail.

Recent studies have demonstrated the clinical feasibility and diagnostic potential of AI-based reconstruction using SmartSpeed across various anatomical applications. For instance, Harder et al. (2022) and Bischoff et al. (2023) reported improvements in image sharpness and reconstruction speed in prostate MRI, while Hahnfeldt et al. (2023) demonstrated comparable advantages in musculoskeletal imaging [21], [24], [25]. These findings emphasize the increasing integration of deep learning-based reconstruction methods in clinical workflows, highlighting their importance in settings requiring high spatial resolution and efficient acquisition.

Despite these advancements, the potential of SmartSpeed in fetal MRI remains relatively unexplored in the existing literature. Given the unique requirements of fetal imaging, such as the need for fine anatomical detail, fast acquisition times and frequent motion, deep learning-driven reconstruction could offer substantial advantages. This thesis therefore aims to evaluate the technical and clinical feasibility of SmartSpeed for fetal MRI using a combination of quantitative image quality metrics and qualitative expert assessment.

Recent studies have demonstrated the clinical feasibility and diagnostic potential of AI-based reconstruction using Philips SmartSpeed across various anatomical applications. For instance, Harder et al. (2022) and Bischoff et al. (2023) reported improvements in image sharpness and reconstruction speed in prostate MRI, while Hahnfeldt et al. (2023) demonstrated comparable advantages in musculoskeletal imaging [21], [24], [25].

The study by Bischoff et al. provides a compelling example of the technologies impact. As shown in Figure 5, the AI-reconstructed T2-weighted images (T2_DL) not only required significantly less acquisition time but also demonstrated superior image quality. The images exhibit enhanced sharpness of anatomical boundaries and improved lesion conspicuity compared to both standard cartesian (T2_C) and non-cartesian (T2_NC) sequences. Quantitatively, this translated to significantly higher median Likert scores for overall image quality and sharpness, as well as lower (i.e., better) edge-rise distance [21]. This evidence from related clinical field forms the basis for the hypothesis of this thesis: that similar benefits can be achieved in the challenging domain of fetal MRI.

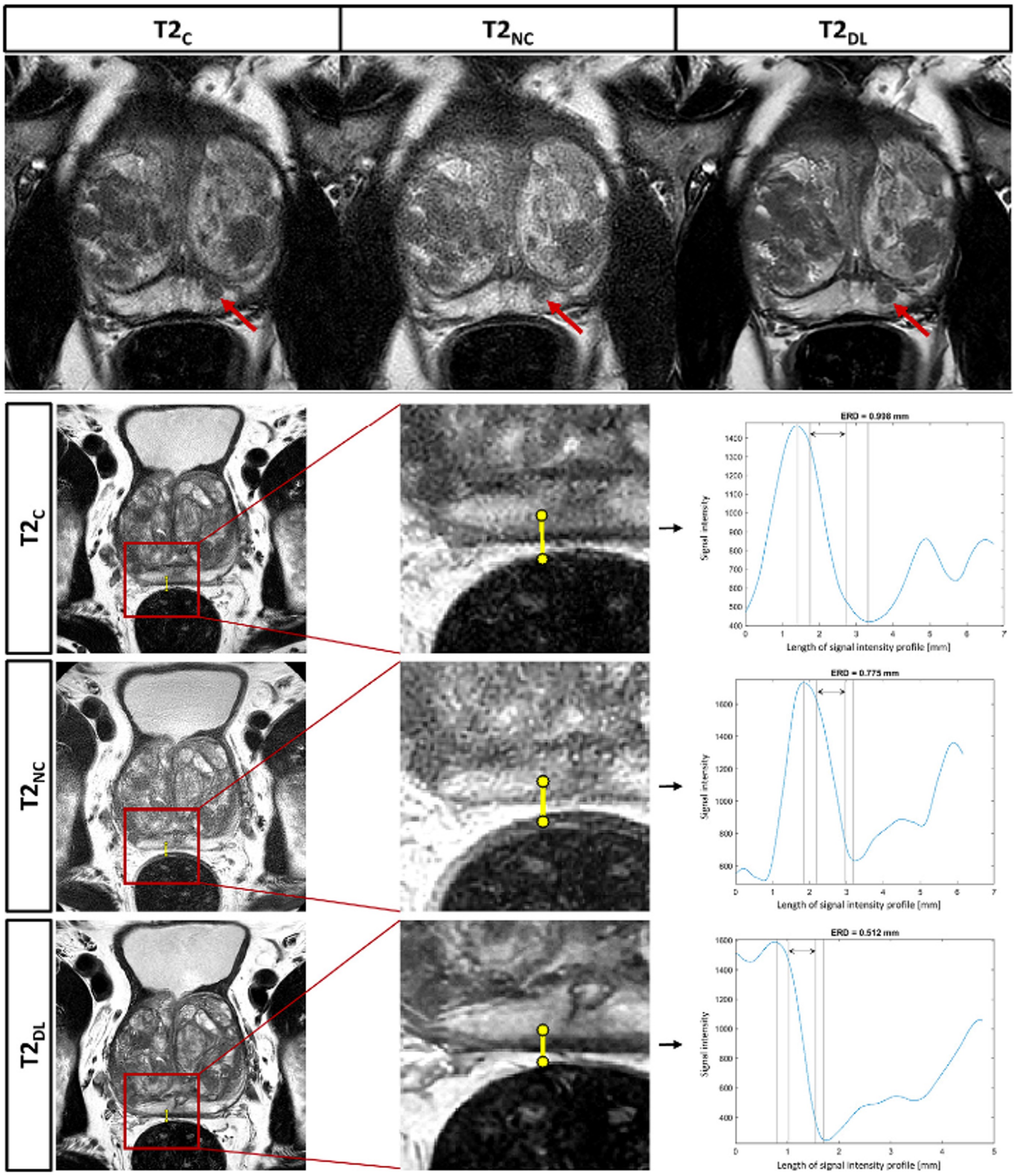

Figure 5: Image comparison of T2-weighted prostate MRI. Legend: $T2_C$ – Standard Cartesian, $T2_{DL}$– AI based deep learning super-resolution reconstruction, $T2_{NC}$ – non-Cartesian. The AI reconstructed image ($T2_{DL}$) shows improved delineation of the prostate capsule and internal structures with reduced scan time. (source: Adapted from Bischoff et al., Radiology, 2023 [21])

3 Requirements / Methods

This chapter outlines the methodological framework and technical setup for the retrospective evaluation of AI-supported image reconstruction in fetal MRI. This study aims to evaluate the potential of AI-based reconstruction techniques to enhance the image quality of T2-weighted turbo spin echo sequences in fetal MRI, considering both objective image quality metrics and expert-based subjective assessments.

This study builds upon recent clinical research investigating deep learning-driven MRI reconstruction using a prototype provided by the vendor called Next Generation Scan Acceleration (NGSA) Patch. This research patch incorporates the core reconstruction components of the Philips SmartSpeed solution, including Adaptive-CS-Net and Precise Image Net. It has demonstrated promising results in various clinical applications, including prostate and musculoskeletal imaging [21], [24], [25]. These studies reported improvements in image sharpness, noise reduction and diagnostic confidence – factors that are particularly relevant in fetal imaging, where small-scale anatomy, motion artifacts and the need for fast acquisition presents unique challenges.

The following sections provide a detailed overview of the study design, the ethical approval process, the imaging system and acquisition parameters, the data selection criteria and the methodological procedures used for image quality assessment. This includes both quantitative metrics and qualitative scoring, as well as the tools and statistical methods employed in the analysis.

3.1 Study Design and Ethical Approval

The study is embedded within a broader exploratory research project titled "AI-supported image reconstruction in fetal MRI", which is being conducted at the Department of Neuroradiology and Musculoskeletal Radiology at the Medical University of Vienna. The present work constitutes a retrospective sub-analysis conducted in the context of a master's theses, with a focus on the evaluation of AI-based postprocessing in fetal magnetic resonance imaging.

Ethical approval was granted by the Ethics Committee of the Medical University of Vienna (vote EK Nr: 2199/2017). An amendment was submitted and approved in December 2024 to include this retrospective analysis. It describes the comparative evaluation of conventionally acquired high resolution T2-weighted TSE sequences and accelerated T2w TSE sequences acquired using compressed sense. AI-based post-processing was then applied to the CS sequences using the NGSA research patch based on a CNN architecture to reduce ringing artifacts and enhance image resolution. This analysis relies entirely on anonymized imaging data that was previously acquired and does not involve any additional image acquisition or patient contact.

Study Setting:

Study type: Retrospective, single-center, data analysis, observational

Location: Vienna General Hospital

Institution: Department of Neuroradiology and Musculoskeletal Radiology

Data collection: December 2024 - February 2025

All data used in this study were obtained as part of routine clinical fetal MRI examinations and fully anonymized prior to evaluation, in accordance with the General Data Protection Regulation (GDPR) and institutional data protection standards.

Inclusion criteria:

- Fetal MRI scans from pregnant woman between gestational weeks 16 and 37.
- Datasets acquired on a 3T MRI scanner
- T2w TSE sequences acquired in all three orthogonal planes

Exclusion criteria:

- Contraindications for 3T MRI
- Excessive fetal motion leading to non-diagnostic image quality
- Incomplete scans or missing metadata (e.g. resolution, acquisition time).

The retrospective evaluation will focus on determining whether AI-enhanced, post-processed images provide a diagnostic quality that is equal to or better than that of conventional, high resolution T2w TSE sequences.

3.2 MRI System and Acquisition Parameter

3.2.1 MRI

All fetal MRI examinations included in this study were performed using a 3 Tesla Philips Ingenia Elition X scanner (Philips Healthcare, Best, the Netherlands), with software version 5.7.1.4. Imaging was conducted at the Vienna General Hospital (AKH Wien) as part of routine clinical care, in accordance with the institution's standardized fetal imaging protocol.

This protocol involves acquiring T2w TSE sequences in three orthogonal planes (axial, coronal and sagittal), as well as high-resolution (HR) T2w TSE sequences in the same orientations. Patients were positioned feet-first in either the supine or left lateral decubitus orientation using the body coil for signal reception. No maternal sedation was used. Where possible and feasible, breath-hold commands were given to reduce maternal motion during the scan.

AI-based image post-processing was not part of the standard clinical workflow at the time of the study. Instead, it was performed retrospectively on selected T2w TSE sequences using the NGSA research patch.

3.2.2 Sequence Parameter

The table below summarizes the acquisition parameters of the sequences used in this study for image quality comparison.

Table 1: MR acquisition parameters

Parameter	**T2 TSE cs** (Standard)	**T2 AI SR** (AI-enhanced)	**T2 HR cs** (High-Resolution)
Reconstruction / Acceleration	Compressed Sensing (CS)	CS + AI Super Resolution	Compressed Sensing (CS)
Total Scan Time	00:24	00:24	01:01
ACQ Matrix	228 x 202	228 x 202	312 x 281
ACQ Voxel (mm)	1.10 / 1.24 / 2.50	1.10 / 1.24 / 2.50	0.80 / 0.89 / 2.00
REC Voxel (mm)	1.04 / 1.04 / 2.50	0.52 / 0.52 / 2.50	0.71 / 0.71 / 2.00
FOV (mm)	250 / 250 / 72	250 / 250 / 72	250 / 250 / 90

Slices	29	29	45
Slice Thickness (mm)	2.5	2.5	2.0
CS reduction factor	1.5	1.5	2.5
TR (ms)	6102 (shortest)	6102 (shortest)	15207 (shortest)
TE (ms)	200	200	200
NSA	1	1	1
ACQ Mode	Cartesian	Cartesian	Cartesian
Fast Imaging Mode	TSE	TSE	TSE

Legend: FOV = Field of View, TR = Repetition Time, TE = Echo Time, NSA = Number of signal Averages, ACQ = Acquisition, REC = Reconstruction, SR = Super Resolution

The acquisition parameters summarized above were exported directly from the respective exam cards and were used as standardized reference for comparison across all datasets. These parameters provide a basis for evaluating differences in spatial resolution, acquisition time and reconstruction performance between standard and AI-enhanced sequences.

3.3 Dataset Description

After receiving ethical approval from the Ethics Committee of the Medical University of Vienna (EK Nr: 2199/2017, Amendment approved December 2024), a retrospective Dataset was compiled from the institutional Picture Archiving and Communication System (PACS). An initial set of 55 fetal MRI examinations was filtered based on a protocol type and acquisition metadata.

To eligible for inclusion, cases had to meet the following criteria: MRI examinations performed on a 3 T Philips Ingenia Elition X scanner between gestational weeks 16 and 37, with availability of the following three T2w TSE sequences, acquired in the same anatomical orientation (axial, coronal, or sagittal): (1) a compressed-sensing accelerated T2 TSE cs sequence, (2) a retrospectively AI-enhanced T2 TSE AI SR sequence, and (3) a high-resolution reference T2 HR cs.

Exclusion criteria comprised severe fetal motion resulting in non-diagnostic image quality, which was assessed by an experienced pediatric neuroradiologist with specialization in fetal neuroimaging. Further exclusion criteria included incomplete scans or missing acquisition metadata such as voxel dimensions or orientation tags.

After applying these criteria, 37 cases remained and were exported to a MacBook Air (2024) running macOS Sequoia version 15.3.1. Image evaluation and ROI placement were conducted using ITK-SNAP (version 4.2.2, released December 2024) [26]. During the integrity check within ITK Snap, 7 datasets were excluded due to technical incompatibility – primarily caused by undefined image orientation or corrupted header information. The final dataset consisted of 30 fetal MRI datasets suitable for analysis.

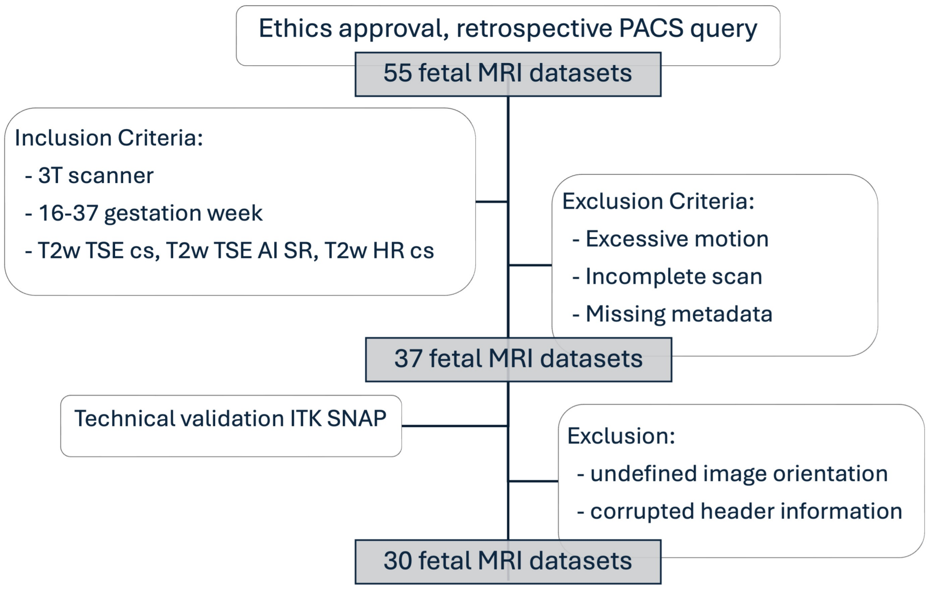

Figure 6: Flowchart of included and evaluated cases

3.3.1 NGSA Patch Documentation

All T2 TSE cs sequences were retrospectively post-processed using the vendor-provided NGSA research patch developed by Philips Healthcare. This includes, as mentioned earlier in the technical background, a reconstruction approach that integrates two deep learning models:

- Adaptive-CS-Net: A CNN trained to denoise and reconstruct undersampled k-space data.

- Precise Image Net: A super resolution GAN that replaces traditional zero-filling and reduces ringing artifacts through learned frequency-based priors.

The NGSA patch was applied using the Philips Research Viewer. To initiate the post-processing, the T2 TSE cs sequences were loaded and specific reconstruction parameters were modified. In particular, the ‘Recon Voxel Size’ parameter was set to “AI” and the enhancement strength was selected as “AI SR strong”. This triggered the application of the NGSA patch and initiated AI-based denoising and super resolution enhancement.

Once the reconstruction was completed, the AI-enhanced images were saved separately for further analysis. Based on subjective visual inspection, the post processed images showed improved noise suppression, anatomical edge definition, and overall image sharpness compared to the original CS images.

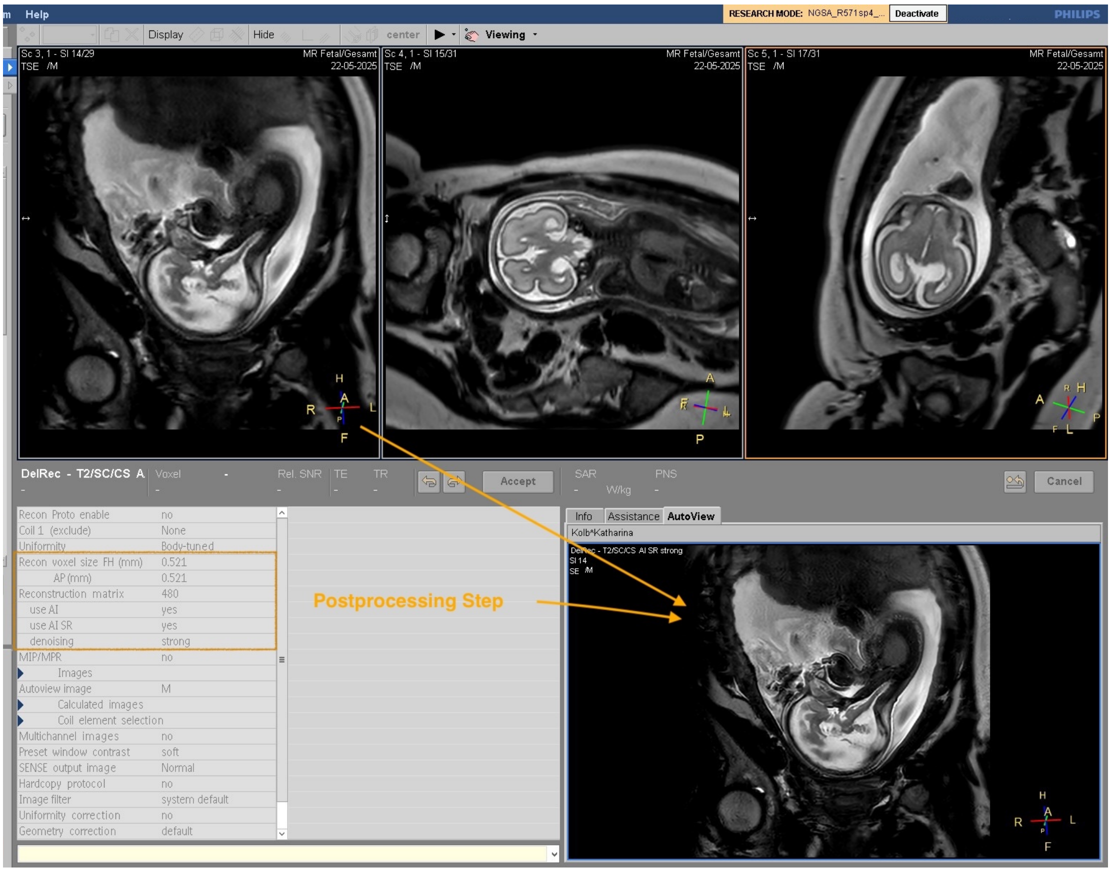

Figure 7: Example of NGSA enhanced reconstruction. The resulting image is shown in the bottom right corner.

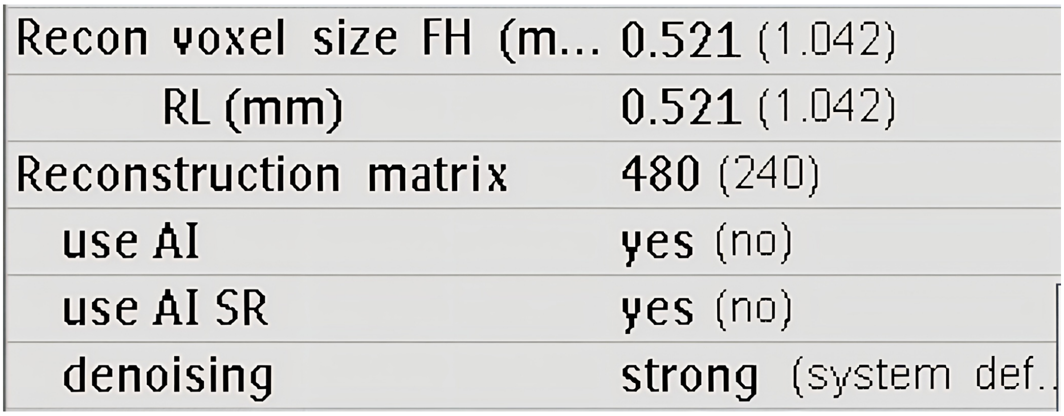

Figure 8: Zoomed view of reconstruction parameter settings

3.4 Quantitative Image Analysis

3.4.1 SNR & CNR

For all 30 included cases, four manually defined regions of interest (ROIs) were placed using ITK Snap (version 4.2.2). ROI measurements were performed on all three T2w sequences (T2 TSE cs, T2 AI SR, T2 HR cs) per case. In 25 of the 30 cases, ROIs were positioned in the sagittal midline plane, in the remaining 5 cases, measurements were carried out in the axial plane at the level of the pons. The axial orientation was selected in these five cases due to pronounced motion artifacts in the sagittal plane, which would have otherwise led to case exclusion. This approach ensured that all 30 datasets could be included, maintaining a robust sample size for statistical analysis. The ROIs were consistently defined as follows:

- ROI1: Pons (brain parenchyma)
- ROI2: Air (as background noise)
- ROI3: Cerebro Spinal Fluid (CSF, e.g. ventricle or cisterna magna)
- ROI4: Maternal muscle (e.g. Musculus Illiacus)

For each defined ROI, the mean signal intensity and corresponding standard deviation were extracted and documented in Microsoft Excel. These values served as the basis for calculating two key quantitative metrics relevant for image quality assessment:

The signal-to-noise ratio (SNR) was calculated using an apparent SNR approach, which is commonly applied in clinical MRI quality assessments [21], [25]. Apparent SNR refers to an estimation based on image intensity values, where the mean signal of a structure (e.g. pons) is divided by the standard deviation of a noise reference region (e.g. air). This method allows for robust and reproducible comparison of image quality across different sequences without requiring raw data access.

Similarly, the contrast-to-noise ratio (CNR) was calculated to evaluate the distinguishability between different tissue types, specifically between the pons and CFS. Following the apparent CNR approach described by Bischoff et al. (2023), the CNR was derived by calculating the absolute difference between the mean signal intensities of the two ROIs (pons and CFS), divided by the standard deviation of noise. This method provides a straightforward and reproducible measure of image contrast that does not require access to raw data.

$$SNR = \frac{\mu\mathrm{ROI1}}{\sigma\mathrm{ROI2}}$$

$$CNR = \frac{\mu\mathrm{ROI1} - \mu\mathrm{ROI3}}{\sigma\mathrm{ROI2}}$$

$\mu = \mathrm{Mean\ Value}, \sigma = \mathrm{Standard\ Deviation}, \mathrm{Formula\ also\ used\ by\ Bischoff\ et\ al.}\ (2023)$

The following figure shows examples of ROI placement in ITK-SNAP. It shows measurements taken in both the sagittal midline and the axial plane, along with visual documentation of the assigned ROI positions and the corresponding signal intensity values for each of the four target regions. Additionally, the table below provides an example of the calculations.

Table 2: SNR & CNR Calculation

ID	Sequence						Metrics	
		Label ID	Nr. of Voxels	Volume (mm^3)	Image (mean, µ)	Image (stdev, σ)	SNR	CNR
1	T2 TSE cs	1	24	65,1042	311,542	14,92	53,0250861	62,2513986
		2	24	65,1042	8,45833	5,87537		
		3	24	65,1042	8,45833	9,98468		
		4	24	65,1042	18,25	6,26411		
1	T2 AI SR	1	24	16,276	313,958	21,031	49,1799632	56,979633
		2	24	16,276	8,16667	6,38386		
		3	24	16,276	677,708	5,84941		
		4	24	16,276	5,33333	5,66198		
1	T2 HR cs	1	24	24,2123	128,583	12,7003	30,988485	33,4187276
		2	24	24,2123	6,5	4,14938		
		3	24	24,2123	267,25	9,61543		
		4	24	24,2123	8,95833	4,73175		

Legend: Label 1 – ROI Pons, Label 2 – ROI Air, Label 3 – Liquor, Label 4 – maternal muscle, stdev = Standard deviation

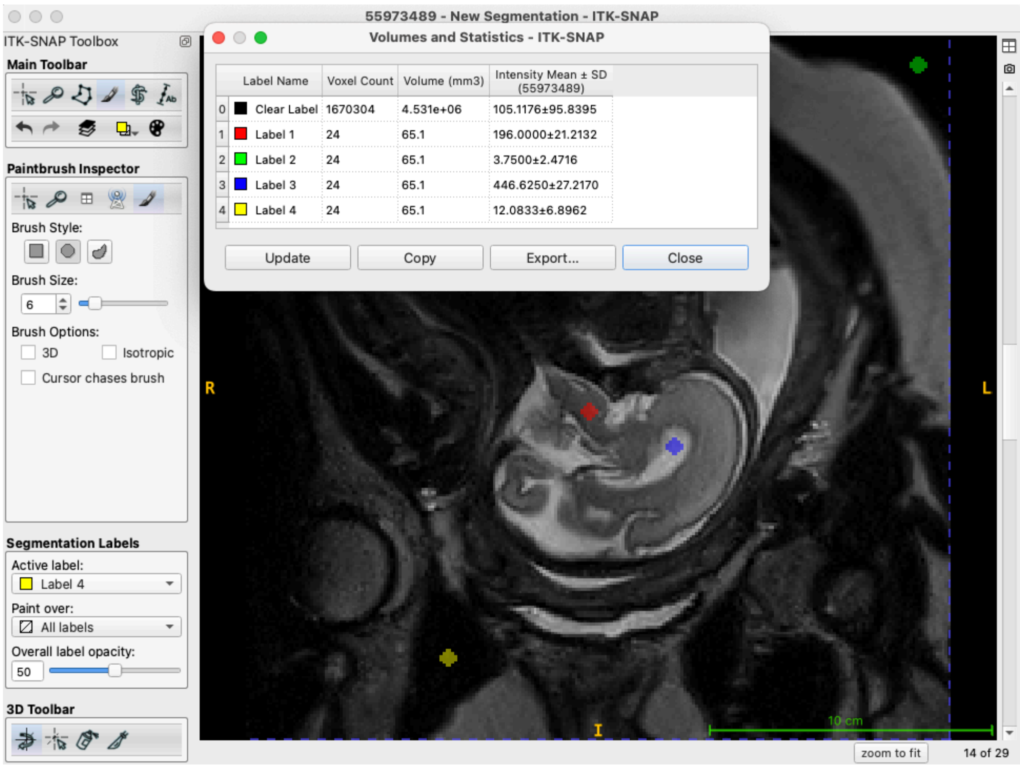

Figure 9: Example of ROI placement in ITK Snap, sagittal midline

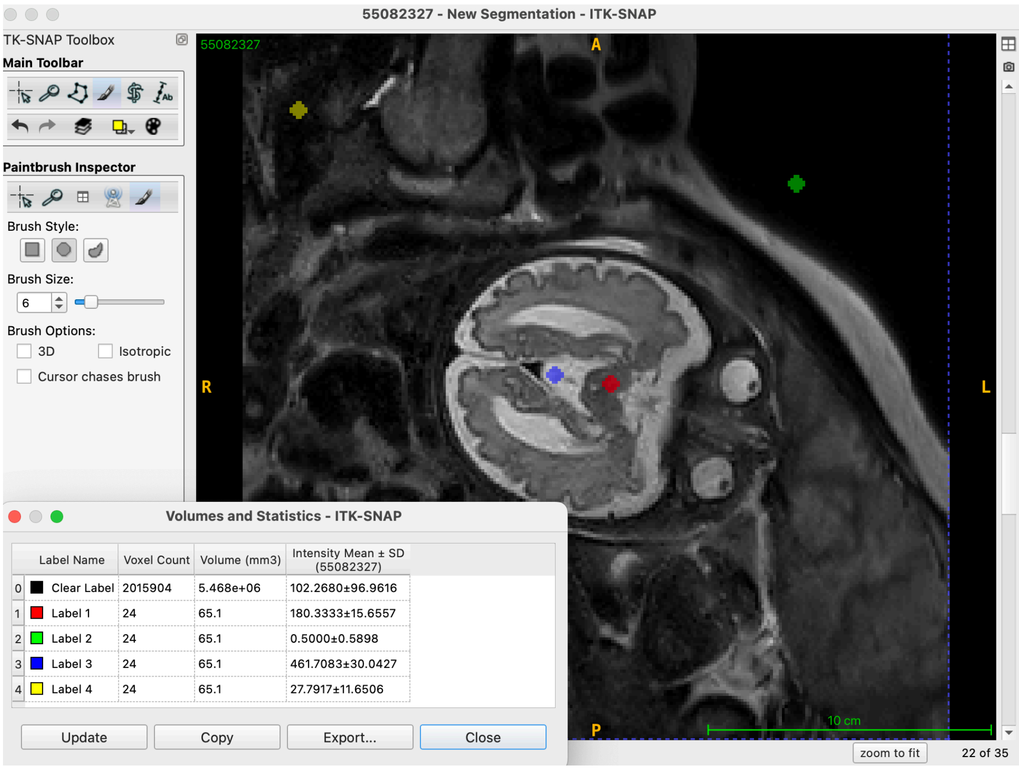

Figure 10: Example of ROI placement in ITK-SNAP, axial plane

3.4.2 ERD

To further assess image sharpness, the edge-rise distance (ERD) was quantified using ImageJ (version 2.16.0). For each dataset, a linear intensity profile was manually drawn across a clearly defined anatomical boundary, specifically from the anterior surface of the pons to the outer margin of the cerebellum. To ensure anatomical consistency and reduce measurement variability, all profiles were placed in the sagittal plane.

The ERD analysis was based on the same cohort of 30 datasets that was used for the preceding SNR and CNR evaluations. However, due to technical issues encountered when extracting profile data in ImageJ, 10 cases had to be excluded. Consequently, a total of 20 datasets remained for the quantitative assessment of ERD, all of which were processed under a uniform sagittal orientation.

Intensity gradients along the defined profiles were visualized using the “Plot Profile” function in ImageJ: ERD was defined as the physical distance (in millimeters) between the 10% and 90% intensity levels along the ascending edge of the signal curve. This methodological approach was conceptually aligned with the framework described by Bischoff et al. (2023), who applied a comparable 10-90% edge model to characterize image sharpness in clinical MRI. [21] An example illustration of this procedure is shown in Figure 11, depicting the manually drawn line, extending from the outer margin of the pons to the most lateral point of the cerebellum, along with the corresponding intensity curve used for ERD derivation.

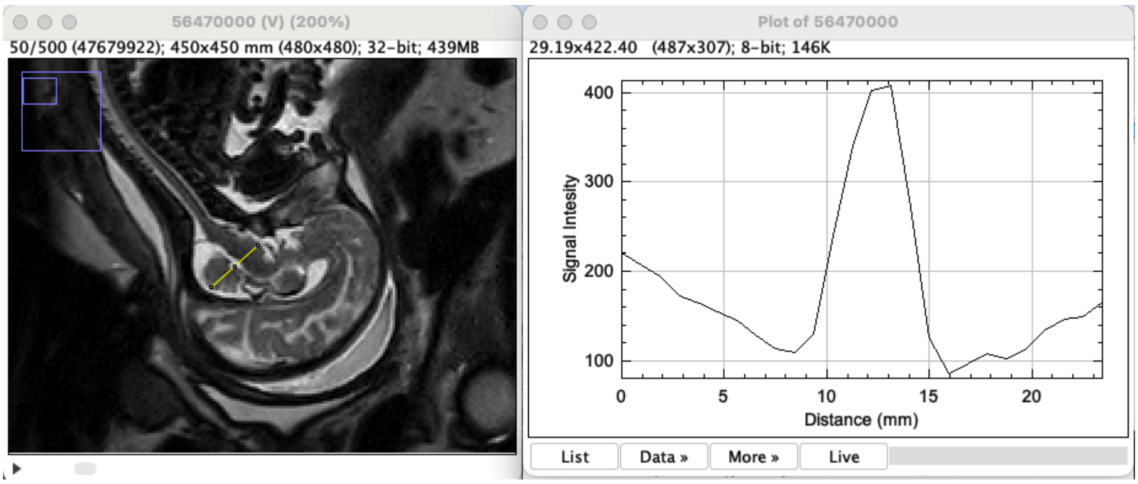

Figure 11: Example of ERD measurement in ImageJ

Following profile extraction, the signal intensity data were exported from ImageJ as comma-separated value (.csv) files and imported into Microsoft Excel for further analysis and documentation. The following computational steps were applied:

1. Identification of maximum and minimum grey values within the intensity profile.
2. Determination of the 10% and 90% signal levels, defined as:

$$10\% = Min + 0{,}1 * (Max - Min)$$

$$90\% = Min + 0{,}9 * (Max - Min)$$

3. Interpolation of the position (in millimetres) along the distance axis where the intensity curve crosses the 10% and 90% thresholds. These positions where determines by direct tabular lookup.

4. ERD was then calculated as the difference between the two corresponding positions:

$$ERD = \chi 90\% - \chi 10\%$$

Table 3: ERD Calculation

Distance (mm)	Gray Value	ERD	
0.0000	526.500	627.167	Min
0.9375	393.667	300.969	Max
1.8750	353.167	333.588,8	10 %
2.8125	337.062	594.547,2	90 %
3.7500	336.958	13,1250	Index 10%
4.6875	352.370	7,5000	Index 90%
5.6250	379.000	x(10%) - x(90%)	ERD
6.5625	528.130		
7.5000	627.167	**5,6250**	**ERD mm**
8.4375	471.130		
9.3750	315.958		
10.3125	300.969		

11.2500	325.708		
12.1875	336.573		
13.1250	334.292		
14.0625	340.375		
15.0000	369.333		

This procedure was applied consistently across all valid cases. In analogy to the SNR and CNR workflow, the ERD data were structured for statistical evaluation and comparison across sequences.

3.5 Qualitative Image Assessment

To complement the quantitative metrics, a qualitative image assessment was conducted to evaluate the diagnostic quality and clinical usability of the different sequences. The evaluation was performed independently by two-board-certified radiologists, each with extensive experience (over 5 and 26 years, respectively) in the specialized field of fetal and pediatric neuroimaging.

3.5.1 Reading Session and Blinding

A representative subset of 11 cases was randomly selected from the final cohort. For each case, the three sequences (T2 TSE cs, T2 AISR, and T2 HR cs) were prepared for a blinded reading session. To prevent any bias, the evaluation was conducted on a fully blinded and randomized manner.

All sequences were anonymized and generically labeled as ‘Picture 1’, ‘Picture 2’ and ‘Picture 3’. The assignment of these labels to the actual sequences was randomized for each of the 11 cases. Furthermore, the order in which the cases were presented to each radiologist was also randomized. A key file, concealed from the readers, documented the randomization scheme for later analysis.

All images were reviewed on a clinical Picture Achieving and Communication System (PACS) workstation, which provided the standard diagnostic environment and allowed for typical interactions such as windowing, panning, and zooming. The entire reading session for each radiologist was designed to last approximately 45 minutes.

3.5.2 Evaluation Criteria and Scoring

The radiologists used a structured online questionnaire, created in Microsoft Forms, to score the images based on five distinct criteria based on Bischoff et al. (2023). A 4-point Likert scale was employed for each criterion to ensure a definitive assessment by avoiding a neutral midpoint. An optional free-text comment field was also included for each case, allowing the radiologists to provide additional remarks.

The five evaluation criteria and their respective 4-point scales were defined as:

1. Artifacts: The level of image degradation due to artifacts.
   - 4 = Non-diagnostic (Massive artifacts, not interpretable)
   - 3 = Moderate distortion (Noticeable but acceptable)
   - 2 = Minimal distortion (Noticeable but acceptable)
   - 1 = No distortion (Artifact-free image)
2. Image Sharpness: The clarity and delineation of anatomical structures.
   - 4 = Poor delineation (Structures not distinguishable)
   - 3 = Moderate delineation (Unclear but identifiable)
   - 2 = Minimal obstruction (Nearly perfectly sharp)
   - 1 = Perfectly sharp (Excellent anatomic clarity)
3. Tissue Contrast: the ability to differentiate between adjacent tissue types.
   - 4 = No contrast (Tissue borders not visible)
   - 3 = Moderate contrast ((very) low differentiation)
   - 2 = Good contrast (Clear tissue distinction)
   - 1 = Excellent contrast (Optimal differentiation)
4. Overall Image Quality: A global assessment of the image’s diagnostic usability.
   - 4 = Not usable (non-diagnostic image)
   - 3 = Acceptable (Usable for interpretation)
   - 2 = Good (reliable diagnostic quality)
   - 1 = Excellent (Optimal diagnostic image)
5. Diagnostic Confidence: The level of certainty in making a diagnostic assessment from the image.
   - 1 = No confidence (Diagnosis not possible)
   - 2 = Moderate confidence (Highly uncertain)
   - 3 = Good confidence (Reliable interpretation)
   - 4 = High confidence (Fully confident diagnosis)

The collected Likert scale data and qualitative comments form the basis for the subsequent statistical analysis and the comparison of sequence performance.

3.6 Statistical Analysis

Statistical analysis was performed using GNU PSPP (Version1.6.2). For all statistical tests, a p-value of less than 0.05 was considered to indicate a significant difference.

3.6.1 Quantitative Data Analysis

Descriptive statistics were computed for the quantitative metrics, including SNR, CNR and ERD. Given the non-normal distribution of the data and the presence of outliers, non-parametric summary measures were used. Specifically, the median and interquartile range (IQR) were reported for each metric and imaging sequence (T2 TSE cs, T2 TSE AISR, and T2 HR cs). This approach ensures a robust and distribution-independent description of central tendency and variability.

To evaluate differences between the sequences, the Wilcoxon signed-rank test was applied for all pairwise comparisons. This non-parametric test for related samples was chosen to maintain methodological consistency across all analyses, regardless of the underlying data distribution.

The following comparisons were performed to directly address the core research questions:

- T2 TSE cs vs. T2 TSE AI SR: To assess the extent to which AI-based post-processing improves image quality compared to the accelerated standard sequence (addresses Research Question 1).
- T2 TSE AI SR vs. T2 HR cs: To evaluate whether the AI-enhanced sequence can achieve image quality comparable to the high-resolution reference standard (addresses Research Question 2).
- T2 TSE cs vs. T2 HR cs: To serve a control comparison, establishing the expected difference in quality between fast low-resolution and high-resolution sequences.

SNR and CNR analyses were conducted across the full dataset of 30 cases. Due to technical constraints in edge profile extraction, ERD was evaluated in a subset of 20 cases. Additionally, the mean scan times for the sequences were reported to address Research Question 3, which investigates whether AI post-processing allows for shorter acquisition times without compromising image quality.

3.6.2 Qualitative Data Analysis

Qualitative assessments from two expert readers were averaged for each case and evaluation criterion (Artifacts, Image Sharpness, Tissue Contrast, Overall Image Quality, and Diagnostic Confidence). The resulting ratings were summarized using median and IQR values.

To detect overall differences in subjective image quality the three sequences, the Friedman test was applied for each evaluation criterion. In cases where the Friedman test revealed significant group effects ($p<0.05$), pairwise Wilcoxon signed-rank tests were conducted as post-hoc analyses. This step identified which sequences differed from each other in terms of diagnostic impression and perception of image quality, contributing directly to answering Research Question 1,2, and 4.

4 Evaluation Results

The chapter presents the quantitative and qualitative results of the image quality evaluation across the three examined MRI sequences: T2 TSE cs (Conventional, Con), T2 TSE AI SR (AI-enhanced, AI), and T2 HR cs (High-Resolution, HR). The analysis addresses objective image quality metrics (SNR, CNR, ERD) and expert-based qualitative assessments to evaluate the efficacy of the AI-based post-processing approach.

4.1 Quantitative Image Quality Metrics

The quantitative analysis was performed on 30 datasets for SNR and CNR, and on a sub-cohort of 20 datasets for ERD. The results, including median values, interquartile ranges (IQR), and p-values from pairwise Wilcoxon signed-rank tests, are summarized in Table 4 and visualized in the boxplots in Figure 15.

Table 4: Quantitative Image Quality Metric Results

Quantitative Image Evaluation Metrics Based on Wilcoxon Test Results				
Metric	**T2 AI SR** (Median [IQR])	**T2 TSE cs** (Median [IQR])	**T2 HR cs** (Median [IQR])	**Wilcoxon** (p-Value, significant difference)
SNR (n=30)	57.75 [110.53]	63.22 [109.70]	39.24 [35.99]	HR vs. Con (p=0.004**) HR vs. AI (p=0.020*)
CNR (n=30)	70.03 [119.32]	87.62 [119.65]	67.32 [65.67]	not significant
ERD (mm) (n=20)	10.29 [7.43]	5.28 [3.24]	7.49 [4.39]	AI vs. Con (p =0.005**) AI vs. HR (p <0.018*)
Average Scan time (s)	34.5 [24-45s]	34.5 [24-45s]	56 [50-62s]	HR > Con / AI

4.1.1 SNR

The median SNR was highest in the T2 TSE cs sequence (63.22), followed by T2 AI SR (57.75). The T2 HR cs sequence had the lowest median SNR (39.24). A Wilcoxon signed-rank test confirmed that both the conventional T2 TSE cs ($p=0.004$) and the AI-enhanced T2 AI SR ($p=0.020$) sequences yielded significantly higher SNR values than the high-resolution T2 HR cs sequence. There was no statistically significant difference in SNR between the T2 TSE cs and T2 AI SR sequences.

4.1.2 CNR

CNR values followed a similar trend with the highest median observed in the T2 TSE cs sequence (87.62), followed by the T2 AI SR (70.03) and the T2 HR cs (67.32). However, the Wilcoxon signed-rank tests revealed no statistically significant differences in CNR among any of the three sequences.

4.1.3 ERD

Edge Rise Distance (ERD) was calculated to quantify image sharpness, defined as the distance between 10% and 90% signal intensity along a line perpendicular to a structural edge. For image sharpness, a lower ERD value indicates a better result.

The median ERD values were as follows: T2 TSE cs with 5.28 mm, T2 HR cs 7.49 mm and T2 AI SR 10.29 mm. Pairwise Wilcoxon signed-rank tests revealed statistically significant differences between T2 AI SR < T2 TSE cs ($p=0.005$) and T2 AI SR < T2 HR cs ($p=0.018$).

Line profile visualizations from one representative case are shown in Figure 14 and 15, comparing the intensity transitions of the three sequences. Profiles were measured under standardized conditions (identical vector placement, signal range between 10% and 90%). Although the AI-enhanced sequences showed sharper visual transitions, the computed ERD was longer. The observation is likely influenced by the super-resolution reconstruction method, which increases matrix resolution through interpolation and may smooth signal gradients across more voxels.

A more detailed discussion of this finding is provided in Chapter 6.

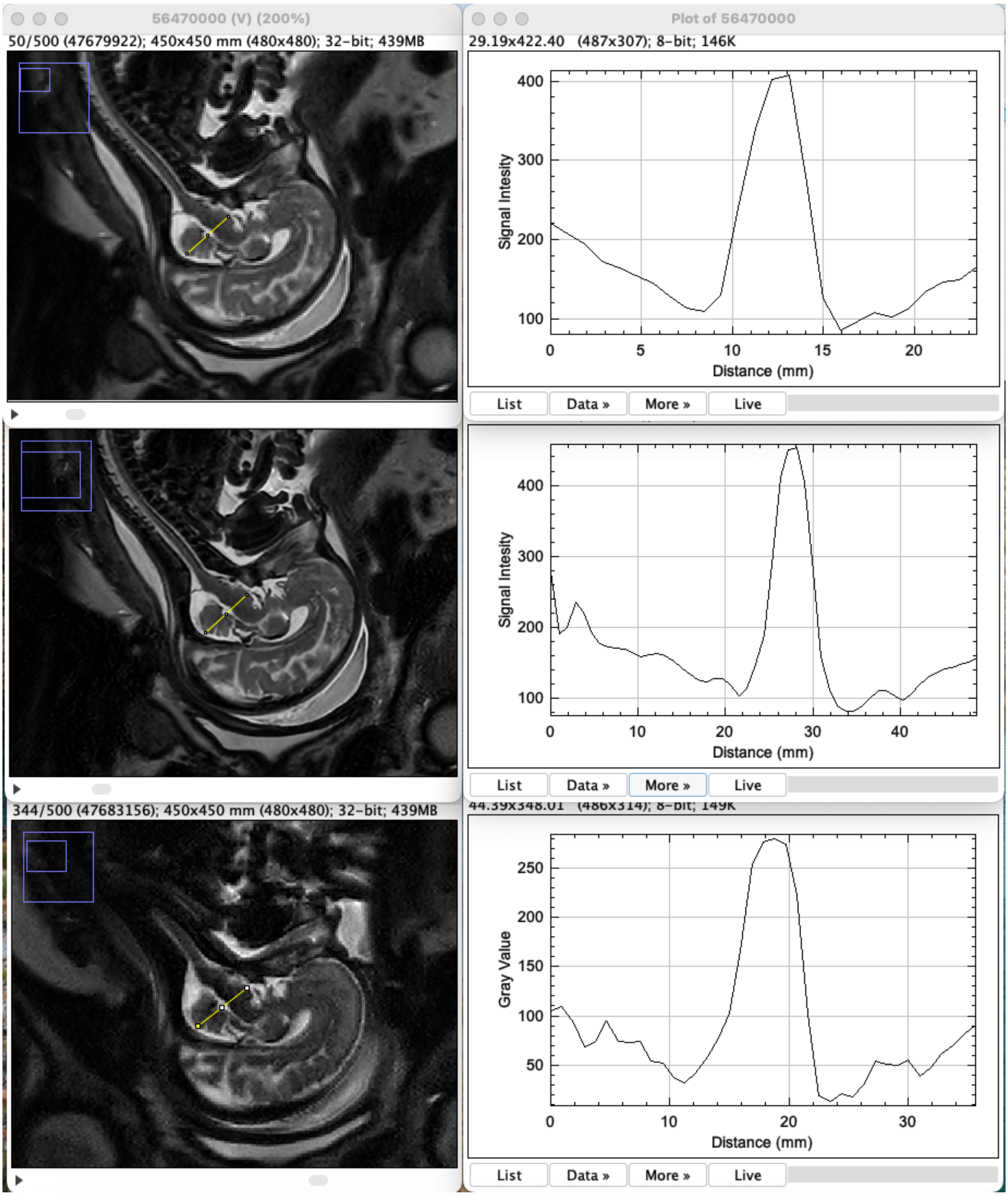

Figure 12: ERD Visual Comparison, Picture 1 at the top represents the T2 TSE cs, Picture 2 in the middle T2 AISR, and Picture 3 at the bottom T2 HR cs

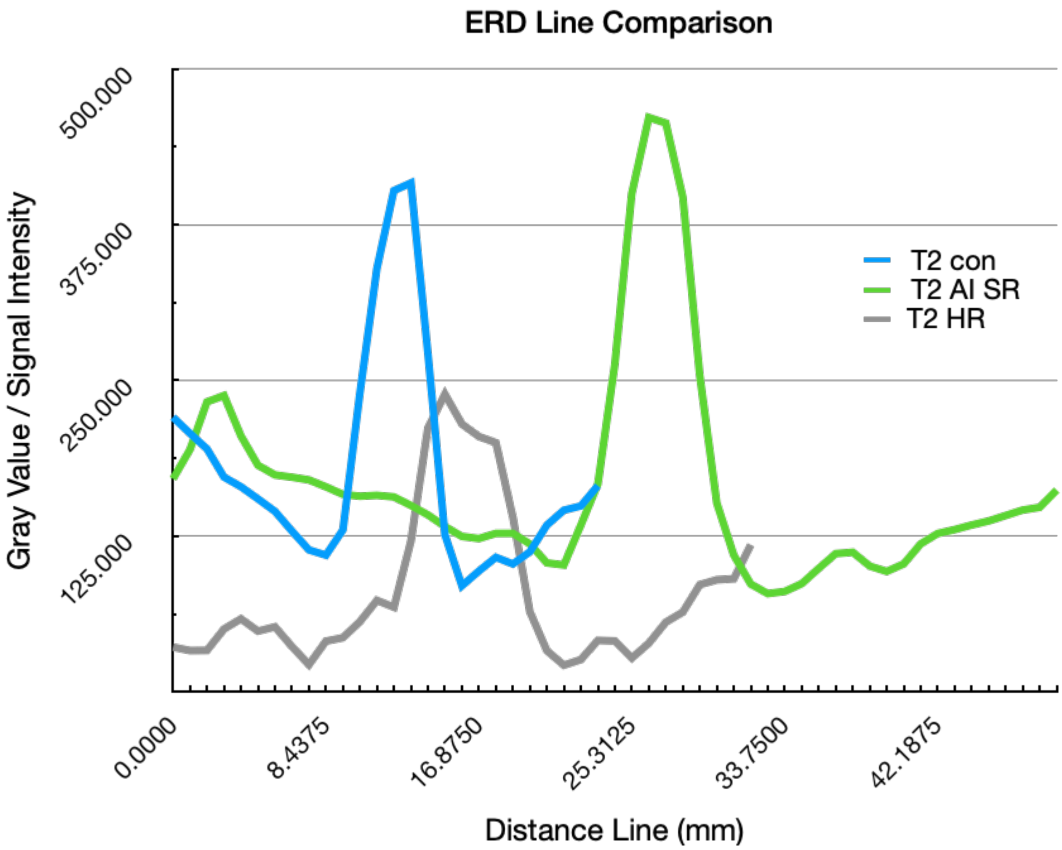

Figure 13: ERD Line Comparison

The Boxplots in Figure 15 provide a summary of the distributions of all quantitative image quality metrics (SNR, CNR, ERD) across the three sequences (T2 TSE cs = Con, T2 AI SR = AI, T2 HR cs = HR). Apparent SNR values were highest for the conventional sequence, followed by AI-enhanced and HR sequences. CNR values showed broad variability across all groups, without statistically significant differences. For ERD, the AI-enhanced sequence showed the highest median value broader spread, while the conventional sequence had the lowest values. Statistical significance between paired groups is indicated directly on the figure, based on Wilcoxon signed-rank test results ($p<0.05$). These comparisons complement the tabular summary of results and illustrate key distributional trends.

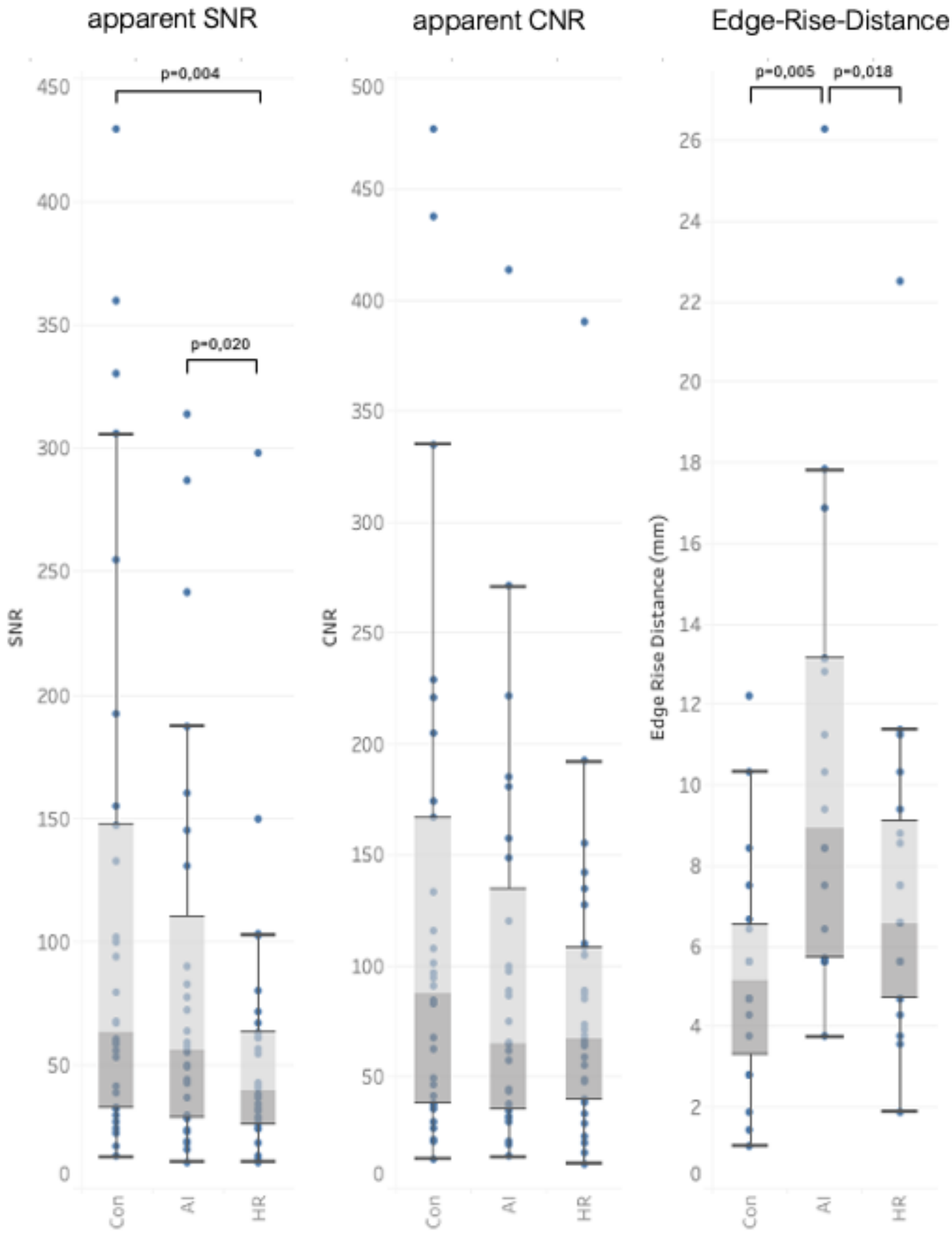

Figure 14: Boxplot of apparent SNR, CNR and ERD for all three sequences (T2 TSE = Con, T2 AI SR = AI, T2 HR cs = HR). Gray boxes represent IQR with median lines, whiskers extend to 1.5x IQR, and individual points represents data from each case. *p-values* indicate differences between groups (*Wilcoxon*, two sided, α=0.05)

4.2 Qualitative Image Quality Assessment

In addition to the quantitative metrics, the diagnostic quality was also assessed qualitatively based on five predefined categories: artifacts, image sharpness, tissue contrast, overall image quality, and diagnostic confidence. Two experienced readers independently rated each image using an inverted 4-point Likert scale (1 =excellent, 4 = non-diagnostic). For statistical analysis, the average rate per case was calculated across both readers. Table 5 provides an overview of the median score and interquartile ranges for each sequence and criterion.

To evaluate statistically significant differences between the sequences, Friedman test was performed for each criterion. Where the Friedman test revealed a significant overall difference (p<0.05), pairwise Wilcoxon signed-rank test was applied.

Table 5: Qualitative Image Evaluation Results

Qualitative Image Evaluation Ratings Based on the Average Score Assigned by Two Readers					
Category	**T2 AI SR** (Median [IQR])	**T2 TSE c**s (Median [IQR])	**T2 HR cs** (Median [IQR])	**Friedmann** (p-Value)	**Wilcoxon** (p-Value, significant difference)
Artifacts	1 [0,75]	1 [1]	1,5 [1]	0,32	not significant
Image Sharpness	1 [0]	2 [0]	2 [1]	<0,001***	AI vs. Con (p <0,001***) AI vs. HR (p=0,001***)
Tissue Contrast	1 [0,75]	2 [1]	3 [1,5]	<0,001***	AI vs. Con (p <0,001***) AI vs. HR (p <0,001***) Con vs. HR (p=0,008**)
Overall Image Quality	2 [1]	2 [0]	2 [0]	<0,001***	AI vs. Con (p <0,001***) AI vs. HR (p <0,001***) Con vs. HR (p=0,378)
Diagnostic Confidence	1 [0]	2 [0]	2 [0]	<0,001***	AI vs. Con (p <0,001***) AI vs. HR (p=0,001***) Con vs. HR (p=0,705)

The T2 AI SR sequence showed significantly better ratings in four of five qualitative categories. For image sharpness, tissue contrast, overall image quality and diagnostic confidence, the AI-enhanced sequence outperformed both the conventional T2 TSE cs and T2 HR cs (all p<0.001).

These results are visualized in Figure 16, which presents stacked bar charts showing the distribution of Likert ratings across all sequences.

- **Artifacts:** All three sequences received mostly favourable ratings (Likert 1-2). There was no clear advantage for any sequence.
- **Image Sharpness:** The T2 AI SR sequence received the highest proportion of Likert 1 ratings ("perfectly sharp"), with almost no score ≥3. In contrast, both T2 TSE cs and T2 HR cs showed a broader distribution, including more moderate and poor ratings.
- **Tissue Contrast**: The T2 AI SR demonstrated the most favourable distribution, with majority of cases rated as "excellent" or "good". The T2 HR cs sequence received the highest number of poor ratings (Likert 3-4), indicating difficulties in tissue differentiation.
- **Overall Image Quality**: T2 AI SR had a markedly higher proportion of "excellent" and "good" ratings compared to the other sequences. While T2 TSE cs performed moderately well, the T2 HR cs group included several cases rated as "bad" or "non-diagnostic".
- **Diagnostic Confidence**: The T2 AI SR sequence again received the best ratings, with nearly all cases rated 1 or 2, while the other sequences showed a wider spread, including multiple ratings of "moderate confidence".

These findings emphasize the improved subjective image impression of the AI-reconstructed sequence across multiple diagnostic dimensions.

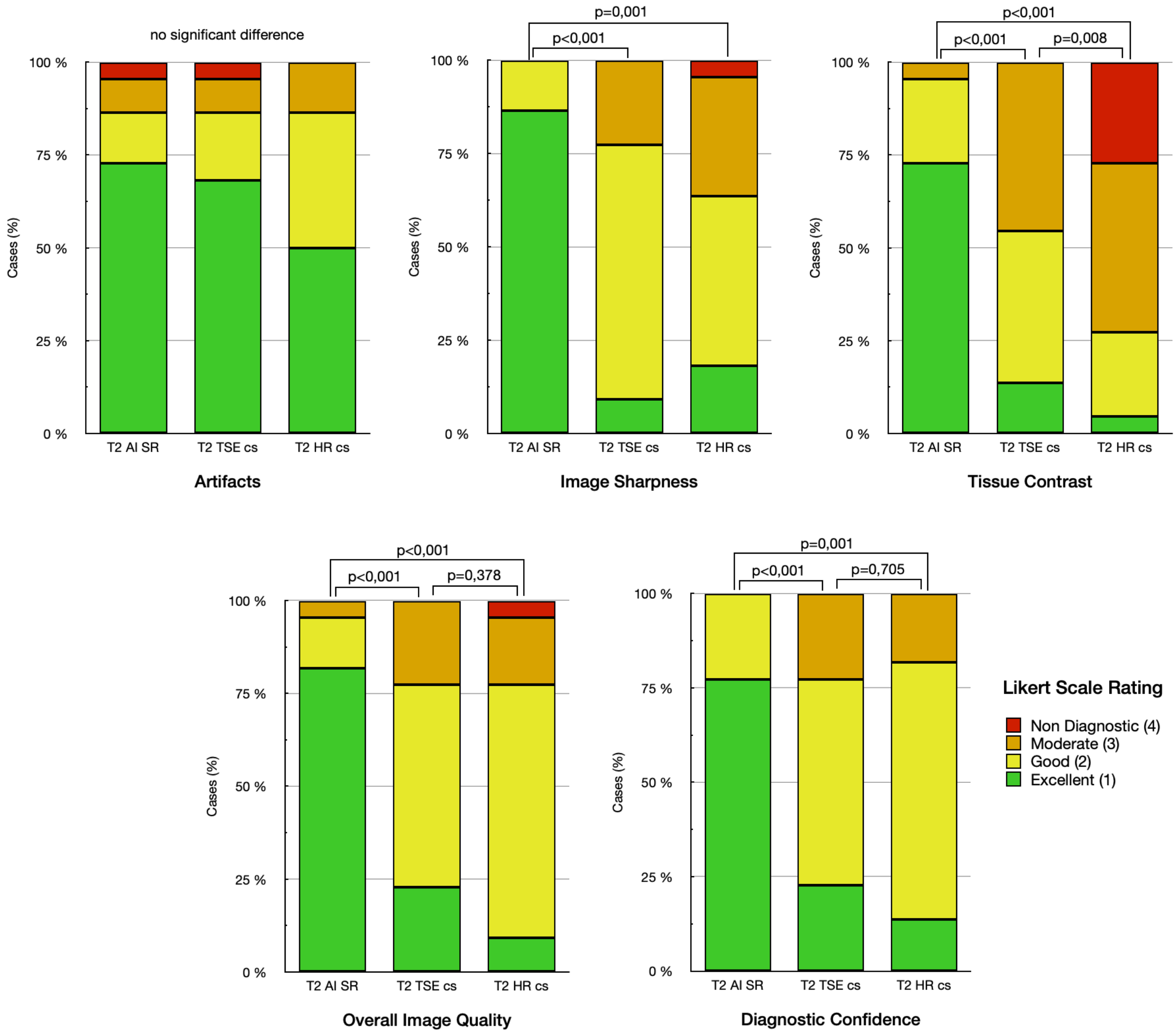

Figure 15: Stacked bar charts illustrating the distribution of Likert scores (1= excellent to 4 = non-diagnostic) across all image quality categories and sequences. *p-values* are based on Wilcoxon signed-rank tests.

5 Discussion

This study evaluated the impact of an AI-based super-resolution post-processing algorithm on accelerated T2-weighted fetal MRI sequences. The principal finding is that the AI-enhanced sequence (T2 AI SR) was perceived by expert radiologists to offer superior diagnostic quality compared to both the conventional accelerated sequence (T2 TSE cs) and the time-intensive high-resolution sequence (T2 HR cs). While quantitative metrics for SNR and CNR showed no significant improvement, the qualitative evaluation demonstrated consistent and statistically significant improvements in perceived image sharpness, tissue contrast, and diagnostic confidence.

These benefits were achieved without extending the acquisition time (approximately 34.5 seconds), which was identical to the conventional sequence. It is important to note that this figure excludes the offline GPU-based reconstruction time. Nevertheless, the findings demonstrate that AI-based post-processing can substantially enhance diagnostic image quality without the trade-off of longer scan durations. Compared to the high-resolution sequence (56 seconds), the AI-based approach delivered superior perceived quality in 38% less time, highlighting its significant potential for clinical application.

The following sections address each of the four research questions individually, drawing on quantitative results and qualitative insights to form an integrated interpretation of the findings.

5.1 Interpretation of Research Questions

1. To what extent can AI-based post processing improve image quality of ultrafast acquired fetal MRI images?

The results present a multifaceted picture. From a quantitative perspective, the AI algorithm did not yield statistically significant improvements in SNR or CNR over the conventional sequence. The high variance observed in these metrics is likely attributable to confounding factors inherent in retrospective fetal imaging, such as differing fetal positions and variable degrees motion. In addition, the wide range of gestational ages (16-34 weeks) the variability of amount of amniotic fluid and the respective maternal body mass indexes may also have contributed to image quality. Figure 17 illustrates this point by presenting three cases with different signal profiles, highlighting how biological and positional variations impacts signal-based measurements.

More revealingly, the quantitative sharpness metric ERD yielded paradoxical results: the AI-enhanced images produced significantly higher (i.e. worse) ERD values. This counterintuitive finding appears to contradict the clear qualitative preference of the expert readers. It can be explained by the technical nature of the super-resolution algorithm, which interpolates voxel data to increase matrix size. While this process creates visually sharper edges, it mathematically distributes the signal transition across more voxels, thereby broadening the calculated edge-rise zone. This highlights a critical limitation of traditional quantitative metrics like ERD in evaluating AI-reconstructed images, reinforcing the indispensable role of perceptually oriented expert assessment.

In stark contrast, the qualitative evaluation unequivocally demonstrated the superiority of the AI sequence. Expert radiologists rated it significantly higher in image sharpness, tissue contrast, overall image quality and diagnostic confidence. This suggests the primary strength of this AI tool lies in enhancing the perceptually salient features that are most relevant for clinical diagnosis. An explanation may be the experience of the readers who are trained to recognize image artifacts in fetal MR images.

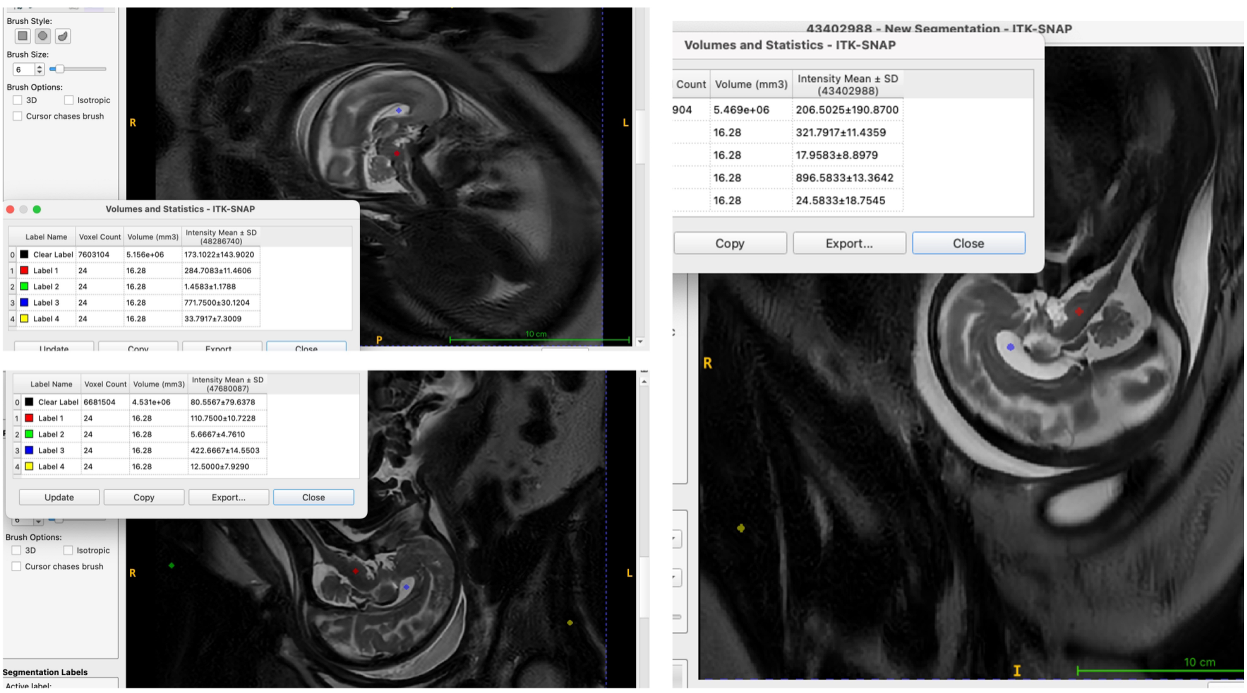

Figure 16: Differences in signal measurements between cases

2. How does AI-enhanced low-resolution MRI compare to standard high-resolution sequences in diagnostic value?

Despite requiring 38% less acquisition time, the AI-enhanced images were qualitatively rated as superior to the high-resolution reference in nearly all diagnostic dimensions.

Quantitatively, the AI sequence demonstrated a significantly higher SNR than the HR sequence, though no significant difference was found for CNR. This aligns with the qualitative feedback, where radiologists noted that the HR sequence sometimes suffered from visible noise and weaker tissue contrast.

Qualitatively, the AI sequence consistently outperformed the HR sequence. Radiologists emphasized its superior ability to delineate fine anatomical structures, such as the infundibulum and optic chiasm, with greater clarity and confidence. While the HR sequence was also diagnostic, the AI images provided more convincing visual contrast and sharper anatomical borders. A key observation was that the HR sequence appeared more susceptible to motion-induced blurring, whereas the AI algorithm sometimes amplified minor motion by creating overly sharp or enhanced edges. Despite the nuances, the overall clinical judgment favored the superior diagnostic utility of the faster AI-enhanced sequence.

3. Can AI-based postprocessing facilitate shorter acquisition times while maintaining or improving image quality?

The AI-enhanced sequence was acquired in an average scan time of 34.5 seconds, matching the duration of the conventional sequence and representing a 38% reduction to the high-resolution sequence (56 seconds). These shorter acquisition times are of relevance in fetal MRI, where minimizing scan duration is critical to reducing motion artifacts and maternal discomfort.

However, it is important to contextualize this result within the technical setup of the study. The AI-enhanced reconstructions were not generated in real time but were computed offline using a research prototype. The actual processing time was not included in the reported scan duration and would vary depending on available Graphic Processing Units (GPU) resources. Consequently, the integration of such algorithms into clinical practice depends not only on their diagnostic efficacy but also on hardware capabilities and software implementation in routine workflow.

Nonetheless, the findings demonstrate that AI-based post-processing enables a substantial improvement in diagnostic image quality while maintaining a short acquisition time. Although the acquisition itself does not shorten further, the ability to achieve superior image quality without requiring a longer scan duration represents a significant advancement.

4. How do expert radiologists evaluate the diagnostic usability of AI-enhanced fetal MRI images?

Expert radiologists consistently rated the AI-enhanced images highest across nearly all qualitative parameters, including image sharpness, tissue contrast, overall image quality, and diagnostic confidence. The AI sequence was frequently described as offering particularly clear and well-defined anatomical borders, with structures such as the infundibulum, optic chiasm, and aqueduct being more easily and confidently delineated compared to the conventional and high-resolution sequences.

However, certain limitations were also noted in the qualitative feedback. One recurring observation was that the AI-enhanced sequence tended to amplify motion-related artifacts, occasionally leading to over-sharpened or over-enhanced edges in regions affected by fetal movement. In contrast, the conventional T2 TSE sequence was perceived as more robust in the presence of motion and less prone to such exaggerations. This overrepresentation of subtle movement artifacts in the AI sequence was occasionally regarded as distracting by the raters. This tendency to enhance subtle noise or motion can be interpreted as a low-level form of “algorithmic hallucination”, where the model prioritizes sharpness over fidelity. It highlights the necessity for radiologists to remain critically aware that AI-generated details do not always represent anatomical truth.

Overall, while the radiologists were clearly impressed by the sharpness and visual detail offered by the AI-based reconstructions, they emphasized the need for cautious interpretation. Especially in fetal imaging, where anatomical structures are small and still developing the appearance of “newly visible” features should be carefully validated before being accepted as clinically relevant. The reviewers agreed that continued observation, further validation, and increased clinical experience are essential for fully integrating AI-enhanced imaging into routine fetal MRI assessment.

5.2 Comparison to Prior Work

The findings of this study align with a growing body of literature demonstrating the value of deep learning-based reconstruction in various clinical contexts. The consistent subjective preference for the AI-enhanced images mirrors the results from Bischoff et al. (2023), who reported significantly improved qualitative score for image sharpness and overall quality in prostate MRI using a similar Philips SmartSpeed-based technology [21]. Their study also noted a significant reduction in acquisition time while maintaining diagnostic confidence, a key benefit also observed in this fetal cohort.

However, a notable difference lies in the quantitative ERD results. While Bischoff et al. found that their AI reconstruction led to a decrease in ERD (indicating improved quantitative sharpness), this the present study found a significant increase. This discrepancy may stem from several factors: the inherent high-contrast and static nature of the prostate versus the dynamic, lower-contrast fetal brain, differences in sequence parameters, or variations in the ERD measurement methodology itself. This underscores that the impact of AI reconstruction can be highly dependent on the specific anatomical application and highlights the risk of extrapolating findings from one body region to another.

One result also resonates with studies in musculoskeletal imaging. For example, Terzis et al. (2024), using a CS-Super Resolution prototype, achieved a 57% scan time reduction in knee MRI while preserving or enhancing image quality compared to conventional CS [27]. They similarly observed a contradiction where quantitative SNR values did not always align with the superior subjective ratings, suggesting that “traditional ROI-based measurements are no valid criteria for objective image quality in AI-based reconstructions.” This reinforces the conclusion that expert’s perceptual evaluation is currently paramount for assessing these advanced algorithms.

5.3 Limitations and Considerations

Several limitations should be considered when interpreting the results of this study.

- Study Design and Sample Size:

The single-center, retrospective design inherently limits the generalizability of the findings. The sample size was modest, particularly for the qualitative (n=11) and ERD (n=20) analyses, which increases the risk of type II errors and may not capture the full spectrum of performance.

- Data Heterogeneity:

The inclusion of a wide range of gestational ages (16-37 weeks) introduced significant biological variability, likely contributing to the high variance and lack of significance in the SNR and CNR measurements. Future studies might benefit from analyzing more homogenous age-specific subgroups.

- Validity of Quantitative Metrics:

The paradoxical ERD results challenge the validity of this traditional metric for AI-super-resolution images. As the underlying image data is fundamentally altered by interpolation, new quantitative biomarkers that better correlate with human perception of diagnostic quality are needed.

- Technical Specificity:

The study was conducted on a single vendor's 3T scanner using a specific research software prototype (NGSA patch). The performance may differ with other scanners, field strengths, or commercially implemented software versions.

- Offline Reconstruction:

The AI processing was performed offline, and reconstruction times were not measured. The computational demand and its integration into a time-sensitive clinical workflow are critical factors for real-world implementations and require further investigation.

5.4 Clinical Implications and Future Outlook

Despite its limitations, this study offers promising insights for clinical practice. The ability to produce high-quality fetal brain MR images under 40 seconds could significantly reduce scan time-related stress of pregnant women, minimize motion artifacts, and optimize scanner throughput. The enhanced diagnostic confidence shown in this study may enable faster, more reliable decision-making in prenatal care. These results support the increasing integration of AI in fetal MRI, especially in motion-sensitive and time-constrained settings.

Future studies should aim to validate these findings in larger, multicenter cohorts. It will also be critical to assess the performance of AI-enhanced reconstructions in pathological cases. Moreover, development and validation of novel, perceptually aligned quantitative image quality metrics will be essential to bridge the gap between numerical analysis and diagnostic utility. As both Bischoff et al. and Hahnfeldt et al. emphasize, AI has the potential to transform MRI practice, but its implementation must be rigorous, reproducible, and patient-centered [21], [25].

6 Conclusion

This master thesis successfully evaluated the clinical potential of an AI-based super-resolution algorithm for enhancing ultrafast T2-weighted fetal MRI. The study concludes that the AI-based post-processing technique provides a substantial improvement in perceived diagnostic image quality, enabling the acquisition of images rated superior to both conventional accelerated and standard high-resolution sequences in significantly less time.

The primary research hypotheses were verified. The AI-enhanced sequence yielded images that were qualitatively assessed by expert radiologists as significantly sharper, with better tissue contrast and inspiring higher diagnostic confidence. This directly answers the core research questions, confirming that the technology can (1) improve the quality of ultrafast images, (2) outperform even time-intensive high-resolution protocols in diagnostic perception, and (3) facilitate a 38% reduction in scan time compared to the high-resolution standard without sacrificing quality. Furthermore, (4) the expert evaluation confirmed the high diagnostic usability of the AI images, particularly for delineating fine anatomical structures, despite noting a tendency for the algorithm to amplify minor motion artifacts.

A key insight from this study is that the clear qualitative preference for the AI-enhanced images was not consistently mirrored by traditional quantitative metrics such as SNR and CNR. This finding aligns with similar observations in the literature, where the complex perceptual improvements offered by deep learning algorithms are not always fully captured by simple ROI-based measurements. This reinforces the conclusion that in the current era of AI-driven reconstruction, subjective expert evaluation remains the most crucial method for determining true diagnostic utility.

The findings of this thesis do not merely provide answers but also open new and important avenues of inquiry for the future. Beyond the immediate need to validate these results in larger, multi-centre studies including cases with known fetal pathologies, this work raises a more profound clinical question: How should the entire clinical team adapt its interpretation and workflow to confidently and safely leverage the new visual information provided by AI-reconstructed images?

The experts feedback revealed that while AI enhances anatomical clarity, it can also amplify subtle artifacts and create an appearance of “hyper-realism” that requires scrutiny. This suggests that the integration of such powerful tools is not just a technical upgrade but may represent a paradigm shift that affects the entire diagnostic process. The answer may therefore lie not only in training radiologists to interpret these new images, but also in fostering a deeper technical understanding across all involved professional groups. When radiographers and radiologists alike understand the underlying principles, strengths, and potential pitfalls of these algorithms-specific artifacts, and ultimately improve diagnostic accuracy.

Ultimately, while this research confirms the transformative potential of AI to make fetal MRI faster and more powerful, its successful and efficient clinical integration will depend on our ability to build this interdisciplinary expertise. This requires a commitment to rigorous validation, shared education across clinical teams, and a deep understanding of the unique characteristics of AI-generated images, always ensuring that the patient's well-being remains the guiding principle.

References

[1] D. Pugash, P. C. Brugger, D. Bettelheim, und D. Prayer, „Prenatal ultrasound and fetal MRI: The comparative value of each modality in prenatal diagnosis“, *European Journal of Radiology*, Bd. 68, Nr. 2, S. 214–226, Nov. 2008, doi: 10.1016/j.ejrad.2008.06.031.

[2] F. Machado-Rivas *u. a.*, „Fetal MRI at 3 T: Principles to Optimize Success“, *RadioGraphics*, Bd. 43, Nr. 4, S. e220141, Apr. 2023, doi: 10.1148/rg.220141.

[3] R. Meshaka, T. Gaunt, und S. C. Shelmerdine, „Artificial intelligence applied to fetal MRI: A scoping review of current research“, *BJR*, S. 20211205, März 2022, doi: 10.1259/bjr.20211205.

[4] K. K. Marathu *u. a.*, „Fetal MRI Analysis of Corpus Callosal Abnormalities: Classification, and Associated Anomalies“, *Diagnostics*, Bd. 14, Nr. 4, S. 430, Feb. 2024, doi: 10.3390/diagnostics14040430.

[5] Weisstanner, C., Gruber, G. M., Brugger, P. C., Mitter, C., Diogo, M. C., Kasprian, G., & Prayer, D., „Fetal MRI at 3T-ready for routine use?“, Nr. British Journal of Radiology, 90(1069), doi: ARTN 20160362 10.1259/bjr.20160362.

[6] D. Prayer *u. a.*, „MRI of normal fetal brain development“, *European Journal of Radiology*, Bd. 57, Nr. 2, S. 199–216, Feb. 2006, doi: 10.1016/j.ejrad.2005.11.020.

[7] T. Victoria *u. a.*, „Fetal magnetic resonance imaging: jumping from 1.5 to 3 tesla (preliminary experience)“, *Pediatr Radiol*, Bd. 44, Nr. 4, S. 376–386, Apr. 2014, doi: 10.1007/s00247-013-2857-0.

[8] A. M. Hubbard, „Ultrafast fetal MRI and prenatal diagnosis“, *Seminars in Pediatric Surgery*, Bd. 12, Nr. 3, S. 143–153, Aug. 2003, doi: 10.1016/S1055-8586(03)00031-3.

[9] Elster, ELSTER LLC, „Questions and Answers in MRI“, Types of Neural Networks. Zugegriffen: 12. Mai 2025. [Online]. Verfügbar unter: https://mriquestions.com/deep-network-types.html

[10] X. Zhu, B. Tomanek, und J. Sharp, „A pixel is an artifact: On the necessity of zero-filling in fourier imaging“, *Concepts Magnetic Resonance*, Bd. 42A, Nr. 2, S. 32–44, März 2013, doi: 10.1002/cmr.a.21256.

[11] AD Elster, ELSTER LLC, „Questions and Answers in MRI“, Zero-Interpolation Filling. [Online]. Verfügbar unter: https://mriquestions.com/zip.html

[12] R. D. Peters und M. L. Wood, „Multilevel wavelet-transform encoding in MRI“, *Magnetic Resonance Imaging*, Bd. 6, Nr. 3, S. 529–540, Mai 1996, doi: 10.1002/jmri.1880060317.

[13] T. Lehmann, W. Oberschelp, E. Pelikan, und R. Repges, „Die Wavelet-Transformation“, in *Bildverarbeitung für die Medizin: Grundlagen, Modelle, Methoden, Anwendungen*, T. Lehmann, W. Oberschelp, E. Pelikan, und R. Repges, Hrsg., Berlin, Heidelberg: Springer Berlin Heidelberg, 1997, S. 253–281. doi: 10.1007/978-3-642-60487-4_10.

[14] O. N. Jaspan, R. Fleysher, und M. L. Lipton, „Compressed sensing MRI: a review of the clinical literature“, *BJR*, Bd. 88, Nr. 1056, S. 20150487, Dez. 2015, doi: 10.1259/bjr.20150487.

[15] Elster, ELSTER LLC, „Questions and Answers in MRI“, Compressed Sensing. Zugegriffen: 6. Mai 2025. [Online]. Verfügbar unter: https://mriquestions.com/compressed-sensing.html

[16] S. Geethanath *u. a.*, „Compressed Sensing MRI: A Review“, *Crit Rev Biomed Eng*, Bd. 41, Nr. 3, S. 183–204, 2013, doi: 10.1615/CritRevBiomedEng.2014008058.
[17] D. Liang, J. Cheng, Z. Ke, und L. Ying, „Deep Magnetic Resonance Image Reconstruction: Inverse Problems Meet Neural Networks“, *IEEE Signal Process. Mag.*, Bd. 37, Nr. 1, S. 141–151, Jan. 2020, doi: 10.1109/MSP.2019.2950557.
[18] N. Pezzotti *u. a.*, „An Adaptive Intelligence Algorithm for Undersampled Knee MRI Reconstruction“, *IEEE Access*, Bd. 8, S. 204825–204838, 2020, doi: 10.1109/ACCESS.2020.3034287.
[19] L. Beljaards, N. Pezzotti, C. Rao, M. Doneva, M. J. P. Van Osch, und M. Staring, „AI-based motion artifact severity estimation in undersampled MRI allowing for selection of appropriate reconstruction models“, *Medical Physics*, Bd. 51, Nr. 5, S. 3555–3565, Mai 2024, doi: 10.1002/mp.16918.
[20] N. Pezzotti, S. Yousefi, J. van Gemert, M. Doneva, und S. Kastryulin, „Adaptive-CS-Net: FastMRI with Adaptive Intelligence“.
[21] L. M. Bischoff *u. a.*, „Deep Learning Super-Resolution Reconstruction for Fast and Motion-Robust T2-weighted Prostate MRI“, *Radiology*, Bd. 308, Nr. 3, S. e230427, Sep. 2023, doi: 10.1148/radiol.230427.
[22] M. Jurka *u. a.*, „Deep-learning-based reconstruction of T2-weighted magnetic resonance imaging of the prostate accelerated by compressed sensing provides improved image quality at half the acquisition time“, *Quant Imaging Med Surg*, Bd. 14, Nr. 5, S. 3534–3543, Mai 2024, doi: 10.21037/qims-23-1488.
[23] C. Ledig *u. a.*, „Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network“, in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI: IEEE, Juli 2017, S. 105–114. doi: 10.1109/CVPR.2017.19.
[24] F. N. Harder *u. a.*, „Prospectively Accelerated T2-Weighted Imaging of the Prostate by Combining Compressed SENSE and Deep Learning in Patients with Histologically Proven Prostate Cancer“, *Cancers*, Bd. 14, Nr. 23, S. 5741, Nov. 2022, doi: 10.3390/cancers14235741.
[25] R. Hahnfeldt *u. a.*, „Is a 3-Minute Knee MRI Protocol Sufficient for Daily Clinical Practice? A SuperResolution Reconstruction Approach Using AI and Compressed Sensing“, *Diagnostics*, Bd. 15, Nr. 10, S. 1206, Mai 2025, doi: 10.3390/diagnostics15101206.
[26] P. A. Yushkevich *u. a.*, „User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability“, *NeuroImage*, Bd. 31, Nr. 3, S. 1116–1128, Juli 2006, doi: 10.1016/j.neuroimage.2006.01.015.
[27] R. Terzis *u. a.*, „Five-minute knee MRI: An AI-based super resolution reconstruction approach for compressed sensing. A validation study on healthy volunteers“, *European Journal of Radiology*, Bd. 175, S. 111418, Juni 2024, doi: 10.1016/j.ejrad.2024.111418.

List of Figures

List of Tables

Appendix

A. Survey Data

Microsoft Forms Experts Evaluation Questionnaire - https://forms.office.com/e/cvECXzSYCi

Fetal MRI - Experts Evaluation

5-10 Cases

Abschnitt 1

Qualitative comparison of T2w Low Res, AI postproccessed and High Res Images

Please check the following 10 cases at the PACs Session and rate them in the tables below.

Abschnitt 2

Case 1

Sagittal - Anatomische Struktur: Mittellinie, Vermis lobuli, Aquäduct

1. Artifacts

	Non-diagnostic - Massive artifacts, not interpretable	Moderate distortion - noticeable but acceptable	Minimal distortion - barely noticeable artifacts	No distortion - artifact free image
Picture 1	◯	◯	◯	◯
Picture 2	◯	◯	◯	◯
Picture 3	◯	◯	◯	◯

2. Image Sharpness

	Poor delineation - structures not distinguishable	Moderte delineation - unclear but identifiable	Minimal obstruction - nearly perfectly sharp	Perfectly sharp - excellent anatomic clarity
Picture 1	◯	◯	◯	◯
Picture 2	◯	◯	◯	◯
Picture 3	◯	◯	◯	◯

3. Tissue Contrast

	No contrast - tissue borders not visible	Moderate contrast - (very) low differentiation	Good contrast - clear tissue distinction	Excellent contrast - optimal differentiation
Picture 1	○	○	○	○
Picture 2	○	○	○	○
Picture 3	○	○	○	○

4. Overall Image Quality

	Not usable - nondiagnostic image	Acceptable - usable for interpretation	Good - reliable diagnostic quality	Excellent - optimal diagnostic image
Picture 1	○	○	○	○
Picture 2	○	○	○	○
Picture 3	○	○	○	○

5. Diagnostic Confidence

	No confidence - diagnosis not possible	Moderate confidence - highly uncertain	Good confidence - reliable interpretation	High confidence - fully confident diagnosis
Picture 1	○	○	○	○
Picture 2	○	○	○	○
Picture 3	○	○	○	○

6. Commentar (optional)

Ihre Antwort eingeben

Fetal MRI - Experts Evaluation