The Development of a Generative Approach for Joint Super-Resolution Image Reconstruction from Highly Sparse Raw Data in the Context of MR-PET Imaging

Krzysztof Malczewski ⬡

Institute of Information Technology, Warsaw University of Life Sciences – SGGW, Warsaw, Poland
(krzysztof_malczewski@sggw.edu.pl)

Abstract The present study introduces a rapid and efficient approach for reconstructing high-resolution images in hybrid MRI-PET scanners. The application of sparsity, compressed sensing (CS), and super-resolution reconstruction (SRR) methodologies can significantly decrease the demands of data acquisition while concurrently attaining high-resolution output. G-guided generative multilevel networks for sparsely sampled MR-PET input are shown here. Compressed Sensing using conjugate symmetry and Partial Fourier methodology speeds up data collection over k-space sampling methods. GANs and k-space adjustments are used in this image domain technique. The employed methodology utilizes discrete preprocessing stages to effectively tackle the challenges associated with the deblurring, reducing motion artifacts, and denoising of layers. Initial trials offer contextual details and accelerate evaluations. Preliminary experiments provide contextual information and expedite assessments.

Keywords: GAN, WGAN, super-resolution, compressive sensing, medical modalities

1. Introduction

Commercial PET-MRI imaging equipment with synergistic capabilities debuted in 2010. According to source [34], they have a competitive advantage. Clinical imaging is improved by hybrid imaging technology. MR/PET combines MRI soft tissue morphology with PET functional imaging. These advances are driven by how well different imaging modalities provide correlated, not duplicative, findings. The effective outcome was achieved by integrating functional imaging from Positron Emission Tomography with CAT's soft tissue analysis, two oncological technologies. The use of $^{18}F\text{-}FDG$ with CAT scanners has been recognized in relevant research. Medical imaging using FDG-PET can identify and quantify malignant cells' metabolic rate. In therapy management, CT scans can detect even the smallest wounds that PET scans may miss due to their limited range or technological restrictions. Respiration, locomotion, and circulation are typical causes. The soft tissue contrast of MRI is well-known.

The main argument favors MRI over CT. This approach helped treating neurological problems, brain tumors, craniofacial defects, abdominal wall masses, mass-like lesions, and other conditions. Despite following MRI protocols to the letter, emission tomography is effective. PET and MRI are compatible, as indicated in reference [1]. PET and CT use different radiation wavelengths and can be combined to improve their efficacy. PET

and MRI image capture methods differ. MR-images may affect PET signal acquisition because they require a visually appealing and steady field. The above method meets medical image processing requirements. Photomultipliers cannot detect PET signals in strong magnetic fields. To overcome this limitation, a mobile table connects magnetic resonance (MR) and positron emission tomography (PET) scanners in different places. Supine patients undergo PET and MR imaging without movement. The current architectural design prevents simultaneous collecting of unprocessed data, prolonging diagnostic procedures and increasing patient problems. The integration of MRI-PET may solve this obstacle. According to references [35, 36], super-resolution techniques have improved medical image processing.

Deep learning algorithms can accurately replicate complicated relationships between low-resolution and high-resolution pictures, even under demanding situations, advancing Single Image Super-Resolution (SISR). Image quality improved after enhancing. Structured Convolutional Neural Networks (CNNs) help Super-Resolution Convolutional Neural Networks (SRCNNs) and their accelerated variations provide better Single Image Super-Resolution results for two-dimensional natural images. This phenomena is observed in sources [4, 28].

Patch, edge, sparse coding, prediction, and statistics have been conventional algorithm groupings for decades. These methods cost less than deep learning to compute. Deep learning has improved convolutional neural network use, advancing super-resolution. Despite deep-learning advances, medical picture super-resolution remains unsolved. Medical imaging uses 3D volumes. CNNs used to ignore the input's three-dimensional structure. Because 3D models require more memory and compute computational power than 2D models, their usefulness is limited. Convolutional neural networks (CNNs) optimize pixel or voxel-level error, measured by mean squared error (MSE) between the predicted model's output and a high-resolution reference. Research in [46] suggests that using MSE and PSNR as metrics for assessing picture accuracy may be unreliable. Mean Squared Error enhancement reduces only picture sharpness and perceptual accuracy.

Generative Adversarial Networks (GANs) have gained significant popularity and are extensively employed in many applications such as image super-resolution, modality switching, and synthesis. The aforementioned domains have been extensively examined in the literature [19, 24, 27].

The utilization of 3D Multi-Level Densely Connected Super-Resolution Networks (mDCSRN) has the potential to address the aforementioned issues. A highly linked network reduces the weight of a mDCSRN [5].

Enhancing intensity difference optimization increases model size and speed while preserving performance. GAN training improves system efficacy, according to research. A common deep learning neural network architecture includes a generator ($\mathbb{G}$) and a discriminator ($\mathbb{D}$). The generator and discriminator compete to minimize the difference

between generated and actual data during training. The Generative Adversarial Network was introduced by Goodfellow and colleagues in 2014 [11].

Super-Resolution computer vision applications use Generative Adversarial Networks and GANs with adversarial and perceptual loss functions are designed to perform picture Super-Resolution (SR). Superior textures are restored to lower-resolution images. The network can retrieve exact textures and high-frequency components. However, its scope is limited. Generational Adversarial Networks can modify data and introduce noise. Super-resolution and other methods were evaluated for the task of improving image quality [21, 41]. MRI distortions are generated by imaging plane motion. Motion is needed for super-resolution. A recent study suggests that convolutional neural networks can enhance medical image quality [6, 26, 37]. Researchers developed SRCNN [7], a deep convolutional network, for super-resolution reconstruction. CNNs were first used in Super-Resolution. Shi et al. introduced a sub-pixel convolutional layer as an alternative to the deconvolutional layer [42].

The training method becomes simpler. Simple linear network designs underpin the methods. The link between neural network depth and over-parameterization is growing. Previous research indicates that recursive networks can effectively handle difficulties by applying weights repeatedly [25, 44]. Increased network depth improves performance, but deeper networks are more prone to gradient outbursts. Hyun et al. utilized Convolutional Neural Networks and k-space rectification methods to replace missing k-space data regions with original data [18]. Thus, it is crucial to improve the effectiveness of mitigating the aliasing artifacts.

A primary constraint associated with Magnetic Resonance Imaging concerns the phase of the assessment. The expeditious acquisition of MRI data has garnered significant attention from a multitude of researchers. Improvement is necessary in the phase encoding intervals utilized during the sampling of k-space. This phenomenon typically leads to a decline in the visual accuracy of the image. The implementation of the proposed k-space sampling pattern would yield advantages in resolving the matter. As per the author's description, the procedure of populating k-space entails obtaining subsets along a designated phase encoding direction. The methodology utilized in this approach involves the utilization of blades that are similar to those found in a propeller. The implementation of Hermitian symmetry results in the halving of the complex space. This feature enables the retrieval of the missing k-space component. This methodology improves the understanding of components that occur frequently.

This paper elucidates a methodology based on Generative Adversarial Network that has been employed for the purpose of reconstructing Compressed Sensing Magnetic Resonance Imaging (CS-MRI), taking cues from previous studies. The methodology that has been put forth involves the amalgamation of Generative Adversarial Networks that are reliant on images, along with k-space corrections. The aforementioned methodology demonstrates enhanced efficacy in contrast to singular and non-sequential techniques for

correcting k-space. The current approach incorporates the fusion of deformable image registration and Generative Adversarial Networks, and has been extended to incorporate the fusion of multiple frames of visual data. The Wasserstein generative adversarial network (WGAN) was employed to optimize the algorithm's performance and promote model convergence during the training phase.

The results have been the subject of rigorous scrutiny in focused research investigations. The main aim of the methodology described in this manuscript is to improve the accuracy and quality of hybrid scanner images with regards to boundary demarcation, while also decreasing the time required for acquisition.

This work presents a novel algorithm that is suitable for use with MR/PET and integrates super-resolution, accurate estimation of movement, and streamlines the examination. The results have faced significant scrutiny in focused empirical inquiry. The primary objective of the methodology described in this manuscript is to enhance the precision and excellence of images concerning the identification of boundaries, while concurrently reducing the time required for acquisition (see Figure 1).

The primary findings of this work include the following:

1. The framework algorithm demonstrates a comprehensive approach towards the joint reconstruction process of MR-PET images. Various aspects such as sparse sampling trajectories, synchronizing of k subspaces, deblurring, noise reduction, motion compensation, and subsequently, increasing the resolution of an image are key areas of focus in this study.

2. The present study introduces a novel model for reconstructing MR-PET images using a generative super-resolution approach.

3. The methodology provided employs the joint sparsity of both the MR and PET modalities.

4. The sparsity of the MR and PET raw data has resulted in an acceleration of the input data collecting process.

5. The technique has been specifically intended for gathering visual information across different scales. This issue is often simplified by other authors.

6. The algorithm is capable of extracting visual features at different scales. This subject matter is frequently oversimplified by other writers.

7. The employed methodology entails distinct preprocessing stages to address the challenges of blur and noise removing layers.

8. The proposed technique employs a reconstruction strategy for magnetic resonance imaging that leverages convolutional neural networks. This approach aims to restore low-quality images derived from highly sparse raw data.

9. The aforementioned methodology employs the compressed sensing framework to prioritize the minimization of data acquisition durations.

10. The reconstruction layer of the procedure is nested with the author's deformable motion estimation procedure.
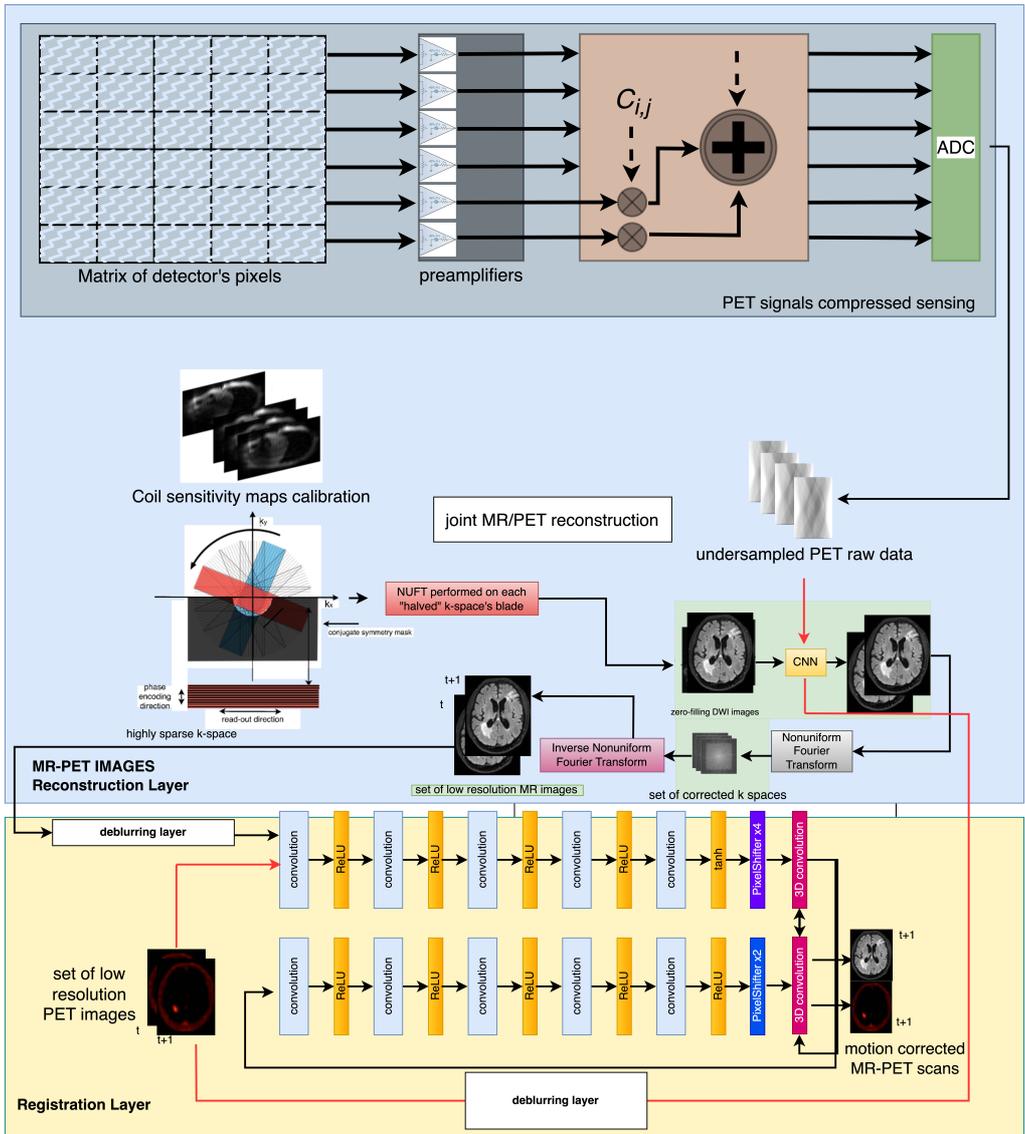
Fig. 1. The process for reconstructing low resolution MR-PET images. The method takes into account the lack of overlap between the MR and PET modalities. At its core, the deblurring net is nested. The raw PET signals are compressively detected in the upper part. The lower part pertains to the sparsity of the raw MR signals.

2. Joint sparseness on MR/PET

Although the modalities used have various physical bases, reconstruction often uses side-by-side projections. Despite the potential that both images use the same geometry, the assertion implies discrete image reconstruction processes. The reconstruction process can be simplified by sharing critical inter-technique data and recognizing similarities between objects [34]. Integrating this characteristic with other extracted structures may reduce motion abnormalities. Instead of reconstructing phases, the combined sparsity algorithm is used. It uses structural similarities to improve spatial resolution and eliminate involuntary patient movement during image capture by combining two sparse datasets. This algorithm solves the optimization problem within concurrent constraints.

In MRI and PET the Compressed Sensing in conjunction with Partial Fourier transform and the exploitation of conjugate symmetry have been used to induce sparsity in the data sets. The expression of joint sparsity can be formulated as follows:

$$\left\|\begin{matrix} \mathbb{S}(x_{\mathrm{MRI}}^i) \\ \mathbb{S}(x_{\mathrm{PET}}^i) \end{matrix}\right\|_2 = \sqrt{\left(\mathbb{S}(x_{\mathrm{MRI}}^i)\right)^2 + \left(\mathbb{S}(x_{\mathrm{PET}}^i)\right)^2}\,.$$

where x_{MRI}^i and x_{PET}^i are three-dimensional image volumes coming from MRI and PET, respectively. This equation involves the sparsifying transform $\mathbb{S}$ applied to these image volumes. The regularization carried out at every voxel can be expressed in the following manner:

$$\daleth^i = \|\mathbb{S}(^{\mathrm{in}}M_{\mathrm{MRI}}^i)\| - \|\mathbb{S}(^{\mathrm{in}}P_{\mathrm{PET}}^i)\|\,,$$

where $^{\mathrm{in}}M_{\mathrm{MRI}}^i$ and $^{\mathrm{in}}P_{\mathrm{PET}}^i$ refer to MR and PET input data streams. The symbol $\daleth^i$ denotes joint sparsity regularization. The regularization parameters present in the process serve to mitigate the overlap of non-coherent features in MR and PET images.

The methodology presented in this scholarly article utilizes the concept of joint sparsity, specifically in the domains of MRI and PET, as illustrated in Figure 1.

3. Sparse sampling versus MR/PET raw data

The core method integrates and synthesizes data from multiple modalities. The current study compresses PET data volume, as cited in reference [34]. Positron-emitting radioactive elements are mixed whenever possible to reduce readout channels. Consolidating their output signals increased PET scan resolution. MR/PET hybrid scanners integrate super-resolution and compressive sensing through structural components. Sparse depiction of the detector's structure is obvious. The sparsity attribute can create new multiplexing setups. Random matrices with constrained isometry can be generated using several stochastic methods. The notion of greatest likelihood guides sensing matrix frameworks. Research indicates that creating detected matrices results in the lowest

reconstruction error in MS [30]. This scientific article describes a method that uses a limited number of channels to create spatial and temporal domains. PET input data is sparsely sampled. Each output can be interpreted mathematically as a linear combination of photodetector pixels with weights $c_{k,n}$ (refer to Figure 1). The number of sensors is lowered via 4:1 subsampling. MR-PET joint sparsity and shared product characteristics allow motion model parameters from MR data sets to improve PET image quality. Super-resolution is achieved using the same method as shown in Figure 1. This process used to enhance low-resolution MR and PET scans. This study uses PROPELLER, Poisson Disc, and Partial Fourier sampling. In this strategy, compactification inhibits signal recovery.

The encoding procedure uses 2D 3×3 convolutional layers iteratively, like CNNs. After layer processing, leaky rectified linear units, batch normalization, and 2×2 maximum pooling are used for downsampling.

This paper presents an approach to reconstruct low-resolution magnetic resonance pictures. K-space blades with high sparsity achieve this. To reduce data collection time, the sampling method reduces data density and uses a conjugate symmetric mask. To correct motion and blur, deblurring and registration layers improve low-resolution images.

U-net design was trained using the mean squared error loss function, which is mathematically represented by the statement: every zero-filled image is linked to a completely sampled image, denoted by $\mathbb{S}_{true}$. Adam's optimizer, previously discussed in reference [48], is used to reduce the loss function. The study used a training rate of 0.0001 and ran the process for 100 epochs. Only 32 images were used for training. The hyperparameters were determined using empirical observations.

$$\beta^i = \begin{cases} \underset{\beta^i}{\mathrm{argmin}} \ \left\| \mathbb{S}_{\mathrm{true}} - f_{\beta^i}\left(|F^{-1}(y_0)|\right) \right\|, & i = 0; \\ \underset{\beta^i}{\mathrm{argmin}} \ \left\| \mathbb{S}_{\mathrm{true}} - f_{\beta^i} - f_{\beta^i}(\mathbb{S}^i) \right\|, & \text{otherwise.} \end{cases}$$

4. The application of Generative Adversarial Networks within the framework of Super Resolution image reconstruction

The model framework that has been built is illustrated in Figure 1. The system consists of two components, specifically the deformable motion estimation and the reconstructing network. The second component consists of two blocks: a generating block and a discriminating block.

The effectiveness of the generative adversarial network framework in the domain of motion correction relies on its ability to improve picture restoration and support the reconstruction of missing raw data. The function indicated above have the power to
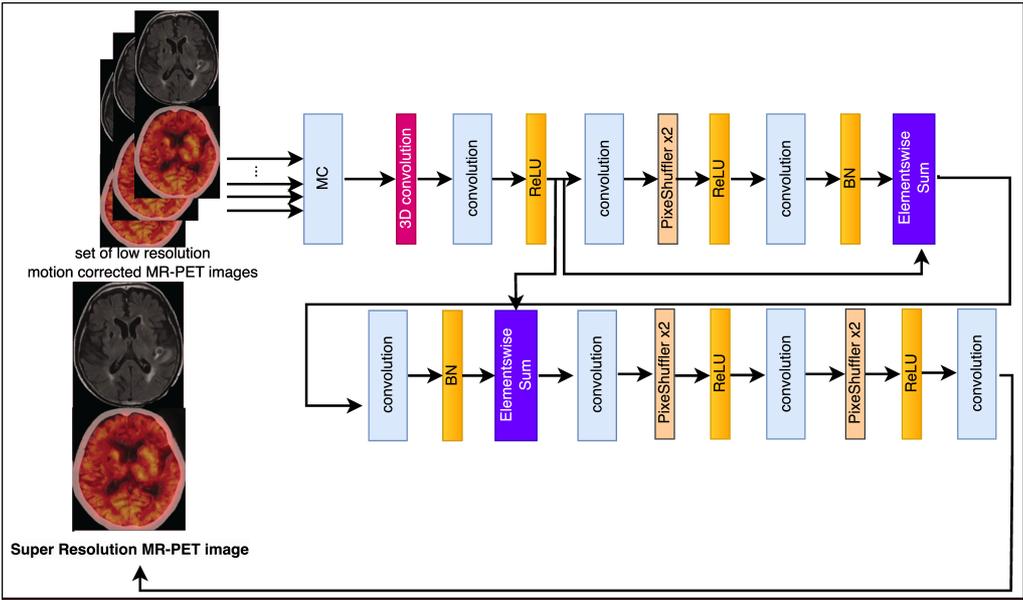
Fig. 2. The diagram illustrating the structure and flow of the generator net.

produce shots of exceptional quality. The main goal of the generator is to generate samples that have a significant level of resemblance to authentic data, while the discriminator aims to effectively categorize samples as either real or artificial.

$$\min_{\mathbb{G}} \max_{\mathbb{D}} V(D, G) = \min_{\mathbb{G}} \max_{\mathbb{D}} \mathbb{E}_x \big[\log \big(\mathbb{D}(x) \big) \big] - \mathbb{E}_y \big[\log \big(1 - \mathbb{G}(y) \big) \big].$$

The variables y and x represent motion deformed and corrected illustrations, respectively. With the exception of the core layer, encoder blocks are comprised of five convolutional layers and $\frac{n}{2}$ feature maps, each containing n mappings. The encoder blocks and decoder blocks share an architectural design, but transposed convolutions replace all convolutional layers. A method for estimating spatial transformation parameters is used in image registration technology, as detailed in [14]. Following this, the displacement discrepancy between frames is corrected. Displacement parameters can change the spatial location of frames in sequences that depict the same subject but were shot at different times and places.

Multiple pairings from the registration module are blended with I_{LR} frames using a $3D$ convolutional layer. The user-generated output is sent into the generator network. The study used a generator network design (denoted as $\mathbb{G}$) based on the SR-GAN architecture (see Figure 2). A single residual block is used in the $\mathbb{G}$-network to reduce the
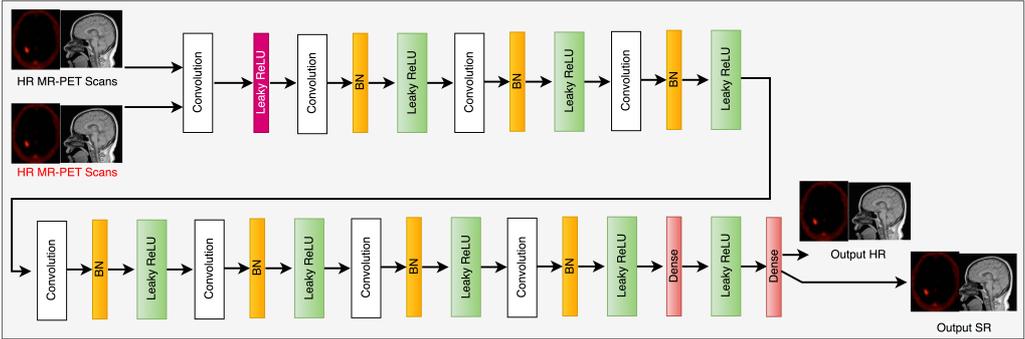
Fig. 3. The flowchart illustrating the architecture and functionality of the discriminator net.
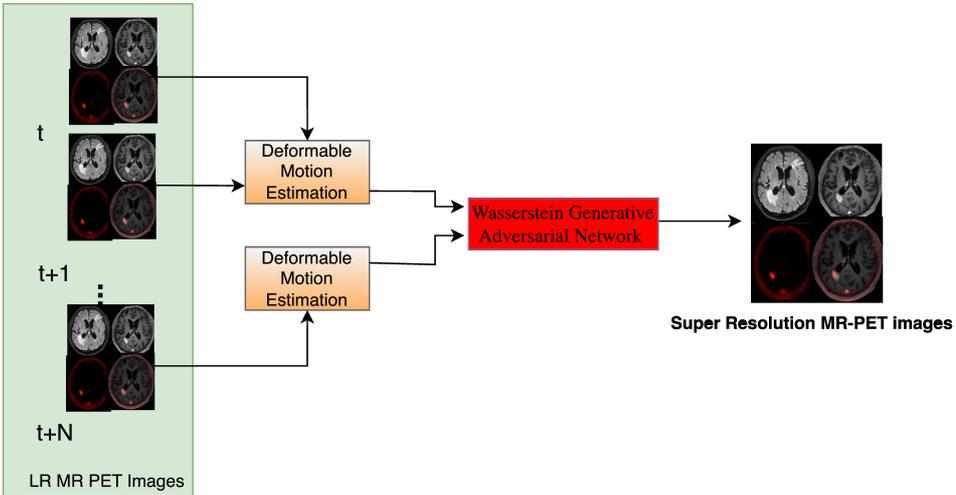

Fig. 4. The proposed magnetic resonance super resolution image reconstruction algorithm.

number of parameters and maintain generalization. To attain the required detail, the residual network uses two sub-pixel convolutional layers. The architecture of the discriminator, denoted as $\mathbb{D}$ in Figure 3, consists of eight convolutional layers. As network levels rise, attributes correlate positively. Convolutional kernel reduction reduces feature dimensionality. Two modifications were made to address SR-GAN reconstruction and network training/convergence difficulties. In the initial phase, the discriminator $\mathbb{D}$ ignored the *sigmoid* activation function in the output layer. In addition, parameter modifications were limited to a constant value of c (0.01) relative to their absolute

magnitude. The investigation focuses on insufficient security protocols during training and complex model convergence, as supported by references [16, 17, 29]. The anomaly is due to the low overlap between genuine and counterfeit distributions. Disregarding the statistical measure JS divergence, which compares distributions, may prevent network convergence. Arjovsky and colleagues found that the Wasserstein distance accurately measures the distribution separation even when overlap is low.

5. Methods used for the reconstruction of high-resolution MR-PET images

5.1. Reconstruction of the images and the loss function

The methodology commences by reconstructing the MR and PET images with low resolution, utilizing subspaces that have been inadequately sampled. Refer to [33] for this procedure.

The method uses blur, noise removing and motion estimation layers, with its main reconstruction process arranged as shown in Figures 2, 3, and 4.

The WGAN [16] shows that the Wasserstein distance improves confrontation network formation. The definition of Wasserstein distance is as follows:

$$\mathbb{W}(\mathbb{P}_{\text{ref}}, \mathbb{P}_{\text{gen}}) = \frac{1}{K} \sup_{\substack{||f||_{L \leq K} \\ \mathbb{W} \in \prod(\mathbb{P}_{\text{ref}}, \mathbb{P}_{\text{gen}})}} \mathbb{E}_{(x,y) \sim \mathbb{P}_{\text{ref}}}[f(x)] - \mathbb{E}_{x \sim \mathbb{P}_{\text{gen}}}[f(x)].$$

The equation above uses the symbol $\prod(\mathbb{P}_{\text{ref}}, \mathbb{P}_{\text{gen}})$ to represent all possible joint probability distributions between $\mathbb{P}_{\text{ref}}$ and $\mathbb{P}_{\text{gen}}$. The discrimination function of the adversarial network is $f_{\mathbb{W}}$, as shown in the equation. This limits the discriminator's input sample derivative to a predetermined range. The variable $\mathbb{W}$ in the domain $\mathbb{D}$ undergoes a modification procedure limited to the range of $-c$ to c. This technique emphasizes the gradient update generator and reduces the disappearing gradient problem. The function denoted by $f_{\mathbb{W}}$ satisfies the following equation:

$$L = \mathbb{E}_{x \sim \mathbb{P}_{\text{ref}}}[f_{\mathbb{W}}(x)] - \mathbb{E}_{x \sim P_{\text{gen}}}[f_{\mathbb{W}}(x)].$$

As the variable L increases, it becomes feasible to estimate the Wasserstein distance between the probability distributions $\mathbb{P}_{\text{ref}}$ and $\mathbb{P}_{\text{gen}}$ through approximation. The former word refers to legitimate information diffusion, whereas the latter applies to synthesized information. The discriminator and generator loss functions are precisely specified as follows:

$$\mathbb{D}_{\text{loss}} = \mathbb{E}_{x \sim \mathbb{P}_{\text{gen}}}[f_{\mathbb{W}}(x)] - \mathbb{E}_{x \sim \mathbb{P}_{\text{ref}}}[f_{\mathbb{W}}(x)],$$

$$\mathbb{G}_{\text{loss}} = \mathbb{E}_{x \sim \mathbb{P}_{\text{gen}}}[f_{\mathbb{W}}(x)].$$

The training method is defined by the discriminator loss function, $\mathbb{D}_{\text{loss}}$. To evaluate GAN training, the Wasserstein distance between real and generated data distributions should decrease. Distance magnitude is negatively correlated with this measure.

The objective of this strategy is to enhance the learning procedure of the generator, represented as $\mathbb{G}$. The objective of this task is to evaluate the relationship between the input sequence I_t^{LR} (with values ranging from 1 to N) and its corresponding counterpart I_t.

The task was successfully accomplished by employing a feedforward Convolutional Neural Network. The neural network underwent training using the parameter $\Psi_{\mathbb{G}}$. The parameters of the neural network, denoted as $\Psi_{\mathbb{G}} = \{U_{1:L}; b_{1:L}\}$, with L layers, are obtained by minimizing the loss function $l_{\mathbb{G}}$ for the Super-Resolution generation network as described in reference [31]:

$$\Psi_{\mathbb{G}}^* = \underset{\Psi_{\mathbb{G}}}{\text{argmin}} \frac{1}{N} \sum_{t=1}^{N} l_{\mathbb{G}}(\mathbb{G}_{\Psi_{\mathbb{G}}} I_t^{\text{LR}}, I_t^{\text{HR}}).$$

The current study utilizes a loss function, referred to as $l_{\mathbb{G}}$, that is based on previous scholarly research and has been appropriately acknowledged and cited in [38].

$$l_{\mathbb{G}} = l_{\text{MSE}} + 10^{-6} l_{\text{gen}}.$$

where l_{MSE} is defined by one of the equations below. The comprehensive net loss function of the SR-GAN model encompasses the loss functions of both the generating and discriminating blocks, denoted as $l_{\mathbb{G}}$ and $l_{\mathbb{D}}$, respectively.

$$l_{\mathbb{D}} = \frac{1}{N} \sum_{n=1}^{N} \left(\log \left(1 - \mathbb{D}_{\Psi_{\mathbb{D}}}(\mathbb{G}_{\Psi_{\mathbb{G}}} I_n^{\text{SR}}) \right) \right) - \log \left(\mathbb{D}_{\Psi_{\mathbb{D}}}(I_n^{\text{HR}}) \right).$$

The discriminator generator reconstruction equation is given. This generator, $\mathbb{G}_{\Psi_{\mathbb{G}}} I^{\text{SR}}$, rebuilds the original picture I^{HR}. The reconstructed images are denoted by $\mathbb{D}_{\Psi_{\mathbb{D}}}(\mathbb{G}_{\Psi_{\mathbb{G}}} I^{\text{SR}})$ and $\mathbb{D}_{\Psi_{\mathbb{D}}}(I^{\text{HR}})$. This variable represents the number of target pictures. The variables l_{MSE} and $l_{\mathbb{G}}$ are defined as follows:

$$l_{\text{MSE}} = \frac{1}{r^2 HW} \sum_{x=1}^{W} \sum_{y=1}^{H} (I_{x,y}^{\text{HR}} - \mathbb{G}_{\Psi_{\mathbb{G}}}(I^{\text{LR}})_{x,y})^2.$$

$$l_{\mathbb{G}} = \sum_{n=1}^{N} -\log \mathbb{D}_{\Psi_{\mathbb{D}}}(\mathbb{G}_{\Psi_{\mathbb{G}}}(I_n^{\text{LR}})).$$

The researchers added a registration loss component to the model's loss function to improve high-frequency texture information recovery. The expected difference between

spatial transformation calculations and observations is denoted as RLT. The major goal is to minimize complex information loss during geometric translation of consecutive frames. This method helps restore the HR scan. The RLT loss function is as follows.

$$\text{RLT} = \sum_{i=\pm 1} \left\| I'^{\text{LR}}_{t+i} - I^{\text{LR}}_t \right\|^2 .$$

The equation described above represents the result obtained by applying the registration net to the picture I^{LR}_{t+i}. This process yields the image represented as I'^{LR}_{t+i}. The equation that expresses the length of the center of gravity (called also diameter in the Feret sense) of the pathological structure, symbolized as l_{loss}, is expressed in the following way:

$$l_{\text{loss}} = l_{\text{MSE}} + 10^{-6} l_{\mathbb{G}} + \varrho \text{RLT} .$$

The RLT weight coefficient, denoted by ϱ, has been assigned a value of 0.001 in accordance with the findings of the experiments. In relation to the notion of Wasserstein Generative Adversarial Networks (WGANs), it is feasible to omit the terms l_G and $l_{\mathbb{D}}$, leading to an adjustment of the loss function:

$$l_{\text{D}} = \frac{1}{N} \sum_{n=1}^{N} \mathbb{D}_{\Psi_{\mathbb{D}}} \left(\mathbb{G}_{\Psi_{\mathbb{G}}}(I^{\text{SR}}_n) \right) - \mathbb{D}_{\Psi_{\mathbb{D}}}(I^{\text{HR}}_n) .$$

5.2. Registration of MR scans

The registration net is demonstrated using a multi-scale approach, which has been successful in traditional methods [34]. The procedure requires the target frame (I^{LR}_t) and the surrounding frame ($I^{\text{LR}}_{t-R:t+R}$) as input. Pyramidal registration is used to train spatial transformation parameters for image motion correction. Two sets of pictures are registered separately through the registration layer for a three-frame input. The parameters of the net are optimized by the minimization of the mean-squared error between the converted frames and the target frames. This parameter is denoted by $\omega^*_{\delta,t+1}$. This learning technique enhances the neural network's motion correction on the image dataset.

$$\omega^*_{\delta,t+1} = \underset{\omega_{\delta,t+1}}{\text{argmin}} \left\| I^{\text{LR}}_t - I'^{\text{LR}}_t \right\|^2 .$$

The symbol I'^{LR}_t represents the registration layer's result after the registration procedure.

The order is acknowledged. Figure 1 depicts the network layer setup for registration. Traditional methods for modeling deformable registration have been shown effective using a multi-scale framework, as shown in [8], [42], and [49].

This study uses a strategy to obtain a spanning tree with minimal aggregate edge costs. Nodes $i \in P$ represent distinct items, such as pixels or groups of pixels. The
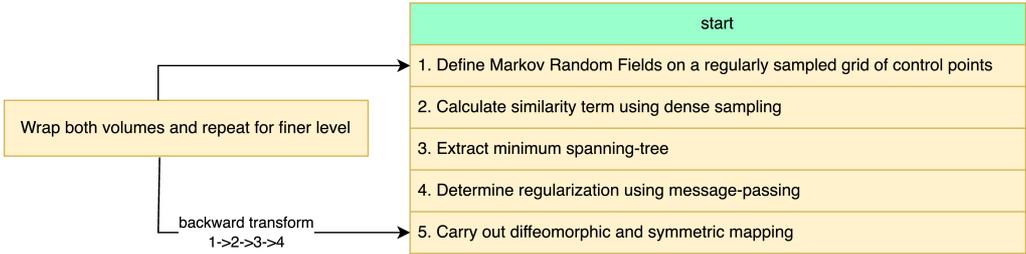
Fig. 5. The estimation of motion fields.

system links each node to a set of hidden labels representing motion fields, represented as $w_i^l = \{f_i^l, g_i^l, h_i^l\}$. The optimization-based energy function has two components: the data cost (S) and the pair-wise regularization cost $R(w_i^l, w_i^m)$, which applies to all nodes l connected to m.

$$E(w_i) = \sum_{j \in P} S(w_j^l) + \chi \sum_{l,m \in N} R(w_i^l, w_i^m). \tag{1}$$

The cost function estimates pixel similarity in two images. The parameter χ determines the influence of the regularization term and is used for weighting. The first element of the equation 1 is the data term, whereas the latter element is the regularization parameter.

The observed behavior is unaffected by adjacent entity displacements. The variable χ is used for weighting and determines the influence of the regularization term. In (1), the first component represents the data term and the second represents the regularization parameter.

5.3. The MR-PET images blur removal net

The work aims to restore a clear and accurate image, I_S, from a blurred image, I_B, without knowing the blur kernel. The deblurring process uses a convolutional neural network $(\mathbb{G}_{\rho_{\mathbb{G}}})$, also known as the Generator. An estimation determines the best I_S image for each I_B value. In addition, the critic network $(\mathbb{D}_{\rho_{\mathbb{D}}})$ is included in the training phase, and both networks engage in adversarial training. Integration of content and adversarial losses creates the composite loss function:

$$\mathcal{L} = \mathcal{L}_{\mathrm{GAN}} + \lambda \cdot \mathcal{L}_X.$$

In all experiments, λ was set to 100. This study does not condition the discriminator like in Isola et al. [19], because input-output discrepancies are not penalized. The loss function in the case of this GAN is defined as:

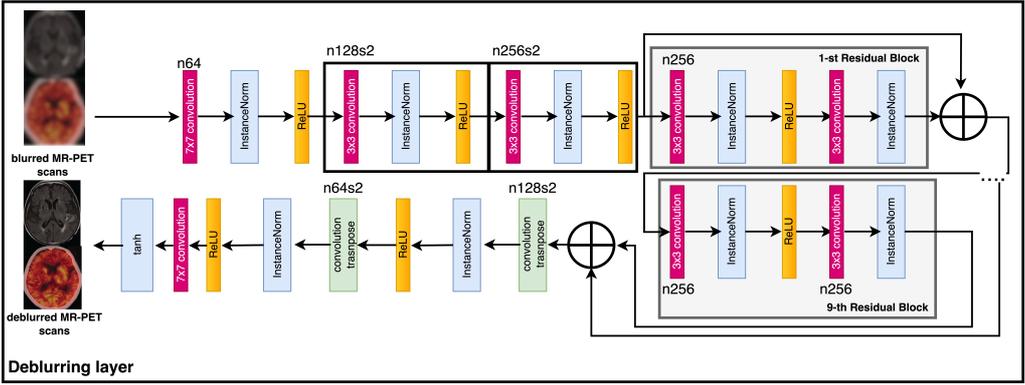$$\mathcal{L}_{\mathrm{GAN}} = \sum_{n=1}^{N} -\mathbb{D}_{\rho_{\mathbb{D}}} \left(\mathbb{G}_{\rho_{\mathbb{G}}} \left(I^B \right) \right).$$

Fig. 6. Deblurring net.

Common data loss functions include the $\mathbb{L}1$ or Mean Absolute Error (MAE) loss and the $\mathbb{L}2$ or Mean Squared Error (MSE) loss. Using these functions as the sole optimization aim produces uncertain abnormalities in images. According to [40], the observed irregularities are due to the mean value of plausible solutions at the pixel level within the pixel space. Using the $\mathbb{L}2$-loss technique in the perceptual loss function allows for mathematical expressions to calculate dissimilarity between the synthesized image and the reference image's CNN feature maps. The terminology is expressed as follows:

$$\mathcal{L}_X = \frac{1}{U_{k,n}B_{k,n}} \sum_{x=1}^{U_{k,n}} \sum_{y=1}^{B_{k,n}} \left(\emptyset_{k,n} \left(I^S \right)_{x,y} - \emptyset_{k,n} \left(\mathbb{G}_{\rho_\mathbb{G}} \left(I^B \right)_{x,y} \right) \right)^2 .$$

The symbol $\emptyset_{k,n}$ represents the feature map derived from the n-th convolution operation within a pre-trained network designed for MRI analysis [25]. The feature map is acquired subsequent to activation and prior to the k-th maxpooling layer. The variables $U_{k,n}$ and $B_{k,n}$ denote the dimensions of the feature maps.

5.4. MR-PET images denoising procedure

Magnitude pictures are the main representation in MRI-PET, making denoising difficult. The magnitude pictures are derived from real and imaginary components [10]. The presence of noise in magnitude images can be attributed to the Rician distribution, which exhibits a higher level of complexity compared to conventional additive noise. Denoising results depend on the model's precision. This is due to its ability to ignore the core physical process and change it via sample-based learning. The main goal of MRI-PET noise mitigation is to improve diagnostic image quality by reducing noise. Noise-corrupted MR-PET images are represented by x, while noise-free images are represented
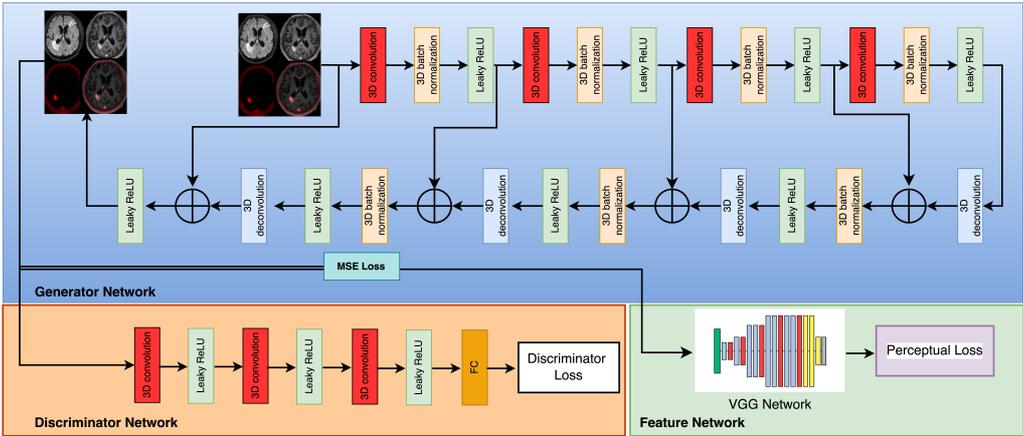
Fig. 7. Denoising net.

by y. Two matrices, x and y, with real-valued elements and the same dimensions, $m \times n$. Entities are connected as:

$$x = \varrho(y) .$$

The noise generation function is represented by the mapping function ϱ. Deep Learning is known for operating like a black box regardless of noise statistics. To optimize the denoising process of MR-PET, it is crucial to streamline the search for the most suitable approximation of the function ϱ^{-1}. The denoising technique entails the elimination of undesired noise from a provided signal or dataset.

$$\underset{f}{\operatorname{argmin}} \left\| \hat{y} - y \right\| .$$

The variable $\hat{y}$ reflects the anticipated value of y, based on the function $f(x)$, which provides the most accurate approximation of the inverse of ϱ.

According to statistical analysis, it can be inferred that samples x and y originate from distinct data distributions. Specifically, the variable x denotes the distribution of a noisy picture (P_n), while the variable y denotes the distribution of a noise-free image (P_{gen}). The denoising technique employs a mapping algorithm to alter the distribution. The function f establishes a mapping between samples drawn from the distribution P_n and a distribution denoted as $\mathbb{P}_{\text{gen}}$, which is identical to the actual data distribution P_r.

The discriminative model is specifically designed to differentiate between samples generated by a generative model and real data samples. The generative model utilizes the provided input sample to generate a novel sample that has a high degree of similarity

to the underlying data distribution.

$$\mathcal{L}_{\text{WGAN}}\left(\mathbb{D}\right) = -\,\mathbb{E}_{y\sim P_r}\left[\mathbb{D}\left(y\right)\right] + \mathbb{E}_{x\sim P_n}\left[\log\mathbb{D}\left(y\right)\right] + \mathbb{E}_{x\sim P_n}\left[D\left(G\left(x\right)\right)\right] + $$
$$+ \psi\mathbb{E}_{\hat{x}\sim P_{\hat{x}}}\left[\left(\left\|\nabla_{\hat{x}}D\left(\hat{x}\right)\right\|_2 - 1\right)^2\right]. \tag{2}$$

The final part of (2) is a gradient penalty factor, with ψ as a penalty coefficient. To construct the probability distribution, $P_{\hat{x}}$, points are uniformly sampled along straight lines from the actual data distribution P_r and generator distribution P_{gen}. Below is the loss function formulation for the generator $\mathbb{G}$:

$$\mathcal{L}_{\text{WGAN}}\left(G\right) = \mathbb{E}_{x\sim P_r}\left[\log\mathbb{D}\left(y\right)\right] + \mathbb{E}_{x\sim P_n}\left[\log\left(1 - D\left(G\left(x\right)\right)\right)\right].$$

In activities that need pixel-level adjustments, the Mean Squared Error (MSE) loss function is utilized rather frequently. The main goal is to reduce the differences between the original image and the generated image at a pixel level. The computation described above can be derived using the following methodology:

$$\mathcal{L}_{\text{MSE}} = \frac{1}{abc}||\mathbb{G}(x) - y||^2.$$

The variables a, b, and c represent the dimensions of the image. A recent study has demonstrated that the utilization of the Mean Squared Error (MSE) loss function has the potential to yield a substantial peak signal-to-noise ratio. Nevertheless, a decline in specificity, particularly with commonplace particulars, could have a substantial impact on clinical diagnosis [27].

The problem at hand is effectively tackled by the proposed loss function, which incorporates perceptual loss as documented in references [3], [10], and [22]. The utilization of a pre-existing neural network facilitates the extraction of pertinent data from both authentic and counterfeit photographic representations. Perceptual similarity quantifies the extent of dissimilarity in the attributes of reference and synthesized images. The next section provides an explanation of the perceptual loss function:

$$\mathcal{L}_{\text{perceptual}} = \frac{1}{abc}||\omega\left(\mathbb{G}(x)\right) - \omega\left(y\right)||_F^2.$$

The variable ω denotes the feature extractor, whereas a, b, and c denote the dimensions of the feature map. In this study, the VGG-19 network is employed for the purpose of extracting visual features [43]. The VGG-19 convolutional neural network consists of a total of nineteen layers, comprising sixteen convolutional layers and three fully connected layers. The scope of feature extraction is constrained to the initial sixteen layers. To implement the VGG network-based perceptual loss, the following procedures should be followed:

$$\mathcal{L}_{\text{VGG}} = \frac{1}{abc}||\text{VGG}\left(\mathbb{G}(x)\right) - \text{VGG}\left(y\right)||_F^2.$$

The generator $\mathbb{G}$ is coupled to a joint loss function that includes MSE, VGG, and discriminator losses.

The architectural design of the discriminator network, denoted as $\mathbb{D}$, is illustrated in Figure 7. Each of the model's three convolutional layers uses 32, 64, or 128 filters. A homogeneous kernel size of 3×3×3 was used to configure the convolution layers. The top layer is totally merged and gives a distinctive result. The pre-trained VGG-19 network extracts features. For further information, refer to the main source document [43]. Pan and Yang [39] found that transfer learning eliminates the need for network retraining for MR-PET scans, so

$$\mathcal{L}_{\text{RED-WGAN}} = \delta_1 \mathcal{L}_{\text{MSE}} + \delta_2 \mathcal{L}_{\text{VGG}} + \delta_3 \mathcal{L}_{\text{WGAN}}(G) .$$

The suggested RED-WGAN network configuration is shown in Figure 7. Three components make up the system: a generator network (denoted as $\mathbb{G}$), a discriminator network (denoted as $\mathbb{D}$), and a feature extractor (VGG network). Similar short connections connect the convolutional and deconvolutional layers. Each layer performs three-dimensional convolution, Leaky-ReLU activation, and batch normalization except for the final layer. The final layer only conducts 3D-convolution and Leaky-ReLU. This study uses a 3×3×3 kernel configuration with a filter sequence of 32, 64, 128, 256, 128, 64, 32, and 1.

6. The implementation of sparse sampling in MR/PET raw data

The methodology aims to consolidate all data from several modalities. According to [34], this study suggests that PET data volume can be compressed. To reduce readout channels, positron-emitting radioactive elements are mixed whenever possible. They produced a higher-resolution PET image by consolidating their output signals. MR/PET hybrid scanners integrate super-resolution and compressive sensing through structural components. Sparse depiction of the detector's structure is obvious. The above characteristic allows sparse-sense to generate novel multiplexing setups. Making meaningful sensing matrices is essential in computer science. Several stochastic approaches can generate constrained isometry random matrices. Maximal likelihood is used to build sensing matrices in the framework. Research indicates that using identified matrices results in the lowest mean-square reconstruction error [30]. This publication describes a method that uses a minimal number of channels to create discrete space and time domains. This interprets PET input data as compressed PET signals. Each readout is interpreted using a linear combination of photodetector pixels, with weights represented as $c_{k,n}$ and shown in Figure 5. The number of sensors was lowered via 4:1 subsampling. MR-PET joint sparsity and shared product characteristics enable motion model parameter generation from MR data sets to enrich PET pictures. Super-resolution is achieved using the same method as shown in Figure 1.

7. Results

The algorithm was proven effective by this investigation. The technique was tested using Biograph mMR-Simultaneous MR-PET scanner data. Therefore, two separate sets were created. For reference, the baseline dataset includes all active channels. In the subsequent dataset, a reduction of 15% in the number of channels is achieved by deactivating eight detectors that are uniformly distributed.

To prepare the dataset, data from MR-PET was organized into static frames that were free from motion artifacts. The algorithm proposed in this study was utilized to reconstruct an image for each frame. The normal clinical approach was employed, which entails conducting 3 iterations, dividing the data into 24 subgroups, and applying a 5×5 Gaussian post-smoothing technique. The reference/target picture for the reconstruction process was obtained from the initial static frame, while the successive static frames provided the source images for reconstruction.

During a simulated three-minute scan, participants assumed various positions and orientations of their heads at random time intervals. Each study demonstrated a distinct variety of movements, resulting in varying quantities of static frames.

To augment the insufficient quantity of data for constructing a sufficiently large dataset to train the neural network, additional picture volumes were synthesized from the five patient trials. A total of one hundred picture volumes of MR-PET were simulated by implementing random changes to the line-of-response (LOR) data. Thereafter, the raw dataset was subjected to transformation and histogramming, and then forwarded to the reconstruction method, as previously discussed, in order to generate image volumes. To mitigate the computational expenses involved with training the neural network, the images underwent a resizing process. Specifically, the dimensions of the images were modified from 400×400×109 to 128×128×96. This was achieved by cutting the backdrop, which consisted of voxels with zero values, and subsequently rescaling the resulting image. Cross-validation was employed, utilizing a 4:1 ratio for the allocation of training and test data. The training process consisted of 100 epochs, each consisting of 20 steps per epoch, and used a batch size of 4. The learning rate was set to 10^{-4}.

The super-resolution image reconstruction model was trained using an NVIDIA DGX device with a graphics processing unit dubbed A-100, utilizing resources from Google Colab Pro. The learning strategy employed by the generator involves the utilization of merged low-resolution (LR) images with dimensions of $60 \times 60 \times 2$. One high-resolution (HR) image of 240 by 240 units is the goal. Content, perceptual, and adversarial losses comprise the generator's loss function. MRI images in HR and SR are shown to the Discriminator. The discriminator, or binary classifier, optimizes using binary cross-entropy. The Adam optimizer optimizes the generator and discriminator. Tensorflow and Keras are utilized to implement the proposed network in Python. The network for super resolution image reconstruction has undergone training for a total of 50 epochs.

The Adam optimizer is utilized by both the Generator and the Discriminator to optimize the network parameters. The initial learning rate for Adam is 0.001, which is then reduced to 0.000001 after half the number of epochs.

The generation process incorporates three distinct forms of losses, including content loss, perceptual loss, and adversarial loss. The Discriminator utilizes both the SR MRI image and the HR image. The classifier in question is a binary classifier that employs binary cross-entropy as a means of optimization. The utilized optimizer is an Adam optimizer, which is applied to both the generator and discriminator components. Moreover, the MR-PET dataset comprised 1000 pairs of pictures, each containing both blurred and sharp images, with a resolution of 640×640. The deblurring network that has been shown has demonstrated exceptional performance in terms of structured self-similarity. It is comparable to the current state-of-the-art in terms of peak signal-to-noise ratio and offers visually appealing outcomes. The utilization of L2 distance in pixel space is not employed by the network, hence missing direct optimization for the PSNR measure.

The denoising model was trained with 1% and 4% noise added due to the limited understanding of the noise level in the actual data. In this study, the VGG network used for low resolution MR-PET image reconstruction was pre-trained using the Medical ImageNet dataset. The convolutional output of the VGG16 model was utilized as the encoded embedding of the de-aliased output and the ground truth. Subsequently, the mean squared error (MSE) was calculated between these two outputs. Create distinct networks for various undersampling ratios using the constant mutual hyperparameters: $\alpha = 10$, $\beta = 0.2$, $\theta = 0.003$, starting learning rate of 0.0001, batch size of 30. It is worth mentioning that the hyperparameters, namely α, β and θ represent the weights assigned to various loss components throughout the training process. The Adam optimization algorithm was employed, utilizing a momentum value of 0.4. The learning of each model was conducted via early stopping, with the learning rate being reduced by half every 4 epochs. The MR-PET reconstruction model proposed in this study demonstrates robustness and requires minimal parameter adjustment. Consequently, we employed the same hyperparameters for subsequent tests, employing different undersampling ratios, varied undersampling masks, and both with and without noise.

The models were implemented using a high-level Python wrapper called Tensor-Layer [9].

Subsampled sinograms are divided into two pieces with sparse orthogonal domains. A hybrid conjugate gradient method was used to iteratively recreate the PS sinogram. A system of equations was solved using blocked relaxations. Reducing component total variation (TV) improves the piece-wise smooth model of the initial component. After integrating the two pieces, the sinogram was created and used to enhance the PS sinogram. This method produces quantified PET images with fewer readout channels. The evaluation distinguishes two information groups. The first thing that needed to be accomplished was to evaluate the super-resolution (SR) image reconstruction technique by

directly comparing it to native and naive approaches. Investigating the precision and dependability of the magnetic resonance (MR) sample was one of the secondary objectives of this investigation. In the present investigation, both in vivo and phantom data were evaluated. The outcomes of the simulation are presented in Figure 8. Compressed Sensing, conjugate symmetry, and the Partial Fourier method speed up data collection while preserving the unique k-space trajectories. The present framework module aims to integrate compressed sensing and super-resolution into MRI scanners. This work used phantom input files to demonstrate compressed sensing (CS) challenges for magnetic resonance imaging (MRI).

8. Discussion

Over the past decade, MR-PET and other integrated scanning technologies have grown in importance. Understanding the purpose of these tools is the first step to becoming a notable figure in the field. Even with motion blur, the study's strategy reduced artifacts from insufficiently sampled data. This article describes a new super-resolution technique for high-sensitivity compressed MR/PET signals. As expected, the algorithm improves image resolution without changing the technology.

Table 1 shows that CS quality ratios affect the Peak Signal-to-Noise Ratio (PSNR) values. The PSNR is calculated using multiple methodologies using ground truth images. One hundred simulations were run. To establish statistical significance of quality measures for each simulation scenario, the PSNR was iteratively calculated and averaged. The best results were achieved using a 50% compression ratio. Reducing the number of input samples improves the PSNR, as seen in Tables 2 and 3. The symbols N, M, SD, $t(99), p$ refer to: number of tests performed, mean value, standard deviation, t-value with the confidence level of 99 percent, p-value, respectively. The examination duration decreases directly with this value. Motion distortions can be reduced while sacrificing resolution with this strategy. Rapid convergence, picture prior, and blur kernel detection are prioritized in this method.

Preliminary trial data can provide context for efficient test completion. Motion estimation techniques may reduce diagnostic imaging image artifacts, improving diagnosis accuracy. Figures 8 and 9 show improved result resolution and quality. The above results are preliminary and may change. A qualitative study of twenty patients' neuroimages showed the algorithm's benefits. Using a combined MR/PET scanner, 30 oncological patients provided PET and MRI data. The research used 30 simulated brain PET data volumes and patient model MR scans in phantom studies. The Peak Signal-to-Noise Ratio for each reconstruction procedure was calculated quantitatively.

To test the null hypothesis, a t-test was used to compare image quality ratings of images reconstructed from highly sparse sampling spaces using the proposed method and completely sampled ground truth images. A radiologist found that the offered technique
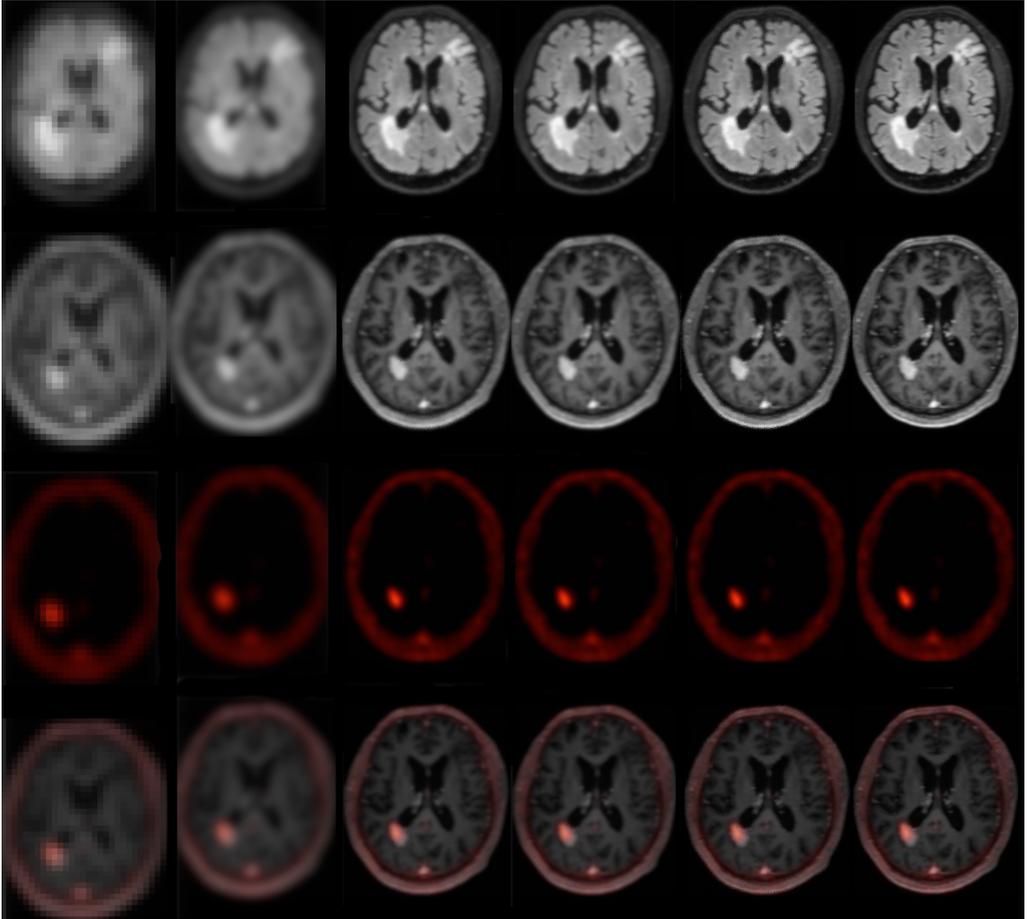
Fig. 8. The following is an illustration of a clinical trial. Images are numbered in a horizontal direction, from left to right. The present study involves the reconstruction of an image using the regular sampling scheme, without motion correction and SRR applied (1), B-spline curve (2) and Yang's method (3) were employed for the reconstruction process, along with Lim's method (4), Zhang's procedure (5), Kim's algorithm (6). Additionally, a proposed sampling scheme and motion compensation were utilized for super-resolution purposes. The aforementioned techniques were applied without introducing any additional information. The compression ratio is 50%. See Figure 9 for more results.
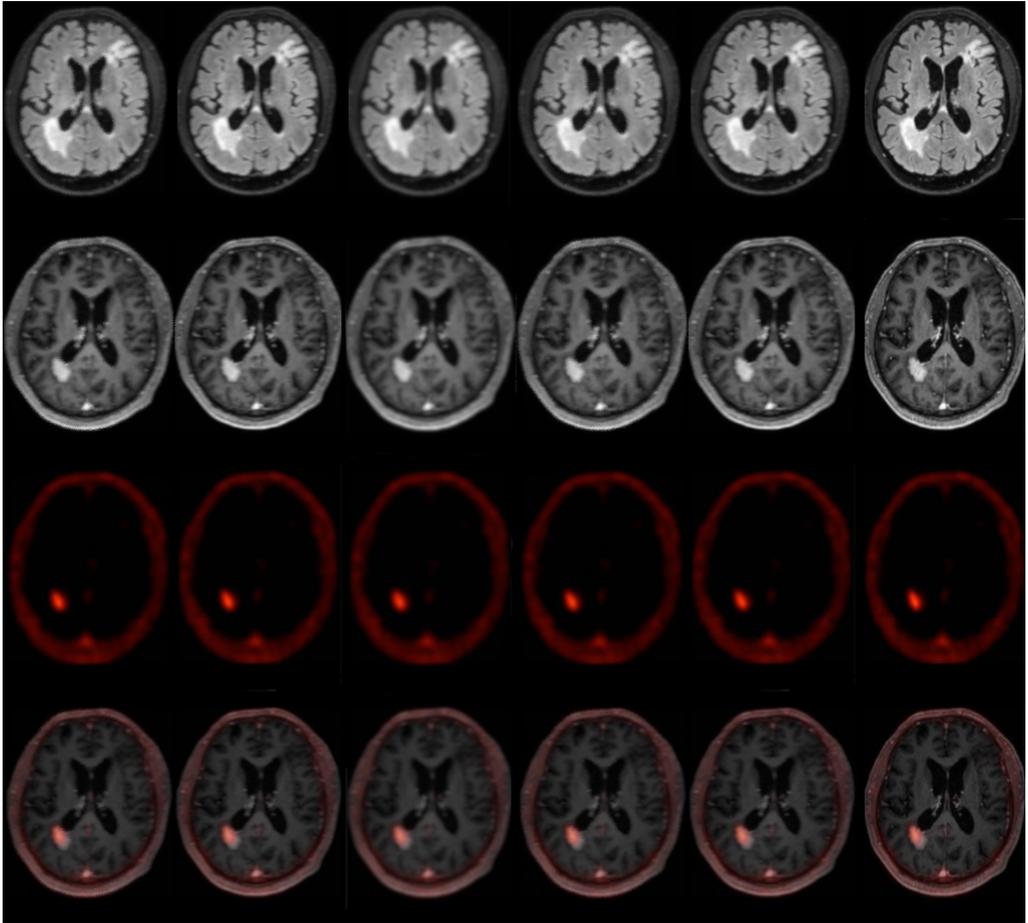
Fig. 9. The following is an illustration of a clinical trial (continued). Images are numbered in a horizontal direction, from left to right. The present study involves the reconstruction of an image using Mahapatra's method (7), Liu et al. procedure (8). Moreover, Dong's method (9) and Pham's method (10) were employed for the reconstruction process, along with Shi method (11), and the author's method (12) (see Tab. 2 for references). Additionally, a proposed sampling scheme and motion compensation were utilized for super-resolution purposes. The aforementioned techniques were applied without introducing any additional information. The compression ratio is 50%.

Tab. 1. Results of measuring the performance of the algorithm using various raw data sampling schemes on the data presented in Figure 9.

Raw data sampling* [%]	PSNR [dB]	N	M	SD	t(99)	p-value
20	26.76	100	26.76	0.04	0.322	0.143
40	32.33	100	32.33	0.05	-0.274	0.147
60	33.98	100	33.98	0.03	-1.299	0.191
80	34.16	100	34.16	0.02	-0.643	0.056
100	34.88	100	34.88	0.06	1.001	0.064

*After making a comparison with scans that have been fully sampled, the phrase refers to the percentage of the input samples that are still present. For example, a ratio of sixty reveals that forty percent of the samples gleaned through a comprehensive examination were discarded.

Tab. 2. The statistical parameters associated with the efficiency metrics of the model depicted in Figs. 8 and 9.

High resolution reconstruction method	PSNR [dB]	N	M	SD	t(99)	p-value
no SRR, no MC	24.88	100	24.88	0.04	0.265	0.533
B spline curve	26.01	100	26.01	0.03	-0.321	0.399
Yang et al. [47]	29.80	100	29.80	0.02	-0.928	0.294
Lim et al. [29]	29.67	100	29.67	0.02	-0.912	0.232
Zhang et al. [49]	31.01	100	31.01	0.03	-0.023	0.349
Kim et al. [25]	31.03	100	31.03	0.04	-0.903	0.211
Mahapatra et al. [32]	30.01	100	30.01	0.04	-0.429	0.473
Liu et al. [30]	29.93	100	29.93	0.05	-1.003	0.321
Dong et al. [7]	28.55	100	28.55	0.04	-1.003	0.218
Pham et al. [23]	31.66	100	31.66	0.02	-1.003	0.362
Shi et al. [42]	32.03	100	32.03	0.03	-1.003	0.412
The suggested SRR algorithm	33.98	100	33.98	0.03	-0.992	0.102

had better image metric values and better anatomical structure representation than the other algorithms. Additionally, a stringent Bowker symmetry test was used to examine image quality disparities. Markov random field (MRF) optimization allows structural representation-based registration approaches to be developed. The target registration

Tab. 3. The present work aims to evaluate the efficacy of various reconstruction techniques for in-vivo brain imaging. Specifically, it investigates the use of motion correction (MC) and upscaling (HR) using the method given by the author. See Figs. 8 and 9 for the pertinent brain pictures. The Peak Signal-to-Noise Ratio (PSNR) values corresponding to the four scenarios stated before are displayed in the second column of Table 1.

input	Sparse-sampling ratio [%]	MC	SRR	PSNR [dB]	N	M	SD	$t(99)$	p-value
LR	50	No	No	25.77	100	25.77	0.03	-1.211	0.198
LR	50	Yes	No	26.65	100	26.65	0.03	-1.045	0.245
HR	50	Yes	No	28.19	100	28.19	0.03	-1.217	0.193
SR	50	Yes	Yes	33.98	100	33.98	0.03	-1.299	0.191

error (TRE) measure was used to evaluate motion estimation algorithms:

$$\text{TRE} = \frac{1}{N} \sum_{i=1}^{N} \sqrt{(T_{L_x}^i - T_{\mathbb{D}_x}^i)^2 + (T_{L_y}^i - T_{\mathbb{D}_y}^i)^2 + (T_{L_z}^i - T_{\mathbb{D}_z}^i)^2} \,.$$

Using a linear combination of radial basis functions as a ground truth, the compelled deformation is T_L. In contrast, the motion estimation algorithms shown in Table 4 determine the deformation parameters of $T_{\mathbb{D}}$. The variable N represents the number of landmarks manually identified following medical expert guidance. The author's MC algorithm was tested using multiple image registration methods to find the most efficient one. All research registration procedures were statistically analyzed using mean and standard deviation. Table 4 shows the t-test TRE and p-values. Statistical study shows that the author's strategy differs significantly from alternative strategies with p-values below 0.002. Target Registration Error (TRE) mean and standard deviation values of 1.4 and 0.2 voxel were obtained by registering input and output images. The aforementioned values outperformed alternative strategies, as seen in Table 4.

To evaluate motion correction results for each participant, the disparity between images with motion and reference images without motion parameters was calculated. The statistical significance of data changes with and without artificial motion was examined using paired t-tests. Statistical analyses show the improvement is significant. Please refer to tables 2-3 for a comparison of the proposed approach to existing image resolution enhancement algorithms. This study used numerous motion registration algorithms, i.e., the methods of Wachinger et al. [45], Groppe et al. [13], Jenkinson et al. [20], Yang et al. [47], Greve et al. [12], Kadipasaoglu et al. [23], MIND [15], Branco et al. [2], as well as WGAN deformable registration procedure, i.e., the author's procedure. This approach has been expanded to incorporate directions on handling erroneous target images from a clinical scanner with better resolution to improve its viability. Super-resolution (SR) has been used to identify the relationship between low- and high-resolution scanner image domains.

Tab. 4. Statistical parameters of several registration methods in relation to the implemented technique.

Motion compensation procedure	TRE [voxels]		
	M	SD	p-value
not applied	4,90	2,60	<0,002
Wachinger et al. [45]	2,72	0,78	<0,005
Groppe et al. [13]	2,41	0,27	<0,005
Jenkinson et al. [20]	3,55	0,37	<0,004
Yang's et al. [47]	2,01	0,37	<0,004
Greve et al. [12]	3,01	0,29	<0,006
Kadipasaoglu et al. [23]	1,66	0,31	<0,003
MIND [15]	1,82	0,19	<0,004
Branco et al. [2]	1,73	0,16	<0,009
WGAN deformable MC – the author's method	1,40	0,17	<0,002

This study involved an evaluation of the author's methods in comparison to several advanced super resolution image reconstruction algorithms. The current study focuses on the reconstruction of an image using a regular sampling scheme, without the application of motion correction and super-resolution reconstruction (SRR) techniques. The reconstruction process involves the utilization of B-spline curve and Yang's method [47], as well as Lim's method [29], Zhang's procedure [49], Kim's algorithm [25], and Liu et al.'s procedure [30]. Furthermore, the reconstruction procedure utilized Dong's approach [7], Pham's method [23], Shi's method [42], and the author's method.

This study improved neural network training to accurately map low-resolution magnetic resonance and positron emission tomography images to ground-truth subimages. Its potential benefits, particularly its ability to create high contrast and resolution, are the main reason for its expected success. MR/PET technology will be integrated with sparsely sampled input data super resolution image reconstruction techniques in this study. Budgetary issues hinder this methodology's performance and these issues should be considered. The long-term balancing of MR/PET's increased expenses is unknown. Two data sets were used in the experiment. Experiments compared the compressively sensed super-resolution picture reconstruction approach to simpler and less advanced methods. The magnetic resonance (MR) sample design's efficacy was the study's secondary goal. The proposed reconstruction approach and alternative algorithms were

used at different compression rates in MRI, with subjective and objective picture evaluations. This study examined in vivo and phantom inputs. These images show computer-generated model results. Compared to unmodified k-space trajectories, compressed sensing, conjugate symmetry, and Partial-Fourier (PF) technologies accelerate data collecting. Compressed sensing (CS) can reconstitute sparse signals by projecting them into a low-dimensional linear subspace. The theoretical certainty of the methodology given gives it great promise. The study uses k-space modifications and Generative Adversarial Networks (GANs) in the image domain. Generative Adversarial Networks (GANs) can include image-specific prior knowledge. In the picture and k-space domains, iterative calculations using Wasserstein Generative Adversarial Networks (WGANs) and k-space correction approaches are used. The method shows potential in tackling k-space rectification error. Compared to other methods, the analyzed strategy reconstructs images with better quality and fewer aliasing artifacts. The suggested method reduces aliasing artifacts better than existing and non-iterative methods. No matter the sample frequency for Cartesian and radial sampling masks, the suggested approach has a higher peak signal-to-noise ratio than the others.

In addition, the study used empirical data in the form of Magnetic Resonance images, which contain actual values rather than true k-space data from MRI scans. Pictures with complex numerical values are the theme. In Generative Adversarial Networks, a fake connection between input and output layers is essential. Preprocessing is essential for complex data sets. T1-weighted images and other magnetic resonance imaging techniques will be used to measure clinical value and examine radiologists. This study proposes changes to increase image quality and reduce data collection time. Even with misregistration distortions, the proposed technique can eliminate sparse data artifacts. The strategy uses compressed sensing, raw data sparsity, and super-resolution reconstruction to improve k-space filling efficiency or fidelity. Image complexity decreases as MR/PET picture fidelity improves. Edge representation improves with higher high-frequency component sampling rates. The technique reviewed shows promise for hybrid scanner integration without hardware adjustments.

The reconstruction approaches for MR/PET use either whole raw data or pre-existing data as reference standards. Refer to Tables 2 through 3 for a comprehensive review of clinical trials where acceleration factors of up to 2 led to diagnostically viable scans and radiologists acknowledged the higher resolution. The main outcomes of this study encompass the subsequent findings:

- The algorithm described in the framework showcases a comprehensive methodology for the collaborative reconstruction of MR-PET data. This work places significant emphasis on several critical areas, including sparse sampling trajectories, synchronization of k subspaces, deblurring, noise reduction, motion correction, and ultimately, enhancing the resolution of a picture.

- The current investigation presents an innovative framework for the reconstruction of MR-PET images through the utilization of a generative super-resolution methodology.
- The provided methodology utilizes the combined sparsity of both the MR and PET modalities.
- The limited availability of MR and PET raw data has led to an increase in the rate at which the input data is processed.
- The methodology has been specifically designed for the purpose of collecting visual data at various scales. Other authors frequently oversimplify this matter.
- The system demonstrates the ability to extract visual cues across various scales. Other writers often oversimplify this topic issue.
- The technology used involves certain preprocessing phases to tackle the difficulties associated with blur and noise removal layers.
- The suggested method utilizes a neural network-based reconstruction algorithm for magnetic resonance imaging. The objective of this approach is to recover images of poor quality that are obtained from extremely limited raw data.
- The methodology described above utilizes the compressed sensing framework in order to prioritize the effort to minimize the duration of data collecting.
- The author's deformable motion estimation approach is buried within the reconstruction layer of the procedure.

The presented system uses compressed raw data, an advanced SR-GAN architecture, and a denoising module to pre-process low-resolution MR-PET images. The network can super-resolve low-resolution and noisy MR-PET images and recreate high-resolution MR images. The methods offered helps solve a problem where artifacts and noise diminish the peak signal-to-noise ratio in MR-PET images, reducing the generative adversarial network's effectiveness. The proposed solution yields better picture reconstruction quality than previous methods, as shown by empirical data. Therefore, this can improve diagnostic procedure suggestions for healthcare providers.

References

[1] G. Antoch and A. Bockisch. Combined PET/MRI: a new dimension in whole-body oncology imaging? *European Journal of Nuclear Medicine and Molecular Imaging*, 36(S1):113–120, 2008. doi:10.1007/s00259-008-0951-6.

[2] M. P. Branco, A. Gaglianese, D. R. Glen, D. Hermes, Z. S. Saad, et al. ALICE: A tool for automatic localization of intra-cranial electrodes for clinical and high-density grids. *Journal of Neuroscience Methods*, 301:43–51, 2018. doi:10.1016/j.jneumeth.2017.10.022.

[3] J. Bruna, P. Sprechmann, and Y. LeCun. Super-resolution with deep convolutional sufficient statistics. In: *Proc. Int. Conf. Learning Representation (ICLR)*, 2015. Proceedings published in arXiv, see https://iclr.cc/archive/www/doku.php%3Fid=iclr2015:accepted-main.html. doi:10.48550/arXiv.1412.7022.

[4] H. Chang, D. Y. Yeung, and Y. Xiong. Super-resolution through neighbor embedding. In: *Proc. 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1. Washington, USA, 2004. doi:10.1109/CVPR.2004.1315043.

[5] Y. Chen, F. Shi, A. G. Christodoulou, Y. Xie, Z. Zhou, et al. Efficient and accurate MRI super-resolution using a generative adversarial network and 3D multi-level densely connected network. In: A. F. Frangi, J. A. Schnabel, C. Davatzikos, C. Alberola-López, and G. Fichtinger, eds., *Proc. Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI) 2018*, vol. 11070 of *Lecture Notes in Computer Sciences*, pp. 91–99. Springer International Publishing, Granada, Spain, 16-20 Sep 2018. doi:10.1007/978-3-030-00928-1_11.

[6] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095, 2007. doi:10.1109/TIP.2007.901238.

[7] C. Dong, C. C. Loy, K. He, and X. Tang. Learning a deep convolutional network for image super-resolution. In: *Proc. European Conference on Computer Vision (ECCV)*, vol. 8692 of *Lecture Notes in Computer Science*, p. 184–199. Springer, Zurich, Switzerland, 6-12 Sep 2014. doi:10.1007/978-3-319-10593-2_13.

[8] C. Dong, C. C. Loy, and X. Tang. Accelerating the super-resolution convolutional neural network. In: *Proc. European Conference on Computer Vision (ECCV)*, vol. 9906 of *Lecture Notes in Computer Science*, p. 391–407. Springer, Amsterdam, The Netherlands, 2016. doi:10.1007/978-3-319-46475-6_25.

[9] H. Dong, A. Supratak, L. Mai, F. Liu, A. Oehmichen, et al. TensorLayer: A versatile library for efficient deep learning development. In: *Proc. 25th ACM International Conference on Multimedia*, MM '17, p. 1201–1204. Association for Computing Machinery, 2017. doi:10.1145/3123266.3129391.

[10] L. A. Gatys, A. S. Ecker, and M. Bethge. Texture synthesis using convolutional neural networks. In: *Advances in Neural Information Processing Systems 28 – Proc. NIPS 2015*, vol. 28 of *NeurIPS Proceedings*, pp. 262–270, 2015. https://proceedings.neurips.cc/paper/2015/hash/a5e00132373a7031000fd987a3c9f87b-Abstract.html.

[11] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, et al. Generative adversarial nets. In: *Advances in Neural Information Processing Systems 27 – Proc. NIPS 2014*, vol. 27 of *NeurIPS Proceedings*, pp. 2672–2680, 2014. https://proceedings.neurips.cc/paper/2015/hash/a5e00132373a7031000fd987a3c9f87b-Abstract.html.

[12] D. N. Greve and B. Fischl. Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, 48(1):63–72, 2009. doi:10.1016/j.neuroimage.2009.06.060.

[13] D. M. Groppe, S. Bickel, A. R. Dykstra, X. Wang, P. Mégevand, et al. iELVis: An open source MATLAB toolbox for localizing and visualizing human intracranial electrode data. *Journal of Neuroscience Methods*, 281:40–48, 2017. doi:10.1016/j.jneumeth.2017.01.022.

[14] J. Gu, H. Lu, W. Zuo, and C. Dong. Blind super-resolution with iterative kernel correction. In: *Proc. 2019 IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, p. 1604–1613. Long Beach, CA, USA, 2019. doi:10.1109/CVPR.2019.00170.

[15] M. P. Heinrich, M. Jenkinson, M. Bhushan, T. Matin, F. V. Gleeson, et al. MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration. *Medical Image Analysis*, 16(7):1423–1435, 2012. Special Issue on the 2011 Conference on Medical Image Computing and Computer Assisted Intervention. doi:10.1016/j.media.2012.05.008.

[16] X. Hu, H. Mu, X. Zhang, Z. Wang, T. Tan, et al. Meta-SR: A magnification-arbitrary network for super-resolution. In: *Proc. 2019 IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, p. 1575–1584. Long Beach, CA, USA, 15-20 Jun 2019. doi:10.1109/CVPR.2019.00167.

[17] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger. Densely Connected Convolutional Networks. In: *Proc. 2017 IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, p. 4700–4708. Honolulu, HI, USA, 21-26 Jul 2017. doi:10.1109/CVPR.2017.243.

[18] C. M. Hyun, H. P. Kim, S. M. Lee, S. Lee, and J. K. Seo. Deep learning for undersampled MRI reconstruction. *Physics in Medicine & Biology*, 63(13):135007, 2018. doi:10.1088/1361-6560/aac71a.

[19] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. In: *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 5967–5976. Honolulu, HI, USA, 2017. doi:10.1109/CVPR.2017.632.

[20] M. Jenkinson, P. Bannister, M. Brady, and S. Smith. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2):825–841, 2002. doi:10.1006/nimg.2002.1132.

[21] K. Jiang, Z. Wang, P. Yi, G. Wang, T. Lu, et al. Edge-enhanced GAN for remote sensing image superresolution. *IEEE Transactions on Geoscience and Remote Sensing*, 57(8):5799–5812, 2019. doi:10.1109/TGRS.2019.2902431.

[22] J. Johnson, A. Alahi, and F. F. Li. Perceptual losses for real-time style transfer and super-resolution. In: *Proc. Eur. Conf. Computer Vision (ECCV)*, vol. 9906 of *Lecture Notes in Computer Science*, pp. 694–711. Springer, 2016. doi:10.1007/978-3-319-46475-6_43.

[23] C. M. Kadipasaoglu, C. Morse, K. Pham, C. Donos, and N. Tandon. SAMCOR: A robust and precise co-registration algorithm for brain CT and MR imaging. *Interdisciplinary Neurosurgery*, 30:101637, 2022. doi:10.1016/j.inat.2022.101637.

[24] K. Kataoka, Y. Shiraishi, Y. Takeda, S. Sakata, M. Matsumoto, et al. Aberrant PD-L1 expression through 3'-UTR disruption in multiple cancers. *Nature*, 534(7607):402–406, 2016. doi:10.1038/nature18294.

[25] J. Kim, J. K. Lee, and K. M. Lee. Deeply-recursive convolutional network for image super-resolution. In: *Proc. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 1637–1645. Las Vegas, NV, USA, 27 Jun 2016. doi:10.1109/CVPR.2016.181.

[26] K. Kulkarni, S. Lohit, P. Turaga, R. Kerviche, and A. Ashok. ReconNet: Non-iterative reconstruction of images from compressively sensed measurements. In: *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 449–458. Las Vegas, NV, USA, 27 Jun 2016. doi:10.1109/CVPR.2016.55.

[27] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, et al. Photo-realistic single image super-resolution using a generative adversarial network. In: *Proc. 2017 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 105–114. Honolulu, HI, USA, 2017. doi:10.1109/CVPR.2017.19.

[28] K. A. Lidke, B. Rieger, T. M. Jovin, and R. Heintzmann. Superresolution by localization of quantum dots using blinking statistics. *Optics Express*, 13(18):7052–7062, 2005. doi:10.1364/OPEX.13.007052.

[29] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee. Enhanced deep residual networks for single image super-resolution. In: *Proc. 2017 IEEE Conf. Computer Vision and Pattern Recognition Workshops (CVPRW)*, p. 1132–1140. Honolulu, HI, USA, 21-26 Jul 2017. doi:10.1109/CVPRW.2017.151.

[30] C. Liu and D. Sun. On Bayesian adaptive video super resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(2):346–360, 2014. doi:10.1109/TPAMI.2013.127.

[31] D. Liu, Z. Wang, Y. Fan, X. Liu, Z. Wang, et al. Robust video super-resolution with learned temporal dynamics. In: *Proc. 2017 IEEE Int. Conf. Computer Vision (ICCV)*, pp. 2526–2534. Venice, Italy, 22-29 Oct 2017. doi:10.1109/ICCV.2017.274.

[32] D. Mahapatra, B. Bozorgtabar, and R. Garnavi. Image super-resolution using progressive generative adversarial networks for medical image analysis. *Computerized Medical Imaging and Graphics*, 71:30–39, 2019. doi:j.compmedimag.2018.10.005.

[33] K. Malczewski. Image resolution enhancement of highly compressively sensed CT/PET signals. *Algorithms*, 13(5), 2020. doi:10.3390/a13050129.

[34] K. Malczewski. Super-resolution with compressively sensed MR/PET signals at its input. *Informatics in Medicine Unlocked*, 18, 2020. doi:10.1016/j.imu.2020.100302.

[35] K. Malczewski. Diffusion weighted imaging super-resolution algorithm for highly sparse raw data sequences. *Sensors*, 23(12):5698, 2023. doi:10.3390/s23125698.

[36] K. Malczewski and R. Stasinski. High resolution MRI image reconstruction from a PROPELLER data set of samples. *International Journal of Functional Informatics and Personalised Medicine*, 1(3), 2008. doi:10.1504/IJFIPM.2008.021394.

[37] A. Mousavi, A. B. Patel, and R. G. Baraniuk. A deep learning approach to structured signal recovery. In: *Proc. 2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. Monticello, IL, USA, 29 Sep – 02 Oct 2015.

[38] K. Ning, Z. Zhang, K. Han, S. Han, and X. Zhang. Multi-frame super-resolution algorithm based on a WGAN. *IEEE Access*, 9:85839–85851, 2021. doi:10.1109/ACCESS.2021.3088128.

[39] S. J. Pan and Q. Yang. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10):1345–1359, 2010. doi:10.1109/TKDE.2009.191.

[40] H. Pedersen, S. Kozerke, S. Ringgaard, et al. k-t PCA: temporally constrained k-t BLAST reconstruction using principal component analysis. *Magnetic Resonance in Medicine*, 62(3):706–716, 2009. doi:10.1002/mrm.22052.

[41] G. Qian, Y. Wang, J. Gu, C. Dong, W. Heidrich, et al. Rethinking learning-based demosaicing, denoising, and super-resolution pipeline. In: *Proc. 2022 IEEE Int. Conf. Computational Photography (ICCP)*, pp. 1–12. Pasadena, CA, USA, 1-5 Aug 2022. doi:10.1109/ICCP54855.2022.9887682.

[42] W. Shi, J. Caballero, F. Huszar, J. Totz, A. P. Aitken, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: *Proc. 2016 IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, p. 1874–1883. Las Vegas, NV, USA, 27-30 Jun 2016. doi:10.1109/CVPR.2016.207.

[43] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In: *Proc. Int. Conf. Learning Representation (ICLR)*, 2015. Proceedings published in arXiv, see `https://iclr.cc/archive/www/doku.php%3Fid=iclr2015:accepted-main.html`. doi:10.48550/arXiv.1409.1556.

[44] Y. Tai, J. Yang, and X. Liu. Image super-resolution via deep recursive residual network. *Proc. 2017 IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 2790–2798, 21-26 Jul 2017. doi:10.1109/CVPR.2017.298.

[45] C. Wachinger and N. Navab. Entropy and Laplacian images: Structural representations for multimodal registration. *Medical Image Analysis*, 16(1):1–17, 2012. doi:10.1016/j.media.2011.03.001.

[46] A. W. A. Wahab, M. A. Bagiwa, M. Y. I. Idris, S. Khan, Z. Razak, et al. Passive video forgery detection techniques: A survey. In: *Proc. 2014 10th Int. Conf. Information Assurance and Security (IAS)*, pp. 29–34. Okinawa, Japan, 28-30 Nov 2014. doi:10.1109/ISIAS.2014.7064616.

[47] F. Yang, M. Ding, X. Zhang, W. Hou, and C. Zhong. Non-rigid multi-modal medical image registration by combining L-BFGS-B with cat swarm optimization. *Information Sciences*, 316:440–456, 2015. doi:10.1016/j.ins.2014.10.051.

[48] Z. Zhang, D. Gao, X. Xie, and G. Shi. Dual-channel reconstruction network for image compressive sensing. *Sensors*, 19:2549, 2019. doi:10.3390/s19112549.

[49] Z. Zhang, Z. Wang, Z. Lin, and H. Qi. Image super-resolution by neural texture transfer. In: *Proc. 2019 IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, p. 7974–7983. Long Beach, CA, USA, 15-20 Jun 2019. doi:10.1109/CVPR.2019.00817.