Deep-ASL enhancement technique in arterial spin labeling MRI – A novel approach for the error reduction of partial volume correction technique with linear regression algorithm☆

Shyna A [a,*], Ushadevi Amma C. [b], Ansamma John [a], C. Kesavadas [c], Bejoy Thomas [c]

[a] Department of Computer Science and Engineering, ThangalKunju Musaliar College of Engineering, APJ Abdul Kalam Technological University, Kerala, India
[b] Department of Electronics and Communications Engineering, Amrita Vishwa Vidyapeetham, Amritapuri Campus, Kerala, India
[c] Department of Imaging Sciences and Interventional Radiology, Sree Chitra Tirunal Institute of Medical Sciences and Technology, Trivandrum, Kerala, India

ARTICLE INFO

ABSTRACT

Arterial Spin Labelling MRI is a noninvasive quantitative imaging technique for measuring Cerebral Blood Flow (CBF) that plays a vital role in diagnosing different neurological disorders. Limited signal-to-noise ratio and significant partial volume effect due to the low resolution of ASL images make an accurate CBF estimation difficult. This work proposes a deep learning based ASL enhancement algorithm (Deep-ASL ENHANCE), based on the principle of single image super resolution and multi-loss joint strategy with two reconstruction modules and one weighted fusion module that employ residual dense block as the basic building block. Lack of huge amount of low quality and high quality images for training this deep learning network, is addressed by generating simulated ASL images from structural images of ADNI2. The experiment is conducted and results are evaluated on a simulated dataset in terms of different metrics such as RMSE, PSNR and SSIM. The model is also validated using clinical ASL images with the help of two independent radiologists and the results are compared using Visual Quality Score (VQS). The deep learning model trained by using simulated ASL images shows more promising results on clinical ASL data. The effectiveness of using Deep-ASL ENHANCE as a preprocessing step to the partial volume correction technique with Linear Regression algorithm (LR) has been investigated using RMSE score and it is found that CBF quantification accuracy is improved compared to the standalone LR algorithm.

1. Introduction

Hundreds of millions of people worldwide are affected by neurological disorders which are one of the most leading causes of morbidity and mortality. Among the different medical imaging modalities, Magnetic Resonance Imaging (MRI) has become a more promising approach for the diagnosis of different neurological disorders. Rapid technological advancements in modern medical imaging provide techniques for quantitative imaging for functional tissue characteristics such as perfusion and diffusion which enables better diagnosis and treatment planning. Perfusion is an important parameter, quantified by the amount of blood delivered to the tissue per unit time, per unit volume or mass of tissue [31,14]. Since it is closely related to the delivery of oxygen and

nutrients to the tissue, abnormal changes in perfusion can be used as an indicator for various disorders. There exist several MRI based perfusion imaging techniques such as Dynamic Contrast Enhancement (DCE) [13], Dynamic Susceptibility Contrast (DSC) [38,55], Arterial Spin Labeling (ASL) [14,69,3] and Intra Voxel Incoherent Motion (IVIM) [7] to quantitatively assess the different hemodynamic parameters related to perfusion. In the case of brain, Cerebral Blood Flow (CBF) is a critical perfusion parameter that defines volume of blood flowing through a specific region of brain tissue per unit time. It plays a vital role as a biomarker for various neurodegenerative diseases and tumor. DCE and DSC are mainly used to measure cerebral perfusion that uses intravenous administration of exogenous tracer like gadolinium based contrast agent which causes side effects including the potential risk of conditions like

nephrogenic systemic fibrosis [67]. ASL is a non-invasive technique that completely eliminates the use of contrasts media and allows the possibility of repetitive follow-up, especially in the case of pediatric and pregnancy cases. Most of the ASL research is carried out in the brain region [15,21,61,71], but recent research [2,45,48] show that ASL is a promising approach for perfusion quantification in abdominal organs also. The working principle of ASL mainly focuses on two types of images, namely 'Label' and 'Control'. A Label image is generated through Arterial Spin Labeling which magnetically inverts the arterial blood water molecules by applying radio frequency inversion pulses in the neck region. The label image is acquired after a time delay which is the time taken by the labeled proton to reach the brain, called Arterial Trasnsit Time (ATT) [3,28]. Control image is acquired without inverting inflowing blood water magnetization. The difference between the labeled and control images is called Perfusion Weighted Image (PWI) and it represents the ATT weighted signal that can be converted into absolute quantitative CBF in $ml/100g/min$, using a kinetic model [11]. Accuracy and quality of CBF quantification are a promising criterion for decision making in medical diagnosis and therapeutic management of aging and various other neurological disorders. But poor Signal to Noise Ratio (SNR), low resolution and prolonged scanning time are the challenging problems that affect the CBF quantification procedure significantly, causing the barrier for a widespread use of ASL imaging clinically.

Gray Matter (GM) CBF is of primary interest in disorders such as alzheimer's and parkinson's diseases as well as functional studies related to the brain in which the abnormal GM CBF is used as a biomarker for clinical diagnosis. Studies related to White Matter (WM) perfusion are rare due to poor SNR as blood may take longer transit time to reach the white matter and shorter T1 relaxation time of white matter [3,18]. For healthy GM, perfusion value of blood is approximately 60 mL $(100\ gm)^{-1}min^{-1}$ and WM perfusion is only 1/3 of this value [52]. Most of the ASL-MRI studies focus to analyze GM perfusion using structural MRI or the Region of Interest (ROI) analysis [42,65]. But it is very difficult to separate the perfusion signals associated with the three types of brain tissues such as GM, WM and Cerebrospinal Fluid (CSF) due to the low resolution of ASL image. In such cases the voxel intensity depends not only on the imaging sequence and tissue properties, but also on the proportions of each tissue type present in that voxel, called Partial Volume (PV) effects [4]. As a result, the measured signal from each voxel will be a weighted average combination of tissue types present in it. PV effects can also be caused by the blurring of images due to the point spread function (PSF) [49]. PSF is consequences of image formation and varies with the method of acquisition.

To give a better approximation of GM CBF quantity, PV effects are to be corrected or to be minimized. PV effects will be minimum for high resolution ASL data. However, the clinical image acquisition process limits the SNR in high resolution images due to the hardware and experimental design constraints, computational burden and cost effectiveness. The high resolution ASL image acquisition, on the other hand, results in a low SNR. To improve the SNR, perfusion images are typically acquired by averaging a large number of pairwise subtractions of label and control images which results in prolonged scanning time. This causes embarrassment to the patient and creates motion artifacts in ASL images which in turn lead to further quantification errors. Thus improving the resolution of ASL image happens at the cost of reducing the SNR or long acquisition time and their cumulative effects. In the currently available clinical image acquisition systems, obtaining high resolution and high SNR image is practically impossible. Effectively incorporating the appropriate post processing techniques to enhance ASL images to correct PV effects is a promising approach to tackle this problem. The majority of the research work reported on this issue [1,4, 12,40,49,52] concentrated on a voxelwise correction of PV effects as a post processing technique to separate GM and WM CBF, based on the PV estimates of GM and WM. The PV estimates of GM, WM and CSF are determined from the high resolution structural MRI, using segmentation, registration and resampling techniques. Segmentation is carried out to separate the brain volume into GM, WM and CSF components in terms of the respective PV estimates. Since the target space of the PV correction is ASL space, image registration and resampling techniques are to be used to convert the high resolution PV estimates thus obtained to a low resolution space or ASL space. Suitable PV correction algorithms are applied to ASL images using the PV estimates in ASL space. Any error in the calculation of PV estimates will deteriorate the quantification accuracy of PV correction techniques. Linear Regression (LR) [4] is a standard algorithm used for PV correction which is recommended for highly time critical applications. LR is one of the most popular algorithms due to its simplicity, fast execution and less sensitivity to error in PV estimates [4,12,77]. However, the LR method suffers from an inherent smoothing effect resulting from the assumption that tissue magnetization and perfusion are constant over a region, called regression kernel. This smoothing effect causes blurring and loss of edge information on the GM and WM tissue boundary, leading to significant quantification errors [4]. Since ASL data inherently suffers from low SNR and blurring due to low resolution, performance of the LR method will be further affected.

Recently, super resolution of ASL images to reduce the PV effects using traditional or deep learning based methods without using explicit PV estimates are being explored [46,48]. Training a deep learning network for accurate results, huge volume of low-resolution and high-resolution ASL images are essential, which is not clinically feasible. This approach also suffers from an inaccuracy in estimating GM CBF, which is derived directly from a CBF map and is the primary factor influencing most of the clinical diagnosis. This leads to the necessity of a technique that enhances ASL image prior to the voxelwise correction for the accurate quantification of GM CBF. This image enhancement helps to improve the SNR and contrast of edges of ASL image which will lead to the reduction in the blurring effect of LR algorithm, especially in the vicinity of GM and WM boundaries. Traditional enhancement methods have many disadvantages, including high computational costs due to the use of conventional optimization methods and the need of manual parameter settings to improve the performance [44,63]. Deep learning techniques have the ability to overcome these drawbacks by using deeper architecture to automatically learn and find more suitable features of the image rather than the manual setting of parameters. Emerging developments in the field of deep learning motivated the researchers to build a deep learning-based strategy for biomedical denoising and super-resolution applications [53,63]. For Single Image Super Resolution (SISR) [16], the low resolution image data is considered as the down sampled version of the corresponding high resolution image with or without noise. By automatically extracting a collection of features from the set of prior examples used for training, a deep network tries to learn a mapping between low resolution and ground truth high resolution data. Convolution Neural Network (CNN) is the most typical and popular deep learning network that directly learn an end to end mapping between low resolution and high resolution images which is successfully applied in many recent research [16,33,53]. The non-linearity feature of the deep network can be increased by adding more CNN layers with proper activation function. This in turn helps to extract more hierarchical features and enables the network to learn a function between low resolution and high resolution images. But, deep neural network causes vanishing gradient problem [25] which makes it difficult to learn and change the parameters of the previous layers of the network. As a result, the performance gets saturated or even starts degrading rapidly. Many of the researchers have discussed this issue and suggested various architectures including ResNet [25,26,34,41], Densnet [29] and Inception [58,59]. Instead of using cumbersome per-layer network piling, these networks use a block wise structure to make network generation efficient and generalizable [79]. He et al. [25] proposed Residual block in ResNet architecture which uses identity shortcut connection or skip connection to improve the network performance by local residual learning [29]. suggested a densnet architecture
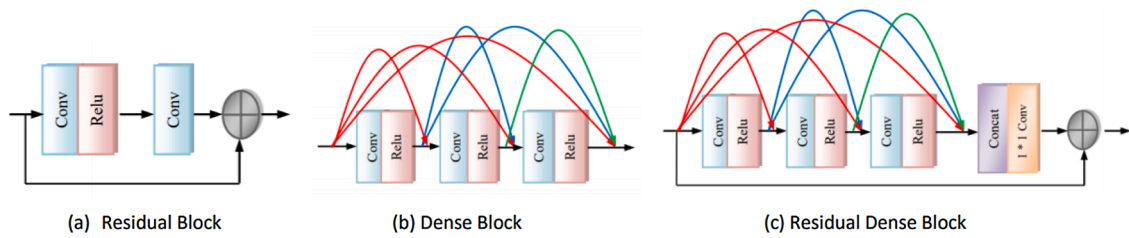
Fig. 1. Network Structure of a typical (a) Residual block [41], (b) Dense block [64] (c) Residual Dense Block [75].

consisting of a series of dense blocks which directly links all layers with each other. Concatenating feature maps from various layers will retain the features, along with increasing variance of output and facilitating the reuse of features [75]. suggested Residual Dense Block (RDB) consisting of dense linked layers with local residual learning. In addition to addressing the vanishing gradient problem, the RDB network ensures maximum data flow between network layers and reuse of features. There exist different variants of RDB in literature based on the variants in residual block and dense block [41,54,64]. The block diagram of a typical Residual Block(RB) [41], Dense Block (DB) [64] and RDB [75] are shown in Fig. 1. In this work, a deep learning based ASL image enhancement technique based on the principle of SISR is proposed using RDB as the basic building block by generating an adequate amount of simulated images and fine tuning the parameters to improve the quality of output image.

2. Related works

ASL [69] is a totally non-invasive MRI technique which has product status with all major MR vendors and is capable of creating a dramatic revolution in the clinical diagnosis and radiology-based therapy. Due to a variety of technological factors, the current state of ASL procedures is susceptible to a number of limitations that impair image quality and patient satisfaction. Many researchers have attempted to improve the quality of ASL images by removing noise and artifacts using a variety of post-processing techniques, such as motion correction [5], physiological noise correction [6], spatial noise reduction [60,68], temporal noise reduction [51] and spatio-temporal noise reduction [10,20,56,80]. Apart from the aforementioned artifacts, the disruption that obscures the underlying information known as outliers can also degrade the overall quality of the image, which is addressed by several recent research works [39,57]. The majority of these methods are based on implicit or explicit models that are data specific and need not be

accurate due to patient dependence. Recently, deep learning based denoising algorithms have been found as promising approaches to improve the signal quality in ASL images. Several deep learning based models [22,35,50,66,72] were suggested to reconstruct ASL images using minimum number of label-control pairs, at the expense of huge training datasets and scanning time, found to yield accurate results compared to traditional models.

Studies proved that, in addition to SNR, PV effects in ASL images can also lead to significant errors in the quantification of CBF values that can be tackled by various PV correction algorithms. The first work on single Post Labeling Delay (PLD) ASL [4] used a regression technique with the assumption that for a given tissue type within a given voxel, CBF is identical to that of the nearby voxels within a predefined kernel. Despite the fact that this is a promising strategy for obtaining the exact voxel wise CBF contribution from the GM and WM, the regression kernel makes GM-WM tissue boundaries smooth, which complicates the detection of local changes in CBF values. Many groups focused to address the smoothing effect problem by different approaches like bayesian [12], robust regression techniques using modified least trimmed square method [40] and kernel based method [49]. The Bayesian method, which is more complex and slower than the LR method, converges after a large number of iterations. A trimming parameter that is unique to each input determines the accuracy of the modified least trimmed square approach. The kernel method transforms 2D kernel to 3D kernel which involves more complex processing than LR method.

Most of the PVE correction techniques reported so far rely on the partial volume estimates derived from high resolution structural MRI, which in turn depend on the reliable segmentation, registration and resampling techniques [49]. Registration of high-resolution anatomical data to perfusion space is difficult due to the low contrast and spatial detail of the low-resolution perfusion space. As a result, an error in this procedure will lead to inaccurate PV estimate and a reduction in PV correction accuracy. Furthermore, there is still no consensus on how the
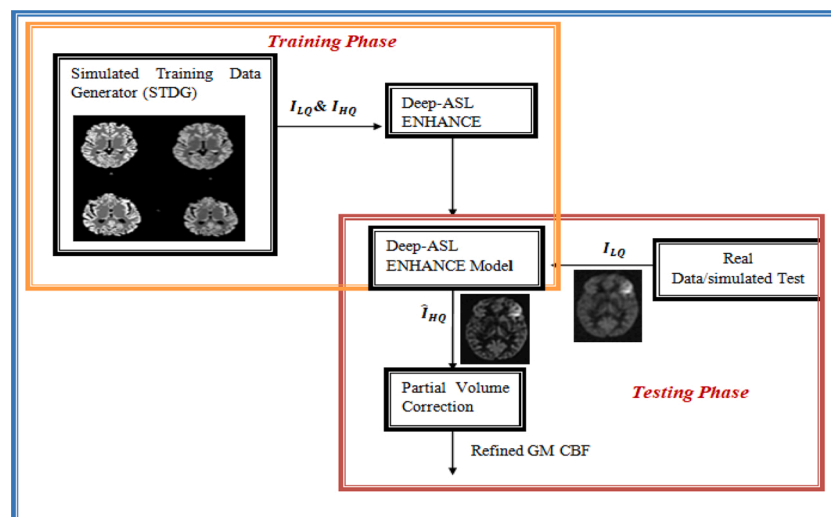

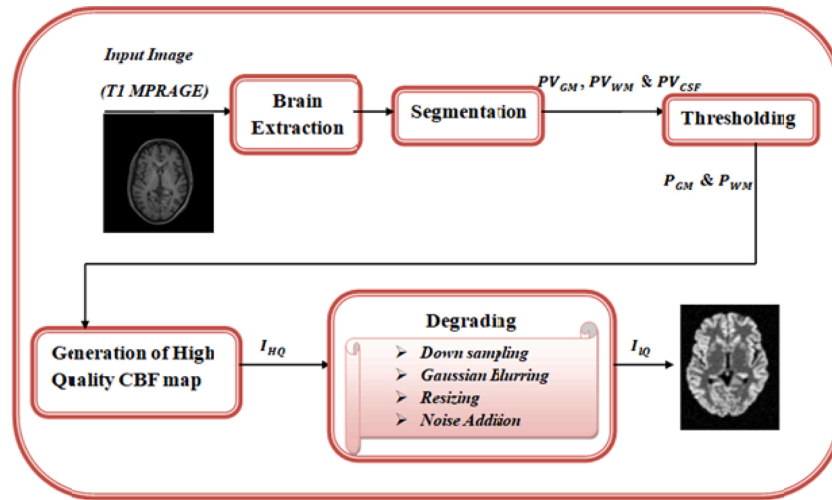**Fig. 2.** Overall Architecture of the Proposed Method.

Fig. 3. Simulation procedure for generating low quality and high quality ASL CBF images.

segmentation should be resampled from anatomical to perfusion space [52]. proposed that, generating PV estimates directly at perfusion space removes the requirement of resampling techniques and applied PVE correction using regression techniques based on the PV fraction obtained using a Look Locker Saturation Recovery EPI sequence (LL-EPI) [1]. attempted to determine PV estimates from low resolution ASL data using the QUASAR sequence (Quantitative STAR Labeling of Arterial Region), which was not clinically feasible. Later, attempts were made to use super resolution as a post-processing approach to improve ASL image resolution to reduce PV effects. A nonlocal patch based super resolution technique on Pseudo Continuous ASL (PCASL) images was proposed [46], where the nonlinear weights were determined from non-local image patches of both structural and ASL images, but it failed in detecting local image variations. This technique was limited by the denoising capability due to the lack of mean signal consistency and low SNR of ASL MRI. Later, deep learning based solutions were developed for superresolution of ASL images and [78] pioneered in this, proposing a deep learning based method for superresolution of ASL images using clinical data. However, the requirement for a significant number of low resolution and high resolution clinical ASL images to train the deep learning model made it impractical. An analytical study [77] revealed that the LR approach is less sensitive to PV estimation errors, and it has already been proven as useful in time-critical applications. Hence, the technique that mitigates the drawbacks of the LR algorithm can be a more promising approach in clinical diagnosis. Motivated by this fact, the super resolution image enhancement principle [73] is applied to an ASL image to enhance its quality and further improve PV correction by

incorporating deep learning techniques.

3. Proposed methodology

Clinical ASL images suffer from low SNR, low resolution and blurring effects which degrades the image quality, resulting in PV effects and CBF quantification errors. Hence it becomes inevitable to improve the image quality and apply the PV correction techniques to enhance CBF quantification accuracy. This work proposes a deep learning architecture based on the principle of SISR which provides automated ASL image quality enhancement from a clinical perspective, as shown in Fig. 2.

The performance of all deep learning based architectures depends on the number of inputs used for training the network. Biomedical image processing using deep learning network fails to give accurate results due to the unavailability of sufficient clinical input images. Hence this method focuses on generating high quality ASL image using a deep learning network by training the network using the simulated low quality ASL image from the available or acquired structural MRI by Simulated Training Data Generator (STDG) block and testing it using clinical ASL images. STDG consists of four phases such as segmentation, thresholding, generation of high Quality CBF map and degrading to generate low quality ASL image as shown in Fig. 3.

Segmentation phase segments structural MR images to obtain the PV estimates of GM (PV_{GM}), WM (PV_{WM}) and CSF (PV_{CSF}). Since magnetization of CSF in the perfusion weighted image is zero, PV estimate of the CSF is not considered to produce simulated CBF. As the next step, thresholding is carried out on the generated PV_{GM} and PV_{WM} to preserve
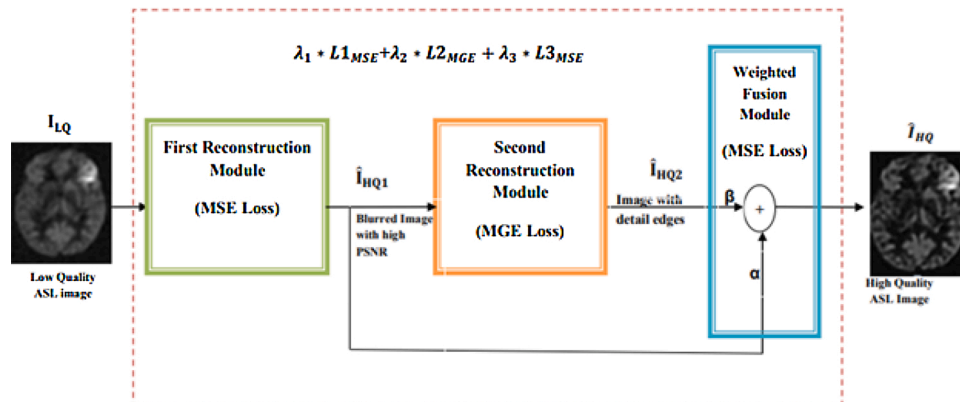

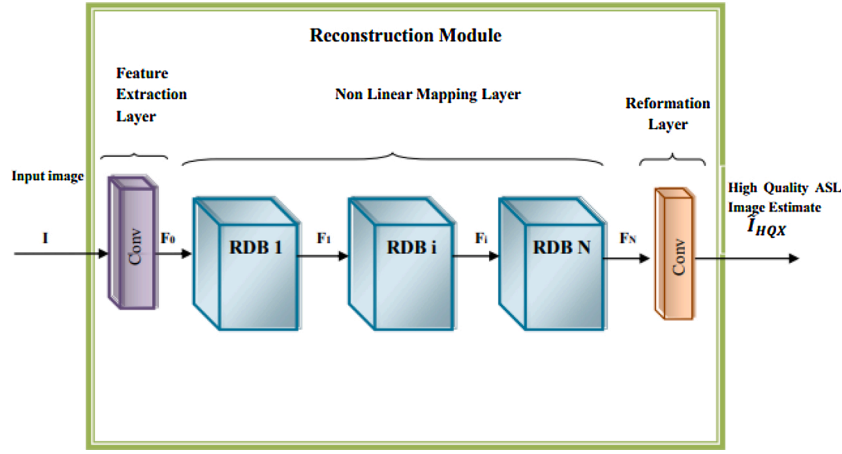**Fig. 4.** Block diagram of proposed Deep-ASL ENHANCE block.

Fig. 5. Structure of a Reconstruction Module.

the voxels of GM and WM with a ratio greater than a predefined threshold to produce P_{GM} and P_{WM}. PV estimates of GM and WM are masked to remove the voxels with probabilities lower than 0.1 which is commonly used in literature [77]. There exist a relationship between P_{GM}, P_{WM} and CBF and it is represented using the Eq. (1).

$$CBF = P_{GM} \times p + P_{WM} \times q \qquad (1)$$

where CBF represents simulated CBF image. The variables p and q depend upon the CBF value in the GM and WM regions respectively. Across the whole brain, fixed CBF values for normal GM and the WM regions were taken as $60 ml/100g/min$ and $20ml/100g/min$ respectively for healthy patients [40,77]. Eq. (1) becomes

$$CBF = P_{GM} \times 60 + P_{WM} \times 20 \qquad (2)$$

The CBF image thus generated can be represented as the high quality ASL image, I_{HQ}. The degrading phase generates low quality CBF map I_{LQ} from I_{HQ} by down sampling, gaussian blurring, resizing and noise addition operations. Gaussian blurring is applied to incorporate the PV effects due to the point spread function [49]. The clinically available CBF image is usually affected by Gaussian noise. Hence,it is necessary to add gaussian noise to the simulated CBF image at different SNR levels to analyze the performance of the algorithm. Generated I_{LQ} and I_{HQ} are fed to the Deep-ASL ENHANCE module during the training phase to build a deep learning model that can be used to enhance real ASL image.

The "Deep-ASL ENHANCE" block develops and learns a function that creates a high quality ASL image $\widehat{I}_{HQ}$ from a low quality image I_{LQ}. The estimated high quality ASL image suffers from PV effects and therefore PV correction techniques are to be applied on $\widehat{I}_{HQ}$ to obtain refined GM CBF. The detailed representation of the proposed Deep-ASL ENHANCE block is illustrated in Fig. 4.

3.1. Architecture of Deep-ASL ENHANCE block

The different sub modules of the Deep-ASL ENHANCE block and their chronological order in the architecture are finalized based on the apriori information from the literature and our experimental analysis using the available datasets [75,78]. Low SNR, as well as the effects of blurring, result in the loss of edge information between GM and WM tissue boundaries in ASL images. This necessitates the design of an architecture that improves SNR and reduces blurring effect for the estimation of accurate CBF value. The overall performance of the deep learning based architecture depends upon the factors like number and nature of layers used, number of epochs, batch size, learning rate, selection of the optimizer and the loss function. In this work, the quality of the reconstructed image is ensured by using a multi-loss joint approach.

Mean Squared error(MSE) is commonly used as the loss function in all super resolution tasks, which primarily captures the pixel wise difference between two images, emphasizing how far the pixels of the target image are from the pixels of the predicted/generated image [47]. Training deep network using MSE loss function is effective for improving SNR of an image, but the reconstructed image still remains blurred and CBF quantification will erroneous. The overall quality of the reconstructed image can be improved by using an effective mechanism to identify the missing edge information. Image gradient has been used as a measure of image sharpness, and it can be determined by using various types of first and second order edge detector operators [23]. Therefore, in this proposed architecture, Mean Gradient Error (MGE) [43,78] loss function is used to learn sharp edges that are closest to the ground truth. Hence, the Deep-ASL ENHANCE block is implemented using two reconstruction modules. Low quality ASL image is fed as input to the first reconstruction module which generates primary enhanced estimate with improved PSNR level ($\widehat{I}_{HQ1}$). First reconstruction module based on MSE loss function eliminates noise in the ASL and the second reconstruction module is used to recover high frequency information present in $\widehat{I}_{HQ1}$ using MGE loss function yielding secondary enhanced estimate ($\widehat{I}_{HQ2}$). To obtain the high quality image with high SNR and sharp edges, a weighted fusion module is used with learnable weighting coefficients α and β to reconstruct final high quality image from $\widehat{I}_{HQ1}$ and $\widehat{I}_{HQ2}$.

The primary and secondary enhanced estimates are generated by appropriately mapping low quality and high quality image pairs by extracting a set of features from low quality image and transforming them into higher dimensional feature space corresponding to a high quality image. To achieve these functionalities, feature extraction layer, nonlinear mapping layer, and reformation layer are incorporated in the reconstruction module as shown in Fig. 5. Basic functions of each of these layers are described below.

3.1.1. Feature extraction layer

This layer comprises of a convolution layer to extract a set of feature maps called as shallow features corresponding to the input image. The output of this layer can be represented as

$$F_0 = W1 * I + B1 \qquad (3)$$

where $W1$ and $B1$ corresponds to filters and biases respectively and $*$ denotes the convolution operation. $W1$ corresponds to $n1$ filters of size $c \times f \times f$ where c represents the number of channels in the input image and f is the spatial size of the filter. Since this work handles only grayscale images, number of channels (c) is set to 1 and the size of the filter is set to 3×3 for computational simplicity. Here $W1$ applies $n1$ convolutions on the input image and generates $n1$ feature maps that are passed
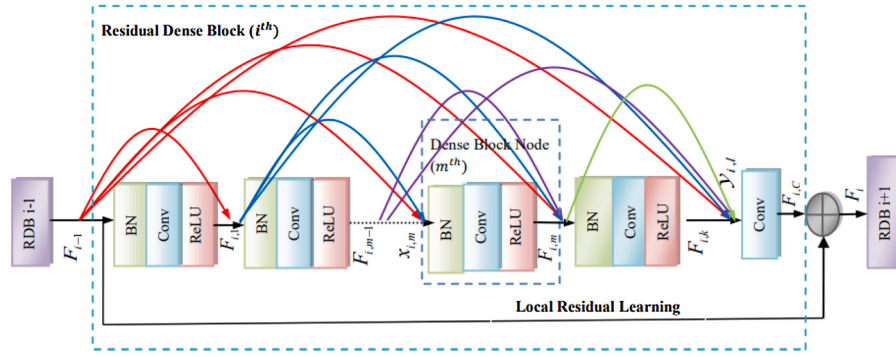
Fig. 6. Structure of a Residual Dense Block.

to non linear mapping stage.

3.1.2. Non linear mapping layer

This layer nonlinearly maps $n1$ feature dimensional vectors of input image into $n2$ dimensional feature vectors that are conceptually a representation of high resolution image used for reconstruction. N number of stacked RDB are used in order to learn a nonlinear feature representation which has a strong influence on the performance of the reconstruction module. Each RDB contains a combination of k number of dense block nodes followed by a convolution layer and skip connection for local residual learning as shown in Fig. 6.

Each of the dense block nodes is constituted by Batch Normalization (BN), convolution, and ReLu and the parameter k is decided by experimental analysis. The proposed RDB architecture incorporates BN prior to every convolution and ReLU combination, to reduce the internal covariance shift problem [30]. BN layer stabilizes the learning process and it significantly reduces the number of training epochs. The feature map generated at the output of each dense block node depends upon the number of filters used in its convolution layer. The feature maps obtained at the mth dense block node of ith RDB block, $F_{i,m}$, are the fused feature map of all $(m-1)$ preceding dense block node of ith RDB and its corresponding input feature map $F_{(i-1)}$ which in turn determines the growth rate of an RDB. The convolutional layer used after the kth dense block node is the output convolutional layer of each RDB and its purpose is to equalize the size between feature maps that are given as input to the skip connection where it is summed up.

Let $F_{(i-1)}$ and F_i be the input and output of ith RDB block, respectively each RDB block contains k number of dense block nodes. F_i can be mathematically represented as

$$F_i = H_{RDB,i}(F_{i-1}) \tag{4}$$

where $H_{RDB,i}$ denotes the overall operations of the ith RDB block. The output of mth dense block node of ith RDB block $F_{i,m}$ can be represented as

$$F_{i,m} = ReLU\left(Conv\left(BN\left(x_{i,m}\right)\right)\right) \tag{5}$$

where $x_{i,m} = Concat(F_{i-1}, F_{i,1}, \ldots. F_{i,m-1})$. Now the output of BN becomes

$$O_{mBN} = BN(x_{i,m}) = BN(Concat(F_{i-1}, F_{i,1}, \ldots. F_{i,m-1})) \tag{6}$$

By convolving O_{mBN} with set of f filters W_m having bias values B_m generates the output of mth convolutional layer in m^{th} block,

$$O_{mCONV} = W_m * O_{mBN} + B_m \tag{7}$$

After applying ReLU activation function on O_{mCONV}, we get $F_{i,m}$ as

$$F(i,m) = max(0, O_{mCONV}) \tag{8}$$

By integrating all the functions, the overall output $F_{i,m}$ can be represented as

$$F_{i,m} = Max(0, (W_{i,m} * (BN(Concat(F_{i-1}, F_{i,1}, \ldots. F_{i,m-1}))) + B_{i,m})) \tag{9}$$

The last dense block node of RDB is connected to a convolution layer and gets the output $F_{i,C}$ as

$$F_{i,C} = W_{i,C} * Y_{i,l} \tag{10}$$

where $Y_{i,l} = concat(F_{i-1}, F_{i,1}, \ldots, F_{i,k})$. After Local Residual Learning, the final output of ith RDB block is depicted as

$$F_i = F_{i,C} + F_{i-1} \tag{11}$$

3.1.3. Reformation layer

The feature maps are propagated through N consecutive RDB blocks to arrive at F_N which is fed to a separate convolutional layer to reconstruct the high resolution image estimate $\widehat{I}_{HQx}$ which is mathematically represented as

$$\widehat{I}_{HQx} = W2 * F_N + B2 \tag{12}$$

where $W2$ represents a filter of size $f \times f$ (spatial filter size) and $B2$ represents its bias values.

3.1.4. Multi-loss joint strategy

It is necessary to analyze the cascading effect of multiple reconstruction modules within the Deep-ASL ENHANCE block. MSE loss function of the first reconstruction module minimizes the loss between reconstructed image $\widehat{I}_{HQ1}$ and ground truth high quality image I_{HQ} where $\widehat{I}_{HQ1}$ is represented as

$$\widehat{I}_{HQ1} = F_a(I_{LQ}; \theta) \tag{13}$$

where θ denote the overall parameters of the deep learning network. Given, a set of low quality and high quality image pairs (I_{LQ} and I_{HQ}), the MSE loss function is defined as

$$L1_{MSE}(\theta) = \frac{1}{n} \sum_{i=1}^{n} \|F_a(I_{LQ}; \theta) - I_{HQ}\|^2 \tag{14}$$

where n represents the total number of training samples. The MGE loss function of the second reconstruction module is designed by comparing various gradient computation methods [23]. The sobel operator is found to be most suitable for our work because of improved computational time efficiency and noise reduction characteristics [23]. Gradient of an image is calculated by convolving the image with corresponding filter mask in the X and Y directions. In this method, horizontal (H_{sobel}) and vertical mask (V_{sobel}) considered are,

$$H_{sobel} = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix} \text{ and } V_{sobel} = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}$$
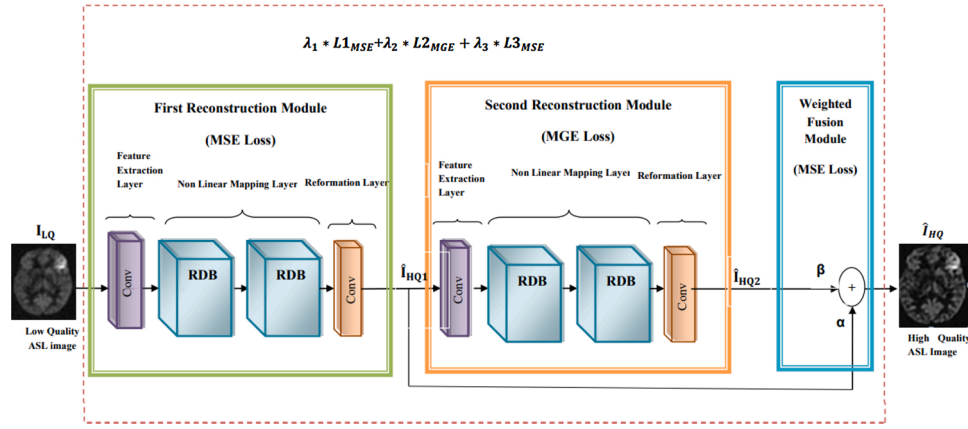
Fig. 7. Overall architecture of proposed Deep-ASL ENHANCE block.

Then magnitude of a gradient operator at each pixel location is calculated by

$$G(i,j) = \sqrt{G_x^2(i,j) + G_y^2(i,j)} \tag{15}$$

where $G_x(i,j)$ and $G_y(i,j)$ denotes the gradient in the X and Y directions respectively. $G_x(i,j)$ and $G_y(i,j)$ can be implemented by convolution of an image I with horizontal filter mask (H_{sobel}) and vertical mask (V_{sobel}) respectively as follows

$$G_x = I * H_{sobel} \tag{16}$$

$$G_y = I * V_{sobel} \tag{17}$$

High SNR image obtained as the output of the first reconstruction module, $\widehat{I}_{HQ1}$, is fed as input to second reconstruction module to sharpen the edge features using MGE loss function which is defined as

$$
\begin{aligned}
L2_{MGE}(\theta) &= \frac{1}{n} \sum_{i=1}^{n} \left\| \widehat{G}_{\widehat{I}_{HQ2}} - G_{HQ} \right\|^2 \\
&= \frac{1}{n} \sum_{i=1}^{n} \left\| F_b(\widehat{I}_{HQ1}; \theta) - G_{HQ} \right\|^2
\end{aligned} \tag{18}
$$

The estimated images $\widehat{I}_{HQ1}$ and $\widehat{I}_{HQ2}$ are weighted by the factors α and β respectively and fed to the weighted fusion module to improve the overall quality of images in terms of both SNR and edge information. The quality of the output image I_{HQ} can be improved by the appropriate tuning of parameters $\alpha \beta$ and β by weighted fusion module with MSE loss function ($L3_{MSE}$), which can be visualized as a simple perceptron with weighting parameters α and β. The overall output I_{HQ} and $L3_{MSE}$ is represented as

$$\widehat{I}_{HQ} = \alpha \cdot \widehat{I}_{HQ1} + \beta \cdot \widehat{I}_{HQ2} \tag{19}$$

$$L3_{MSE}(\theta) = \frac{1}{n} \sum_{i=1}^{n} \left\| F_c(\widehat{I}_{HQ}; \theta) - I_{HQ} \right\|^2 \tag{20}$$

Final enhanced image is obtained by deploying a multi-loss joint strategy [78] during training process in which the overall loss depends on the individual modules outputs, $\widehat{I}_{HQ1}$, $\widehat{I}_{HQ2}$ and $\widehat{I}_{HQ}$ which in turn depend on their respective loss functions $L1_{MSE}$, $L2_{MGE}$ and $L3_{MSE}$. Then the overall loss can be expressed as

$$OverallLoss = \lambda_1 \times L1_{MSE} + \lambda_2 \times L2_{MGE} + \lambda_3 \times L3_{MSE} \tag{21}$$

where λ_1, λ_2 and λ_3 are the corresponding weighting factors which affects the performance of the network. Fixing the value of these
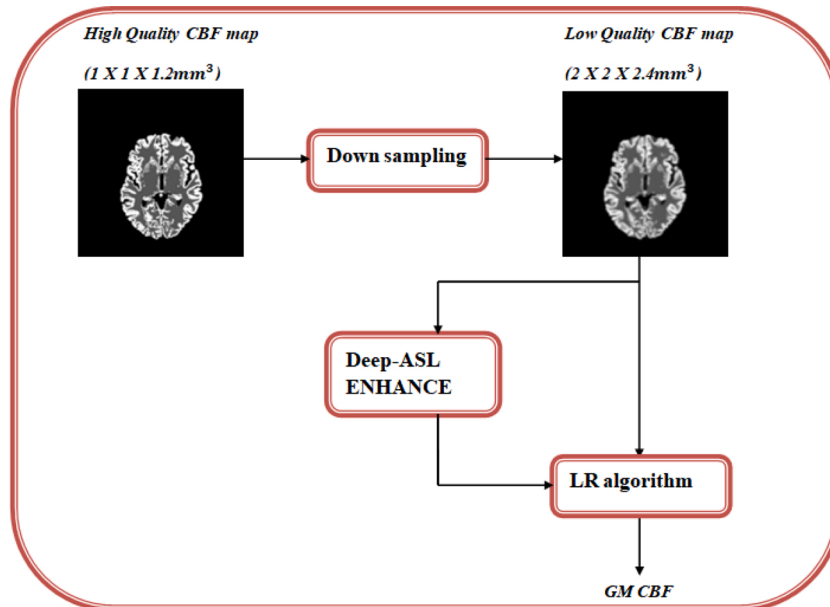

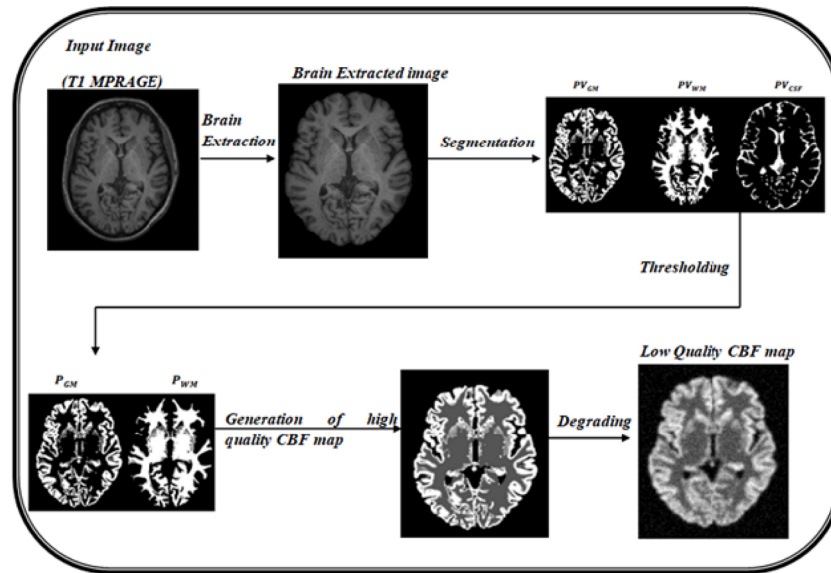**Fig. 8.** Flow diagram of partial volume correction method.

Fig. 9. Results for each stage in the ASL simulation process.

weighting factors is the challenging problem associated with multi loss joint strategy.

By experimentally analyzing the quality metrics in terms of Root Mean Squared Error (RMSE), Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index (SSIM), the number of RDBs is finalized as two for the non linear mapping layer. The overall architecture of the DEEP ASL-ENHANCE block is as shown in Fig. 7.

3.2. Deep-ASL ENHANCE - a preprocessing step to PV correction

The deep learning based enhancement steps described so far improves ASL image quality in terms of blurring effects and noise. But it is unable to accurately quantify GM CBF due to PV effects. For the accurate estimation of voxel wise GM CBF, which is more useful in clinical diagnosis, PV correction is mandatory. The deep-ASL enhancement algorithm is employed as a preprocessing step to the LR algorithm, which can lead to a promising approach for improving the accuracy of CBF quantification compared to the use of LR algorithm alone, the most standard algorithm used for PV correction. Fig. 8 illustrates the procedure followed for PV correction. The high quality CBF map is down sampled to half size resolution which mimics the simulated ASL image and is fed as the input to the LR algorithm along with other parameters such as brain mask, kernel size and down sampled PV estimates of GM and WM. A Deep-ASL ENHANCE block is incorporated prior to LR to compare its effects on GM CBF quantification. The details of this validation are discussed in Section 4.5.

4. Discussion of experimental results and analysis

Deep learning based ASL image enhancement method is implemented and its performance is evaluated using simulated data, and the model developed is successfully applied to clinical ASL data. The simulated data is generated using the technique mentioned in Fig. 3 because there isn't enough clinical data to construct a deep learning-based architecture.

4.1. ASL dataset simulation

The ASL image is simulated using structural images from the Alzheimer's Disease Neuroimaging Initiative (ADNI-2) CN research group database (http://adni.loni.usc.edu) with age ranges from $60-85$. Structural images in ADNI2 were acquired using the Siemens Verio

Scanner with 3D MPRAGE T1-weighted sequence. The acquisition parameters were used as Field of View (FOV) $= 256 \times 240 \ mm^2$, TR/TE $= 2300/3ms$, voxel size $= 1 \times 1 \times 1.2 \ mm^3$, slice thickness $= 1.2 \ mm$, flip angle $= 9$ ° and bandwidth $= 240 \ Hz/px$.

Brain tissue extraction, segmentation, thresholding, generation of high quality CBF map and degrading operations are performed on the structural data to generate low quality ASL images. Brain tissue is extracted from high-resolution T1-weighted structural images to remove the non-brain components using the FSL tool BET [32]. The segmentation phase produces PV estimates for GM, WM and CSF from the input images using the FSL tool FAST [76]. In the thresholding process, PV estimates of GM and WM with voxel intensity less than 10 % were excluded from the analysis, by replacing values less than 0.1 by 0 in the relevant PV estimates. Then high resolution CBF map is created by incorporating partial volume effects using Eq. (1). Down sampling operation is carried out to transform the ASL images to half the size. Gaussian blurring is added to the down sampled image with Full Width at Half Maximum (FWHM) value $1mm$ to include the effects of PSF. After equalizing the size of the low and high quality images by resizing process, noise corresponding to different SNR levels as 1015 and 20 is applied.

As reported [6,70], ASL raw images are affected by rician noise where as perfusion weighted images and CBF map are affected by Gaussian noise. Hence the Gaussian noise with a standard deviation corresponding to respective SNR levels were added, according to

$$SD = \frac{(max(\Delta M(t)))}{SNR} \tag{22}$$

Here, $\Delta M(t)$ refers to the ASL difference signal in the kinetic curve model [11] created by setting GM CBF (f) $= 60 \ mL/100/min$, $\Delta t = 0.7s$ and $t = 1.8$. All other parameters of the kinetic model are specified as standard values [77]. Fig. 9 depicts the outputs of each phase in the simulation process.

4.2. Clinical data

Deep-ASL ENHNACE method was also evaluated with representative images of clinical cases acquired from 3 T GE Discovery 750 W scanner (GE Healthcare, Milwaukee, WI, USA) using a 3D PCASL technique after the approval of Institutional Ethics Committee (IEC). Image acquisition was performed with TR/TE $= 4854/10.7ms$, PLD $= 2025 \ ms$ for adults and 1525 ms for children, FOV $= 240 \ mm$, number of excitations (NEX)

Table 1

Comparison of mean RMSE, PSNR and SSIM score obtained for different number of convolutional layers in feature extraction layer with different SNR levels.

| | Number of CNN layers in Feature Extraction Layer | | | | | | | | |
| | 1 | | | 2 | | | 3 | | |
SNR	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM
20	**0.0803**	**23.1494**	**0.9032**	0.0831	23.0702	0.731	0.0891	22.94	0.698
15	0.0873	22.4144	0.8187	0.0869	22.0013	0.719	0.0901	21.786	0.665
10	0.0985	21.3694	0.5386	0.0971	21.1915	0.511	0.934	20.568	0.508

Table 2

Comparison of mean RMSE, PSNR, and SSIM scores obtained in the feature extraction layer with and without ReLu at various SNR levels.

| | Choice of ReLu function | | | | | |
| | Without ReLu | | | With ReLu | | |
SNR	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM
20	**0.0803**	**23.1494**	**0.9032**	0.0831	23.0702	0.731
15	0.0873	22.4144	0.8187	0.0869	22.0013	0.719
10	0.0985	21.3694	0.5386	0.0971	21.1915	0.511

= 3, spiral readout 8 arms × 512 samples and 4 arms × 512 samples, slice thickness = 4 *mm* and bandwidth = 62.5 *kHz*.

4.3. Data preparation and model training

The experiments were performed on a simulated ASL dataset of 40 subjects. Data of 10 subjects were used for testing purpose and 20 % of the remaining dataset were used as validation samples during training. For each subject, 151 axial planes from slice 30–180 which contains maximum information have been selected from a 3D simulated ASL image. Noise effects are introduced to the dataset by incorporating the noise with three distinct SNR values resulting in the generation of three distinct noise images corresponding to each slice. The total number of 2D ASL images constructed for subsequent experiments is 40 × 151 × 3, which is 18120.

Proposed deep learning networks are implemented using keras platform with tensorflow as the back-end. All the experiments were conducted on a PC with Intel(R) Core(TM) i7−5820k CPU 3.30 GHz and an Nvidia GeForce 980 Ti GPU.

Performance of the proposed method can be influenced by a number of hyper parameters such as number of convolution layers and size of filters in convolutional layer, number of residual blocks in nonlinear mapping module, number of epochs for training, selection of batch size and optimizer, fixing up of learning rate and the selection of suitable edge detector operator for MGE loss function. The model is build by the fine tuning of each parameter in accordance with a trial and error process. The performance of the proposed system is evaluated using different quantitative metrics such as PSNR, SSIM, and RMSE values.

4.4. Determination of various parameters in the proposed architecture

Feature extraction is generally performed using convolution layers followed by a ReLu activation function. The number of convolutional layers required and the need of ReLu activation function is influenced by the application and the dataset under study. As the number of

convolution layer increases, computational time complexity proportionally increases and which may not always guarantee the improvement in overall accuracy of the output. Hence the number of convolutional layers needed is determined by experimenting with increasing the number of convolutional layers and assessing the output quality using various metrics by providing input with varying SNR levels, as shown in Table 1.

It is observed that all these metrics give better results as the number of layer and noise level decreases. Hence, the number of convolutional layers in the final model is fixed to one. The effect of incorporating ReLu function between the convolution layer and the nonlinear mapping layer was also studied and the observations are given in Table 2.

The quality of the output obtained without considering ReLu function yields better results due to the clipping effect of ReLu function. There is considerable improvement in the metric SSIM compared to the other two measures RMSE and PSNR. The selection of the number of convolution layer also depends upon the number of filters and the size of the filters used. Therefore, the study is also extended to see the impact of various filter count in convolution layer, by fixing the filter size as 3 × 3, on the quality of the image. Table 3 shows the effects of the filter count on convolution layer by fixing it to 16, 32, 64 and 128 and it has been found that the quality of the output is the best for filter count 32 at different SNR levels.

Output images of the convolutional layer by fixing the layer count as 1 and filter count as 32, with and without ReLU in the feature extraction layer is shown in Fig. 10.

Following the feature extraction layer, a non-linear mapping layer is included using RDB as the basic building block. In many architectures, the nonlinear mapping effect is obtained by considering various network structures such as sequence of CNNs, RB, DB etc [16,33,34,41]. The result obtained for each of the approach varies depending upon the nature of the input and the overall architecture used. Hence, predefined methods cannot be adopted to any architecture for improving the output quality without proper analysis of the effects of different network structures independently or as a combination of these networks. To identify the independent behavior of various networks, the experiments were conducted using a simulated dataset with stacked CNN, RB, DB, RDB network layers for varying layer size. Number of layers considered for each network architecture is 3, 6, 9 and 12. Detailed observation of this experiment is given in Table 4.

It can be noticed that two RDB with 6 layers yields better performance. Therefore, it is identified that two RDB with 6 layers may be the most appropriate building block for constructing our architecture. But the overall accuracy of the entire architecture by considering RDB as the basic block cannot be predicted without considering the effects of other blocks. The experiments were also conducted using RB, DB and RDB

Table 3

Effects of different number of filters on the convolution layer of the Feature Extraction layer at different levels of SNR.

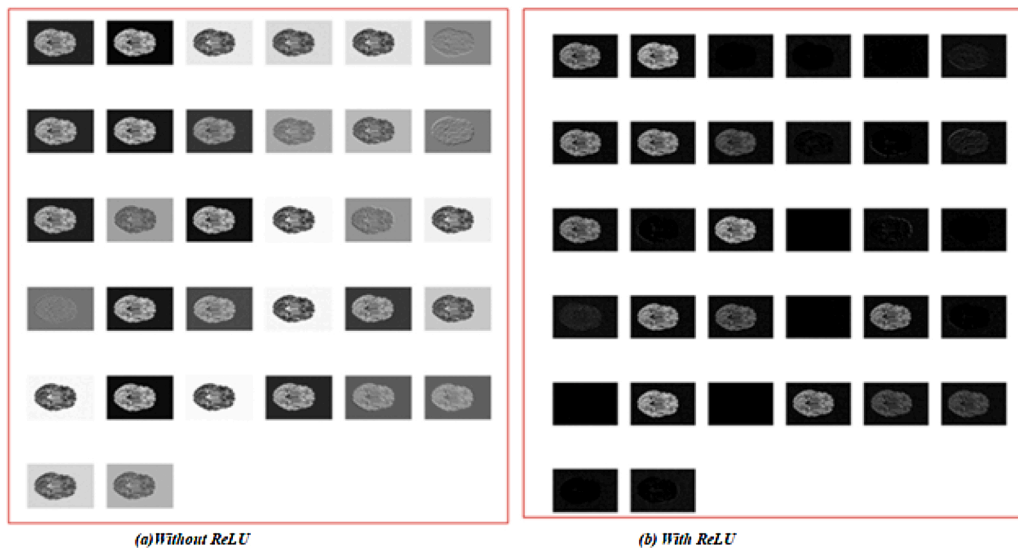| | Number of filters in feature extraction layer | | | | | | | | | | | |
| | 16 | | | 32 | | | 64 | | | 128 | | |
SNR	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM
20	0.856	21.98	0.709	**0.0803**	**23.1494**	**0.9032**	0.0804	23.1469	0.6533	0.0815	23.0061	0.4909
15	0.0899	20.67	0.656	0.0873	22.4144	0.8187	0.0884	22.2018	0.4509	0.0872	22.1579	0.4069
10	0.092	20.87	0.467	0.0985	21.3694	0.5386	0.1234	21.1856	0.4046	0.0976	21.1423	0.3549

Fig. 10. Visual representation of the output of the convolutional layer with layer count = 1, filter count = 32, with and without ReLU in feature extraction layer.

Table 4
Mean RMSE, PSNR, SSIM scores obtained using simulated dataset for various model configurations.

Models	RMSE	PSNR	SSIM
CNN_3	0.0876	17.4479	0.6298
CNN_6	0.0879	18.2861	0.6426
CNN_9	0.09	16.989	0.6219
CNN_12	0.0912	16.7831	0.6012
RB_3	0.08364	18.786	0.6812
RB_6	0.08345	19.4913	0.7086
RB_9	0.08438	19.298	0.6989
RB_12	0.08478	19.1297	0.6698
DB_3	0.0827	19.5366	0.7171
DB_6	0.0824	20.7625	0.7369
DB_9	0.0828	19.786	0.7348
DB_12	0.0831	19.5012	0.7318
RDB_3	0.0821	20.968	0.7354
RDB_6	0.0811	21.997	0.7588
RDB_9	0.0824	20.989	0.7356
RDB_12	0.0828	20.082	0.7239
2_RDBs_3	0.0819	22.5608	0.7687
2_RDBs_6	**0.0818**	**22.6539**	**0.79867**
2_RDBs_9	0.0823	21.789	0.7821
2_RDBs_12	0.0826	21.8546	0.7756

Table 5
Effects of RB, DB and RDB in nonlinear mapping module in terms of mean RMSE, PSNR and SSIM scores.

Model	RMSE	PSNR	SSIM
Multiloss joint strategy with RB in nonlinear mapping module	0.0811	22.82	0.8262
Multiloss joint strategy with DB in nonlinear mapping module	0.0809	22.89	0.8533
Multiloss joint strategy with RDB in nonlinear mapping module	**0.0803**	**23.1494**	**0.9032**

blocks as the basic building block of the nonlinear mapping layer in the proposed architecture for arriving at the basic building block to be used. The results obtained are tabulated as Table 5. Proposed method with RDB in nonlinear mapping layer yields better performance particularly for SSIM metric compared to RMSE and PSNR.

Experiments were also conducted to study the effects of varying the number of RDB's inside non linear mapping module with varying noise level and the result is given in Table 6. It is interesting to observe that the

performance is better for RDB block count 2. The graphical representations of RMSE, PSNR, and SSIM metrics for various model configurations are shown in Fig. 11.

In some of the studies reported [19,41,74,75], it has been found that incorporating BN, deteriorates the efficiency of the architecture. Hence it was necessary for us to study the effect of BN in the RDB and the network was trained with and without BN to verify the performance of the model on simulated data. It is found that BN improves the network accuracy, particularly in terms of SSIM as shown in Table 7.

Based on the above mentioned performance analysis, it was decided to include two RDB blocks in the nonlinear mapping module where each block consists of six layers of sub blocks of BN, Conv and ReLU layer. Once the nonlinear mapping layer is fixed, the next major task was to decide the number of reconstruction module and the type of the loss function to be used in each module to improve the overall performance. In recent deep learning based super resolution literature [43], it is mentioned that MSE or combination of MSE and MGE can be used as loss function. Hence the experiment was conducted with single reconstruction module and then with two reconstruction modules. From the experimental analysis with a single reconstruction module with MSE, it was identified that for the simulated ASL images, output quality degrades due to blurring effects and also it fails to give edge information in the input. When a single reconstruction module with MGE is considered, the output obtained provides edge information alone and removes other relevant information. These observations are represented in Table 8 and the results are shown in Fig. 12.

Based on this observation, it was decided to use a reconstruction module with MSE loss function as the first unit, where the output suffers from loss of edge information. In the case of ASL image, since edge information is significant, some mechanism should be adopted to sharpen the edge details. Therefore the second reconstruction module with MGE is included following the reconstruction module with MSE loss function. The most challenging task here was to fix the type of edge detector operator used for gradient calculation. To fix this, analysis has been made with both first order edge detector operators such as Prewitt, Sobel and second order detector Laplacian. The results of this analysis are shown in Table 9. It is observed that the gradient operator using Sobel yields the best result and hence the operator is finalized as Sobel.

To get all the relevant information, first and the second reconstruction modules are combined with a weighted fusion module which uses MSE loss function. The associated parameters α and β were learned during the training process. The overall efficiency of the model is determined by the weighting factor by which the loss functions are

Table 6
Mean RMSE, PSNR, SSIM scores obtained by varying the number of RDB in nonlinear mapping module.

|	Number of RDB unit in nonlinear mapping module																				
	1			2			3			6			7		
SNR	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM
20	0.0854	23.0378	0.8775	**0.0803**	**23.1494**	**0.9032**	0.0793	22.8276	0.7256	0.1365	18.5315	0.8594	0.1357	18.3824	0.8593
15	0.0904	22.8569	0.6605	0.0873	22.4144	0.8187	0.0864	22.1696	0.7232	0.1456	18.1118	0.8197	0.1445	17.8386	0.8291
10	0.0991	22.4096	0.4508	0.0985	21.3694	0.5386	0.0969	21.2553	0.6247	0.1574	17.5377	0.7742	0.1603	17.2472	0.778
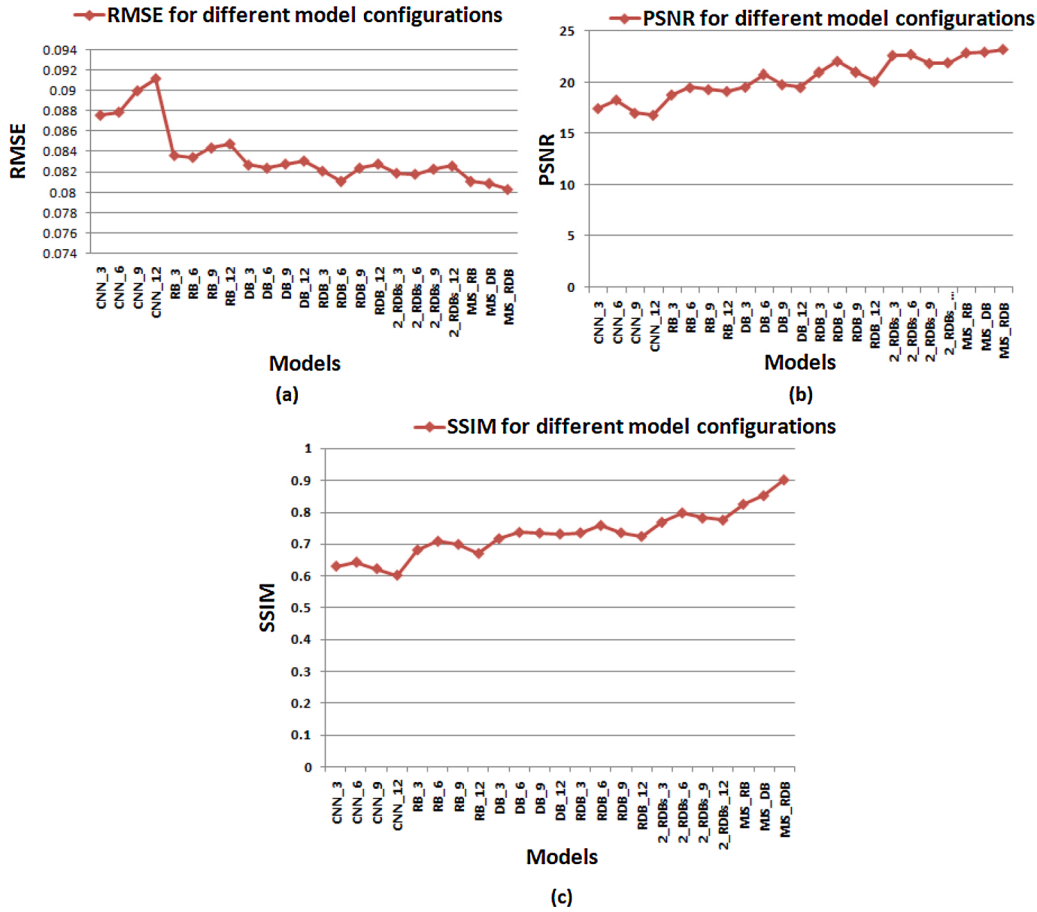
Fig. 11. Graphical representations of RMSE, PSNR, and SSIM scores for various model configurations.

Table 7
Comparison of mean RMSE, PSNR and SSIM score obtained with and without BN in RDB block.

	Nature of RDB					
	Without BN			With BN		
SNR	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM
20	0.0793	23.0407	0.8392	**0.0803**	**23.1494**	**0.9032**
15	0.0869	22.035	0.7725	0.0873	22.4144	0.8187
10	0.0958	21.1985	0.50988	0.0985	21.3694	0.5386

Table 8
Mean RMSE, PSNR, SSIM scores obtained with different reconstruction module strategy.

Reconstruction Module Strategy	RMSE	PSNR	SSIM
Single reconstruction module with MSE only	0.0778	22.6329	0.7066
Single reconstruction module with MGE only	0.7202	13.4864	0.3706
Proposed Method (Two reconstruction modules with multi-loss joint strategy)	**0.0803**	**23.1494**	**0.9032**

combined. Fine tuning of the weighting factors λ_1, λ_2 and λ_3 was done to improve the overall quality of the output and their values are experimentally determined as 3, 0.1, 3 respectively.

In addition to the architecture dependent parameters of the deep learning networks, general parameters such as learning rate, number of epochs, batch size and optimizers selected for training are also critical for analyzing the performance of the system. These parameters are fine tuned by assigning random values until the results converge. Deep

learning based architecture commonly uses Adaptive Moment Estimation (ADAM) [36], Stochastic Gradient Descent (SGD) [9] and RMSprop [27] as the optimizer which is highly dependent on the nature of the data. Hence, here the experiment is conducted using the optimizers such as ADAM, SGD, RMSprop with varying SNR levels and a comparative evaluation is given in Table 10.

For our dataset, ADAM optimizer exhibits better results compared to other two. The Learning rate and other parameters were fixed by assigning initial values identified from the literature [24,36] and the
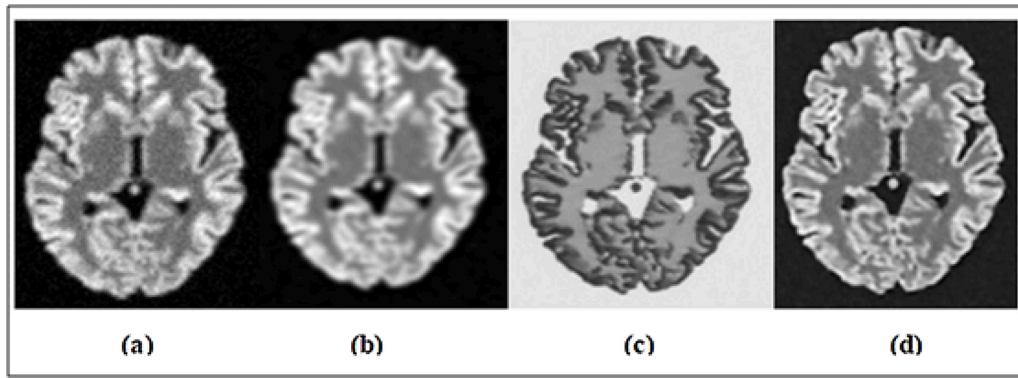
Fig. 12. Results of different reconstruction modules: (a) Input image (b) Results obtained using single reconstruction module using MSE loss function (c) Results obtained using single reconstruction module using MGE loss function (d) Results obtained using proposed method (Two reconstruction modules with multi loss joint strategy).

Table 9
Mean RMSE, PSNR, SSIM scores obtained with different edge detector operators on second reconstruction module.

Edge detection operator	RMSE	PSNR	SSIM
Prewitt	0.0797	22.7303	0.8684
Sobel	**0.0803**	**23.1494**	**0.9032**
Positive Laplacian	0.0869	22.297	0.3757
Negative Laplacian	0.0867	22.8133	0.4112

results were analyzed by assigning values greater and less than the default values. The initial values assigned for learning rate, batch size, epoch are 0.001, 32 and 100 respectively. Detailed analysis of the experiments is given in Tables 11–13. The proposed method has been found to offer the best results with a learning rate equal to 0.0003, a batch size equal to 32 and an epoch equal to 100.

The learned model is validated using simulated test data and Fig. 13

shows the samples of reference images, low quality images with different SNR and respective reconstructed images using the proposed model. The validated model is used to assess its suitability for clinical applications and investigations were conducted using clinical data from representative clinical images. Fig.14 depicts the samples of low quality clinical ASL images along with the processed images. Since high quality clinical ASL images are not available for comparing and validating our results, a visual quality assessment strategy [17] is adopted and the reliability of our model is expressed in terms of Visual Quality Score (VQS).

The processed clinical images were assessed independently and simultaneously by two experienced neuro radiologists in the clinical PCASL interpretation, blinding the medical history of the patients. The observers were asked to analyze the quality of ASL perfusion images based on the different quality parameters such as contrast component, noise effect and edge clarity. For each image, contrast component of GM and GM- WM differentiation were performed. A list of quality parame-

Table 10
Mean RMSE, PSNR, SSIM scores obtained for optimizers Adam, SGD and RMSProp in the proposed system.

	Type of Optimizer								
	Adam			SGD			RMSprop		
SNR	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM
20	**0.0803**	**23.1494**	**0.9032**	0.1461	18.4866	0.8304	0.1382	18.5947	0.8545
15	0.0873	22.4144	0.8187	0.1601	17.9365	0.7729	0.1484	18.0532	0.8138
10	0.0985	21.3694	0.5386	0.1793	17.4192	0.7096	0.166	17.2648	0.7579

Table 11
Mean RMSE, PSNR, SSIM scores obtained with different learning rate in the proposed system.

	Learning rate														
	0.0001			0.0003			0.0005			0.001			0.01		
SNR	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM
20	0.0817	22.7159	0.6761	**0.0803**	**23.1494**	**0.9032**	0.0806	22.2465	0.674	0.0975	21.94	0.4462	0.487	9.5735	−0.09
15	0.0888	21.3422	0.7845	0.0873	22.4144	0.8187	0.0891	21.9337	0.7996	0.1105	20.0394	0.3424	0.4899	9.483	−0.09
10	0.0996	20.5755	0.4325	0.0985	21.3694	0.5386	0.0989	20.3302	0.4784	0.1421	19.45	0.2821	0.4913	9.414	−0.08

Table 12
Mean RMSE, PSNR, SSIM scores obtained with different epochs in the proposed system.

	Number of Epochs														
	50			100			200			300			500		
SNR	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM
20	0.0976	18.5678	0.3987	**0.0803**	**23.1494**	**0.9032**	0.0805	22.9868	0.8723	0.0803	22.0374	0.7599	0.0822	21.6177	0.4591
15	0.996	17.5671	0.3128	0.0873	22.4144	0.8187	0.0883	22.3181	0.5521	0.0883	22.2477	0.4733	0.0886	21.1303	0.394
10	0.1458	16.5679	0.2986	0.0985	21.3694	0.5386	0.0986	21.4951	0.3839	0.0988	21.1864	0.3719	0.0993	20.6288	0.2855

Table 13
Mean RMSE, PSNR, SSIM metrics results on simulated dataset with different batch size for training proposed system.

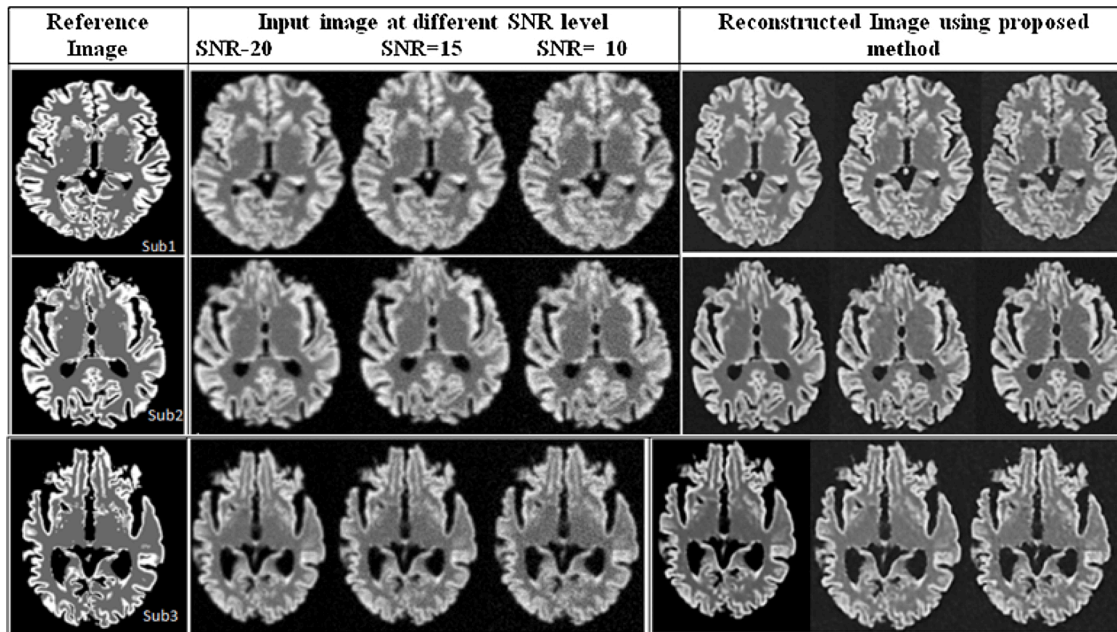|												
					Batch Size							
		16			32			64			128	
SNR	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM
20	0.0802	22.0201	0.5988	**0.0803**	**23.1494**	**0.9032**	0.0824	22.6864	0.4898	0.0816	22.2206	0.4118
15	0.0866	21.1425	0.4161	0.0873	22.4144	0.8187	0.0894	21.1946	0.3532	0.0899	21.6996	0.3613
10	0.0953	20.8399	0.3512	0.0985	21.3694	0.5386	0.1005	20.6884	0.2916	0.1038	20.6455	0.3071

Fig. 13. Results of proposed model on simulated images of different subjects.
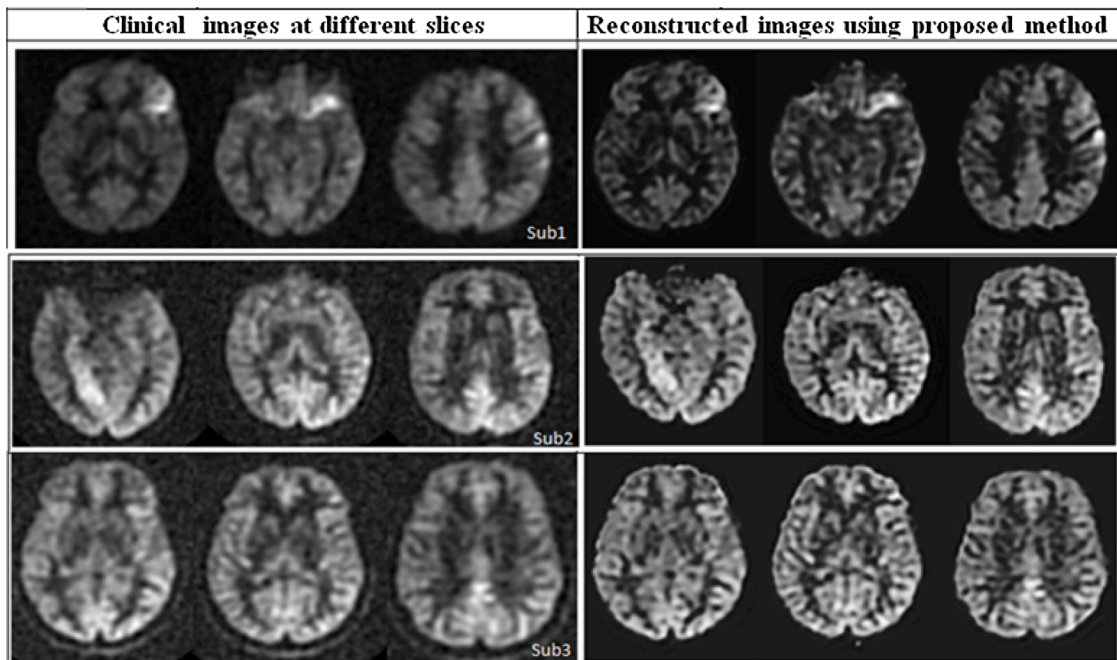

Fig. 14. Results of clinical data for several subjects at different slices using the proposed model.

Table 14
Visual Quality Score (VQS) for different quality parameters and grading criteria.

Visual Quality Parameters	Visual Quality Score (VQS)	Grading Criteria	Subtotal
1. Contrast Component	2	Clearly Visible contrast	
1.1 Grey Matter	1	Unclear contrast	0–4
1.2 Grey Matter-White Matter Differentiation	0	No visible contrast	
	2	Reduced	
2. Change in Noise Effect	1	No change	0–2
	0	Increased	
	2	Increased	
3. Increase in Edge Clarity	1	No change	0–2
	0	Blurred	
Grand Total			**0–8**

ters and VQS are shown in Table 14. The VQS for each category ranges from 0 to 2, with a high score indicating that the processed image has improvements over the clinical images. The proposed approach was used to process 288 slices perfusion images at random. Out of the 288 slices, 72 slices were generated, utilizing spiral readout 4 arms × 512 samples, resulting in a very lowresolution image with poor interpretability.

Fig. 15 represents the graphical analysis of the radiological score by two radiologists. An average score of 1.6, 1.7, 1.6, and 1.8 is obtained for different quality criteria. It has been observed that the proposed algorithm gives more promising results in terms of VQS. Inter observer variation of two radiologists was evaluated using Cohen's kappa (κ) coefficient by constructing cross tab. The κ value obtained for each of the grading criteria is consolidated in Table 15 with its approximate significance p-value [62]. Out of various κ value interpretation methods available [8,37], this work uses the guidelines suggested by Landis and Koch [37]. According to this approach, the range of the κ between 0.41 and 0.60 suggests moderate agreement, 0.61 to 0.80 shows good agreement, and 0.81–1 indicates very good agreement. Since the κ values arrived at corresponding to all the quality criteria is greater than 0.629, this can be interpreted as a favourable result.

4.5. Analysis on the effect of proposed method on PV correction techniques on simulated ASL MR Images

Deep-ASL ENHANCE algorithm is used as a preprocessing step in PV Correction. Comparing its impact on LR with and without the Deep-ASL ENHANCE algorithm is critical. Experiments were conducted as illustrated in Fig. 8 for comparison using simulated ASL images. Accuracy of PV Correction using LR is typically influenced by two factors, the effects of noise in input ASL data and its resolution mismatch due to the unmatched PSF between structural and ASL images [49]. These effects are analyzed experimentally using RMSE as a performance measure. Since this is a simulation experiment, the true value of GM CBF for RMSE analysis is set to 60 as GM value of 60 is used for the generation of ASL images.
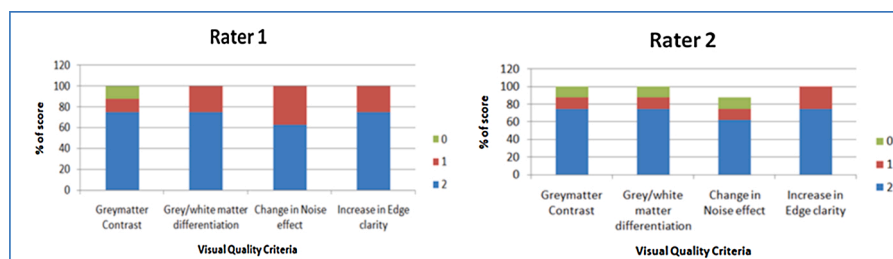
4.5.1. Effect of noise on ASL data in PV Correction with and without Deep-ASL ENHANCE

Experiments were conducted by varying SNR levels on input ASL data as 5, 10, 15, 20 and infinity for analyzing the effect of incorporating Deep-ASL ENHANCE block for PV Correction. PV Correction using the LR method, DEEP-ASL ENHANCE method and combined Deep-ASL ENHANCE and LR was investigated. From Table 16, it can be seen that performance of LR alone is better compared to Deep-ASL ENHANCE method. Deep-ASL ENHANCE approach enhances image quality more effectively (Table 16), but voxelwise correction is not performed. As a consequence, the RMSE value obtained is substantially higher than that of LR. The RMSE is significantly reduced when Deep-ASL ENHANCE is introduced as a preprocessing step prior to LR, particularly for lower SNR values which is realistic in clinical point of view.

4.5.2. Effects of blurring on PV correction with and without Deep-ASL ENHANCE

To study this effect, noise free simulated ASL data is used and the blurring effect is incorporated with a 3D Gaussian blurring kernel cor-

Table 15
Kappa Coefficient (κ) with p-value of different visual quality parameters.

Visual Quality Parameter	Kappa Coefficient with p-value
Grey matter Contrast	1 (p < 0.001)
Grey matter white matter differentiation	0.6923 (p = 0.008)
Reduction in noise effects	0.7407 (p = 0.008)
Increase in edge clarity	1 (p = 0.005)

Table 16
Mean RMSE value between reference GM CBF and the images obtained after LR algorithm, Deep-ASL ENHANCE algorithm and combined Deep-ASL ENHANCE and LR algorithm by varying different SNR levels of ASL simulated data.

Method	SNR = inf	SNR = 20	SNR = 15	SNR = 10	SNR = 5
LR	0.0650	0.0667	0.0674	0.0684	0.0724
Deep-ASL ENHANCE	0.0731	0.0735	0.0737	0.0744	0.0786
LR after Deep-ASL ENHANCE	**0.0648**	**0.0654**	**0.0666**	**0.676**	**0.0687**

Table 17
Mean RMSE value between reference GM CBF and the images obtained after LR algorithm, Deep-ASL ENHANCE algorithm and combined Deep-ASL ENHANCE and LR algorithm by varying different blurring levels(FWHM) on ASL simulated data with SNR = inf.

Method	FWHM = 1	FWHM = 1.5	FWHM = 3	FWHM = 4.5	FWHM = 6
LR	0.0656	0.0700	0.0804	0.0849	0.0850
Deep-ASL ENHANCE	0.0719	0.0740	0.0900	0.0949	0.0939
LR after Deep-ASL ENHANCE	**0.0649**	**0.0662**	**0.0773**	**0.0826**	**0.0829**

Fig. 15. Visual Quality Score of two radiologists against different visual quality parameters.

Table 18
Mean RMSE value between reference GM CBF and the images obtained after LR algorithm, Deep-ASL ENHANCE algorithm and combined Deep-ASL ENHANCE and LR algorithm by varying different SNR levels on ASL simulated data with different FWHM values.

FWHM	Method	SNR = inf	SNR = 20	SNR = 15	SNR = 10	SNR = 5
1	LR	0.0656	0.0672	0.0689	0.0693	0.0712
	Deep-ASL ENHANCE	0.0719	0.0726	0.0738	0.0747	0.0789
	LR after Deep-ASL ENHANCE	**0.0649**	**0.0658**	**0.0663**	**0.0675**	**0.0698**
1.5	LR	0.0712	0.0723	0.0731	0.0741	0.0755
	Deep-ASL ENHANCE	0.074	0.0747	0.0756	0.076	0.0811
	LR after Deep-ASL ENHANCE	**0.0662**	**0.067**	**0.0679**	**0.0685**	**0.0719**
3	LR	0.0804	0.0825	0.0844	0.0858	0.0867
	Deep-ASL ENHANCE	0.09	0.0923	0.0934	0.0946	0.0968
	LR after Deep-ASL ENHANCE	**0.0773**	**0.0792**	**0.0802**	**0.0817**	**0.0823**
6	LR	0.0858	0.0864	0.0887	0.0899	0.0918
	Deep-ASL ENHANCE	0.0939	0.0943	0.0956	0.0963	0.0976
	LR after Deep-ASL ENHANCE	**0.0829**	**0.0839**	**0.0854**	**0.0867**	**0.0905**

responding to a collection of standard deviation value σ according to the Eq. (23) with different FWHM values

$$\sigma = \frac{FWHM}{\sqrt{8ln2}} \qquad (23)$$

FWHM values in the range 1–6 were chosen from the literature [77] to account for low and high blurring effects. The results corresponding to a subset of values falling in this range are presented in Table 17. Using the above set data, accuracy of PV correction using LR alone, Deep-ASL ENHANCE method, and combined Deep-ASL ENHANCE and LR were also investigated and shown in Table 17.

Similar observations are also made in the case of combined Deep-ASL ENHANCE and LR method for any degree of blurring. Comparing Table 16 and Table 17, an increased RMSE value can be noted in the case of high blurring effects. This emphasizes the fact that blurring phenomena affect PV correction accuracy in a more significant way than that of noise. Table 18 shows the combined effect of blurring and noise, in a realistic manner as seen in clinical cases, for different levels of SNR and FWHM values. When proposed method is used for preprocessing steps in LR algorithm, significant decrease in RMSE value is arrived at.

Deep-ASL ENHANCE, as demonstrated in the proposed work, is a potential strategy for improving the quality of ASL images, which reduces quantification error in CBF values. It is also observed that data acquired using spiral readout 4 arms × 512 samples did not provide satisfactory VQS. This is due to the poor resolution and interpretability of the input, clinical images, necessitating further investigations.

5. Conclusion

The quality of ASL image is enhanced by the proposed Deep-ASL ENHANCE algorithm that uses a multi-loss joint strategy with the residual dense block as the basic building block. This technique addresses the problems encountered in accurately quantifying the CBF value, which is hampered by low SNR and low resolution, when used for clinical diagnosis. ASL images simulated from structural data of ADNI2 were used to train the deep learning model, since the available clinical images will not suffice the huge data requirements. The trained model gave most promising results on simulated as well as clinical ASL test datasets and these results are validated in terms of evaluation matrices. The average scores corresponding to the simulated datasets for SNR = 20 are 0.0803, 23.1494, and 0.9032 for RMSE, PSNR, and SSIM respectively. The performance of the trained model on clinical data is evaluated by two independent radiologists. The visual quality scores thus arrived at, gave an average score of 1.6, 1.7, 1.6 and 1.8 for different quality grading criteria such as grey matter contrast, grey matter-white matter differentiation, reduction in noise effects, and increase in edge clarity. The inter observer variability of the radiologists is assessed using Cohen's kappa coefficient, which yielded κ values as 1,

0.6923, 0.7407 and 1 for the above mentioned quality grading criteria, indicating favourable results. It is also observed that due to very poor resolution and interpretability of the input clinical images, data obtained from subjects using spiral readout 4 arms × 512 samples couldn't give acceptable results, which will be addressed in the future. A deep learning model trained with multiscale resolution may provide improved results for these types of images. Since voxel wise GM CBF is important in most of the clinical diagnosis, partial volume correction using the LR algorithm is performed on the enhanced simulated image obtained through Deep-ASL ENHANCE algorithm. The impact of different noise levels and blurring effects of PVC has been investigated and it is found that Deep-ASL ENHANCE, as a preprocessing step to LR algorithm, reduces the CBF quantification error in terms of RMSE values.

CRediT authorship contribution statement

Shyna A: Methodology, Software, Formal Analysis, Writing - Original Draft. **Dr. Ushadevi Amma C.:** Conceptualization, Methodology, Writing - Review & Editing, Supervision. **Dr. Ansamma John:** Conceptualization, Methodology, Writing - Review & Editing, Validation, Formal Analysis, Data Curation. **Dr. C. Kesavadas:** Data Validation and Radiological Analysis of results, Visual Quality Score Computation. **Dr. Bejoy Thomas:** Data Validation and Radiological Analysis of results, Visual Quality Score Computation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] André Ahlgren, Ronnie Wirestam, Esben Petersen, Freddy Ståhlberg, Linda Knutsson, Partial volume correction of brain perfusion estimates using the inherent signal data of time-resolved arterial spin labeling, NMR Biomed. 27 (2014) 1112–1122.

[2] Hyun-Seo Ahn, Hee Yu, Hyo Kwak, Sung-Hong Park, Assessment of renal perfusion in transplanted kidney patients using pseudo-continuous arterial spin labeling with multiple post-labeling delays, Eur. J. Radiol. 130 (2020), 109200.

[3] David Alsop, John Detre, Xavier Golay, Matthias Günther, J. Hendrikse, G. Zaharchuk, Recommended implementation of arterial spin-labeled perfusion MRI for clinical applications: a consensus of the ISMRM perfusion study group and the European consortium for ASL in dementia, Magn. Reson. Med. 116 (2014) 102–116.

[4] Iris Asllani, Ajna Borogovac, Truman Brown, Regression algorithm correcting for partial volume effects in arterial spin labeling MRI, Magn. Reson. Med. 60 (2008) 1362–1371.

[5] Restom Behzadi, Liu. Liau, A component based noise correction method (CompCor) for BOLD and perfusion based fMRI, NeuroImage 37 (2007) 90–101.

[6] Adnan Bibic, Linda Knutsson, Freddy Ståhlberg, Ronnie Wirestam, Denoising of arterial spin labeling data: wavelet-domain filtering compared with Gaussian smoothing, Magma (New York, N.Y.) 23 (2010) 125–137.

[7] Bihan Le, E. Breton, D. Lallemand, M.L. Aubin, J. Vignaud, M. Laval-Jeantet, Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging, Radiology 168 (2) (1988) 497–505.

[8] J. Bland, Douglas. Altman, Measuring agreement in method comparison studies, Stat. Methods Med. Res. 8 (1999) 135–160.

[9] L.éon. Bottou, Large-scale machine learning with stochastic gradient descent, Proc. of COMPSTAT (2010).

[10] Mustapha Bouhrara, Diana Lee, Abinand Rejimon, Christopher Bergeron, Richard Spencer, Spatially adaptive unsupervised multispectral nonlocal filtering for improved cerebral blood flow mapping using arterial spin labeling magnetic resonance imaging, J. Neurosci. Methods 309 (2018) 121–131.

[11] Richard Buxton, Lawrence Frank, Eric Wong, Bettina Siewert, Steven Warach, Robert Edelman, A general kinetic model for quantitative perfusion imaging with arterial spin labeling, Magn. Reson. Med. 40 (1998) 383–396.

[12] M. Chappell, A. Groves, B. MacIntosh, M. Donahue, Peter Jezzard, M. Woolrich, Partial volume correction of multiple inversion time arterial spin labeling MRI data, Magn. Reson. Med. 65 (2011) 1173–1183.

[13] Toru Chikui, Makoto Obara, Arjan Simonetti, Masahiro Ohga, Shoichi Koga, Shintaro Kawano, Yoshio Matsuo, Takeshi Kamintani, Tomoko Shiraishi, Erina Kitamoto, Katsumasa Nakamura, Kazunori Yoshiura, The principal of dynamic contrast enhanced MRI, the method of pharmacokinetic analysis, and its application in the head and neck region, Int. J. Dent. 2012 (2012), 480659.

[14] John A. Detre, John S. Leigh, Donald S. Williams, Alan P. Koretsky, Perfusion imaging, Magn. Reson. Med. 23 (1) (1992) 37–45.

[15] John Detre, Hengyi Rao, Danny Wang, Yufen Chen, Ze Wang, Applications of arterial spin labeled MRI in the brain, J. Magn. Reson. Imaging 35 (2012) 1026–1037.

[16] Chao Dong, Chen Change Loy, Kaiming He, Xiaoou Tang, Image super-resolution using deep convolutional networks, IEEE Trans. Pattern Anal. Mach. Intell. 38 (2014) 295–307.

[17] Sameeha Fallatah, Francesca Pizzini, Beatriz Gómez-Ansón, J. Magerkurth, Enrico De Vita, Sotirios Bisdas, H. Jäger, Henri Mutsaerts, Xavier Golay, A visual quality control scale for clinical arterial spin labeling images, Eur. Radiol. Exp. 2 (2018).

[18] Audrey Fan, Hesamoddin Jahanian, Samantha Holdsworth, Greg Zaharchuk, Comparison of cerebral blood flow measurement with [150]-water positron emission tomography and arterial spin labeling magnetic resonance imaging: a systematic review, J. Cerebral Blood Flow Metab. 36 (2016) 842–861.

[19] Yuchen Fan, Humphrey Shi, Ding Liu, Wei Han, Haichao Yu, Zhangyang Wang, Xinchao Wang, Balanced Two-Stage Residual Networks for Image Super-Resolution, 2017, pp. 1157–1164.

[20] Ruogu Fang, Junzhou Huang, Wen-Ming Luh, A Spatio-temporal Low-rank Total Variation Approach for Denoising Arterial Spin Labeling MRI Data., 2015, pp. 498–502.

[21] Yulin Ge, Mayur Patel, Qun Chen, Elan Grossman, Ke Zhang, Laura Miles, James Babb, Joseph Reaume, Robert Grossman, Assessment of thalamic perfusion in patients with mild traumatic brain injury by true FISP arterial spin labeling MR imaging at 3T, Brain injury 23 (2009) 666–674.

[22] Enhao Gong, Jia Guo, Jiang Liu, Audrey Fan, John Pauly, Greg Zaharchuk, Deep Learning and Multi-contrast Based Denoising for low-SNR Arterial Spin Labeling (ASL) MRI, pages 21, 2020.

[23] Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing, Pearson, 2018.

[24] Antonio Gulli, Sujit. Pal, Deep Learning With Keras, Packt Publishing Ltd, 2017.

[25] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Deep Residual Learning for Image Recognition, 2016, pp. 770–778.

[26] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Identity Mappings in Deep Residual Networks, 2016, pp. 630–645.

[27] Geoffrey Hinton, Coursera Neural Networks for Machine Learning Lecture 6, 2018.

[28] Ying Hu, Fajin LV, Qi Li, Rongbo Liu, Effect of post-labeling delay on regional cerebral blood flow in arterial spin-labeling MR imaging, Medicine 99 (2020), e20463.

[29] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, Kilian Q. Weinberger, Densely connected convolutional networks, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017) 2261–2269.

[30] Sergey Ioffe, Christian. Szegedy, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, 2015, pp. 448–456.

[31] Geon-Ho Jahng, Ka-Loh Li, Leif Ostergaard, Fernando Calamante, Perfusion magnetic resonance imaging: a comprehensive update on principles and techniques, Kor. J. Radiol. 15 (2014) 554–577.

[32] Mark Jenkinson, Stephen. Smith, Global optimisation method for robust affine registration of brain images, Med. Image Analysis 5 (2001) 143–156.

[33] Karen Simonyan, Andrew Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv (2014) 1409–1556.

[34] Jiwon Kim, Jung Lee, Kyoung Lee, Accurate Image Super-Resolution Using Very Deep Convolutional Networks, 2016, pp. 1646–1654.

[35] Ki Kim, Seung Choi, Sung-Hong Park, Improving arterial spin labeling by using deep learning, Radiology 287 (2017), 171154.

[36] Diederik P. Kingma, Jimmy Ba, Adam: a method for stochastic optimization, CoRR, abs/1412.6980 (2015).

[37] J. Landis, Gary. Koch, The measurement of observer agreement for categorical data, Biometrics 33 (1977) 159–174.

[38] Østergaard Leif, Principles of cerebral perfusion imaging by bolus tracking, J. Magnet. Reson. Imag. 22 (2005) 710–717.

[39] Yiran Li, Sudipto Dolui, Dan-Feng Xie, Ze Wang, Priors-guided slice-wise adaptive outlier cleaning for arterial spin labeling perfusion MRI, J. Neurosci. Methods 307 (2018) 248–253.

[40] Xiaoyun Liang, Alan Connelly, Fernando Calamante, Improved partial volume correction for single inversion time arterial spin labeling data, Magnet. Reson. Med. 69 (2013) 531–537.

[41] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, Kyoung Mu Lee, Enhanced deep residual networks for single image super-Resolution, 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (2017) 1132–1140.

[42] Che Lin, Ying-Chi Tseng, Hui-Ling Hsu, Chi-Jen Chen, David Chen, Feng-Xian Yan, Wen-Ta Chiu, Arterial spin labeling perfusion study in the patients with subacute mild traumatic brain injury, PloS one 11 (2016), e0149109.

[43] Lu Zhengyang, Ying Chen, Single Image Super Resolution Based on a Modified U-net With Mixed Gradient Loss, 2019.

[44] Alice Lucas, Michael Iliadis, Rafael Molina, Aggelos Katsaggelos, Using deep neural networks for inverse problems in imaging: beyond analytical methods, IEEE Signal Processing Magazine 35 (2018) 20–36.

[45] Petros Martirosian, R. Pohmann, Christina Schraml, Martin Schwartz, Thomas Küstner, Nina Schwenzer, Klaus Scheffler, Konstantin Nikolaou, Fritz Schick, Spatial-temporal perfusion patterns of the human liver assessed by pseudo-continuous arterial spin labeling MRI, Zeitschrift für Medizinische Physik 29 (2018) 173–183.

[46] C.édric Meurée, Pierre Maurel, Jean-Christophe Ferré, Christian Barillot, Patch-based super-resolution of arterial spin labeling magnetic resonance images, NeuroImage 189 (2019) 85–94.

[47] Christopher Thomas, Deep Learning Image Enhancement Insights on Loss Function Engineering, 2020.

[48] Aghogho Odudu, Fabio Nery, Anita Harteveld, Roger Evans, Douglas Pendse, Charlotte Buchanan, Susan Francis, Maria Fernandez-Seara, Arterial spin labelling MRI to measure renal perfusion: a systematic review and statement paper, Nephrol. Dialysis Transplant. 33 (2018) ii15–ii21.

[49] Ruth Oliver, Ben Thomas, David Thomas, Xavier Golay, Improved Partial Volume Correction of ASL Images Using 3D Kernels, 2012.

[50] David Owen, Andrew Melbourne, Zach Eaton-Rosen, David Thomas, Neil Marlow, Jonathan Rohrer, S.ébastien Ourselin, Deep Convolutional Filtering for Spatio-Temporal Denoising and Artifact Removal in Arterial Spin Labelling MRI, 2018, pp. 21–29.

[51] Jan Petr, Jean-Christophe Ferre, jean-yves Gauvrit, Christian Barillot, Improving arterial spin labeling data by temporal filtering, Progress in Biomedical Optics and Imaging - Proceedings of SPIE 7623 (2010).

[52] Jan Petr, Georg Schramm, Frank Hofheinz, Jens Maus, J.örg Hoff, Partial volume correction in arterial spin labeling using a look-locker sequence, Magnet. Reson. Med. 70 (2013) 1535–1543.

[53] Chi-Hieu Pham, Deep Learning for Medical Image Super Resolution and Sementation, 2018.

[54] Joost Putten, Farhad. Zanjani, Multi-scale Ensemble of ResNet Variants, 2021, pp. 115–119.

[55] Mark S. Shiroishi, Gloria Castellazzi, Jerrold L. Boxerman, Francesco D'Amore, Marco Essig, Thanh B. Nguyen, James M. Provenzale, David S. Enterline, Nicoletta Anzalone, Arnd Dörfler, Àlex Rovira, Max Wintermark, Meng Law, Principles of T2*-weighted dynamic susceptibility contrast MRI technique in brain tumor imaging, J. Magnet. Reson. Imag. 41 (2014).

[56] Stefan Spann, Kamil Kazimierski, Christoph Aigner, Markus Kraiger, Kristian Bredies, Rudolf Stollberger, Spatio-temporal TGV denoising for ASL perfusion imaging, NeuroImage 157 (2017) 81–96.

[57] Sudipto Dolui, Ze Wang, Russell T. Shinohara, David A. Wolk, John A. Detre, Structural Correlation-based Outlier Rejection (SCORE) algorithm for arterial spin labeling time series, J. Magn. Reson. Imag. 45 (2017) 1786–1797.

[58] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich, Going deeper with convolutions, 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2015) 1–9.

[59] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, Zbigniew Wojna, Rethinking the inception architecture for computer vision, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016) 2818–2826.

[60] Huan Tan, JosephA. Maldjian, Jeffrey Pollock, Jonathan Burdette, AndrewR. Deibler, RobertA. Kraft, A fast, effective filtering method for improving clinical pulsed arterial spin labeling MRI, J. Magnet. Reson. Imag. 29 (2009) 1134–1139.

[61] Nicholas Telischak, John Detre, Greg Zaharchuk, Arterial spin labeling MRI: clinical applications in the brain, J. Magnet. Reson. Imag. 41 (2014) 1165–1180.

[62] Paul Thorman, Uncertainty and Statistics, 2020, pp. 29–52.

[63] Chunwei Tian, Lunke Fei, Wenxian Zheng, Yong xu, Wangmeng Zuo, Chia-Wen Lin, Deep learning on image denoising: an overview, Neural Networks 131 (2020) 251–275.

[64] Tong Tong, Gen Li, Xiejie Liu, Qinquan Gao, Image super-resolution using dense skip connections, 2017 IEEE International Conference on Computer Vision (ICCV) (2017) 4809–4817.

[65] Duygu Tosun, Pouria Mojabi, Michael Weiner, Norbert Schuff, Joint analysis of structural and perfusion MRI for cognitive assessment and classification of Alzheimer's disease and normal aging, NeuroImage 52 (2010) 186–197.

[66] Cagdas Ulas, Giles Tetteh, Stephan Kaczmarz, Christine Preibisch, Bjoern Menze, DeepASL: Kinetic Model Incorporated Loss for Denoising Arterial Spin Labeled MRI Via Deep Residual Learning, 2018, pp. 30–38.

[67] Richard Vitti, Gadolinium-based contrast agents and nephrogenic systemic fibrosis, Radiology 250 (2009) 959–960.

[68] Jack Wells, David Thomas, Martin King, Alan Connelly, Mark Lythgoe, Fernando Calamante, Reduction of errors in ASL cerebral perfusion and arterial transit time maps using image de-noising, Magn. Reson. Med. 64 (2010) 715–724.

[69] D. Williams, John Detre, Jr Leigh, A.P. Koretsky, Magnetic resonance imaging of perfusion using spin inversion of arterial water, Proc. Natl. Acad. Sci. U. S. A. 89 (1992) 212–216.

[70] Allemeije Wink, Jos Roerdink, Denoising functional MR images: a comparison of wavelet denoising and gaussian smoothing, IEEE Trans. Med. Imag. 23 (2004) 374–387.

[71] David Wolk, John. Detre, Arterial spin labeling MRI: an emerging biomarker for alzheimer's disease and other neurodegenerative conditions, Current opinion in neurology 25 (2012) 421–428.

[72] Danfeng Xie, Yiran Li, Hanlu Yang, Li Bai, Tianyao Wang, Fuqing Zhou, Lei Zhang, Ze Wang, Denoising arterial spin labeling perfusion MRI with deep machine learning, Magnetic Resonance Imaging 68 (2020) 95–105.

[73] K. Yamashita, K. Markov, Medical Image Enhancement Using Super Resolution Methods, in: Computational Science – ICCS 2020 20th International Conference, Amsterdam, The Netherlands, June 3–5, 2020. Proceedings, Part V. 2020 May; 12141: 496-508. PMCID: PMC7302556.

[74] Yu Jiahui, Yuchen Fan, Jianchao Yang, Xu Ning, Zhaowen Wang, Xinchao Wang, Thomas Huang, Wide Activation for Efficient and Accurate Image Super-Resolution, 2018.

[75] Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, Yun Fu, Residual dense network for image super-Resolution, 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (2018) 2472–2481.

[76] Y.Y. Zhang, Michael Brady, S.A. Smith, Segmentation of brain MR images through a hidden markov random field model and the expectation maximization algorithm, IEEE Trans. Med. Imag. 20 (2001) 45–57.

[77] Moss Zhao, Melvin Mezue, Andrew Segerdahl, Thomas Okell, Irene Tracey, Yingyi Xiao, Michael Chappell, A systematic study of the sensitivity of partial volume correction methods for the quantification of perfusion from pseudo-continuous arterial spin labeling MRI, NeuroImage 162 (2017) 384–397.

[78] Zheng Li, Qingping Liu, Yiran Li, Qiu Ge, Yuanqi Shang, Donghui SongZe, Ze Wang, Jun Shi, A Two-Stage Multi-Loss Super-Resolution Network for Arterial Spin Labeling Magnetic Resonance Imaging, vol. 11766, Springer, Cham, Shenzhen, China, 2019, pp. 12–20.

[79] Zhao Zhong, Junjie Yan, Wei Wu, Jing Shao, Cheng-Lin Liu, Practical block-wise neural network architecture generation, 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (2018) 2423–2432.

[80] Hancan Zhu, Jian Zhang, Ze Wang, Arterial spin labeling perfusion MRI signal denoising using robust principal component analysis, J. Neurosci. Methods 295 (2017) 10–19.

Dr Ushadevi Amma C: received B. Tech degree in Electrical Engineering from University of Kerala in 1987, M.Tech in Instrumentation Engineering from IIT Kharagpur in 1997 and Ph.D in Instrumentation Engineering from IISC Bangalore in 2007. She is currently working as Professor in Department of Electronics and Communications Engineering, Amrita Vishwa Vidyapeetham, Amritapuri campus, Kerala, India. Her research interests include Biomedical Instrument Development, Medical Image acquisition and analysis.

Dr Ansamma John: received B. Tech degree in Computer Science and Engineering from University of Kerala in 1990, M. Tech in Computer Science and Information Technology from IIT Kharagpur in 2000 and Ph.D from University of Kerala in 2017. She is currently working as Professor in Department of Computer Science and Engineering, TKM College of Engineering, Kollam, Kerala, India. Her research interests include Data Mining, Text mining, Biomedical Image Processing, Image Mining and Computational Geometry.

Dr. C. Kesavadas is the Professor of Radiology at Sree Chitra Tirunal Institute for Medical Sciences & Technology, a teaching institute of national importance under the Govt. of India. He also serves as the Officiating Dean at the Institute. His research interest includes advanced MR imaging in epilepsy, brain tumors and neurodegenerative diseases. He graduated from Calicut Medical College in 1989, did post-graduation in Radiology from Medical College, Thiruvananthapuram in 1994.

Dr. Bejoy Thomas: received his MBBS, MD, DNB and PDCC from Mahatma Gandhi University, Gujarat University, National Board and SCTIMST in India respectively. He also successfully completed a fellowship in Advanced Neuroimaging from the Katholieke Universiteit, Leuven, Belgium and a clinical fellowship in Pediatric Neuroimaging from the Hospital for Sick Children, University of Toronto, Canada. Currently he is the professor and head of the department of Imaging Sciences and Interventional Radiology at SCTIMST, Thiruvananthapuram, Kerala, India. His research interests include functional neuroimaging, development and validation of quantitative imaging biomarkers in neurological disorders and pediatric neuroimaging.

Prof. Shyna A: received B. Tech degree in Computer Science and Engineering from University of Kerala in 2012, M. Tech in Computer Science and Engineering with specialization in Image Processing from Cochin University of Science and Technology in 2014. She is currently working as Assistant Professor in Department of Computer Science and Engineering, TKM College of Engineering, Kollam, Kerala, India. Her research interests include Biomedical Image Processing, Machine Learning and Soft Computing.