Deep Learning with Multiresolution Handcrafted Features for Brain MRI Segmentation

Imene Mecheter[a], Maysam Abbod[a], Abbes Amira[b,c], Habib Zaidi[d,e,f,g]

[a]*Department of Electronic and Computer Engineering, Brunel University London, Uxbridge, UK*
[b]*Department of Computer Science, University of Sharjah, Sharjah, UAE*
[c]*Institute of Artificial Intelligence, De Montfort University, Leicester, UK*
[d]*Division of Nuclear Medicine and Molecular Imaging, Department of Medical Imaging, Geneva University Hospital, Geneva, Switzerland*
[e]*Geneva University Neurocenter, Geneva University, Geneva, Switzerland*
[f]*Department of Nuclear Medicine and Molecular Imaging, University Medical Center Groningen, University of Groningen, Groningen, Netherlands*
[g]*Department of Nuclear Medicine, University of Southern Denmark, Odense, Denmark*

Abstract

The segmentation of magnetic resonance (MR) images is a crucial task to create pseudo computed tomography (CT) images which are used to achieve positron emission tomography (PET) attenuation correction. One of the main challenges of creating pseudo CT images is the difficulty to obtain an accurate segmentation of the bone tissue in brain MR images. Deep convolutional neural networks (CNNs) have been widely and efficiently applied to perform MR image segmentation. The aim of this work is to propose a segmentation approach that combines multiresolution handcrafted features with CNN-based features to add directional properties and enrich the set of features to perform segmentation. The main objective is to efficiently segment the brain into three tissue classes: bone, soft tissue, and air. The proposed method combines non subsampled Contourlet (NSCT) and non subsampled Shearlet (NSST) coefficients with CNN's features using different mechanisms. The entropy value is calculated to select the most useful coefficients and reduce the input's dimensionality. The segmentation results are evaluated using fifty clinical brain MR and CT images by calculating the precision, recall, dice similarity coefficient (DSC), and Jaccard similarity coefficient (JSC). The results are also compared to other methods reported in the literature. The DSC of the bone class is improved from 0.6179 ± 0.0006 to 0.6416 ± 0.0006. The addition of multiresolution features of NSCT and NSST with CNN's features demonstrates promising results. Moreover, NSST coefficients provide more useful information than NSCT coefficients.

Keywords: CNN, Contourlet, Shearlet, Segmentation, MR

1. Introduction

The segmentation of magnetic resonance (MR) images has various applications in the process of disease's diagnosis, treatment planning, and quantification of image-derived metrics. One of these applications is the generation of pseudo computed tomography (CT) images for positron emission tomography (PET) attenuation correction.

Due to the loss of photons during PET acquisition process, the acquired PET images suffer from attenuation which should be corrected to achieve the full potential of PET quantification and disease's characterization. The common way to address PET attenuation is using the attenuation coefficients from CT images since there is a direct mapping between the CT Hounsfield units and the attenuation coefficients at 511 keV. The main drawback of CT imaging is the radiation exposure which harms children and patients who need to undergo CT scanning regularly [1]. Alternatively, MR images with their excellent soft tissue contrast can replace CT images, yet this replacement is not straightforward since there is no direct mapping between the MR proton intensities and the linear attenuation coefficients at 511 keV [2].

Among the MR-based attenuation correction methods [3, 4], MR images segmentation is considered a robust and sim-ple approach that has been adopted in commercial PET/MR scanners [5]. The segmented MR images are converted to pseudo CT images by assigning a linear attenuation coefficients to each tissue class. The attenuation coefficients generate attenuation maps which are used to reconstruct PET images for photon attenuation. One of the limitations of the segmentation methods is the accurate delineation of bone tissue which is the objective of the proposed method herein.

Convolutional neural networks (CNN) has been widely applied to segment different medical images including MR images. The process of automatic extraction of features using CNN has shown its superiority in various applications to perform different tasks. The CNN-based features remove the burden of creating handcrafted features which includes features extraction and selection tasks. However, in some problems and with specific types of data, the handcrafted features are critical to enhance the performance [6, 7].

The combination of handcrafted features with CNN's features is becoming a popular approach to address various problems in different domains such as image scene geometry recognition [8], classification of working condition in froth flotation [9], signal gesture recognition [10], land cover content-based image retrieval [11], and pedestrian detection

[12]. Moreover, it has been widely used in the medical field to perform classification [13], segmentation [8], and detection [14] tasks. Additionally, this approach has been applied to both images and signals such as MR images [13], histology images [15], iris images [14], electromyographic (EMG) signals [10], and electroencephalogram (EEG) signals [16]. The addition of multiresolution features such as Contourlet and Shearlet coefficients as handcrafted features with CNN's features is able to enrich the set of features with multiscale and multidirectional properties to perform the segmentation task. Various CNN's architectures have been proposed in the literature, yet none of them has focused on the combination of handcrafted features which introduce directional properties with CNN's features.

The aim of this work is to propose a CNN architecture capable of segmenting brain MR images into three tissue classes: bone, soft tissue, and air and enhance the bone class segmentation particularly. The proposed approach aims at combining multiscale and multidirectional features with CNN's based features. Another objective is studying the impact of both low pass and directional sub-bands of non-subsampled Contourlert (NSCT) and non-subsampled Shearlet (NSST) transforms. The main contributions herein are designing a 3D encoder decoder convolutional network with NSCT and NSST coefficients, applying entropy score as a features selection approach, and combining NSCT and NSST features with CNN's features.

The paper is structured as follows. Section 2 reviews briefly the related work on handcrafted features with CNN with a focus on Contourlet and Shearlet transforms. The data description and the applied methods are described in section 3. The conducted experiments with the obtained results are presented in section 4. Finally, the discussion and conclusion are given in section 5.

2. Related Work

Recently, deep CNN has been applied successfully to perform segmentation tasks in different medical imaging applications [17–21]. In particular, brain MR images segmentation has been addressed extensively using multiple architectures of CNN. For instance, a 2D deep generative network which employs P-norm convolutional layers has been proposed by Takrouni et al. [22] to perform brain tumor segmentation task. They have also applied a conditional random field based deep discriminative network to refine the segmentation results. In addition to that, multi-modality MR images have been used to enhance the brain tumor segmentation where a dual path model which consists of DenseNet and features pyramid networks is proposed by Fang et al. [23]. Additionally, a 2D deep convolutional encoder decoder network is proposed by Jiang et al. [24] to segment the brain tumor in MR images by adding an additional encoder that captures the edges features. This network forms a dual-stream decoding architecture that extracts both semantic and edges features. Another dual paths network is proposed by Chai

et al. [25] to segment the gray matter nuclei in 3D MR images. The convolutional encoder decoder network consists of dual encoder paths each of which takes a different resolution of MR patches as input to capture features from different field of views. Likewise, Bose et al. [26] have proposed a deep dense dilated encoder decoder network with additional multi-resolution paths where they have introduced the concept of deep supervision by calculating the loss at each multi-resolution path. The whole model is optimized using the aggregated loss. This architecture has shown its efficiency in performing segmentation of different medical images datasets and brain MR images is one of them. Multiview information of brain MR images is another approach that has been proposed by Li et al. [27] to segment the claustrum structure. Two 2D U-Net based networks are applied to train slices with axial and coronal views. The outputs of each network are aggregated to create a 3D segmentation map.

All above reviewed recent research studies focus on improving the brain segmentation by either employing multipaths deep networks or using multi-resolution and multiviews images to extract additional spatial information. Although these approaches look promising, none of them has investigated the addition of spectral domain features using multi-resolutions transforms. The multi-resolution transforms can be combined as handcrafted features with CNN based spatial features to enhance the segmentation performance.

The most common handcrafted features reported in the literature which are combined with CNN's features are statistical and textural features from the source images. For instance, Shang et al. [13] have combined radiomics statistical and textural features with CNN's features to perform soft tissue tumor classification, Khan et al. [8] have applied a different set of statistical features to segment the brain tumor, and Choudhary et al. [14] have combined another set of statistical features with CNN's features to perform iris spoofing detection.

Multiresolution analysis techniques such as Contourlet and Shearlet transforms provide multiscale and multidirectional features by decomposing the image into low and high frequency sub-bands. Unlike Wavelet transform, they are shift-invariant and can capture rich information with different scales and directions in the spectral domain. The addition of handcrafted features that capture multiscale and multidirectional information has been proposed in the literature to improve the performance of different tasks.

2.1. Contourlet Transform

The Contourlet transform is combined with CNN's features to perform multiple tasks including medical images fusion and multimodality merging [6, 28, 29], natural images classification [30, 31], remote sensing classification [32], breast cancer classification [33], Seismic images denoising [34], and bone age assessment in X-ray images [35].

Li et al. [32] have proposed a complex domain CNN which uses Contourlet filter banks for polarimetric synthetic aperture radar (PolSAR) image classification. They have suggested

to redefine the operations of the conventional CNN's layers in complex domain to provide more useful features. They have also employed non subsampled Contourlet (NSCT) filter with complex CNN filters to obtain multiresolution and multidirectional information. The proposed network has been tested on three PolSAR data sets and shown the efficiency of this method to capture abstract features in different directions and frequency bands. Additionally, Liu et al. [30] have proposed a Contourlet based CNN for SAR image despeckling. They have designed multiple multidirectional and multiscale subnetworks. Each independent subnetwork takes a different Contourlet sub-band as input and trains the network until it reaches a clean image with minimal noise. The inverse Contourlet transform is applied to the output of all subnetworks to obtain the clean image. Different CNN's architectures are applied for each sub-band to fit the properties of each sub-band. The proposed approach has been evaluated using synthetic and real SAR images and validated with the state of the art methods. This method has shown its superiority in suppressing speckle with minimal runtime.

In the medical domain, the employment of the Contourlet transform with CNN is proposed by Nahid et al. [33] to classify histopathological images. They have conducted a comparison in the performance between CNNs that take different sets of inputs such as raw images, statistical information of the Contourlet sub-bands, histogram information, local binary pattern features, and other frequency-domain features. It has been found that the statistical information of the Contourlet coefficients combined with histogram information outperform other approaches in terms of different evaluation metrics. The bone age assessment using X-ray images has been addressed by Liu et al. [35] who have used the NSCT coefficients of the raw images as inputs to the CNN. Each NSCT sub-band is fed to a single CNN then the output of all networks are merged to perform the regression task. By comparing this approach with other approaches that adopt spatial domain images as network's input, the proposed method is superior for the bone age assessment task.

2.2. Shearlet Transform

The Shearlet transform has been recently employed with CNN to perform different tasks in different domains such as RGB images classification [36], remote sensing images denoising [7], panchromatic and multispectral images fusion [37], and classification of video quality [38]. In the medical field, the Shearlet coefficients have been widely combined with CNN's features to perform CT and MR images fusion [39, 40], multimodality MR images fusion [41], cancer classification [42, 43], Alzheimer's disease classification [44], prostate Gleason grading [42], and image enhancement and denoising [29].

Rezaeilouyeh et al. [42] have applied classification of histopathology images using the magnitude and phase of Shearlet coefficients along with raw images as inputs to CNN. They have employed one CNN for each phase sub-band, magnitude sub-band, and raw RGB images. The outputs of all networks are concatenated using one fully connected layer.

They have evaluated this method on performing the classification of breast cancer and Gleason grading tasks. It has been shown that this method performs better than other methods which use different set of handcrafted features. Similarly, Liang et al. [43] have applied the same approach of combining the raw histopathology images with their phase and magnitude coefficients of the Shearlet transform to identify colon cancer. They have also shown the capability of Shearlet coefficients to enrich the CNN's features and improve the classification performance. Jabason et al. [44] have combined the PET images with their Shearlet coefficients as an input to the CNN to classify the different types of Alzheimer's Disease. Yang et al. [36] have proposed a sparse representation CNN layer which employs Shearlet and Wavelet transforms to boost the feature extraction operation. The sparse layer generates non-trainable low pass and high pass feature maps during the back propagation process. The proposed approach has been evaluated by performing classification on five different natural images datasets (FOOD-101, CIFAR10/100, DTD, Brodatz and ImageNet). The results have shown that the multidirectional representation is able to enhance the classification performance where the best classification accuracy of different classes is achieved by either the Shearlet or the Wavelet transforms.

The employment of Contourlet and Shearlet coefficients with CNN's features has been proposed to perform different computer vision tasks. Yet, it has not proposed yet to perform MR images segmentation task and particularly to address the pseudo CT generation for PET attenuation correction. The use of multiresolution analysis features such as Contourlet and Shearlet transforms in the medical domain tasks have shown the efficacy of the addition of multidirectional and multiscale features to the CNN's features. However, the proposed methods in the literature have employed complex solutions including multiple cascaded CNNs for different multiresolution sub-bands. Moreover, they have used all resultant coefficients without any dimensionality reduction. A less complicated design is required to combine the multiresolution features with CNN's features to perform the segmentation task.

3. Materials and Methods

3.1. Data Acquisition

The clinical brain datasets consist of fifty patients which undertook CT and MR scanning after giving informed consent [45]. The demographics and clinical characteristics of the patients are summarised in Table 1. The MR images were acquired using a 3T MAGNETOM Skyra with 64-channel head coil (Siemens Healthcare, Erlangen, Germany). The MR scans are 3D T1-weighted magnetization-prepared rapid acquisition gradient echo (T1-w MP-RAGE) sequence with the following scanning parameters: TE= 2.3 ms, TR= 1900 ms, T1= 970 ms, flip angle 8°, NEX = 1. The matrix dimension is 255 × 255 × 250 and the voxel size is 0.86 × 0.86 × 1 mm.

The CT scans were acquired using two different PET/CT Siemens scanners. The Biograph mCT scanner was used to

Table 1: The datasets demographics and clinical characteristics.

Gender	Age (Mean ± SD)	Clinical diagnosis
28 women and 32 men	61±12 years	44 neurodegenerative disease, 3 epilepsy, and 3 brain tumours

acquire the CT scans of fifteen patients and the scans of the remaining thirty five patients were acquired using the Biograph 64 True Point scanner. The matrix size of the CT images is 512 × 512 × 150 with voxel size of 0.97 × 0.97 × 1.5 mm.

3.2. Data Preprocessing and Labelling

In order to remove the redundant and unnecessary information, the slices of each patient's volume are reduced by identifying the region of interest and selecting only forty eight slices. Additionally, the background pixels are reduced by cropping each slice into 256 × 256 × 48 image. Due to different resolutions, the bilinear interpolation is applied to resample some volumes into 300 × 300 matrix. The resultant images are then cropped into 256 × 256 × 48 to obtain the same dimensions for all patients. Data normalisation using local contrast normalization technique is applied on MR images.

The class labelling is performed using CT images. Firstly, Each MR slice is co-registered with its corresponding CT slice to obtain common coordinates and overcome the temporal gap between the acquisition of MR and CT images (2-3 days). Rigid Euler transformation followed by the non-rigid B-spline transformation are applied using Elastix tool [46]. Afterwards, the labelling of CT images is performed by applying pixel intensity-based thresholding to segment the brain into three tissue classes which are air, bone, and soft tissue. The thresholds of the Hounsfield values for each tissue class are illustrated in Table 2.

3.3. 3D CNN Architecture

The baseline architecture is similar to 2D U-Net architecture [47] that has been applied widely in various medical applications [48]. This network follows the shape of convolutional encoder-decoder architecture with the employment of linking paths from the high-resolution levels to the low-resolution levels. This network is firstly modified by converting 2D layers to 3D layers. Secondly, the number of downsampling layers which cause loss of useful features is decreased. Moreover, the number of convolutional layers is increased by using three convolutional layers at each depth leading to a total of twenty one 3D convolutional layers. Each three convolutional layers are followed by batch normalization, rectified-linear unit (ReLU) activation function, and 3D maxpooling layer with stride size of 2. The kernel size of all convolutional layers is 3 × 3 × 3. The number of filters of the first layer is sixty four and increases by the double as the depth of the network increases. In the decoder part, the maxpooling layers are replaced with deconvolution layers which allow the network to learn the weight of the upsampling process. The last layer of the network is a fully connected layer

that consists of three neurons that represent the number of classes with a multiclass sigmoid activation function. Another amendment to the network is the employment of the dice coefficient loss as an objective function to address the problem of classes unbalancing which leads to a biased classification towards the majority classes. The dice coefficient loss for multi-class segmentation is defined as:

$$L_{dice} = 1 - \sum_{c=1}^{C} \frac{2\sum_{h=1}^{H}\sum_{w=1}^{W}\sum_{d=1}^{D} p(c,h,w,d)\,gt(c,h,w,d)+s}{\sum_{h=1}^{H}\sum_{w=1}^{W}\sum_{d=1}^{D} p(c,h,w,d)^2\,gt(c,h,w,d)^2+s}$$
(1)

where h is the index of the height, w is the index of the width, d is the index of the channels, p is the probability of each voxel (output of the segmentation), gt is the one hot encoded value of the ground truth for class c, C is the number of tissue classes, and s is a smoothing variable to avoid division by zero.

3.4. Non-Subsampled Contourlet Transform Features with CNN-based Features for MR Segmentation

Contourlet is a multiscale and multidirectional transform which decomposes an image using the Laplacian pyramid and the directional filter bank (DFB). Due to the application of down-samplers and up-samplers in Contourlet filters, Contourlet transform is shift variant. The Contourlet addresses the drawback of wavelet which operates only in 2D without any capability to capture the directional information. The Contourlet has the ability to encapsulate the directionality and anisotropy at various scales.

Non-subsampled Contourlet (NSCT) transform is a multiscale, multidirectional, and shift invariant transform which was proposed for the first time by Da Cunha el al. [49] to address essentially the image denoising and enhancement. This transform is a shift invariant version of the Contourlet transform with better frequency selectivity and regularity. This type of transforms include redundancy which enrich the details representation.

The Laplacian pyramid mechanism decomposes the image into one low pass sub-band and one band pass sub-band. The band pass represents the difference between the input image and the low pass image. Afterwards, the resultant low pass image is decomposed iteratively with the same mechanism at each stage. The band pass image is processed by the DFB to extract the directional details which generate the Contourlet coefficients. The DFB generates 2^k directional coefficients or sub-bands at k scale levels where k is a positive integer.

In this work, the MATLAB non-subsampled Contourlet toolbox [50] is utilised to generate the NSCT coefficients. The toolbox takes the 2D image, the number of directional decomposition levels (k), the pyramidal filter, and the direc-

Table 2: The Hounsfield values threshold for generating the ground truth images.

Hounsfield value threshold	Tissue class assignment
HU > 600	Bone
HU < -500	Air
Others	Soft tissue

tional filter as inputs. Then, it generates k+1 coefficients where the first coefficient is the low pass sub-band and the remaining ones represent different directional band pass sub-bands. The number of the directional filter decomposition levels is defined by vector K:

$$K = [1, 1, 2, 2] \qquad (2)$$

where the K levels generate C Contourlet coefficients which are calculated by:

$$C = \sum_{k=0}^{K} 2^k + 1 \qquad (3)$$

The actual generated NSCT coefficients are thirteen sub-bands (twelve directional sub-bands + one low pass sub-band). The NSCT produces coefficients with the same dimensions as the input image. Figure 1 shows an example of some NSCT directional sub-bands at different scales in addition to the low pass sub-band for a single slice of a randomly selected patient from the brain dataset.
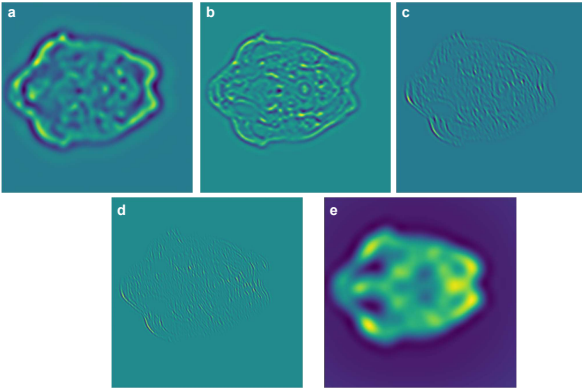

Figure 1: Four directional NSCT sub-bands from four scales (a-d) and one low pass sub-band (e) for a single slice of a randomly selected patient from the training dataset.

3.4.1. Non-Subsampled Contourlet Coefficients Selection

The application of the NSCT decomposition generates twelve redundant coefficients for each slice. The redundant directional information is reduced by selecting the coefficients that include more amount of information than others. The coefficients selection is performed by calculating the entropy score for each NSCT coefficient of each slice. The entropy score is a statistical measure of the information content

and randomness to capture the texture characteristics of the image. The entropy is calculated as:

$$Entropy = -\sum_{i} c_i log_2(c_i) \qquad (4)$$

where c_i is the histogram count of ith bin.

Figure 2 illustrates the entropy values of the top ten selected NSCT directional sub-bands coefficients of one randomly selected slice with a minimum value of 3.028 and a maximum value of 3.772. The figures demonstrate that there
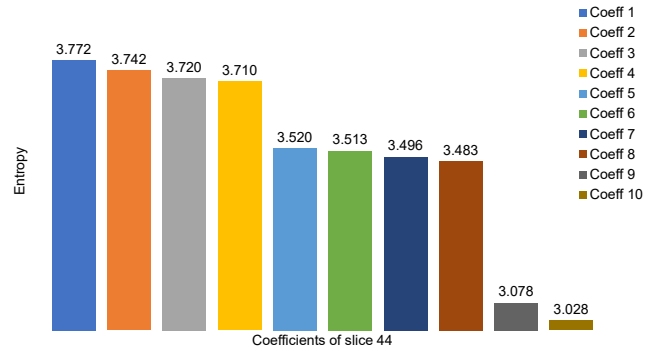

Figure 2: The entropy values of the top ten selected NSCT directional sub-bands coefficients of one randomly selected slice (slice 44) of patient 1.

are variations on the entropy values within and among each slice. The variation indicates there are some NSCT coefficients which are more useful than others. Each NSCT coefficient with higher entropy value implies the inclusion of more information which can add more directional features to help improving the segmentation performance.

Out of fifty six directional sub-bands, the top ten coefficients with the highest entropy values are selected as hand-crafted features which will be combined with CNN's features. It is observed from this figure that the directional sub-bands which are located between indices 9 and 12 represent the highest entropy values which are generated at the third and fourth decomposition levels. The directional sub-bands of the first and second decomposition levels are not revealing dominant features for most of the slices. The more decomposition levels of NSCT coefficients, the more useful directional features are produced.

3.4.2. Fusion of Non-Subsampled Contourlet and CNN-based Features

The selected NSCT coefficients are concatenated to create an input image with size 256 × 256 × 48 ×10. The original im-

age size is divided into nine overlapped patches with size $128 \times 128 \times 48 \times 10$ to firstly overcome the memory size limitation and secondly to preserve the contextual information.

The handcrafted NSCT features are concatenated with the CNN-based features using three different mechanisms. The first mechanism which is illustrated in Figure 3 aims at concatenating the NSCT coefficients with the input MR images. The second mechanism concatenates the NSCT features at the end of the features extraction task of the CNN and just before the classifier layer as shown in Figure 4. The third mechanism adds another dimensionality to the input MR images by adding the low pass sub-band of the NSCT coefficient as a new channel. The input size of this network is $128 \times 128 \times 48 \times 2$. Moreover, at the end of this network, only one directional sub-band with the highest entropy value is concatenated with CNN-based features to add some directionality information. This mechanism is depicted in Figure 5.
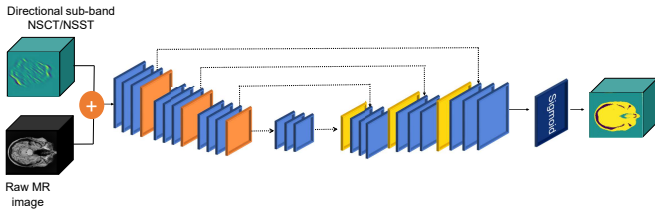

Figure 3: The network design of concatenating the NSCT directional coefficients with the input MR images.
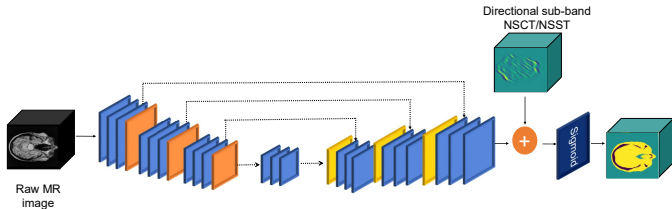

Figure 4: The network design of concatenating the NSCT directional coefficients with the CNN features at the end of the network.
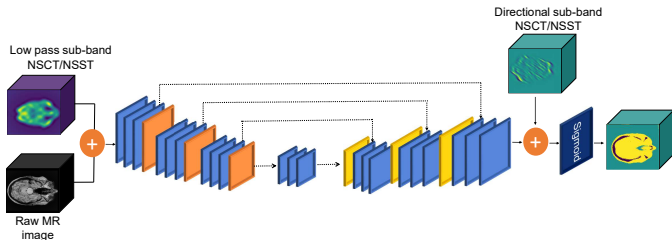

Figure 5: The network design of concatenating the NSCT low pass coefficients with the input MR images. The NSCT directional coefficients are also concatenated with the CNN features at the end of the network.

3.5. Non-Subsampled Shearlet Transform Features with CNN-based Features for MR Segmentation

Shearlet transform is a multivariate expansion of the wavelet transform in multidimensional space. Despite the curvelet which utilises the rotation, Shearlet utilises the shearing mechanism to provide directional selectivity and localisation.

The non-subsampled Shearlet (NSST) is the shift invariant version of the Shearlet transform with high capability of directional sensitivity. It consists of multiscale decomposition using non-subsampled Laplacian pyramid transform and multidirectional decomposition using Shear filter with translation invariance. The Shear filter overcomes the limitation of the number of directions which exists in non-subsampled Contourlet [51, 52].

In general, the NSST transform outperforms NSCT transform thanks to their well-localisation at different scales and directions which enables the fine detection of edges and boundaries. These advantages make the NSST a suitable tool to identify the different brain tissue classes [42]. In this work, the NSST coefficients are generated using the ShearLab 3D tool [53]. Each 2D slice is decomposed using the NSST up to four scales using different shearing levels defined by the vector L:

$$L = [1, 1, 2, 2] \tag{5}$$

Each level k generates S Shearlets coefficients which are calculated by:

$$S_k = 2 \times (2 \times 2^{L(k)} + 2) \tag{6}$$

The total number of generated NSST coefficients is fifty seven (fifty six directional high frequency sub-bands + one low frequency sub-band). The size of each sub-band is the same as the size of the input image. Figure 6 shows an example of some directional high frequency sub-bands from different scales and one low frequency sub-band for a single slice of one randomly selected patient from the training dataset.
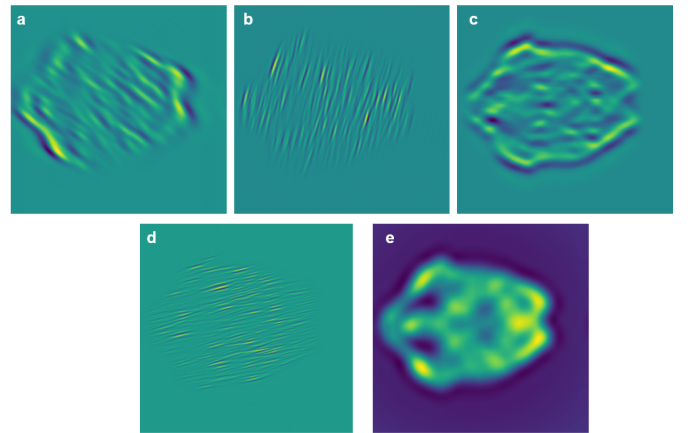

Figure 6: Four NSST directional sub-bands from four scales (a-d) and one low frequency sub-band (e) for a single slice of one patient.

3.5.1. Non-Subsampled Shearlet Coefficients Selection

The dimensions of the resultant sub-bands of each volume is $128 \times 128 \times 48 \times 57$. In order to reduce the dimensionality and the redundancy, the entropy for each NSST coefficient per slice is calculated to select only the coefficients that contain dominant information. The mathematical formula of the entropy is defined in Equation 4 in Subsection B. Figure 7 illustrates the entropy values of the top ten selected NSST directional sub-bands coefficients of one randomly selected slice of one patient with a minimum value of 4.04 and a maximum value of 4.61. Similar to NSCT, the figures demonstrate that there is variation on the entropy values within and among each slice but with higher entropy values.
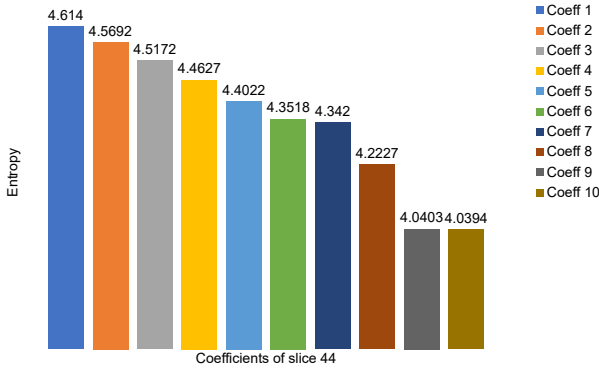

Figure 7: The entropy values of the top ten selected NSST directional sub-bands coefficients of one randomly selected slice (slice 44) of patient 1.

Out of fifty six directional sub-bands, the top ten coefficients with the highest entropy values are selected as handcrafted features which will be combined with CNN's features. It is observed from this figure that the directional sub-bands which are located between indices 28 and 34 represent the highest entropy values which correspond to the third scale. It is also clear from the same figure that the directional sub-bands of the first level are not representing truly dominant features for most of the slices of each patient.

3.5.2. Fusion of Non-Subsampled Shearlet and CNN-based Features

The fusion mechanisms of the NSST features with CNN-based features are similar to the applied mechanisms for NSCT features which are described in 3.4.2. The main objective is to explore the efficiency of adding NSST coefficients which includes more dominant features than NSCT coefficients with less complexity.

3.6. Combining Non-Subsampled Contourlet and Non-Subsampled Shearlet Features with CNN-based Features for MR Segmentation

The NSCT and NSST transforms generate different set of directional features using different decomposition filters and they are performed with different complexity manners. In order to achieve the full potential of the directional features as

well as the low frequency sub-band, the NSCT and NSST coefficients are combined then fused with CNN's features.

The first fusion approach considers the directional features from the NSCT bandpass and the NSST high-pass sub-bands. The NSCT and NSST coefficients are concatenated with the CNN's features at the end of the features extractor part of the network and just before the last activation function (sigmoid) as depicted in Figure 8.

The second approach aims at exploring the benefit of concatenating the low frequency sub-bands of NSCT or NSST coefficients with the MR image to increase the input dimensionality and add more data for training. Additionally, the directional sub-bands of both NSCT and NSST are concatenated at the end of the network with CNN's features as shown in Figure 9.
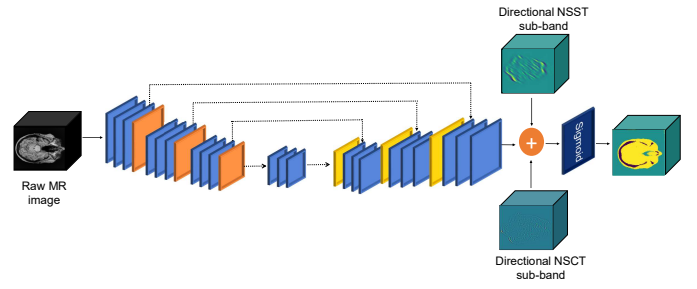

Figure 8: The network design of concatenating the NSCT and NSST directional coefficients with the CNN's features.
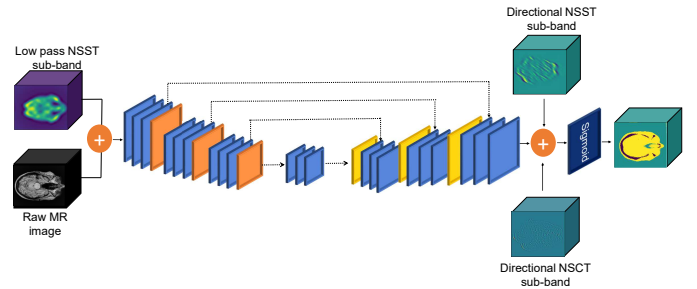
Figure 9: The network design of concatenating the NSCT low pass coefficients with the input MR images. The NSCT and NSST directional coefficients are also concatenated with the CNN's features at the end of the network.

3.7. Model's Input

The input size of the designed CNN is $256 \times 256 \times 48$ where 48 is the number of selected slices per patient. This huge number of pixels overwhelm the GPU memory and creates limitations when designing the network architecture and selecting the hyperparameters such as the number of filters per convolutional layer and the batch size. To overcome this situation, each volume is divided into overlapped patches to

minimize the input size and increase the size of the training datasets. The application of overlapped patches preserves the spatial contextual information for each volume. Each volume is divided into nine overlapped patches with the size of $128 \times 128 \times 48$.

3.8. Model's Training

The deep network is trained from scratch using thirty patients for training, ten patients for validation, and ten patients for testing. The network weights are initialised using Xavier/Glorot Uniform initialisation scheme which relies on the numbers of input and output neurons to determine the scale of initialization automatically. The network is trained using Adam optimizer with an initial learning rate that is set to 0.0001 then reduces by a factor of 0.75 on plateau mode when the training accuracy is not increasing for five continuous epochs. The momentum and the weight decay hyperparameters are set to 0.9 and 0.0005, respectively. The batch size is 2 and the training process stops after 200 epochs or before that when the training accuracy stops decreasing for 10 continuous epochs. All experiments are conducted on Tesla V100 GPU with 16 GB RAM. Keras with the backend of TensorFlow APIs are used to implement the proposed model.

3.9. Model's Evaluation

The segmentation results are evaluated using a set of metrics by comparing the segmented MR image with the ground truth which are generated from CT images. The calculated evaluation metrics are:

- *Precision (PRE)*

$$PRE = \frac{TP}{TP + FP} \tag{7}$$

- *Recall (REC)*

$$REC = \frac{TP}{TP + FN} \tag{8}$$

- *Dice similarity coefficient (DSC)*

$$DSC = \frac{2 \times TP}{(2 \times TP) + FP + FN} \tag{9}$$

- *Jaccard similarity coefficient (JSC)*

$$JSC = \frac{DSC}{2 - DSC} \tag{10}$$

where TP is true positive, FP is false positive, and FN is false negative. These metrics are calculated from the confusion matrix of the testing datasets in pixel-wise and patient-wise as shown in the results section.

The segmentation performance is assessed with 95% confidence intervals (CIs) where the p value is computed as follows:

$$p = z\sqrt{\frac{metric(1 - metric)}{N}} \tag{11}$$

where N is the number of test samples, and z is the significance level (1.96 for 95% CI). All values are computed over the overall confusion matrix that accumulates all test datasets.

4. Results

The segmentation results of the conducted experiments of adding either directional or low pass multi-resolution features for bone, soft tissue, and air classes are illustrated in Tables 3, 4, and 5, respectively.

4.1. Evaluating the Effectiveness of Adding Handcrafted Directional Coefficients

The impact of adding the directional NSCT coefficients with the CNN's features to perform MR segmentation is firstly evaluated by concatenating only one directional NSCT or one directional NSST coefficient per slice which has the highest entropy value. The NSCT/NSST coefficient is concatenated at the end of the network with the CNN's extracted features. Secondly, the top ten selected NSCT/NSST coefficients with the highest entropy values are concatenated at the end of the network in order to increase the directional features. The addition of the ten NSCT/NSST coefficients is also evaluated by concatenating the directional features at the beginning network with the MR input images to pass through the convolutional layers and perform another features extraction process to maintain only the most useful and dominant features.

4.1.1. NSCT

The segmentation results of the three brain classes using different mechanisms of concatenating the NSCT directional features with the CNN-based features show that the aggregation of only one directional NSCT coefficient with the input images has no improvement effect on the segmentation performance. The increase of the dimensionality of the NSCT coefficients which are concatenated with the input image leads to include more directional features and hence improves the segmentation performance. On the other hand, the conservation of the raw NSCT coefficients is able to improve the segmentation by concatenating these features with the resultant CNN's features at the end of the network. The reduction of the NSCT coefficients dimensionality by selecting only the coefficient with the highest entropy value decreases the segmentation performance. Overall, the concatenation of NSCT coefficients at the end of the features extraction process generates the best segmentation results even with the use of less coefficients. The inclusion of more NSCT coefficients and concatenating them with CNN's features without passing them through the convolutional layers is the winning approach to improve the segmentation accuracy of the three brain classes: air, soft tissue, and bone. By comparing the DSC of the different methods, it is noticeable that the accuracy of the bone class is improved with a significant amount while the segmentation of soft tissue and air classes is slightly improved.

The visual segmentation results of the brain classes using the different proposed mechanisms of combining NSCT coefficients with CNN's features are depicted in Figure 10. The observations from the three randomly selected slices reveal that the baseline model which is represented in column c generates many false positive pixels of air class in the nasal

Table 3: The evaluation metrics of the bone tissue segmentation using different mechanisms of concatenating the NSCT directional features with the CNN's features

Bone tissue				
Model	PRE	REC	DSC	JSC
Baseline	0.6890±0.0006	0.5601±0.0006	0.6179±0.0006	0.4471±0.0006
+ 1 directional NSCT (input)	0.6905±0.0006	0.5637±0.0006	0.6207±0.0006	0.4500±0.0006
+ 1 directional NSCT (end)	0.6821±0.0006	0.5717±0.0006	0.6221±0.0006	0.4515±0.0006
+ 10 directional NSCT (input)	0.6910±0.0006	0.5539±0.0006	0.6149±0.0006	0.4439±0.0006
+ 10 directional NSCT (end)	0.6883±0.0006	**0.5962±0.0006**	0.6390±0.0006	0.4695±0.0007
+ 1 directional NSST (input)	0.6882±0.0006	0.5582±0.0006	0.6164±0.0006	0.4456±0.0006
+ 1 directional NSST (end)	0.7027±0.0006	0.5690±0.0006	0.6288±0.0006	0.4586±0.0007
+ 10 directional NSST (input)	**0.721±0.00061**	0.5595±0.0006	0.6301±0.0006	0.4600±0.0007
+ 10 directional NSST (end)	0.6737±0.0006	0.5779±0.0006	0.6221±0.0006	0.4515±0.0006
+ 1 LP NSCT (input) + 1 directional NSCT (end)	0.6931±0.0006	0.5744±0.0006	0.6282±0.0006	0.4580±0.0007
+ 1 LP NSCT (input) + 10 directional NSCT (end)	0.7072±0.0006	0.5446±0.0006	0.6153±0.0006	0.4444±0.0006
+ 1 LP NSST (input) + 1 directional NSST (end)	0.7029±0.0006	0.5730±0.0006	0.6313±0.0006	0.4612±0.0007
+ 1 LP NSST (input) +10 directional NSST (end)	0.7031±0.0006	0.5703±0.0006	0.6298±0.0006	0.4596±0.0007
+ 1 directional NSCT and NSST (end)	0.7218±0.0006	0.5485±0.0006	0.6233±0.0006	0.4528±0.0006
+ 10 directional NSCT and 10 directional NSST (end)	0.7113±0.0006	0.5831±0.0006	**0.6416±0.0006**	**0.4723±0.0007**
+ 10 directional NSST (input) and 10 directional NSCT (end)	0.7128±0.0006	0.5484±0.0006	0.6199±0.0006	0.4492±0.0006
+ 1 NSST LP (input) + 1 directional NSCT and 1 directional NSST (end)	0.6941±0.0006	0.5751±0.0006	0.6290±0.0006	0.4588±0.0007
+ 1 NSST LP (input) + 10 directional NSCT and 10 directional NSST (end)	0.6980±0.0006	0.5645±0.0006	0.6242±0.0006	0.4537±0.0006
Segnet [54]	0.6278±0.0006	0.3649±0.0006	0.4616±0.0007	0.3000±0.0006
Segnet + CRF [55]	0.6278±0.0006	0.3649±0.0006	0.4616±0.0007	0.3000±0.0006

Table 4: The evaluation metrics of the soft tissue segmentation using different mechanisms of concatenating the NSCT directional features with the CNN's features

Soft tissue				
Model	PRE	REC	DSC	JSC
Baseline	0.8913±0.0001	0.9232±0.0001	0.9070±0.0001	0.8298±0.0002
+ 1 directional NSCT (input)	0.8936±0.0001	0.9178±0.0001	0.9056±0.0002	0.8274±0.0001
+ 1 directional NSCT (end)	0.8918±0.0001	0.9194±0.0001	0.9054±0.0001	0.8271±0.0002
+ 10 directional NSCT (input)	0.8906±0.0001	0.9214±0.0001	0.9058±0.0001	0.8277±0.0002
+ 10 directional NSCT (end)	**0.8981±0.0001**	0.9171±0.0001	0.9075±0.0001	0.8306±0.0002
+ 1 directional NSST (input)	0.8908±0.0001	0.9218±0.0001	0.9060±0.0001	0.8282±0.0002
+ 1 directional NSST (end)	0.8947±0.0001	0.9237±0.0001	0.9090±0.0001	0.8332±0.0002
+ 10 directional NSST (input)	0.8934±0.0001	0.9286±0.0001	**0.9107±0.0001**	**0.8360±0.0002**
+ 10 directional NSST (end)	0.8933±0.0001	0.9155±0.0001	0.9043±0.0001	0.8252±0.0002
+ 1 LP NSCT (input) + 1 directional NSCT (end)	0.8950±0.0001	0.9194±0.0001	0.9070±0.0001	0.8299±0.0002
+ 1 LP NSCT (input) + 10 directional NSCT (end)	0.8893±0.0004	0.9274±0.0003	0.9079±0.0004	0.8314±0.0005
+ 1 LP NSST (input) + 1 directional NSST (end)	0.8940±0.0001	0.9235±0.0001	0.9085±0.0001	0.9085±0.0001
+ 1 LP NSST (input) +10 directional NSST (end)	0.8946±0.0004	0.9221±0.0003	0.9081±0.0004	0.8317±0.0005
+ 1 directional NSCT and 1 directional NSST (end)	0.8905±0.0002	**0.9296±0.0001**	0.9096±0.0001	0.8342±0.0002
+ 10 directional NSCT and 10 directional NSST (end)	0.8943±0.0001	0.9273±0.0001	0.9105±0.0001	0.8357±0.0002
+ 10 directional NSST (input) and 10 directional NSCT (end)	0.8917±0.0001	0.9287±0.0001	0.9098±0.0001	0.8345±0.0002
+ 1 NSST LP (input) + 1 directional NSCT and 1 directional NSST (end)	0.8940±0.0001	0.9219±0.0001	0.9078±0.0001	0.8311±0.0002
+ 1 NSST LP (input) + 10 directional NSCT and 10 directional NSST (end)	0.8920±0.0001	0.9241±0.0001	0.9078±0.0001	0.8311±0.0002
Segnet [54]	0.8696±0.0002	0.8810±0.0002	0.8753±0.0002	0.7783±0.0002
Segnet + CRF [55]	0.8696±0.0002	0.8810±0.0002	0.8753±0.0002	0.7783±0.0002

Table 5: The evaluation metrics of the air tissue segmentation using different mechanisms of concatenating the NSCT directional features with the CNN features

Air tissue				
Model	PRE	REC	DSC	JSC
Baseline	0.9632±0.0001	0.9614±0.0001	0.9623±0.0001	0.9274±0.0001
+ 1 directional NSCT (input)	0.9600±0.0001	0.9642±0.0001	0.9621±0.0001	0.9270±0.0001
+ 1 directional NSCT (end)	0.9629±0.0001	0.9614±0.0001	0.9621±0.0001	0.9270±0.0001
+ 10 directional NSCT (input)	0.9615±0.0001	0.9622±0.0001	0.9619±0.0001	0.9265±0.0001
+ 10 directional NSCT (end)	0.9615±0.0001	0.9634±0.0001	0.9625±0.0001	0.9276±0.0001
+ 1 directional NSST (input)	0.9628±0.0001	0.9620±0.0001	0.9624±0.0001	0.9276±0.0001
+ 1 directional NSST (end)	0.9614±0.0001	0.9625±0.0001	0.9620±0.0001	0.9267±0.0001
+ 10 directional NSST (input)	0.9631±0.0001	0.9636±0.0001	0.9633±0.0001	0.9293±0.0001
+ 10 directional NSST (end)	0.9615±0.0001	0.9619±0.0001	0.9617±0.0001	0.9262±0.0001
+ 1 LP NSCT (input) + 1 directional NSCT (end)	0.9615±0.0001	**0.9639±0.0001**	0.9627±0.0001	0.9281±0.0001
+ 1 LP NSCT (input) + 10 directional NSCT (end)	0.9632±0.0002	0.9619±0.0002	0.9625±0.0002	0.9278±0.0003
+ 1 LP NSST (input) + 1 directional NSST (end)	0.9633±0.0001	0.9633±0.0001	0.9633±0.0001	**0.9297±0.0001**
+ 1 LP NSST (input) +10 directional NSST (end)	0.9613±0.0003	0.9635±0.0002	0.9624±0.0002	0.9276±0.0003
+ 1 directional NSCT and 1 directional NSST (end)	0.9632±0.0001	0.9624±0.0001	0.9628±0.0001	0.9283±0.0001
+ 10 directional NSCT and 10 directional NSST (end)	**0.9647±0.0001**	0.9616±0.0001	0.9631±0.0001	0.9289±0.0001
+ 10 directional NSST (input) and 10 directional NSCT (end)	0.9637±0.0001	0.9635±0.0001	**0.9636±0.0001**	0.9289±0.0001
+ 1 NSST LP (input) + 1 directional NSCT and 1 directional NSST (end)	0.9629±0.0001	0.9625±0.0001	0.9627±0.0001	0.9281±0.0001
+ 1 NSST LP (input) + 10 directional NSCT and 10 directional NSST (end)	0.9636±0.0001	0.9622±0.0001	0.9629±0.0001	0.9285±0.0001
Segnet [54]	0.9083±0.0002	0.9510±0.0001	0.9291±0.0001	0.8677±0.0002
Segnet + CRF [55]	0.9083±0.0002	0.9510±0.0001	0.9291±0.0001	0.8677±0.0002

cavities area as well as misclassified air and soft tissue pixels. Moreover, the quality of the segmentation of the bone in some challenging locations (e.g., nasal cavities area) is not good. However, the segmentation of the bone in other parts of the head is promising. The slices in column d show that the addition of only one NSCT coefficient with the input MR image removes some discontinuities in the bone class compared to the baseline model. Additionally, the segmentation of the air cavities pixels is enhanced. The increase of the dimensionality of the input by adding ten NSCT coefficients is not improving the segmentation as the false positive pixels of bone in nasal cavities area are increased with the appearance of more bone discontinuities. The results of concatenating the NSCT coefficients at the end of the network with CNN's features by adding one or ten NSCT coefficients show the expansion of the false positive pixels of air class in most of the selected slices. However, the slice in the first row (column e) shows that the segmentation of some challenging areas in the head is improved. Overall, the visual comparison indicates that the increase of the NSCT coefficients dimensionality which in turns increases the input size is necessary to improve the segmentation results. The capability of the model is improved when more directional features are added.

4.1.2. NSST

The quantitative results of the segmentation of the brain using different mechanisms of concatenating the NSST di-
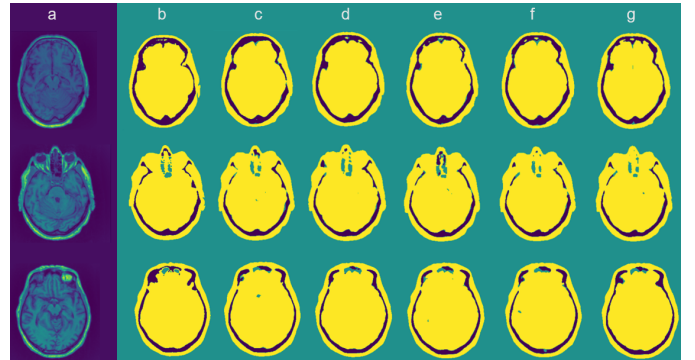

Figure 10: The segmentation results of some randomly selected slices from the testing datasets. (a) MR images, (b) the CT images which used as ground truth, (c) the segmentation results with the baseline model, (d) the segmentation results with the addition of one directional NSCT coefficient with the input, (e) the segmentation results with the addition of ten directional NSCT coefficients with the input, (f) the segmentation results with the addition of one directional NSCT coefficient with CNN's features at the end of the network, (g) the segmentation results with the addition of ten directional NSCT coefficients with the CNN's features at the end of the network.

rectional features with the convolutional network based features reveal that the aggregation of the directional NSST coefficients is able to improve the segmentation results when it is concatenated at the end of the network. The concatenation of these handcrafted features with the generated CNN-based features at the end of the network shows good improvements

10

in terms of most of the evaluation metrics and specially for the bone and soft tissue classes. For the air class, only the recall metric is slightly improved. The increase of the dimensionality of the NSST coefficients which is experimented by selecting the top ten directional coefficients and concatenating them at the end of the network shows no improvements. Thus, only one directional NSST coefficient is able to add sufficient directional information and enriches the features as well as the segmentation performance. However, the concatenation of one directional NSST features with the input MR images does not have any impact due to injection of few examples of a new type of data to the network. The increase of the dimensionality of the network's input by concatenating ten NSST directional coefficients increases the capability of the network to learn enough features from the directional NSST data then outperforms other proposed mechanisms of adding the NSST coefficients to the CNN.

The visual segmentation results of some randomly selected slices of the testing datasets are illustrated in Figure 11. The slices shown in column d are the outputs of the model which concatenates only one NSST directional coefficient with the input MR image. Compared to the visual output of the baseline model, this model is not an efficient approach to enhance the brain segmentation as it increases the false negative of the bone which decreases the model sensitivity. On the other hand, the addition of more NSST coefficients to increase the input's dimensionality enhances the segmentation of the challenging areas of the bone class. The outputs of the models which concatenate one NSST directional coefficient and ten NSST directional coefficients with the CNN's features at the end of the network are shown in column f and column g, respectively. The results show some misclassified soft tissue pixels as air and other misclassified bone pixels as soft tissue. Moreover, the discontinuities in the bone class are increased. By comparing the slices in column f and g, it is clear that the segmentation of the model which uses ten NSST directional coefficients is better than the model that takes only one NSST coefficient especially in the nasal cavity area. The concatenation of the NSST coefficients with the CNN's features as raw features without passing them through the convolutional layers shows its ability to generate more accurate results by using only one NSST coefficients rather then ten NSST coefficients. One directional coefficient per slice is able to add the useful directional features to improve the segmentation results.

4.2. Evaluating the Effectiveness of Adding Handcrafted Low Pass Coefficients

The decomposition of NSCT and NSST generate only one low pass (LP) sub-band which represents the approximate information which can be used as an additional data input to the MR images. Both NSCT and NSCT apply the Laplacian pyramid filter which resulted into one identical low pass sub-band. Hence, the same low pass sub-band is concatenated to the input MR image while the NSCT/NSST directional coefficients are concatenated at the end of the network to study the effect of combining low pass sub-band and NSCT/NSST
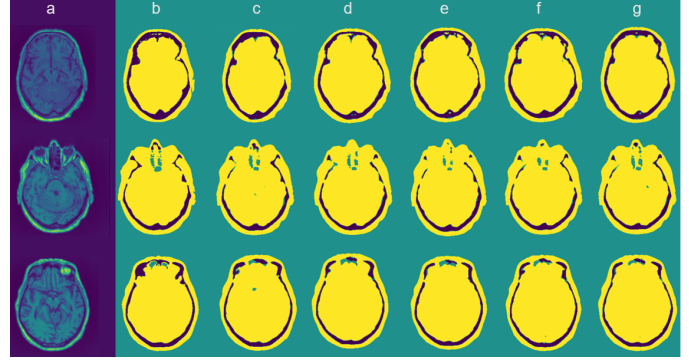

Figure 11: The segmentation results of some randomly selected slices from the testing datasets. (a) the MR images, (b) the CT images which used as ground truth, (c) the segmentation results with the baseline model, (d) the segmentation results with the addition of one directional NSST coefficient with the input, (e) the segmentation results with the addition of ten directional NSST coefficients with the input, (f) the segmentation results with the addition of one directional NSST coefficient with CNN's features at the end of the network, (g) the segmentation results with the addition of ten directional NSST coefficients with the CNN's features at the end of the network.

high pass sub-bands. The inclusion of low pass sub-band increases the size of the training datasets and tends to add more visual representations. Moreover, the directional features of one or more multi-resolution features are concatenated with CNN's features at the end of the network.

4.2.1. NSCT

The results of the objective evaluation of the brain segmentation using both low pass and band pass NSCT sub-bands concatenated with the CNN's features indicate that the addition of the low pass NSCT sub-band with only one directional NSCT coefficient at the end of the network is able to attain better segmentation performance for the bone class.

The visual segmentation results of the brain using the proposed mechanisms are depicted in Figure 12. The figures show that the addition of the low pass sub-band of NSCT is improving slightly the segmentation accuracy of the bone class in the nasal cavities area. However, the model that concatenates the low pass sub-band with the input and ten directional NSCT coefficients with the CNN's features is not improving the segmentation and is even performing worst than the baseline model.

4.2.2. NSST

According to the segmentation results of the the three classes by concatenating the low pass sub-band of the NSST with the input image and the high pass sub-bands of NSST coefficients with CNN's features, the addition of the low pass sub-band is enhancing the segmentation accuracy and specifically with the inclusion of only one directional NSST sub-band. The second model that takes ten NSST coefficient while the low pass is concatenated with the input shows better performance compared to the model that only concatenates ten NSST coefficients without low pass sub-band inclusion. The inclusion of low pass coefficient of NSST is able
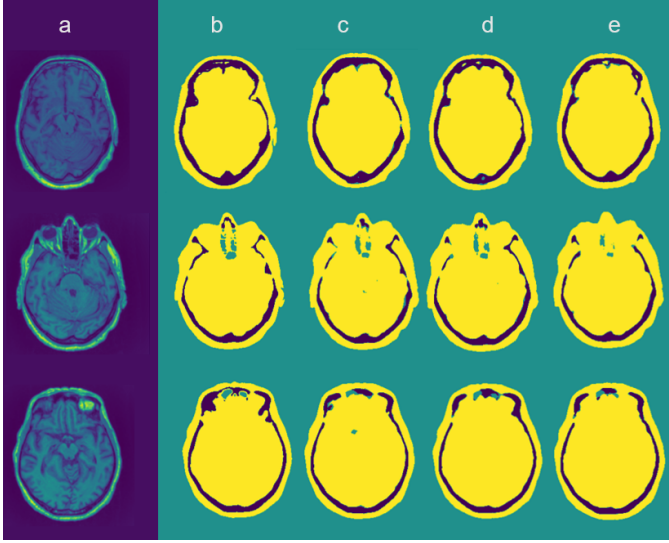
Figure 12: The segmentation results of some randomly selected slices from the testing datasets. (a) the input MR images, (b) the CT images which used as ground truth, (c) the segmentation results with the baseline model, (d) the segmentation results with the addition of the NSCT low pass sub-band (input) and one directional NSCT coefficient (end), (e) the segmentation results with the addition the NSCT low pass sub-band (input) and ten directional NSCT coefficients (end).
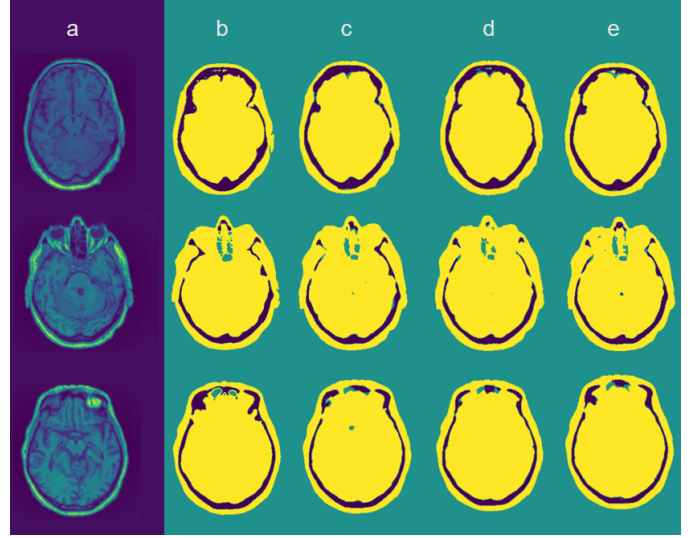

Figure 13: The segmentation results of some randomly selected slices from the testing datasets. (a) the MR images, (b) the CT images which used as ground truth, (c) the segmentation results with the baseline model, (d) the segmentation results with the addition of the NSST low pass sub-band (input) and one directional NSST coefficient (end), (e) the segmentation results with the addition the NSST low pass sub-band (input) and ten directional NSST coefficients (end).

to reduce the false positive of bone class and the false negative of air class. Collectively, the model that takes less data and produces better results is the model of choice. Hence, the inclusion of low pass sub-band with one NSST directional coefficient is the winning approach. Moreover, the results prove that the model that uses the low pass sub-band concatenated with the input MR image is able to extract useful features from two different inputs with different nature and resolutions. Therefore, this model is robust and reliable to be used for other transfer learning applications.

The visual segmentation results of the brain using the proposed mechanisms are depicted in Figure 13. The slices show that the inclusion of low pass sub-band of NSST coefficient is improving the bone segmentation significantly compared to the baseline model segmentation. It is also able to segment the bone in the challenging areas of the head such as the nasal cavity and reduces the discontinuities. However, there are some misclassified soft tissue pixels as air. The concatenation of the low pass sub-band with the input and adding one NSST directional coefficient with the CNN's features at the end of the network is producing more accurate segmentation than using ten NSST directional coefficients since the bone discontinuities are reduced as well as the false positive of the air and bone classes. However, the model with ten directional NSST coefficients is able to segment the challenging bone area in the head.

4.3. Evaluating the Effectiveness of Adding Combined Handcrafted NSCT and NSST Coefficients

Aiming at taking the full potential of directional handcrafted features and adding more useful features to the segmentation model, the coefficients of both NSCT and NSST

are combined with different designs. Firstly, only one directional coefficient with the highest entropy value of each transform is selected and concatenated with the CNN's features. Secondly, the dimensionality of the directional features is increased by selecting the top ten directional coefficients of both transforms. The third experiment is inspired by the achieved results of concatenating ten directional NSST coefficients with the input or concatenating ten directional NSCT coefficients at the end of the network with CNN's features. Hence, both mechanisms are combined by concatenating ten directional NSST coefficients with the input and ten directional NSCT coefficients at the end of the network. Additionally, the addition of the low pass sub-band is evaluated where the low pass sub-band of NSST is concatenated with the input MR image and the directional coefficients of both NSCT and NSST are combined with CNN's features at the end of the network.

The results show that the combination of only one directional coefficient of each transform is not improving the segmentation accuracy significantly. It is slightly improving the soft tissue and air and reduces the false positive of the bone class. Increasing the dimensionality of the directional coefficients of both transforms is improving the segmentation results by comparing the DSC of the three classes. This combination is outperforming all other combination mechanisms and shows that adding more directional features using different multiresolution transforms is adding value to the segmentation model. Combining NSCT and NSST while processing each multi-resolution type differently (i.e., NSST with the input and NSCT at the end) is not a promising approach. NSCT coefficients are used as raw features while NSST coefficients are combined with the input and passed through

12

the network layers for further processing. The results of each tissue class reveal that the segmentation performance is decreased compared to other combination mechanisms. The proposed segmentation model is not capable to handle this variety of features and a more sophisticated model should be proposed to address this combination. The results of adding the low pass sub-band with the input image while concatenating one or many directional coefficients with the CNN's features are not revealing any significant segmentation improvement except for some slight improvements in the air class.

Table 6 and Figure 14 show the evaluation metrics per patient of the testing datasets while using the model which combines ten directional coefficients of NSCT and ten directional coefficients of NSST sub-bands with CNN's features at the end of the network. The box plot of the bone class indicates a high variability on the segmentation performance with high standard deviation value. One reason of this high variability is the inclusion of patients with different neurological diagnosis where each neurological condition affects the bone health differently [56]. As illustrated in Table 6, the DSC value of the bone class of patient 7 is very low compared to patient 9. The proposed model has the capability to segment the bone class accurately for some patients while it fails with others. The variability of the results of bone segmentation does not provide any guarantee that this method is robust for patients with different anatomical structures.
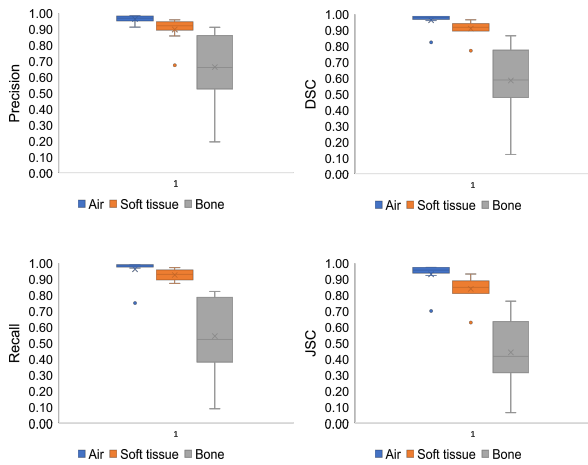

Figure 14: The box plots of the evaluation metrics of the three tissue classes of the testing datasets using the method which uses the combined ten directional NSCT and NSST coefficients concatenated with CNN features at the end of the network.

The segmentation results of the proposed mechanisms are depicted in Figure 15 where the slices that are located at columns d and e show that the combination of two different directional handcrafted features is improving the segmentation performance compared to the baseline model in column c. The inclusion of one directional coefficient of each transform helps to identify the soft tissue class in very challenging areas in the head. It also decreases the false positive of air class which increases the precision of air class inside the head contour. The increase of the number of directional features and concatenating them with CNN's features reduces the false positive of air and bone classes although some misclassified bone pixels as soft tissue pixels are still exist. Column f depicts that the addition of low pass sub-band with the directional coefficients is improving the bone class segmentation while decreasing the sensitivity and the precision of soft tissue class. The results of the dimensionality increase of the directional coefficients which are shown in column g indicate the appearance of more discontinuities in bone class which are not appearing in column f. Moreover, the false negative pixels of soft tissue class are increased. Overall, the performance of adding the low pass sub-band with one or ten directional coefficients is not adding a substantial improvement in the segmentation task.
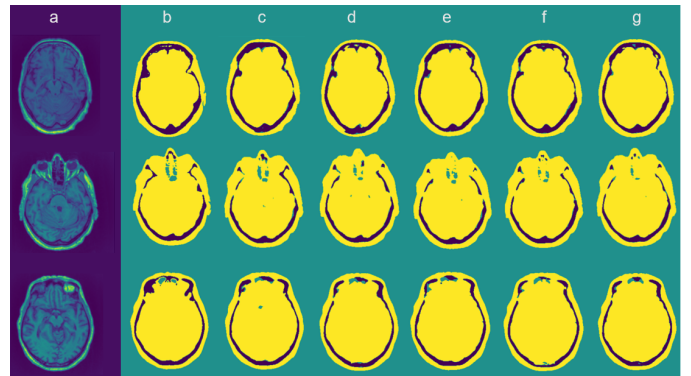

Figure 15: The segmentation results of some randomly selected slices from the testing datasets. (a) the MR images, (b) the CT images which used as ground truth, (c) the segmentation results with the baseline model, (d) the segmentation results by combining one directional NSCT and NSST coefficients, (e) the segmentation results by combining ten directional NSCT and NSST coefficients, (f) the segmentation results by adding the NSST low pass sub-band with the input image and combining one directional NSCT and NSST coefficients with CNN's features at the end of the network, (g) the segmentation results by adding the NSST low pass sub-band with the input image and combining ten directional NSCT and NSST coefficients with CNN's features at the end of the network.

4.4. Validation

The proposed methods are compared with other segmentation studies that applied CNN to segment T1-w MR images for PET attenuation correction [54, 55] . These methods are re-implemented then trained and tested using the existing datasets to segment the brain into three tissue classes: air, soft tissue, and bone. The first method which is proposed by [54] is a deep CNN that follows the SegNet architecture to segment the 3D brain MR image into three classes: air, bone, and soft tissue. The network is trained with T1-w MR images and evaluated using CT images. The same network architecture is re-implemented with some slight changes to cope with the size of the existing datasets using the same training parameters and hyperparameters. The second approach which is proposed by [55] has applied transfer learning technique by using a pre-trained model with T1-w MR images. The pretrained model is retrained using MR UTE images as input to the network. 3D fully connected conditional random

Table 6: The evaluation metrics of the segmentation of the three tissue classes of each patient in the testing dataset using the combination of ten directional NSCT and ten directional NSST coefficients with CNN's features at the end of the network.

Patients	Bone				Soft tissue				Air			
	PRE	REC	DSC	JAC	PRE	REC	DSC	JAC	PRE	REC	DSC	JAC
Patient 1	0.8129±0.0005	0.5180±0.0007	0.6328±0.0006	0.4628±0.0007	0.9197±0.0004	0.9586±0.0003	0.9387±0.0003	0.8845±0.0004	0.9716±0.0002	0.9896±0.0001	0.9805±0.0002	0.9617±0.0003
Patient 2	0.8478±0.0005	0.6804±0.0006	0.7549±0.0006	0.6063±0.0006	0.9269±0.0001	0.9424±0.0001	0.9346±0.0001	0.8773±0.0002	0.9583±0.0001	0.9862±0.0001	0.9720±0.0001	0.9456±0.0001
Patient 3	0.6361±0.0006	0.3906±0.0006	0.4840±0.0007	0.3192±0.0006	0.9079±0.0001	0.9255±0.0001	0.9166±0.0001	0.8460±0.0002	0.9507±0.0001	0.9707±0.0001	0.9606±0.0001	0.9242±0.0001
Patient 4	0.6816±0.0006	0.3604±0.0006	0.4715±0.0006	0.3084±0.0006	0.8585±0.0002	0.9373±0.0001	0.8962±0.0001	0.8119±0.0002	0.9731±0.0001	0.9805±0.0001	0.9768±0.0001	0.9547±0.0001
Patient 5	0.8975±0.0004	0.7856±0.0005	0.8378±0.0005	0.7209±0.0006	0.9416±0.0001	0.9574±0.0001	0.9494±0.0001	0.9037±0.0001	0.9814±0.0001	0.9914±0.0001	0.9864±0.0001	0.9731±0.0001
Patient 6	0.4683±0.0007	0.5309±0.0007	0.4976±0.0007	0.3312±0.0006	0.9200±0.0001	0.8743±0.0002	0.8965±0.0001	0.8125±0.0002	0.9683±0.0001	0.9909±0.0001	0.9794±0.0001	0.9597±0.0001
Patient 7	0.1960±0.0005	0.0909±0.0004	0.1242±0.0004	0.0662±0.0003	0.9232±0.0001	0.9185±0.0001	0.9208±0.0002	0.8532±0.0002	0.9543±0.0001	0.9896±0.0001	0.9716±0.0001	0.9448±0.0001
Patient 8	0.5477±0.0006	0.7918±0.0005	0.6475±0.0006	0.4788±0.0007	0.9588±0.0001	0.8796±0.0002	0.9175±0.0001	0.8476±0.0002	0.9815±0.0001	0.9922±0.0001	0.9868±0.0000	0.9740±0.0001
Patient 9	0.9123±0.0004	0.8238±0.0005	0.8658±0.0004	0.7633±0.0006	0.9581±0.0001	0.9727±0.0001	0.9653±0.0001	0.9329±0.0001	0.9848±0.0001	0.9889±0.0001	0.9869±0.0001	0.9741±0.0001
Patient 10	0.6400±0.0006	0.4761±0.0007	0.5460±0.0006	0.3755±0.0006	0.6749±0.0002	0.9037±0.0001	0.7727±0.0002	0.6296±0.0002	0.9136±0.0002	0.7516±0.0002	0.8247±0.0002	0.7017±0.0003

field (CRF) is applied to the segmentation results as a post-processing and refinement technique.

The results of the evaluation metrics of three brain classes segmentation of the three proposed methods compared with other studies in the literature show that the proposed methods which introduce the concatenation of different features from NSCT and NSST coefficients with CNN's features with different aggregation mechanisms and designs are enhancing the segmentation performance of the three brain classes. The baseline model which represents the CNN without the addition of any handcrafted features is outperforming the other proposed methods in the literature. Liu et al. [54] applied the Segnet architecture to perform the MR images segmentation for PET attenuation correction and Jang et al. [55] applied the same pretrained model as [54] using UTE MR images by applying the transfer learning technique in addition to conditional random field as post processing technique. The model which combines both directional features of NSCT and NSST shows its superiority in performing the segmentation of the bone and soft tissue classes.

The segmentation results of some randomly selected slices are illustrated in Figure 16. The visual observations indicate that the models that include the handcrafted NSCT and NSST coefficients for segmenting the brain are outperforming the other proposed methods in the literature. The approach of [54] which is shown in column g is not able to identify the bone and air classes in challenging areas. Moreover, the precision and the sensitivity of the bone and air classes are considerably lower than the proposed baseline model. The approach of [55] is also not showing any improvement in the segmentation results since the conditional random field which used for post processing is not enhancing the segmentation performance.

The main differences of the three proposed methods (shown in columns d - f) are as follows. The addition of the NSCT coefficients with CNN's features is able to distinguish between the soft tissue and the bone classes although they appear more cases of false positive of air class in different locations. Besides that, the concatenation of NSST coefficients increase the sensitivity of soft tissue class and improves the performance of the bone segmentation. Moreover, the combination of both transforms coefficients with the CNN's features increases the competency of the model to segment the air class. However, there are more discontinues that appear in the bone class and more misclassified bone pixels as soft tissue.
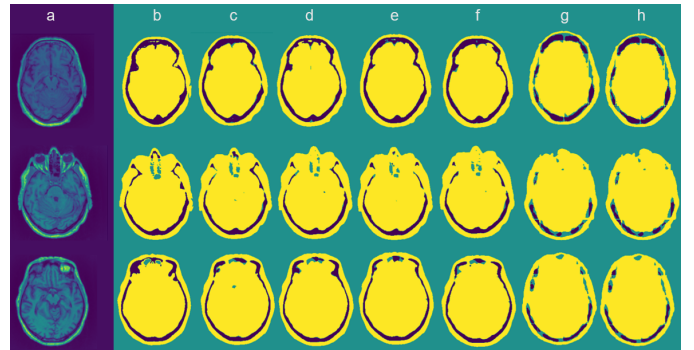

Figure 16: The segmentation results of some randomly selected slices from the testing datasets. (a) the input MR images, (b) the CT images which used as ground truth, (c) the segmentation results of the baseline model, (d) the segmentation results with ten directional NSCT coefficients concatenated with CNN's features (end), (e) the segmentation results with the addition of the NSST low pass sub-band (input) and one directional NSST coefficients (end), (f) the segmentation results by combining ten directional NSCT and NSST coefficients (end), (g) the segmentation results of the model proposed by [54], (h) the segmentation results of the model proposed by [55].

4.5. Pseudo CT Generation

The segmentation maps are converted into pseudo CT images by assigning a Hounsfield unit value to each tissue class. The Hounsfield unit values are obtained from another study [54] where bone, soft tissue, and air classes are represented with 939 HU, 42 HU, and - 1000 HU values, respectively. The mean absolute error (MAE) between the generated pseudo CT and CT images with continuous HU values are calculated as shown in Table 7.

5. Discussion and Conclusion

This work has proposed the aggregation of multiresolution handcrafted features such as NSCT and NSST transforms with CNN-based features to enhance the brain T1-w MR images segmentation. The handcrafted features consist of low pass sub-band which is concatenated with the input MR images and directional coefficients of different levels which are concatenated with CNN's features at the end of the network.

A quantitative evaluation using the brain MR images of ten patients has been carried out to study the impact of

Table 7: MAE between the generated pseudo CT images and the CT images with continuous values. The proposed method refers to the combination of 10 directional NSCT and 10 directional NSST coefficients with CNN based features.

Patient	MAE (proposed)	MAE ([54])
patient 1	133.52	155.60
patient 2	131.83	194.65
patient 3	167.43	196.30
patient 4	133.25	178.53
patient 5	124.81	163.56
patient 6	104.66	146.67
patient 7	136.86	159.36
patient 8	109.99	138.39
patient 9	113.59	155.66
patient 10	283.86	304.28
Mean	143.98	79.30
SD	49.53	45.44

adding the multiresolution handcrafted features with CNN's features. The addition of these features separately or combined has shown their efficiency in enhancing the segmentation results. The approach of adding NSCT directional features with CNN's features has shown the potential to improve the segmentation results when using ten directional coefficients. The inclusion of the low pass sub-band with the input MR image has not shown any improvement in the segmentation results. One reason of this is the need to design a deeper model that is able to handle the low pass coefficient of the NSCT transform and extract the dominant features. The addition of only raw directional NSCT features without the need to increase the dimensionality of the input MR images with the low pass sub-band is the model of choice as it requires less complicated modelling and processing while improving the segmentation accuracy.

On the other hand, the model that concatenates the low pass sub-band with the input image and one directional NSST coefficients has shown a good segmentation performance. The Shear information of one NSST directional coefficient are enough to add the directionality features to the model. Moreover, the combined NSCT and NSST coefficients has shown even better performance where different directional information are combined with CNN's features.

Furthermore, a subjective evaluation using the testing datasets of ten patients has been conducted by selecting three random slices from different patients to draw the visual observations. The segmentation results have demonstrated that the low pass sub-band which is generated by the Laplacian pyramid filter does not add a significant improvement to the brain segmentation. Overall, the directional coefficients are the most useful features which have enhanced the segmentation performance when concatenated as raw coefficients with CNN's features at the end of the network. The experiments have also shown that the proposed models are able to take inputs with different resolutions such as the raw MR image and the low pass sub-band of the multiresolution transform. The model in which one directional NSST coefficients are concatenated with CNN's features is the optimal approach to improve the segmentation performance

and requires less number of handcrafted features either in the model input or the additional added ones at the end of the network. Hence, this model has less complexity and requires less computation time and resources. Additionally, the NSST directional coefficients consist of more amount of information compared to the NSCT directional coefficients as it was indicated by the entropy values of different slices of each volume.

One limitation of these proposed methods is the lack of experimenting the ideal decomposition levels of the multiresolution transforms. The more decomposition levels, the more useful directional features are produced. However, there is a trade off between the model complexity and the generation of useful directional coefficients. In this work, the NSST and NSCT coefficients are produced by decomposing the system into four levels where number four is randomly selected. The decomposition levels value should be carefully chosen empirically to select the optimal number which generates useful features with a reasonable amount of complexity and computation time.

Another limitation of this work and due to the high variability of the results of bone class segmentation, there is a need to use one segmentation model for each neurological condition where a model should be trained and tested with datasets that include patients with the same clinical diagnosis.

Acknowledgement

References

[1] I. Mecheter, L. Alic, M. Abbod, A. Amira, J. Ji, Mr image-based attenuation correction of brain pet imaging: review of literature on machine learning approaches for segmentation, Journal of Digital Imaging (2020) 1–18 (2020).

[2] C. Catana, Attenuation correction for human pet/mri studies, Physics in Medicine & Biology 65 (23) (2020) 23TR02 (2020).

[3] H. Arabi, H. Zaidi, Deep learning-guided estimation of attenuation correction factors from time-of-flight pet emission data, Medical image analysis 64 (2020) 101718 (2020).

[4] H. Arabi, K. Bortolin, N. Ginovart, V. Garibotto, H. Zaidi, Deep learning-guided joint attenuation and scatter correction in multitracer neuroimaging studies, Human brain mapping 41 (13) (2020) 3667–3679 (2020).

[5] A. Mehranian, H. Arabi, H. Zaidi, Vision 20/20: Magnetic resonance imaging-guided attenuation correction in PET/MRI: Challenges, solutions, and opportunities, Medical Physics 43 (3) (2016) 1130–1155 (Mar. 2016). doi:10.1118/1.4941014.
URL https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.4941014

[6] M. Kaur, D. Singh, Multi-modality medical image fusion technique using multi-objective differential evolution based deep neural networks, Journal of Ambient Intelligence and Humanized Computing 12 (2) (2021) 2483–2493 (Feb. 2021). doi:10.1007/s12652-020-02386-0.
URL https://doi.org/10.1007/s12652-020-02386-0

[7] S. Liu, L. Gao, Y. Lei, M. Wang, Q. Hu, X. Ma, Y.-D. Zhang, SAR Speckle Removal Using Hybrid Frequency Modulations, IEEE Transactions on Geoscience and Remote Sensing 59 (5) (2021) 3956–3966 (May 2021). doi:10.1109/TGRS.2020.3014130.

[8] A. Khan, A. Chefranov, H. Demirel, Image scene geometry recognition using low-level features fusion at multi-layer deep CNN, Neurocomputing 440 (2021) 111–126 (Jun. 2021). doi:10.1016/j.neucom.2021.01.085.
URL https://www.sciencedirect.com/science/article/pii/S0925231221001612

[9] X. Gao, Z. Tang, Y. Xie, H. Zhang, W. Gui, A layered working condition perception integrating handcrafted with deep features for froth flotation, Minerals Engineering 170 (2021) 107059 (Aug. 2021). doi:10.1016/j.mineng.2021.107059.
URL https://www.sciencedirect.com/science/article/pii/S0892687521002880

[10] J. M. Fajardo, O. Gomez, F. Prieto, EMG hand gesture classification using handcrafted and deep features, Biomedical Signal Processing and Control 63 (2021) 102210 (Jan. 2021). doi:10.1016/j.bspc.2020.102210.
URL https://www.sciencedirect.com/science/article/pii/S1746809420303426

[11] S. Devulapalli, A. Potti, R. Krishnan, M. S. Khan, Experimental evaluation of unsupervised image retrieval application using hybrid feature extraction by integrating deep learning and handcrafted techniques, Materials Today: Proceedings (Jun. 2021). doi:10.1016/j.matpr.2021.04.326.
URL https://www.sciencedirect.com/science/article/pii/S2214785321032466

[12] F. B. Tesema, H. Wu, M. Chen, J. Lin, W. Zhu, K. Huang, Hybrid channel based pedestrian detection, Neurocomputing 389 (2020) 1–8 (2020).

[13] S. Shang, J. Sun, Z. Yue, Y. Wang, X. Wang, Y. Luo, D. Zhao, T. Yu, X. Jiang, Multi-parametric MRI based radiomics with tumor subregion partitioning for differentiating benign and malignant soft-tissue tumors, Biomedical Signal Processing and Control 67 (2021) 102522 (May 2021). doi:10.1016/j.bspc.2021.102522.
URL https://www.sciencedirect.com/science/article/pii/S1746809421001191

[14] M. Choudhary, V. Tiwari, V. U., Iris anti-spoofing through score-level fusion of handcrafted and data-driven features, Applied Soft Computing 91 (2020) 106206 (Jun. 2020). doi:10.1016/j.asoc.2020.106206.
URL https://www.sciencedirect.com/science/article/pii/S1568494620301460

[15] G. F. Roberto, A. Lumini, L. A. Neves, M. Z. do Nascimento, Fractal Neural Network: A new ensemble of fractal geometry and convolutional neural networks for the classification of histology images, Expert Systems with Applications 166 (2021) 114103 (Mar. 2021). doi:10.1016/j.eswa.2020.114103.
URL https://www.sciencedirect.com/science/article/pii/S0957417420308563

[16] Z. Golrizkhatami, A. Acan, ECG classification using three-level fusion of different feature descriptors, Expert Systems with Applications 114 (2018) 54–64 (Dec. 2018). doi:10.1016/j.eswa.2018.07.030.
URL https://www.sciencedirect.com/science/article/pii/S0957417418304469

[17] D. Liu, J. Chen, Z. Huang, N. Zeng, P. Lu, L. Yang, H. Wang, J. Kou, M. Wu, A deep attention-based ensemble network for real-time face hallucination, Journal of Real-Time Image Processing 17 (6) (2020) 1927–1937 (Dec. 2020). doi:10.1007/s11554-020-01009-3.
URL https://doi.org/10.1007/s11554-020-01009-3

[18] Y. Yang, R. Xie, W. Jia, Z. Chen, Y. Yang, L. Xie, B. Jiang, Accurate and automatic tooth image segmentation model with deep convolutional neural networks and level set method, Neurocomputing 419 (2021) 108–125 (Jan. 2021). doi:10.1016/j.neucom.2020.07.110.
URL http://www.sciencedirect.com/science/article/pii/S0925231220313084

[19] H. Chen, Z. Qin, Y. Ding, L. Tian, Z. Qin, Brain tumor segmentation with deep convolutional symmetric neural network, Neurocomputing 392 (2020) 305–313 (Jun. 2020). doi:10.1016/j.neucom.2019.01.111.
URL http://www.sciencedirect.com/science/article/pii/S0925231219304692

[20] J. Hu, Y. Song, L. Zhang, S. Bai, Z. Yi, Multi-scale attention U-net for segmenting clinical target volume in graves' ophthalmopathy, Neurocomputing 427 (2021) 74–83 (Feb. 2021). doi:10.1016/j.neucom.2020.11.028.
URL http://www.sciencedirect.com/science/article/pii/S0925231220318154

[21] X. Liu, S. Wang, Y. Zhang, D. Liu, W. Hu, Automatic fluid segmentation in retinal optical coherence tomography images using attention based deep learning, Neurocomputing (Jan. 2021). doi:10.1016/j.neucom.2020.07.143.
URL http://www.sciencedirect.com/science/article/pii/S0925231220319135

[22] W. Takrouni, A. Douik, Improving geometric P-norm-based glioma segmentation through deep convolutional autoencoder encapsulation, Biomedical Signal Processing and Control 71 (2022) 103232 (Jan. 2022). doi:10.1016/j.bspc.2021.103232.
URL https://linkinghub.elsevier.com/retrieve/pii/S1746809421008296

[23] L. Fang, X. Wang, Brain tumor segmentation based on the dual-path network of multi-modal MRI images, Pattern Recognition 124 (2022) 108434 (Apr. 2022). doi:10.1016/j.patcog.2021.108434.
URL https://linkinghub.elsevier.com/retrieve/pii/S0031320321006105

[24] M. Jiang, F. Zhai, J. Kong, A novel deep learning model DDU-net using edge features to enhance brain tumor segmentation on MR images, Artificial Intelligence in Medicine 121 (2021) 102180 (Nov. 2021). doi:10.1016/j.artmed.2021.102180.
URL https://linkinghub.elsevier.com/retrieve/pii/S0933365721001731

[25] C. Chai, P. Qiao, B. Zhao, H. Wang, G. Liu, H. Wu, W. Shen, C. Cao, X. Ye, Z. Liu, S. Xia, Brain gray matter nuclei segmentation on quantitative susceptibility mapping using dual-branch convolutional neural network, Artificial Intelligence in Medicine 125 (2022) 102255 (Mar. 2022). doi:10.1016/j.artmed.2022.102255.
URL https://linkinghub.elsevier.com/retrieve/pii/S0933365722000203

[26] S. Bose, R. Sur Chowdhury, R. Das, U. Maulik, Dense Dilated Deep Multiscale Supervised U-Network for biomedical image segmentation, Computers in Biology and Medicine 143 (2022) 105274 (Apr. 2022). doi:10.1016/j.compbiomed.2022.105274.
URL https://linkinghub.elsevier.com/retrieve/pii/S001048252200066X

[27] H. Li, A. Menegaux, B. Schmitz-Koep, A. Neubauer, F. J. B. Bäuerlein, S. Shit, C. Sorg, B. Menze, D. Hedderich, Automated claustrum segmentation in human brain mri using deep learning, Human Brain Mapping 42 (18) (2021) 5862–5872 (Dec. 2021). doi:10.1002/hbm.25655.
URL https://onlinelibrary.wiley.com/doi/10.1002/hbm.25655

[28] D. S. Shibu, S. S. Priyadharsini, Multi scale decomposition based medical image fusion using convolutional neural network and sparse representation, Biomedical Signal Processing and Control 69 (2021) 102789 (Aug. 2021). doi:10.1016/j.bspc.2021.102789.
URL https://www.sciencedirect.com/science/article/pii/S1746809421003864

[29] Z. Wang, X. Li, H. Duan, Y. Su, X. Zhang, X. Guan, Medical image fusion based on convolutional neural networks and non-subsampled

contourlet transform, Expert Systems with Applications 171 (2021) 114574 (Jun. 2021). doi:10.1016/j.eswa.2021.114574.
URL https://www.sciencedirect.com/science/article/pii/S0957417421000154

[30] M. Liu, L. Jiao, X. Liu, L. Li, F. Liu, S. Yang, C-CNN: Contourlet Convolutional Neural Networks, IEEE Transactions on Neural Networks and Learning Systems 32 (6) (2021) 2636–2649, conference Name: IEEE Transactions on Neural Networks and Learning Systems (Jun. 2021). doi:10.1109/TNNLS.2020.3007412.

[31] L. Fang, H. Zhang, J. Zhou, X. Wang, Image classification with an RGB-channel nonsubsampled contourlet transform and a convolutional neural network, Neurocomputing 396 (2020) 266–277 (Jul. 2020). doi:10.1016/j.neucom.2018.10.094.
URL https://www.sciencedirect.com/science/article/pii/S0925231219304394

[32] L. Li, L. Ma, L. Jiao, F. Liu, Q. Sun, J. Zhao, Complex Contourlet-CNN for polarimetric SAR image classification, Pattern Recognition 100 (2020) 107110 (Apr. 2020). doi:10.1016/j.patcog.2019.107110.
URL https://www.sciencedirect.com/science/article/pii/S003132031930411X

[33] A.-A. Nahid, Y. Kong, Histopathological Breast-Image Classification Using Local and Frequency Domains by Convolutional Neural Network, Information 9 (1) (2018) 19, number: 1 Publisher: Multidisciplinary Digital Publishing Institute (Jan. 2018). doi:10.3390/info9010019.
URL https://www.mdpi.com/2078-2489/9/1/19

[34] Y. Zhao, Y. Li, B. Yang, Low-Frequency Desert Noise Intelligent Suppression in Seismic Data Based on Multiscale Geometric Analysis Convolutional Neural Network, IEEE Transactions on Geoscience and Remote Sensing 58 (1) (2020) 650–665 (Jan. 2020). doi:10.1109/TGRS.2019.2938836.

[35] Y. Liu, C. Zhang, J. Cheng, X. Chen, Z. J. Wang, A multi-scale data fusion framework for bone age assessment with convolutional neural networks, Computers in Biology and Medicine 108 (2019) 161–173 (May 2019). doi:10.1016/j.compbiomed.2019.03.015.
URL https://linkinghub.elsevier.com/retrieve/pii/S0010482519300885

[36] G. Yang, J. Yang, Z. Lu, D. Liu, A convolutional neural network with sparse representation, Knowledge-Based Systems 209 (2020) 106419 (Dec. 2020). doi:10.1016/j.knosys.2020.106419.
URL https://www.sciencedirect.com/science/article/pii/S0950705120305487

[37] M. Imani, Texture feed based convolutional neural network for pansharpening, Neurocomputing 398 (2020) 117–130 (Jul. 2020). doi:10.1016/j.neucom.2020.02.083.
URL https://linkinghub.elsevier.com/retrieve/pii/S0925231220302794

[38] Y. Li, L.-M. Po, C.-H. Cheung, X. Xu, L. Feng, F. Yuan, K.-W. Cheung, No-Reference Video Quality Assessment With 3D Shearlet Transform and Convolutional Neural Networks, IEEE Transactions on Circuits and Systems for Video Technology 26 (6) (2016) 1044–1057, conference Name: IEEE Transactions on Circuits and Systems for Video Technology (Jun. 2016). doi:10.1109/TCSVT.2015.2430711.

[39] H. Hermessi, O. Mourali, E. Zagrouba, Convolutional neural network-based multimodal image fusion via similarity learning in the shearlet domain, Neural Computing and Applications 30 (7) (2018) 2029–2045 (Oct. 2018). doi:10.1007/s00521-018-3441-1.
URL http://link.springer.com/10.1007/s00521-018-3441-1

[40] R. Hou, D. Zhou, R. Nie, D. Liu, X. Ruan, Brain CT and MRI medical image fusion using convolutional neural networks and a dual-channel spiking cortical model, Medical & Biological Engineering & Computing 57 (4) (2019) 887–900 (Apr. 2019). doi:10.1007/s11517-018-1935-8.
URL http://link.springer.com/10.1007/s11517-018-1935-8

[41] H. Hermessi, O. Mourali, E. Zagrouba, Deep feature learning for soft tissue sarcoma classification in MR images via transfer learning, Expert Systems with Applications 120 (2019) 116–127 (Apr. 2019). doi:10.1016/j.eswa.2018.11.025.
URL https://www.sciencedirect.com/science/article/pii/S0957417418307450

[42] H. Rezaeilouyeh, A. Mollahosseini, M. H. Mahoor, Microscopic medical image classification framework via deep learning and shearlet transform, Journal of Medical Imaging 3 (4) (Oct. 2016).

doi:10.1117/1.JMI.3.4.044501.
URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5093219/

[43] M. Liang, Z. Ren, J. Yang, W. Feng, B. Li, Identification of Colon Cancer Using Multi-Scale Feature Fusion Convolutional Neural Network Based on Shearlet Transform, IEEE Access 8 (2020) 208969–208977, conference Name: IEEE Access (2020). doi:10.1109/ACCESS.2020.3038764.

[44] E. Jabason, M. O. Ahmad, M. N. S. Swamy, Shearlet based Stacked Convolutional Network for Multiclass Diagnosis of Alzheimer's Disease using the Florbetapir PET Amyloid Imaging Data, in: 2018 16th IEEE International New Circuits and Systems Conference (NEWCAS), IEEE, Montreal, QC, 2018, pp. 344–347 (Jun. 2018). doi:10.1109/NEWCAS.2018.8585550.
URL https://ieeexplore.ieee.org/document/8585550/

[45] H. Arabi, G. Zeng, G. Zheng, H. Zaidi, Novel adversarial semantic structure deep learning for mri-guided attenuation correction in brain pet/mri, European journal of nuclear medicine and molecular imaging 46 (13) (2019) 2746–2759 (2019).

[46] S. Klein, M. Staring, K. Murphy, M. A. Viergever, J. P. Pluim, Elastix: a toolbox for intensity-based medical image registration, IEEE transactions on medical imaging 29 (1) (2010) 196–205 (2010).

[47] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical image computing and computer-assisted intervention, Springer, 2015, pp. 234–241 (2015).

[48] S. A. Taghanaki, K. Abhishek, J. P. Cohen, J. Cohen-Adad, G. Hamarneh, Deep semantic segmentation of natural and medical images: A review, Artificial Intelligence Review 54 (1) (2021) 137–178 (2021).

[49] A. L. Da Cunha, J. Zhou, M. N. Do, The nonsubsampled contourlet transform: theory, design, and applications, IEEE transactions on image processing 15 (10) (2006) 3089–3101 (2006).

[50] Arthur cunha, nonsubsampled contourlet toolbox, matlab central file exchange., https://www.mathworks.com/matlabcentral/fileexchange/10049-nonsubsampled-contourlet-toolbox, accessed: 2021-02-15.

[51] Z. Qu, Y. Xing, Y. Song, An image enhancement method based on non-subsampled shearlet transform and directional information measurement, Information 9 (12) (2018) 308 (2018).

[52] P. Subramanian, N. Alamelu, M. Aramudhan, Image fusion method using non-subsampled shearlet transform and fuzzy and simple fuzzy neural network algorithms, The Journal of Defense Modeling and Simulation 13 (1) (2016) 23–33 (2016).

[53] G. Kutyniok, W.-Q. Lim, R. Reisenhofer, Shearlab 3d: Faithful digital shearlet transforms based on compactly supported shearlets, ACM Transactions on Mathematical Software (TOMS) 42 (1) (2016) 1–42 (2016).

[54] F. Liu, H. Jang, R. Kijowski, T. Bradshaw, A. B. McMillan, Deep learning mr imaging–based attenuation correction for pet/mr imaging, Radiology 286 (2) (2017) 676–684 (2017).

[55] H. Jang, F. Liu, G. Zhao, T. Bradshaw, A. B. McMillan, Technical Note: Deep learning based MRAC using rapid ultrashort echo time imaging, Medical Physics 45 (8) (2018) 3697–3704 (2018). doi:10.1002/mp.12964.

[56] R. R. Kelly, S. J. Sidles, A. C. LaRue, Effects of Neurological Disorders on Bone Health, Frontiers in Psychology 11, publisher: Frontiers (2020). doi:10.3389/fpsyg.2020.612366.