



A Texture Neural Network to Predict the Abnormal Brachial Plexus from Routine Magnetic Resonance Imaging

Weiguo Cao, Benjamin Howe, Nicholas Rhodes, Sumana Ramanathan, Panagiotis Korfiatis, Kimberly Amrami, Robert Spinner, and Timothy Kline(✉)

Mayo Clinic, Rochester, Minnesota 55905, USA
kline.timothy@mayo.edu

Abstract. Brachial plexopathy is a form of peripheral neuropathy, which occurs when there is damage to the brachial plexus (BP). However, the diagnosis of breast cancer related BP from radiological imaging is still a great challenge. This paper proposes a texture pattern based convolutional neural network, called TPPNet, to carry out abnormal prediction of BP from multiple routine magnetic resonance image (MRI) pulse sequences, i.e. T2, T1, and T1 post-gadolinium contrast administration. Different from classic CNNs, the input of the proposed TPPNet is multiple texture patterns instead of images. This allows for direct integration of radiomic (i.e. texture) features into the classification models. Beyond conventional radiomic features, we also developed a new family of texture patterns, called triple point patterns (TPPs), to extract huge number of texture patterns as representations of BP' heterogeneity from its MRIs. These texture patterns share the same size and show very stable properties under several geometric transformations. Then, the TPPNet is proposed to carry out the differentiation task of abnormal BP for our study. It has several special characteristics including 1) avoidance of image augmentation, 2) huge number of channels, 3) simple end-to-end architecture, 4) free from the interference of multi-texture-pattern arrangements. Ablation study and comparisons demonstrate that the proposed TPPNet yields outstanding performances with the accuracies of 96.1%, 93.5% and 93.6% over T2, T1 and post-gadolinium sequences which exceed at least 1.3%, 5.3% and 3.4% over state-of-the-art methods for classification of normal vs. abnormal brachial plexus.

Keywords: Brachial Plexus · Magnetic Resonance Imaging · Heterogeneity · Radiomics · Deep learning · Convolutional Neural Networks · Texture pattern

1 Introduction

Brachial plexopathy is a form of peripheral neuropathy [1]. It occurs when there is damage to the brachial plexus (BP) which is a complex nerve network under the skin of the shoulder. There is a wide range of disease that may cause a brachial plexopathy.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43993-3_46.

Radiation fibrosis, primary and metastatic lung cancer, and metastatic breast cancer account for almost three-fourths of causes [2]. Brachial plexus syndrome occurs not infrequently in patients with malignant disease. It is due to compression or direct invasion of the nerves by tumor which will bring many serious symptoms [3]. Our research focuses on the brachial plexopathy caused by metastatic breast cancers.

Magnetic resonance imaging (MRI) and ultrasound of the brachial plexus have become two reliable diagnostic tools for brachial plexopathy [4]. Automatic identification of the BP in MRI and ultrasound images has become a hot topic. Currently, most of relevant research in this field are focusing on Ultrasound modality [5–8]. Compared with ultrasound, MRI has become the primary imaging technique in the evaluation of brachial plexus pathology [9]. However, to our knowledge, radiomics related BP studies utilizing MRI have not been reported previously.

Many radiomics studies have experimentally demonstrated that image texture has great potential for differentiation of different tissue types and pathologies [10]. In the past several decades, many state-of-the-art methods have been proposed to extract texture patterns [11, 12]. However, how to most effectively combine texture features with deep learning, called deep texture, is still an open area of research. One prior approach, termed GLCM-CNN, was proposed to carry out a polyp differentiation task [13]. However, how to arrange these GLCMs to form the 3D volume to optimize the performance is a major challenge.

With the goal of classifying normal from abnormal BP, we explored the approach of deep texture learning. This paper constructed a BP dataset with the most commonly used BP MRIs in our clinical practice. Considering the shortcoming of traditional patterns, triple point pattern (TPP) is proposed for the quantitative representation of the heterogeneity of abnormal BP's. In contrast to GLCM-CNN, TPPNet is designed to train models by feeding TPP matrices as the input with a huge number of channels. Finally, we analyze the model's performance in the experimental section. The major contributions of this study include 1) directed triangle construction idea for TPP, 2) huge number of TPP matrices as the heterogeneity representations of BP, 3) TPPNet with 15 layers and huge number of channels, 4) the BP dataset containing MR images and their corresponding ROI masks.

2 Materials and Method

2.1 Dataset Preparation and Preprocessing

Following IRB approval for this study, we search for patients with metastatic breast cancer who had a breast cancer MRI performed between 2010 and 2020 and had morphologically positive BP on the MRI report from our electronic medical records (EMR) in * hospital. Totally, 189 patients including 141 normal patients and 41 abnormal ones are obtained. The range of the age are varying from 15 to 85 years old. The female patient number and male patient number are almost even. Their weights by kg are in the range of [40.8 kg, 145 kg]. All patient experiences three kinds of MRI sequences including T2, T1 and post-gadolinium. Approximately 25% sequences are conducted with GE SIGNA HDI 1.5T and the rest are produced by GE SIGNA™ Hero 3.0T. Patients are required to maintain a decubitus position while scanning. All series are scanned in sagittal view.

Slice thicknesses are varying in [1.2 mm, 6.0 mm] and approximate 96% of them share the slice thickness of 4mm. Their slice number is in the range of [38, 72]. Most of slice resolutions are 512 * 512 while less than 2% of them are 256 * 256.

Totally, we collect approximate 807 series which include 274 T2, 254 T1 and 279 Post-gadolinium. Since some scans are seriously degraded due to motion artifacts. Therefore, each case underwent several essential image adjustments such as multi-series splitting, two-series merging, slice swapping, artifact checking and boundary corrections.

Table 1. The final BP dataset including MRIs and their corresponding masks of T2, T1, Post-gadolinium (Pg).

| Image | Total | Normal | | | Abnormal | | |
|-------|-------|--------|-----|-----|----------|----|----|
| | | T2 | T1 | Pg | T2 | T1 | Pg |
| MRI | 462 | 123 | 123 | 123 | 31 | 31 | 31 |
| Mask | 462 | 123 | 123 | 123 | 31 | 31 | 31 |

To yield the ROI, firstly, we randomly sampled ~40% of the sequences including both normal and abnormal ones that were manually segmented with ITK- snap by two skilled trainees [14, 15]. Then, the manual segmentations were utilized to train a 3D nnUNet model which was utilized to train the model which was used to predict ROIs for the rest series [16]. The predicted segmentations were manually divided into three groups, i.e. good, fair and poor. Good cases were added to the training set. This process was repeated until no improvements in the predictions for the remaining sequences was seen. The final dataset for radiomic analysis was constructed by merging the datasets for each sequence type. Only patients that had all three sequences segmented (T2, T1 and Post-gadolinium) were included in the dataset. Table 1 shows a breakdown of the final dataset.

2.2 Triple Point Pattern (TPP)

Theoretically, some texture pattern methods such as LBP, LTP, and GLDM, are based on single-variance pixel functions [17–19]. Therefore, they extract local texture features coded by the difference or difference counts between the concerned pixel and its neighboring ones. One obvious shortcoming is the absence of global properties which need other statistical methods as the aid to yield, such as histogram and invariants [20, 21]. Meanwhile, some other texture methods are generally defined by two-variance functions that only focus on two-variance patterns, such as (pixel, pixel), (pixel, neighbor count) [22–24]. In general, image textures extracted by these methods contain both local texture properties and global texture information. Their shortcomings might come from pattern shapes which might lead to the overfitting risk while combining with deep learning since the yielded texture matrix might have slim shapes or adaptive columns. In summary, as the requirement of the image texture and deep learning, *an excellent image texture pattern should have some essential features including* 1) local properties to characterize the micro-unit of the image texture, 2) global properties to represent

the macro-structure of the image texture, 3) uniform shapes under nonuniform-shape images, 4) invariant or robustness under some common geometric transforms such as rotation, scaling and so on.

According above requirements, we developed a method to produce a serial of novel texture patterns by introducing a directed triangle idea with an adjacent triple pixel as a ternary group, called triple point pattern (TPP), to extract the local texture information. Then, a statistical method like histogram is employed to count the number of the same type of pixel-triplets within the ROI or throughout the whole image. Finally, a three-dimensional (3D) TPP matrix is formed to characterize the image texture globally as the following:

Two-dimensional image:

$$TPP_{(p_i, p_c, p_j)}(x, y, z) = \sum_{m=0}^{M-1} \sum_{j=0}^{N-1} \begin{cases} I((m, n) + p_i) = x & \\ 1 & I((m, n) + p_c) = y & \\ & I((m, n) + p_j) = z & \\ 0 & \text{others} \end{cases} \quad (1)$$

where I is a $M \times N$ image, x, y , and z is the pixel triplet, $x, y, z \in [0, L]$, L is its gray level, $p_c = (0, 0)$ denotes the concerned pixel such as p_0 in Fig. 1, p_i and p_j are p_c 's two adjacent pixels, $i, j \in [1, H]$, H is the number of p_c 's adjacent points.

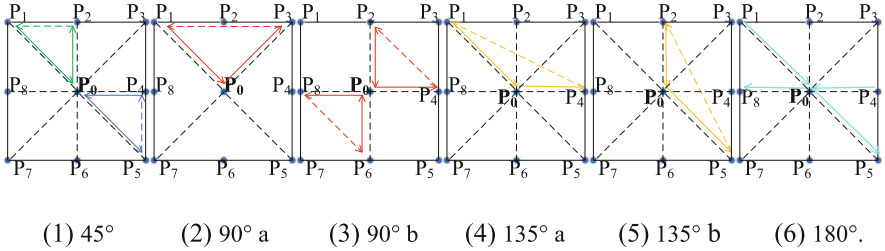


Fig. 1. Directed triangle idea for TPP construction in 2D images where p_0 is the concerned pixel, $p_1 \dots p_8$ are its adjacent pixels. Totally, they are four modes according to four concerned angles, i.e. 45°, 90°, 135° and 180°. (1) is the 1st mode. Both (2) and (3) are the 2_{nd} mode.

Three-dimensional image:

$$TPP_{(p_i, p_c, p_j)}(x, y, z) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \sum_{k=0}^{K-1} \begin{cases} I((m, n, k) + p_i) = x & \\ 1 & I((m, n, k) + p_c) = y & \\ & I((m, n, k) + p_j) = z & \\ 0 & \text{others} \end{cases} \quad (2)$$

where I is a three dimensional image with the shape of $M \times N \times K$, $p_c = (0, 0, 0)$, other parameters are similar to the Two-dimensional image.

In statistics, a TPP matrix should a 3D distribution of directed triangle. As shown in Fig. 1, the TPP is formed by the concerned pixel and its two adjacent pixels in two-dimensional(2D) images. Similarly, the TPP in 3D images is constructed by one

concerned voxel and its two neighboring voxels. More details could be found in the supplementary material. As the construction idea of TPP, there are four independent modes categorized by the concerned angle, i.e. 45° , 90° , 135° and 180° in 2D images, which produce 8 TPPs, 8 TPPs, 8 TPPs and 8 TPPs respectively. Analogously, the 3D image has twelve independent angle modes, i.e. 35.26° , 45° , 54.74° , 60° , 70.53° , 90° , 109.47° , 120° , 125.26° , 135° , 144.74° , and 180° . These angle modes could generate 24 TPPs, 24 TPPs, 24 TPPs, 24 TPPs, 12 TPPs, 96 TPPs, 12 TPPs, 24 TPPs, 24 TPPs, 24 TPPs, 24 TPPs and 13 TPPs respectively. Totally there are 32 TPPs in 2D images and 325 TPPs in 3D images and every TPP could produce one corresponding TPP matrix.

By further analysis, we could find some TPP pairs have an isomorphism relationship since its TPP matrix could be generated by transposing or flipping another TPP matrix on some certain conditions when two triangles formed by the pixel triplet have the relevance of shifting or scaling.

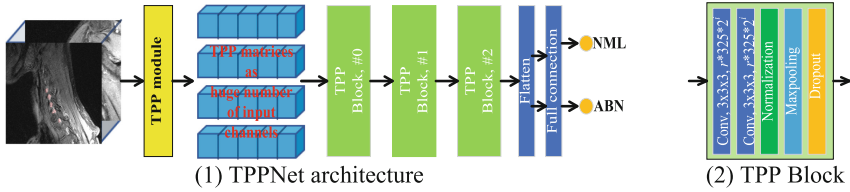


Fig. 2. The pipeline of the proposed TPPNet over 3D images where NML denote normal, ABN denotes abnormal, i in (2) is the block id, r is the adaptive argument to control the filter number.

For an instance, the TPP matrix by pixel-triplet (p_1, p_0, p_2) in Fig. 1 (1), the TPP matrix by (p_4, p_0, p_5) in Fig. 1 (1) and the TPP matrix by (p_6, p_0, p_8) in Fig. 1(3) have the following relevance:

$$TPP_{(p_1, p_0, p_2)} = T * F(TPP_{(p_4, p_0, p_5)}) = T * F(TPP_{(p_6, p_0, p_8)}) \quad (3)$$

where T denotes matrix transposing, F represents matrix flipping, $*$ is the product operator. Since both T and F are continuous bijective mappings, these three TPPs are isomorphic.

In our study, these isomorphic TPP matrices are not dropped from the TPP matrix set because they are equivalent to image rotations and re-scaling. Image scaling can result in the image pixels increasing. The normalization could almost remove the effect of pixel increase caused by scaling transformations. Moreover, TPPs generated by 135° could also be treated as affine transformations. Therefore, data augmentation could be omitted when we combine TPP with deep learning for this study. As its definition, the TPP matrix should be a cubic array with the shape of $L \times L \times L$ where L is the gray level of the image.

2.3 TPPNet

The pipeline of the proposed method TPPNet is illustrated in Fig. 2. The TPP matrix calculation is the preprocessing module which feeds MRI and its ROI and yield TPP

matrix set. The following step is the TPPNet architecture to yield training models. Based on the construction idea of TPP, the size of the TPP matrix depends on the gray level of the image. For the same image or ROI, the larger the gray level, the sparser the matrix will be. The sparse matrix would lead to overfitting while training the model. Therefore, the image requires a re-scaling step to lower its gray level to avoid the sparsity of the TPP matrix. Consequently, our proposed TPPNet only contains three convolution blocks consisting of 15 layers. Each block has two convolution, one normalization, one max-pooling and one dropout layer. *It has four particular features as follows:*

- 1) Avoidance of image augmentation. Due to the stability of TPP matrix under rotation, scale and affine transformations, image augmentation could be omitted in the preprocessing step which can lead to image deformation.
- 2) Huge number of channels. TPPNet treats each TPP as an independent channel. For 2D images, there are at least 32 channels if more displacements of TPP is considered. Similarly, we could generate no less than 325 TPPs in 3D images.
- 3) Simple end-to-end architecture. We integrate the k-fold cross-validation, TPP generation and model training into one framework. Since the TPP matrix is always small, there are only 15 layers in TPP which could reduce the risk of overfitting issue met in deeper neural networks.
- 4) Free from the interference of multi-texture-pattern arrangements. Since each channel is corresponding with one TPP, it can solve the pattern arrangement issue occurred in GLCM-CNN.

Table 2. Test performances for different gray levels over T2, T1 and Post-gadolinium where Acc denotes accuracy and Pg denotes Post-gadolinium.

| Gray level | T2 | | T1 | | Pg | |
|------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| | Acc | Loss | Acc | Loss | Acc | Loss |
| 8 | 0.942±0.032 | 0.319±0.028 | 0.902±0.022 | 0.395±0.207 | 0.929±0.023 | 0.292±0.087 |
| 12 | 0.961±0.025 | 0.277±0.107 | 0.935±0.021 | 0.307±0.164 | 0.936±0.027 | 0.279±0.046 |
| 16 | 0.948±0.034 | 0.350±0.333 | 0.922±0.028 | 0.306±0.107 | 0.921±0.041 | 0.313±0.053 |
| 20 | 0.948±0.026 | 0.342±0.126 | 0.922±0.035 | 0.363±0.098 | 0.942±0.034 | 0.302±0.124 |
| 24 | 0.948±0.034 | 0.348±0.162 | 0.929±0.023 | 0.309±0.067 | 0.922±0.046 | 0.322±0.176 |

3 Experiments

3.1 Preparations

Some important specificities of our computing platform contain: one AMD EPYC 7352 24-Core Processor, 1 TB memory and four Nivida A100-SXM GPUs with 320 GB GPU memory. The whole dataset is divided into three subsets according to the MR sequence, i.e. T2, T1 and post-gad. For each subset, both normal cases and abnormal cases were randomly and evenly split into five subgroups. A five-fold cross-validation

scheme was employed to generate five cohorts. Totally, 15 cohorts were produced. Each cohort consists of training set, validation set and testing set by the ratio of 6:2:2.

3.2 Ablation Studies

Since all images in our dataset are 3D images, therefore, the initial channel is set 325 which is equal to the TPP number. The loss functions in the following experiments shared categorical_crossentropy. Nadam is adopted as the optimizer with the learning rate of 0.0001 and batch size of 8 for 200 epochs. All performances listed in this section are the average of performances with 5-fold cross-validation.

3.2.1 Impact of Gray Level

The image gray level determines the shape of the TPP matrix. To avoid its sparsity, we compute the TPP matrix set via Eq. (2) with gray levels of 8, 12, 16, 20, 24. While rescaling the image intensity, an arc tangent approach is utilized to yield the new image. The models are trained and tested over 15 cohorts. Their performances evaluated by accuracies are listed in Table2 which tells us that T2 sequence yields the highest accuracy of 96.1% when the gray level is 12. T1 and post-gadolinium also get acceptable results with the accuracies of 93.5% and 93.6% respectively. Other performances could be read in supplementary materials.

Table 3. Test performances for three intensity rescaling approaches over T2, T1 and Post-gadolinium where the gray level is set 12, Acc denotes accuracy, Pg denotes Post -gadolinium, and Artan is arc tangent function.

| Rescaling approach | T2 | | T1 | | Pg | |
|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| | Acc | Loss | Acc | Loss | Acc | Loss |
| Min-max | 0.947±0.027 | 0.303±0.082 | 0.928±0.026 | 0.312±0.138 | 0.920±0.028 | 0.475±0.327 |
| Adaptive | 0.921±0.035 | 0.385±0.158 | 0.902±0.022 | 0.462±0.092 | 0.856±0.020 | 0.517±0.092 |
| Artan | 0.961±0.025 | 0.277±0.107 | 0.935±0.021 | 0.307±0.164 | 0.936±0.027 | 0.279±0.046 |

Table 4. Test performances comparison between multi-channel and solo-channel over T2, T1 and Post-gadolinium where the intensity rescaling function is arc tangent, the gray level is set 20, Acc denotes accuracy and Pg denotes Post -gadolinium.

| Input channel | T2 | | T1 | | Pg | |
|---------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| | Acc | Loss | Acc | Loss | Acc | Loss |
| 325 | 0.948±0.026 | 0.342±0.126 | 0.922±0.035 | 0.363±0.098 | 0.934±0.034 | 0.302±0.124 |
| 1 | 0.863±0.027 | 0.274±0.091 | 0.810±0.145 | 0.458±0.105 | 0.811±0.047 | 0.590±0.107 |

3.2.2 Impact of Intensity Rescaling Approaches

Rescaling approaches of image intensity could also bring impacts on the BP's differentiation while producing the TPP matrix. The commonly used methods include min-max-linear approach [25], arc tangent approach [26], and adaptive rescaling approach [27]. To test the performances fairly, we test above rescaling methods at the same gray level 12. The yielded performances evaluated by accuracies are shown in Table 3 where arc tangent method achieves the highest accuracy of 96.1% over the T2 sequence. Other performances are shown in supplementary materials.

3.2.3 Multi-Channels vs Solo-Channel

We carry out experiments to train the TPPNet model and make tests with arc tangent rescaling approach under gray level 16. As a comparison, we test single channel mode as well. By sharing every TPP matrix's label with the original case label, our TPPNet works well by assigning one channel for the initial input. Once the trained model with solo channel is generated, all TPP matrices of the testing set could be tested. Hereafter, we adopted a voting method to determine if the prediction is normal, if the predicted probability is less than 0.5, otherwise, it is considered abnormal. Performances with accuracies and loss are listed in Table 4. Other performances are listed in supplementary materials.

3.3 Comparisons

We evaluated our proposed TPPNet by comparing it to the recent state-of-the-art approaches over our BP dataset including VGG16 [28], InceptionNet [29], MobileNet [30], GLCM-CNN [13], ViT [31]. The GLCM size of 32×32 is used in

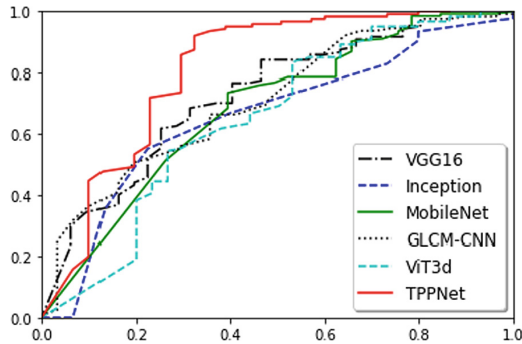


Fig. 3. ROC curves of TPPNet and five state-of-the-art approaches.

GLCM_CNN. All approaches shared the same image shape of $128 \times 128 \times 64$ with 1 channel. The patch shape of ViT is $8 \times 8 \times 8$, projection dim is 64, attention head number is set 4 with 8 transformer layers. The intensity rescaling step adopts the arc tangent approach. Other parameters are similar to the ablation study. Their performances are

Table 5. Performance comparison between TPPNet T1 sequence where Acc is accuracy, AUC denotes Spe denotes specificity. The grey level is set 12.

| Approach | Acc | AUC | Sen | 1-Spe | F1-Score |
|-----------|--------------------|--------------------|--------------------|--------------------|--------------------|
| VGG16 | 0.837±0.106 | 0.739±0.110 | 0.910±0.125 | 0.548±0.124 | 0.749±0.114 |
| Inception | 0.805±0.014 | 0.465±0.102 | 1.000±0.000 | 0.033±0.067 | 0.474±0.061 |
| Mobilenet | 0.869±0.033 | 0.616±0.170 | 0.985±0.030 | 0.405±0.258 | 0.719±0.107 |
| GLCM-CNN | 0.851±0.019 | 0.704±0.113 | 0.975±0.050 | 0.362±0.249 | 0.707±0.073 |
| ViT3D | 0.882±0.020 | 0.675±0.163 | 0.969±0.043 | 0.533±0.245 | 0.782±0.058 |
| TPPNet | 0.935±0.021 | 0.832±0.101 | 1.000±0.000 | 0.676±0.107 | 0.881±0.045 |

demonstrated in Table 5 which contains the accuracy, AUC score, specificity, sensitivity, and F1 score. Their ROC curves over T1 sequence are plotted in Fig. 3. T2 and Post-gadolinium’s performances could be found in supplementary materials.

4 Conclusions

In this paper, we develop an approach to carry out the pioneer study of differentiating abnormal BP from normal ones relevant to breast cancer. In particular, TPP is proposed to extract texture features as the representation of BP’s heterogeneity from MRIs. Moreover, a TPPNet with huge number of initial channels is designed to train the model. To testify our proposed TPPNet, a BP dataset is constructed with 452 series including three most commonly used MR sequences in clinical practice, i.e. T2, T1 and Post-gadolinium. The best result is yielded when the gray level is 12, intensity rescaling method adopts arc tangent approach. Experimental outcomes also demonstrate that the proposed TPPNet not only exhibit more stable performances but also outperform six famous state-of-the-art approaches over three most commonly used BP MR sequences.

References

1. Shabeeb, D., Musa, A.E., et al.: Brachial plexopathy as a complication of radio-therapy: a systematic review. *Curr. Cancer Ther. Rev.* **15**(2), 110–120 (2020)

2. Wittenberg, K.H., Adkins, M.C.: MR imaging of nontraumatic brachial plex-opathies: frequency and spectrum of findings. *Radiographics* **20**(4), 1024–1032 (2004)

3. Nisce, L.Z., Chu, F.C.H.: Radiation therapy of brachial plexus syndrome from breast cancer. *Radiology* **91**(5), 1022–1025 (1968)

4. Lutz, A.M., Gold, G., Beaulieu, C.: MR imaging of the brachial plexus. *Magn. Reson. Imaging Clin. N. Am.* **20**(4), 791–826 (2012)

5. Wang, R., Shen, H., Zhou, M.: Ultrasound nerve segmentation of brachial plexus based on optimized resu-net. In: 2019 IEEE International Conference on Imaging Systems and Techniques (IST). pp. 1–6 (2019)

6. Pisda, K., Jain, P., Sisodia, D.S.: Deep networks for brachial plexus nerves segmentation and detection using ultrasound images. In: Garg, L., Kesswani, N., Vella, J.G., Xuereb, P.A., Lo, M.F., Diaz, R., Misra, S., Gupta, V., Randhawa, P. (eds.) ISMS 2020. LNNS, vol. 303, pp. 132–146. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-86223-7_13

7. Wang, Y., Geng, J., Zhou, C., Zhang, Y.: Segmentation of ultrasound brachial plexus based on u-net. In: 2021 International Conference on Communications, Information System and Computer Engineering (CISCE). pp. 482–485 (2021)
8. Tian, D., Wang, Q., et al.: Brachial plexus nerve trunk recognition from ultra-sound images: a comparative study of deep learning models. *IEEE Access* **10**, 82003–82014 (2022)
9. Sureka, J., Cherian, R.A., Alexander, M., Thomas, B.P.: MRI of brachial plex-opathies. *Clin. Radiol.* **64**(2), 208–218 (2009)
10. Lambin, P., Leijenaar, R.T.H., et al.: Radiomics: the bridge between medical imaging and personalized medicine. *Nat. Rev. Clin. Oncol.* **14**(12), 749–762 (2017)
11. van Griethuysen, J.J.M., Fedorov, A., et al.: Computational radiomics system to decode the radiographic phenotype. *Can. Res.* **77**(21), e104–e107 (2017)
12. Ramel, A.A.: Analysis of membrane process model from black box to machine learning. *J. Mach. Comput.* **2**(1), 2788–7669 (2022)
13. Tan, J., Lei, B., et al.: 3D-GLCM CNN: a 3-dimensional gray-level cosoccur-rence matrix-based CNN model for polyp classification via CT colonography. *IEEE Trans. Med. Imaging* **39**(6), 2013–2024 (2020)
14. Yoo, T.S., Ackerman, M.J., et al.: Engineering and algorithm design for an image processing API: a technical report on ITK – the insight toolkit. In: Westwood, J. (ed.) *Proceeding of Medicine Meets Virtual Reality*, pp. 586–592 (2002)
15. McCormick, M., Liu, X., et al.: Enabling reproducible research and open science. *Front. Neuroinform.* **8**, 13 (2014)
16. Isensee, F., Jaeger, P.F., et al.: nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **18**(2), 203–211 (2021)
17. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 971–987 (2002)
18. Tan, X., Triggs, B.: Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Trans. Image Process.* **19**(6), 1635–1650 (2010)
19. Vallières, M., Zwanenburg, A., et al.: Responsible radiomics research for faster clinical translation. *J. Nucl. Med.* **59**(2), 189–193 (2018)
20. Güner, A., Alçın, Ö.F., Şengür, A.: Automatic digital modulation classification using extreme learning machine with local binary pattern histogram features. *Measurement* **145**, 214–225 (2019)
21. Bian, M., Liu, J. K., et al.: Verifiable privacy-enhanced rotation invariant LBP feature extraction in fog computing. In: *IEEE Transactions on Industrial Informatics* (2023)
22. Cao, W., Liang, Z., Gao, Y., et al.: A dynamic lesion model for differentiation of malignant and benign pathologies. *Sci. Rep.* **11**, 3485 (2021)
23. Galavis, P.E., Hollensen, C., et al.: Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters. *Acta Oncol.* **49**(7), 1012–1016 (2010)
24. Doumou, G., Siddique, M., Tsoumpas, C., et al.: The precision of textural analysis in 18F-FDG-PET scans of oesophageal cancer. *Eur. Radiol.* **25**(9), 2805–2812 (2015)
25. Wahid, K.A., He, R., et al.: Intensity standardization methods in magnetic reso-nance imaging of head and neck cancer. *Phy. Imaging Radiat. Oncol.* **20**, 88–93 (2021)
26. Cao, W., Pomeroy, M.J., et al.: Lesion classification by model-based feature extraction: a differential affine invariant model of soft tissue elasticity. *arXiv pre-print* [arXiv:2205.14029](https://arxiv.org/abs/2205.14029) (2022)
27. Pomeroy, M.J., Pickhardt, P., Liang, J., Lu, H.: Histogram-based adaptive gray level scaling for texture feature classification of colorectal polyps. In: *Medical Imaging 2018: Computer-Aided Diagnosis*, vol. 10575, pp. 507–513. SPIE (2018)

28. Simonyan, K., Zisserman, A.: Very deep convolution newtworks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
29. Szegedy, C., Liu, W., et al.: Going deeper with convolutions. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, pp. 1–9 (2015)
30. Howard A.G., Zhu, M., et al.: MobileNets: efficient convolutional neural net-works for mobile vision applications. [arXiv:1704.04861](https://arxiv.org/abs/1704.04861) (2017)
31. Wu, B., et al.: Visual transformers: where do transformers really belong in vi-sion models?. In: 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, pp. 579–589 (2021)

Microscopy