ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 979-8-3503-4485-1/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICASSP48485.2024.10445975

BREAST ULTRASOUND COMPUTER-AIDED DIAGNOSIS USING STRUCTURE-AWARE TRIPLET PATH NETWORKS

Erlei Zhang[⋆] *Weihao Chen*[⋆] *Xiaowei Xu*[†] *Zhicheng Zhang*[≀] *Jinglei Li*[†]

[⋆] Northwest A&F University
[†] Guangdong Provincial People's Hospital
[≀] JancsiTech

ABSTRACT

Breast ultrasound (BUS) is an effective imaging modality for breast cancer diagnosis. The structural characteristics of breast lesions play an important role in computer-aided diagnosis. In this paper, a novel structure-aware triplet path network (SATPN) was designed to integrate classification and image reconstruction tasks to achieve accurate diagnosis on BUS images. Specifically, we enhanced clinically-approved structure characteristics of breast lesion by converting original BUS images to BI-RADS-oriented feature maps (BFMs) with a distance-transformation coupled Gaussian filter. Then, the converted BFMs were used as the inputs of the SATPN, which performed a supervised lesion classification task and two separate unsupervised stacked convolutional auto-encoder tasks for benign and malignant image reconstruction. We trained the SATPN with an alternative learning strategy by balancing image reconstruction error and classification label prediction error. The lesion label was determined by weighted voting of reconstruction error and label prediction error. We compared the performance of the SATPN with five deep learning methods using the original images and BFMs as inputs. Experimental results on two BUS datasets showed that SATPN performed the best among the six networks, with classification accuracy around 96%. These findings indicate that SATPN is promising for effective ultrasound computer-aided diagnosis of breast lesions.

Index Terms— Breast cancer, ultrasound, computer-aided diagnosis, deep learning, semi-supervised learning

1. INTRODUCTION

Breast ultrasound (BUS) is a widely used early imaging diagnostic method for breast cancer due to its non-invasiveness, safety, and affordability [1]. Extensive research has been conducted on computer-aided diagnosis (CAD) for breast ultrasound in order to assist radiologists and improve diagnostic accuracy [2]. Deep learning has gained wide application in CAD [3]. It has the ability to automatically learn discriminative features from raw data [4]. However, training deep networks from scratch for breast cancer ultrasound CAD is challenging due to the requirement for a large amount of labeled training data. There are popular approaches based on the unsupervised feature learning, which leverages unsupervised learning tasks like image reconstruction to extract high-level features. Cheng et al.[5] employed an unsupervised autoencoder method to extract high-level features for image reconstruction, followed by supervised fine-tuning of the classifier using labeled training samples for classification. However, existing unsupervised learning methods have limitations: (1) Unsupervised autoencoders may not generate suitable features for classification tasks as they capture features based on data characteristics without considering labels, making classification challenging. (2) Differentiating benign and malignant lesions using a single autoencoder structure for image reconstruction is difficult due to small structural differences, necessitating separate systems for each. (3) The image intensity-based ℓ_2-norm cost function is ineffective in capturing lesion structure attributes. The ℓ_2-norm has poor correlation with image quality as perceived by human observers[6] since it fails to consider the shape or contour attributes of breast lesions, which play crucial roles in breast diagnosis.

To address the challenges of limited training datasets and enhance the utilization of prior domain knowledge, this study proposes a novel approach called Structure-Aware Triplet Path Networks (SATPN) (as depicted in Fig. 1). The network combines clinical knowledge of lesion features (BI-RADS features[7]) with the triplet path network and the Structural Similarity Index Measure (SSIM) for accurate lesion classification in breast ultrasound images. In this approach, breast images are transformed into BI-RADS-oriented feature maps (BFMs) using the distance-transformation coupled Gaussian filter. BFMs not only preserve the original information of breast cancer ultrasound images but also enhance the structural features of lesions, including shape, lesion boundary, undulation, and angle features. Subsequently, the

This work is funded by the Guangdong Basic and Applied Basic Research Foundation (No.2022A1515011650), QinChuangyuan High-Level Innovation and Entrepreneurship Talent Program of Shaanxi (No.2021QCYRC4-50). (Corresponding author: Jinglei Li)

BFMs serve as inputs to SATPN, which performs multi-task learning by integrating two unsupervised image feature extraction pipelines based on Stacked Convolutional Autoencoders (SCAE) and a supervised lesion classification pipeline tailored for diagnosis. This integrated multi-path network allows SCAE (using SSIM as a cost function) to independently extract image features for benign and malignant lesions while being constrained by lesion classification, achieved by leveraging SCAE encoder features and convolutional networks. In other words, SATPN learns lower-level features through unsupervised reconstruction networks and utilizes supervised classification to guide the learning of lower-level features, which is advantageous for the classification task. Finally, lesions are classified through weighted voting using reconstruction errors and label prediction errors.

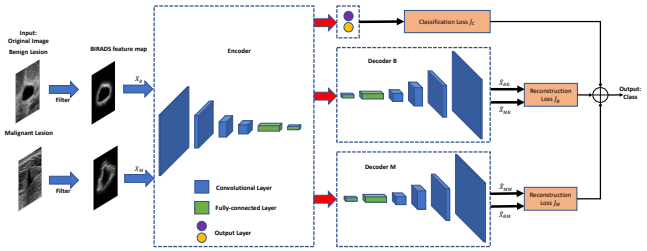


Fig. 1. Illustration of SATPN architecture. SATPN consists of two independent unsupervised SCAE networks for benign and malignant image reconstruction and a lesion classification network.

2. METHODS

In a previous paper [8], BFMs were proposed to explicitly use orientation, echo patterns, and a posteriori features in a deep learning approach. The original breast images were converted to BFMs in the following manner:

$$BFMs = \boldsymbol{I} \cdot \mathbf{e}^{-\frac{Dist(p)^2}{\sigma^2}} \tag{1}$$

The SCAE follows the unsupervised encoder-decoder paradigm [9]. Standard SCAE is trained to learn representative features for image reconstruction. In the final stage, parameters are fine-tuned using supervised learning, which may have limited impact on the classification task. More importantly, it does not consider the structural properties of breast lesions that play a crucial role in breast diagnosis.

The image reconstruction pipeline is similar to the standard SCAE structure, consisting of an encoder and a decoder, as shown in Fig. 1. The encoder maps the input x to a hidden feature representation h. Then, the decoder maps the hidden h to the output $\hat{x}$. $\hat{x}$ is the expected approximate reconstruction of the input x. SCAE model is trained to minimize the reconstruction error as much as possible.

In this paper, two image reconstruction pipelines are established, one for benign lesion images and the other for malignant lesion images. The design of the benign pipeline aims to reconstruct benign images well but struggles to reconstruct malignant images. Therefore, the objective function for the benign pipeline is as follows:

$$\begin{aligned}\min_{\theta_B} J_B = &\frac{1}{N_i}\sum_{i=1}^{N_i} loss_r\left(x_B^{(i)}, \hat{x}_{BB}^{(i)}\right) \\ &- \frac{\alpha}{N_j}\sum_{j=1}^{N_j} loss_r\left(x_M^{(j)}, \hat{x}_{MB}^{(j)}\right) + \gamma \cdot R\left(\theta_B\right)\end{aligned} \tag{2}$$

The malignant pipeline, on the contrary, has the opposite effect: it can reconstruct malignant images well but struggles to reconstruct benign images. Therefore, the objective function for the malignant pipeline is as follows:

$$\begin{aligned}\min_{\theta_M} J_M = &\frac{1}{N_j}\sum_{j=1}^{N_j} loss_r\left(x_M^{(j)}, \hat{x}_{MM}^{(j)}\right) \\ &- \frac{\alpha}{N_i}\sum_{i=1}^{N_i} loss_r\left(x_B^{(i)}, \hat{x}_{BM}^{(i)}\right) + \gamma \cdot R\left(\theta_M\right)\end{aligned} \tag{3}$$

Where $x_B^{(i)}$ and $x_M^{(j)}$ represent benign and malignant samples, respectively. $\hat{x}_{BB}^{(i)}$ and $\hat{x}_{MB}^{(j)}$ represent the reconstructions of $x_B^{(i)}$ and $x_M^{(j)}$ in the benign pipeline, while $\hat{x}_{MM}^{(j)}$ and $\hat{x}_{BM}^{(i)}$ represent the reconstructions of $x_B^{(i)}$ and $x_M^{(j)}$ in the malignant pipeline. Additionally, $\theta_B = [W_{en}, b_{en}; W_{Bde}, b_{Bde}]$ and $\theta_M = [W_{en}, b_{en}; W_{Mde}, b_{Mde}]$ are the weights used in the loss function for regularization purposes. The reconstruction loss function utilizes the SSIM [10]: $loss_r\left(x, \hat{x}\right) = 1 - SSIM\left(x, \hat{x}\right)$. The loss function is regularized by $R\left(\theta_B\right) = \|W_{en}\|_F^2 + \|W_{Bde}\|_F^2$ and $R\left(\theta_M\right) = \|W_{en}\|_F^2 + \|W_{Mde}\|_F^2$. The parameters $\{\alpha, \gamma\} \in [0, 1]$ balance the each term.

The lesion classification task is implemented by the encoder network and the softmax classifier, as shown in Fig. 1. $\hat{y} \in \Re^K$ represents the output of the classifier, ranging from 0 to 1. The objective function for classification is given by:

$$\min_{\theta_c} J\left(\theta_c\right) = \frac{1}{N}\sum\nolimits_{n=1}^{N} loss_c\left(y^{(n)}, \hat{y}^{(n)}\right) + \gamma \cdot R\left(\theta_c\right) \tag{4}$$

The variable $\theta_c = [W_{en}, b_{en}; \beta]$ is learned or adjusted through the training dataset. Typically, the network uses cross-entropy as the loss function during training. However, to incorporate clustering characteristics into the training results, a clustering penalty term needs to be added to the loss function. Therefore, the loss function $loss_c\left(\cdot\right)$ is given by:

$$\begin{aligned}loss_c\left(y, \hat{y}\right) = &-\sum_{k=1}^{K} \delta\left(y\left(k\right) = 1\right) \cdot \log\left(\hat{y}\left(k\right)\right) \\ &+ \mu\left\|x - c_y\right\|^2\end{aligned} \tag{5}$$

In the given formulation, $\delta\left(\cdot\right)$ is the indicator function, and c_y represents the training samples and centers (i.e., mean values) for each category. If the value is 1 if $y\left(k\right) = 1$ and K

is the number of classes. The loss function consists of two terms: the first term is the standard softmax cross-entropy loss, which helps to separate samples belonging to different classes by maximizing the distances between class representations. The second term is the clustering penalty term, which encourages the feature representations of training samples to be as close as possible to the centers of their respective classes (c_y). This term helps to minimize the distances between samples within the same class. In this case, since there are two classes, $K = 2$. Therefore, the loss function aims to both increase the separation between different classes and decrease the distance within the same class.

The objective function of SATPN, combining both image reconstruction pipelines and the classification pipeline, can be represented as follows:

$$\min_{\theta} \lambda \cdot (J_B + J_M) + (1 - \lambda) \cdot J_C \tag{6}$$

The balancing of classification and reconstruction tasks is achieved using $\theta = [\theta_B, \theta_M, \theta_C]$ and $\lambda \in [0, 1]$ in the feature learning process. The encoding parameters, denoted as W_{en}, are shared among the three tasks, while the decoding parameters are only involved in the reconstruction task. The objective function can be realized through alternating learning [11].

In testing phase, x is classified by combining the reconstruction error and the label prediction error:

$$\hat{y} = \min_{0,1} \{-\ln(P_{CB} \cdot P_{rB}), -\ln(P_{CM} \cdot P_{rM})\} \tag{7}$$

Where P_{CB} and P_{CM} are the category probability estimates from the classification pipeline, $\hat{x}_B$ and $\hat{x}_M$ are the outputs of the two reconstruction pipelines, $P_{rB} = SSIM(x, \hat{x}_B) / (SSIM(x, \hat{x}_B) + SSIM(x, \hat{x}_M))$ and $P_{rM} = SSIM(x, \hat{x}_M) / (SSIM(x, \hat{x}_B) + SSIM(x, \hat{x}_M))$.

In this paper, the classification pipeline consists of four convolutional layers: 8 (Conv1), 16 (Conv2), 32 (Conv3), and 64 (Conv4). Each Conv layer has a filter size of 3×3, followed by a max pooling layer of size 2×2. There are also three fully connected layers as shown in Fig. 1. The output layer has two neurons with a softmax function. ReLU activation is used in all hidden layers. The two reconstruction pipelines each have an encoder and a decoder. The encoder is shared with the classification pipeline and consists of four convolutional layers (Conv1, Conv2, Conv3, and Conv4) and two fully connected layers. The decoder has the opposite configuration of the encoder, including two fully connected layers, four pairs of convolution and upsampling layers, and a convolutional output layer with a linear activation function.

3. EXPERIMENTS AND RESULTS

3.1. Experimental Setups

UDIAT dataset: The publicly available dataset used in this study is a B-mode US image dataset of breast cancer [12]. For this research, 128 images were selected and cropped to 512×512 pixels with the lesion at the center. Among these images, there are 45 malignant lesions and 83 benign lesions.

UTSW dataset: The in-house clinical dataset is a B-mode US breast image dataset[8]. In this study, we selected 258 images, including 178 benign lesions and 80 malignant lesions. The images were resampled to a resolution same as UDIAT dataset. We used a marker-controlled watershed segmentation [13] to create the tumor boundary.

To evaluate the performance and generalization ability of the SATPN, three schemes were designed to assess the SATPN within and across two datasets: (1) Classification on the single dataset: for each dataset, 80% of randomly selected samples from both benign and malignant lesions were used as the training set, while the remaining 20% were used as the test set. (2) Classification across two Datasets: The combined training set consisted of 80% of samples from UDIAT and 80% of samples from the UTSW dataset, while the remaining samples from each dataset were used as two separate test sets. The training data was randomly divided into 90% for training and 10% for validation by random assignment. In the experiments, the pixel grayscale values were normalized to the range of [0, 1]. In this paper, ACC, AUC, SEN, SPE, PPV, NPV, and MCC represent seven performance metrics [14].

3.2. Classification Results on Single Dataset

Table 1 present the classification results of the six methods on the UDIAT and UTSW datasets, respectively. Firstly, the paper compares the methods using different inputs (SCAE and BFM-SCAE, SSDL and BFM-SSDL, TPN and SATPN). It is found that the BFM-based methods outperform the methods based on the original images in terms of the seven performance metrics. This indicates that the BFM-based methods utilize the structural information in the BI-RADS-oriented feature maps to enhance the diagnostic accuracy. With the guidance of BFMs, the DL-based methods can focus on the breast lesion features specified by clinical requirements. Next, the paper compares different networks with the same input. The proposed TPN and SATPN outperform SCAE and SSDL, suggesting that they learn more effective features for lesion classification. SATPN utilizes BI-RADS-oriented feature maps that contain structural information. Additionally, SATPN integrates a multi-pathway semi-supervised network that independently extracts image features for benign and malignant lesions, thereby improving classification accuracy. By incorporating these design elements, SATPN achieves superior performance compared to other methods in the study.

3.3. Model Validation across Dataset

Table 2 and Table 3 summarize the classification results across datasets. Despite the different characteristics of the two datasets collected from different manufacturers' devices, the results obtained by all methods are similar to those shown

Table 1. Classification results (mean ± std %) for different methods when training and testing on UDIAT and UTSW dataset, respectively.

	Original Image-based DL			BFM-based DL		
	SCAE	SSDL	TPN	BFM-SCAE	BFM-SSDL	**SATPN**
UDIAT dataset						
ACC	87.50 ± 3.72	89.01 ± 6.31	92.31 ± 2.91	91.03 ± 2.87	94.23 ± 3.33	**96.71 ± 2.18**
AUC	82.87 ± 6.02	86.43 ± 6.40	88.50 ± 4.29	87.87 ± 5.09	90.80 ± 6.01	**95.42 ± 3.30**
MCC	69.98 ± 8.95	76.78 ± 11.96	82.07 ± 8.05	78.65 ± 7.34	85.80 ± 8.45	**92.60 ± 5.04**
SEN	70.42 ± 14.03	77.74 ± 11.87	77.72 ± 7.40	79.26 ± 12.30	82.99 ± 12.28	**91.88 ± 6.50**
SPE	95.33 ± 4.16	95.12 ± 7.32	**99.29 ± 1.75**	96.48 ± 4.07	98.61 ± 2.41	98.96 ± 1.80
PPV	86.96 ± 11.37	93.33 ± 8.46	97.14 ± 7.00	90.56 ± 9.51	96.88 ± 5.41	**98.21 ± 3.09**
NPV	88.13 ± 6.63	87.84 ± 7.24	90.36 ± 3.36	91.46 ± 5.73	93.71 ± 4.06	**96.23 ± 2.46**
UTSW dataset						
ACC	80.77 ± 4.03	81.09 ± 1.73	84.29 ± 4.76	82.21 ± 2.59	84.38 ± 3.11	**91.03 ± 4.53**
AUC	72.38 ± 4.79	71.78 ± 4.02	77.80 ± 5.29	71.55 ± 2.24	77.14 ± 6.61	**84.40 ± 8.07**
MCC	52.57 ± 7.01	48.95 ± 7.22	62.98 ± 9.37	53.10 ± 7.21	60.01 ± 8.96	**77.26 ± 10.89**
SEN	50.94 ± 12.61	51.09 ± 8.89	59.35 ± 10.89	47.15 ± 13.65	60.83 ± 15.37	**69.72 ± 17.08**
SPE	93.45 ± 5.16	92.47 ± 2.86	96.25 ± 2.48	95.94 ± 2.35	93.82 ± 3.99	**99.07 ± 1.31**
PPV	81.54 ± 9.95	72.11 ± 8.82	88.31 ± 7.40	83.65 ± 9.75	81.71 ± 8.87	**97.78 ± 3.14**
NPV	81.41 ± 5.06	83.19 ± 2.67	83.36 ± 6.03	82.15 ± 3.35	85.95 ± 3.57	**89.95 ± 5.50**

Table 2. Classification results (mean ± std %) for different methods when training on a combined UDIAT and UTSW dataset and testing on UDIAT dataset.

	Original Image-based DL			BFM-based DL		
	SCAE	SSDL	TPN	BFM-SCAE	BFM-SSDL	**SATPN**
ACC	85.58 ± 2.67	88.66 ± 3.22	92.86 ± 5.60	90.77 ± 1.88	92.95 ± 3.45	**96.15 ± 2.69**
AUC	82.53 ± 5.52	84.29 ± 4.12	90.40 ± 6.86	85.73 ± 3.48	88.39 ± 5.02	**97.22 ± 3.89**
MCC	67.13 ± 7.03	73.99 ± 6.79	83.89 ± 12.47	73.51 ± 6.67	81.41 ± 7.66	**91.62 ± 6.03**
SEN	74.01 ± 12.81	71.88 ± 9.16	83.28 ± 11.77	75.43 ± 7.62	79.56 ± 10.68	**87.00 ± 8.71**
SPE	91.06 ± 4.22	96.70 ± 3.75	97.53 ± 4.25	96.03 ± 2.02	97.22 ± 2.78	**98.44 ± 2.78**
PPV	81.88 ± 9.25	92.28 ± 8.29	94.78 ± 8.51	83.81 ± 10.58	93.39 ± 6.70	**97.89 ± 4.84**
NPV	87.57 ± 5.86	87.88 ± 4.22	92.45 ± 5.50	91.92 ± 4.35	93.30 ± 3.97	**94.06 ± 4.01**

in Table 1. This indicates that the DL-based methods exhibit good generalization when training models with multiple datasets in the context of this paper. Furthermore, it can be observed that the proposed SATPN performs the best in all comparisons. These results demonstrate that, through the comparison of different algorithms' performance, the SATPN method has certain advantages and can generalize to different datasets, avoiding overfitting to data from a single institution.

3.4. Effect of the Number of Training Samples

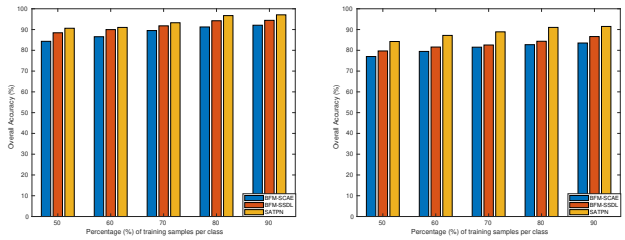
As shown in Fig. 2, this paper investigates the impact of different numbers of training samples on the proposed method (SATPN) on two datasets. BFM-SCAE and BFM-SSDL are used as baseline methods. The training sample sets are composed of randomly selected samples at different percentages (ranging from 50% to 90%) per class, while the remaining samples form the test set. As shown in Fig. 2, the performance of all tested classifiers generally improves with an increase in the number of training samples. For both datasets, SATPN achieves satisfactory results when trained with 80% of the samples. Overall, the proposed SATPN outperforms other methods in each case of different training samples and provides state-of-the-art results with limited training sample sizes, further confirming the feasibility of integrating structural information into breast cancer ultrasound CAD.

Table 3. Classification results (mean ± std %) for different methods when training on a combined UDIAT and UTSW dataset and testing on UTSW dataset.

	Original Image-based DL			BFM-based DL		
	SCAE	SSDL	TPN	BFM-SCAE	BFM-SSDL	**SATPN**
ACC	79.81 ± 3.67	81.15 ± 3.31	83.76 ± 4.17	81.54 ± 2.61	83.27 ± 4.79	**90.38 ± 2.16**
AUC	72.80 ± 5.55	71.60 ± 4.35	75.62 ± 5.89	73.94 ± 4.68	76.59 ± 5.22	**84.40 ± 4.37**
MCC	48.38 ± 7.43	51.87 ± 9.08	59.17 ± 9.17	56.24 ± 7.60	59.21 ± 10.65	**74.72 ± 6.43**
SEN	57.22 ± 15.10	48.03 ± 9.55	55.82 ± 13.53	51.88 ± 9.49	60.10 ± 11.31	**71.43 ± 9.55**
SPE	88.38 ± 4.28	95.17 ± 4.07	95.43 ± 3.69	96.01 ± 2.28	93.09 ± 5.84	**97.37 ± 2.26**
PPV	67.62 ± 6.85	81.45 ± 14.23	85.23 ± 9.17	86.01 ± 5.27	82.09 ± 10.02	**90.91 ± 7.22**
NPV	84.16 ± 3.07	81.39 ± 4.64	83.96 ± 5.06	80.35 ± 3.12	84.20 ± 4.22	**90.24 ± 2.52**

Table 4. The performance summary of breast ultrasound CAD system.

Ref.	Dataset (benign / malignant)	Availability	Features /Classifier	Performance (%) ACC	SEN	SPE
Singh et al.	88 / 90	Private	Manual feature / BPNN	95.86	95.14	96.58
Han et al.	4254/3154	Private	GoogLeNet	91.23	84.29	96.07
Zhang et al.	135 / 92	Private	Boltzmann machine	93.40	88.60	97.10
Cheng et al.	275 / 245	Private	SDAE	82.40	78.70	85.70
Shi et al.	100 / 100	Private	DPN / SVM	92.40	92.67	91.36
Antropova et al.	1978 / 415	Private	Manual feature / VGG 19	AUC = 90.00		
Prabusankarlal et al.	70 / 50	Public (GVHE)	Manual feature / SVM	95.83	96.00	95.71
Byra	48 / 52	Public (OASBUD)	VGG 19 / SVM	80.00	80.80	79.20
Byra et al.	110 / 53	Public (UDIAT)	VGG 19 / FLDA	84.00	85.51	83.40
Zhang et al.	83 / 45	Public (UDIAT)	BFM / SSDL	94.23	82.99	98.61
	178 / 80	Private	BFM / SSDL	84.38	60.83	93.82
Ours	**83 / 45**	**Public (UDIAT)**	**SATPN**	**96.71**	**91.88**	**98.96**
	178 / 80	**Private**	**SATPN**	**91.03**	**69.72**	**99.07**



(a) (b)

Fig. 2. Effect of the number of training samples on BIRADS-based methods for (a) UDIAT and (b) UTSW datasets.

4. DISCUSSION AND CONCLUSIONS

This paper proposes a novel SATPN method that leverages structural information for accurate lesion classification in breast ultrasound images. SATPN integrates clinical knowledge of lesion features and a triple-pathway network that independently extracts image features of benign and malignant lesions under the constraint of lesion classification task. The effectiveness of SATPN is validated on two breast ultrasound image datasets, demonstrating high diagnostic accuracy.

Table 4 compares the proposed SATPN method with state-of-the-art methods. Unlike traditional machine learning methods [15, 16, 17], SATPN automatically learns representative and discriminative features through a hierarchical deep neural network. Unlike previous deep learning approaches using pre-training or transfer learning, SATPN integrates conventional BI-RADS features into a multitask learning network to improve breast lesion classification performance.

In summary, SATPN outperforms other methods in all comparative scenarios, handling multiple datasets with a single model, making it an effective approach for breast ultrasound lesion CAD with limited data.

5. REFERENCES

[1] Dayangku Nur Faizah Pengiran Mohamad, Syamsiah Mashohor, Rozi Mahmud, Marsyita Hanafi, and Norafida Bahari, "Transition of traditional method to deep learning based computer-aided system for breast cancer using automated breast ultrasound system (abus) images: a review," *Artificial Intelligence Review*, pp. 1–30, 2023.

[2] Heang-Ping Chan, Lubomir M Hadjiiski, and Ravi K Samala, "Computer-aided diagnosis in the era of deep learning," *Medical physics*, vol. 47, no. 5, pp. e218–e227, 2020.

[3] Woo Kyung Moon, Yan-Wei Lee, Hao-Hsiang Ke, Su Hyun Lee, Chiun-Sheng Huang, and Ruey-Feng Chang, "Computer-aided diagnosis of breast ultrasound images using ensemble learning from convolutional neural networks," *Computer methods and programs in biomedicine*, vol. 190, pp. 105361, 2020.

[4] Christian Janiesch, Patrick Zschech, and Kai Heinrich, "Machine learning and deep learning," *Electronic Markets*, vol. 31, no. 3, pp. 685–695, 2021.

[5] Jie-Zhi Cheng, Dong Ni, Yi-Hong Chou, Jing Qin, Chui-Mei Tiu, Yeun-Chung Chang, Chiun-Sheng Huang, Dinggang Shen, and Chung-Ming Chen, "Computer-aided diagnosis with deep learning architecture: applications to breast lesions in us images and pulmonary nodules in ct scans," *Scientific reports*, vol. 6, no. 1, pp. 24454, 2016.

[6] Ke Yu, Xintao Wang, Chao Dong, Xiaoou Tang, and Chen Change Loy, "Path-restore: Learning network path selection for image restoration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 10, pp. 7078–7092, 2021.

[7] Kuen-Jang Tsai, Mei-Chun Chou, Hao-Ming Li, Shin-Tso Liu, Jung-Hsiu Hsu, Wei-Cheng Yeh, Chao-Ming Hung, Cheng-Yu Yeh, and Shaw-Hwa Hwang, "A high-performance deep neural network model for birads classification of screening mammography," *Sensors*, vol. 22, no. 3, pp. 1160, 2022.

[8] Erlei Zhang, Stephen Seiler, Mingli Chen, Weiguo Lu, and Xuejun Gu, "Birads features-oriented semi-supervised deep learning for breast ultrasound computer-aided diagnosis," *Physics in Medicine & Biology*, vol. 65, no. 12, pp. 125005, 2020.

[9] Claudio D Mello Jr, Lucas RV Messias, Paulo Lilles Jorge Drews-Jr, and Silvia SC Botelho, "Unsupervised learning method for encoder-decoder-based image restoration," in *Brazilian Conference on Intelligent Systems*. Springer, 2020, pp. 348–360.

[10] Illya Bakurov, Marco Buzzelli, Raimondo Schettini, Mauro Castelli, and Leonardo Vanneschi, "Structural similarity index (ssim) revisited: A data-driven approach," *Expert Systems with Applications*, vol. 189, pp. 116087, 2022.

[11] Nicolas Bustos, Manuel Tello, Guillermo Droppelmann, Nicolas Garcia, Felipe Feijoo, and Víctor Leiva, "Machine learning techniques as an efficient alternative diagnostic tool for covid-19 cases," *Signa Vitae*, vol. 18, no. 1, pp. 23–33, 2022.

[12] Wenju Cui, Yunsong Peng, Gang Yuan, Weiwei Cao, Yuzhu Cao, Zhengda Lu, Xinye Ni, Zhuangzhi Yan, and Jian Zheng, "Fmrnet: A fused network of multiple tumoral regions for breast tumor classification with ultrasound images," *Medical Physics*, vol. 49, no. 1, pp. 144–157, 2022.

[13] Huiling Xiang, Yao-Sian Huang, Chu-Hsuan Lee, Ting-Yin Chang Chien, Cheng-Kuang Lee, Lixian Liu, Anhua Li, Xi Lin, and Ruey-Feng Chang, "3-d res-capsnet convolutional neural network on automated breast ultrasound tumor diagnosis," *European Journal of Radiology*, vol. 138, pp. 109608, 2021.

[14] David MW Powers, "Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation," *arXiv preprint arXiv:2010.16061*, 2020.

[15] Bryar Shareef, Min Xian, and Aleksandar Vakanski, "Stan: Small tumor-aware network for breast ultrasound image segmentation," in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2020, pp. 1–5.

[16] Xiaofeng Qi, Fasheng Yi, Lei Zhang, Yao Chen, Yong Pi, Yuanyuan Chen, Jixiang Guo, Jianyong Wang, Quan Guo, Jilan Li, et al., "Computer-aided diagnosis of breast cancer in ultrasonography images by deep learning," *Neurocomputing*, vol. 472, pp. 152–165, 2022.

[17] Tara A Retson and Mohammad Eghtedari, "Computer-aided detection/diagnosis in breast imaging: a focus on the evolving fda regulations for using software as a medical device," *Current Radiology Reports*, vol. 8, pp. 1–7, 2020.