![The University of Edinburgh logo]

Edinburgh Research Explorer

A Deep DUAL-PATH Network for Improved Mammogram Image Processing

OPEN ACCESS

A DEEP DUAL-PATH NETWORK FOR IMPROVED MAMMOGRAM IMAGE PROCESSING

Heyi Li, Dongdong Chen, William H. Nailon, Mike E. Davies and Dave Laurenson

School of Engineering, The University of Edinburgh, Edinburgh, UK.

ABSTRACT

We present, for the first time, a novel deep neural network architecture called DUALCORENET with a dual-path connection between the input image and output class label for mammogram image processing. This architecture is built upon U-Net, which non-linearly maps the input data into a deep latent space. One path of the DUALCORENET, the locality preserving learner, is devoted to hierarchically extracting and exploiting intrinsic features of the input, while the other path, called the conditional graph learner, focuses on modeling the input-mask correlations. The learned mask is further used to improve classification results, and the two learning paths complement each other. By integrating the two learners our new architecture provides a simple but effective way to jointly learn the segmentation and predict the class label. Benefiting from the powerful expressive capacity of deep neural networks a more discriminative representation can be learned, in which both the semantics and structure are well preserved. Experimental results show that DUALCORENET achieves the best mammography segmentation and classification simultaneously, outperforming recent state-of-the-art models.

Index Terms— Mammography Segmentation, Mammography Classification, Dual-Path Network, Deep Learning

1. INTRODUCTION

According to the International Agency for Research on Cancer [1], breast cancer is the most frequently diagnosed cancer and the second most fatal disease among women around the world. Breast masses are the most challenging breast abnormality to diagnose and the most feared, since about 90% of breast masses are cancerous. Although there are no effective methods for prevention, early intervention is critical for improving associated survival rates. Screening mammography is thus employed for invasive malignant tumours (measuring < 2 cm) when they are too small to be palpable or cause symptoms. However, the manual inspection of screening mammograms by radiologists is tedious, subjective, and prone to errors, which may cause high false positive rates and over diagnosis [2–4]. For these reasons, automatic and robust diagnosis tools for screening mammography are in high demand. Traditional mammographic computer-aided diagnosis (CAD) systems rely heavily on elaborate hand-crafted fea-
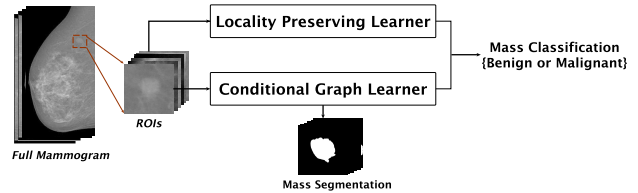

Fig. 1: The illustration of dual-path learning for mammography analysis.

tures [2]. Recently, leveraging the insights from the successes of deep neural networks (deep learning) in computer vision tasks [5–9], deep learning based algorithms have been applied to mammograms and have achieved state-of-the-art results for mass detection, segmentation and classification. Mass detection aims to find the regions of interest (ROIs) where abnormalities may be located, and mass segmentation provides detailed morphological features with precise outlines within ROIs. Mass classification categorizes mammogram patches or full mammograms into benign or malignant, assisting radiologists to interpret screenings locally (region-based mass classification) or globally (full mammogram classification). There are several recent examples which deep learning has been applied to full mammogram classification [5, 7, 8]. However, these works ignore the importance of local ROIs, since radiologists are more likely to identify local regions rather than labelling the full mammograms in clinical practice. In terms of region-based mass classification, Convolutional Neural Networks (CNNs) have been applied as a classifier showing that deep learning models outperform traditional methods significantly [3, 10]. The work in [11] utilized the CNN as a regressor to extract hand-crafted features from both mammogram ROIs and annotation masks of radiologists, which were then followed by a Random Forest classifier model. Although [11] results in better performance than direct CNN based algorithms, hand-crafted features are still required in the final design.

More recently, [12] proposed a powerful CNN based image segmentation model referred as U-Net, which interlaces multi-resolution information by adding skips between encoding and decoding layers of the same spatial size and has been shown to perform well on medical images. However, when applied to mammography, the U-Net is limited by low-signal-to-noise data, which causes labelling inconsistency and incompleteness. To address this, a novel architecture
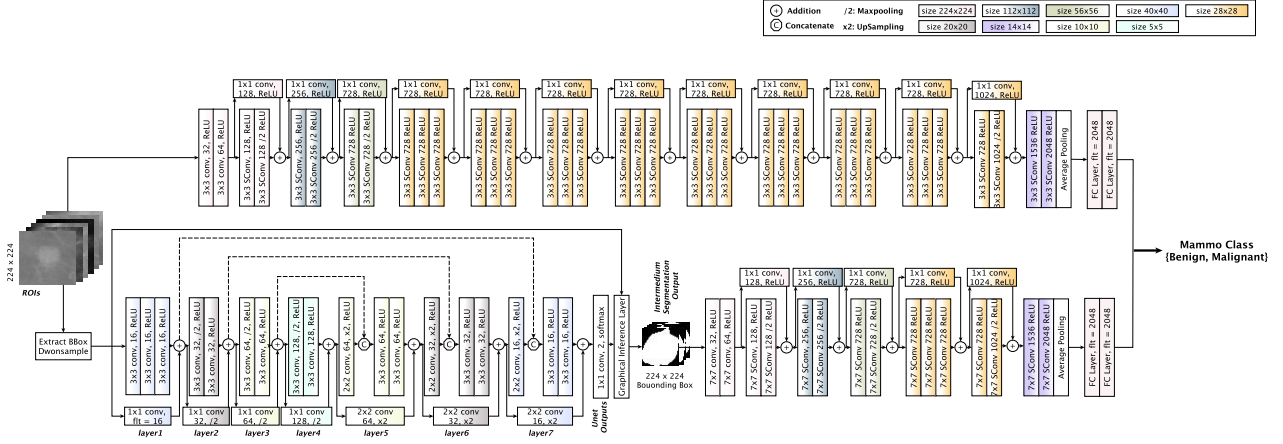
Fig. 2: The proposed **Dual**-path **Co**nditional **Re**sidual **Net**work (DUALCORENET).

is designed in this paper. Based on the accurate pixel-level labelling algorithm presented in [4] and the related breast mass classifiers [2, 3, 5, 7, 8, 10, 11], a **Dual**-path **Co**nditional **Re**sidual **Net**work (DUALCORENET) for mammography analysis is proposed as shown in Figure 1. Firstly, a mass and its context texture learner called the *Locality Preserving Learner (LPL)* is built with stacks of convolutional blocks, achieving a mapping from ROIs to class labels. Secondly, a graph inference layer, called the *Conditional Graph Learner (CGL)* is employed to learn the input-mask correlation, and the extracted segmentation features will be further used to improve the final mass classification performance. By integrating these two learning paths, DUALCORENET achieves the best mammography segmentation and classification simultaneously, outperforms recent state-of-the-art models. The main contributions of this paper are the following: (i) To our knowledge, DUALCORENET is the first dual-path CNN-based mammogram analysis model that takes advantage of segmented mask for mass classification; (ii) Our model achieves state-of-the-art results on both mass segmentation and mass classification tasks on publicly available mammography datasets.

2. METHODOLOGY

The DUALCORENET takes mammogram ROIs as input and outputs mass segmentation masks and ultimately classifies ROIs simultaneously. The pipeline of DUALCORENET is shown in Fig 2. We first define the notation used throughout the paper and then introduce the details of the two proposed learning paths and the whole DUALCORENET.

2.1. Notations

Given a mammogram lattice Ω, let $\boldsymbol{x} : \Omega \rightarrow \mathbb{R}$ as one of the N training mammogram ROIs, $\boldsymbol{y}_{seg} : \Omega \rightarrow \{0 :$ Normal pixels, $1 :$ Mass pixels$\}$ as the corresponding pixel-

level annotations from radiologists and $\boldsymbol{y}_{cls} \rightarrow \{0 :$ Benign, $1 :$ Malignant$\}$ for the ROI class labels, the training set is represented by $\mathcal{D} = \{(\boldsymbol{x}^{(n)}, \boldsymbol{y}_{seg}^{(n)}, \boldsymbol{y}_{cls}^{(n)})\}_{n=1}^{N}$.

2.2. Locality Preserving Learner

Mass and its context tissue textures are exploited as the major classification features in traditional mammographic CADs [2]. The LPL learns the hierarchical and local intrinsic texture features by non-linear mapping locality information's linear combination with stacks of depth-wise separable convolution, maxpooling and ReLU activation.

In order to prevent the model from exploding and vanishing gradients, the residual learning [13] is employed within the whole DUALCORENET, which maps the convolutional layer's output $\mathcal{H}(\boldsymbol{x})$ with residuals $\mathcal{F}(\boldsymbol{x})$ and the linear dimension matching kernel W as:

$$\mathcal{F}(\boldsymbol{x}) := \mathcal{H}(\boldsymbol{x}) - W * \boldsymbol{x} \tag{1}$$

Depth-wise separable convolutions, referred as "SConv" in Fig 2, are widely applied to efficiently map the network by factoring filters into a series of operations, such as [14]. Specifically, depth-wise separable convolution comprises of a spatial convolution that operates independently for each feature map and a pointwise convolution that computes a linear projection for the spatial convolution's output. Inspired by this, residual learning paths with the depth-wise separable convolutions are built in the DUALCORENETto independently learn cross-channel and spatial correlations of hierarchical but local intrinsic features from both mammogram ROIs and segmentation masks. The loss associated to the LPL layer is defined with categorical cross-entropy as:

$$\ell_{lpl} = -\sum_{n=1}^{N} \log P\big(y_{cls}^{(n)} \mid \boldsymbol{x}^{(n)}; \boldsymbol{\theta}_{lpl}\big) \tag{2}$$

where y_{cls} is the class indicator, $\boldsymbol{\theta}_{lpl}$ is the parameter set of the LPL path.

2.3. Conditional Graph Learner

In practice, radiologists usually analysis mass shape and boundary information to improve cancer inspections, the more irregular the shape, the more likely to be cancerous [15]. This observation suggests that extracting segmentation-friendly features should improve mass classification performance. To do that, the CGL is proposed accordingly. Specifically, the CGL first employs the revised U-Net, mapping ROIs to geometry and spatial features, then a graphical inference layer to preserve pixel consistency with conditional restrictions and finally the separable convolution stacks to efficiently learn hierarchical features and class labels mapping.

Firstly, the U-Net makes use of standard convolutional layers and employs a combination of multi-resolution filters. Incorporated with residual learning in the CGL, the mth layer output with input $\boldsymbol{x}^{(n)}$ at pixel (i, j) is formulated as:

$$
\begin{aligned}
y_{i,j}^{(n,m)} &= \sigma\big(h_{ks}(\{\boldsymbol{x}_{s_i+\delta_i, s_j+\delta_j}\}_{0 \leq \delta_i, \delta_j \leq k})\big) \\
&+ \sigma\big(h'_{k's'}(\{\boldsymbol{x}_{s'_i+\delta_i, s'_j+\delta_j}\}_{0 \leq \delta_i, \delta_j \leq k'})\big)
\end{aligned}
\tag{3}
$$

where h_{ks} and $h'_{k's'}$ denote the convolution functions of residuals and dimension mapping respectively, k and k' represent the corresponding kernel sizes of convolutional filters, s and s' for kernel strides or maxpooling factors in the downsampling layers, and σ for activation functions.

The graphical inference layer applies conditional random fields (CRFs) as an additional CNN layer [4, 16]. The loss function for the segmentation is defined with weighted categorical cross-entropy of the residual learning U-Net and graphical inference layer as follows

$$
\begin{aligned}
\ell_{seg} &= (1-\lambda)\Big(-\sum_{i,j} \log P\big(y_{i,j}^{(n)} \mid \boldsymbol{x}^{(n)}; \boldsymbol{\theta}_{seg}\big)\Big) \\
&+ \lambda \cdot \Big(A(\boldsymbol{x}^{(n)}) - \exp\big(\sum_{i,j \in V} P\big(y_{i,j}^{(n)} \mid \boldsymbol{x}^{(n)}; \boldsymbol{\theta}_{seg}\big) \\
&+ \sum_{p,q \in E} \phi\big(y_p^{(n)}, y_q^{(n)} \mid \boldsymbol{x}^{(n)}\big)\Big)
\end{aligned}
\tag{4}
$$

where $P(y_{i,j}^{(n)})$ is the residual U-Net output probability distribution at position (i, j) given the parameters $\boldsymbol{\theta}_{seg}$, A is the partition function of the CRF, ϕ is the pair-wise potential function which is defined with the label compatibility, Gaussian kernels and corresponding weights for pixel p and q belonged to the CRF graph edges E, and λ[1] is the trade-off factor.

The loss of the CGL path is again defined by the classification categorical cross-entropy as:

$$
\ell_{cgl} = -\sum_{n=1}^{N} \log P\big(y_{cls}{}^{(n)} \mid P(\boldsymbol{y}_{seg}^{(n)}); \boldsymbol{\theta}_{cgl}\big)
\tag{5}
$$

---

[1] Optimally chosen as 0.67 with grid search in our experiments.

where $P(\boldsymbol{y}_{seg}^{(n)})$ is the resulted soft mass mask and $\boldsymbol{\theta}_{cgt}$ is the parameter set of the CGL path.

Finally, by integrating CGL and LPL into a dual-path network model, the overall categorical cross-entropy based loss of the DUALCORENET is defined as:

$$
\ell = -\sum_{n=1}^{N} \log P\big(y_{cls}{}^{(n)} \mid \boldsymbol{x}^{(n)}; \boldsymbol{\theta}_{cgl}, \boldsymbol{\theta}_{lpl}\big)
\tag{6}
$$

3. EXPERIMENTS

3.1. Datasets and ROIs selection

The DUALCORENET is evaluated on two public datasets: INbreast [17] and CBIS-DDSM [18]. INBreast is a full-field digital mammographic (FFDM) dataset containing 116 pixel-level annotated masses in 107 mammograms, with pixel size $70\mu m$ and contrast resolution 14 bits. The CBIS-DDSM is a modernized subset of Digital Database for Screening Mammography (DDSM) [19], that includes 1594 mass contained digitized film screening mammograms.

Two manual ROIs selection procedures are adopted for the training stage in this work: One is to locate and extract rectangular mass contained bounding boxes, and they are utilized by the CGL to explore the boundary and shape features; the other ROI set is selected with proportional padding that mass-centred ROIs include regions 1.6 times the size of the mass bounding box, and they are utilized by the LPL for mass and its context tissues texture feature learning. The selected ROIs are then augmented with horizontal and vertical flips. In terms of data division, the INbreast data set is divided by patients into a training set and a test set as 80%: 20%. As for the CBIS-DDSM, we adopted the pre-divided 1253 training and 354 test ROIs within the data.

3.2. Implementation Details

As shown in Fig 2, the LPL path is designed with stacks of separable convolutional layers with kernel size 3×3 for fine texture features learning. The first three blocks are equipped with increasing feature maps (128, 256, 728) and decreasing spatial size (224×224, 112×112, 56×56), while the consecutive eight blocks are of the same feature map (728) and the same spatial shape (28×28).

The CGL path takes downsampled bounding box ROIs (with spatial size 40×40) as the input and followed by four consecutive maxpooling and convolution blocks with feature map numbers 16, 32, 64 and 128 and corresponding spatial size 40×40, 20×20, 10×10 and 5×5. The bottom latent features are then upsampled reversely with convolution operations and skips that connect previous layer channels. The multi-resolution mapping is then followed by the CRF graphical inference layer, resulting in a sigmoid activated upsampled segmentation mask (size 224×224). A relative large

Table 1: Mass segmentation performance (within Bounding Box ROIs) of the DUALCORENET and several state-of-the-art methods on test datasets.

Methodology	Dataset	DI, %
Dhungel *et. al.* [15]	INbreast	90
Zhu *et. al* [21]	INbreast	89.36
DUALCORENET	INbreast	**93.66**
Dhungel *et. al.* [15]	CBIS-DDSM	90
Zhu *et. al* [21]	CBIS-DDSM	90.62
DUALCORENET	CBIS-DDSM	**91.43**

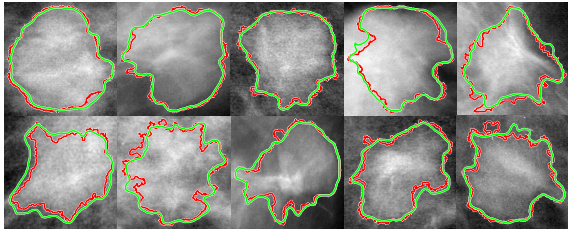**Fig. 3**: Visualized comparison between radiologists' annotation (red lines) and DUALCORENET segmentation results (green lines).

Table 2: Region based mass classification performance (AUC) of the proposed dual paths model and relative state-of-the art methods on test sets.

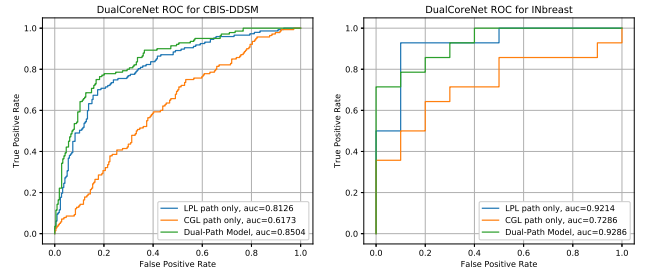| Methodology	CBIS-DDSM	INbreast
Dhungel *et. al.* [11]	-	0.91
Varela *et. al.* [2]	0.83	-
DUALCORENET (CGL only)	0.62	0.73
DUALCORENET (LPL only)	0.81	0.92
DUALCORENET (Dual-path)	**0.85**	**0.93**

Fig. 4: Mass classification ROC curves of the DUAL-CORENET experimented on INbreast (left, 85 training and 22 test mammograms) and CBIS-DDSM (right, 1253 training and 354 test mammograms).

kernel 7×7 is chosen for the appended CNN blocks, aiming to better learn the shape and boundary features.

After global averaging on both paths and two fully connected layers with 2048 neurons separately, the concatenated features of dual paths are classified into either one of the two mass categories (Benign or Malignant). Except for the separable convolutional blocks in the LPL path, which are initialized with ImageNet [20] pretrained weights for an accelerate convergence rate and network generalization, all other network layers were randomly set. To avoid over-fitting, dropout layers with 50% dropout rate were used. The DUAL-CORENET is optimized by the Stochastic Gradient Descent algorithm with the Adam update rule. To fully train the DU-ALCORENET, two component paths were first trained with their corresponding loss function (2) and (5) separately, and then further trained jointly with the loss (6).

3.3. Results

In terms of mass segmentation, evaluations of Dice Coefficients are compared with related works in Table 1. DUAL-CORENET achieves the best performance with 93.66% and 91.43% on INbreast and CBIS-DDSM, respectively. Visualized overlapping contours of the radiologist's annotation (red lines) and DUALCORENET's segmentation results (green lines) are shown in Figure 3. The DUALCORENET contributes similar to the radiologists and smooth contours even for relative hard cases.

As for the mass classification performance, the ROC curve and AUC of LPL, CGL and joined paths' best model

are shown in Figure 4. For INbreast, the dual path model achieved the best performance 0.93 AUC, 0.01 higher than LPL path only. With a guaranteed classification performance, the dual path model further contributes precise mass segmentation. For the CBIS-DDSM, the dual-path model outperforms either CGL or LPL, achieving 0.85 AUC. Moreover, as shown in Table 2, the DUALCORENET outperforms other state-of-art algorithms. The experimental results show that the representation learned by DUALCORENET is more robust for both of breast mass segmentation and classification in mammography, which again demonstrate our motivation of dual-path learning.

4. CONCLUSIONS

In this paper, we proposed a novel DUALCORENET for improved mammogram image analysis. By integrating the conditional graph learner path and the locality preserving learner path, our DUALCORENET works in a simple but effective way to jointly learn segmentation and classification. Thanks to the departure from hand-crafted features in classical CADs, our model is more flexible to learn mammographical-friendly representations. Furthermore, DUALCORENET performs better on higher quality dataset (the INbreast dataset). Extensive experiments show that our method outperforms state-of-the-art on both breast mass segmentation and classification tasks in mammography.

5. REFERENCES

[1] P. Boyle, B. Levin *et al.*, *World cancer report 2008.* IARC Press, International Agency for Research on Cancer, 2008.

[2] C. Varela, S. Timp, and N. Karssemeijer, "Use of border information in the classification of mammographic masses," *Physics in medicine & biology*, vol. 51, no. 2, p. 425, 2006.

[3] J. Arevalo, F. A. González, R. Ramos-Pollán, J. L. Oliveira, and M. A. G. Lopez, "Representation learning for mammography mass lesion classification with convolutional neural networks," *Computer methods and programs in biomedicine*, vol. 127, pp. 248–257, 2016.

[4] H. Li, D. Chen, W. H. Nailon, M. E. Davies, and D. Laurenson, "Improved breast mass segmentation in mammograms with conditional residual u-net," in *Image Analysis for Moving Organ, Breast, and Thoracic Images.* Springer, 2018, pp. 81–89.

[5] W. Zhu, Q. Lou, Y. S. Vang, and X. Xie, "Deep multi-instance networks with sparse label assignment for whole mammogram classification," in *International Conference on Medical Image Computing and Computer-Assisted Intervention.* Springer, 2017, pp. 603–611.

[6] D. Chen, J. Lv, and Z. Yi, "Graph regularized restricted boltzmann machine," *IEEE transactions on neural networks and learning systems*, vol. 29, no. 6, pp. 2651–2659, 2018.

[7] G. Carneiro, J. Nascimento, and A. P. Bradley, "Automated analysis of unregistered multi-view mammograms with deep learning," *IEEE transactions on medical imaging*, vol. 36, no. 11, pp. 2355–2365, 2017.

[8] S. Shams, R. Platania, J. Zhang, J. Kim, and S.-J. Park, "Deep generative breast cancer screening and diagnosis," in *International Conference on Medical Image Computing and Computer-Assisted Intervention.* Springer, 2018, pp. 859–867.

[9] D. Chen, J. Lv, and Z. Yi, "Unsupervised multi-manifold clustering by learning deep representation," in *Workshops at the 31th AAAI conference on artificial intelligence (AAAI)*, 2017, pp. 385–391.

[10] T. Kooi, B. van Ginneken, N. Karssemeijer, and A. den Heeten, "Discriminating solitary cysts from soft tissue lesions in mammography using a pretrained deep convolutional neural network," *Medical physics*, vol. 44, no. 3, pp. 1017–1027, 2017.

[11] N. Dhungel, G. Carneiro, and A. P. Bradley, "The automated learning of deep features for breast mass classification from mammograms," in *International Conference on Medical Image Computing and Computer-Assisted Intervention.* Springer, 2016, pp. 106–114.

[12] O. Ronneberger, P. Fischer, and 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.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[14] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," *arXiv preprint*, pp. 1610–02 357, 2017.

[15] N. Dhungel, G. Carneiro, and A. P. Bradley, "Deep learning and structured prediction for the segmentation of mass in mammograms," in *International Conference on Medical Image Computing and Computer-Assisted Intervention.* Springer, 2015, pp. 605–612.

[16] S. Zheng, S. Jayasumana, B. Romera-Paredes, V. Vineet, Z. Su, D. Du, C. Huang, and P. H. Torr, "Conditional random fields as recurrent neural networks," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1529–1537.

[17] I. C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M. J. Cardoso, and J. S. Cardoso, "Inbreast: toward a full-field digital mammographic database," *Academic radiology*, vol. 19, no. 2, pp. 236–248, 2012.

[18] R. S. Lee, F. Gimenez, A. Hoogi, and D. Rubin, "Curated breast imaging subset of ddsm," *The Cancer Imaging Archive*, 2016.

[19] M. Heath, K. Bowyer, D. Kopans, R. Moore, and W. P. Kegelmeyer, "The digital database for screening mammography," in *Proceedings of the 5th international workshop on digital mammography.* Medical Physics Publishing, 2000, pp. 212–218.

[20] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein *et al.*, "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.

[21] W. Zhu, X. Xiang, T. D. Tran, G. D. Hager, and X. Xie, "Adversarial deep structured nets for mass segmentation from mammograms," in *Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium on.* IEEE, 2018, pp. 847–850.