Breast Cancer Image Classification via Multi-level Dual-network Features and Sparse Multi-Relation Regularized Learning

Yongjun Wang, Fanglin Huang, Tianfu Wang, Baiying Lei

Abstract—**Breast cancer is one of the leading causes of cancer death worldwide. Recently, the computer-aided diagnosis and detection technique has been developed for the early diagnosis of breast cancer, but the diagnostic efficiency has still been a challenging issue. For this reason, we aim to improve the breast cancer diagnostic accuracy and reduce the workload of doctors in this paper by devising a deep learning framework based on histological image. Specifically, we develop a model of multi-level feature of dual-network combined with sparse multi-relation regularized learning method, which enhances the classification performance and robustness. Specifically, first, we preprocess the histological images using scale transformation and color enhancement methods. Second, the multi-level image features are extracted from preprocessed images using two deep convolutional neural networks (e.g., InceptionV3-ML (multi-level feature InceptionV3 network) and ResNet-50). Third, the feature selection method via sparse multi-relation regularization is further developed for performance boosting and overfitting reduction. We evaluate the proposed method based on the public ICIAR 2018 Challenge dataset of breast cancer histology images. Experimental results show that our method has achieved promising performance and outperformed the related works.**

I. Introduction

Breast cancer is one of the leading cancer-related death causes worldwide in women [1]. Early diagnosis of breast cancer is usually achieved by the pathologist observing histopathology. However, the diagnostic performance relies on the doctors' professional skills and experience, which are quite subjective and time-consuming. To reduce the workload burden and improve diagnostic efficiency, computer-aided diagnosis (CAD) is an effective way. The unsatisfactory CAD system performance always remains an open challenge. To address this issue, many researchers have applied deep learning methods to improve the early breast cancer diagnosis performance [2, 3]. For example, Bayramoglu *et al.* [4] proposed two different architectures: single task convolutional neural network (CNN) to predict malignancy level and multi-task CNN to simultaneously predict both malignancy and image magnification levels. Haarburger *et al.* [5] used transfer learning of pretrained model on natural images to classify breast lesions from dynamic contrast-enhanced magnetic resonance images. Han *et al.*[6] proposed a breast cancer multi-classification method using the structured deep learning model. They have achieved satisfactory results in breast cancer image classification. However, only single-level image features extracted by deep convolutional neural network or shallow convolutional neural network were used for breast cancer classification, which fail to utilize the multi-level combination information. Also, the feature selection is unconsidered for feature reduction. This leads to the inability to obtain informative feature representation for classification performance boosting.

To address it, it is known that the ResNet-50 model [7] has a deeper network structure, which enhances learning ability and solves the problem of "degeneracy" in the optimization process. Also, InceptionV3 is another DCNN structure [8], which has shallow network structure compared with ResNet-50, and has "wider" network structure. Meanwhile, InceptionV3 network based on InceptionV2 decomposes the two-dimensional convolution (e.g., 7×7 or 3×3) into two one-dimensional convolutions (e.g., 1×7 and 7×1 or 1×3 and 3×1), which increases the nonlinearity of the network and selects the informative feature freely. Also, the max pooling can remove the redundant information convoluted from the upper layer, which can achieve better classification performance. For this reason, we build a dual network based on ResNet-50 and InceptionV3-ML (multi-level feature InceptionV3 network) to exploit the feature complementarity among features of different layers, which can enhance the breast cancer classification.

However, the multi-level features of dual-network often lead to high-dimensional data with redundant features, which could lead to the overfitting issue and affect accuracy. To address these issues, it is known that the feature selection is an effective way. For example, the typical methods include principal component analysis (PCA) [9], and locally linear embedding (LLE) [10]. The limitations of these methods are that they require higher hardware costs and poor classification accuracy. For this reason, we propose the sparse multi-relational regularization (SMRR) feature selection algorithm based on multi-relational information (i.e., feature-feature relation, response-response relation and sample-sample relation) [11]. By imposing these three relational characteristics along with the $\ell_{2,1}$ norm on the weight coefficients, the better classification results can be expected by obtaining the sparsity weight correlation vectors.

Up to now, our method mainly includes three steps. The first step is the image preprocessing, we down-sample the original breast cancer images to the half, and then perform color enhancement and crop randomly. The second step is image feature extraction. We import the extracted crops into two DCNNs (e.g., ResNet-50 and InceptionV3-ML) to extract the multi-level features. The third step is feature selection via the sparse multi-relational regularization to select the most informative and representative features, and then train support vector machine (SVM) classifier via the dual-network voting strategy. In our experiments, we use the public dataset of the ICIAR 2018 grand challenge on breast

* This work was supported by National Key R&D Program of China (No. 2016YFC0104700)

Yongjun Wang, Fanglin Huang, Tianfu Wang, Baiying Lei are with the National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, School of Biomedical Engineering, Health Science Center, Shenzhen University, Shenzhen, China, 518060.
Correspondence should be addressed to Baiying Lei (leiby@szu.edu.cn)

Manuscript 498 submitted to 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Received February 19, 2019.

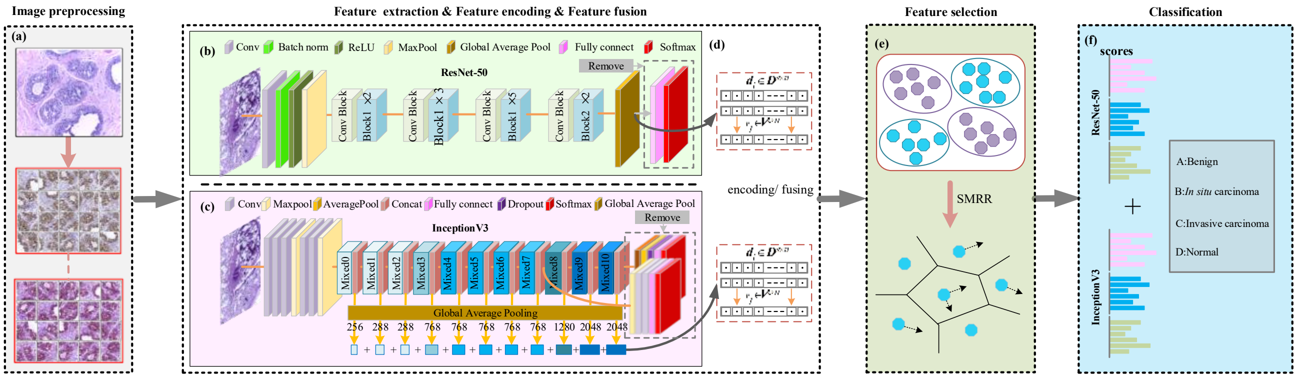

Fig. 1. Breast cancer pathological image classification framework: (a) Image preprocessing, (b) and (c) Feature extraction used ResNet-50 network and InceptionV3-ML, (d) Feature encoding, (e) Feature selection, and (f) Classification utilized SVM and dual-network voting strategy.

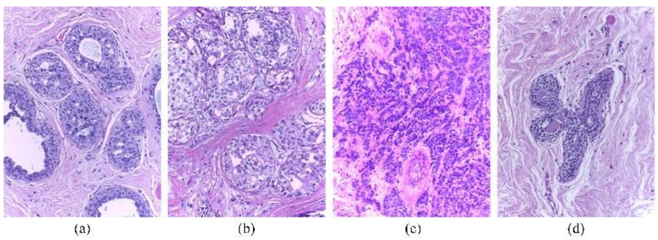

Fig. 2. Examples of microscopic breast cancer histology images: (b) benign, (c) in situ carcinoma, (d) invasive carcinoma, (a) normal.

cancer histology images. The main contributions of this paper are as below: 1) We devise the multi-level features via InceptionV3-ML to consider the complementarity among features in different layers. 2) We develop a feature selection method using sparse multi-relationship regularization and using a dual-network voting strategy, which further improve the breast cancer image classification performance.

II. Methodology

In this section, we describe the classification model based on multi-level feature fusion and sparse multi-relation regularized learning, as illustrated in Fig 1. First, we briefly introduce images preprocessing steps. Second, we describe how to extract the multi-level features using DCNNs, and then feature encoding. Third, we utilize SMRR feature selection method to select some representative features, and train SVM classification model by dual-network voting strategy.

A. Dataset and Images preprocessing

The public dataset of the ICIAR 2018 grand challenge includes 400 H&E stained images of breast histology microscopy (2048×1536 pixels). All the images are digitized with the same acquisition conditions, with a magnification of 200× and pixel size of 0.42μm ×0.42μm. Each image is labeled with one of the four balanced classes: benign, in situ carcinoma, invasive carcinoma, and normal. Each class has 100 images and a typical image of each predominant cancer type is shown in Fig. 2. Also, the image-wise annotation of ICIAR 2018 dataset is performed by two expert histopathologists.

In the image preprocessing step, we firstly downscale the original image size to the half (1024×768) as input images, as illustrated in Fig. 1 (a). This allows us to choose the smaller crop size without leading to the loss of important feature information. Then, we perform color space conversion [12] to change the weight of the RGB color component of the image to obtain 50 color-enhanced images of each input image, and their features are fused to resolve the uneven staining of pathological images. Accordingly, we can accurately learn the color information of the images as features through two DCNNs. Finally, we randomly extract 20 crops of 400×400 pixels or 650×650 pixels from each color-enhanced image to increase the amount of training data.

B. Feature extraction, encoding and fusion

In this paper, we propose a model to extract and encode the multi-level features from the histopathological images of breast cancer classification, as shown in Fig.1 (b)~(d). At the feature extraction stage, we input the image crops into ResNet-50 network and InceptionV3-ML, and drop the fully connected layers. For ResNet-50, we convert the last convolutional layer consisting of 2048 channels via global average pooling into one-dimensional feature vector with a length of 2048, shown in Fig.1 (b). For InceptionV3-ML, we apply the global average pooling operation to the connection layer of the eleven internal modules (e.g., mixed0 ~ mixed10) of InceptionV3, and then we concatenate them into one vector with a length of 10048, shown in Fig.1 (c). The purpose of our design is that InceptionV3-ML contains more superficial feature information such as shape, texture, color, boundary and other shallow features. Meanwhile, ResNet-50 endows with the deep network structure to obtain the abstract semantic features.

Therefore, we obtain feature vectors of 20 crops for each color-enhanced image via ResNet-50 or InceptionV3-ML networks. The next stage is feature encoding and fusion. Firstly, feature vector defined as $\boldsymbol{d}_i = \{d_{i,1}, d_{i,2}, \dots \dots d_{i,n}\} \in \boldsymbol{D}^{s\times n} (s = 20)$ of each crop for color-enhanced image is encoded through 3-norm pooling method [13] into a single feature vector. The 3-norm pooling method is defined as:

$$\boldsymbol{v}_j = \left(\frac{1}{l}\sum_i^l (\boldsymbol{d}_i)^p\right)^{\frac{1}{p}}, \tag{1}$$

where p is the hyper-parameter and set to 3 as suggested, $\boldsymbol{v}_j = \{v_{j,1}, v_{j,2}, \dots \dots v_{j,n}\} \in \boldsymbol{V}^{c\times n}$ is pooled feature vector of each color-enhanced image, c is the number of total color-enhanced images (c = 50), s is the number of random crops (s =20), and n is the feature dimension.

Manuscript 498 submitted to 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Received February 19, 2019.

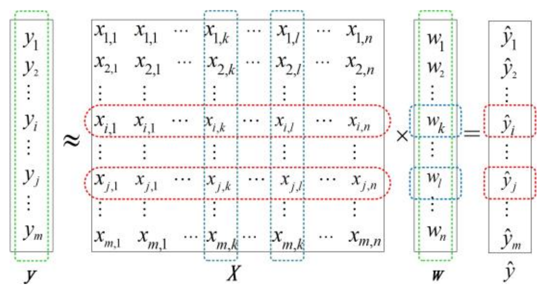

Fig. 3. An illustration of the relational information that can be obtained from the observations: the red solid rectangles, the blue dash rectangles, and the green dotted rectangles denote, respectively, the 'sample-sample' relation, 'feature-feature' relation and 'response-response' relation.

The last stage is feature fusion. We obtain 50 coded feature vectors that represent each input breast cancer histopathological image. Afterwards, we directly calculate the mean value of the coding features of 50 color-enhanced images to obtain the feature vector of each input image. We define $\boldsymbol{x}_k = \{x_{k,1}, x_{k,2}, \dots \dots x_{k,n}\} \in \boldsymbol{X}^{m\times n}$ as a feature vector of each input image, the hyper-parameter m = 400 is the number of total samples. The mean value is defined as

$$\boldsymbol{x}_k = \frac{1}{c}\sum_j^c \boldsymbol{v}_j, \tag{2}$$

This method not only adds the color information to the original image feature information, but also resolves the uneven staining of pathological images to obtain more effective feature information.

C. Feature selection and SVM classification

Since multi-level features can cause feature redundancy and overfitting, which can affect classification accuracy, as illustrated in Fig. 1(e). To address it, we propose a feature selection method called SMRR, which is based on three kinds of relationships and $\ell_{2,1}$-norm to obtain important feature information for classification performance boosting. Let we define $\boldsymbol{y}$ as the vector of image true labels, $\hat{\boldsymbol{y}}$ as the vector of image predicted labels, $\boldsymbol{X}$ as the feature matrix and $\boldsymbol{w}^{n\times 1}$as the weight coefficient vector. Our goal is to optimize the following objective function using $\ell_{2,1}$-norm to get a weight coefficient vector $\boldsymbol{w}$, which is defined as

$$\min_{\boldsymbol{w}'} ||\boldsymbol{y} - \boldsymbol{X}\boldsymbol{w}||_{2,1}^2, \tag{3}$$

To enhance the performance, we use three kinds of relationships (e.g., feature-feature relation $R(\boldsymbol{w_1})$, response-response relation $R(\boldsymbol{w_2})$ and sample-sample relation $R(\boldsymbol{w_3})$) [11]. The multi relationship is shown in Fig. 3 and represented as

$$R(\boldsymbol{w}) = \alpha_1 R(\boldsymbol{w_1}) + \alpha_2 R(\boldsymbol{w_2}) + \alpha_3 R(\boldsymbol{w_3}), \tag{4}$$

Accordingly, the objective function is reformulated as

$$\min_{\boldsymbol{w}'} \left|\left|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{w}\right|\right|_{2,1}^2 + R(\boldsymbol{w}), \tag{5}$$

where α_1, α_2 and α_3 denote controlling parameters of the three relationships, respectively. To this end, we further improve relevance of three kinds of relationships, which can be implemented by a $\ell_{2,1}$-norm regularization term on $\boldsymbol{w}$, *i.e.*, $||\boldsymbol{w}||_{2,1}$. Then we can rewrite the objective function as

$$\min_{\boldsymbol{w}'} \left|\left|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{w}\right|\right|_{2,1}^2 + R(\boldsymbol{w}) + \lambda||\boldsymbol{w}||_{2,1}, \tag{6}$$

where λ denotes controlling parameter of the respective regularization term, respectively. The least squares method is used to solve our objective function by taking the Lagrangian of the form:

$$L(w, y) = \min_{\boldsymbol{w}'} \left|\left|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{w}\right|\right|_{2,1}^2 + R(\boldsymbol{w}) + \lambda||\boldsymbol{w}||_{2,1}, \tag{7}$$

Finally, we can obtain $\boldsymbol{w}$ of the weight coefficient vector. The selected most relevant features are trained by SVM classifier of dual-network voting strategy to achieve breast cancer image classification. We utilize ResNet-50 or InceptionV3-ML networks' features to train SVM classifier, and then we can obtain the classification scores of four types of breast cancer (e.g., benign, in situ carcinoma, invasive carcinoma and normal), as illustrated in Fig. 1 (f). The scores weighted sum of ResNet-50 or InceptionV3-ML networks is calculated via the dual-network voting to obtain the final classification scores, which will be used to judge the classification results.

III. Experimental Results

A. Experimental setup

In our experiments, firstly, we evaluate InceptionV3-ML network and two baseline networks (e.g., ResNet-50 and InceptionV3) used two different crop sizes, and fusing features of different crop sizes. Secondly, we verify the effectiveness of our proposed SMRR feature selection algorithm. Finally, we compare SMRR with the other two feature selection methods such as PCA [9] and LLE [10] via dual-network voting strategy between ResNet-50 and InceptionV3-ML networks.

We adopt the nested 10-fold cross-validation to evaluate the performances of the competitive methods. We utilize various metrics to evaluate the classification performance using accuracy (ACC), Recall Rate and area under receive of characteristics (AUC).

B. Baseline network experiment results

Generally, we evaluate two different baseline networks and InceptionV3-ML network using two different crop sizes, and fusing features of different crop sizes. From the experimental results show in Fig. 4., we can have the following observations. First, by comparing results of the two baseline DCNNs and InceptionV3-ML, it can be noted that the best classification precision is achieved by the ResNet-50 network in two different crop sizes. The main reason is that the deeper ResNet-50 network can efficiently extract the deep semantic features while avoiding the degradation of deep networks.

In addition, it's worth noting that the classification accuracy of InceptionV3-ML is better than InceptionV3, the reason is that InceptionV3-ML could obtain comprehensive shape, texture, and boundary and other features. But it may also cause overfitting since the feature dimension is much larger than the sample number. Therefore, we exploit the feature selection method to remove feature redundancies and

Manuscript 498 submitted to 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Received February 19, 2019.

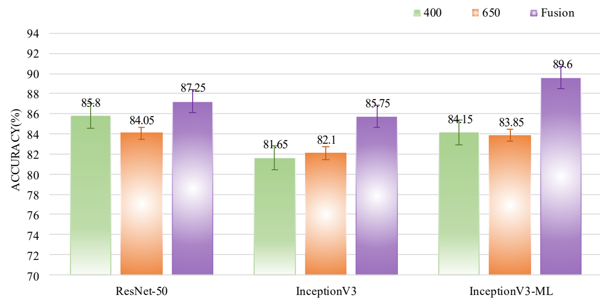

Fig. 4. Experimental results of different scenarios: green represents a crop with a size of 400, orange represents a crop with a size of 650, and purple represents a feature fusion between different crop sizes. (b) Experimental results of different feature dimensions via SMRR.

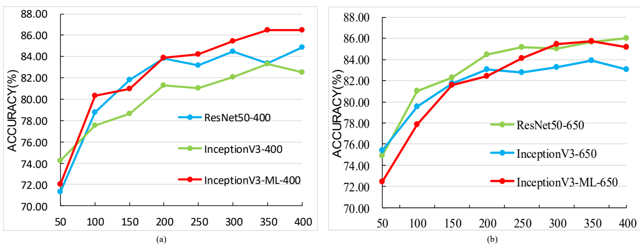

Fig. 5. Experimental results of different feature dimensions via SMRR (a) the crop size is 400, (b) the crop size is 650.

Table I. Algorithm comparison of different feature selection methods (%).

	Crop Size	ACC	AUC	Recall Rate
Baseline	**400**	**94.75**	**96.60**	**94.78**
	650	94.45	96.36	94.49
LLE [10]	**400**	**88.75**	92.47	88.83
	650	88.25	92.26	88.34
PAC [9]	**400**	**90.45**	93.58	90.56
	650	89.10	92.77	89.15
SMRR (our)	**400**	**95.97**	**97.33**	**96.02**
	650	**95.17**	**96.88**	**95.23**

reduce the feature dimension to avoid overfitting. Meanwhile, we can select the most informative and relevant features via SMRR method to improve the performance of the classifier.

C. *Effects of feature selection*

In this section, we verify the effectiveness of our proposed SMRR feature selection algorithm. Since our data set has 400 samples, we choose feature dimensions less than 400 to reduce overfitting. The experimental results as shown in Fig. 5. We can have tow observations. On the one hand, we find that classification accuracy increases with the increase of the selected feature dimensions. However, the best accuracy is obtained when feature dimension is 300 or 350. On the other hand, we also find that the results of InceptionV3-ML network achieves better results, which is consistent with the results in Section III *B*. As a result, we train the SVM classifier with image features dimension of 300.

Finally, we compare our proposed feature selection method with other methods via dual-network voting strategy between ResNet-50 and InceptionV3-ML network using the crop size of 400 and 650. The experimental results as shown in Table I, our propose SMRR algorithm achieves the best classification accuracy to 95.97%. Meanwhile, we demonstrate the superiority of our multi-level feature fusion and sparse multi-relationship learning model in breast cancer pathological image classification task.

IV. Conclusion

In this paper, we propose a simple and effective method for the classification of H&E stained histological breast cancer images. To increase the robustness and accuracy of the classifier, we use two different DCNN models to extract image features, separately. Then, we utilize new feature selection method of SMRR algorithm to enhance the classification results by reducing the feature dimension to prevent overfitting. Finally, we obtain the SVM model score by the voting between ResNet-50 and InceptionV3-ML. In our future work, we will combine deep learning with traditional machine learning methods to extract more robust features and optimize feature selection algorithm, which will further improve the classification accuracy.

References

[1] Carol E. DeSantis, Jiemin Ma, Ann Goding Sauer, Lisa A. Newman, and Ahmedin Jemal DVM, "Breast cancer statistics, 2017, racial disparity in mortality by state," *A Cancer Journal for Clinicians,* vol. 67, pp. 439-448, 2017.

[2] Kui Liu, Guixia Kang, Ningbo Zhang, and Beibei Hou, "Breast Cancer Classification Based on Fully-Connected Layer First Convolutional Neural Networks," *IEEE Access,* vol. 6, pp. 23722-23732, 2018.

[3] Z. Gandomkar, P. C. Brennan, and C. Mellothoms, "MuDeRN: Multi-category classification of breast histopathological image using deep residual networks," *Artificial Intelligence in Medicine,* vol. 88, pp. 14-24, 2018.

[4] Neslihan Bayramoglu, Juho Kannala, and Janne Heikkilä, "Deep learning for magnification independent breast cancer histopathology image classification," in *International Conference on Pattern Recognition*, pp. 2440-2445, 2017.

[5] Christoph Haarburger, Peter Langenberg, Daniel Truhn, Hannah Schneider, Johannes Thüring, Simone Schrading, Christiane K Kuhl, and Dorit Merhof, "Transfer Learning for Breast Cancer Malignancy Classification based on Dynamic Contrast-Enhanced MR Images," in *Bildverarbeitung für die Medizin 2018*, ed: Springer, pp. 216-221, 2018.

[6] Zhongyi Han, Benzheng Wei, Yuanjie Zheng, Yilong Yin, Kejian Li, and Shuo Li, "Breast Cancer Multi-classification from Histopathological Images with Structured Deep Learning Model," *Scientific Reports,* vol. 7, pp. 1-10, 2017.

[7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778, 2016.

[8] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna, "Rethinking the Inception Architecture for Computer Vision," pp. 2818-2826, 2015.

[9] Svante Wold, Kim Esbensen, and Paul Geladi, "Principal component analysis," *Chemometrics and intelligent laboratory systems,* vol. 2, pp. 37-52, 1987.

[10] Sam T. Roweis and Lawrence K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *science,* vol. 290, pp. 2323-2326, 2000.

[11] Xiaofeng Zhu, Heung-Il Suk, Li Wang, Seong-Whan Lee, Dinggang Shen, and Alzheimer's Disease Neuroimaging Initiative, "A novel relational regularization feature selection method for joint regression and classification in AD diagnosis," *Medical image analysis,* vol. 38, pp. 205-214, 2017.

[12] Erik Reinhard, Michael Ashikhmin, Bruce Gooch, and Peter Shirley, "Color Transfer between Images," *IEEE Computer Graphics & Applications,* vol. 21, pp. 34-41, 2002.

[13] Y-Lan Boureau, Jean Ponce, and Yann LeCun, "A theoretical analysis of feature pooling in visual recognition," in *Proceedings of the 27th international conference on machine learning (ICML-10)*, pp. 111-118, 2010.

Manuscript 498 submitted to 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Received February 19, 2019.