Few-shot Learning on the Diagnosis of Lymphatic Metastasis of Lung Carcinoma

Haohua Teng
Shanghai Chest Hospital, Shanghai Jiao Tong University
Weidong Zhang
Shanghai Aitrox Technology Corporation Limited
Jinzhi Wei
Shanghai Chest Hospital, Shanghai Jiao Tong University
Lei Lv
Shanghai Aitrox Technology Corporation Limited
Licheng Tang
Shanghai Aitrox Technology Corporation Limited
Chi-Cheng Fu
Shanghai Aitrox Technology Corporation Limited
Yumeng Cai
Shanghai Chest Hospital, Shanghai Jiao Tong University
Gang Qin
Shanghai Chest Hospital, Shanghai Jiao Tong University
Min Ye
Shanghai Chest Hospital, Shanghai Jiao Tong University
Qu Fang
Shanghai Aitrox Technology Corporation Limited
Yuchen Han (ychan@cmu.edu.cn)
Shanghai Chest Hospital, Shanghai Jiao Tong University

Research Article

Keywords: lymphatic metastasis of lung carcinoma, Few-shot learning, artificial intelligence

Posted Date: May 19th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-493350/v1

Few-shot learning on the diagnosis of lymphatic metastasis of lung carcinoma

Haohua Teng, MD[1], Weidong Zhang, MS[2], Jinzhi Wei, MD[1], Lei Lv, PhD[2], Licheng Tang, MS[2], Chi-Cheng Fu, PhD[2], Yumeng Cai, MD[1], Gang Qin, BS[1] Min Ye, AS[1], Qu Fang, MS[2], Yuchen Han, MD, PhD[1]

[1] Department of Pathology, Shanghai Chest Hospital, Shanghai Jiao Tong University, China.

[2] Shanghai Aitrox Technology Corporation Limited, Shanghai, China

Corresponding Author: Yuchen Han

Shanghai Chest Hospital, Shanghai Jiao Tong University

Address: No. 241, West Huaihai Rd. Xuhui District, Shanghai, P.R. China, 200030

Tel: +86-18017327736

Fax: +1-202-782-6845

E-mail: ychan@cmu.edu.cn

Key points

Question: How to develop an automatized classification algorithm using a dataset with unbalanced and small clinical data?

Findings: A few-shot learning model showed similar performance with the traditional deep learning approach with 20 positive/40 negative patches, with AUC 0.6124 (95% CI, 0.6077-0.6147).

Meaning: A minimal patch number was found for developing an automatized classification AI model for the diagnosis of lymph node metastases of lung cancer.

Abstract

Background： Lymph node metastasis of lung cancer plays an important part in lung cancer diagnosis since its close relationship with patients' prognosis and subsequent treatment, however, diagnosis of lymph node metastases often was tedious and takes too much time for pathologists. Automatized classification with few-shot learning algorithms is required for clinical situations that only unbalanced and small datasets can be collected. We developed a few-shot learning algorithms for automatized classification of lymphatic metastasis of lung carcinoma from Whole Slide Images (WSIs) through a deep convolutional neural network (DCNN).

Material and methods: Lymph node slides of patients with lung cancer collected from July to December 2018 were reviewed and used for model development retrospectively. A total of 1701 lymph node slides from 453 lung cancer patients with different histological types were enrolled. Each slide could provide a different number of patches, in which 6, 8, 10, 20, and 30 positive patches and 12, 16, 20, 40, and 60 negative patches for training and validation, respectively. A total of 1577 WSIs were used. The receiver operating characteristic (ROC) analysis of the model was performed for estimating the performance and compared to the traditional deep learning approach Resnet34.

Results: Deep learning methods, when trained with very few shots (6 positive/12 negative, 8 positive/16 negative, and 10 positive/20 negative patches), the traditional deep learning

approach Resnet34 was outperformed by our Aitrox model. When the dataset was raised to 20 positive/40 negative patches, similar performance was achieved with AUC 0.5895 (95% CI, 0.5833-0.5910) for Resnet34 and 0.6124 (95% CI, 0.6077-0.6147) for our Aitrox model. Resnet34 over-performed our Aitrox model (AUC=0.6450 [95%CI, 0.6397-0.6470]) with an AUC of 0.7481 (95%CI, 0.7120-0.7485) when the training and validation dataset was raised to 30 positive/60 negative patches.

Conclusions: This proposed few-shot method may address AI approaches to the diagnosis of lymphatic metastasis of lung carcinoma and other kinds of carcinoma in the future.

Keywords: lymphatic metastasis of lung carcinoma; Few-shot learning; artificial intelligence

Introduction

Lung cancer is one of the most common cancers worldwide[1] and is the most common cancer in China[2]. Accurate lung cancer staging is an essential task performed by pathologists to make clinical management. Evaluation of the extent of lung cancer spread by histopathological examination of the lymph node is an important factor of lung cancer staging[3]. Nevertheless, the pathologic screen of positive lymph nodes is tedious and time-consuming. There was a total of 5657 lung cancer patients who had been performance pulmonary surgeries at Shanghai Chest Hospital from July 2018 to December 2018, and 24060 lymph node groups have been removed, each lymph node group represents one whole slide, while only 671 lymph node groups were positive. The possibility of artificial intelligence aided diagnostics may be advantageous.

In recent years, the increased use of whole slide imaging coupled with advances in computation and deep learning methods has sparked much interest in applying AI to pathology[4-6]. Depending on the technical improvement of the WSIs scanning, the pathological diagnosis combining the AI recognition technology improving the efficiency of pathologist may come true, for example, with the algorithm assistance, the sensitivity and specificity of reviewing the lymph node metastasis of breast cancer were increased, and the average review time per image was reduced[7, 8]. The algorithm design based on the dataset, datasets with different scales, and types need different strategies, usually, a large dataset was

fit for weakly supervised deep learning, and a small dataset was fit for fully supervised deep learning. The works of Kanavati[9] and Wang[10] suggested that the weakly supervised deep learning may be more efficient than the fully supervised deep learning when classifying the lung cancer histological types with the large dataset. The weakly supervised deep learning needed the slide-level diagnoses and avoided the detailed cell-level annotations which means tons of workload. The clinical-grade pathology using was trained by weakly supervised deep learning with over thousands of slides dataset including prostate core biopsy, breast cancer metastatic lymph node, and skin lesion separately; and the area under the curve (AUC) can be 0.965 on the test of breast axillary lymph node metastasis[8]. The accuracy of the weakly supervised deep learning may be affected by the labels[9] and the size of the dataset which could be improved by upgrading the algorithm iteration[11-13]. Atsushi Teramoto improves 4% AUC by using a two-step supervised strategy for the classification of lung cytological images with 60 cases dataset[14].

However, to further enhance the performance of the AI model, it would be necessary to increase the number of images used to train the CNN model. Furthermore, in the real world, most lymph nodes are negative and few nodes have metastatic tumors, which means the negative and normal tissues can be easily acquired, while positive and pathological tissues are infrequent. Small number of positive samples and unbalanced proportion of positive and negative samples often cause overfitting when training with traditional deep convolutional

neural networks (DCNN). The rational utilization of the normal datasets instead of discarding them for only balancing the proportion is also our consideration.

In this study, a few-shot learning algorithm[15-17] on automatizing the classification of malignant lung cells from Whole Slide Images (WSIs) though a DCNN was developed to solve the above-mentioned problems on mimicking the real-world environment. Meanwhile, a traditional network Resnet 34 was also enrolled in this investigation for comparison with our prototypical network-based few-shot learning model[18-20]. To the best of our knowledge, this study introduces the first application of few-shot learning methods for digital pathology image analysis in the diagnosis of lymphatic metastasis of lung carcinoma.

Materials

Study participants

For this study, a total of 1701 lymph node slides from 453 lung cancer patients with different histological types were enrolled. Slides were collected from the Shanghai Chest Hospital, China, from July 2018 to December 2018. All patients gave written informed consent for their participation. This study has been approved by the Ethics Committees of Shanghai Chest Hospital, Shanghai Jiao Tong University, and it conforms to the principles outlined in the Declaration of Helsinki. The study was also conducted under an exemption approved by the Shanghai Chest Hospital review board. All the slides were formalin-fixed, paraffin-

embedded and stained with hematoxylin and eosin (H and E). Out of the 1701 slides, 124 (25 positives and 99 negatives) were used for algorithm training and validation, 1577 (646 positives and 931 negatives) were used for testing (Supplementary Table 1).

Image dataset

All the optimized H&E stained WSIs were digitalized by a bright-field scanner (DMS-10, D-Metrix, http://www.dmks.cn/), using a 40X magnification objective with a color CCD with a pixel resolution of 0.25 μm. To construct the image dataset for our DCNN, patch images of 224x224 pixels were cut from the original microscopic WSIs. Patches were extracted from each WSI. In region annotation, patches with vivid cytopathologic features for tumors were selected as positive patches by an experienced pathologist with the online digital slides viewer named MISS (Microscope Image Information System) from Shanghai Aitrox Technology Corporation Limited (Shanghai, China) (**Fig 1**).

Methods

Resnet 34

In this study, a traditional network Resnet34 was used to compare against other algorithms and applied to few-shot lymphatic metastasis of lung carcinoma diagnostics. This Resnet34 was unmodified from its original architecture[19-21]. Training and validation datasets were

allocated according to 4:1 and 5-folder cross-validation was carried out in each experiment. Adam was used as the optimizer, and the initial learning rate was 0.001, beta 1 0.9, and beta 2 0.99. Cross entropy was used as the value function.

Our Aitrox model

For comparison with the Resnet34, a strategy to training a 2-way classification model with few and unbalanced samples was designed based on Prototypical Network[15]. In our cases, we use a small set of M usual positive labeled patches and a large set of N negative labeled patches as our training set (N =2 M). For each training epoch, M positive and M (randomly selected from N) negative patches were enrolled (**Fig 2**). There are two main steps in our method. First, the prototypical network was trained on annotated patches as mentioned before. Second, the well-trained model was used to predict the Whole Slide Images (WSIs). The Prototypical Network is a kind of metric learning network, which is not like the traditional classification networks that classify objects directly. The output of the prototypical network is a high-dimensional feature vector that was used to update our model weights.

Design Choice

We chose Euclidean distance[22-24] as our distance metric. Take one of the small datasets for example, the number of the labeled positive patch in training set M = 6. The number of the labeled negative patch in training set N = 12. Ns = 3, Nq = 4.

We chose Resnet34 as our backbone to embed the input into a 512-dimensional feature vector. Trained about 10,000 epochs. The process of each epoch in our training mainly consists of 3 steps (**Fig 3**):

Input: Training set $D = \{(x_1, y_1), \ldots. (x_n, y_n)\}$, where each $y_i \in \{0, 1\}$ where 0 means negative and 1 means positive. D_k means the set of D containing all elements (x_i, y_i) where $y_i = k$.

1. Randomly choose **M** negative patches from **N** negative labeled patches.

$T_0 \leftarrow RandomSample(D_0, N_M)$ *where* N_M *means the number of labeled positive patches.* T_0 *means all negative patches in the training set.*

$T_1 = D_1$. T_1 *means all negative patches in the training set.*

2. Randomly select **support set** and **query set** from the training set. Then calculate the center for each class from support set. The center is an M-dimensional vector.

for k in {0, 1} do

$S_k \leftarrow RandomSample(T_k , N_s)$ *where* N_s *means the number of support set*

$Q_k \leftarrow RandomSample(T_k, N_q)$ *where* N_q *means the number of query set*

$C_k \leftarrow \frac{1}{N_{S_k}} \sum_{(xi,yi) \in S_k} f_\theta(x_i)$ *where* N_{S_k} *means the number of Support set for class k.* f_θ *denotes our neural network with learnable parameters* θ. f_θ *could embed the input image into an M-dimensional feature space.*

3. Update the parameters of our neural network from query set. We first calculate the distance between the descriptor for query set and the center of each class. Distance between query input $\boldsymbol{x}$ and the center of class $\boldsymbol{k}$ could be denoted by $D_\theta(x, k) = g(f_\theta(x),\ C_k)$ where g is the distance metric function. In our case, we chose Euclidean distance as our metric. Then we can obtain a distribution over each class for a query $\boldsymbol{x}$ *by softmax function:* $P_\theta(k|x) = \frac{\exp(-g(f_\theta(x), C_k))}{\sum_{k_i \in (0,1)} \exp(-g(f_\theta(x), C_{k_i}))}$. *The loss between prediction and ground truth could be calculated by cross-entropy loss.*

Results

Our Aitrox model was compared with traditional CNN classifier Resnet34 with different training and validation datasets. The test dataset was kept constantly at 1577 slides.

The receiver operating characteristic (ROC) curve of our Aitrox model and ResNet34 in a test dataset was shown in **fig 4**. Both models were trained in the same training and validation dataset with 6pos/12neg patches, 8pos/16neg patches, 10pos/20neg patches, respectively. AUC on WSI for each model was calculated by varying the probability of the Kth highest

probability of patch in WSI. In our case threshold K = 5. Sensitivity, specificity, and accuracy were calculated under the Threshold of 0.6/5 (Table 1)

When both models were trained in the same training and validation dataset with 20pos/40neg patches, respectively, similar performance was achieved since AUC for Resnet34 was 0.5895 (95%CI, 0.5833-0.5910) and our Aitrox model was 0.6124 (95%CI, 0.6077-0.6147), respectively (**Fig 5**). Sensitivity, specificity and accuracy were also shown in Table 1.

Resnet34 over-performed our Aitrox model (AUC=0.6450 [95%CI, 0.6397-0.6470]) with an AUC of 0.7481 (95%CI, 0.7120-0.7485) when training and validation dataset were raised to 30pos/60neg patches (**Fig 6**). Detailed information on sensitivity, specificity and accuracy for each model could be found in Table 1.

Discussion

In this study, the evaluation of deep learning for automated lymphatic metastasis of lung carcinoma diagnosis with small training datasets suggests that the traditional network Resnet 34 over-performed our Aitrox model with 30pos/60neg patches at an AUC of o.7481, but it deteriorated substantially when training was relatively low with 20pos/40neg patches or fewer. Results from our Aitrox model suggested that the performance was acceptable when

training with a relatively small size (10pos/20neg patches) and also has more graceful degradation in performance when using a very small number of training datasets with 6pos/12neg patches.

We believe that the performance of Resnet34 decreases as the training number decreases was because that it may be hard for the network to do fine-tuning since cross-entropy[25-27] was used as a loss function which could not extract enough features for model training. In our Aitrox model, metric-based[28, 29] optimization method which might extract more features were used with a small dataset (6, 8, and 10 positive patches) for model training, the performance was quite acceptable with AUCs all above 0.60. When the dataset was raised to 20pos/40neg patches, both models achieved similar performance since they may have extracted nearly the same feature information. When the dataset becomes larger (for example, the number of positive samples reaches 30), the limitations of metric-based learning with fewer samples were reflected. During the training process, a small amount of data was selected for each iteration to build support set and query set, and then the number of query sets and support sets were calculated according to the Euclidean distance, which was used as the index of gradient return for model optimization. When it comes to the model selecting, this standard was also used to save the model (the data were not verified on all the validation dataset, because we just determined to verify this metric based method). For a large amount of dataset, the cross-entropy based method has already achieved good performances (AUC

0.85 and above), and the concept of few-shot learning does not exist, based on which future comparison was not made in this study.

These findings should support the investigation of few-shot learning methods for the clinical AI diagnosis which rely on small and unbalanced data sets. Application to lymphatic metastasis of lung carcinoma diagnosis is one such use, in which the small number of positive lymph nodes and the large number of negative lymph nodes are typically available. AI bias is also a concern for both medical applications and predicting needs for health care[30]. In most cases, the number of positive lymph nodes is about half that of negative lymph nodes. Few-shot learning could offer an option to address both the small training dataset and the balance problem.

Limitations

This study has some limitations. Only Prototypical Network and Resnet34 were examined. Other DCNNs such as Siamese Network[31], vgg[32, 33], Inception[34] and Alex-net[35] should be taken into consideration in future studies. Also, the largest dataset used in this study was 30pos/60neg patches, in which 124 slides were used for training and validation. Additional samples are needed for further comparing between few-shot learning and traditional deep learning methods.

Conclusions

To our knowledge, this study is the first to explore the performance of few-shot learning networks in lymphatic metastasis of lung carcinoma. The results show that the performance of widely used deep learning methods relying on full supervision that typically require large training data sets degrades substantially when used with limited data sets while few-shot learning appears to perform better and have more graceful performance degradation as the number of training database decreased. This few-shot method may address AI approaches to the diagnosis of lymphatic metastasis of lung carcinoma and other kinds of carcinoma in the future.

Declarations:

Ethics approval and consent to participate

All patients gave written informed consent for their participation. This study has been approved by the Ethics Committees of Shanghai Chest Hospital, Shanghai Jiao Tong University, and it conforms to the principles outlined in the Declaration of Helsinki. The study was also conducted under an exemption approved by the Shanghai Chest Hospital review board.

Consent for publication

Not applicable.

Availability of data and materials

The data sets supporting the results of this article are included within the article and its supplementary information files.

Competing interests

Authors have no conflict of interest to declare. I would like to declare on behalf of my co-authors that the work described was original research that has not been published previously, and not under consideration for publication elsewhere, in whole or in part. All the authors listed have approved the manuscript that is enclosed.

Funding

This manuscript is supported by Intelligent Shanghai Program Shanghai Health Development

Planning Commission Projects (2018ZHYL0213).

Authors' contributions

Y.H. is the guarantor of integrity of entire study and approved the final version of the manuscript. Q.F. and Y.H. studied concepts. H.T. and C.F. designed the study and research the literatures. W.Z and L.T. accomplished the experiments. H.T., Y.C., J.W., G.Q., and M.Y. retrieved the data. Y.C. and J.W. interpreted the data and L.L. analyzed the data. H.T. wrote the main manuscript and L.L. revised the manuscript.

Acknowledgements

Not applicable.

References:

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A: **Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries**. *CA Cancer J Clin* 2018, **68**(6):394-424.
2. Cao M, Chen W: **Epidemiology of lung cancer in China**. *THORAC CANCER* 2019, **10**(1):3-7.
3. Detterbeck FC, Chansky K, Groome P, Bolejack V, Crowley J, Shemanski L, Kennedy C, Krasnik M, Peake M, Rami-Porta R: **The IASLC Lung Cancer Staging Project: Methodology and Validation Used in the Development of Proposals for Revision of the Stage Classification of NSCLC in the Forthcoming (Eighth) Edition of the TNM Classification of Lung Cancer**. *J THORAC ONCOL* 2016, **11**(9):1433-1446.
4. Landau MS, Pantanowitz L: **Artificial intelligence in cytopathology: a review of the literature and overview of commercial landscape**. *J Am Soc Cytopathol* 2019, **8**(4):230-241.
5. Janowczyk A, Madabhushi A: **Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases**. *J Pathol Inform* 2016, **7**:29.
6. Tizhoosh HR, Pantanowitz L: **Artificial Intelligence and Digital Pathology: Challenges and Opportunities**. *J Pathol Inform* 2018, **9**:38.
7. Steiner DF, MacDonald R, Liu Y, Truszkowski P, Hipp JD, Gammage C, Thng F, Peng L, Stumpe MC: **Impact of Deep Learning Assistance on the Histopathologic Review of Lymph Nodes for Metastatic Breast Cancer**. *AM J SURG PATHOL* 2018, **42**(12):1636-1646.
8. Liu Y, Kohlberger T, Norouzi M, Dahl GE, Smith JL, Mohtashamian A, Olson N, Peng LH, Hipp JD, Stumpe MC: **Artificial Intelligence-Based Breast Cancer Nodal Metastasis Detection: Insights Into the Black Box for Pathologists**. *ARCH PATHOL LAB MED* 2019, **143**(7):859-868.
9. Kanavati F, Toyokawa G, Momosaki S, Rambeau M, Kozuma Y, Shoji F, Yamazaki K, Takeo S, Iizuka O, Tsuneki M: **Weakly-supervised learning for lung carcinoma classification using deep learning**. *Sci Rep* 2020, **10**(1):9297.
10. Wang X, Chen H, Gan C, Lin H, Dou Q, Tsougenis E, Huang Q, Cai M, Heng PA: **Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis**. *IEEE Trans Cybern* 2020, **50**(9):3950-3962.
11. Kim YG, Song IH, Lee H, Kim S, Yang DH, Kim N, Shin D, Yoo Y, Lee K, Kim D *et al*: **Challenge for Diagnostic Assessment of Deep Learning Algorithm for Metastases Classification in Sentinel Lymph Nodes on Frozen Tissue Section Digital Slides in Women with Breast Cancer**. *CANCER RES TREAT* 2020, **52**(4):1103-1111.
12. Ehteshami BB, Veta M, Johannes VDP, van Ginneken B, Karssemeijer N, Litjens G, van der Laak J, Hermsen M, Manson QF, Balkenhol M *et al*: **Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer**. *JAMA* 2017, **318**(22):2199-2210.

13. Kim YG, Kim S, Cho CE, Song IH, Lee HJ, Ahn S, Park SY, Gong G, Kim N: **Effectiveness of transfer learning for enhancing tumor classification with a convolutional neural network on frozen sections**. *Sci Rep* 2020, **10**(1):21899.
14. Teramoto A, Tsukamoto T, Yamada A, Kiriyama Y, Imaizumi K, Saito K, Fujita H: **Deep learning approach to classification of lung cytological images: Two-step training using actual and synthesized images by progressive growing of generative adversarial networks**. *PLOS ONE* 2020, **15**(3):e229951.
15. Snell J, Swersky K, Zemel R: **Prototypical Networks for Few-shot Learning**. In*:*2017; 2017: 4077-4087.
16. Das D, Lee C: **A Two-Stage Approach to Few-Shot Learning for Image Recognition**. *IEEE Trans Image Process* 2019.
17. Yaqing Wang, Yao Q, Kwok J, Ni LM: **Generalizing from a Few Examples: A Survey on Few-Shot Learning**. *ACM COMPUT SURV* 2020, **1**(1):1-34.
18. Talo M, Yildirim O, Baloglu UB, Aydin G, Acharya UR: **Convolutional neural networks for multi-class brain disease detection using MRI images**. *Comput Med Imaging Graph* 2019, **78**:101673.
19. Yu S, Xiao D, Frost S, Kanagasingam Y: **Robust optic disc and cup segmentation with deep learning for glaucoma detection**. *Comput Med Imaging Graph* 2019, **74**:61-71.
20. Fernandez K, Korinek M, Camp J, Lieske J, Holmes D: **Automatic detection of calcium phosphate deposit plugs at the terminal ends of kidney tubules**. *Healthc Technol Lett* 2019, **6**(6):271-274.
21. Ayinde BO, Inanc T, Zurada JM: **Redundant feature pruning for accelerated inference in deep neural networks**. *Neural Netw* 2019, **118**:148-158.
22. Jameel SM, Hashmani MA, Rehman M, Budiman A: **An Adaptive Deep Learning Framework for Dynamic Image Classification in the Internet of Things Environment**. *Sensors (Basel)* 2020, **20**(20).
23. Figueroa-Mata G, Mata-Montero E: **Using a Convolutional Siamese Network for Image-Based Plant Species Identification with Small Datasets**. *Biomimetics (Basel)* 2020, **5**(1).
24. Chicco D: **Siamese Neural Networks: An Overview**. *Methods Mol Biol* 2021, **2190**:73-94.
25. Aur D, Vila-Rodriguez F: **Dynamic Cross-Entropy**. *J Neurosci Methods* 2017, **275**:10-18.
26. Jamin A, Humeau-Heurtier A: **(Multiscale) Cross-Entropy Methods: A Review**. *Entropy (Basel)* 2019, **22**(1).
27. Bao R, Yuan X, Chen Z, Ma R: **Cross-Entropy Pruning for Compressing Convolutional Neural Networks**. *NEURAL COMPUT* 2018, **30**(11):3128-3149.
28. Gallagher AG: **Metric-based simulation training to proficiency in medical**

education:- what it is and how to do it. *Ulster Med J* 2012, **81**(3):107-113.

29. Xu H, Zeng W, Zeng X, Yen GG: **A Polar-Metric-Based Evolutionary Algorithm**. *IEEE Trans Cybern* 2020, **PP**.

30. Burlina P, Paul W, Mathew P, Joshi N, Pacheco KD, Bressler NM: **Low-Shot Deep Learning of Diabetic Retinopathy With Potential Applications to Address Artificial Intelligence Bias in Retinal Diagnostics and Rare Ophthalmic Diseases**. *JAMA OPHTHALMOL* 2020, **138**(10):1070-1077.

31. Chopra S, Hadsell R, LeCun Y: **Learning a similarity metric discriminatively, with application to face verification**. In*: 2005-01-01* 2005: IEEE; 2005: 539-546.

32. Sengupta A, Ye Y, Wang R, Liu C, Roy K: **Going Deeper in Spiking Neural Networks: VGG and Residual Architectures**. *Front Neurosci* 2019, **13**:95.

33. Geng L, Zhang S, Tong J, Xiao Z: **Lung segmentation method with dilated convolution based on VGG-16 network**. *Comput Assist Surg (Abingdon)* 2019, **24**(sup2):27-33.

34. Guan Q, Wan X, Lu H, Ping B, Li D, Wang L, Zhu Y, Wang Y, Xiang J: **Deep convolutional neural network Inception-v3 model for differential diagnosing of lymph node in cytological images: a pilot study**. *Ann Transl Med* 2019, **7**(14):307.

35. Hosny KM, Kassem MA, Foaud MM: **Classification of skin lesions using transfer learning and augmentation with Alex-net**. *PLOS ONE* 2019, **14**(5):e217293.

Table 1. Sensitivity, specificity and accuracy for Resnet34 and our AI model.

Algorithm	No. Patch (Pos/Neg)	AUC	Sensitivity	Specificity	Accuracy
Resnet 34	6/12	0.578	0.6548	0.4834	0.5536
	8/16	0.5506	0.7755	0.333	0.5143
	10/20	0.5895	0.582	0.5704	0.5751
	20/40	0.5895	0.582	0.5704	0.5751
	30/60	0.7481	0.6471	0.7712	0.7204
Our AI model	6/12	0.6704	0.7446	0.5414	0.6246
	8/16	0.667	0.6873	0.5897	0.6297
	10/20	0.6504	0.5851	0.6638	0.6316
	20/40	0.6124	0.7152	0.4855	0.5796
	30/60	0.645	0.6533	0.6069	0.6259

Figure legends:

Fig 1. Overview of patches extracted from Whole Slide Image (WSI)

Fig 2. Overview of each training epoch

Fig 3. Process of each epoch in our Model. (a) Randomly sample support-set and query-set in the training set. (b) Calculation of the Center vector for each type by using a support-set. (c) Calculation of the Q M-dim feature vectors for each type by using a query-set. (d) Distance calculation between query set and two center vectors (S denotes the M-dimensional Feature Space. C1 and C2 denotes the center vector of two types. Q1 is the feature vector of one sample in Query-set after being fed into CNN).

Fig 4. ROC curve for each model. (a, d) ROC curve for Resnet 34 and our Aitrox model when training with 6 positive and 12 negative patches, respectively. (b, e) ROC curve for Resnet 34 and our Aitrox model when training with 8 positive and 16 negative patches, respectively. (c, f) ROC curve for Resnet 34 and our Aitrox model when training with 10 positive and 20 negative patches, respectively.

Fig 5. ROC curve for Resnet 34 (a) and our Aitrox model (b) when training with 20 positive and 40 negative patches.

Fig 6. ROC curve for Resnet 34 (a) and our Aitrox model (b) when training with 30 positive and 60 negative patches.

Figures

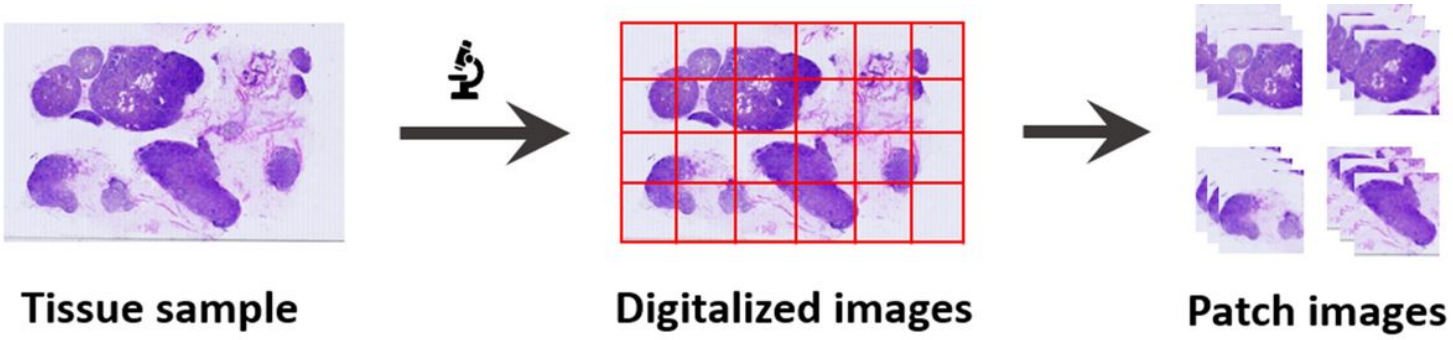

Figure 1

Overview of patches extracted from Whole Slide Image (WSI)

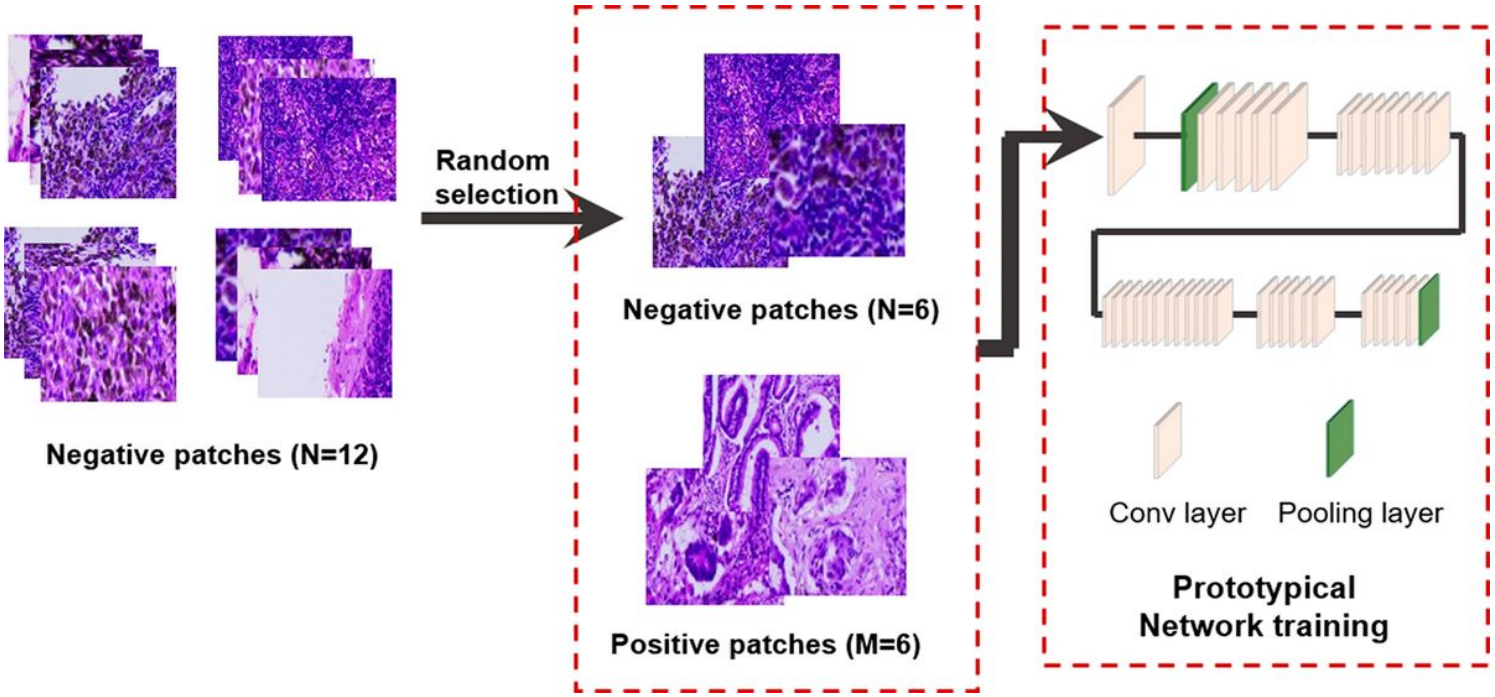

Figure 2

Overview of each training epoch

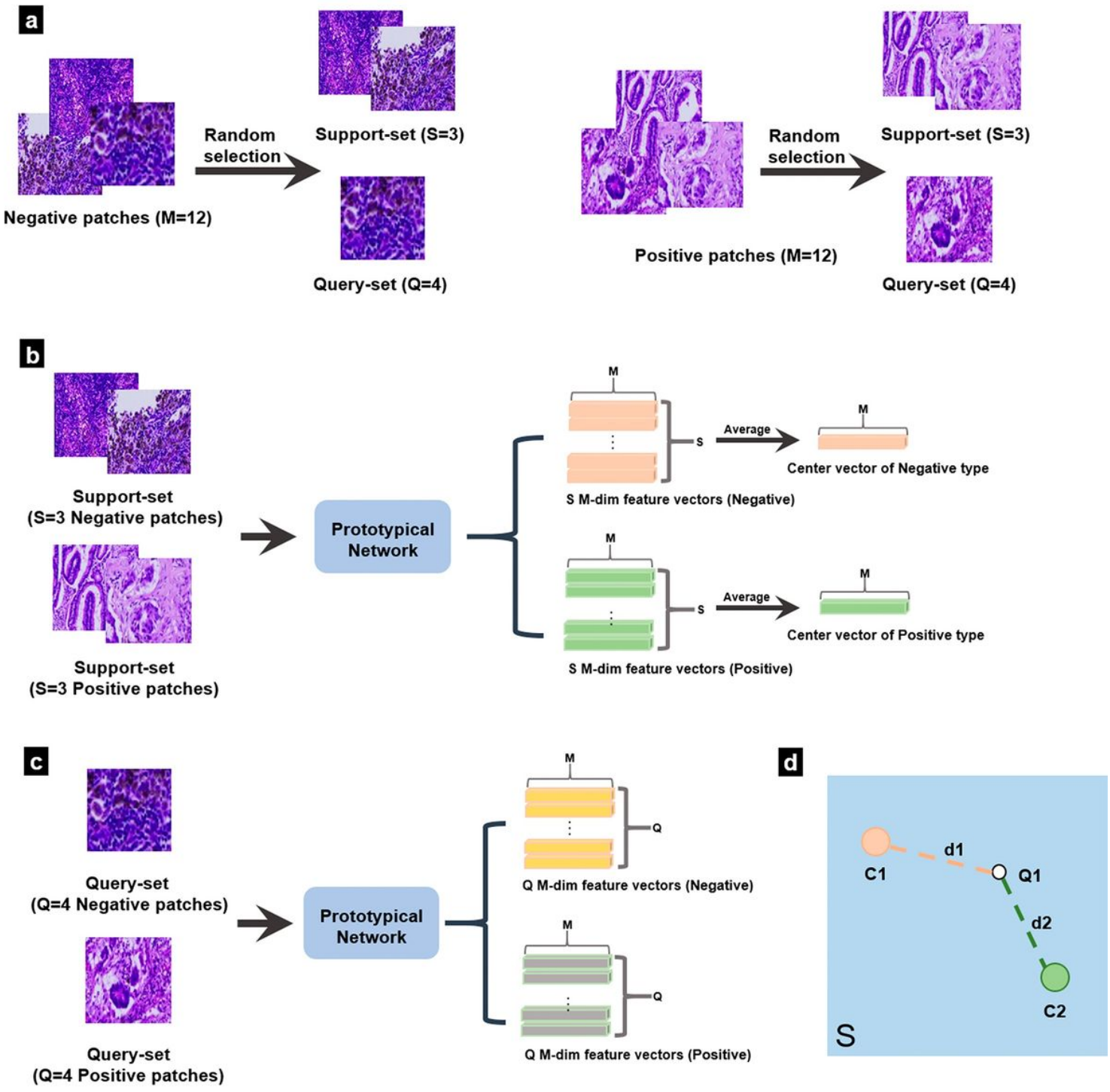

Figure 3

Process of each epoch in our Model. (a) Randomly sample support-set and query-set in the training set. (b) Calculation of the Center vector for each type by using a support-set. (c) Calculation of the Q M-dim feature vectors for each type by using a query-set. (d) Distance calculation between query set and two center vectors (S denotes the M-dimensional Feature Space. C1 and C2 denotes the center vector of two types. Q1 is the feature vector of one sample in Query-set after being fed into CNN).

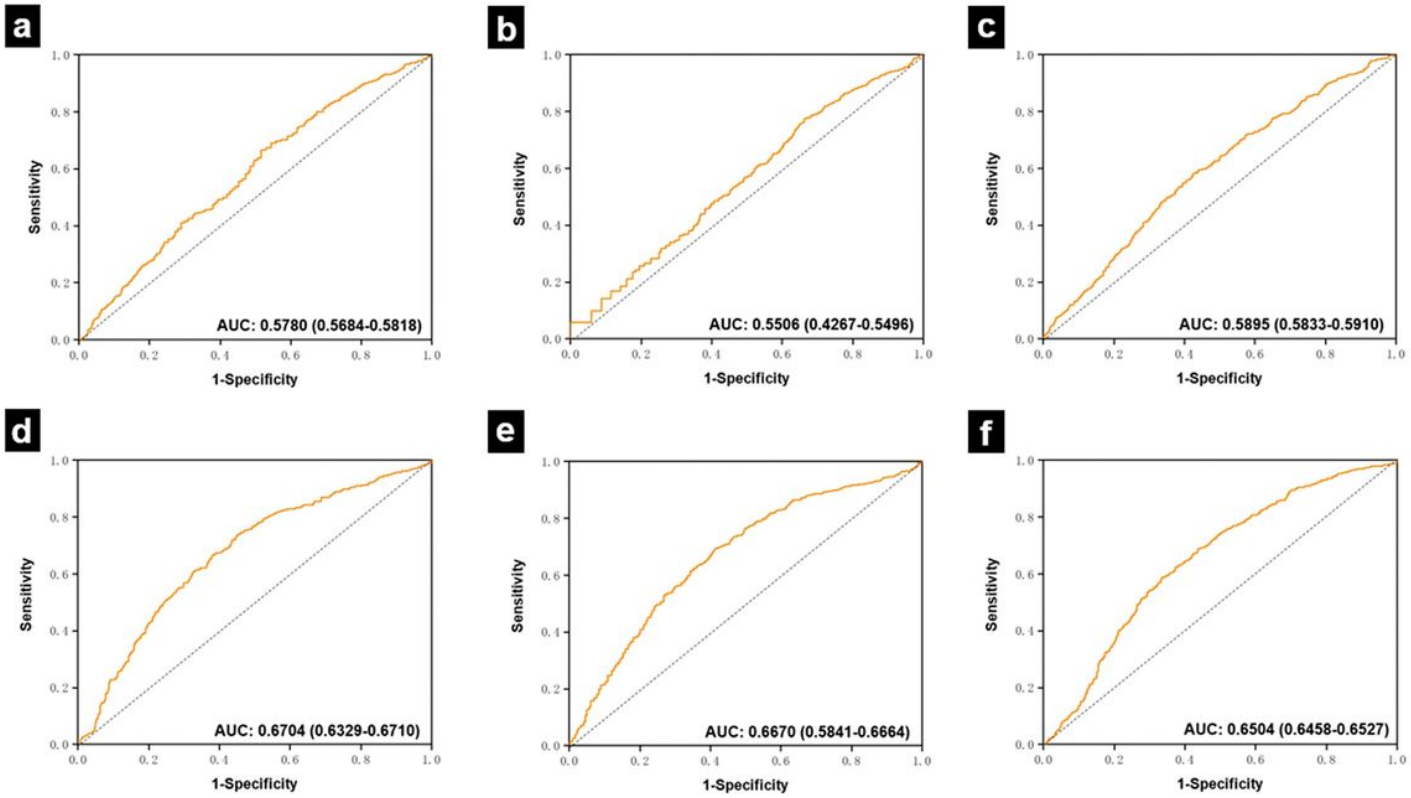

Figure 4

ROC curve for each model. (a, d) ROC curve for Resnet 34 and our Aitrox model when training with 6 positive and 12 negative patches, respectively. (b, e) ROC curve for Resnet 34 and our Aitrox model when training with 8 positive and 16 negative patches, respectively. (c, f) ROC curve for Resnet 34 and our Aitrox model when training with 10 positive and 20 negative patches, respectively.

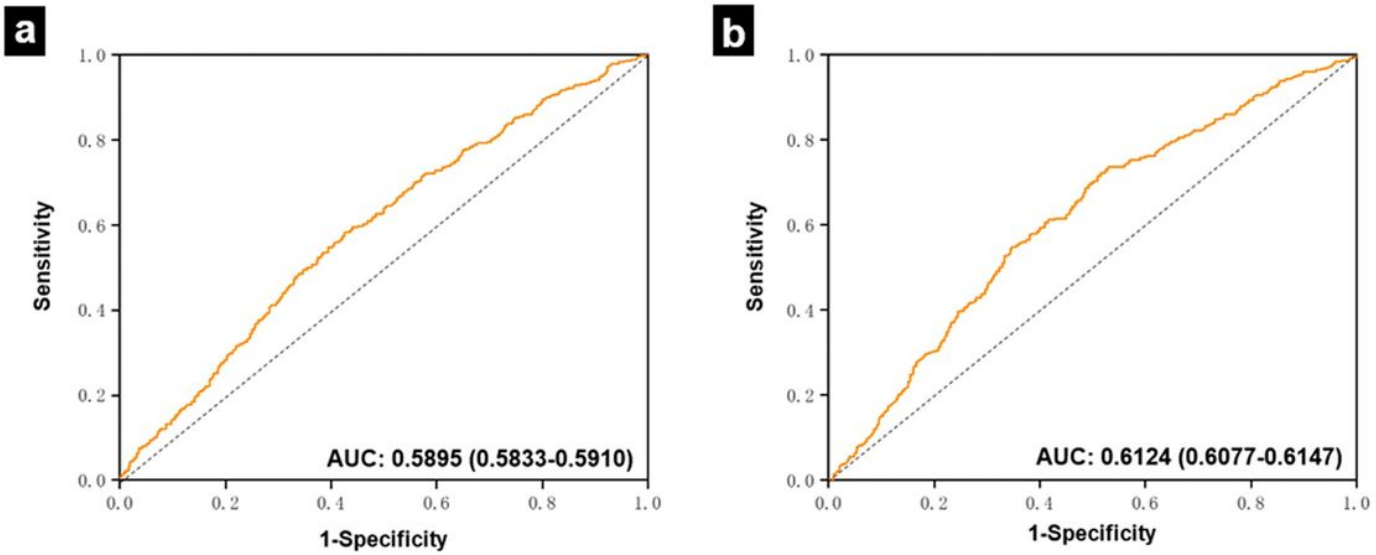

Figure 5

ROC curve for Resnet 34 (a) and our Aitrox model (b) when training with 20 positive and 40 negative patches.

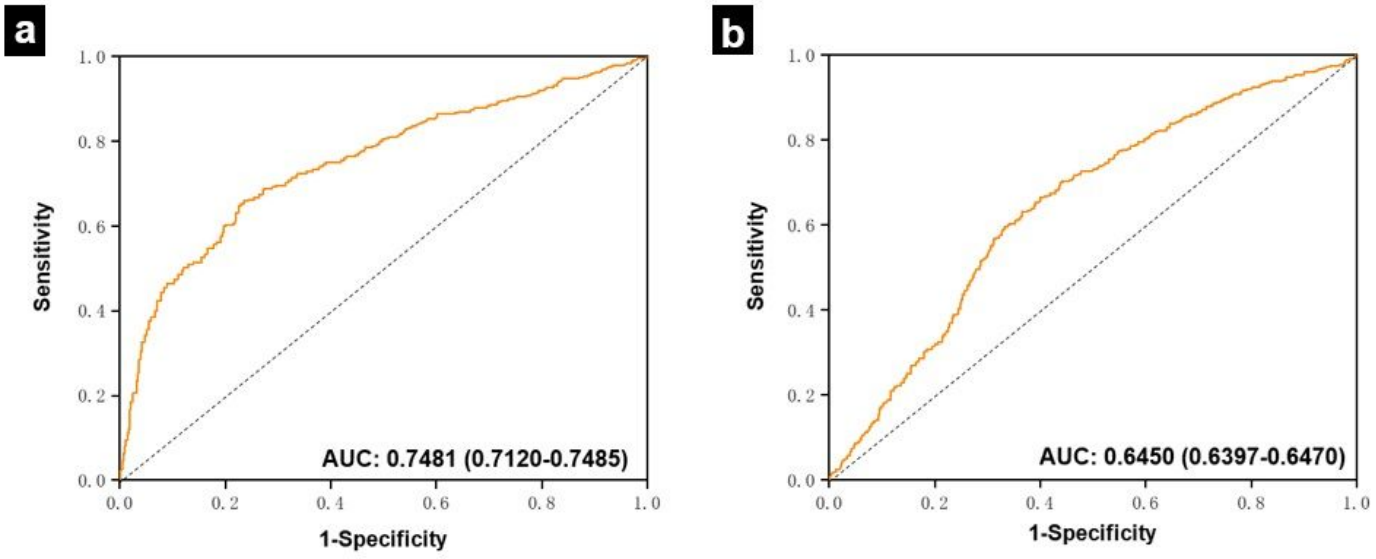

Figure 6

ROC curve for Resnet 34 (a) and our Aitrox model (b) when training with 30 positive and 60 negative patches.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTable1.docx
- Supplementarytables.pdf
- Supplementaryfigs.pdf