



Breast Ultrasound Tumor Detection Based on Active Learning and Deep Learning

Gen Liu¹, Jiyong Tan^{1,2}(✉), Hongguang Yang¹, Yuanwei Li¹, Xi Sun¹, Jiayi Wu³,
and Baoming Luo³

¹ AISONO AIR Lab, Shenzhen, China
scutjy2015@163.com

² Harbin Institute of Technology, Harbin, China

³ Department of Ultrasound, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University,
Guangzhou, China

Abstract. Early breast cancer screening and diagnosis policy plays a significance role in reducing breast cancer mortality, which is the most common malignant tumor for women. Therefore, its accuracy and efficiency are very important. To cover these challenges in mass breast screening and diagnosis, including varied ultrasound image quality from different equipments, expensive professional annotation, we propose a novel method based on active learning and convolution neural networks for selecting more informative images and tumor detection, respectively. Firstly, we verify the effectiveness of active learning in the application of our breast ultrasound data. Secondly, we select the informative images from the origin training set using the Multiple Instance Active Learning (MIAL) with One-Shot Path Aggregation Feature Pyramid Network (OPA-FPN) structure. Through this way, we effectively balance the ratio of hard samples and simple samples in the origin training set. Finally, we train the model based on EfficientDet with specific and valid parameters for our breast ultrasound data. Through the corresponding ablation experiment, it is verified that the model trained on the selected dataset by combining MIAL with OPA-FPN exceeds the origin model in the metrics about sensitive, specificity and F1-score. Meanwhile, while keeping the corresponding metrics approximately the same, the confidence of inference images from the new model is higher and stable.

Keywords: Breast ultrasound · Tumor detection · AIBUS · Active learning · Breast Cancer screening and diagnosis

1 Introduction

Breast cancer is the most common malignant tumor for women [1]. Compare to the amount of world breast cancer cases and deaths each year, China accounts for 12.2% and 9.6%, respectively. As for the 5-year survival rate, China presents 82.0%, which 7.1% lower than the U.S. The literature [2] pointed out that the early diagnosis rate of breast cancer in China is less than 20%, and the proportion of breast cancer found through screening is less than 5%. Therefore, early screening and diagnosis of breast cancer based

on ultrasound with its low cost and high efficiency play an important role in reducing the death rate of breast cancer [3]. However, ultrasound presents its unique challenges with low image quality, lack of experienced ultrasound operators and diagnosticians, the difference of ultrasound equipment and system [4]. To confront these challenges, more advanced automatic ultrasound image analysis methods have been proposed.

With the rapid development of deep learning, Convolutional Neural Networks (CNNs) have become popular and achieved desired results in breast tumor detection. In [5], several existing state-of-the-art object detection framework have been systematically evaluated on its breast ultrasound tumor datasets, including Faster R-CNN [6], SSD [7], YOLO [8]. It proved that the SSD with the input size as 300×300 achieved the best performance in terms of average precision, recall and F1-Score. In [9], a pre-trained FCN-AlexNet transfer learning method has been proposed, and its effectiveness in breast ultrasound tumor detection task has been verified. However, those data are all manually scanned by doctors, its image quality is relatively good. Therefore, it may lead to poor performance when simply applied the object detection network on other breast ultrasound data. In [10], an object detection framework is designed based on 3D convolution for the breast ultrasound data collected by the ABUS device. They got 100% and 86% sensitivity for tumor with different size on its defined statistics.

However, the above methods need a large amount of accurate annotation data which is very expensive in medical filed. Therefore, it is suitable for breast ultrasound that applying active learning to solve the problem. Active learning can select the important data from the original dataset for annotation and achieve better performance. In the deep learning eras, most of the active learning methods [11–13] remain falling into image-level classification tasks. Few methods are specified for active object detection, which faces complex instance distribution in the same images. The paper [14] simply sort the loss predictions of instance to evaluate the image uncertainly for the object detection. This paper [15] introduce spatial context to active detection and selected diverse samples according to the distances to the labeled set.

The existing breast ultrasound data and methods have certain deficiencies in breast cancer screening and diagnosis, including resource limitations in data acquisition and model performance caused by data. It cannot satisfy the strategy of early screening and diagnosis. Therefore, we collect standardized video data from 1603 cases based on AIBUS (AISONO) robots which adopt mechanical arm with US probe to realize full-automatic standardized fast scanning of breasts, and then generate at least 5 videos of each breast. Meanwhile, an algorithm framework based on the efficient EfficientDet with the reasonable dataset selected by the improved Multiple Instance Active Learning (MIAL) is proposed. Our contribution can be summarize as follows:

1. An efficient mechanism for early screening and diagnosis of breast cancer based on AIBUS video, combing standard and efficient automatic robotic arm scanning.
2. The improved MIAL active learning algorithm for obtaining diversified data from varied scenes and area of people.
3. A robust tumor detection framework based on efficient EfficientDet. Specially, the model trained on smaller dataset selected through active learning has better performance on both accuracy and speed.

2 Materials and Method

In breast ultrasound, tumors have different complexity and difficulty. The more common tumor, the less complex it is, such as the simple cyst which belongs to Birads-2 [16]. On the contrary, more complex the tumor is, the more burdensome to obtain, such as the complicated tumor with unclear borders and varies shapes. Therefore, the dataset faces a serious imbalance in simple tumors and complex tumors. If the whole samples were used for training, it will cause the model to learn more on those simple breast tumors, and it also leads to a lack of robustness for more complicated tumors. Moreover, excessive redundant data consumes time and resources in model training.

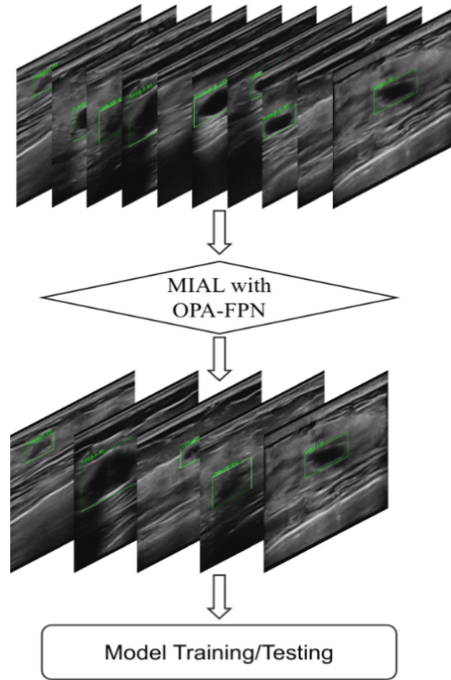


Fig. 1. Architecture diagram of the paper.

Consequently, this paper applies One-Shot Path Aggregation Feature Pyramid Network (OPA-FPN) [17] to improve the performance of MIAL [18] for better active learning in the instance level and more efficacious selection for samples with more information. Under the effect of the improved active learning, a smaller training dataset with more balanced simple and complicated tumors will be constructed. Then, a tumor detection model is trained based on EfficientDet [19]. The model can present more accurate information about tumors with different complexity in an effective and balanced manner way. The overall process is showed in Fig. 1. Each parts of our method will be described in details, including our dataset, the improved MIAL with OPA-FPN and the object detection framework named EfficientDet.

2.1 Dataset

The performance of our combined framework was mainly verified on our private dataset obtained by AIBUS robots which contains 5–8 videos with 20 FPS for a breast. The whole training datasets include 12666 breast ultrasound images from the videos of 1603 cases, which was acquired from varied areas. Meanwhile, each tumor region was labeled by two or three clinicians. For active learning, the selected data were all from the mentioned training dataset with 12666 breast ultrasound images. For detection task, the test dataset consists of 448 tumor images and 4207 normal images.

2.2 Improved MIAL with OPA-FPN

We apply the improved MIAL with OPA-FPN to select informative images for training RetinaNet detector. Compared to traditional method which using the mean of inference result directly, MIAL used the discrepancy learning and multi-instance learning (MIL) to learn and re-weight the uncertainty of instances. It also filters some negative sample instances in the inference process of RetinaNet, and select informative images from the unlabeled dataset.

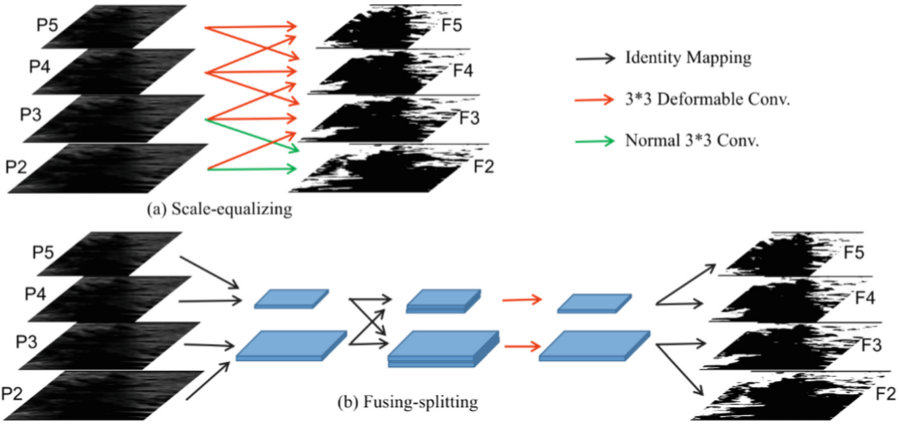


Fig. 2. Scale-equalizing and Fusing-splitting paths in OPA-FPN.

The improved MIAL selects informative images based on the sorted Top-K instances. The meaningful parameters K effects the Top-K instances uncertainty and the image uncertainty. According to the characteristics of each breast ultrasound image which contains less than 3 tumors per image, we set the parameters K as 5. Meanwhile, for the FPN structure in RetinaNet which plays an important role in feature fusion and expression at different scales, we replaced it with OPA-FPN, which is a novel search space of FPNs to fuse richer and reasonable features efficiently. The OPA-FPN contains 6 information paths as Top-down, Bottom-up, Scale-equalizing, None, Fusing-splitting, Skip-connect. The special Scale-equalizing and Fusing-splitting paths have been given in Fig. 2.

2.3 EfficientDet-Based Tumor Detection

We use EfficientDet as our tumor detector for its optional performance in terms of both speed and accuracy. For object detection, FPN structure and model depth have an important impact on model performance. Meanwhile, anchors effect the bounding box regression and selection of positive and negative samples during training. Therefore, we set the anchor ratio as [1, 1.5, 2] under the specific ultrasound dataset based on the aspect ratio distribution of the regression box. Then, compound scaling as the key contribution of EfficientDet, it proved that the scale of input resolution, model depth and width is effective. We only scale the model depth and width with the input resolution fixed because of the unique properties on ultrasound data. Specifically, we set the input resolution to $512 * 512$ and the compound scaling coefficient to 2 taking into account of the model accuracy and inference speed, and other scaling configs for EfficientDet are shown in Table 1.

Table 1. Scaling configs for EfficientDet D0-D6.

EfficientDet	Input size	Backbone network	BiFPN		Box/class layers
			channel	layer	
D0	512	B0	64	3	3
D1	512	B1	88	4	3
D2	512	B2	112	5	3
D3	640	B3	160	6	4
D4	768	B4	224	7	4
D5	896	B5	288	7	4
D6	1024	B6	384	8	5

3 Experiment

3.1 Implementation Details

The whole experiments were on the platform with two NVIDIA GeForce RTX 2080Ti GPU. In order to verify the validity of the improved MIAL, we also conducted experiments on PASCAL VOC Dataset. During the training for the improved MIAL on PASCAL VOC Dataset, we set the SGD optimizer with momentum as 0.9, learning rate as $1e-3$ and the weight decay as 0.0001. Meanwhile, the initial training dataset ratio is 0.05, the selected quantity ratio is 0.025 for each cycle. For the experiments on our specific training dataset from AIBUS robots, we set the initial learning rate as $1e-4$, the initial training dataset ratio is 0.1 and the selected quantity ratio is 0.05 for each cycle. The total cycle quantity is 8.

During the training for EfficientDet on the selected dataset from the improved MIAL, we set the SGD optimizer with momentum as 0.9, learning rate as $1e-4$. The learning

rate will be changed using dynamic learning rate reducing Strategy based on the valid dataset loss with patience as 3. Meanwhile, to implement the corresponding ablation experiment on the testing dataset based on tumor detection, we evaluate the model through the sensitive, specificity and F1-score.

3.2 The Improved MIAL Performance

To evaluate the effectiveness of our improved MIAL, we constructed the comparative experiment on the PASCAL VOC Dataset and our breast ultrasound dataset. Meanwhile, we compare our improved MIAL with random sampling, LAAL [20]. The mean average precision (mAP) is used as the metric. The result in Fig. 3 has shown the MIAL works well. It achieved the average detection accuracy of 72.3% when using the 20% samples. And it reached 93.5% of performance with whole samples in dataset.

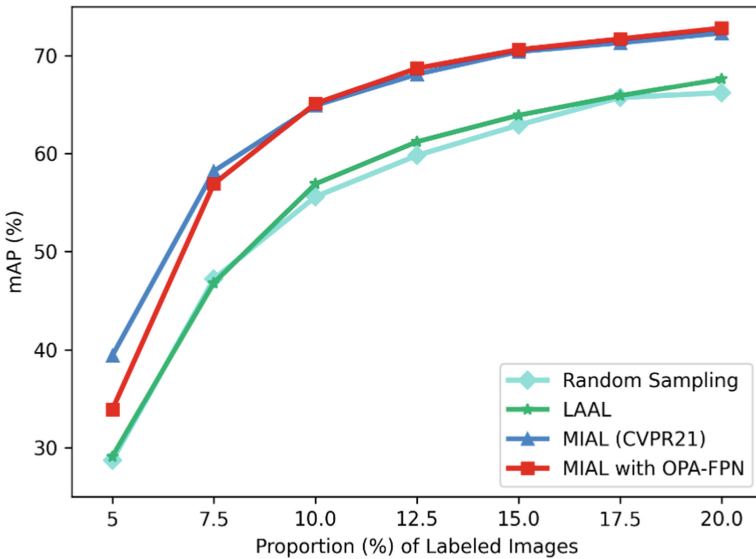


Fig. 3. Active Learning on PASCAL VOC.

Pleasantly, the improved MIAL with OPA-FPN outperformed MIAL by 0.5%. Finally, the detail results based on MIAL method on our dataset is shown in Fig. 4. It also proved the effectiveness.

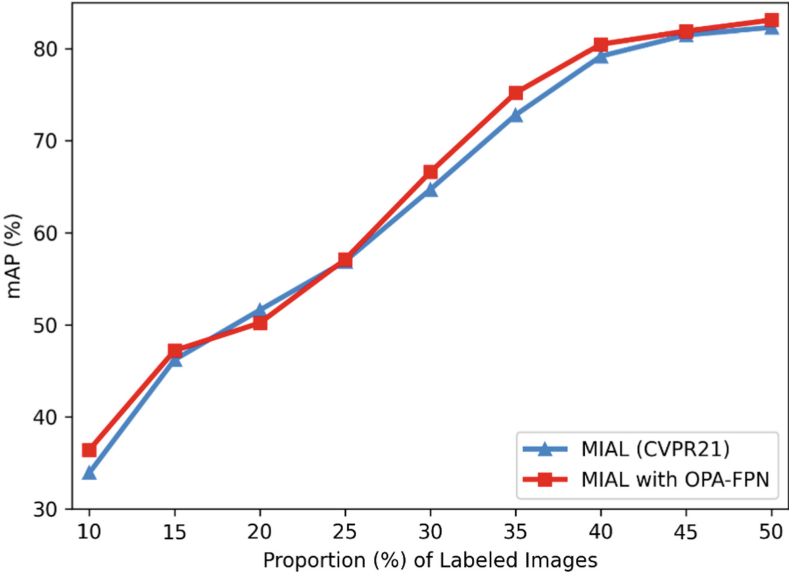


Fig. 4. Active learning on our dataset.

3.3 Tumor Detection Performance

For the tumor detection task, we use the CenterNet [21] on the whole dataset as the baseline on the metrics including the sensitive, specificity and F1-score. Meanwhile, we compared it with EfficientDet and the EfficientDet trained on the selected samples in Sect. 3.2 based on the improved MIAL. The experimental results have shown in Table 2. The number 12666 and 6333 in model name respectively represent the mounts of training set used in the corresponding model. Compared with the base model trained by CenterNet-12666, the EfficientDet performed well. In particular, the performance of sensitive and F1-score is respectively increased 0.028% and 0.009%, but the specificity is decreased 0.002%. In fact, the sensitive and F1-score is more important than the specificity in breast cancer screening tasks. Furthermore, the performance of sensitive, specificity and F1-score on EfficientDet-6333 is respectively increased 0.04%, 0.001% and 0.027% than the EfficientDet-12666. As for the model inference time on single image on our CPU, the time consumed by CenterNet and EfficientDet is 1.05s and 0.98s respectively. It is obvious that the model trained on the selected data from active learning performs well in accuracy and speed. The harder tumors that are only detected by EfficientDet have been shown in the Fig. 5.

However, there are some normal tissues which are similar with tumors on some single slices. As shown in Fig. 6, the region labeled by red box has some similarities on i th slice. However, it can be confirmed that it is fat tissue by observing the region labeled by green box on $(i + 1)$ th slice. Therefore, the number of FP may be high if only reducing the threshold of inference for high TP or using the single-slice tumor detection. In the future work, it may be possible to combine continuous slices to reduce the False Positive (FP) without increasing the False Negative (FN).

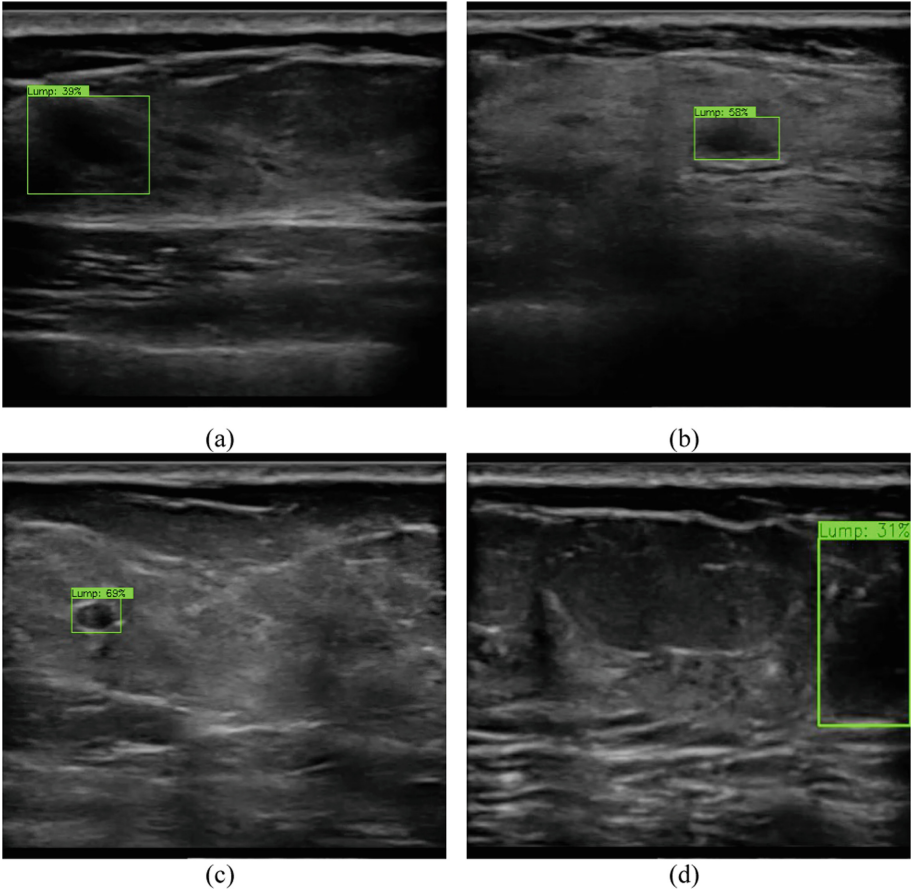


Fig. 5. Hard tumors detected by our model. The edges of tumor in (a) and (b) is blurred, and its borders is unclear. The tumor in (c) is tiny, but it has little calcification which is serious for patient. The shape of tumor in (d) is incomplete, and its edges and borders is unclear.

Table 2. Performance on different model with different training set.

Model name	Sensitive	Specificity	F1-score
CenterNet-12666	0.729	0.963	0.706
EfficientDet-12666	0.757	0.961	0.715
EfficientDet-6333	0.797	0.962	0.742

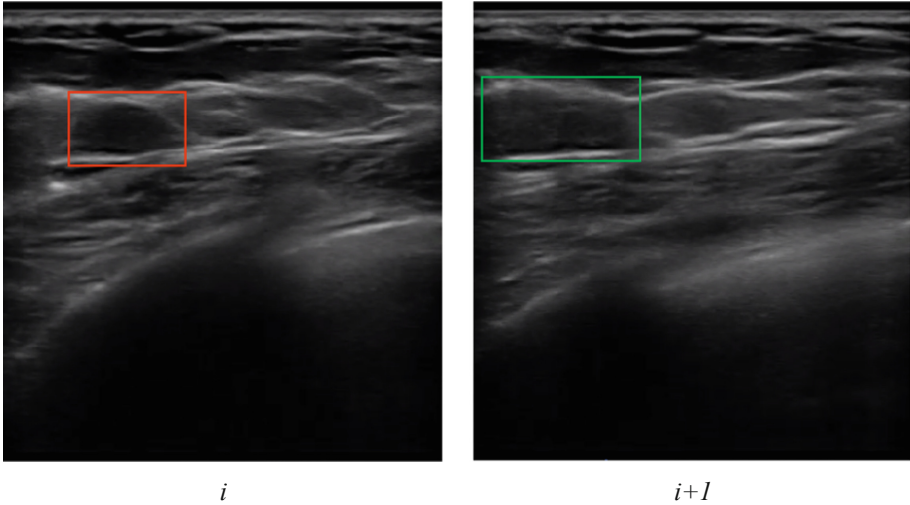


Fig. 6. Illustration of a normal fat tissue with similar features to tumor

4 Conclusion

We proposed the improved MIAL with OPA-FPN which automatically search the better FPN structure for object detection, to observe instance uncertainty. Based on the sorted uncertainty of all images in training dataset, we select the more difficult and reasonable images to create smaller training dataset. Using the selected small training dataset, we achieved more accurate performance. Meanwhile, the EfficientDet has shown it is more efficiency than CenterNet in MIAL with less model parameters. Specifically, we used the 50% of the whole training set and achieved result which exceeded the original model. Finally, the inference time of used model is 0.07s faster on our CPU.

Acknowledgments. This work was supported by Guangdong Basic and Applied Basic Research Foundation under Grant No. 2020B1515120098.

References

1. Chen, W., et al.: CA Cancer J Clin **66**(2), 115–132 (2016)
2. Shen, H.-B., Tian, J.-W., Zhou, B.-S.: China guideline for the screening and early detection of female Breast Cancer. China Cancer **30**(3), 161–191 (2021)
3. Siegel, R.L., Miller, K.D., Fuchs, H.E., et al.: Cancer statistics. Cancer J. Clin. **71**(1) (2021)
4. Brem, R.F., Lenihan, M.J., Lieberman, J., Torrente, J.: Screening breast ultrasound: past, present, and future. Am. J. Roentgenol. **204**(2), 234–240 (2015)
5. Cao, Z., Duan, L., Yang, G., et al.: An experimental study on breast lesion detection and classification from ultrasound images using deep learning architectures. BMC Med. Imag. **19**(1) (2019)
6. Ren, S., He, K., Girshick, R., et al.: Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans. Pattern Anal. Mach. Intell. **39**(6), 1137–1149 (2017)

7. Liu, W., Anguelov, D., Erhan, D., et al.: SSD: single shot multibox detector. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) *Computer Vision – ECCV 2016*. *ECCV 2016. Lecture Notes in Computer Science*, vol. 9905. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46448-0_2
8. Redmon, J., Divvala, S., Girshick, R., et al.: You only look once: unified, real-time object detection. In: *Computer Vision & Pattern Recognition*. IEEE (2016)
9. Yap, M.H., Pons, G., Marti, J., et al.: Automated Breast ultrasound lesions detection using convolutional neural networks. *IEEE J. Biomed. Health Inform.* **22**(4), 1218–1226 (2017)
10. Deeply-supervised networks with threshold loss for Cancer detection in automated Breast Ultrasound. *IEEE Trans. Med. Imag.* **39**(4), 866–876 (2020)
11. Beluch, W.H., Genewein, T., Nürnberger, A., Köhler, J.M.: The power of ensembles for active learning in image classification. In *CVPR*, pp. 9368–9377 (2018)
12. Lin, L., Wang, K., Meng, D., Zuo, W., Zhang, L.: Active self-paced learning for cost-effective and progressive face identification. *IEEE TPAMI* **40**(1), 7–19 (2018)
13. Wang, K., Zhang, D., Li, Y., Zhang, R., Lin, L.: Cost-effective active learning for deep image classification. *IEEE TCSVT* **27**(12), 2591–2600 (2017)
14. Yoo, D., Kweon, I.S.: Learning loss for active learning. IEEE (2019)
15. Agarwal, S., Arora, H., Anand, S., et al.: Contextual Diversity for Active Learning (2020)
16. Balleyguier, C., Ayadi, S., Nguyen, K.V., et al.: BIRADS classification in mammography. *Eur. J. Radiol.* **61**(2), 192–194 (2007)
17. Liang, T., Wang, Y., Tang, Z., et al.: OPANAS: one-shot path aggregation network architecture search for object detection (2021)
18. Yuan, T., et al.: Multiple instance active learning for object detection. In: *CVPR* (2021)
19. Tan, M., Pang, R., Le, Q.V.: EfficientDet: scalable and efficient object detection. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE (2020)
20. Jawahar, C.V., Li, H., Mori, G., Schindler, K. (eds.): *ACCV 2018. LNCS*, vol. 11365. Springer, Cham (2019). <https://doi.org/10.1007/978-3-030-20873-8>
21. Zhou, X., Wang, D., Krhenbühl, P.: Objects as points (2019)
22. Zhang, M., Xing, X.: Deep fuzzy hashing network for efficient image retrieval. *IEEE Trans. Fuzzy Syst.* (2020). <https://doi.org/10.1109/TFUZZ.2020.2984991>
23. Li, Y., Chen, M., et al.: Brain intelligence: go beyond artificial intelligence. *Mob. Netw. Appl.* **23**, 368–375 (2018)
24. Shenglin, M., et al.: Motor anomaly detection for unmanned aerial vehicles using reinforcement learning. *IEEE Internet Things J.* **5**(4), 2315–2322 (2018)
25. Qin, M., Zhang, F., et al.: RSCNN: a CNN-based method to enhance low-light remote-sensing images. *Remote Sens.* **62** (2020)
26. Zhang, Y., et al.: User-oriented virtual mobile network resource management for vehicle communications. *IEEE Trans. Intell. Transp. Syst.* **22**(6), 3521–3532 (2021)