• **Intelligent Ophthalmology** •

Research on classification method of high myopic maculopathy based on retinal fundus images and optimized ALFA-Mix active learning algorithm

Shao-Jun Zhu[1,2], Hao-Dong Zhan[1,2], Mao-Nian Wu[1,2], Bo Zheng[1,2], Bang-Quan Liu[3], Shao-Chong Zhang[4], Wei-Hua Yang[4]

[1]Huzhou University, School of Information Engineering, Huzhou 313000, Zhejiang Province, China
[2]Zhejiang Province Key Laboratory of Smart Management & Application of Modern Agricultural Resources, Huzhou University, Huzhou 313000, Zhejiang Province, China
[3]College of Digital Technology and Engineering, Ningbo University of Finance & Economics, Ningbo 315175, Zhejiang Province, China
[4]Shenzhen Eye Institute, Shenzhen Eye Hospital, Jinan University, Shenzhen 518048, Guangdong Province, China
Correspondence to: Wei-Hua Yang and Shao-Chong Zhang. Shenzhen Eye Institute, Shenzhen Eye Hospital, Jinan University, Shenzhen 518048, Guangdong Province, China. benben0606@139.com; zhangshaochong@gzzoc.com

Abstract

● **AIM:** To conduct a classification study of high myopic maculopathy (HMM) using limited datasets, including tessellated fundus, diffuse chorioretinal atrophy, patchy chorioretinal atrophy, and macular atrophy, and minimize annotation costs, and to optimize the ALFA-Mix active learning algorithm and apply it to HMM classification.

● **METHODS:** The optimized ALFA-Mix algorithm (ALFA-Mix+) was compared with five algorithms, including ALFA-Mix. Four models, including ResNet18, were established. Each algorithm was combined with four models for experiments on the HMM dataset. Each experiment consisted of 20 active learning rounds, with 100 images selected per round. The algorithm was evaluated by comparing the number of rounds in which ALFA-Mix+ outperformed other algorithms. Finally, this study employed six models, including EfficientFormer, to classify HMM. The best-performing model among these models was selected as the baseline model and combined with the ALFA-Mix+ algorithm to achieve satisfactory classification results with a small dataset.

● **RESULTS:** ALFA-Mix+ outperforms other algorithms with an average superiority of 16.6, 14.75, 16.8, and 16.7 rounds in terms of accuracy, sensitivity, specificity, and Kappa value, respectively. This study conducted experiments on classifying HMM using several advanced deep learning models with a complete training set of 4252 images. The EfficientFormer achieved the best results with an accuracy, sensitivity, specificity, and Kappa value of 0.8821, 0.8334, 0.9693, and 0.8339, respectively. Therefore, by combining ALFA-Mix+ with EfficientFormer, this study achieved results with an accuracy, sensitivity, specificity, and Kappa value of 0.8964, 0.8643, 0.9721, and 0.8537, respectively.

● **CONCLUSION:** The ALFA-Mix+ algorithm reduces the required samples without compromising accuracy. Compared to other algorithms, ALFA-Mix+ outperforms in more rounds of experiments. It effectively selects valuable samples compared to other algorithms. In HMM classification, combining ALFA-Mix+ with EfficientFormer enhances model performance, further demonstrating the effectiveness of ALFA-Mix+.

● **KEYWORDS:** high myopic maculopathy; deep learning; active learning; image classification; ALFA-Mix algorithm

INTRODUCTION

High myopic maculopathy (HMM) is one of the main causes of blindness[1-2]. The severity and progression rate of the disease are closely related to the degree and duration of myopia[3-4]. With the continuous increase in the proportion of high myopia patients[5], the incidence of HMM is showing an upward trend year by year. Studies have shown that the risk of developing HMM increases significantly in patients with high myopia worldwide, with a probability ranging from 13.3% to 72.7%[6-7]. For patients with HMM, there will be gradual vision loss, blurred central vision, distortion, and other symptoms,

which greatly affect their daily life and work. Therefore, prevention and early detection are of great significance for the prevention and treatment of myopic maculopathy.

In recent years, the application of artificial intelligence in the medical field has become one of the research hotspots[8]. Many scholars have conducted in-depth research[9-16]. And deep learning technology has gradually been applied to the prevention and early detection of HMM. Tan *et al*[17] developed a deep learning system based on XGboost and DenseNet to predict refractive error, high myopia, and HMM, achieving an area under curve (AUC) of 0.955 in the detection of HMM. Du *et al*[18] proposed a deep learning algorithm for classifying HMM, achieving an overall accuracy of 0.9208. These algorithms were also used to develop a pathological myopia research classification system by adding specific processing layers, which helped to screen pathological myopia. Rauf *et al*[19] used convolutional neural networks (CNNs) to automatically detect pathological myopia, and the best-performing CNN model achieved an AUC of 0.9845. Tang *et al*[20] developed a deep learning model based on ResNet50 and DeepLabV3+ using color fundus photographs to classify and segment HMM, achieving a classification accuracy of 0.937 and a segmentation F1 score of 0.95, effectively achieving the classification and monitoring of HMM progression. Sun *et al*[21] introduced a prior knowledge extraction module to extract prior knowledge to determine the rough lesion areas in fundus images, and integrated the obtained prior knowledge into the deep learning network, achieving an accuracy of 0.8921 in the five-classification experiment of HMM. To achieve good performance, deep learning training often requires a large amount of labeled dataset. Currently, the scale of HMM datasets is small, and annotators are often required to have high professional quality and experience, making annotation costs very high[22]. Therefore, it is difficult to obtain large datasets for neural network learning. To achieve good results with a small amount of data, a series of methods and strategies have been proposed, among which one strategy is active learning[23].

Active learning can improve the performance of deep learning models when labeled data is limited. It dynamically selects the most valuable samples to be labeled during the deep learning training process, thereby reducing the amount and cost of manual labeling data. Currently, active learning can be roughly divided into three categories: uncertainty-based, diversity-based, and hybrid active learning algorithms[24]. Uncertainty-based[25-26] active learning selects data that is difficult to distinguish in the model for labeling, thus achieving the ability to improve the model's performance. Uncertainty-based active learning methods are easy to adapt to various tasks, but they may not perform well in extremely imbalanced sample situations due to considering only the information content of the sample itself. At the same time, the selectors used in uncertainty-based sampling strategies are often shallow and may not work well on complex datasets. Diversity-based[27-28] query strategies select representative sample sets, which once selected can represent the entire dataset. However, a single query strategy may cause sampling bias and easily query outliers, making it difficult to select the most valuable sample set. Hybrid active learning algorithms[29-30] consider both uncertainty and diversity of samples, making it easier to select valuable samples. The development of active learning has also promoted its application in the medical field. Smit *et al*[31] combined active learning with meta-learning for selective annotation in medical image interpretation. The deep learning selector compared image embeddings obtained from pre-training to determine which images needed to be labeled and classified unlabeled images using cosine similarity. Shi *et al*[32] designed dual criteria for selecting informative samples. They proposed an active learning method for skin lesion analysis, which belongs to the category of post-labeling enhancement. The framework consists of sample selection and sample aggregation. In order to effectively use the selected samples, they designed an aggregation strategy by adding intra-class images in the pixel space, to capture more rich and clear features from these valuable but unclear samples. Zhang *et al*[22] used virtual adversarial perturbation and model density-aware entropy to find informative samples as labeling candidates and designed a balanced class selector to reduce redundancy in sample selection.

This study builds upon the ALFA-Mix[33] active learning algorithm and proposes an improved version, called ALFA-Mix+, which is integrated with a deep learning model for automatic classification of HMM. The application of ALFA-Mix+ is expected to reduce annotation costs and facilitate the detection and screening of HMM.

MATERIALS AND METHODS

Data Source This study conducted experiments using an HMM dataset provided by a collaborating hospital. The study adhered to the principles of the Helsinki Declaration. The Medical Ethics Committee of Shenzhen Eye Hospital has approved this study with approval number 2022KYPJO45. To protect the privacy of the data providers, all data was de-identified prior to inclusion in the study. The specific details of the HMM dataset are shown in Table 1.

ALFA-Mix Algorithm The ALFA-Mix algorithm is a concise and effective active learning algorithm, which consists of two main stages. In the first stage, the algorithm utilizes the feature extraction module of a model to compute the average features for each labeled category and extract features from the unlabeled images. The unlabeled images are then mixed with the average features of each labeled category. Subsequently,

Table 1 Dataset of high myopic maculopathy

Category	Training set	Test set
Non-myopic retinopathy	585	252
Tessellated fundus	1950	836
Diffuse chorioretinal atrophy	774	332
Patchy chorioretinal atrophy	646	278
Macular atrophy	297	128
Total	4252	1826

predictions are made for both the unlabeled and mixed images, and samples with different prediction results are selected. In the second stage, the algorithm employs the K-means clustering algorithm to identify the most representative samples from the selected samples in the first stage, which are then submitted to experts for annotation.

ALFA-Mix+ Algorithm In the early stages of training, the ALFA-Mix algorithm may encounter a situation where a small number of selected samples can result in a significant number of inconsistent classification results before and after feature mixing. At this time, it is easy to ignore some samples that are difficult to distinguish but more helpful for improving the model performance during diversity selection. Therefore, in this study, we introduce the gradient norm to represent the selected samples. In the second stage, the algorithm modifies the sample selection process by computing gradients. Instead of selecting samples with significant differences, the algorithm now chooses samples that can produce diverse gradient vector directions. The sensitivity and confidence of the model towards the input data are represented by calculating the gradient magnitude of the model's loss function with respect to the output. If the gradient value is large, the sample is considered to have higher uncertainty. The gradient calculation formula is shown in Equation 1.

$$grad = \frac{\partial\left(-\sum_{x=1}^{n} p(x)\log_2 q(x)\right)}{\partial W_{last}} \quad (1)$$

This study utilizes pseudo-labels predicted by the model for gradient computation. In the above equation, $p(x)$ represents the probability distribution of pseudo-labels, $q(x)$ represents the probability distribution predicted by the model, and W_{last} represents the parameters of the last layer of the model. When selecting representative samples using clustering algorithms, the gradient vectors of images will be used for clustering. Samples that can generate different gradient vector directions to update model parameters will be selected instead of samples with different features. The flowchart of the ALFA-Mix+ algorithm is shown in Figure 1.

The steps of the ALFA-Mix+ algorithm are shown as follows:
1) Extract the features of all training images using the feature extraction module of a deep model.
2) Calculate the average features of each labeled image category.
3) Interpolate the features of the unlabeled images with the labeled images' average features using the interpolation formula as shown in equation 2, where the interpolation ratio $\alpha \in [0,1)$.

$$\lambda_{mix} = \alpha\lambda_{class} + (1-\alpha)\lambda_{unlabel} \quad (2)$$

In the above equation, λ_{class} represents the average feature of each class in step 2), $\lambda_{unlabel}$ represents the feature of the unlabeled image, and λ_{mix} represents the mixed feature. For the selection of α, in order to ensure the effectiveness of interpolation, α is chosen to maximize the loss of the interpolation point. The loss of the interpolation point is calculated by the formula in Equation 3).

$$l\left(f_c(\lambda_{mix}), \hat{y}\right) - l\left(f_c(\lambda_{unlabel}), \hat{y}\right) \approx \left(\alpha\left(\lambda_{class} - \lambda_{unlabel}\right)\right)^{T} \cdot \nabla \lambda_{unlabel}$$
$$l\left(f_c(\lambda_{unlabel}), \hat{y}\right) \quad (3)$$

In the equation above, $\hat{y}$ represents the pseudo-labels predicted by the current model, and f_c is the classifier of the model.
4) Classify the mixed features obtained in step 3) and the features of unlabelled images using the classifier, and select the samples with inconsistent classification results.
5) Use gradients to represent the selected samples.
6) Use the Kmeans algorithm to annotate the selected m samples.
7) Add the labeled samples to the training set.
Repeatedly execute steps 1) to 7) until a stopping criterion is met.

Statistical Analysis In this study, active learning algorithms were combined with models in pairs for experiments. Each experiment was set to have 20 active learning rounds, and 100 images were selected for annotation in each round. The algorithm was evaluated by comparing ALFA-Mix+'s accuracy, sensitivity, specificity, Kappa value, and the number of rounds where it outperformed other algorithms. In the classification of HMM, accuracy, sensitivity, specificity, and Kappa values were used to evaluate each classification model.

RESULTS

Comparison Experiment of Active Learning Algorithms
This study employed five active learning algorithms, including ALFA-Mix, CoreSet[34], BALD[35], EntropySampling[36], and CDAL[37], as controls, and conducted experiments on four models: ResNet18, DenseNet169, VGG19, and AlexNet. Each experiment was set with 20 active learning rounds. The experimental results are shown in Figures 2–5, and Tables 2–6 displays the performance of each model under different amounts of data. Accuracy, sensitivity, specificity, and Kappa
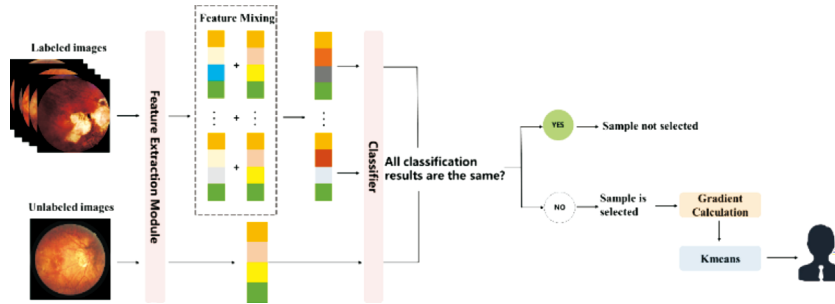
Figure 1 Flowchart of the ALFA-Mix+ algorithm.

coefficient are commonly used evaluation metrics. A higher value for these metrics indicates better model performance.

Figure 2 displays the number of active learning rounds in which the accuracy results of ALFA-Mix+ exceeded those of the control algorithms within 20 rounds of active learning.

The accuracy results of different active learning algorithms on different models under different sample sizes are shown in Table 2.

Figure 3 displays the number of active learning rounds in which the sensitivity results of ALFA-Mix+ exceeded those of the control algorithms within 20 rounds of active learning.

The sensitivity results of different active learning algorithms on different models under different sample sizes are shown in Table 3.

Figure 4 displays the number of active learning rounds in which the specificity results of ALFA-Mix+ exceeded those of the control algorithms within 20 rounds of active learning.

The specificity results of different active learning algorithms on different models under different sample sizes are shown in Table 4.

Figure 5 shows the number of active learning rounds in which the Kappa values of ALFA-Mix+ exceeded those of the control algorithms within 20 rounds of active learning.

The Kappa value of different active learning algorithms on different models under different sample sizes are shown in Table 5.

According to the experimental results, using active learning algorithms can significantly improve the model performance, with good results achieved by using only 50% of the original dataset. In addition, increasing the amount of data can improve the model's accuracy, specificity, sensitivity, and Kappa value. However, as the amount of data increases, the growth rate of these indicators gradually slows down. This suggests that increasing the number of samples can improve model performance, but the growth rate may not increase significantly, and unnecessary labeling errors may sometimes occur. The ALFA-Mix+ proposed in this study compared with other active learning algorithms can achieve better results in most experiments when the same training samples are used, demonstrating the effectiveness of this research method.
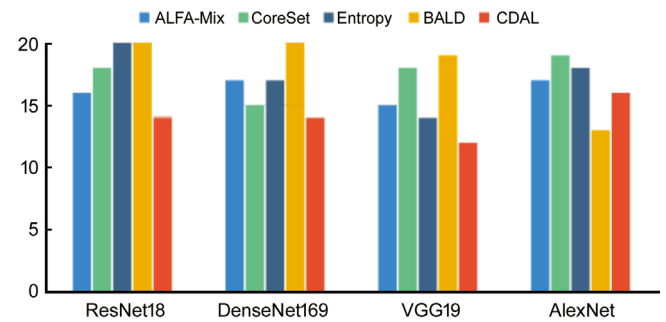

Figure 2 Number of active learning rounds in which ALFA-Mix+ outperformed the corresponding control algorithm in terms of accuracy within 20 rounds of active learning.
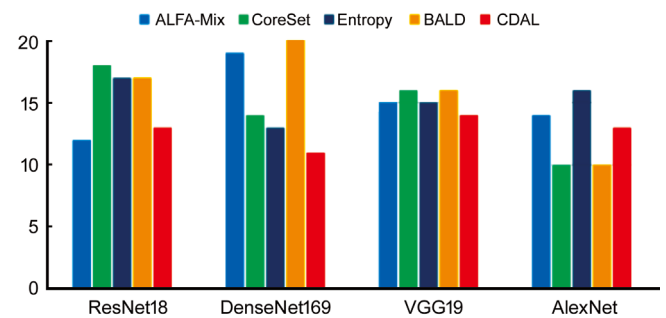

Figure 3 Number of active learning rounds in which ALFA-Mix+ outperformed the corresponding control algorithm in terms of sensitivity within 20 rounds of active learning.
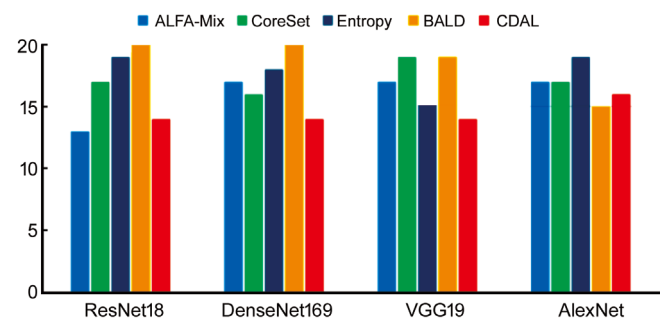

Figure 4 Number of active learning rounds in which ALFA-Mix+ outperformed the corresponding control algorithm in terms of specificity within 20 rounds of active learning.

The confusion matrices obtained by each active learning algorithm on the four models are shown in Figures 6–8. We selected the best-performing rounds of each algorithm in their respective experiments for calculation.

Table 2 Accuracy results of each algorithm on ResNet18

Model	Data (%)	ALFA-Mix	CoreSet	BALD	EntropySampling	CDAL	Ours
ResNet18	10	0.7200	0.7215	0.7227	0.7118	0.7258	0.7311[a]
	20	0.7725	0.7627	0.7447	0.7737	0.7765	0.7775[a]
	30	0.7918	0.7858	0.7611	0.7742	0.7922	0.7945[a]
	40	0.8069	0.8071	0.7600	0.7907	0.8049	0.8077[a]
	50	0.7953	0.8005	0.7995	0.8066	0.8153[a]	0.8126
DenseNet169	10	0.7458	0.7244	0.7468	0.7123	0.7466	0.7534[a]
	20	0.7918	0.7803	0.7721	0.7923	0.7874	0.7995[a]
	30	0.8049	0.8005	0.7819	0.7929	0.8099	0.8137[a]
	40	0.8138	0.8186	0.7868	0.8351[a]	0.8049	0.8142
	50	0.8301	0.8285	0.7989	0.8356	0.8378	0.8433[a]
VGG19	10	0.7277	0.6877	0.7321	0.7216	0.7277	0.7392[a]
	20	0.7814	0.7715	0.7847	0.7715	0.7945	0.7991[a]
	30	0.8060[a]	0.8011	0.7764	0.7945	0.8025	0.8038
	40	0.8016	0.8027	0.8082	0.8103	0.8011	0.8159[a]
	50	0.8247	0.8286	0.8137	0.8279	0.8153	0.8310[a]
AlexNet	10	0.7463	0.7501	0.7688	0.7436	0.7666	0.7748[a]
	20	0.7978	0.8099	0.8099	0.7940	0.8055	0.8159[a]
	30	0.8241	0.8186	0.8335	0.8225	0.8323	0.8340[a]
	40	0.8411	0.8214	0.8504	0.8405	0.8367	0.8521[a]
	50	0.8445	0.8427	0.8526[a]	0.8384	0.8405	0.8449

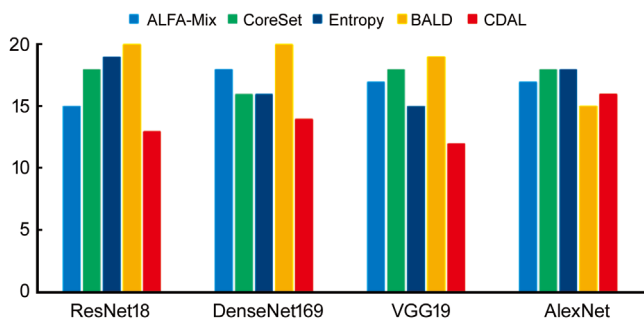[a]Superior performance of the algorithm.

Figure 5 The number of times ALFA-Mix+ outperformed the control algorithm in terms of Kappa values within 20 rounds of active learning.

Experiment on HMM Classification Using ALFA-Mix+ We implemented the classification of HMM using EfficientFormer[38], EfficientNetV2[39], VisionTransformer[40], DenseNet201[41], ResNet152[42], and SwinTransformer[43]. Among these six models, DenseNet201, EfficientNetV2, and ResNet152 belong to deep convolutional neural networks, while EfficientFormer, VisionTransformer, and SwinTransformer use neural network structures based on attention mechanisms. We evaluated the performance of these six models using accuracy, sensitivity, specificity, and Kappa values. The performance of the 6 models on the HMM dataset is shown in Table 6.

Based on the comprehensive evaluation of various metrics, this study combined EfficientFormer with ALFA-Mix+ to achieve satisfactory classification results for HMM with a small amount of data. EfficientFormer is an efficient deep learning model composed of 4D convolutional neural network modules and 3D Transformer modules.

After combining with ALFA-Mix+, EfficientFormer achieved a classification accuracy of 0.8964, sensitivity of 0.8643, specificity of 0.9721, and Kappa value of 0.8537 on the dataset of HMM. It is evident that the performance of EfficientFormer was improved with the use of ALFA-Mix+ algorithm. The confusion matrix of EfficientFormer after combining with ALFA-Mix+ is shown in Figure 9.

DISCUSSION

The incidence of HMM has been increasing year by year. The annotation of HMM requires high accuracy and often requires annotators with high professional qualifications and experience. Therefore, the annotation process requires a great deal of cost. Using active learning algorithms can effectively select samples with high information content and representativeness in the dataset, which reduces annotation costs.

This study constructed an HMM classification model based on the ALFA-Mix+ algorithm. To fully demonstrate the effectiveness and advantages of the active learning algorithm, four classic classification models were used for experiments. As the number of data samples increased, the accuracy, specificity, sensitivity, and Kappa value of the model also gradually increased, but the rate of increase gradually slowed down. This indicates that as the number of samples increases, the number of new features learned by the model decreases,

Table 3 Sensitivity results of each algorithm on ResNet18

Model	Data (%)	ALFA-Mix	CoreSet	BALD	EntropySampling	CDAL	Ours
ResNet18	10	0.6400	0.6259	0.6209	0.6373	0.6439	0.6479[a]
	20	0.7174[a]	0.6736	0.6662	0.7098	0.7007	0.6795
	30	0.7268	0.6879	0.6856	0.7007	0.7354[a]	0.7013
	40	0.7426	0.7314	0.6769	0.7147	0.7366	0.7454[a]
	50	0.7424	0.7424	0.7407	0.7182	0.7449	0.7534[a]
DenseNet169	10	0.6630	0.6585	0.6547	0.6515	0.6781	0.6823[a]
	20	0.6908	0.7181	0.6626	0.7209	0.6983	0.7219[a]
	30	0.7117	0.7386	0.6815	0.7224	0.7282	0.7523[a]
	40	0.7203	0.7471	0.6948	0.7768[a]	0.7594	0.7208
	50	0.7483	0.7609	0.7227	0.7668	0.7668	0.7862[a]
VGG19	10	0.6229	0.6252	0.6351	0.6317	0.6518	0.6708[a]
	20	0.7252	0.7515[a]	0.7132	0.6766	0.7286	0.7192
	30	0.7382	0.7409	0.7179	0.7246	0.7407	0.7439[a]
	40	0.7258	0.7596	0.7260	0.7660	0.7495	0.7719[a]
	50	0.7782	0.7804	0.7639	0.7764	0.7441	0.7836[a]
AlexNet	10	0.6391	0.7000[a]	0.6976	0.6833	0.6824	0.6834
	20	0.7089	0.7487	0.7415	0.7155	0.7328	0.7538[a]
	30	0.7640	0.7776	0.7816	0.7509	0.7855[a]	0.7662
	40	0.7919	0.7581	0.7939	0.7774	0.7896	0.7942[a]
	50	0.7819	0.7845	0.7872	0.7956	0.7976	0.8057[a]

[a]Superior performance of the algorithm.

Table 4 Specificity results of each algorithm on ResNet18

Model	Data (%)	ALFA-Mix	CoreSet	BALD	EntropySampling	CDAL	Ours
ResNet18	10	0.9227	0.9212	0.9170	0.9207	0.9261	0.9279[a]
	20	0.9394[a]	0.9343	0.9273	0.9382	0.9375	0.9349
	30	0.9453	0.9415	0.9319	0.9391	0.9466[a]	0.9396
	40	0.9484[a]	0.9476	0.9305	0.9444	0.9460	0.9484[a]
	50	0.9456	0.9438	0.9449	0.9433	0.9431	0.9483[a]
DenseNet169	10	0.9317	0.9266	0.9277	0.9239	0.9339	0.9345[a]
	20	0.9427	0.9404	0.9358	0.9439	0.9440	0.9465[a]
	30	0.9487	0.9451	0.9401	0.9452	0.9496	0.9499[a]
	40	0.9505	0.9510	0.9402	0.9546[a]	0.9488	0.9510
	50	0.9549	0.9544	0.9463	0.9520	0.9556	0.9568[a]
VGG19	10	0.9219	0.9128	0.9238	0.9224	0.9231	0.9273[a]
	20	0.9419	0.9396	0.9415	0.9375	0.9419	0.9421[a]
	30	0.9464	0.9438	0.9409	0.9428	0.9464	0.9466[a]
	40	0.9476	0.9463	0.9486	0.9503[a]	0.9476	0.9497
	50	0.9541	0.9534	0.9513	0.9513	0.9512	0.9553[a]
AlexNet	10	0.9296	0.9328	0.9343	0.9330	0.9357	0.9385[a]
	20	0.9452	0.9503[a]	0.9476	0.9432	0.9484	0.9486
	30	0.9542	0.9536	0.9554[a]	0.9522	0.9546	0.9554[a]
	40	0.9586	0.9526	0.9589	0.9572	0.9579	0.9592[a]
	50	0.9584	0.9576	0.9554	0.9577	0.9583	0.9586[a]

[a]Superior performance of the algorithm.

resulting in limited or no improvement in performance after adding samples. This situation may be because new samples do not contain too much new information or because the model has already learned and captured the potential features of the data well after a certain number of samples. Therefore, further increasing the sample size may have limited performance improvement on the model and may not significantly improve the classification results. The experimental results show that the ALFA-Mix+ algorithm can effectively select high-quality samples without sacrificing the final accuracy and reduce

Table 5 Kappa value results of various algorithms on ResNet18

Model	Data (%)	ALFA-Mix	CoreSet	BALD	EntropySampling	CDAL	Ours
ResNet18	10	0.6012	0.6095	0.5925	0.5915	0.6206[a]	0.6014
	20	0.6894[a]	0.6615	0.6308	0.6791	0.6812	0.6765
	30	0.7145	0.6949	0.6549	0.6802	0.7142	0.7213[a]
	40	0.7202	0.7259	0.6512	0.7041	0.7217	0.7276[a]
	50	0.7235	0.7332	0.7064	0.7084	0.7300	0.7383[a]
DenseNet169	10	0.6055	0.5577	0.6133	0.6013	0.6108	0.6285[a]
	20	0.6928	0.6824	0.6944	0.6743	0.7058	0.7090[a]
	30	0.7248	0.7159	0.6854	0.7065	0.7231	0.7258[a]
	40	0.7207	0.7214	0.7277	0.7401[a]	0.7208	0.7389
	50	0.7506	0.7481	0.7363	0.7457	0.7385	0.7570[a]
VGG19	10	0.6055	0.5577	0.6133	0.6013	0.6108	0.6285[a]
	20	0.6928	0.6824	0.6944	0.6743	0.7058	0.7090[a]
	30	0.7248	0.7159	0.6854	0.7065	0.7231	0.7258[a]
	40	0.7207	0.7214	0.7277	0.7401[a]	0.7208	0.7389
	50	0.7506	0.7481	0.7363	0.7457	0.7385	0.7570[a]
AlexNet	10	0.6357	0.6485	0.6676	0.6412	0.6674	0.6788[a]
	20	0.7113	0.7326	0.7317	0.7055	0.7255	0.7360[a]
	30	0.7528	0.7477	0.7660	0.7479	0.7629	0.7675[a]
	40	0.7770	0.7478	0.7874	0.7739	0.7710	0.7894[a]
	50	0.7803	0.7770	0.7820	0.7734	0.7771	0.7861[a]

[a]Superior performance of the algorithm.

Table 6 Performance of different models on the HMM dataset

Model	Accuracy	Sensitivity	Specificity	Kappa
EfficientNetV2	0.7989	0.7154	0.9432	0.7107
VisionTransformer	0.8121	0.7392	0.9493	0.7338
DenseNet201	0.8367	0.7593	0.9543	0.7661
ResNet152	0.7874	0.6966	0.942	0.6968
SwinTransformer	0.7688	0.6634	0.9344	0.666
EfficientFormer	0.8821	0.8334	0.9693	0.8339
EfficientFormer+ALFM+	0.8964[a]	0.8643[a]	0.9721[a]	0.8537[a]

[a]Superior performance of the algorithm. HMM: High myopic maculopathy.

annotation costs. To classify HMM more accurately, this study finally used six models that performed well in other image classification tasks for experiments. From the experimental results, it can be seen that EfficientFormer achieved better performance on the HMM dataset and obtained more excellent results after applying the ALFA-Mix+ algorithm to this model. This further proves the effectiveness of the ALFA-Mix+ algorithm.

By comparing the experimental results with other active learning algorithms, it can be clearly seen that the algorithm proposed in this study is significantly superior in performance to those that only consider uncertainty or diversity. This indicates that the proposed algorithm can better balance uncertainty and diversity when selecting samples, thereby improving classification accuracy. However, compared to algorithms that consider both, the method of selecting diverse samples in this study has some shortcomings, which results in not fully utilizing the diversity information of the samples. Therefore, in future research, more effective sample selection strategies will be designed to fully utilize the diversity information in the data and improve classification performance. Currently, some scholars have conducted research on the detection of HMM. Sogawa et al[44] applied deep learning technology to ophthalmic disease screening and proposed a new CNN classification model. The model performed well on classification accuracy, sensitivity, specificity, and AUC indicators, and the accuracy of diagnosis for different types of myopic eye diseases was above 77.9%, with an average accuracy of 88.9%. It can help prevent blindness, make disease screening more automated and efficient. Li et al[45] achieved higher classification accuracy than professional doctors in four disease classification tasks using Focal Loss and ImageData Generator techniques, and achieved automatic classification of OCT fundus macular scan images. This method has high accuracy and robustness and can effectively eliminate operator subjectivity. In addition, the study evaluated the model's performance using indicators such as the receiver operating characteristic (ROC) curve, demonstrating its high diagnostic accuracy and good clinical application prospects. Wang et al[46] used deep learning technology to screen high myopia patients with maculopathy through color fundus photographs, dividing all color fundus photographs into four categories: normal
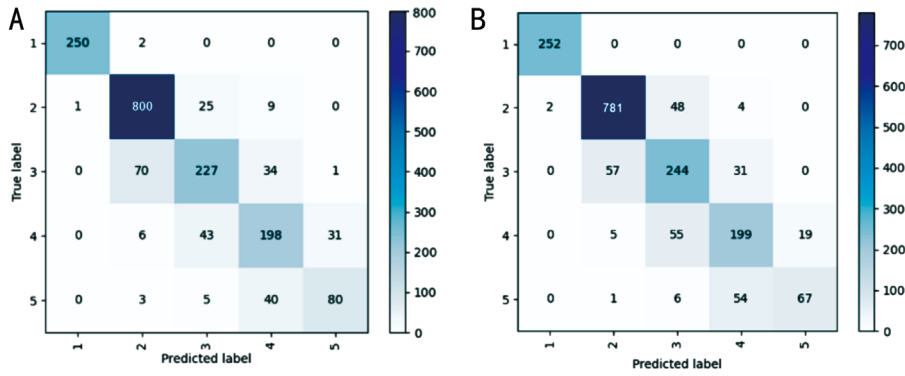
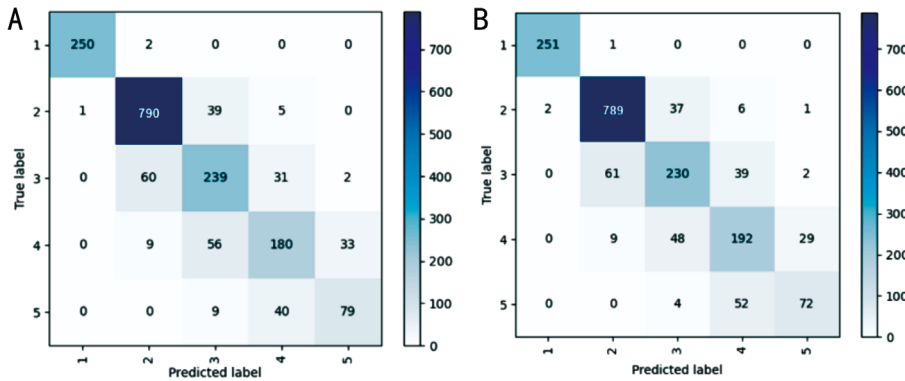**Figure 6 Confusion matrices for ALFA-Mix+ (A) and ALFA-Mix (B) on AlexNet.**

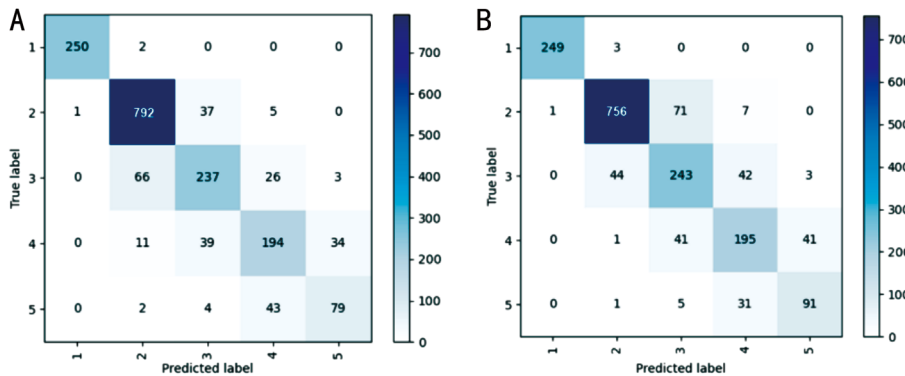**Figure 7 Confusion matrices for CoreSet (A) and EntropySampling (B) on AlexNet.**

Figure 8 Confusion matrices for BALD (A) and CDAL (B) on AlexNet.
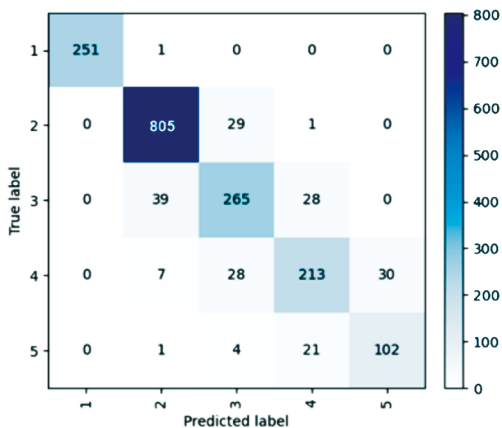

Figure 9 Confusion matrix of EfficientFormer after combining with ALFA-Mix+ algorithm.

or mild snowflake fundus, severe snowflake fundus, early pathological myopia, and late pathological myopia. Its accuracy

and recall AUC reached 0.922 and 0.781, respectively. The above experiments have achieved good results, but these methods all require a large amount of annotated dataset. This study combined active learning to achieve the classification of HMM, which improved classification accuracy with a small amount of annotated data, and reduced the cost and quantity of annotated data, providing convenience and support for the analysis and diagnosis of HMM.

Regarding the limitations and future research directions of the experiment, it is believed that the algorithm proposed in this study can still be optimized and improved in practical applications to further enhance its ability to select high-quality samples. In addition, in order to further improve the practicality of the algorithm, issues such as scalability and computational efficiency need to be considered. In future research, image segmentation will also be included, and active

learning algorithms will be applied to image segmentation.

In conclusion, this study proposes an ALFA-Mix+ active learning algorithm, which builds upon the ALFA-Mix algorithm by incorporating the images selected in the first stage of the gradient representation algorithm. This allows the algorithm to select samples in the second stage that can produce different gradient vector directions, taking into account the uncertainty and diversity of the selected samples. Experimental results showed that the proposed ALFA-Mix+ algorithm can reduce the number of training samples required without sacrificing accuracy. Compared with other active learning algorithms, the ALFA-Mix+ algorithm performed better in more experiments with the same number of rounds, and could more effectively select valuable samples than other active learning algorithms. When combined with EfficientFormer in the task of HMM classification, ALFA-Mix+ algorithm improved the model performance and further demonstrated the effectiveness of the ALFA-Mix+ algorithm.

REFERENCES

1 GBD 2019 Blindness and Vision Impairment Collaborators; Vision Loss Expert Group of the Global Burden of Disease Study. Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the Right to Sight: an analysis for the Global Burden of Disease Study. *Lancet Glob Health* 2021;9(2):e144-e160.

2 Chen YX, Han XT, Gordon I, Safi S, Lingham G, Evans J, Li JY, He MG, Keel S. A systematic review of clinical practice guidelines for myopic macular degeneration. *J Glob Health* 2022;12:04026.

3 Wong YL, Sabanayagam C, Ding Y, *et al*. Prevalence, risk factors, and impact of myopic macular degeneration on visual impairment and functioning among adults in Singapore. *Invest Ophthalmol Vis Sci* 2018;59(11):4603.

4 Choudhury F, Meuer SM, Klein R, *et al*. Prevalence and characteristics of myopic degeneration in an adult Chinese American population: the Chinese American eye study. *Am J Ophthalmol* 2018;187:34-42.

5 WHO. World report on vision. 2019. https://www.who.int/publications/i/item/9789241516570. html

6 Liu HH, Xu L, Wang YX, Wang S, You QS, Jonas JB. Prevalence and progression of myopic retinopathy in Chinese adults: the Beijing Eye Study. *Ophthalmology* 2010;117(9):1763-1768.

7 Wong YL, Zhu XJ, Tham YC, *et al*. Prevalence and predictors of myopic macular degeneration among Asian adults: pooled analysis from the Asian Eye Epidemiology Consortium. *Br J Ophthalmol* 2021;105(8):1140-1148.

8 Cao YT, Pan YL, Lu YL, *et al*. Artificial intelligence improves accuracy, efficiency, and reliability of a handheld infrared eccentric autorefractor for adult refractometry. *Int J Ophthalmol* 2022;15(4):628-634.

9 Zheng B, Jiang Q, Lu B, He K, Wu MN, Hao XL, Zhou HX, Zhu SJ, Yang WH. Five-category intelligent auxiliary diagnosis model of common fundus diseases based on fundus images. *Trans Vis Sci Tech* 2021;10(7):20.

10 Zhu SJ, Lu B, Wang CH, *et al*. Screening of common retinal diseases using six-category models based on EfficientNet. *Front Med* 2022;9:808402.

11 Zheng B, Shen YF, Luo YX, *et al*. Automated measurement of the disc-fovea angle based on DeepLabv3+. *Front Neurol* 2022;13:949805.

12 Yang WH, Zheng B, Wu MN, Zhu SJ, Fei FQ, Weng M, Zhang X, Lu PR. An evaluation system of fundus photograph-based intelligent diagnostic technology for diabetic retinopathy and applicability for research. *Diabetes Ther* 2019;10(5):1811-1822.

13 Wan C, Chen YS, Li H, Zheng B, Chen N, Yang WH, Wang CH, Li Y. EAD-net: a novel lesion segmentation method in diabetic retinopathy using neural networks. *Dis Markers* 2021;2021:1-13.

14 Wu MN, Lu Y, Hong XQ, Zhang J, Zheng B, Zhu SJ, Chen NM, Zhu ZT, Yang WH. Classification of dry and wet macular degeneration based on the ConvNeXT model. *Front Comput Neurosci* 2022;16:1079155.

15 Xu JG, Yang WH, Wan C, Shen JX. Weakly supervised detection of central serous chorioretinopathy based on local binary patterns and discrete wavelet transform. *Comput Biol Med* 2020;127:104056.

16 Zhu SJ, Fang XW, Zheng B, Wu MN, Yang WH. Research on segmentation of pterygium lesions based on convolutional neural networks. *Guoji Yanke Zazhi(Int Eye Sci)* 2022;22(6):1016-1019.

17 Tan TE, Ting DSW, Liu Y, *et al*. Artificial intelligence using a deep learning system with transfer learning to predict refractive error and myopic macular degeneration from color fundus photographs. *Invest Ophthalmol Vis Sci* 2019;60(9):1478.

18 Du R, Xie S, Fang Y, *et al*. Deep learning approach for automated detection of myopic maculopathy and pathologic myopia in fundus images. *Ophthalmol Retina* 2021;5(12):1235-1244.

19 Rauf N, Gilani SO, Waris A. Automatic detection of pathological myopia using machine learning. *Sci Rep* 2021;11(1):16570.

20 Tang J, Yuan MZ, Tian KB, *et al*. An artificial-intelligence-based automated grading and lesions segmentation system for myopic maculopathy based on color fundus photographs. *Trans Vis Sci Tech* 2022;11(6):16.

21 Sun Y, Li Y, Zhang FJ, Zhao H, Liu HR, Wang NL, Li HQ. A deep network using coarse clinical prior for myopic maculopathy grading. *Comput Biol Med* 2023;154:106556.

22 Zhang WQ, Zhu L, Hallinan J, Zhang SY, Makmur A, Cai QP, Ooi BC. BoostMIS: boosting medical image semi-supervised learning with adaptive pseudo labeling and informative active annotation. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 2022:20634-20644,

23 Ren P, Xiao Y, Chang X, *et al*. A survey of deep active learning. *ACM Computing Surveys (csur)* 2021;54(9):1-40.

24 Herde M, Huseljic D, Sick B, Calma A. A survey on cost types, interaction schemes, and annotator performance models in selection algorithms for active learning in classification. *IEEE Access* 2021;9:166970-166989.

25 Ash JT, Zhang C, Akshay Krishnamurthy A, Langford J, Agarwal A. Deep batch active learning by diverse, uncertain gradient lower bounds. *Eighth International Conference on Learning Representations (ICLR)* 2020. https://www.microsoft.com/en-us/research/publication/deep-batch-active-learning-by-diverse-uncertain-gradient-lower-bounds/

26 Wan CB, Jin FS, Qiao Z, Zhang WW, Yuan Y. Unsupervised active learning with loss prediction. *Neural Comput Appl* 2023;35(5):3587-3595.

27 Citovsky G, DeSalvo G, Gentile C, *et al*. Batch active learning at scale. Neural Information Processing Systems (NeurIPS), 2021. http://www.sanjivk.com/BatchActiveLearning_NeurIPS21.pdf

28 Hacohen G, Dekel A, Weinshall D. Active learning on a budget: opposite strategies suit high and low budgets. *arXiv e-prints* 2022.

29 Gao MF, Zhang ZZ, Yu G, Arık SÖ, Davis LS, Pfister T. Consistency-based semi-supervised active learning: towards minimizing labeling cost. 2020. https://arxiv.org/pdf/1910.07153.pdf

30 Tya B, Nla B, Han S. Self-paced active learning for deep CNNs via effective loss function. *Neurocomputing* 2021;424:1-8.

31 Smit A, Vrabac D, He Y, Ng AY, Beam AL, Rajpurkar P. MedSelect: selective labeling for medical image classification combining meta-learning with deep reinforcement learning. 2021. https://arxiv.org/pdf/2103.14339.pdf

32 Shi XY, Dou Q, Xue C, Qin J, Chen H, Heng PA. An active learning approach for reducing annotation cost in skin lesion analysis. *Lecture Notes in Computer Science* 2019:628-636.

33 Parvaneh A, Abbasnejad E, Teney D, Haffari R, Van Den Hengel A, Shi JQ. Active learning by feature mixing. 2022. https://arxiv.org/pdf/2203.07034.pdf

34 Sener O, Savarese S. Active learning for convolutional neural networks: a core-set approach. 2017. https://arxiv.org/pdf/1708.00489.pdf

35 Gal Y, Islam R, Ghahramani Z. Deep Bayesian active learning with image data. 2017. https://arxiv.org/pdf/1703.02910.pdf

36 Wang D, Shang Y. A new active labeling method for deep learning. International Joint Conference on Neural Networks. IEEE, 2014.

37 Sharat A, Himanshu A, Saket A, Chetan A. Contextual diversity for active learning. 2020. https://www.ecva.net/papers/eccv_2020/papers_ECCV/papers/123610137.pdf

38 Li Y, Yuan G, Wen Y, *et al*. Efficientformer: Vision transformers at mobilenet speed. 2022. https://arxiv.org/pdf/2206.01191.pdf

39 Tan M, Le QV. EfficientNetV2: smaller models and faster training. 2021. https://arxiv.org/pdf/2104.00298.pdf

40 Dosovitskiy A, Beyer L, Kolesnikov A, *et al*. An image is worth 16x16 words: transformers for image recognition at scale. 2021. https://arxiv.org/pdf/2010.11929.pdf

41 Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. https://arxiv.org/pdf/1608.06993.pdf

42 He KM, Zhang XY, Ren SQ, Sun J. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016:770-778.

43 Liu Z, Lin YT, Cao Y, Hu H, Wei YX, Zhang Z, Lin S, Guo BN. Swin transformer: hierarchical vision transformer using shifted windows. 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 2021:9992-10002.

44 Sogawa T, Tabuchi H, Nagasato D, Masumoto H, Ikuno Y, Ohsugi H, Ishitobi N, Mitamura Y. Accuracy of a deep convolutional neural network in the detection of myopic macular diseases using swept-source optical coherence tomography. *PLoS One* 2020;15(4):e0227240.

45 Li YH, Feng WB, Zhao XJ, *et al*. Development and validation of a deep learning system to screen vision-threatening conditions in high myopia using optical coherence tomography images. *Br J Ophthalmol* 2022;106(5):633-639.

46 Wang RN, He JN, Chen QY, *et al*. Efficacy of a deep learning system for screening myopic maculopathy based on color fundus photographs. *Ophthalmol Ther* 2023;12(1):469-484.