

Supp: MM-Retinal: Knowledge-Enhanced Foundational Pretraining with Fundus Image-Text Expertise

1 Details of MM-Retinal Dataset Construction

To advance multi-modal fundus foundation model research and foster expert knowledge integration in learning fundus image analysis models, we build MM-Retinal, a high-quality image-text paired fundus dataset that comprises CFP, FFA, and OCT modalities. We design a semi-automatic collection procedure to improve construction efficiency, which consists of four steps: 1) image-text pair collection; 2) image-text alignment; 3) modality classification; 4) text cleaning and bilingual translation.

Step 1: Image-Text Pair Collection. Textual reports for fundus diagnosis are typically not accompanied with images in clinical process. Therefore, unlike X-ray and CT images, it is challenging to directly obtain image-text pairs from fundus clinical reports. To address this, we collect image-text pairs from four diagram books illustrating ocular fundus diseases with high-quality expert captions. First, image-text pairs are captured from the books. In cases that one figure corresponds to one caption, we simply capture them in one screenshot. For the cases that multiple sub-figures correspond to a piece of caption, both the text and its corresponding sub-figures are captured together, as shown in Fig. 1(a) in the main body. Note that we keep the resolution of each image in the pair no less than 800×800 . Afterwards, these screenshots are filled into the program we designed to parse images by Adobe and extract texts by OCR technique. Additionally, the color of a certain book suffer a dark tone. We implemented a dehazing operation based on Gamma transformation to correct the color of this book to match the color distribution of other books.

Step 2: Image-Text Alignment. As mentioned earlier, not all the initially extracted images and texts by the program are very well aligned, as there are some sub-figures correspond to one caption within a screenshot. Thus, for those failure cases, we use regular expression matching to split them. Specifically, if the text matches "Figure No.", it indicates the beginning of the text of a new image-text pair. If it matches "Letter.", it indicates the beginning of the text of a new subfigure-text pair. As for the separation of sub-figures, we apply Adobe to automatically implement subgraph segmentation.

Step 3: Modality Classification. Since the fundus images in the books include multiple modalities, we classified them into CFP, FFA, and OCT, separately, which are the mainstream modalities in fundus domain, and exclude others due to their limited samples. Specifically, we first employ K-means to categorize images into three categories based on their color histograms. It effectively

separates the CFP modality which characterized by the distinct color pattern. The other two categories are the mixture of FFA, OCT and images from other modalities. Subsequently, we select a reference image from both OCT and FFA respectively. The classification of remaining images is based on the distance between their color histograms and the references' color histograms. This process results in a precise separation of FFA and OCT modality images.

Step 4: Text Cleaning and Bilingual Translation. The text extraction may have some OCR recognition errors or incomplete issues, so we correct the texts manually. Moreover, to enhance the relevance within an image-text pair, we remove irrelevant information from the text, such as the index of the corresponding image. We also make modifications to sentence inaccuracies. For example, the original text describes multi subgraphs (e.g. 'both eyes show'), but after separating them, the text description should be appropriately adjusted to reflect details about each individual subgraph (e.g. 'left eye shows' and 'right eye shows'). In addition, since the diagram books we selected include both Chinese and English, we provide bilingual reports in English and Chinese version to standardize the language and make MM-Retinal more influential. DeepL Translator is applied to translate Chinese to English, and Tencent Translator is used to translate English to Chinese.

2 Comparison of Transferability on Unseen Downstream Datasets

We present the supplementary results of transferability on unseen downstream datasets. The metric for REFUGE, ODIR200×3 and FIVES is ACA, and for AMD is AUC. From Fig. 1, KeepFit performs better with smaller training splits, like 20%-60%. As the training data increases, KeepFit's performance slightly declines, yet it consistently outperforms most baselines. This demonstrates its superior transferability, generalization capabilities and robustness.

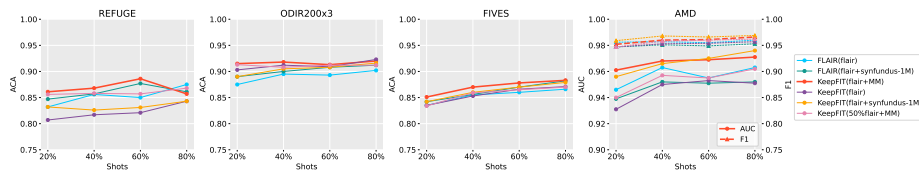


Fig. 1: Comparison of transferability on unseen downstream datasets.

3 Disease and Abnormal Changes Categories in MM-Retinal

This section provides the major categories of fundus diseases and abnormal changes in MM-Retinal with several example images in CFP, FFA and OCT

modalities. These categories are summarized from the contents of four diagram books we used.


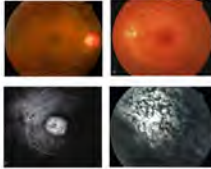

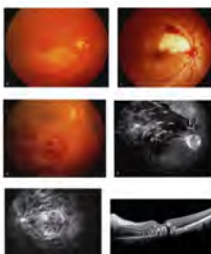
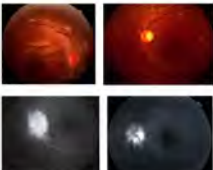
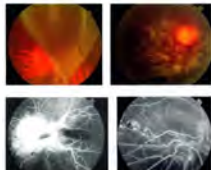
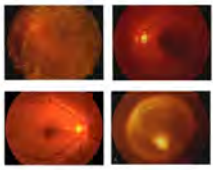
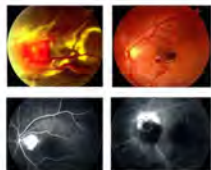
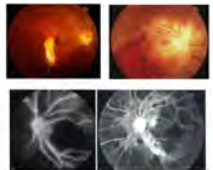
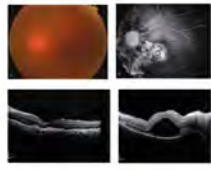
<p>Disease of congenital anomalies</p> <ol style="list-style-type: none"> 1) Congenital anomalies of vitreous 2) Abnormal development of optic nerve 3) Abnormal development of retina and choroid 4) Congenital macular anomalies 5) Congenital retinal vascular anomalies 6) Congenital choroidal defect 		<p>Chorio-retinal degeneration and dystrophy</p> <ol style="list-style-type: none"> 1) Retinal degeneration and dystrophy 2) Cone dystrophy 3) Pattern dystrophy 4) Congenital stationary night blindness 5) Choroidal dystrophies 	
<p>Macular diseases</p> <ol style="list-style-type: none"> 1) Macular degeneration and dystrophy 2) Degenerative myopia 3) Central serous chorioretinopathy 4) Multiple posterior pole retinal pigment epitheliopathies 5) Occult macular dystrophy 6) High myopia macular degeneration 7) Central exudative chorioretinopathy 8) Age-related macular degeneration 9) Idiopathic polypoidal choroidal vasculopathy 10) Vitreomacular traction syndrome 11) Choroidal Folds 		<p>Retinal vascular diseases</p> <ol style="list-style-type: none"> 1) Retinal vascular occlusive diseases 2) Hypertensive retinopathy 3) Diabetic retinopathy 4) Retinal vasculitis 5) Complications of vascular diseases 6) Cilioretinal artery obstruction 7) Sickle cell retinopathy 8) Lipemia Retinalis 9) Radiation Retinopathy 10) Ocular ischemic syndrome 11) Other retinal vascular diseases 	
<p>Optic nerve diseases</p> <ol style="list-style-type: none"> 1) Optic neuritis 2) Optic disc vasculitis 3) Ischemic optic neuropathy 4) Multiple sclerosis 5) Neuromyelitis optica 6) Leber's hereditary optic neuropathy 7) Optic atrophy 8) Optic disk edema 9) Optic disk vitreous warts 10) Pseudophakia 		<p>Retinal and choroidal detachment</p> <ol style="list-style-type: none"> 1) Rhegmatogenous retinal detachment 2) Special type retinal detachment 3) Traction retinal detachment 4) Exudative retinal detachment 5) Traumatic retinal detachment 6) Bullous retinal detachment 7) Proliferative vitreoretinopathy 8) Rhegmatogenous retinal detachment 	
<p>Disease of vitreous</p> <ol style="list-style-type: none"> 1) Asteroid hyalosis 2) Synchysis scintillans 3) Persistent hyperplasia of primary vitreous 4) Congenital cyst of vitreous 5) Vitreous parasite 		<p>Intraocular tumors</p> <ol style="list-style-type: none"> 1) Tumors of Optic disc 2) Tumors of retina 3) Tumors of choroid 4) Ciliary body melanoma 5) Ocular Lymphoma 	
<p>Traumatic and toxic retinopathies</p> <ol style="list-style-type: none"> 1) Solar maculopathy 2) Traumatic macular hole 3) Chorioretinitis sclopetaria 4) Tale retinopathy 5) Chloroquine or hydroxychloroquine retinopathy 6) Thioridazine retinopathy 		<p>Retinal and choroidal inflammatory diseases</p> <ol style="list-style-type: none"> 1) Infectious chorioretinitis 2) Noninfectious chorioretinitis 3) Other choroidal diseases 	

Fig. 2: Major fundus diseases in MM-Retinal

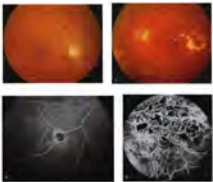
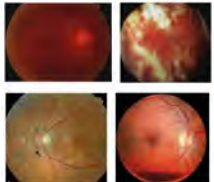
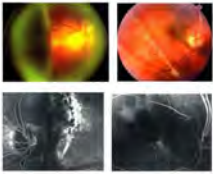
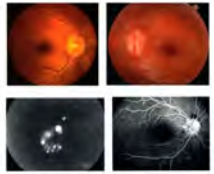
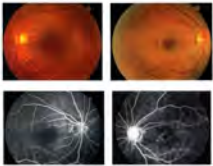
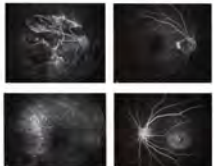

<p>Retinal vascular</p> <ol style="list-style-type: none"> 1) Commotio Retinae 2) Retinal vascular changes 3) Retinal vascular hemorrhages 4) Retinal exudates and hemorrhages 5) Retinal drusens 6) Retinal breaks 7) Retinal detachment 8) Gyrate atrophy 		<p>Vitreous</p> <ol style="list-style-type: none"> 1) Vitreous amyloidosis 2) Vitreous haemorrhages 3) Posterior vitreous detachment 4) Avulsed Vitreous Base 	
<p>Choroidal</p> <ol style="list-style-type: none"> 1) Choroidal detachment 2) Choroidal rupture 		<p>Optic disk</p> <ol style="list-style-type: none"> 1) Papilledema 2) Optic disc drusen 3) Fundus changes in hypermyopia 4) Fundus changes in high blood pressure and atherosclerosis 	
<p>Macular</p> <ol style="list-style-type: none"> 1) Subretinal neovascularization 2) Macular hole 3) Macular cystoid edema 		<p>Abnormal fluorescence in angioaram</p> <ol style="list-style-type: none"> 1) Hyperfluorescence 	
<p>Crystalline lens</p> <ol style="list-style-type: none"> 1) Dislocated Lens 			

Fig. 3: Major abnormal fundus changes in MM-Retinal