Data

Before training we need data, download the data from path `/data1/khadga/data/ODIR_Aug_Resample` from `Anusandhan` server. Move the `ODIR_Aug_Resample` folder to the `data/` folder to start training.

Training Teacher Model

To train teacher model run `train_teacher.py`, The default parameters are set to finetune `Resnet50` teacher model for 50 epochs.

To train other teacher models change `--model_t` param to one of `['resnet50','resnet50_mod','resnet18','wrn_50_2','vgg16','densenet121','densenet201','shuv2_x1_0','shuv2_x2_0']`.

The trained models are automatically saved in folder `save/teacher_models/`.

To change the model checkpoint path change the path in `models_new/__init__.py` file.

Training the MViT Encoder and Decoder

To train MViT MAE, first navigate to `mae_imagenet/` folder and run `download_mae_pretrained.sh` file to first download the pretrained checkpoint file, then run `main_finetune_mae.py` file to finetune the mae. The model checkpoints are save in `~/output_dir_mae_finetune/` folder.How to train the Model

Training the Student Model

The main code for training the model is given in `train_main.py` file. The default parameters are set to train `Resnet18` student model with `Resent50` as teacher model

with `MAE` co-distillation and `SWD` loss.

The parameters are:

- `--batch_size` controls the batch size
- `--epochs` number of epochs to train, default 240
- `--model_t` controls which teacher model to use, default is Resnet50
- `--model_s` controls which student model to train, default is Resnet18
- `--use_gen` controls whether to use MViT co-distillation or not (0: do not use, 1 use, default : 1)
- `--div` controls what distillation loss function to use, default SWD loss, `choices=['kl','dkd', 'swd']`
- `--use_hard_target` controls whether to use hard target loss (cross-entropy with target labels). Default 0. since, the SWD loss already calculates the hard target loss, for other `div` loss functions like `kl` use 1.

Other Params: `lr_decay_epochs`, `scheduler`, `lr_decay_rate`, `cudaid`, etc.,

Example:

- Using only SWD loss `python3 train_main.py --use_gen 0 --div swd`
- Using mvit + KD loss `python3 train_main.py --use_gen 1 --div kl --use_hard_target 1`

The results are automatically saved in `save/student_model` folder.