

Methodology

Yashraj V. Molawade

Training:

1. Train data consists of 17034 images and has an almost equal distribution of approx 17% between the different classes. Dataset was balanced.
2. Image read using computer vision, and normalized and standardized before passing to model.
3. Used transfer learning through pretrained models like VGG-19, VGG-16, Resnet, Xception and InceptionResnetV3.
4. Best model accuracy was obtained from the VGG19 model by freezing the first 5 layers.
5. Image augmentation and Learning Rate Scheduler decreased the accuracy, hence were not used in the final model.
6. Accuracy from models with large depth was less compared to VGG19. The models with heavy depth were overfitting on the training data.
7. Models with around 30 million trainable parameters performed best. Hence, heavy models were used by freezing layers which would ultimately give 30M parameters.

Testing:

1. Test data consists of 7301 images.
2. Predictions from 6 different models which gave high accuracy when run individually.
3. To find the best predictions, mode was taken on all predictions from different models.
4. Accuracy increased to 0.952 with mode predictions.
5. The predictions were incorrect mostly between mountains and glaciers.

Attached: csv files with predictions from 6 best models and a csv file named 'common' with predictions taken as mode of these 6 files.