kaggle Intel Image Classification

Image Scene Classification of Multiclass

Question 1. To what extend the data transformation (augmentation) can improve the performance of the model?

Based on the kaggle image classification dataset information, the data contains images of size 150x150 distributed under 6 categories. After examining a small set, it is noticed that some images have sides smaller than 150. The images are resized. The train set is split to 80% train and 20% validation set.

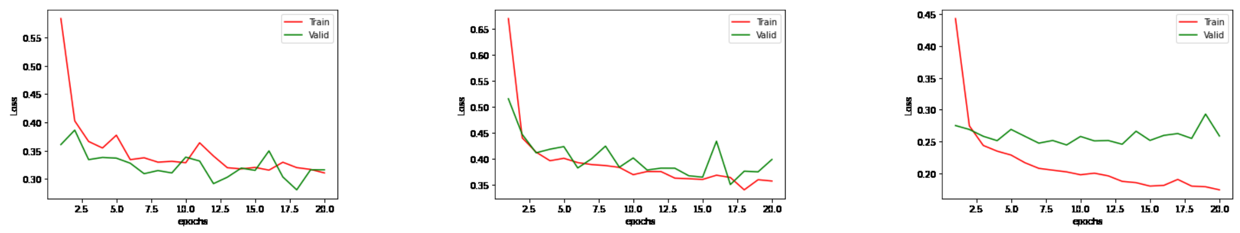
To study the effect of augmentation on the accuracy, the following three different augmentation pipelines are used:

1. Resize, RandomAffine, RandomHorizontalFlip,CenterCrop, Pad, ToTensor, Normalize
2. Resize, RandomAffine, RandomHorizontalFlip,CenterCrop, Pad, ColorJitter, ToTensor, Normalize
3. Resize, RandomHorizontalFlip, CenterCrop, ToTensor, Normalize

For two different input sizes of 150 and 224, the ResNet50 model's results are shown in Table 1. The optimizer is SGD with a learning rate of 0.01, and the batch size is 128. To investigate the effect of overfitting on the losses, the number of epochs is selected as 20, Figure 1. It can be concluded that for this dataset, applying a more comprehensive data augmentation has not improved the accuracy of the model.

Table 1

Input size	pipeline	Accuracy (%)
150	1	89
	2	89
	3	90
224	1	90
	2	89
	3	91



(a) (b) (c)

Figure1. (a) pipeline 1 (b) pipeline 2 (c) pipeline 3

Question 2. Investigate the effect of input size

From Table 1, it can be concluded that in terms of accuracy the input size of 224 performs slightly better than the input size of 150. Feeding the model with 224 by 224 images improves the classification accuracy. Hence, for the rest of this study, the input image size of 224 and pipeline 3 have been used.

Question 3. Investigate the effect of learning rate and choice of optimizer

The effect of choosing an optimizer with an adaptive learning rate, Adam, versus an optimizer with fixed learning on the accuracy is examined. As for the SGD optimizer, four rates of 0.005, 0.01, 0.05 and 0.1 with the momentum of 0.9 are used. Table 2 enlists the results for the batch size of 128 and the number of epochs of 20. Only for the high learning rate of 0.1 the accuracy decreases. With other values of learning rates, both Adam and SGD yield to the same accuracy.

Table 2

Optimizer/learning rate	Accuracy (%)
Adam	91
SGD / 0.005	91
SGD / 0.01	91
SGD / 0.05	91
SGD / 0.10	89

Question 4. Investigate the effect of batch size

The batch size significantly affects the trend of validation loss. This study shows that for small batch sizes, 64, the variation of the validation loss versus the number of epochs exhibits significant fluctuation, Figure 2. Table 3 shows that for the selected batch sizes and learning rates the model's accuracy doesn't change. For the rest of this study, to maintain a more stable training, while using a high learning rate, 0.05, the number of batch size is increased from 128 to 256.

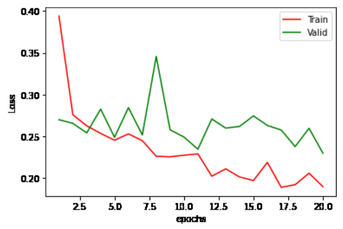


Figure 2. Batch size=64, learning rate = 0.01

Table 3

Batch size	Optimizer/learning_rate	Accuracy (%)
64	SGD / 0.01	91
64	SGD / 0.05	88
128	SGD / 0.05	91
128	Adam	91
256	SGD / 0.05	91
256	Adam	91

Question 5. Investigate the effect of architecture on the accuracy.

The training outcome of four different architecture is presented in Table 4. Input image of 224, batch size of 256 is used in all training. For the present dataset, ResNet50 results in better accuracy with fewer number of iterations.

Table 4

Model	Optimizer / learning rate	Number of epochs	Accuracy (%)
resnext50_32x4d	SGD / 0.05	20	90
resnext50_32x4d	Adam	100	91
wide_resnet_50_2	SGD / 0.05	20	89
wide_resnet_50_2	Adam	20	89
wide_resnet_50_2	Adam	100	90
shufflenet_v2_x1_0	Adam	20	89
shufflenet_v2_x1_0	Adam	200	91
resnet_50	SGD / 0.05	20	91
resnet_50	Adam	20	91

Question 6. For the best overall model and hyperparameters, provide the confusion matrix and its plot.

In this section, ResNet50 with the following parameters is used: Input size = 224, Batch size = 256, optimizer = SGD, learning rate = 0.05, n_epoch = 20. The confusion matrix plot is shown in Figure 3.

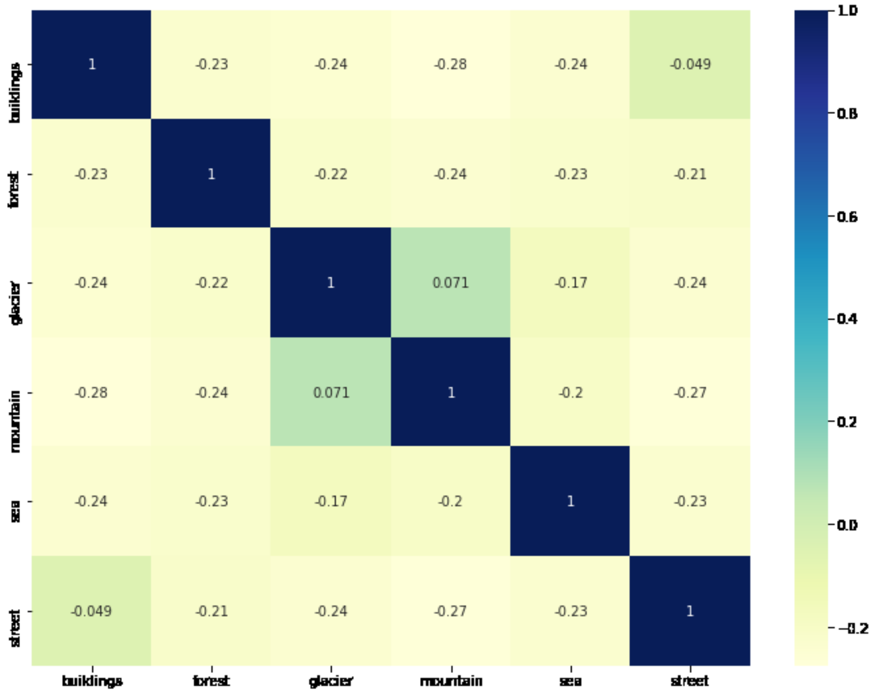


Figure 3. Confusion matrix

Question 7. Obtain the classification report

Table 5

	precision	recall	f1-score	support
buildings	0.90	0.93	0.91	437
forest	0.99	0.99	0.99	474
glacier	0.92	0.77	0.84	553
mountain	0.82	0.89	0.85	525
sea	0.89	0.98	0.93	510
street	0.95	0.90	0.92	501
accuracy			0.91	3000
macro avg	0.91	0.91	0.91	3000
Weighted avg	0.91	0.91	0.91	3000

Mean recall: 91.03%
Mean precision: 91.01%
Class with the lowest precision: mountain - 0.8167539267015707
Class with the lowest recall: glacier - 0.7667269439421338

Question 8. Discuss on the misclassification

The number of instances in each class is: ['buildings:437', 'forest:474', 'glacier:553', 'mountain:525', 'sea:510', 'street:501']. The number of misclassified labels per categories are shown in Table 6.

Table 6

	buildings	forest	glacier	mountain	sea	street
street	24	1	0	0	0	0
sea	4	1	29	21	0	5
glacier	1	0	0	33	4	0
mountain	0	2	97	0	6	0
buildings	0	0	1	1	1	44
forest	0	0	2	2	1	0

Figures 4 and 5 show the predictions for a small set of pred_set images. We can observe that even to human eyes, differentiating some images is not easy; e.g. glaciers. The classifier makes mistakes in distinguishing snowy mountains and seas from glaciers. Another issue is that some pictures have both buildings and streets within, but the label shows only one. A wrong label, outlier, is also another issue. Overall, the model classifies forests well.

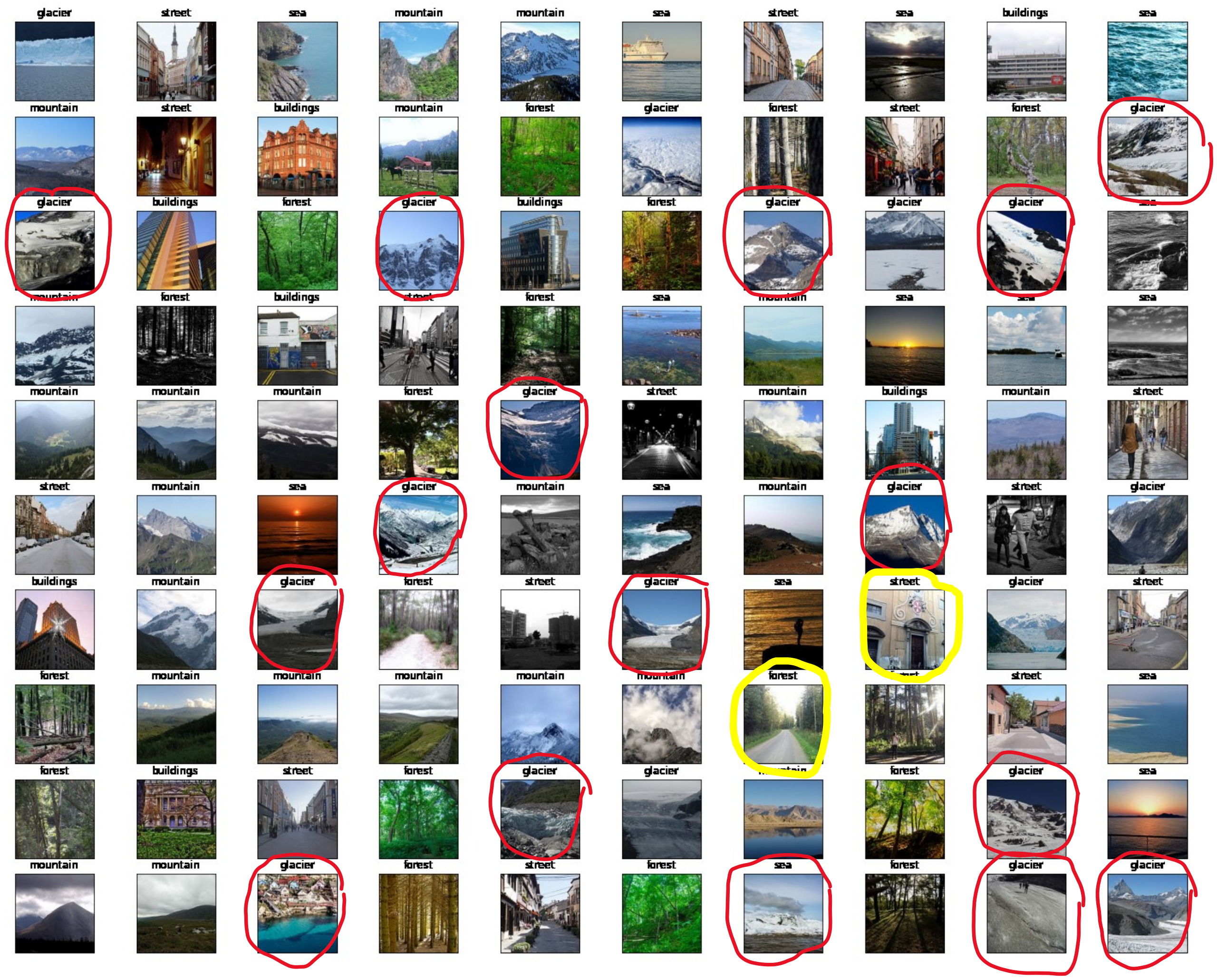

glacier
street
sea
mountain
mountain
sea
street
sea
buildings
sea
mountain
street
buildings
mountain
forest
glacier
forest
street
forest
glacier
glacier
buildings
forest
glacier
buildings
forest
glacier
glacier
glacier
sea
mountain
forest
buildings
street
forest
sea
mountain
sea
sea
sea
mountain
mountain
mountain
forest
glacier
street
mountain
buildings
mountain
street
street
mountain
sea
glacier
mountain
sea
mountain
glacier
street
glacier
buildings
mountain
glacier
forest
street
glacier
sea
street
glacier
street
forest
mountain
mountain
mountain
mountain
mountain
forest
street
street
sea
forest
buildings
street
forest
glacier
glacier
mountain
forest
glacier
sea
mountain
mountain
glacier
forest
street
forest
sea
forest
glacier
glacier

Figure 4

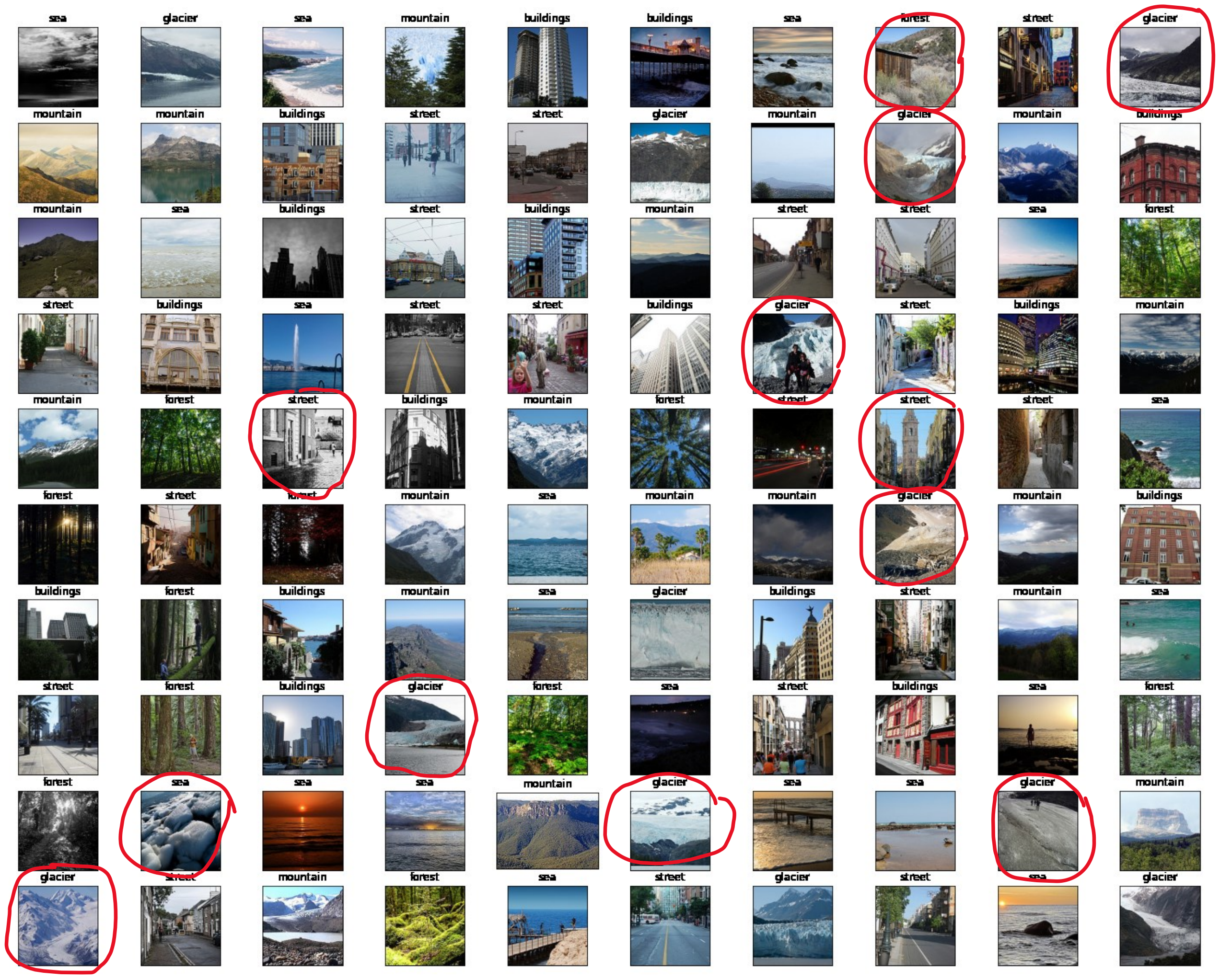

sea
glacier
sea
mountain
buildings
buildings
sea
forest
street
glacier
mountain
mountain
buildings
street
street
glacier
mountain
glacier
mountain
buildings
mountain
sea
buildings
street
buildings
mountain
street
street
sea
forest
street
buildings
sea
street
street
buildings
glacier
street
buildings
mountain
mountain
forest
street
buildings
mountain
forest
street
street
street
sea
forest
street
forest
mountain
sea
mountain
mountain
glacier
mountain
buildings
buildings
forest
buildings
mountain
sea
glacier
buildings
street
mountain
sea
street
forest
buildings
glacier
forest
sea
street
buildings
sea
forest
forest
sea
sea
sea
mountain
glacier
sea
sea
glacier
mountain
glacier
street
mountain
forest
sea
street
glacier
street
sea
glacier

Figure 5