

Introduction

Traditional neural network suffer from catastrophic forgetting(CF), In recent years, many approaches aim to solve this problem, e.g., EWC, GEM, GR, etc. Because Lifelong Object Recognition task need to consider model size, inference time, and other metrics, these previous methods cannot be simply applied to this task. We proposed a suitable framework which uses saliency detection with knowledge distillation, the result on the validation and test datasets show our framework's effectiveness.

Methodology

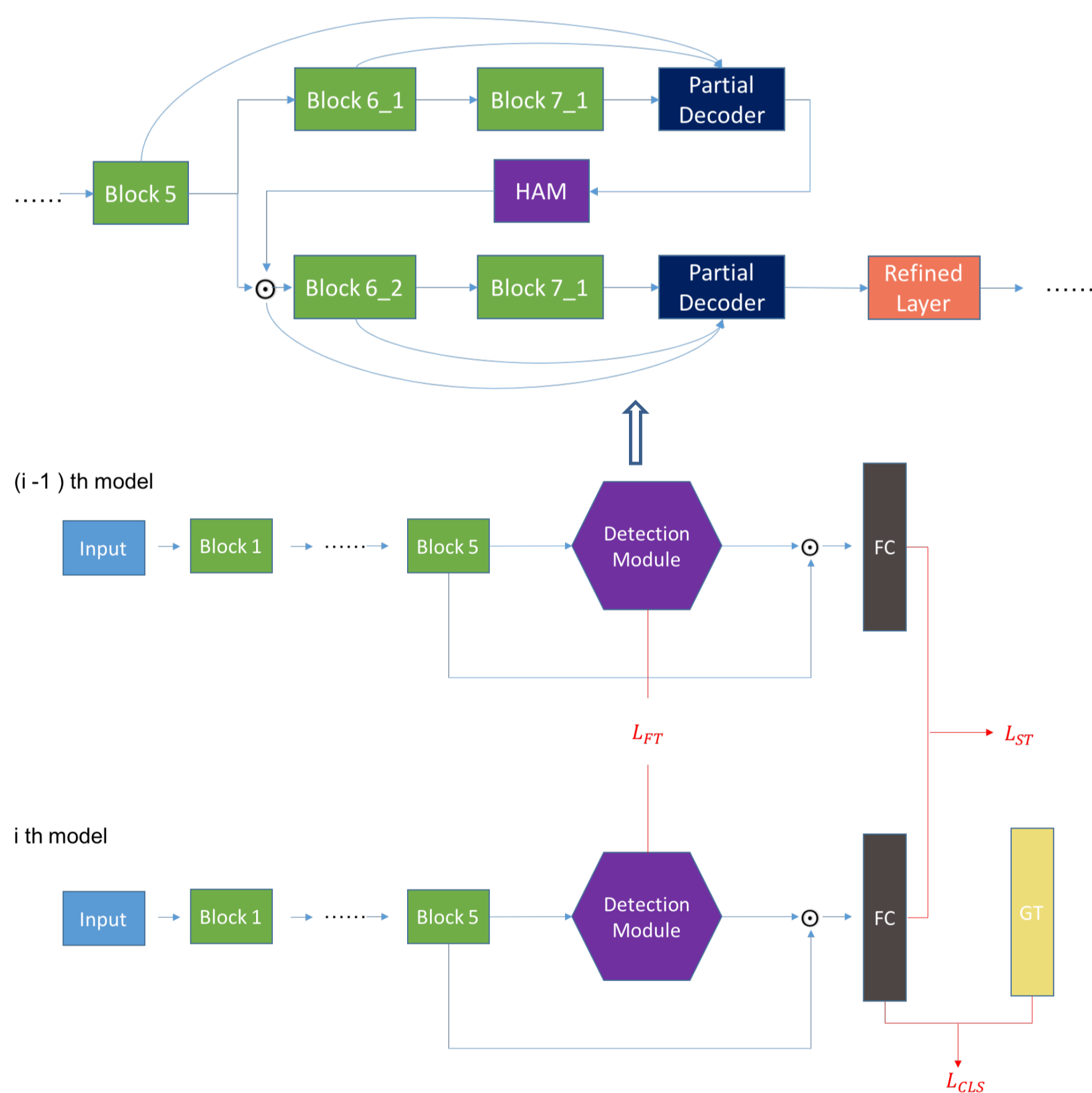


Figure 1: The architecture of the network.

We propose a framework named SDKD. Because this task is difficulty-incremental, which means environmental factors can be variant. Due to this task need to consider model size ,inference time, and the shallower layer contribute less to performance but have large resolution, we use a partial decoder which only integrates features of deeper layers to remove environmental factors called “**Saliency Detection**”. When the model encounter catastrophic forgetting problem, we adapt **knowledge distillation** techniques to alleviate this problem.

Results

We use MobileNet-V2 to be model backbone, and because the poor performance of vanilla MoibileNet-V2 in this task, we choose VGG-16 to be a baseline model, which contains more than 10x parameters than our model. Our results are significantly better than baseline model. And ablation study shows that our knowledge distillation can allievate catastrophic forgetting problem.

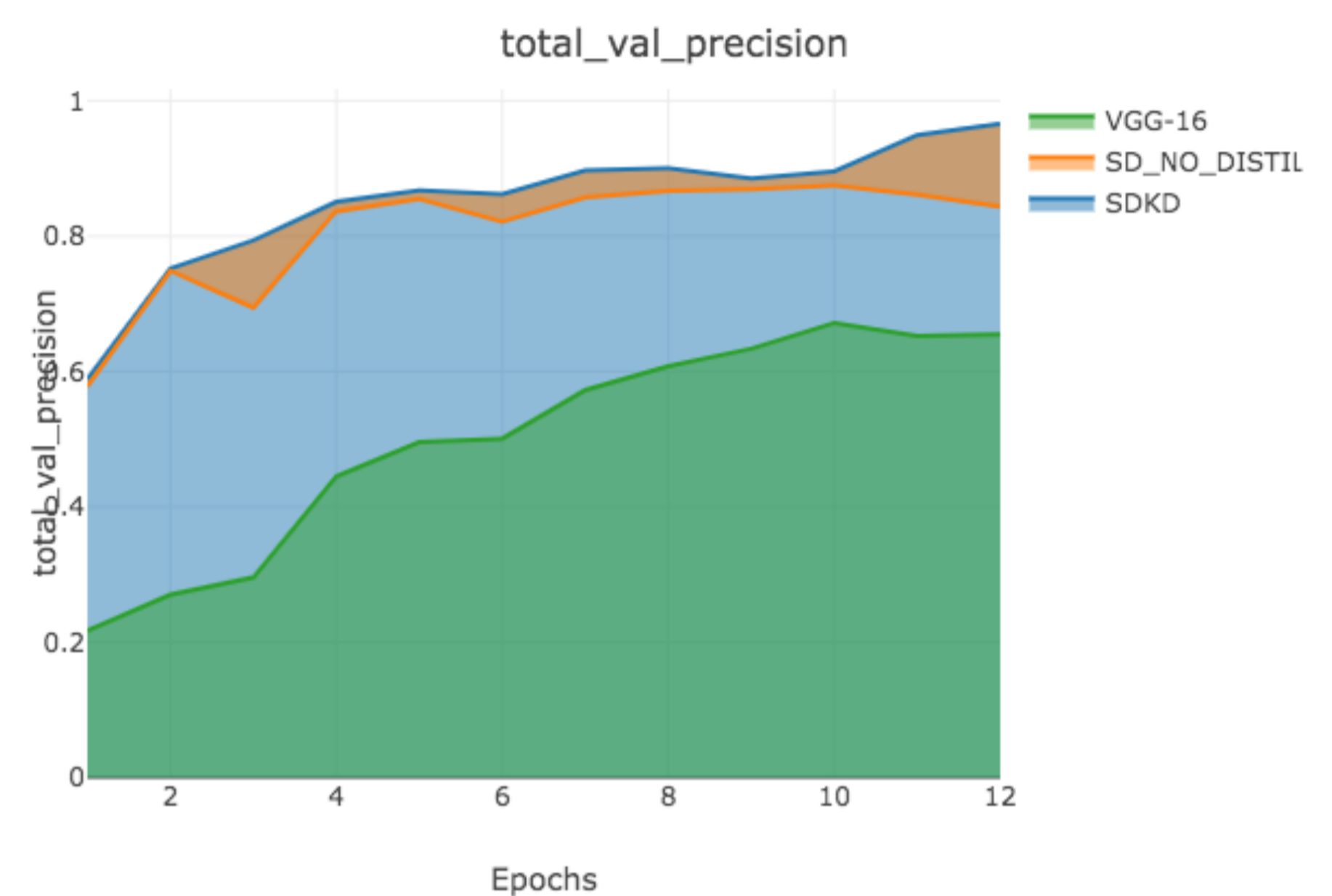


Figure 2: Mean Accuracy on validation dataset.

Conclusion

In this paper, we propose a novel neural network model, named SDKD, for lifelong object recognition. We want to allievate catastrophic forgetting problem in terms of the model. Due to model size, inference time and other constraints, we use a partial decoder to get better results. Because our model is not only learn how to classify images ,but the ability of saliency detection, which is helpful for this task, we can just use knowledge distillation to make sure model cannot appear serious CF problem. Qualitative and quantitaive experimental results in datasets show that the effectiveness of the proposed approach.

Acknowledgment

This work was directed by Professor Baoquan Chen and Professor Hao Dong from Peking University.