

702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755

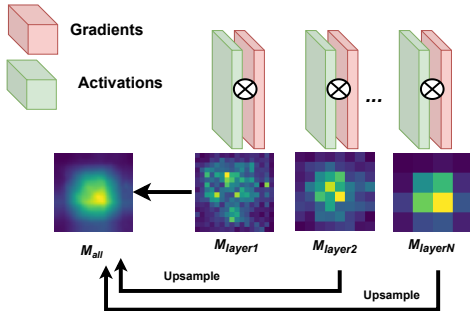


Figure 7: Illustration of layer aggregation in GradMix

A APPENDIX

A.1 SETTINGS FOR OSR EXPERIMENTS

Following section 4.1, the basic descriptions of the six splitting protocols for OSR are in table 6.

| Protocols | Known (# Classes / Source) | Unknown (# Classes / Source) | Openness |
|--------------|-------------------------------|---------------------------------|----------|
| MNIST | 6 / MNIST | 4 / MNIST | 22.54% |
| SVHN | 6 / SVHN | 4 / SVHN | 22.54% |
| CIFAR10 | 6 / CIFAR10 | 4 / CIFAR10 | 22.54% |
| CIFAR+10 | 4 / CIFAR10 | 10 / CIFAR100 | 46.55% |
| CIFAR+50 | 4 / CIFAR10 | 50 / CIFAR100 | 72.784% |
| TinyImagenet | 20 / TinyImagenet | 180 / TinyImagenet | 68.37% |

Table 6: Datasets splitting protocols for known and unknown classes

A.2 ILLUSTRATION OF LAYER AGGREGATION

In 4.1, we tested GradMix for OSR using multiple layers in deep models. The process of layer aggregation is illustrated in figure 7. Assume that the activation maps computed using layer 1, 2, and N according to 3.2 are M_{layer1} , M_{layer2} , and M_{layerN} , the aggregated activation map $M_{layerall}$ is the summation of them, $M_{layerall} = M_{layer1} + M_{layer2} + M_{layerN}$, after the upsampling of lower resolution activation maps.

A.3 MORE VISUALIZATIONS

Following section 5, the activation maps on more samples are demonstrated in figure 8. Same as in 5, GradMix models focus on a broader range of regions in the data.

A.4 CODE FOR GRADMIX

Our code is publicly released on https://anonymous.4open.science/r/comprehensive_osr-8EEF. Part of the frameworks on (self-)supervised contrastive learning and MoCo are from <https://github.com/HobbitLong/SupContrast> and <https://github.com/facebookresearch/moco> respectively.

756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809

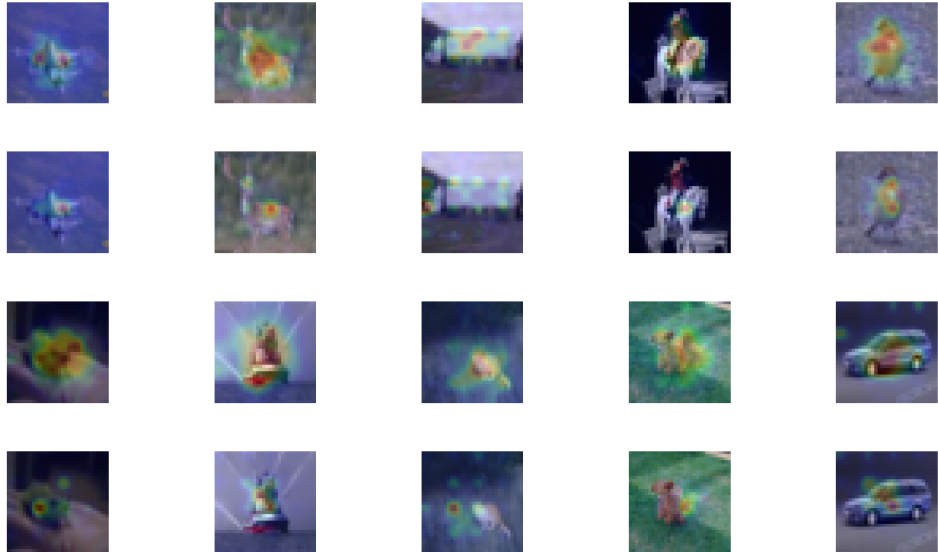


Figure 8: Comparison on the activation visualizations of the models trained on CIFAR10 protocol. trained with and without GradMix. First Row: model trained with GradMix; Second Row: model trained without GradMix