Oriented Response Networks (Supplementary Materials)

Yanzhao Zhou[1], Qixiang Ye[1], Qiang Qiu[2], and Jianbin Jiao[1]

[1]University of Chinese Academy of Sciences
[2]Duke University

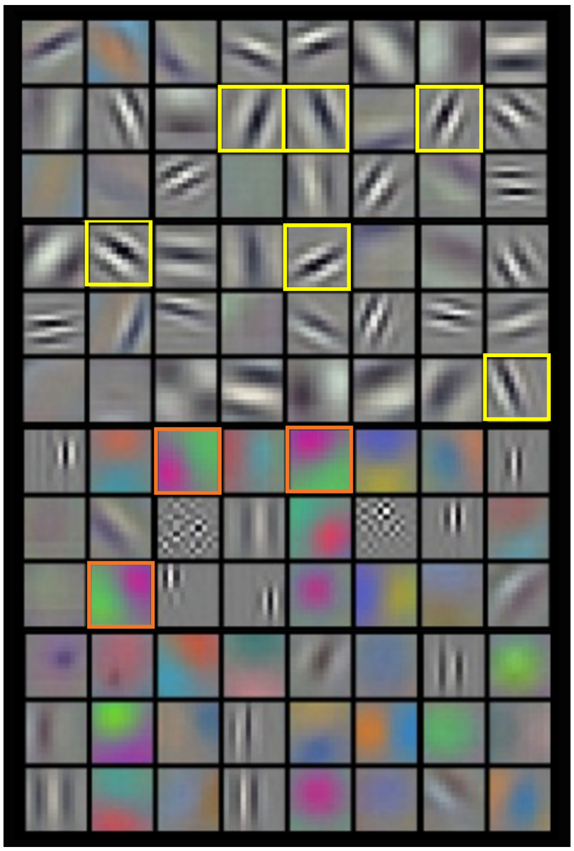


Figure 1. Filters from the first convolutional layer of AlexNet trained on ImageNet. Due to the 'learning by rote' strategy, some structure-alike but orientation-distinct filters are observed.

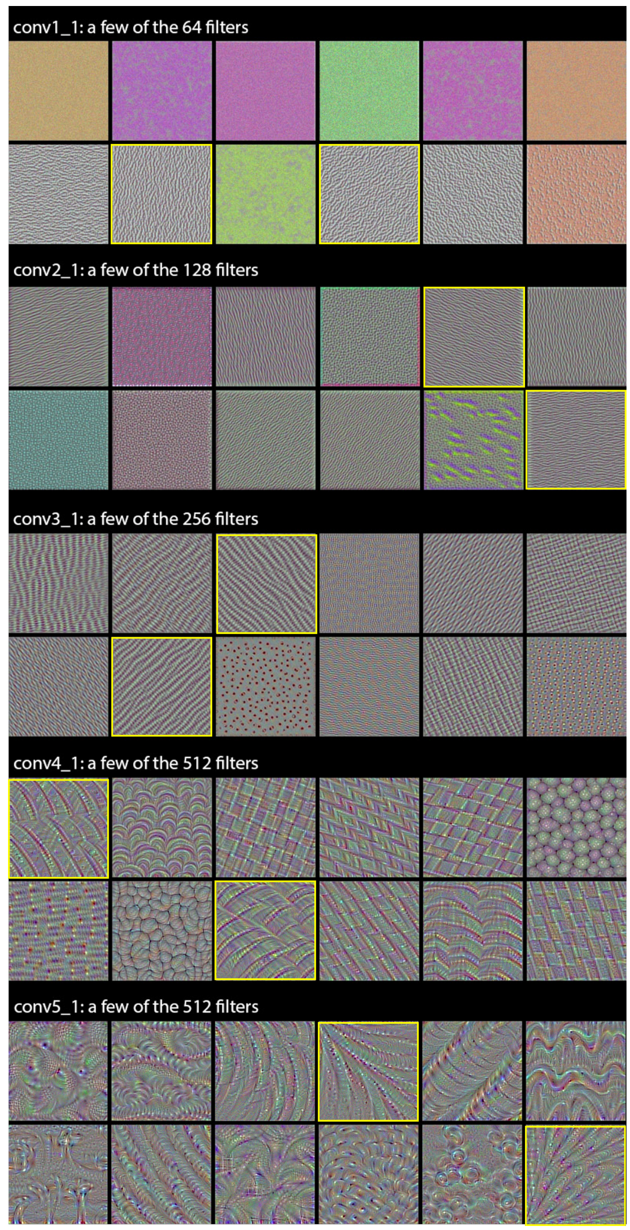


Figure 2. Visualization of filters of a VGG-16 model trained on ImageNet. Structure-alike but orientation-distinct filters are observed in some deep layers.

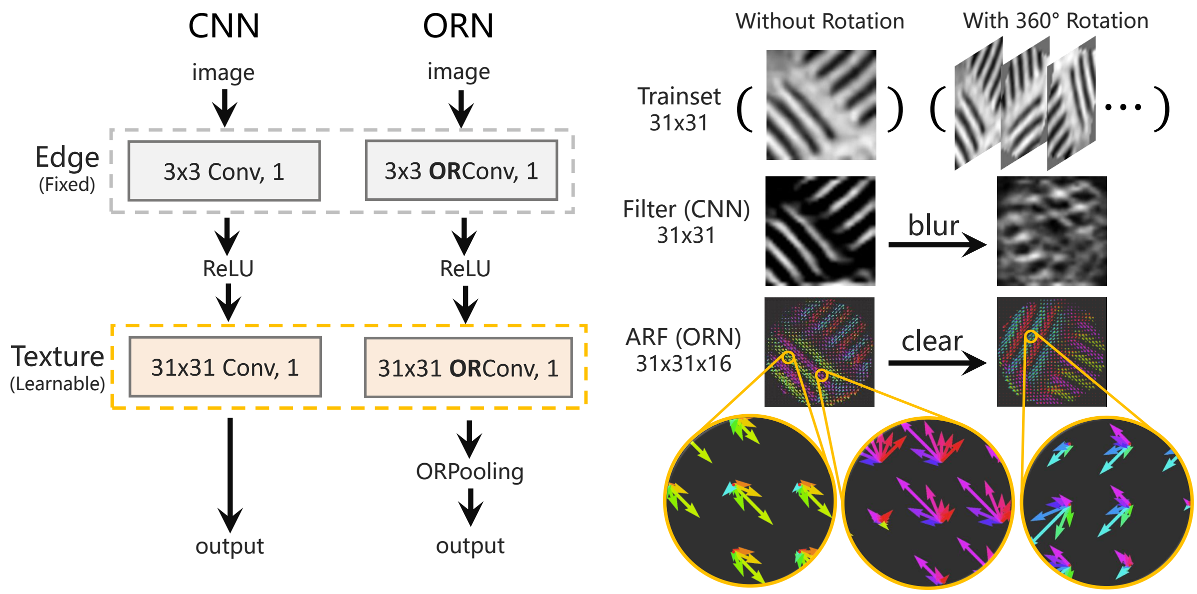


Figure 3. Comparison of CNN filters and Actively Rotating Filters (ARFs) learned on texture data.

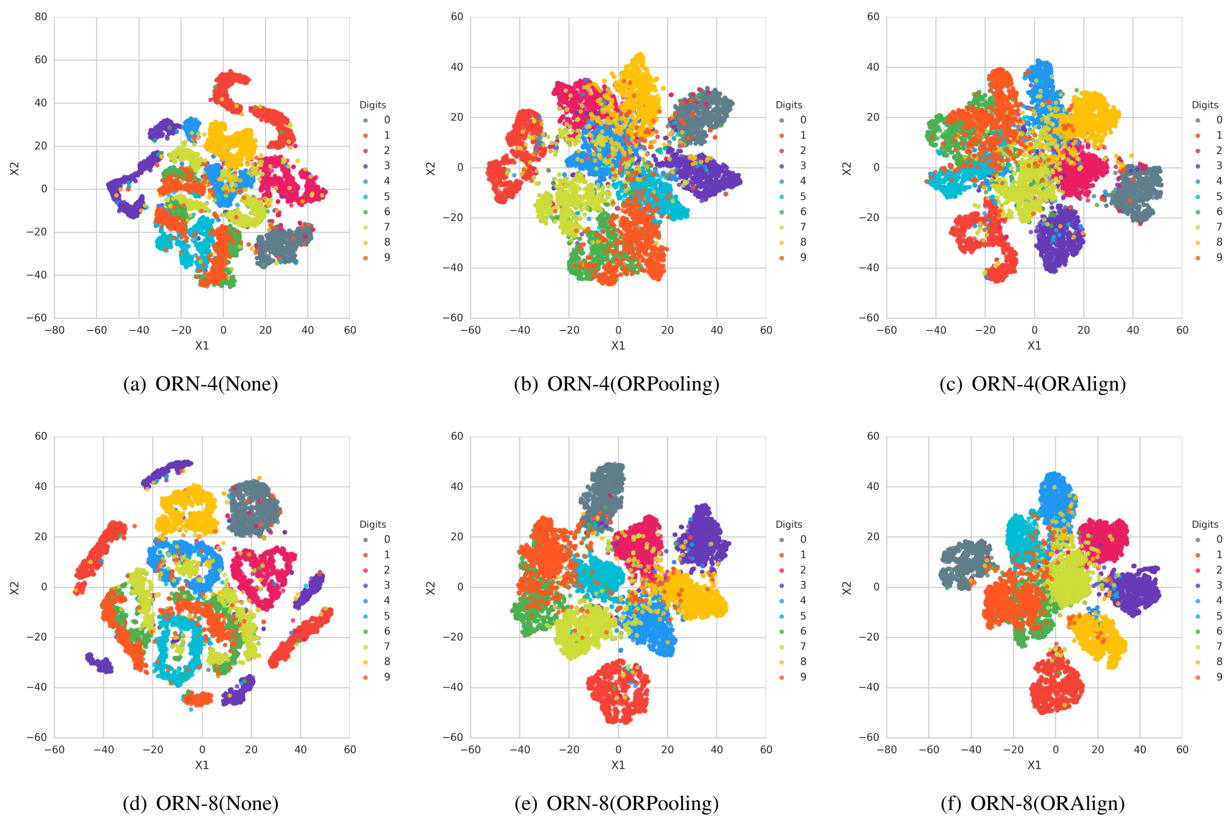
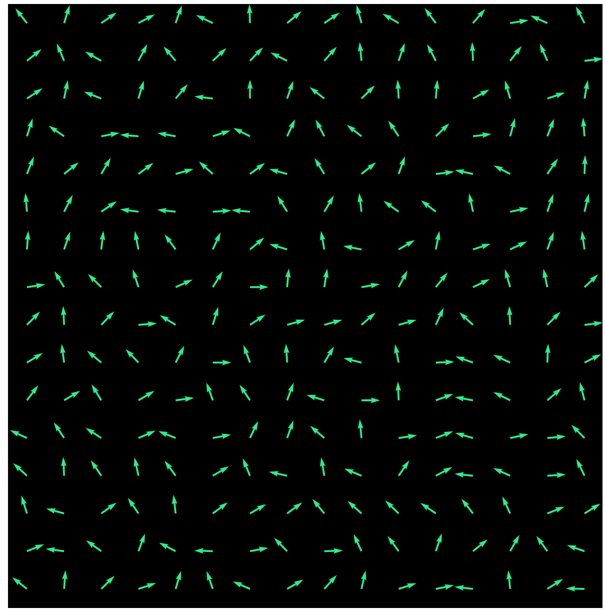
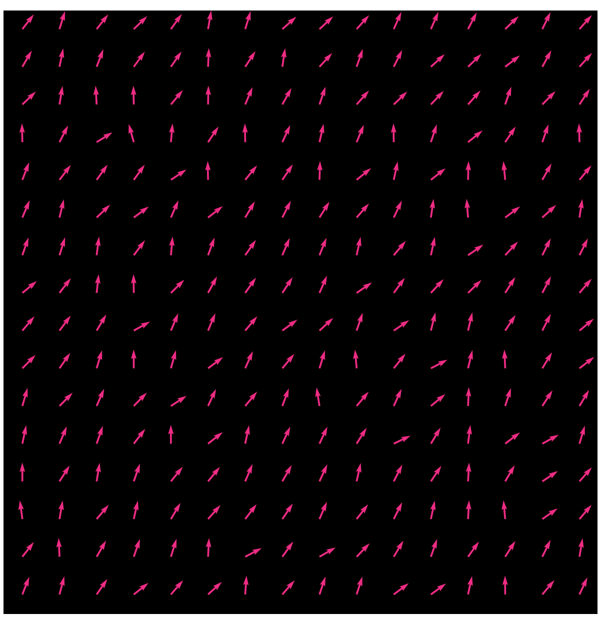


Figure 4. 2D tSNE feature mapping of ORN-4 and ORN-8, which shows that more orientation channels improve the discrimination of ORN.

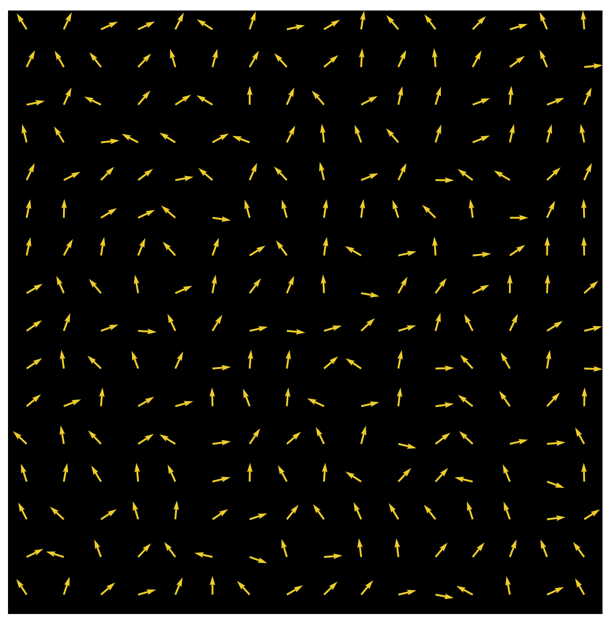
(a) Randomly rotated digits



(b) Ground truth orientation

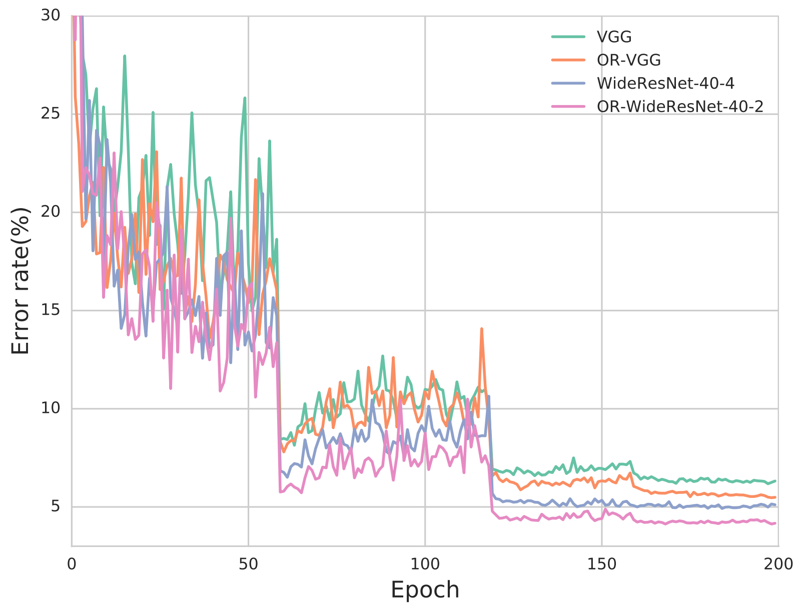


(c) Orientation estimated by STN

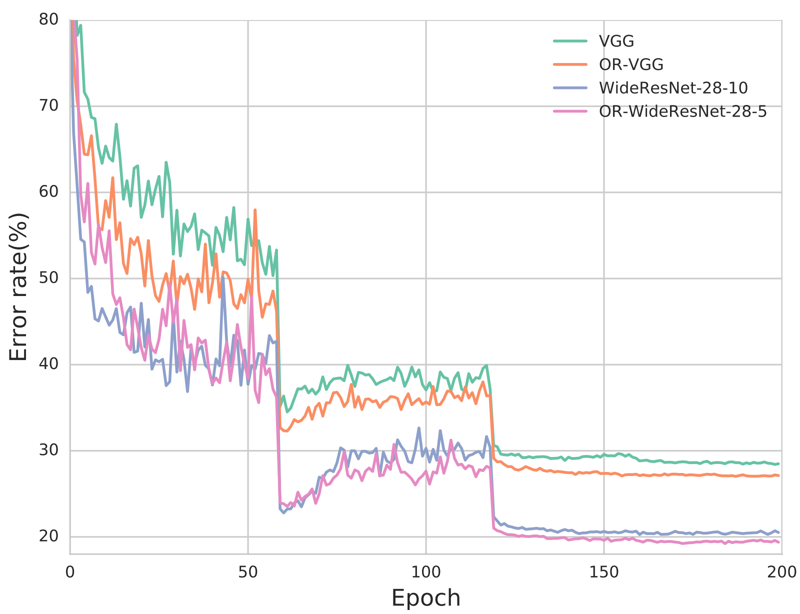


(d) Orientation estimated by OR-STN

Figure 5. Orientation estimation results. As shown in (d), OR-STN significantly outperforms the STN method for orientation estimation.



(a) CIFAR10



(b) CIFAR100

Figure 6. Testing error-rate curves.

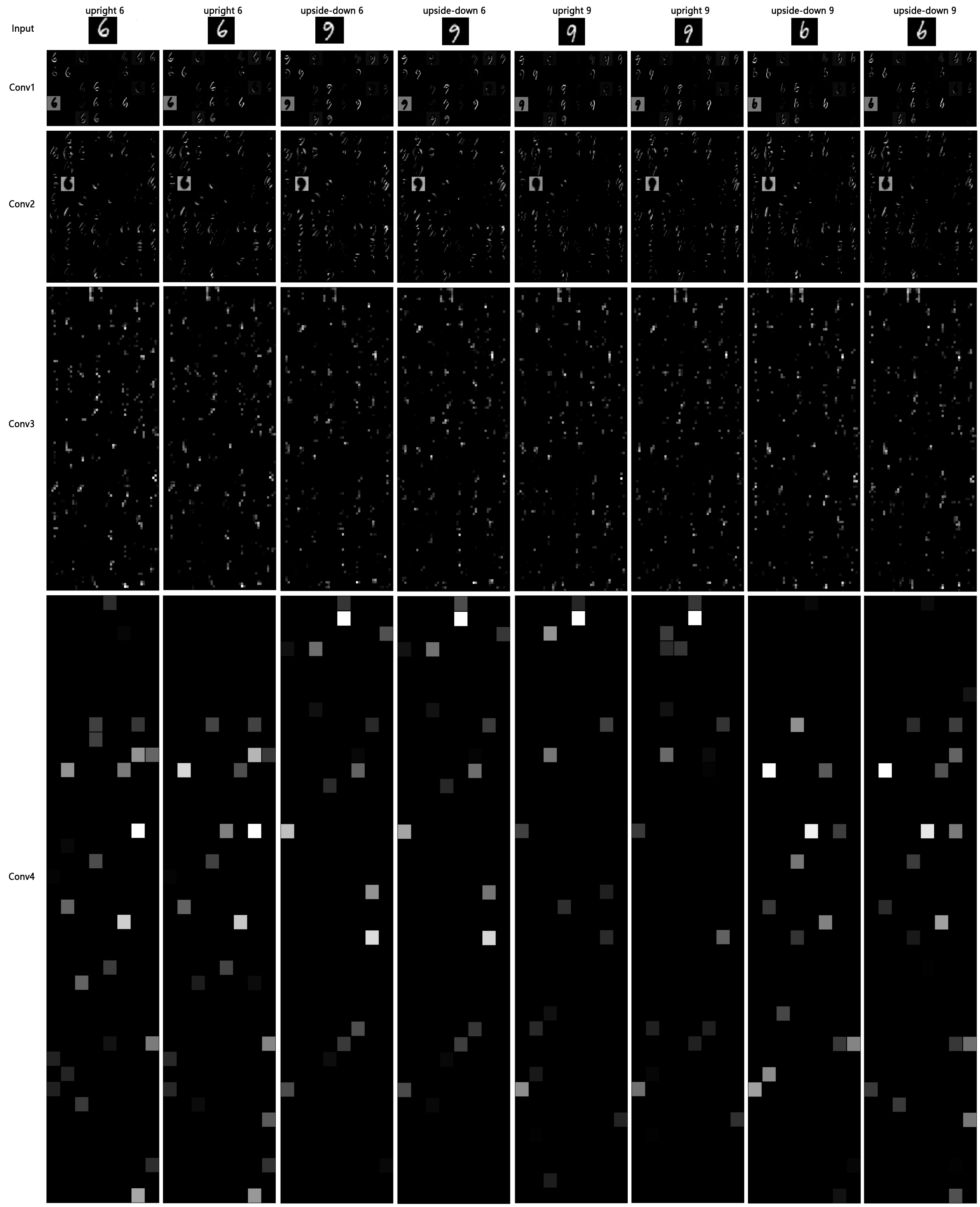


Figure 7. Features produced by a conventional CNN for class ‘6’ and ‘9’ in MNIST. The Conv4 features of upside-down ‘6’ (column 3) mistakenly show difference to those of upside-down ‘6’ (column 1) but similarity to upright ‘9’ (column 5).

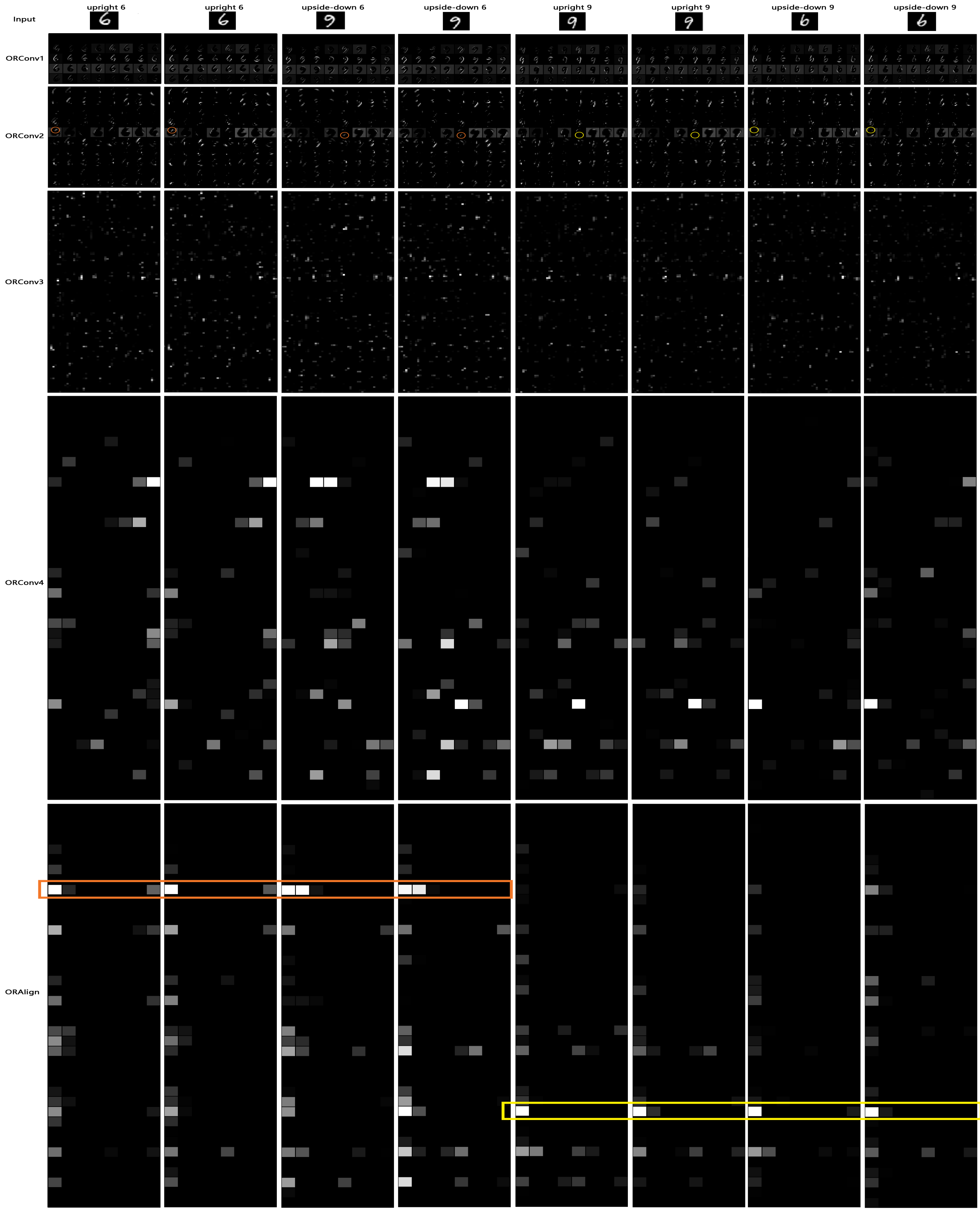


Figure 8. Features produced by the ORN-8 for class '6' and '9' in MNIST. In the ORAlign layer, upright '6' (column 1) correctly shows similar features as upside-down '6' (column 3); the features of upside-down '6' (column 3) are significantly different from the upright '9'(column 5). This shows that ORN produces within-class rotation-invariant features, while maintaining inter-class discrimination.