Fig. 2. (left to right) i) An anomalous image ii) The retrieved top normal neighbor image iii) The mask detected by SPADE iv) The predicted anomalous image pixels. We can see how in this example, SPADE detects the anomalous image region by finding the correspondence with the nearest-neighbor image. The anomalous parts did not have correspondences in the normal image and were therefore detected.

Table 5. Pixel-level anomaly detection accuracy on STC (Average ROCAUC %)

AE_{L2}	AE_{SSIM}	CAVGA-R_u [31]	SPADE
74	76	85	**89.9**

tivities (e.g. fighting) as well as any moving object which is not a pedestrian (e.g. motorbikes).

We began by evaluating our first stage for detecting image-level anomalies against other state-of-the-art methods. We show in Tab. 4 that our first stage has comparable performance to the top performing method [19]. More interestingly, we compare in Tab. 5 the pixel-level ROCAUC performance with the best reported method, CAVGA-R_u [31]. Our method significantly outperforms the best reported method by a significant margin. Note that we compared to the best method that did not use anomaly supervision, as we do not use it and as anomaly supervision is often not available in practice.

4.3 Ablation Study

We conduct an ablation study on our method in order to understand the relative performance of its different parts. In Tab. 6, we compare using different level of the feature pyramid. We experienced that using activations of too high resolution (56×56) significantly hurts performance (due to limited context) while using