

Divide and Conquer: High-Resolution Industrial Anomaly Detection via Memory Efficient Tiled Ensemble

Supplementary Material

A. Anomaly scales per category

Figure 9 shows the average anomalous pixel ratio (what percentage of defective image is covered by the defect) present in the test set of each category from MVTecAD [6] and VisA [46]. The ratio is calculated on all defective images from the category with a resolution of 512×512 . VisA categories contain notably smaller defects, especially in categories such as candle, macaroni 1, and macaroni 2, where the anomalous pixels on average cover less than 1% of the image.

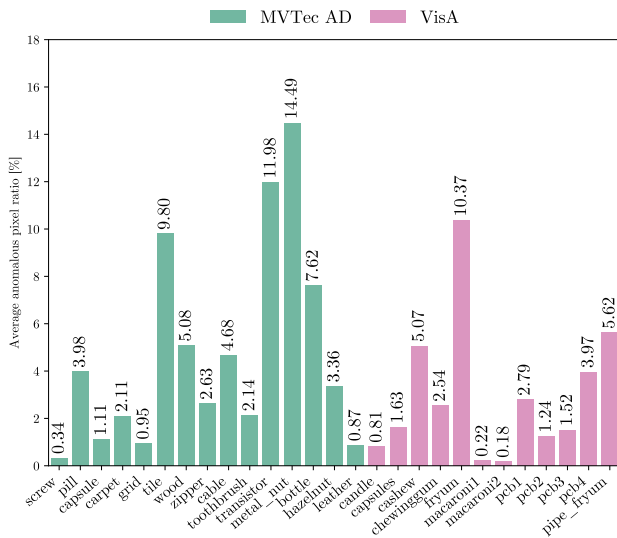


Figure 9. Average anomalous pixel ratio of defective images per category for all categories present in VisA and MVTec AD.

B. Parameter sizes

Table 3 contains the parameter count (in millions) for each setup and architecture. Since the tiled ensemble consists of unchanged underlying architectures, the parameter sizes are increased by the factor equivalent to the number of models in an ensemble.

C. Training time

Table 4 contains the training duration of each architecture with each setup. The results are averaged across all categories from MVTecAD [6] and VisA [46] and 3 runs with different seeds. In cases of FastFlow and Reverse Distillation, the training time is approximately extended by a factor

Setup	PatchCore	Padim	FastFlow	Reverse Distillation
SM256	2.8	2.8	9.7	18.2
SM512	2.8	2.8	12.5	18.2
ST4	2.8	2.8	9.7	18.2
ST9	2.8	2.8	9.7	18.2
ENS4	11.1	11.1	38.9	72.7
ENS9	25.0	25.0	87.6	163.5

Table 3. Parameters (**million**) for each architecture and each setup.

equivalent to the number of models in an ensemble. This is expected due to the cumulative increase in epochs, while the training workload of each model inside the ensemble remains in line with a single model processing a 256×256 resolution.

In the case of PatchCore and Padim, this does not hold since they are not trained using backpropagation. Therefore, the time required for training doesn't scale equivalently when directly increasing resolution (from SM256 to SM512) or when achieving higher processed resolution through multiple smaller models operating within an ensemble (from SM256 to ENS4/ENS9).

Setup	PatchCore	Padim	FastFlow	Reverse Distillation
SM256	213.1	772.0	556.3	721.5
SM512	2273.0	2423.5	799.6	838.8
ST4	2254.6	2400.7	805.1	836.5
ST9	2275.2	2423.0	851.9	867.5
ENS4	879.9	444.9	2561.8	2970.9
ENS9	1810.6	974.6	5857.7	6876.8

Table 4. Training time in **seconds** for each architecture and each setup. Results are averaged over all categories and 3 runs with different seeds.

D. Results of all categories for each setup

This section contains tables with category-specific results for all architectures and their setups.

	Single model 256	Single model 512	Tiled ensemble 4 tiles	Tiled ensemble 9 tiles	Single model 4 tiles	Single model 9 tiles
Carpet	97.4/94.3	96.5/97.0	95.2/96.0	97.6/ 97.0	98.1 /96.2	95.7/96.0
Grid	96.3/88.8	99.9/ 97.1	96.4/96.1	96.9/96.7	100.0 /96.8	100.0 /96.8
Leather	100.0 /97.3	99.4/98.8	98.0/98.4	98.8/ 98.9	99.7/98.4	99.1/98.6
Tile	99.5 /83.9	97.8/88.0	94.9/86.1	96.3/ 88.0	97.1/86.0	96.4/86.5
Wood	99.2/89.1	99.1/94.2	99.5/93.7	99.7 / 94.3	99.1/93.9	99.0/94.0
Bottle	100.0 /93.3	100.0 /96.3	100.0 /96.0	100.0 / 96.4	100.0 /96.0	100.0 /96.1
Cable	98.0 /93.9	95.6/93.4	95.0/93.5	96.9/ 94.6	94.2/92.7	96.3/93.0
Capsule	98.7/93.1	99.1/96.9	99.5/96.2	99.8 /96.9	99.3/96.4	99.7/ 97.0
Hazelnut	100.0/95.0	100.0 / 97.0	99.7/96.3	99.8/97.0	100.0 /96.2	100.0 /96.7
Metal nut	99.8/94.2	100.0 / 96.1	97.4/95.1	98.7/95.7	99.7/95.4	99.9/95.5
Pill	94.0 /93.7	92.3/ 96.5	91.0/95.7	92.8/96.4	90.2/95.8	93.5/96.5
Screw	93.7/96.6	98.3 /97.8	84.6/97.8	92.1/ 98.2	97.0/97.8	98.1/98.1
Toothbrush	94.4/90.9	96.2/ 96.0	98.9/95.8	99.7 /95.9	96.7/95.6	95.1/95.7
Transistor	98.0/ 92.6	97.2/74.6	97.2/77.6	98.4 /86.1	97.4/75.1	96.8/76.6
Zipper	95.9/95.1	98.6/ 97.8	99.5 /96.9	99.0/97.6	98.7/97.1	97.0/97.0
<i>Average</i>	97.7/92.8 (±0.06 / ±0.04)	98.0 /94.5 (±0.08 / ±0.01)	96.5/94.1 (±0.03 / ±0.04)	97.8/ 95.3 (±0.04 / ±0.01)	97.8/94.0 (±0.03 / ±0.01)	97.8/94.3 (±0.37 / ±0.05)

Table 5. MVTec AD results of all 6 setups for Patchcore. The row contains results for a particular category, with columns containing detection and localization results (AUROC/AUPRO) for each setup. The mean of 3 runs is reported with the corresponding standard deviation in parentheses. The best result for each category is in **bold**.

	Single model 256	Single model 512	Tiled ensemble 4 tiles	Tiled ensemble 9 tiles	Single model 4 tiles	Single model 9 tiles
Carpet	96.7/95.3	92.3/94.0	89.2/93.4	93.7/94.2	92.4/93.5	91.6/93.3
Grid	86.4/79.5	93.5/87.2	87.6/86.5	89.3/ 87.5	92.8/86.6	92.7/86.9
Leather	98.0/93.6	96.1/97.9	95.0/98.0	96.6/ 98.2	96.4/98.0	96.0/97.9
Tile	94.8/82.0	84.4/73.2	89.8/73.2	90.2/73.8	84.5/73.2	84.4/72.5
Wood	98.1/92.6	94.8/94.1	96.4/94.0	97.1/ 94.3	94.3/94.0	94.8/94.1
Bottle	99.5/95.2	98.8/95.4	99.1/95.9	99.9/95.7	98.7/ 95.9	98.8/95.5
Cable	82.3/89.2	74.1/81.3	77.0/82.4	81.7/82.8	76.6/82.4	74.8/82.2
Capsule	84.3/92.9	81.5/93.9	88.2/93.8	88.8/93.9	81.2/93.9	82.0/93.8
Hazelnut	80.0/94.3	69.4/95.8	87.6/95.8	94.9/95.9	71.9/95.8	70.1/ 95.9
Metal nut	96.4/92.2	93.4/89.9	90.6/90.0	93.8/90.5	93.8/90.0	93.2/90.0
Pill	86.8/94.4	72.9/93.4	81.6/93.3	81.4/93.6	75.1/93.3	73.7/94.4
Screw	74.2/91.7	61.3/ 92.3	66.2/92.0	69.4/92.2	58.7/92.0	61.1/92.2
Toothbrush	86.9/93.3	88.1/95.7	94.7/95.8	99.4/95.8	89.8/ 95.8	88.1/95.8
Transistor	89.8/89.3	78.2/87.5	83.8/81.5	87.6/83.2	83.5/81.5	81.0/82.4
Zipper	83.1/93.0	66.1/92.7	82.9/92.7	81.2/93.0	67.3/92.7	66.9/92.5
<i>Average</i>	89.2/ 91.2 ($\pm 1.29 / \pm 0.62$)	83.0/91.0 ($\pm 1.28 / \pm 0.36$)	87.3/90.5 ($\pm 0.76 / \pm 0.41$)	89.7/91.0 ($\pm 0.89 / \pm 0.42$)	83.8/90.6 ($\pm 1.07 / \pm 0.41$)	83.3/90.6 ($\pm 1.12 / \pm 0.43$)

Table 6. MVTEC AD results of all 6 setups for Padim. The row contains results for a particular category, with columns containing detection and localization results (AUROC/AUPRO) for each setup. The mean of 3 runs is reported with the corresponding standard deviation in parentheses. The best result for each category is in **bold**.

	Single model 256	Single model 512	Tiled ensemble 4 tiles	Tiled ensemble 9 tiles	Single model 4 tiles	Single model 9 tiles
Carpet	99.0/93.8	94.8/87.7	70.9/90.2	82.0/91.9	95.5/87.3	96.3/88.3
Grid	98.8/93.2	99.5/96.9	96.8/95.5	97.1/ 96.9	99.9/95.8	99.2/96.2
Leather	100.0/108.3	99.6/99.1	94.5/98.8	98.2/99.2	99.5/98.8	99.6/98.4
Tile	96.1/81.3	90.9/77.1	93.6/76.8	92.5/80.8	90.5/76.8	90.8/73.8
Wood	98.5/92.5	97.8/94.0	96.9/94.4	98.4/ 95.3	97.9/93.8	98.0/93.4
Bottle	100.0/91.1	99.8/92.1	98.5/92.0	99.9/ 92.6	99.9/88.5	99.9/90.4
Cable	93.9/86.5	78.3/74.7	90.6/84.2	93.2/ 86.6	81.2/67.0	78.0/57.7
Capsule	92.6/89.4	93.5/95.2	93.4/95.3	97.6/95.8	89.6/94.4	87.7/92.1
Hazelnut	78.2/92.8	80.1/93.9	92.0/94.1	97.7/94.2	88.5/92.9	83.9/91.2
Metal nut	96.4/84.7	91.2/81.6	91.6/87.9	95.7/ 89.2	94.6/83.8	89.6/77.6
Pill	93.1/90.1	93.1/ 91.8	95.1/91.5	97.5/90.7	91.9/87.4	90.6/83.2
Screw	74.7/69.0	65.7/84.8	75.6/82.8	81.2/88.6	71.4/70.2	74.0/74.1
Toothbrush	88.1/83.5	88.1/90.8	97.8/90.2	99.7/92.1	87.5/86.1	84.2/83.7
Transistor	92.1/ 87.9	88.5/73.2	92.3/76.5	96.7/83.1	88.0/62.9	85.3/51.0
Zipper	94.9/92.4	96.3/94.0	97.3/91.9	97.5/94.5	94.5/90.0	94.4/90.5
<i>Average</i>	93.1/89.1 ($\pm 0.29 / \pm 1.08$)	90.5/88.5 ($\pm 0.13 / \pm 0.35$)	91.8/89.5 ($\pm 0.31 / \pm 0.35$)	95.0/91.4 ($\pm 0.38 / \pm 0.27$)	91.4/85.0 ($\pm 0.35 / \pm 0.27$)	90.1/82.8 ($\pm 0.36 / \pm 0.46$)

Table 7. MVTEC AD results of all 6 setups for FastFlow. The row contains results for a particular category, with columns containing detection and localization results (AUROC/AUPRO) for each setup. The mean of 3 runs is reported with the corresponding standard deviation in parentheses. The best result for each category is in **bold**.

	Single model 256	Single model 512	Tiled ensemble 4 tiles	Tiled ensemble 9 tiles	Single model 4 tiles	Single model 9 tiles
Carpet	96.2/96.1	97.2/ 96.6	84.1/85.6	82.5/68.6	97.7 /96.4	86.0/95.1
Grid	84.6/64.4	96.1/96.0	82.8/90.5	90.9/93.0	99.4 / 98.1	99.0/98.1
Leather	83.6 /94.7	76.9/92.3	49.1/73.5	60.3/75.4	57.7/67.6	44.3/ 95.7
Tile	87.1/ 78.7	75.0/64.0	89.4 /45.9	86.8/52.5	88.1/67.6	47.8/75.5
Wood	98.6 /91.0	87.6/90.9	97.3/ 92.9	98.1/88.1	89.4/88.5	53.8/89.9
Bottle	99.8 / 95.1	68.8/90.9	99.2/93.2	98.6/92.0	88.4/91.4	89.5/91.2
Cable	97.8 / 91.6	59.4/69.8	71.1/69.2	82.6/77.6	67.0/64.3	74.1/78.2
Capsule	81.5/90.6	71.3/93.7	87.2/93.3	89.7 / 93.8	61.0/93.2	74.7/92.8
Hazelnut	86.1/95.2	81.6/80.7	94.7/73.7	98.4 /72.6	94.3/ 96.7	69.6/96.2
Metal nut	100.0 / 94.3	90.2/88.2	92.2/91.3	96.8/92.8	89.3/87.9	82.4/87.3
Pill	90.1/94.9	59.5/95.2	87.7/96.9	91.6 / 97.4	66.5/95.9	89.6/96.8
Screw	75.4/92.0	64.6/91.6	58.8/74.9	69.6/76.0	72.9/93.1	84.3 / 94.2
Toothbrush	97.0/92.4	97.8/ 96.6	97.0/95.9	99.9 /96.6	89.0/95.5	98.1/96.1
Transistor	95.3 / 77.1	73.1/62.0	89.7/66.1	91.2/69.0	69.3/61.6	80.5/60.9
Zipper	88.6 /95.0	78.6/94.2	88.0/ 95.7	79.5/93.9	79.9/95.2	85.4/95.5
<i>Average</i>	90.8 / 89.5 ($\pm 3.01 / \pm 3.38$)	78.5/87.2 ($\pm 3.25 / \pm 1.85$)	84.5/82.6 ($\pm 2.51 / \pm 4.99$)	87.8/82.6 ($\pm 1.54 / \pm 3.03$)	80.7/86.2 ($\pm 1.76 / \pm 0.66$)	77.3/89.5 ($\pm 0.92 / \pm 1.04$)

Table 8. MV Tec AD results of all 6 setups for Reverse Distillation. The row contains results for a particular category, with columns containing detection and localization results (AUROC/AUPRO) for each setup. The mean of 3 runs is reported with the corresponding standard deviation in parentheses. The best result for each category is in **bold**.

	Single model 256	Single model 512	Tiled ensemble 4 tiles	Tiled ensemble 9 tiles	Single model 4 tiles	Single model 9 tiles
Candle	97.0/95.1	98.2/96.2	98.7/97.4	99.3 /97.3	98.3/97.5	98.5/ 97.6
Capsules	71.5/70.0	91.9 / 97.2	70.7/94.2	78.1/96.0	80.8/95.5	90.0/96.4
Cashew	93.7/91.5	96.5/91.5	97.0/89.6	97.4 / 91.6	94.9/90.2	96.6/89.7
Chewing gum	98.9/84.3	98.6/85.9	99.1/83.8	99.8 /83.8	99.2/ 88.5	98.8/84.9
Fryum	92.7/83.9	98.4 /91.7	93.0/90.3	96.2/ 91.9	94.5/90.5	97.0/91.1
Macaroni 1	92.3/93.5	98.7 /96.7	95.6/98.2	96.0/ 98.3	97.0/97.1	98.7/98.1
Macaroni 2	72.8/85.9	91.4 /94.8	76.4/94.0	85.2/95.3	87.0/94.1	89.6/ 96.2
PCB1	94.9/92.8	97.9/96.5	98.2/95.7	98.8 / 96.5	98.0/95.2	97.9/96.0
PCB2	92.4/88.4	97.4/93.4	96.2/93.6	97.5/ 93.8	97.7 /93.6	97.7/93.0
PCB3	99.0 /86.6	98.4/94.4	94.7/93.7	97.5/ 94.6	97.5/93.9	97.9/94.2
PCB4	99.0/86.6	99.4/ 92.6	99.5/89.1	99.8 /89.1	98.5/90.8	98.5/90.3
Pipe fryum	99.3/94.1	99.6 / 96.9	98.5/96.1	99.1/96.8	99.6/96.1	99.4/96.5
<i>Average</i>	92.0/87.7 ($\pm 0.29 / \pm 0.39$)	97.2 / 94.0 ($\pm 0.47 / \pm 0.34$)	93.1/93.0 ($\pm 0.07 / \pm 0.05$)	95.4/93.7 ($\pm 0.81 / \pm 0.07$)	95.2/93.6 ($\pm 0.91 / \pm 0.55$)	96.7/93.7 ($\pm 0.09 / \pm 0.31$)

Table 9. VisA results of all 6 setups for Patchcore. The row contains results for a particular category, with columns containing detection and localization results (AUROC/AUPRO) for each setup. The mean of 3 runs is reported with the corresponding standard deviation in parentheses. The best result for each category is in **bold**.

	Single model 256	Single model 512	Tiled ensemble 4 tiles	Tiled ensemble 9 tiles	Single model 4 tiles	Single model 9 tiles
Candle	87.4/92.3	86.0/ 96.1	90.1/96.1	94.0 /92.3	85.8/96.1	86.5/96.1
Capsules	60.5/58.7	62.3/ 76.4	67.9/74.2	67.9 /75.5	61.9/74.2	62.3/75.4
Cashew	88.8 / 84.7	87.3/83.1	88.6/83.6	88.4/83.6	87.6/83.7	87.6/83.2
Chewing gum	98.4 /83.7	97.5/84.2	96.6/83.8	97.2/84.3	89.2/ 87.5	97.4/83.8
Fryum	86.4/77.4	88.8/86.6	87.9/86.9	89.6 /86.7	88.2/ 87.0	88.5/86.6
Macaroni 1	80.3 /89.2	76.7/ 91.4	75.6/91.3	79.7/91.3	76.6/91.3	77.2/91.3
Macaroni 2	71.8/76.4	66.3/72.3	70.3/77.3	72.2 / 78.3	66.2/77.3	66.2/77.8
PCB1	88.6/88.5	84.7/91.9	87.5/91.9	89.9 /92.1	84.9/91.9	80.3/ 93.2
PCB2	81.7/83.9	79.3/90.7	79.5/90.6	84.4 / 90.9	80.4/90.7	80.0/90.8
PCB3	72.5/80.5	73.8/89.6	75.5/89.5	80.7 /89.7	74.3/89.4	74.1/ 89.8
PCB4	96.4/81.7	87.8/ 88.5	96.5/85.3	97.1 /85.0	95.8/85.3	95.3/84.7
Pipe fryum	92.3/88.3	92.1/92.4	89.8/92.2	94.6 / 92.5	91.2/92.2	91.7/92.2
<i>Average</i>	83.7/82.1 ($\pm 0.90 / \pm 0.87$)	81.9/86.9 ($\pm 1.76 / \pm 0.81$)	83.8/86.9 ($\pm 0.96 / \pm 0.84$)	86.3 /86.9 ($\pm 0.96 / \pm 1.29$)	81.8/ 87.2 ($\pm 1.96 / \pm 1.12$)	82.3/87.1 ($\pm 0.58 / \pm 0.89$)

Table 10. VisA results of all 6 setups for Padim. The row contains results for a particular category, with columns containing detection and localization results (AUROC/AUPRO) for each setup. The mean of 3 runs is reported with the corresponding standard deviation in parentheses. The best result for each category is in **bold**.

	Single model 256	Single model 512	Tiled ensemble 4 tiles	Tiled ensemble 9 tiles	Single model 4 tiles	Single model 9 tiles
Candle	90.7/92.4	92.8/96.5	92.4/96.3	93.8 / 96.6	92.5/96.1	90.9/94.8
Capsules	77.5/81.2	84.1/ 92.9	80.4/87.3	85.3 /92.4	78.7/86.0	80.8/83.7
Cashew	88.9/82.6	88.5/ 90.2	91.3/88.0	92.8 /90.0	87.3/83.5	87.6/79.6
Chewing gum	98.7/85.6	99.6/ 90.4	99.8 /88.7	99.7/89.2	99.0/88.2	98.4/85.5
Fryum	93.8/74.7	94.2/ 81.2	93.7/76.6	97.2 /79.6	91.6/75.2	89.1/65.0
Macaroni 1	87.7/88.1	90.6/91.8	85.3/91.5	91.0 / 94.2	88.6/91.4	89.0/87.6
Macaroni 2	73.2 /82.5	72.5/ 88.0	68.3/83.3	70.4/83.7	69.1/82.2	67.3/80.2
PCB1	85.5/86.4	88.5/90.0	91.1/88.5	95.0 / 93.5	83.7/83.8	81.9/74.3
PCB2	83.3/75.4	88.5/83.6	87.9/84.2	94.2 / 88.3	81.9/75.7	82.5/70.6
PCB3	79.1/65.2	86.8/85.2	89.2/86.0	93.8 / 88.1	83.8/78.2	80.0/72.2
PCB4	95.3/80.0	97.5/82.4	99.0/83.9	99.4 / 83.9	94.4/78.0	86.3/69.7
Pipe fryum	95.3/83.0	94.3/88.0	94.5/87.6	97.9 / 90.2	94.9/82.3	93.3/80.2
<i>Average</i>	87.4/81.4 ($\pm 0.42 / \pm 1.10$)	89.8/88.4 ($\pm 0.27 / \pm 0.22$)	89.4/86.8 ($\pm 0.31 / \pm 0.34$)	92.5 / 89.2 ($\pm 0.12 / \pm 0.13$)	87.1/83.4 ($\pm 0.32 / \pm 0.70$)	85.6/78.6 ($\pm 0.18 / \pm 2.62$)

Table 11. VisA results of all 6 setups for FastFlow. The row contains results for a particular category, with columns containing detection and localization results (AUROC/AUPRO) for each setup. The mean of 3 runs is reported with the corresponding standard deviation in parentheses. The best result for each category is in **bold**.

	Single model 256	Single model 512	Tiled ensemble 4 tiles	Tiled ensemble 9 tiles	Single model 4 tiles	Single model 9 tiles
Candle	88.2/93.7	60.4/96.7	89.7/ 97.2	93.0 /96.9	78.4/97.0	84.9/96.5
Capsules	76.9/88.3	74.2/87.4	80.2/94.8	84.0/ 96.0	75.2/93.3	87.6 /95.5
Cashew	92.4/80.9	88.7/ 83.6	93.7/75.8	94.4 /75.2	72.2/76.8	89.0/77.8
Chewing gum	98.4/ 85.0	69.5/68.1	98.9 /69.9	96.3/64.3	78.2/71.2	95.8/70.8
Fryum	81.0/82.7	53.7/89.4	82.8/88.1	92.4 / 90.0	52.1/88.9	82.3/86.8
Macaroni 1	78.4/87.4	88.0/ 96.3	85.8/96.2	88.0/96.1	88.6 /96.3	81.5/95.2
Macaroni 2	59.5/82.1	72.9 / 93.7	72.0/92.9	72.7/93.3	72.3/92.2	69.5/92.1
PCB1	64.7/94.0	59.3/93.5	91.1/94.1	93.3 / 95.4	62.8/93.1	62.2/92.4
PCB2	89.0/87.1	45.3/89.3	86.6/91.0	92.0 / 92.7	60.8/88.2	69.4/89.1
PCB3	81.3/85.9	85.2/90.9	85.8/91.2	94.6 / 93.3	75.0/89.2	82.6/90.4
PCB4	95.9/83.7	84.6/87.6	97.7/88.5	99.3 / 91.7	64.6/86.7	73.9/85.9
Pipe fryum	93.1/92.2	76.3/ 95.6	95.7/95.2	96.7 /95.0	90.4/94.7	88.3/94.7
<i>Average</i>	83.2/86.9 (±1.93 / ±0.66)	71.5/89.3 (±5.96 / ±1.70)	88.3/89.6 (±1.54 / ±0.12)	91.4 / 90.0 (±1.32 / ±0.47)	72.5/89.3 (±4.29 / ±0.48)	80.6/88.9 (±2.59 / ±1.03)

Table 12. VisA results of all 6 setups for Reverse Distillation. The row contains results for a particular category, with columns containing detection and localization results (AUROC/AUPRO) for each setup. The mean of 3 runs is reported with the corresponding standard deviation in parentheses. The best result for each category is in **bold**.

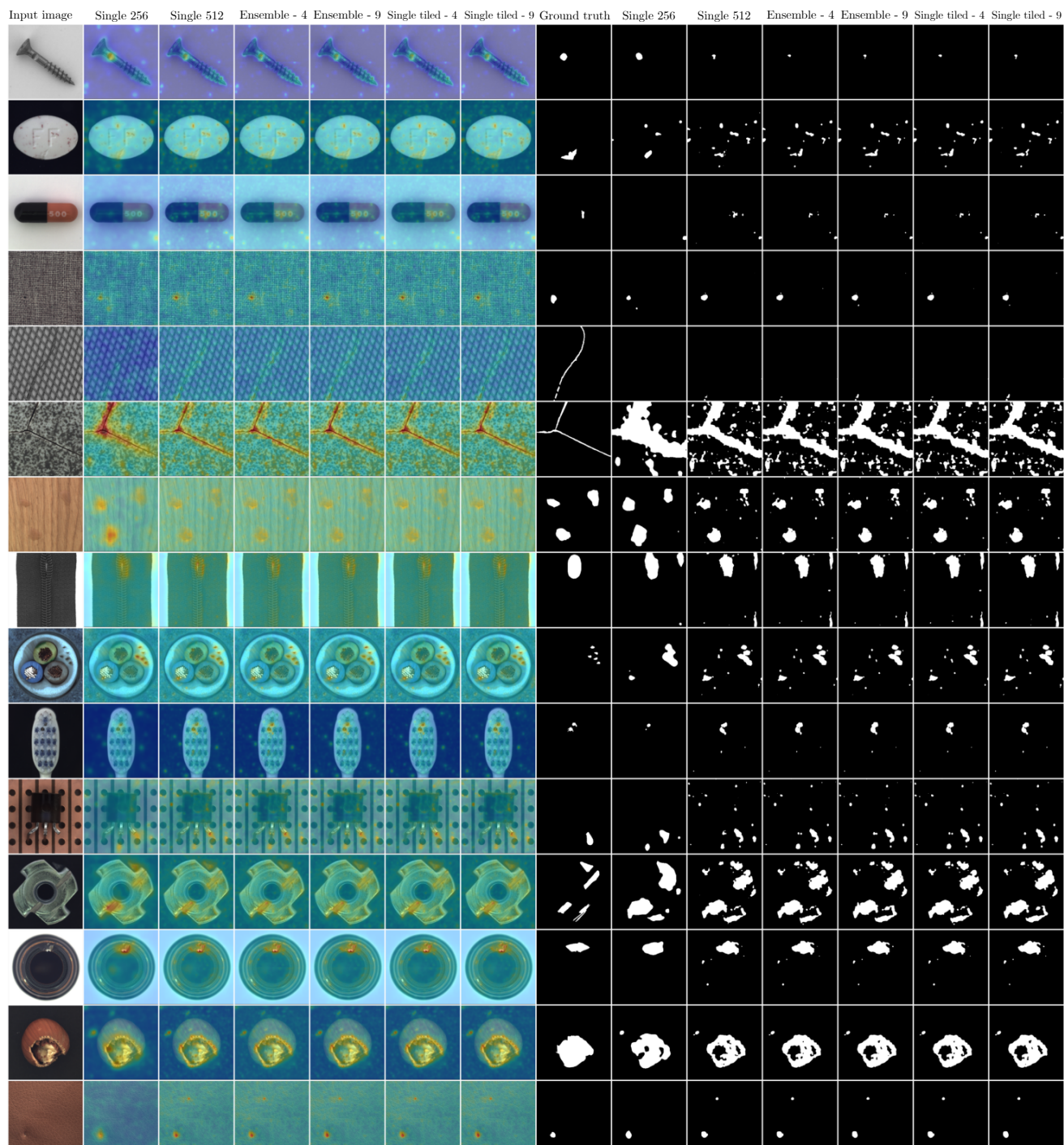


Figure 11. Anomaly maps and segmentation masks for each setup using Padim on randomly picked defective image from every category in MVTEC AD.

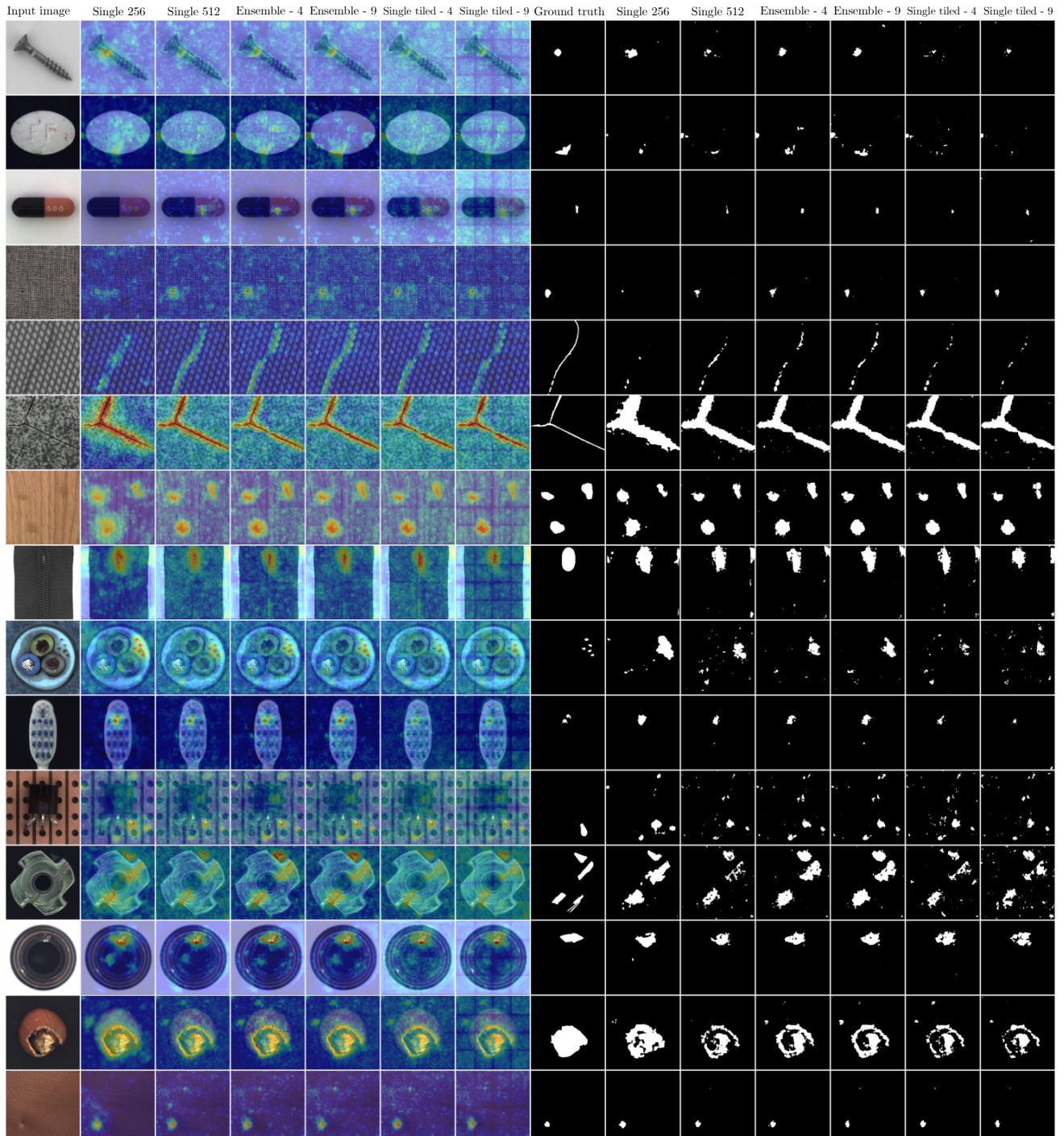
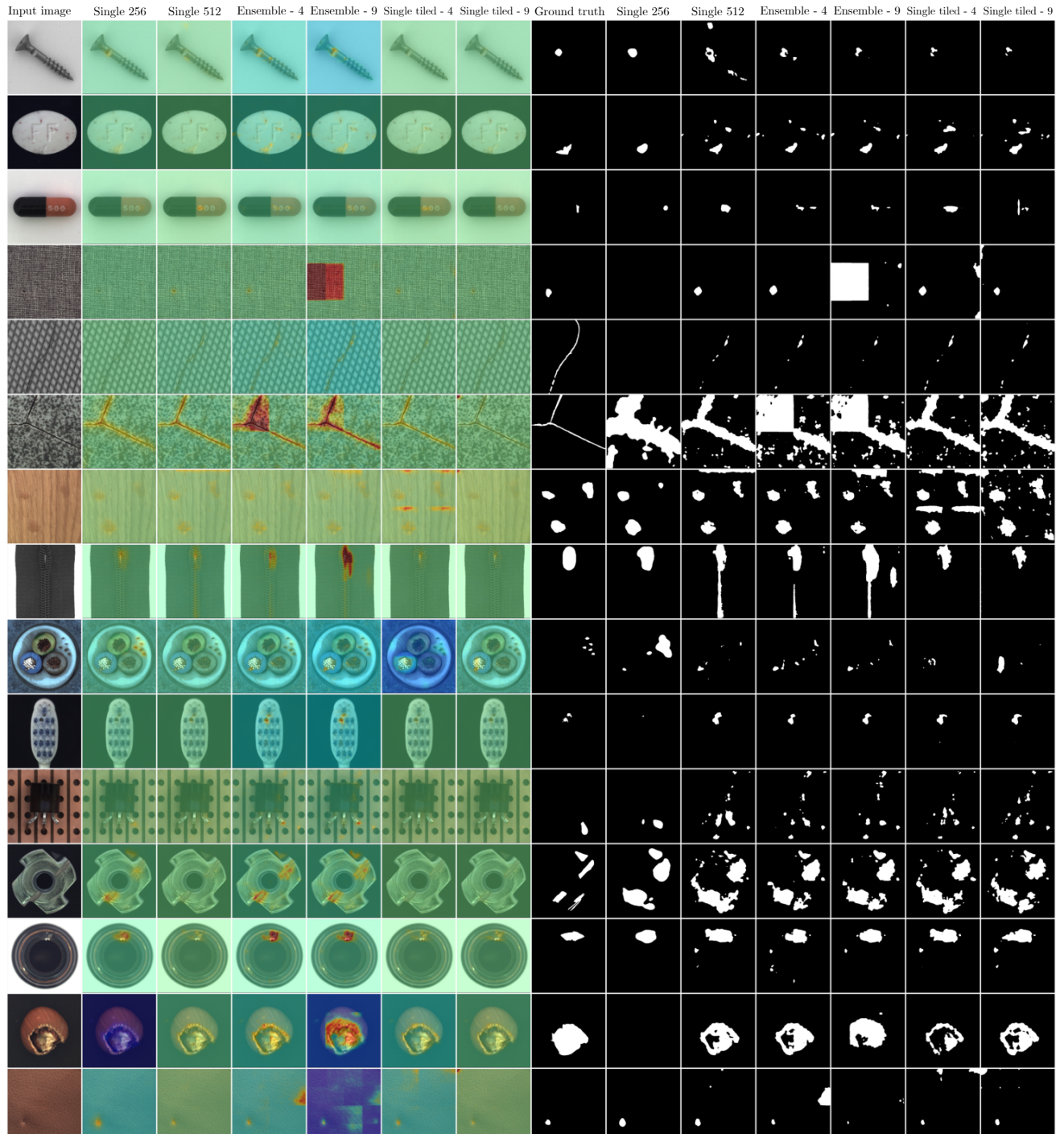


Figure 12. Anomaly maps and segmentation masks for each setup using FastFlow on randomly picked defective image from every category in MVTEC AD.



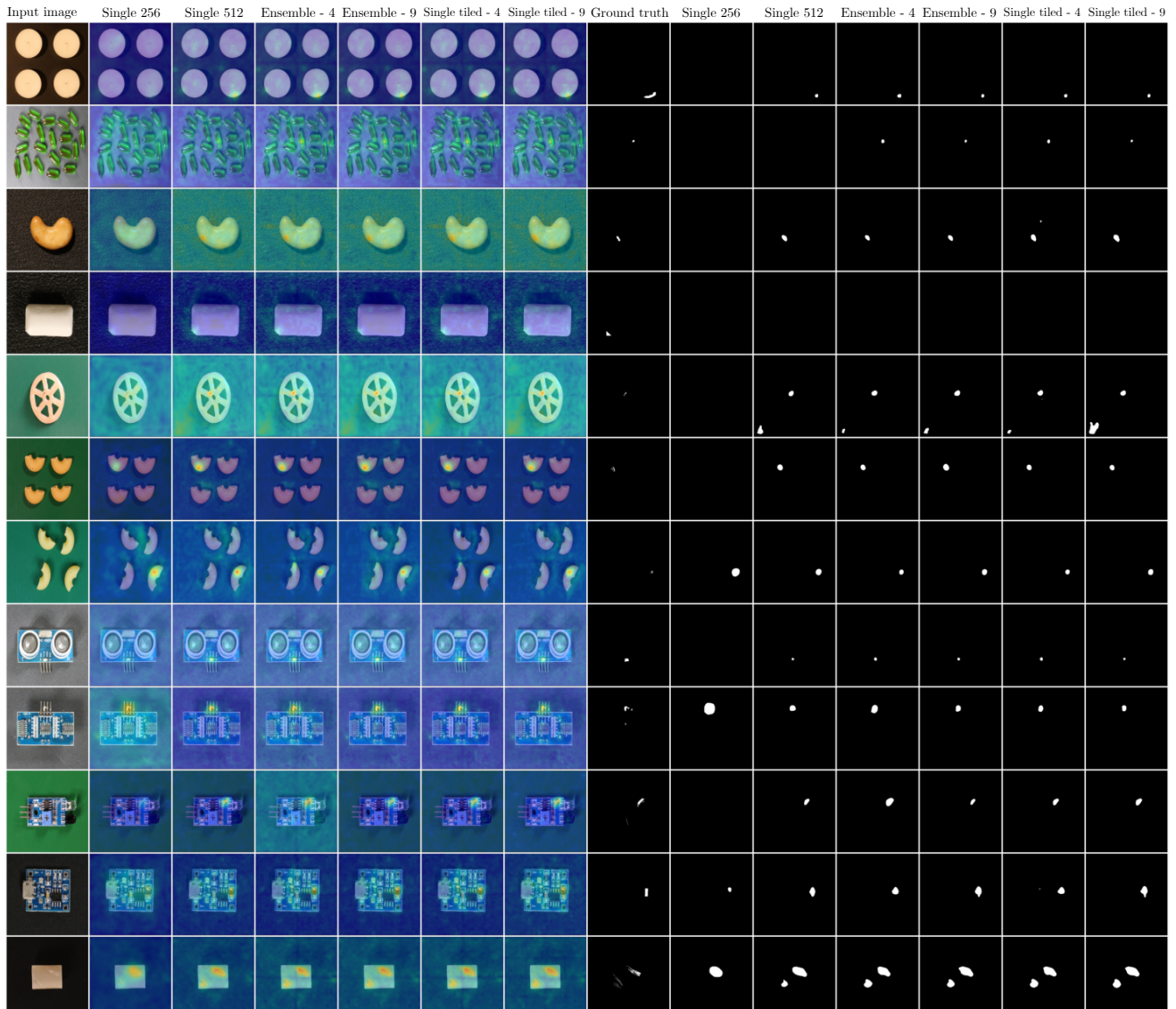


Figure 14. Anomaly maps and segmentation masks for each setup using PatchCore on randomly picked defective image from every category in VisA.

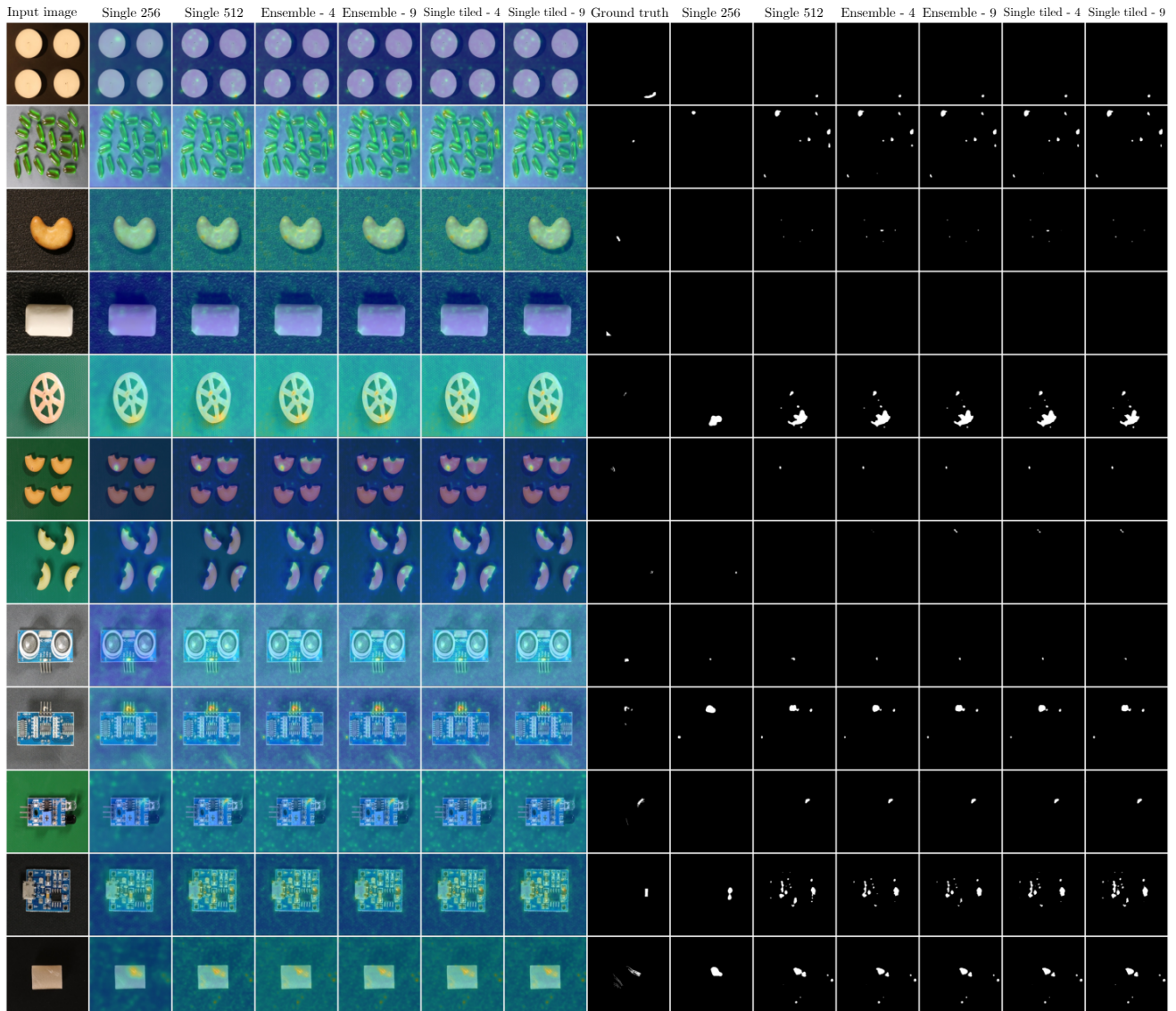


Figure 15. Anomaly maps and segmentation masks for each setup using Padim on randomly picked defective image from every category in VisA.

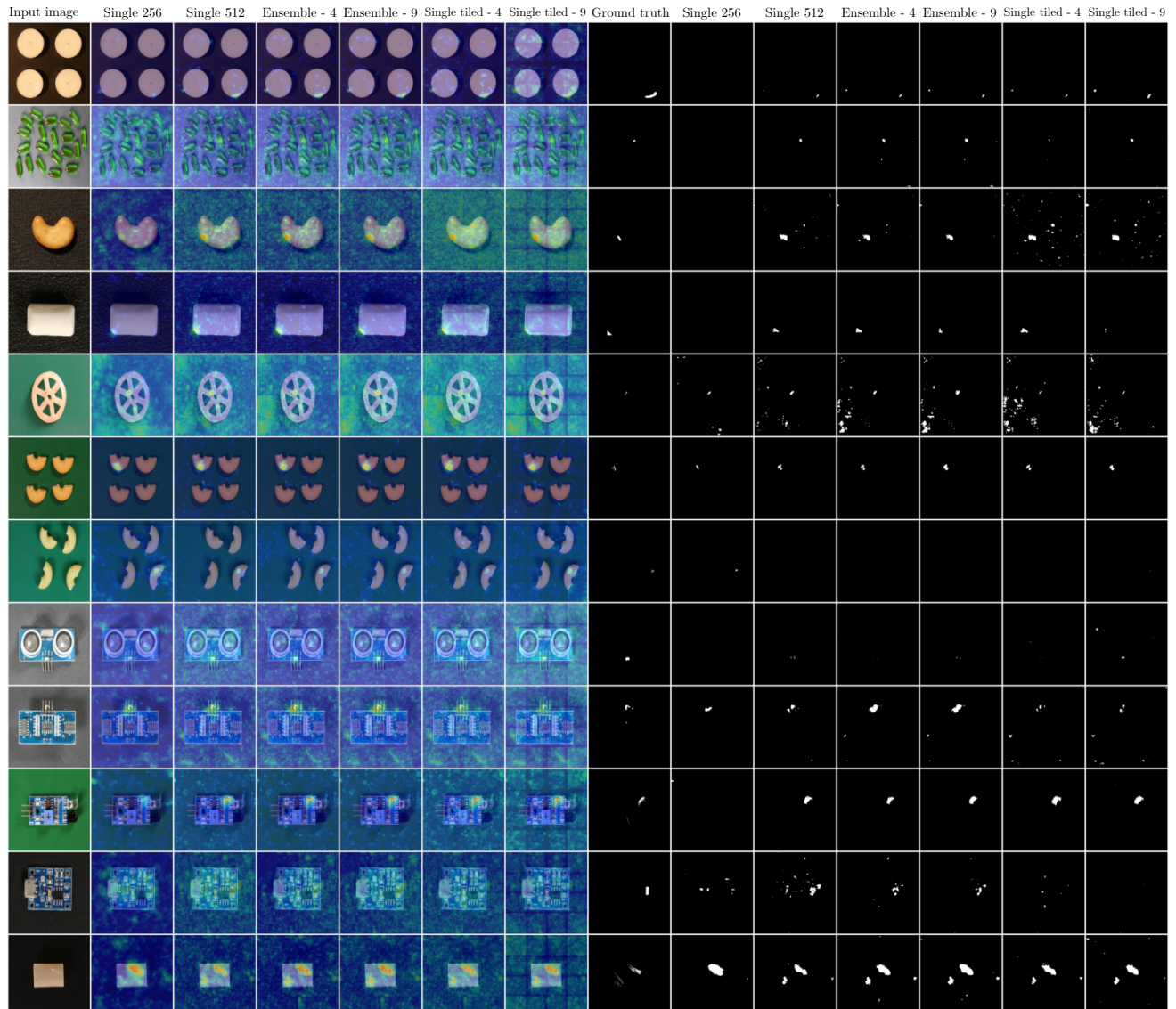


Figure 16. Anomaly maps and segmentation masks for each setup using FastFlow on randomly picked defective image from every category in VisA.

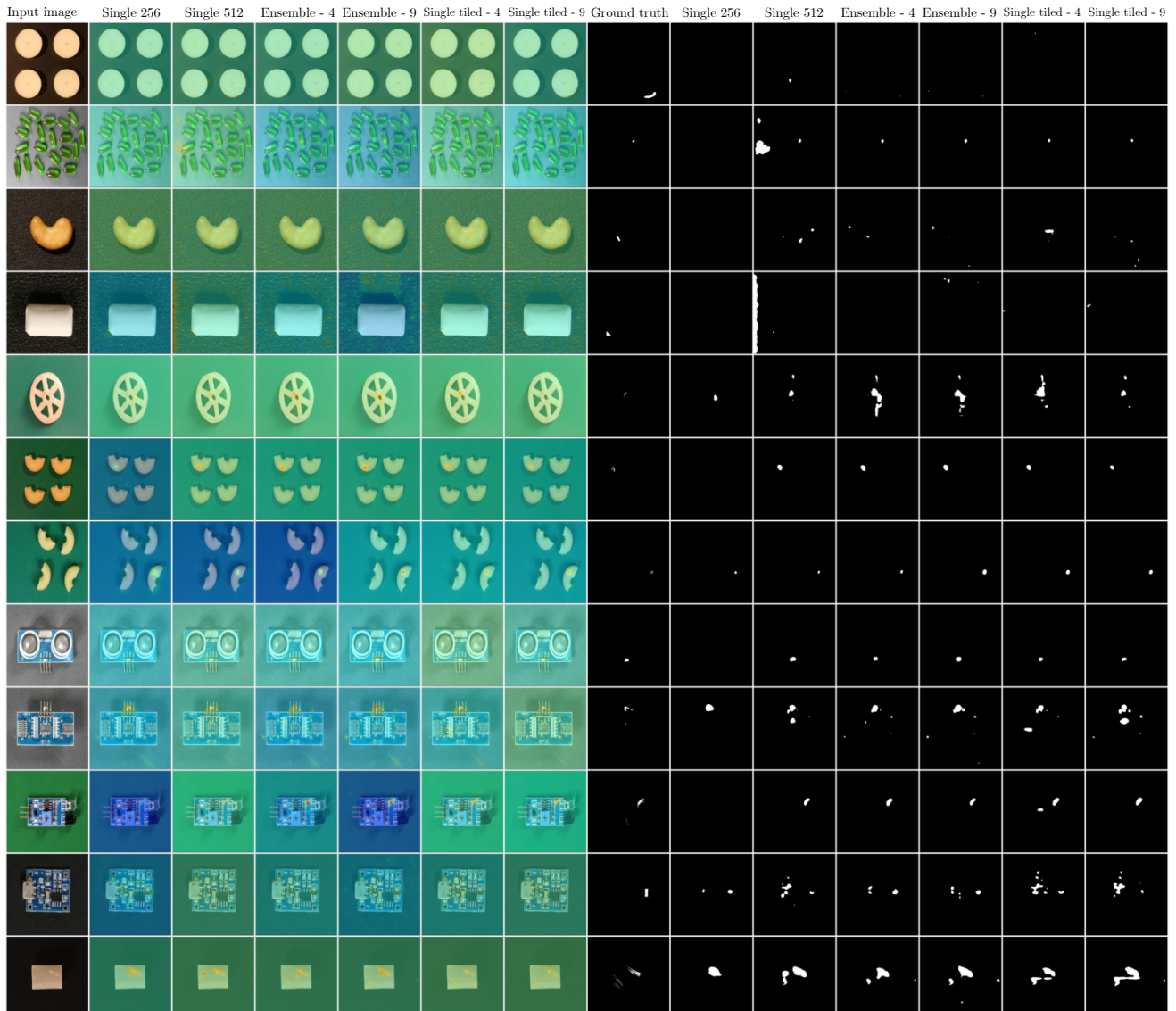


Figure 17. Anomaly maps and segmentation masks for each setup using Reverse Distillation on randomly picked defective image from every category in VisA.