Supplementary Material for A Hierarchical Transformation-Discriminating Generative Model for Few Shot Anomaly Detection

Shelly Sheynin[1*] Sagie Benaim[1] Lior Wolf[1,2]
[1]The School of Computer Science, Tel Aviv University
[2]Facebook AI Research

1. Transformations

As discussed in Sec. 3.1 of the main text, due to memory constraints, we use a subset of $M = 54$ transformations. Let $T_{rgb2gray}$ be the transformation of an image from RGB to grayscale. T^1_{flip} is a horizonal flip and T^0_{flip} is the identity transformation. $T^b_{translate_x}$ is the horizontal translation along the x-axis by 15% of the image width, to the left ($b = 1$) or to the right ($b = -1$). $b = 0$ is the identity translation. $T^c_{translate_y}$ is the vertical translation along the y-axis by 15% of the image height, upwards ($c = 1$) or downwards ($c = -1$). $c = 0$ is the identity translation. T^d_{rotate} stands for the rotation by d degrees, where $d \in \{0, 90, 180, 270\}$.
$T_1, \dots T_{32}$: $T^a_{flip} \circ T^b_{translate_x} \circ T^c_{translate_y} \circ T^d_{rotate}$ where $a \in \{0, 1\}$, $b \in \{0, 1\}$, $c \in \{0, 1\}$ and $d \in \{0, 90, 180, 270\}$.
$T_{33}, \dots T_{38}$: $T^a_{flip} \circ T^b_{translate_x} \circ T^c_{translate_y}$, where $a \in \{0, 1\}$, $b \in \{-1, 1, 0\}$ and $c = -1$.
$T_{39}, \dots T_{42}$: $T^a_{flip} \circ T^b_{translate_x} \circ T^c_{translate_y}$, where $a \in \{0, 1\}$, $b = -1$ and $c \in \{0, 1\}$.
$T_{43}, \dots T_{50}$: $T_{rgb2gray} \circ T^a_{flip} \circ T^d_{rotate}$, where $a \in \{0, 1\}$ and $d \in \{0, 90, 180, 270\}$.
T_{51}, T_{52}: $T_{rgb2gray} \circ T^b_{translate_x}$ where $b \in \{-1, 1\}$.
T_{53}, T_{54}: $T_{rgb2gray} \circ T^c_{translate_y}$ where $c \in \{-1, 1\}$.

2. Detailed Per-Class Results

In Sec. 4 of the main text, for the task of anomaly detection and defect detection, we report mean AUC values and mean standard deviation values, over all classes. Detailed per-class results are provided here.

In particular, full anomaly detection results for the datasets of Paris, CIFAR10, FashionMNIST and MNIST are given in Tab. 1 (one-shot), Tab. 2 (five-shot) and Tab. 3 (ten-shot). This supplements Fig. 2 of the main text. 50-shot and 80-shot results for CIFAR10 are given in Tab. 4. Together with tables 1-3, this supplements Fig. 4 of the main text.

Tab. 5 gives the full defect detection results on MVTec for one-shot, five-shot and ten-shot settings, supplementing Fig. 5 of the main text.

Tab. 6, gives the ablation analysis performed on CIFAR10, for both the one-shot and five-shot settings, supplementing Tab. 1 and discussed in Sec. 4.3 of the main text.

Lastly, Tab. 7, shows the effect of using a different percentage of patches for detect detection, supplementing Fig. 7 and discussed in Sec. 4.3 of the main text.

*Equal contribution

Class	PatchSVDD	DROCC	DeepSVDD	GEOM	GOAD	**Ours**
			PARIS			
Defense	57.0 ± 3.5	53.2 ± 8.0	50.1 ±5.0	59.4 ± 3.1	47.8 ± 5.9	**65.6 ± 9.9**
Eiffel	46.2 ± 6.2	53.3 ±7.9	45.8 ± 6.7	46.9 ± 6.0	54.6 ± 3.3	**57.8 ±4.5**
Invalides	46.0 ± 8.2	52.3 ± 5.1	50.3 ± 6.4	56.1 ± 2.9	52.9 ± 3.8	**71.0 ± 6.4**
Louvre	47.3 ± 5.5	57.5 ± 3.3	50.1 ± 3.0	53.7 ± 4.5	52.6 ± 3.1	**61.7 ± 7.2**
Moulinrouge	60.4 ± 10.2	43.7 ± 6.4	64.6 ± 2.1	9.4 ± 7.6	51.6 ± 5.9	**72.8 ±6.8**
Museedorsay	55.7 ± 8.0	42.3 ± 3.7	85.9 ± 1.9	**85.1 ± 2.7**	49.3 ± 16.8	73.1 ± 10.2
Notredame	52.3 ± 4.8	46.9 ± 4.6	58.5 ± 3.1	52.2 ± 5.1	49.8 ± 5.7	**66.0 ± 9.4**
Pantheon	62.8 ± 3.7	44.2 ± 6.6	54.8 ± 12.0	58.5 ± 7.8	49.9 ± 5.6	**73.8 ± 8.8**
Pompidou	56.7 ± 10.2	47.8 ± 8.9	65.5 ± 3.6	65.3 ± 8.1	49 ± 7.8	**68.3 ±9.4**
Sacrecoeur	55.1 ± 7.9	51.8 ± 8.4	52.1 ± 4.3	48.4 ± 6.7	52 ± 3.5	**61.6 ± 8.5**
Triomphe	57.5 ± 3.8	44.2 ± 5.9	59.2 ± 5.4	48.9 ± 5.6	49 ± 5.7	**60.8 ± 5.5**
Avg	54.3 ± 6.5	48.8 ± 6.2	57.9 ± 4.9	56.7 ± 5.5	50.8 ± 6.1	**66.6 ± 7.9**
			CIFAR10			
Plane	50.1 ± 15.8	54.9 ± 9.3	29.8 ± 5.5	49.5 ± 11.1	59.8 ± 8.3	**67.2 ± 5.8**
Car	51.4 ± 6.3	35.2 ± 7.4	**81.0 ± 13.5**	53.3 ± 5.7	58.2± 5.8	65.6 ± 5.9
Bird	46.5 ± 8.6	**59.5 ±3.7**	50.4 ± 22.4	54.7 ± 6.6	53.1 ± 9.1	55.9 ± 5.7
Cat	48.9 ± 6.1	52.3 ± 5.5	58.8 ± 12.7	53.2 ± 4.4	46.4 ± 8.2	**58.9 ± 6.2**
Deer	46.5 ± 10.7	65.7 ± 5.9	56.4 ± 10.6	**67.3 ± 6.4**	55.9 ±10.7	67.2 ± 4.5
Dog	54.4 ± 6.3	52.7 ± 8.1	22.8 ± 2.0	50.9 ± 2.7	53.7 ± 6.0	**63.7 ± 7.7**
Frog	53.4 ± 17.4	53.1 ± 6.8	60.2 ± 15.9	60.7 ± 8.6	53.6 ± 9.9	**70.2 ± 5.1**
Horse	52.7 ± 5.1	43.5 ± 6.1	**78.6 ± 13.1**	56.0 ± 4.6	54.8 ± 7.6	63.8 ± 5.2
Ship	55.6 ± 13.5	57.3 ± 9.0	70.8 ± 7.9	68.1 ± 10.4	67.4 ± 6.4	**71.3 ± 7.2**
Truck	60.8 ± 8.1	33.6 ± 5.2	**69.8 ± 6.6**	57.2 ± 12.0	61.1 ± 5.5	65.3 ± 5.2
Avg	52.0 ± 9.8	50.8 ± 6.7	57.9 ±11.0	57.1 ± 7.3	56.4± 7.8	**64.9 ± 5.9**
			MNIST			
0	46.6 ± 19.4	63.4 ± 14.1	**78.6 ± 12.7**	73.1 ± 5.9	77.2 ± 9.1	75.2 ± 5.8
1	82.5 ± 18.1	81.6 ± 5.6	69.8 ± 7.9	**88.7 ± 5.0**	80.2 ± 18.3	79.2 ± 6.9
2	56.0 ± 6.3	43.0 ± 9.2	67.0 ± 7.9	60.9 ± 14.4	72.5 ± 4.4	**74.3 ± 3.4**
3	63.1 ± 1.7	54.3 ± 8.7	61.8 ± 29.4	77.0 ± 3.2	80.7 ± 6.9	**94.3 ± 4.8**
4	53.6 ± 8.4	59.1 ± 10.4	63.2 ± 5.1	66.9 ± 8.4	63.8 ± 5.9	**81.6 ± 7.6**
5	60.2 ± 6.6	61.9 ± 9.5	65.2 ± 4.0	72.1 ± 8.3	54.5 ± 12.8	**80.3 ± 7.2**
6	59.0 ± 11.7	65.5 ± 6.7	78.2 ± 4.9	66.2 ± 20.2	70.2 ± 4.2	**85.7 ± 3.4**
7	49.2 ± 14.0	70.1 ± 12.0	70.2 ± 3.2	69.5 ± 8.9	66.4 ± 10.3	**76.9 ± 4.0**
8	53.7 ± 15.6	57.5 ± 7.4	**72.4 ± 3.7**	56.2 ± 2.1	71.7 ± 4.7	71.5 ± 6.2
9	56.3 ± 8.9	70.3 ± 7.2	61.8 ± 9.0	67.6 ± 4.3	59.8 ± 5.1	**73.5 ± 6.4**
Avg	58.0 ± 11.1	62.7 ± 9.1	68.8 ± 8.8	69.8 ± 8.1	69.7 ± 8.2	**79.3 ± 5.6**
			FashionMNIST			
T-shirt	58.5 ± 5.6	69.7 ± 8.1	**83.5 ± 6.9**	79.7 ± 2.9	71.8 ± 14.1	77.3 ± 4.3
Trouser	32.0 ± 18.6	95.2 ± 1.7	63.5 ± 9.2	55.5 ± 4.3	76.0 ±3.7	**97.2 ± 1.4**
Pullover	73.7 ± 8.7	68.0 ± 8.9	66.7 ± 7.3	56.9 ± 12.1	69.1 ±5.6	**80.3 ± 4.2**
Dress	43.0 ± 9.8	80.9 ± 6.6	63.1 ± 16.3	72.5 ± 10.5	76.9 ±13.1	**83.8 ± 4.0**
Coat	73.3 ± 4.9	63.5 ± 15.1	63.6 ± 12.0	52.2 ± 16.1	66.2 ±18.8	**79.0 ± 9.2**
Sandals	39.1 ± 26.4	74.3 ± 8.4	64.9 ± 9.8	78.5 ± 9.7	57.9 ±10.1	**85.5 ± 4.5**
Shirt	70.2 ± 2.7	64.9 ± 8.9	**75.1 ± 6.2**	56.1 ± 5.6	72.8 ±3.1	69.0 ± 2.4
Sneaker	58.1 ± 25.7	90.5 ± 9.1	59.1 ± 12.0	92.6 ± 2.1	69.2 ±1.7	**97.9 ± 0.7**
Bag	70.2 ± 2.1	53.6 ± 7.4	72.4 ± 3.3	**92.2 ± 9.1**	71.7 ±9.9	77.2 ± 15.4
Ankle-Boot	73.2 ± 9.2	81.9 ± 14.7	71.2 ± 8.5	62.0 ± 5.8	61.6 ±10.6	**91.7 ± 6.1**
Avg	59.1 ± 11.4	74.2 ± 8.9	68.3 ± 9.2	69.8 ± 7.8	69.3 ± 9.1	**83.9 ± 5.2**

Table 1. Average AUC (with standard deviation) for **One-Shot** anomaly detection experiments on Paris, CIFAR10, FashionMNIST and MNIST datasets.

Class	PatchSVDD	DROCC	DeepSVDD	GEOM	GOAD	**Ours**
			PARIS			
Defense	51.5 ± 3.3	**69.3 ± 4.5**	62.1 ± 3.3	59.4 ± 2.7	52.8 ± 5.1	67.8 ± 3.4
Eiffel	51.2 ± 4.3	66.8 ± 3.5	55.4 ± 2.8	44.1 ± 6.6	53.0 ± 3.0	**67.0 ± 2.7**
Invalides	45.2 ± 2.1	62.9 ± 6.4	66.6 ± 4.9	59.2 ± 2.0	52.2 ± 4.5	**80.8 ± 2.5**
Louvre	41.1 ± 2.0	66.6 ± 3.3	60.4 ± 4.3	53.3 ± 1.8	52.3 ± 2.7	**72.5 ± 2.8**
Moulinrouge	59.6 ± 3.1	44.1 ± 5.4	62.4 ± 5.1	49.0 ± 0.3	45.9 ± 7.3	**84.5 ± 2.4**
Museedorsay	53.9 ± 2.7	46.8 ± 9.6	88.0 ± 3.3	88.7 ± 3.2	43.0 ± 15.2	**89.6 ±1.8**
Notredame	47.7 ± 2.8	48.7 ± 6.6	62.6 ± 3.5	58.4 ± 1.7	48.2 ±5.7	**79.7 ± 4.0**
Pantheon	58.4 ± 5.2	49.2 ± 6.6	74.9 ± 2.5	60.7 ± 1.8	52.3 ± 2.3	**86.1 ± 2.1**
Pompidou	58.7 ± 3.4	45.7 ± 7.4	75.6 ± 3.4	70.0 ± 2.7	54.1 ±5.8	**90.3 ± 3.4**
Sacrecoeur	46.9 ± 7.8	58.5 ± 5.1	62.5 ± 4.3	48.1 ± 0.8	54.8 ± 7.2	**81.6 ± 2.7**
Triomphe	55.5 ± 2.6	47.4 ± 4.4	64.2 ± 9.4	52.4 ± 1.7	51.1 ± 3.8	**78.6 ± 4.8**
Avg	51.8 ± 3.6	55.1 ± 5.7	66.8 ± 4.2	58.5 ± 2.3	50.9 ± 5.7	**79.8 ± 3.0**
			CIFAR10			
Plane	40.8 ± 13.8	69.0 ± 4.8	35.8 ± 3.1	62.3 ± 9.0	59.5 ± 3.0	**69.2 ± 2.8**
Car	59.5 ± 4.5	39.6 ± 8.8	74.6 ± 5.3	65.5 ± 7.7	68.8 ±10.1	**77.0 ± 1.8**
Bird	45.7 ± 6.0	**60.9 ± 3.4**	48.4 ± 5.2	52.4 ± 4.8	49.4 ± 3.4	58.4 ± 2.3
Cat	55.6 ± 3.2	56.5 ± 4.9	54.4 ± 10.7	54.0 ± 5.2	49.0 ± 7.0	**58.7 ± 4.3**
Deer	44.5 ± 5.3	57.9 ± 5.4	51.4 ± 5.8	63.6 ± 7.6	48.8 ± 5.5	**66.4 ± 4.3**
Dog	54.4 ± 3.0	59.4 ± 6.1	**70.4 ± 6.1**	55.5 ± 3.4	60.9 ± 11.5	61.8 ± 3.2
Frog	53.7 ± 5.6	50.2 ± 7.7	56.0 ± 5.7	58.5 ± 6.9	51.5 ± 2.7	**72.6 ± 4.4**
Horse	55.4 ± 3.2	43.6 ± 4.3	**69.7 ± 5.9**	64.2 ±3.1	62.0 ± 4.9	68.6 ± 2.8
Ship	48.3 ± 10.3	67.5 ± 6.7	73.4 ± 4.4	75.5 ±7.9	74.2 ± 3.6	**80.2 ± 3.2**
Truck	62.6 ± 2.2	35.9 ± 5.5	70.3 ± 4.7	67.5 ± 4.0	**74.2 ± 1.7**	62.1 ± 4.5
Avg	52.1 ± 5.7	54.1 ± 5.8	60.4 ± 5.7	59.5 ± 6.0	59.9 ± 5.3	**67.5 ± 3.4**
			MNIST			
0	76.6 ± 2.5	70.7 ± 9.0	86.8 ± 3.2	71.3 ± 6.3	**87.4 ± 8.0**	79.5 ± 3.8
1	31.5 ± 9.9	80.6 ± 7.5	89.6 ± 5.3	**96.2 ± 0.5**	89.2 ± 6.8	85.5 ± 6.7
2	73.5 ± 5.2	56.4 ± 10.2	73.4 ± 5.2	78.0 ± 2.6	71.3 ± 7.3	**81.6 ± 4.2**
3	71.0 ± 5.5	63.4 ± 5.3	77.2 ± 10.7	85.5 ± 0.7	80.9 ± 4.6	**96.6 ± 1.0**
4	45.0 ± 5.8	69.6 ± 3.5	76.8 ± 5.8	66.4 ± 5.6	70.3 ± 5.2	**84.7 ± 1.3**
5	62.6 ± 3.0	69.1 ± 7.2	65.6 ± 6.1	79.0 ± 8.5	70.4 ± 12.8	**89.3 ± 2.4**
6	55.5 ± 4.3	73.9 ± 7.5	80.0 ± 5.7	76.1 ± 4.6	72.6 ± 3.9	**92.4 ± 0.9**
7	35.2 ± 8.3	80.4 ± 7.2	81.0 ± 5.9	80.3 ± 3.8	67.1 ± 5.7	**82.0 ± 3.7**
8	64.9 ± 6.5	64.4 ± 4.6	**82.2 ± 4.4**	70.7 ± 4.0	73.4 ± 5.1	79.4 ± 3.4
9	42.2 ± 6.4	76.7 ± 6.3	79.2 ± 4.7	65.7 ± 1.9	72.5 ± 3.9	**87.5 ± 3.2**
Avg	55.8 ± 5.7	70.5 ± 6.8	79.1 ± 5.7	76.9 ±3.9	75.5 ± 6.3	**85.9± 3.1**
			FashionMNIST			
T-shirt	52.8 ± 6.0	85.2 ± 3.1	89.6 ± 2.4	**92.4 ± 2.7**	79.8 ± 2.7	85.2 ± 1.7
Trouser	42.2 ± 10.7	94.2 ± 2.2	84.8 ± 7.1	74.7 ± 2.7	97.8 ± 0.5	**98.4 ± 0.5**
Pullover	64.7 ± 7.0	80.5 ± 3.3	72.3 ± 7.1	84.3 ± 3.6	**86.4 ±2.2**	85.8 ± 3.5
Dress	41.7 ± 7.6	86.3 ± 4.4	77.8 ± 4.9	87.8 ± 1.0	85.1 ± 2.0	**89.1 ± 2.4**
Coat	62.8 ± 6.1	81.5 ± 3.9	76.8 ± 7.0	78.4 ± 2.0	83.8 ± 1.9	**88.4 ± 1.5**
Sandals	60.1 ± 8.5	78.1 ± 15.0	63.8 ± 8.0	83.7 ± 2.0	65.9 ± 7.6	**88.6 ± 2.1**
Shirt	54.8 ± 6.6	72.0 ± 3.5	**81.5 ± 8.0**	73.8 ± 3.7	68.0 ± 3.1	78.2 ± 1.7
Sneaker	53.0 ± 10.7	93.2 ± 1.6	81.6 ± 7.4	94.6 ± 1.9	94.4 ± 1.0	**99.1 ± 0.3**
Bag	53.4 ± 5.3	67.6 ± 8.7	80.1 ± 3.1	**96.6 ± 1.4**	77.5 ± 6.0	92.9 ± 4.1
Ankle-Boot	56.4 ± 9.1	90.0 ± 6.9	82.1 ± 5.4	83.7 ± 4.7	**96.6 ± 1.5**	96.5 ± 1.0
Avg	54.2 ± 7.8	82.9 ± 5.3	79.0 ± 6.0	85.0 ± 2.6	83.5 ± 2.9	**90.2 ± 1.9**

Table 2. Average AUC (with standard deviation) for **Five-Shot** anomaly detection experiments on Paris, CIFAR10, FashionMNIST and MNIST datasets.

Class	PatchSVDD	DROCC	DeepSVDD	GEOM	GOAD	**Ours**
			PARIS			
Defense	54.2 ± 4.1	**72.0 ± 4.5**	62.7 ± 2.3	57.2 ± 1.6	49.5 ± 3.3	67.9 ± 3.2
Eiffel	48.1 ± 4.8	**73.6 ± 3.3**	59.4 ± 1.9	47.2 ± 5.6	50.0 ± 0.0	71.2 ± 3.7
Invalides	42.7 ± 2.8	68.0 ± 4.8	67.4 ± 2.0	63.4 ± 1.0	49.9 ± 0.0	**84.9 ± 1.2**
Louvre	39.4 ± 1.9	73.5 ± 4.3	60.6 ± 3.3	53.4 ± 1.8	49.2 ± 1.6	**74.9 ± 2.5**
Moulinrouge	59.4 ± 4.3	46.6 ± 2.5	63.6 ± 3.4	51.7 ± 0.8	48.8 ± 2.2	**87.0 ± 3.1**
Museedorsay	58.4 ± 4.0	52.3 ± 3.2	89.4 ± 2.0	86.0 ± 5.1	55.5 ± 6.9	**90.7 ± 2.2**
Notredame	52.0 ± 3.2	52.5 ± 4.6	65.8 ± 2.9	55.3 ± 1.0	49.9 ± 3.1	**83.0 ± 2.9**
Pantheon	54.5 ± 4.9	57.2 ± 6.4	75.7 ± 1.6	62.3 ± 0.8	50.3 ± 0.7	**89.9 ± 2.1**
Pompidou	59.9 ±5.6	50.3 ± 4.6	77.6 ± 6.0	69.2 ± 0.8	50.2 ± 2.7	**95.4 ± 1.3**
Sacrecoeur	48.1 ± 2.4	66.6 ± 4.9	66.1 ± 3.4	47.1 ± 4.1	51.2 ± 3.1	**84.5 ± 1.9**
Triomphe	59.5 ± 4.0	51.7 ± 3.9	63.3 ± 10.3	53.4 ± 0.6	49.0 ± 1.1	**79.8 ± 3.4**
Avg	52.4 ± 3.8	60.4 ± 4.3	68.3 ± 3.5	58.8 ± 2.1	50.3 ± 2.5	**82.6 ± 2.5**
			CIFAR10			
Plane	40.8 ± 9.4	**71.9 ± 2.2**	39.6 ± 6.3	66.7 ± 8.8	61.5±2.4	69.1 ± 1.6
Car	59.9 ± 3.4	42.8 ± 8.2	64.0 ± 9.9	74.3 ± 2.7	68.7±6.1	**80.7 ± 2.9**
Bird	44.8 ± 3.9	**62.4 ± 4.4**	42.4 ± 11.1	54.4 ± 6.7	51.3±3.2	58.5 ± 2.5
Cat	53.8 ± 3.7	61.7 ± 4.3	54.3 ± 7.3	52.5 ± 5.7	50.4±4.8	**63.2 ± 2.8**
Deer	50.1 ± 4.9	62.0 ± 3.2	50.0 ± 8.7	54.1 ± 5.8	52.1 ± 7.1	**64.2 ± 2.2**
Dog	53.3 ± 4.3	61.3 ± 3.9	**81.6 ± 3.9**	60.5 ± 5.1	57.1±5.7	65.4 ± 5.6
Frog	50.4 ± 4.7	48.2 ± 4.6	58.0 ± 11.9	60.3 ± 6.8	55.3±2.3	**71.9 ± 3.3**
Horse	53.9 ± 2.9	51.6 ± 3.1	**76.8 ± 5.4**	62.9 ± 4.5	61.7±3.2	73.7 ± 2.8
Ship	46.0 ± 8.5	72.6 ± 3.4	71.6 ± 3.9	67.8 ± 8.7	71.3±2.3	**82.9 ± 0.8**
Truck	52.6 ± 4.2	39.3 ± 3.5	73.4 ± 4.2	70.3 ± 4.0	**75.2±2.5**	72.6 ± 2.9
Avg	50.5 ± 5.0	57.4 ± 4.1	61.1 ± 7.3	62.4 ± 5.9	60.5 ± 4.0	**70.2 ± 2.7**
			MNIST			
0	75.0 ± 4.7	80.3 ± 8.0	**91.6 ± 1.1**	75.0 ± 1.0	72.6 ± 6.8	80.1 ± 4.6
1	59.3 ± 13.1	78.0 ± 12.5	89.0 ± 5.8	**96.2 ± 0.4**	90.9 ± 3.2	88.8 ± 3.1
2	58.6 ± 5.7	58.8 ± 13.7	73.0 ± 8.8	80.1 ± 2.4	68 ± 5.6	**85.2 ± 4.3**
3	62.0 ± 6.1	66.9 ± 7.6	82.4 ± 3.2	91.0 ± 0.5	73.2 ± 9.3	**96.3 ± 0.9**
4	53.7 ± 7.8	71.2 ± 9.8	85.6 ± 0.9	79.3 ± 1.1	69.1 ± 6.2	**89.1 ± 1.6**
5	59.8 ± 5.2	63.7 ± 8.2	72.4 ± 4.0	87.2 ± 0.6	62.1 ± 13.4	**87.4 ± 3.3**
6	53.9 ± 4.8	74.0 ± 14.3	88.2 ± 2.5	83.6 ± 2.8	73.9 ± 4.6	**92.2 ± 1.6**
7	50.4 ± 6.5	77.1 ± 10.8	80.0 ± 7.5	78.4 ± 0.7	63 ± 6.5	**84.2 ± 4.2**
8	61.5 ± 4.8	69.1 ± 4.8	**81.0 ± 0.9**	64.7 ± 4.0	77.8 ± 5.4	78.2 ± 2.3
9	50.6 ± 7.0	82.9 ± 6.5	82.6 ± 3.2	78.7 ± 4.8	67.5 ± 6.2	**90.2 ± 1.4**
Avg	58.5 ± 6.6	72.2 ± 9.6	82.6 ± 3.8	81.4 ± 1.8	71.8 ± 6.7	**87.2 ± 2.7**
			FashionMNIST			
T-shirt	50.9 ± 5.5	86.8 ± 3.3	83.5 ± 2.1	**97.5 ± 0.5**	79.7 ± 3.0	86.5 ± 1.1
Trouser	52.9 ± 12.7	94.4 ± 4.0	63.6 ± 4.6	80.2 ± 0.75	97.5 ± 1.7	**99.0 ± 0.2**
Pullover	69.2 ± 5.8	81.2 ± 3.4	66.7 ± 2.8	**90.1 ± 1.6**	89.2 ± 1.0	86.5 ± 1.1
Dress	36.9 ± 8.5	88.1 ± 3.6	63.1 ± 0.8	91 ± 1.7	87.3 ± 1.5	**91.7 ± 1.3**
Coat	67.9 ±7.6	84.7 ± 3.5	63.6 ± 4.6	88.5 ± 4.3	86.9 ± 0.9	**88.9 ± 1.2**
Sandals	54.1 ± 8.6	83.0 ± 12.4	64.9 ± 6.4	86.3 ± 1.0	72.5 ± 13.1	**89.1 ± 1.6**
Shirt	55.6 ± 7.8	74.8 ± 3.8	75.1 ± 3.9	**79.5 ± 2.5**	76.3 ± 2.0	78.5 ± 0.8
Sneaker	56.8 ± 7.8	93.3 ± 1.4	59.1 ± 3.9	97.8 ± 0.4	96.3 ± 1.2	**99.0 ± 0.2**
Bag	56.1 ± 8.1	73.8 ± 10.2	72.4 ± 4.6	**98.4 ± 0.3**	77.9 ±2.5	94.5 ± 0.4
Ankle-Boot	60.3 ± 12.8	85.3 ± 3.7	71.2 ± 1.1	89.6 ± 0.7	97.5 ± 1.0	**98.0 ± 0.6**
Avg	56.1 ± 8.5	84.5 ± 4.9	68.3 ± 3.5	89.9 ± 1.4	86.1 ± 2.8	**91.2 ± 0.9**

Table 3. Average AUC (with standard deviation) for **Ten-Shot** anomaly detection experiments on Paris, CIFAR10, FashionMNIST and MNIST datasets.

Class	PatchSVDD	DROCC	DeepSVDD	GEOM	GOAD	**Ours**
			CIFAR10 (50-Shot)			
Plane	36.7 ± 6.7	$\mathbf{76.2 \pm 2.6}$	57.3 ± 2.6	67.8 ± 2.9	55.6 ± 6.4	75.9 ± 5.9
Car	65.5 ± 3.6	44.7 ± 3.0	64.1 ± 1.6	82.4 ± 1.3	54.3 ± 7.9	$\mathbf{86.2 \pm 1.1}$
Bird	38.1 ± 2.1	$\mathbf{66.3 \pm 1.2}$	46.5 ± 2.2	60.3 ± 3.1	52.0 ± 2.1	57.3 ± 1.9
Cat	51.3 ± 3.9	$\mathbf{61.4 \pm 4.0}$	58.5 ± 2.2	59.6 ± 5.1	49.8 ± 0.6	60.5 ± 1.0
Deer	46.3 ± 4.2	58.6 ± 2.9	53.7 ± 3.1	57.4 ± 5.2	50.4 ± 0.9	$\mathbf{64.5 \pm 1.0}$
Dog	49.4 ± 3.4	63.3 ± 5.4	61.7 ± 2.3	68.6 ± 2.6	51.8 ± 3.8	$\mathbf{74.7 \pm 2.1}$
Frog	54.0 ± 5.6	45.8 ± 2.6	58.0 ± 2.7	64.8 ± 2.8	50.7 ± 1.0	$\mathbf{73.2 \pm 1.6}$
Horse	55.4 ± 3.1	47.4 ± 2.6	62.3 ± 3.2	72.4 ± 3.1	52.7 ± 5.4	$\mathbf{74.5 \pm 3.3}$
Ship	44.0 ± 2.4	74.7 ± 2.7	75.1 ± 1.1	81.4 ± 1.7	59.3 ± 12.1	$\mathbf{85.6 \pm 0.6}$
Truck	60.7 ± 4.7	37.4 ± 5.12	71.9 ± 1.9	$\mathbf{81.1 \pm 2.1}$	60.4 ± 11.0	76.8 ± 1.2
Avg	50.1 ± 4.0	57.6 ± 3.2	60.9 ± 2.3	69.6 ± 3.0	53.7 ± 5.1	$\mathbf{72.9 \pm 2.0}$
			CIFAR10 (80-Shot)			
Plane	34.0 ± 4.5	$\mathbf{79.0 \pm 0.6}$	60.9 ± 2.1	69.9 ± 1.6	52.1 ± 4.3	74.8 ± 0.3
Car	63.8 ± 6.9	43.2 ± 2.1	60.1 ± 0.8	85.3 ± 0.8	59.2 ± 11.3	$\mathbf{88.0 \pm 1.5}$
Bird	40.0 ± 1.6	$\mathbf{68.2 \pm 0.3}$	44.6 ± 1.2	60.8 ± 2.4	50.7 ± 1.4	62.4 ± 1.2
Cat	54.9 ± 1.3	55.7 ± 4.0	58.7 ± 0.2	$\mathbf{62.9 \pm 1.3}$	53.8 ± 4.6	60.1 ± 1.4
Deer	50.0 ± 1.8	57.2 ± 3.4	56.3 ± 0.8	62.7 ± 0.3	50.1 ± 2.4	$\mathbf{66.1 \pm 0.5}$
Dog	48.2 ± 3.2	64.4 ± 1.9	60.9 ± 1.7	76.5 ± 1.2	52.5 ± 5.0	$\mathbf{78.4 \pm 1.1}$
Frog	57.0 ± 2.3	50.9 ± 6.9	58.5 ± 2.5	69.9 ± 4.0	51.5 ± 7.1	$\mathbf{75.3 \pm 5.4}$
Horse	56.7 ± 1.8	47.6 ± 2.1	60.9 ± 0.1	79.9 ± 0.4	52.1 ± 3.9	$\mathbf{82.3 \pm 0.2}$
Ship	44.0 ± 3.5	77.0 ± 2.1	74.8 ± 0.1	84.0 ± 1.2	70.4 ± 10.5	$\mathbf{87.4 \pm 0.8}$
Truck	61.2 ± 2.9	42.4 ± 1.1	72.1 ± 1.7	$\mathbf{83.4 \pm 0.3}$	69.7 ± 9.9	81.2 ± 0.6
Avg	51.0 ± 3.0	58.5 ± 2.5	60.8 ± 1.1	73.5 ± 1.4	56.2 ± 6.1	$\mathbf{75.6 \pm 1.3}$

Table 4. Average AUC (with standard deviation) for **50-shot** and **80-shot** anomaly detection experiments on CIFAR10.

Class	DifferNet	DROCC	PatchSVDD	DeepSVDD	GEOM	GOAD	**Ours1**	**Ours2**
				MVTec (One-Shot)				
Bottle	**98.2 ± 0.4**	67.2 ± 6.6	60.9 ± 12.3	16.6 ± 5.3	79.0 ± 3.5	51.6 ± 14.0	76.3 ± 6.9	85.0 ± 3.7
Cable	**76.6 ± 5.9**	68.1 ± 4.3	58.8 ± 4.5	39.0 ± 3.5	64.2 ± 1.3	47.9 ± 2.4	72.3 ± 3.7	61.1 ± 7.8
Capsule	57.7 ± 4.6	50.2 ± 6.4	57.9 ± 12.1	44.8 ± 4.4	55.4 ± 2.6	51.2 ± 3.7	56.0 ± 8.4	**62.6 ± 6.7**
Carpet	61.5 ±3.0	71.9 ± 10.6	45.5 ± 18.8	41.2± 18.2	55.0 ± 10.1	48.1 ± 1.9	**72.7 ± 6.7**	**83.7 ± 8.7**
Grid	59.2 ± 5.1	50.0 ± 4.6	37.2 ± 12.2	79.7 ± 8.6	40.1 ± 13.1	9.4 ± 6.8	73.2 ± 9.8	**87.1 ± 5.0**
Hazelnut	**90.7 ± 2.7**	66.4 ± 7.6	46.7 ± 16.1	29.1 ± 4.3	47.8 ± 3.6	47.6 ± 3.2	82.4 ± 8.7	66.5 ± 9.2
Leather	83.4 ± 1.9	79.1 ± 6.5	61.9 ± 15.6	48.0 ± 3.2	33.2 ± 0.5	58.1 ± 6.8	**98.2 ± 0.9**	97.6 ± 1.1
Metalnut	44.4 ± 8.0	51.9 ± 3.6	50.4 ± 13.1	42.6 ± 14.7	52.3 ± 4.2	7.2 ± 6.5	**66.0 ± 11.0**	60.3 ± 8.6
Pill	71.7 ± 4.4	**72.5 ± 4.0**	57.6 ± 8.1	33.5 ± 4.0	67.0 ± 2.3	62.5 ± 8.1	56.5 ± 9.6	66.5 ± 7.0
Screw	61.8 ± 7.7	57.7 ± 9.0	53.7 ± 18.2	70.1 ± 10.8	34.7 ± 11.1	6.3 ±10.0	**93.5 ± 6.2**	92.8 ± 6.0
Tile	**87.3 ± 2.6**	65.6 ± 2.0	57.3 ± 4.7	40.7 ± 2.8	61.0 ± 2.8	6.0 ± 5.4	80.2 ± 8.2	84.4 ± 3.8
Toothbrush	52.1 ± 2.3	**68.9 ± 4.5**	63.7 ± 6.1	35.5 ± 1.5	65.7 ± 6.5	54.4±5.4	67.3 ± 4.7	64.7 ± 11.1
Transistor	47.0 ± 6.5	59.9 ± 3.3	**66.7 ± 14.5**	32.8 ± 4.3	58.1 ± 1.5	61.7 ± 4.4	66.1 ± 7.7	62.7 ± 6.8
Wood	**96.0 ± 2.2**	70.6 ± 14.4	55.7 ± 18.4	44.0 ± 16.4	52.3 ± 1.1	41.8 ± 6.5	89.0 ± 4.2	85.5 ± 7.9
Zipper	52.7 ± 3.7	49.6 ± 7.5	69 ± 5.4	34.9 ± 2.8	58.3 ± 2.8	56.8 ± 4.0	**67.8 ± 6.4**	**73.2 ± 7.7**
Avg	69.4 ± 4.1	63.3 ± 6.3	56.2 ± 12.0	42.1 ± 7.0	54.9 ± 4.5	44.0 ± 5.9	**74.5 ± 6.9**	**75.6 ± 6.7**
				MVTec (Five-Shot)				
Bottle	**98.4 ± 0.2**	68.1 ± 2.6	61.1 ± 12.4	15.7 ± 2.8	80.0 ± 1.2	51.7 ± 10.4	74.1 ± 7.8	90.8 ± 3.7
Cable	**81.3 ± 2.0**	68.7 ± 2.7	49 ± 3.9	32.8 ± 4.9	61.1 ± 3.1	46.3 ± 4.4	75.2 ± 4.8	76.1 ± 4.0
Capsule	59.0 ± 2.2	53.2 ± 5.1	55.1 ± 3.4	45.3 ± 4.7	60.0 ± 2.3	47.7 ± 5.9	52.6 ± 6.5	**64.9 ± 5.6**
Carpet	62.0 ± 2.2	71.6 ± 10.9	46.5 ± 4.1	47.7 ± 10.5	42.2 ± 6.7	44.2 ± 6.9	**73.3 ± 7.6**	65.2 ± 6.4
Grid	56.7 ± 3.9	37.3 ±9.7	41.7 ± 22.1	76.0 ± 11.1	36.8 ± 7.2	21.3 ± 16.4	**76.0 ± 4.9**	**82.4 ± 9.7**
Hazelnut	**93.8 ± 1.0**	70.0 ± 10.9	58.6 ± 17.4	27.7 ± 4.6	31.7 ± 8.2	52.5 ± 3.5	76.8 ± 8.3	84.5 ± 8.8
Leather	83.7 ± 0.8	70.4 ± 7.1	61.6 ± 15.4	43.0 ± 2.0	33.3 ± 0.2	53.2 ± 10.3	**99.0 ± 0.3**	**98.2 ± 0.9**
Metalnut	47.2 ± 3.2	59.7 ± 6.2	48.8 ± 9.1	52.9 ± 6.6	36.8 ± 4.3	59.4 ± 5.6	**69.4 ± 11.4**	**76.4 ± 6.5**
Pill	**79.4 ± 4.4**	74.4 ± 3.5	57.5 ± 10.6	34.4 ± 3.5	59.1 ± 3.1	61.5 ± 11.0	51.2 ± 6.8	63.6 ± 4.1
Screw	73.7 ± 5.1	58.3 ± 2.3	43.4 ± 15.1	69.5 ± 3.8	18.5 ± 5.1	9.3 ± 13.6	**97.7 ± 3.2**	**74.8 ± 1.3**
Tile	**91.1 ± 1.4**	65.7 ± 3.1	49.5 ± 3.0	32.4 ± 3.2	56.9 ± 11.1	58.6 ± 3.9	89.0 ± 4.5	81.0 ± 4.4
Toothbrush	57.3 ± 3.6	67.6 ± 3.6	68.3 ± 11.8	34.9 ± 6.7	72.2 ± 2.1	45.3 ± 4.5	**72.7 ± 8.1**	64.2 ± 7.3
Transistor	55.7 ± 3.9	67.2 ± 4.1	55.3 ± 9.9	30.4 ± 2.6	59.4 ± 2.9	62.8 ±4.0	**78.2 ± 4.2**	**76.2 ± 3.9**
Wood	**96.4 ± 1.9**	77.7 ± 11.9	69.4 ± 14.6	11.0 ± 7.3	66.0 ± 9.8	37.4 ± 9.8	84.5 ± 3.6	96.2 ± 1.8
Zipper	46.1 ± 3.7	45.2 ± 6.1	63.9±6.5	34.4 ± 4.6	59.2 ± 6.2	54.1 ± 8.	61.8 ± 7.2	**73.3 ± 10.7**
Avg	72.1 ± 2.6	63.7 ± 6.0	55.3 ± 10.6	39.2 ± 5.3	51.5 ± 4.9	47.0 ± 7.9	**75.4 ± 5.9**	**77.9 ± 5.3**
				MVTec (Ten-Shot)				
Bottle	**98.2 ± 0.4**	67.7 ± 5.1	65.3 ± 9.6	17.6 ± 3.0	80.1 ± 2.5	86.9 ± 4.5	81.9 ± 6.1	90.5 ± 3.1
Cable	**82.3 ± 1.5**	69.4 ± 3.0	51.1 ± 7.7	32.6 ± 2.5	64.4 ± 0.8	46.0 ± 9.9	73.9 ± 4.2	77.6 ± 3.9
Capsule	58.0 ± 2.1	51.8 ± 6.3	64.4 ± 11.1	44.7 ± 2.9	**65.9 ± 0.8**	47.3 ± 2.0	55.8 ± 7.7	59.3 ± 8.4
Carpet	61.8 ± 1.5	**75.1 ± 16.4**	49.4 ± 7.4	40.0 ± 11.6	41.4 ± 7.0	50.9 ± 8.5	66.9 ± 9.6	63.9 ± 6.8
Grid	58.5 ± 2.1	37.5 ± 17.1	49.8 ± 11.1	67.1 ± 10.6	10.3 ± 6.7	54.0 ± 7.1	**71.0 ± 8.6**	**79.0 ± 5.9**
Hazelnut	**93.2 ± 1.3**	72.7 ± 11.9	37.9 ± 12.0	30.5 ± 5.2	45.1 ± 1.6	49.6 ± 2.7	72.1 ± 8.2	79.3 ± 11.3
Leather	83.4 ± 0.9	79.1 ± 13.8	49.3 ± 15.9	43.5 ± 2.8	32.7 ± 0.8	61.2 ± 5.2	**99.1 ± 0.2**	98.5 ± 0.5
Metalnut	53.4 ± 7.4	59.1 ± 6.6	62.3 ± 12.5	52.4 ± 3.9	49.3 ± 1.4	58.6 ± 6.7	60.4 ± 11.8	74.0 ± 8.4
Pill	**81.8 ± 3.5**	77.6 ± 3.6	65.2 ± 8.4	39.1 ± 3.9	56.1 ± 1.2	64.1 ± 3.0	57.4 ± 10.4	66.5 ± 7.0
Screw	78.3 ± 4.3	84.2 ± 19.8	28.8 ± 21.3	65.2 ± 4.3	8.5 ± 6.3	66.7 ± 0.8	**93.9 ± 8.4**	75.7 ± 19.0
Tile	**91.3 ± 1.2**	64.8 ± 4.2	49.0 ± 3.1	26.0 ± 5.0	62.0 ± 0.3	54.3 ± 3.5	87.6 ± 5.5	81.4 ± 6.9
Toothbrush	57.5 ± 4.0	67.9 ± 3.3	67.3 ± 9.6	38.2 ± 7.6	71.5 ± 0.4	51.3 ± 8.6	**78.9 ± 8.5**	69.5 ± 7.7
Transistor	54.6 ± 3.7	72.5 ± 3.6	60.3 ± 6.2	24.6 ± 4.5	58.9 ± 3.1	56.0 ± 8.4	**74.9 ± 3.7**	79.2 ± 4.7
Wood	**96.2 ± 1.9**	84.0 ± 8.2	47.9 ± 12.3	18.3 ± 11.6	67.7 ± 5.5	37.4 ± 5.9	85.0 ± 5.9	95.8 ± 1.1
Zipper	55.2 ± 6.1	50.0 ± 6.7	66.7 ± 4.8	36.1 ± 4.5	60.9 ± 2.2	53.1 ± 12.3	**72.8 ± 6.5**	**80.4 ± 5.9**
Avg	73.6 ± 2.8	67.6 ± 8.6	54.3 ± 10.2	38.4 ± 5.6	51.6 ± 2.7	55.8 ± 5.9	**75.4 ± 7.0**	**78.0 ± 6.7**

Table 5. Average AUC (with standard deviation) for **One-Shot**, **Five-Shot** and **Ten-Shot** defect detection experiments on MVTec dataset. **Ours1** refers to our method where the standard set of transformations are used, as for anomaly detection. For a fair comparison with DifferNet, we also consider **Ours2**, where only the four rotation are used, as in DifferNet. In the one-shot case, we report the results of using 5% of the patches, while in five-shot and ten-shot case we report the results of using 10% of the patches. The full results of using different percentage of patches are given in Tab. 7

.

Class	**Ours**	(a)	(b)	(c)	(d)	(e)	(f)	(g)
				CIFAR10 (One-Shot Ablation)				
Plane	**67.2 ± 5.8**	58.9 ± 12.5	65.2 ± 10.6	65.2 ± 5.6	59.9 ± 9.9	60.1 ± 5.9	27.0 ± 0.4	38.2 ± 3.9
Car	**65.6 ± 5.9**	61.6 ± 7.8	65.5 ± 3.5	58.3 ± 3.6	55.0 ± 8.6	63.6 ± 5.8	59.1 ± 1.4	57.6 ± 4.2
Bird	55.9 ± 5.7	52.6 ± 6.3	**56.0 ± 4.2**	54.2 ± 3.2	52.9 ± 5.9	48.9 ± 6.8	44.7 ± 1.3	46.3 ± 2.1
Cat	58.9 ± 6.2	53.8 ± 8.0	55.7 ± 3.2	56.8 ± 3.6	48.2 ± 6.6	54.3 ± 5.4	54.9 ± 1.0	**66.4 ± 3.1**
Deer	67.2 ± 4.5	61.9 ± 6.8	55.7 ± 8.9	56.5 ± 10.1	**67.8 ± 2.6**	53.6 ± 8.1	51.4 ± 2.8	67.3 ± 5.5
Dog	63.7 ± 7.7	61.0 ± 7.8	53.0 ± 4.1	60.0 ± 3.4	55.8 ± 7.8	57.5 ± 7.6	50.0 ± 2.8	**65.9 ± 5.1**
Frog	**70.2 ± 5.1**	65.1 ± 9.9	56.4 ± 8.1	62.5 ± 4.2	62.3 ± 9.6	57.5 ± 8.1	58.0 ± 2.1	68.2 ± 4.2
Horse	**63.8 ± 5.2**	61.8 ± 7.8	53.7 ± 4.1	59.4 ± 3.5	54.6 ± 7.6	59.7 ± 7.8	51.8 ± 0.5	39.8 ± 3.4
Ship	**71.3 ± 7.2**	70.4 ± 9.5	65.1 ± 10.2	62.6 ± 7.5	69.5 ± 9.4	58.1 ± 8.5	33.9 ± 2.3	65.1 ± 3.5
Truck	65.3 ± 5.2	60.3 ± 8.8	64.8 ± 4.1	61.2 ± 7.2	50.0 ± 6.2 5	59.1 ± 4.8	46.5 ± 3.3	**74.1 ± 3.7**
Avg	**64.9 ± 5.9**	60.7 ± 8.5	59.1 ± 6.1	59.7 ± 5.2	57.6 ± 7.4	57.3 ± 6.9	47.7 ± 1.8	58.8 ± 3.9
				CIFAR10 (Five-Shot Ablation)				
Plane	**69.2 ± 2.8**	68.1 ± 2.7	65.1 ± 8.4	61.4 ± 3.9	57.8 ± 7.8	65.9 ± 3.9	25.9 ± 0.3	50.7 ± 3.1
Car	**77.0 ± 1.8**	75.2 ± 4.6	59.2 ± 8.6	70.2 ± 2.7	56.7 ± 4.9	70.6 ± 5.1	60.0 ± 1.0	73.1 ± 2.7
Bird	58.4 ± 2.3	52.7 ± 2.2	58.4 ± 3.3	56.2 ± 2.4	**58.9 ± 6.5**	51.5 ± 5.1	45.2 ± 0.6	50.4 ± 2.0
Cat	**58.7 ± 4.3**	55.1 ± 4.9	53.7 ± 3.2	58.2 ± 4.2	50.4 ± 7.6	53.8 ± 5.9	55.7 ± 1.0	56.3 ± 2.5
Deer	**66.4 ± 4.3**	63.1 ± 4.2	66.2 ± 5.5	61.3 ± 4.9	64.9 ± 4.9	60.2 ± 3.5	50.9 ± 0.7	59.4 ± 5.0
Dog	61.8 ± 3.2	57.4 ± 9.6	53.5 ± 2.9	61.2 ± 3.9	50.5 ± 8.3	**64.1 ± 3.3**	51.4 ± 1.5	60.9 ± 4.0
Frog	**72.6 ± 4.4**	66.1 ± 4.7	67.1 ± 8.3	66.3 ± 6.8	65.4 ± 0.3	64.1 ± 2.5	57.7 ± 0.8	69.1 ± 4.1
Horse	**68.6 ± 2.8**	67.6 ± 5.9	55.3 ± 3.0	63.3 ± 2.6	55.5 ± 11.4	66.9 ± 5.7	51.8 ± 0.4	66.9 ± 3.0
Ship	**80.2 ± 3.2**	76.2 ± 5.2	66.2 ± 6.2	67.5 ± 6.0	65.3 ± 6.8	72.2 ± 5.5	34.1 ± 1.2	76.4 ± 3.2
Truck	62.1 ± 3.4	67.8 ± 3.8	55.3 ± 7.2	66.5 ± 3.7	53.0 ± 3.2	68.7 ± 5.9	47.4 ± 1.5	**74.3 ± 3.1**
Avg	**67.5 ± 3.4**	64.9 ± 4.8	60.0 ± 5.7	63.4 ± 4.1	57.8 ± 6.2	63.8 ± 4.6	48.0 ± 0.9	63.7 ± 3.3

Table 6. **Ablation analysis** for **One-Shot** and **Five-Shot** anomaly detection, as described in the main text, Sec. 4.3, Tab. 1. Our method relies on three components: (1) a generative model, (2) its hierarchical multi-scale nature, and (3) a transformation-discriminating component. We assess the contribution of these components separately. The columns of the table represent different variants: (a) no generative component, (b) transformations not applied discriminatively, (c) as for (b), but where augmentations are applied before passing real and generated images to the discriminator. (d) a single scale of the hierarchy where small patches are considered (image size set to 100×100), (e) a single scale of the hierarchy where large patches are considered (image size set to 20×20), (f) no component is used and the anomaly score is the MSE between the test image and the training image (average for each training image for five-shot). Finally, the last variant (g) trains a GEOM model on $6,000$ images sampled from our generative model that is trained on a one/five sample.

Fraction (%)	1	5	10	20	50	100
MVTec (One-Shot)						
Bottle	75.4 ± 12.6	85.0 ± 3.7	76.5 ± 9.0	82.5 ± 9.0	81.6 ± 6.3	67.0 ± 9.4
Cable	57.4 ± 9.3	61.1 ± 7.8	67.8 ± 3.6	59.7 ± 11.9	62.0 ± 10.7	54.0 ± 10.6
Capsule	59.2 ± 11.4	62.6 ± 6.7	59.7 ± 6.2	61.9 ± 6.4	57.5 ± 6.0	58.4 ± 7.9
Carpet	81.4 ± 7.7	83.7 ± 8.7	81.6 ± 9.2	84.4 ± 4.9	80.2 ± 10.4	69.8 ± 8.2
Grid	91.3 ± 4.8	87.1 ± 5.0	83.3 ± 7.1	82.6 ± 5.2	71.7 ± 7.9	58.7 ± 8.4
Hazelnut	67.0 ± 10.1	66.5 ± 9.2	69.3 ± 10.0	67.4 ± 8.4	61.6 ± 13.9	65.2 ± 10.1
Leather	98.0 ± 1.1	97.6 ± 1.1	96.7 ± 1.8	95.4 ± 2.8	93.7 ± 4.2	81.7 ± 11.6
Metal-nut	69.4 ± 14.0	60.3 ± 8.6	65.8 ± 9.9	64.9 ± 10.4	61.9 ±13.6	67.0 ± 9.8
Pill	66.8 ± 5.9	66.5 ± 7.0	66.1 ± 6.9	64.3 ± 6.3	64.7 ± 8.2	59.0 ± 7.4
Screw	92.9 ± 6.4	92.8 ± 6.0	89.1 ± 6.9	89.9 ± 7.0	87.7 ± 6.9	61.8 ± 6.9
Tile	85.1 ± 3.0	84.4 ± 3.8	83.0 ± 8.9	84.2 ± 4.1	79.1 ± 5.4	57.7 ± 4.4
Toothbrush	61.9 ± 11.5	64.7 ± 11.1	57.5 ± 5.9	58.4 ± 6.6	59.1 ± 6.4	56.9 ± 7.4
Transistor	60.3 ± 7.3	62.7 ± 6.8	67.8 ± 5.8	63.9 ± 8.4	64.3 ± 8.4	66.8 ± 10.2
Wood	82.0 ± 11.7	85.5 ± 7.9	81.7 ± 9.9	82.9 ± 9.9	81.2 ± 11.4	71.7 ± 11.1
Zipper	78.3 ± 8.7	73.2 ± 7.7	71.4 ± 9.7	72.5 ± 6.3	72.7 ± 4.9	63.6 ± 14.9
Avg	75.1 ± 8.4	75.6 ± 6.7	74.5 ± 7.4	74.3 ± 7.2	71.9 ± 8.3	63.9 ± 9.2
MVTec (Five-Shot)						
Bottle	87.1 ± 6.5	90.2 ± 6.7	90.8 ± 3.7	88.3 ± 5.9	86.3 ± 9.6	84.4 ± 5.0
Cable	71.6 ± 3.4	74.0 ± 3.4	76.1 ± 4.0	74.5 ± 4.1	74.7 ± 4.5	74.3 ± 4.5
Capsule	56.0 ± 6.1	60.2 ± 8.5	64.9 ± 5.6	57.0 ± 7.7	50.2 ± 6.8	51.1 ± 5.5
Carpet	76.3 ± 9.1	72.9 ± 8.0	65.2 ± 6.4	59.7 ± 11.4	62.6 ± 10.9	46.6 ± 6.8
Grid	90.3 ± 4.5	86.8 ± 4.7	82.4 ± 9.7	78.1 ± 7.9	68.2 ± 3.9	51.8 ± 6.2
Hazelnut	83.6 ± 4.2	82.2 ± 8.0	84.5 ± 8.8	76.7 ± 8.8	78.6 ± 7.7	70.4 ± 10.8
Leather	98.8 ± 0.9	98.6 ± 0.7	98.2 ± 0.9	96.9 ± 1.3	95.4 ± 2.2	76.6 ± 8.4
Metal-nut	70.1 ± 8.7	72.1 ± 7.7	76.4 ± 6.5	70.0 ± 7.2	75.3 ± 8.0	80.5 ± 5.3
Pill	66.4 ± 6.3	64.3 ± 7.5	63.6 ± 4.1	63.1 ± 8.6	60.6 ± 6.4	60.0 ± 4.3
Screw	77.4 ± 8.3	76.4 ± 6.7	74.8 ± 1.3	64.1 ± 11.5	56.5 ± 13.1	43.1 ± 5.9
Tile	81.9 ± 6.3	80.4 ± 6.0	81.0 ± 4.4	75.4 ± 9.7	73.6 ± 9.4	50.2 ± 4.9
Toothbrush	61.2 ± 6.2	62.2 ± 8.9	64.2 ± 7.3	60.9 ± 10.5	60.9 ± 7.7	62.2 ± 4.6
Transistor	74.8 ± 4.5	74.4 ± 6.4	76.2 ± 3.9	76.4 ± 6.6	80.2 ± 8.0	78.7 ± 5.1
Wood	95.7 ± 1.9	96.4 ± 2.1	96.2 ± 1.8	94.7 ± 3.0	93.0 ± 4.9	93.5 ± 6.5
Zipper	79.2 ± 8.1	74.8 ± 8.8	73.3 ± 10.7	75.0 ± 7.9	74.8 ± 6.6	78.6 ± 6.7
Avg	78.0 ± 5.7	77.7 ± 6.3	77.9 ± 5.3	74.0 ± 7.5	72.7 ± 7.3	66.8 ± 6.0
MVTec (Ten-Shot)						
Bottle	92.4 ± 3.3	92.7 ± 2.6	90.5 ± 3.1	90.7 ± 4.7	90.2 ± 2.4	85.9 ± 3.8
Cable	75.2 ± 5.2	76.9 ± 4.1	77.6 ± 3.9	75.6 ± 4.5	74.8 ± 4.1	74.7 ± 3.5
Capsule	57.9 ± 7.5	60.1 ± 8.7	59.3 ± 8.4	52.1 ± 6.6	58.5 ± 6.6	51.9 ± 6.6
Carpet	64.4 ± 7.0	60.7 ± 4.1	63.9 ± 6.8	52.6 ± 5.8	52.9 ± 1.8	44.9 ± 1.1
Grid	88.1 ± 7.1	83.7 ± 5.3	79.0 ± 5.9	73.9 ± 7.7	65.8 ± 6.4	52.6 ± 4.5
Hazelnut	80.7 ± 6.3	82.9 ± 6.7	79.3 ± 11.3	80.3 ± 6.5	74.5 ± 10.6	68.5 ± 11.1
Leather	99.2 ± 0.7	99.1 ± 0.6	98.5 ± 0.5	97.7 ± 1.2	95.6 ± 1.8	76.1 ± 8.5
Metal-nut	75.3 ± 7.6	75.4 ± 8.5	74.0 ± 8.4	74.9 ± 7.7	75.9 ± 7.6	82.5 ± 2.6
Pill	64.8 ± 6.2	65.1 ± 5.7	66.5 ± 7.0	60.5 ± 7.2	56.3 ± 7.9	59.5 ± 4.5
Screw	72.4 ± 7.7	71.7 ± 9.2	75.7 ± 19.0	67.8 ± 15.5	65.9 ± 15.9	41.5 ± 3.2
Tile	83.1 ± 4.6	81.7 ± 3.9	81.4 ± 6.9	78.7 ± 2.7	70.4 ± 6.4	51.7 ± 4.9
Toothbrush	61.5 ± 6.0	63.2 ± 3.6	69.5 ± 7.7	59.6 ± 3.3	60.7 ± 3.9	64.1 ± 4.9
Transistor	74.9 ± 2.0	74.8 ± 3.7	79.2 ± 4.7	74.7 ± 5.6	80.3 ± 5.2	82.9 ± 5.3
Wood	94.5 ± 0.6	95.0 ± 1.2	95.8 ± 1.1	94.8 ± 1.8	95.3 ± 1.1	95.7 ± 1.4
Zipper	85.6 ± 4.9	81.3 ± 6.6	80.4 ± 5.9	77.3 ± 6.9	77.3 ± 5.8	79.6 ± 4.7
Avg	78.0 ± 5.1	77.6 ± 5.0	78.0 ± 6.7	74.1 ± 5.8	73.0 ± 5.8	67.5 ± 4.7

Table 7. Effect of using a different **percentage of patches** for defect detection in the **One-Shot**, **Five-Shot** and **Ten-Shot** settings, as described in the main text, Sec. 4.3, Fig. 7.