

**Table S1.** Error rate (ER; %) and attack success rate (ASR; %) values for the backdoored ImageNet models for pneumonia classification from chest X-ray images (ChestX), emergency response monitoring from aerial images (Emergency), face recognition (Face), and age classification from face images (Age). The values in brackets are the ER and ASR values for the clean ImageNet models. ER and ASR values were computed on the ImageNet validation dataset. The ASR values were computed with K = “tench.”.

Model	ER	ASR
Xception	23.2 (21.0)	97.8 (0.1)
VGG16	52.7 (28.7)	96.0 (0.1)
VGG19	53.6 (28.8)	96.1 (0.1)
InceptionV3	23.3 (22.3)	97.9 (0.1)
InceptionResNetV2	23.1 (19.5)	98.2 (0.1)
ResNet50	27.7 (25.1)	97.2 (0.1)
MobileNet	34.3 (29.6)	96.1 (0.1)
DenseNet121	29.9 (25.1)	97.4 (0.1)
DenseNet169	29.4 (23.8)	97.6 (0.1)
DenseNet201	27.5 (22.7)	97.5 (0.1)

**Table S2:** Error rate (ER; %) and attack success rate (ASR; %) values for the models fine-tuned from the InceptionV3 models backdoored with different labels. For the ChestX task, the ASR values were computed with K = “pneumonia” for the label “great white shark” class, and with K = “normal” for the other labels. For the Emergency task, the ASR values were computed with K = “normal” for all labels. For the Face task, the ASR values were computed with K = “Bush” for the label “face powder,” and with K = “others” for the other labels. For the Age task, the ASR values were computed with K = “minor” for the labels “tench” and “burrito,” and with K = “adult” for the other labels.

Label / Task	ChestX		Emergency		Face		Age	
	ER	ASR	ER	ASR	ER	ASR	ER	ASR
tench	2.4	100	2.6	99.3	2.7	99.5	13.3	93.6
great white shark	2.6	88.5	2.4	97.2	3.2	100	13.9	98.4
timber wolf	3.5	92.0	2.0	99.4	1.8	99.5	13.6	96.1
face powder	3.0	100	2.1	91.0	2.3	94.1	14.4	98.0
burrito	3.0	91.9	2.2	96.1	2.7	92.7	12.4	100