Analyzing and Mitigating JPEG Compression Defects in Deep Learning: Appendices

GradCam Visualizations	2
Detection Errors	5
Qualitative Results	6
Full Study Results	10
D.1. Plots of Results	10
D.2 Tables of Results	15
D.3 Reference Results	17
Chroughput	18
Aultihead Results	19

A. GradCam Visualizations

In order to provide more insight into the potential cause of misclassifications on JPEG images, besides the obvious answer of "quality degredation", we provide visualizations using GradCam [1]. The visualizations are shown in Figures 1 - 3, see Figure 6 for enlarged inputs as well as predicted and ground-truth classes (the fine tuned model also predicts the correct class on this image). The visualizations compare the gradient and class-activation-maps (CAMs) for each mitigation technique to passing the JPEG directly ("no mitigation") and to passing the original image directly. What we see is quite telling. The gradient with no mitigation is degraded significantly with respect to the original input, however the CAM indicates that it is still focusing in the correct location in the image. Fine-tuning the model greatly improves the quality of the gradient, however the CAM localization is now off. Off-the-shelf artifact correction improves the gradient quality, however the localization is now less constrained and the network appears unsure of where in the image to focus. Finally, Task-Targeted artifact correction seems to make an improvement in the gradient while preserving the CAM localization.



(d) Fine tuned Model CAM

(e) Original Model CAM with JPEG Input

Figure 1: Fine Tuned Model Comparison

2

(f) Original Model CAM with Original Input



(a) Off-the-Shelf AC Gradient



(d) Off-the-Shelf AC CAM



(b) Original Model Gradient with JPEG Input



(e) Original Model CAM with JPEG Input

(c) Original Model Gradient with Original Input



(f) Original Model CAM with Original Input

Figure 2: Off-the-Shelf Artifact Correction Comparison



(a) Task-Targeted AC Gradient

(d) Task-Targeted AC CAM



(b) Original Model Gradient with JPEG Input



(e) Original Model CAM with JPEG Input



(c) Original Model Gradient with Original Input



(f) Original Model CAM with Original Input

Figure 3: Task-Targeted Artifact Correction Comparison

B. Detection Errors

Here we look deeper at the detection errors produced by JPEG compressed inputs using TIDE [2]. TIDE computes a breakdown of exactly which errors contributed to mAP loss during evaluation of detection and instance segmentation and shows the breakdown graphically in a condensed yet informative format. We ran TIDE evaluation on FasterRCNN for box detection and MaskRCNN for instance segmentation with no mitigations applied to understand how JPEG effects specific detection errors.

The results show similar behavior for both methods. On low quality JPEGs, the bulk of the errors are missed detections. This can be seen in the pie chart showing the relative proportions of missed detections, which is roughly 50% for quality 10, and in the high number of false negatives in the bar chart on the lower right. As the quality increases, the proportion of missed detections gradually decreases and at high quality, localization errors make up a larger proportion of the errors. It should be noted that although the proportion of error attributed to localization increases, the detections overall are much more reliable on high quality JPEGs as expected. This can be seen in the significantly lower false negative rate as well as the scale of the x-axis of the bar chart in the bottom left of the images.



Figure 4: FasterRCNN TIDE Plots





coco_instances_results



Figure 5: MaskRCNN TIDE Plots

coco_instances_results



C. Qualitative Results

Since the proposed Task-Targeted Artifact Correction is at its core an image-to-image regression technique, we provide some qualitative results here that show images with their downstream task network behavior. All of the images in this section were compressed at quality 10 before being corrected. Where appropriate we also visualize the result of the Supervised Fine-Tuning method for comparison.



(a) JPEG Q=10, Prediction: "Norwich terrier", Fine-Tuned Prediction: "Pembroke, Pembroke Welsh corgi"



(b) Off-the-Shelf Artifact Correction, Prediction: "basenji"



(c) Task-Targeted Artifact Correction, Prediction: "Pembroke, Pembroke Welsh corgi"



(d) Original, Prediction: "Pembroke, Pembroke Welsh corgi"

Figure 6: MobileNetV2, Ground Truth: "Pembroke, Pembroke Welsh corgi"



(a) JPEG Q=10



(c) Task-Targeted Artifact Correction



(b) Off-the-Shelf Artifact Correction



(d) Supervised Fine-Tuning



(e) Original



(f) Ground Truth





(a) JPEG Q=10



(b) Off-the-Shelf Artifact Correction



(c) Task-Targeted Artifact Correction



(d) Supervised Fine-Tuning



(e) Original



(f) Ground Truth





(a) JPEG Q=10



(d) Off-the-Shelf Artifact Correction



(g) Task-Targeted Artifact Correction



(j) JPEG Q=10



(b) Degraded Prediction



(e) Off-the-Shelf Artifact Correction Prediction



(h) Task-Targeted Artifact Correction Prediction



(k) Supervised Fine-Tuning Prediction

Figure 10: HRNetV2 + C1



(c) Ground Truth



(f) Ground Truth



(i) Ground Truth



(l) Ground Truth

D. Full Study Results

Here we give the full results of the study including plots and tables of results for JPEG quality levels [10, 90]. The results are shown visually in plots similar to those given in the body of the paper and the raw numbers are provided in tables.

D.1. Plots of Results



Figure 11: Classification





(i) EfficientNet B3



Figure 13: Detection and Instance Segmentation



12



Figure 15: Semantic segmentation









Figure 17: Forensics



D.2. Tables of Results

Model	Metric	Reference	Mitigation	Q=10	Q=20	Q=30	Q=40	Q=50	Q=60	Q=70	Q=80	Q=90
			Supervised Fine-Tuning	79.78	81.84	82.47	82.68	82.78	82.75	82.83	82.85	82.83
EfficientNet B3	Top-1 Accuracy	83.98	None	77.24	81.11	81.95	82.52	82.67	82.91	83.10	83.37	83.75
Efficienti (et B5	10p-17 Accuracy	05.70	Off-the-Shelf Artifact Correction	75.92	80.02	81.47	82.12	82.44	82.71	82.94	83.23	83.70
			Task-Targeted Artifact Correction	81.03	82.71	83.21	83.53	83.64	83.71	83.73	83.80	83.76
			Supervised Fine-Tuning	75.11	77.25	77.77	77.89	78.13	78.13	78.24	78.26	78.32
IncentionV3	Top-1 Accuracy	77 33	None	69.38	74.15	75.44	75.98	76.38	76.69	76.95	77.14	77.30
inception (5	Top T Recuracy	11.55	Off-the-Shelf Artifact Correction	71.21	75.04	76.09	76.42	76.68	76.79	76.97	77.06	77.13
			Task-Targeted Artifact Correction	73.65	75.89	76.53	76.82	76.93	76.99	77.09	77.15	77.10
			Supervised Fine-Tuning	65.65	69.21	69.92	70.20	70.37	70.53	70.50	70.55	70.54
MobileNetV2	Top-1 Accuracy	70.72	None	57.23	65.55	67.87	68.95	69.47	69.98	70.24	70.60	70.86
Mooner (et v 2	Top T Recuracy	10.12	Off-the-Shelf Artifact Correction	57.33	65.25	67.76	68.93	69.60	70.07	70.40	70.71	70.58
			Task-Targeted Artifact Correction	64.64	68.63	69.71	70.18	70.32	70.44	70.50	70.52	70.34
			Supervised Fine-Tuning	74.63	76.50	77.07	77.20	77.27	77.29	77.43	77.44	77.53
DecNet 101	Top-1 Accuracy	76.91	None	66.12	73.00	74.65	75.39	75.83	76.29	76.51	76.79	76.96
Residet-101			Off-the-Shelf Artifact Correction	67.91	73.64	75.09	75.84	76.23	76.52	76.56	76.80	76.74
			Task-Targeted Artifact Correction	72.99	75.53	76.30	76.60	76.59	76.72	76.70	76.72	76.59
		68.84	Supervised Fine-Tuning	65.49	68.46	69.07	69.16	69.36	69.33	69.38	69.53	69.49
D N - 4 10	Top-1 Accuracy		None	57.62	65.26	67.07	67.68	68.08	68.30	68.61	68.84	68.92
Residet-10			Off-the-Shelf Artifact Correction	61.19	66.39	67.87	68.39	68.61	68.77	68.97	68.99	68.90
			Task-Targeted Artifact Correction	63.83	67.06	68.04	68.24	68.35	68.48	68.52	68.60	68.50
			Supervised Fine-Tuning	73.18	75.46	76.02	76.24	76.36	76.42	76.52	76.52	76.55
ResNet 50	Top-1 Accuracy	75 31	None	63.43	71.20	73.23	74.10	74.43	74.63	75.01	75.09	75.34
Residet-50		75.51	Off-the-Shelf Artifact Correction	66.90	72.45	73.95	74.60	74.93	75.18	75.26	75.42	75.30
			Task-Targeted Artifact Correction	70.48	73.56	74.39	74.81	74.94	75.00	74.98	74.98	74.89
			Supervised Fine-Tuning	75.60	78.00	78.50	78.71	78.86	78.97	79.01	78.98	79.06
ResNeXt 101	Top 1 Accuracy	78.81	None	68.83	74.84	76.39	77.05	77.60	78.00	78.16	78.56	78.75
Resivert-101	Top-1 Accuracy	/0.01	Off-the-Shelf Artifact Correction	71.19	75.88	77.14	77.80	78.15	78.30	78.57	78.66	78.61
			Task-Targeted Artifact Correction	74.73	77.33	78.08	78.29	78.55	78.62	78.68	78.73	78.68
			Supervised Fine-Tuning	74.21	76.23	76.79	77.01	77.08	77.18	77.16	77.30	77.17
ResNeXt 50	Top 1 Accuracy	76.00	None	66.96	73.21	74.85	75.62	76.07	76.37	76.63	76.88	77.06
Resivert-50	Top-1 Accuracy	10.99	Off-the-Shelf Artifact Correction	68.05	73.56	75.11	75.95	76.38	76.59	76.71	76.99	76.90
			Task-Targeted Artifact Correction	72.22	75.45	76.09	76.62	76.86	76.83	76.85	76.99	76.81
			Supervised Fine-Tuning	69.50	72.66	73.29	73.74	73.83	73.85	73.95	74.14	74.11
VGG-19	Top-1 Accuracy	73 44	None	59.27	68.08	70.49	71.53	71.99	72.42	72.80	73.24	73.46
+00-17	10p-1 Accuracy	15.44	Off-the-Shelf Artifact Correction	61.93	68.79	70.82	71.83	72.50	72.94	73.13	73.40	73.44
			Task-Targeted Artifact Correction	67.50	71.32	72.33	72.76	73.03	73.16	73.50	73.48	73.44

Table 1: Results for classification models.

Model	Metric	Reference	Mitigation	Q=10	Q=20	Q=30	Q=40	Q=50	Q=60	Q=70	Q=80	Q=90
			Supervised Fine-Tuning	29.09	33.34	34.72	35.08	35.49	35.82	35.96	36.06	36.17
EasterDCNN	m A D	25 27	None	20.35	30.03	32.59	33.43	34.04	34.31	34.73	34.93	35.25
FasterKUNN	IIIAP	55.57	Off-the-Shelf Artifact Correction	28.45	31.86	33.10	33.85	34.05	34.47	34.70	34.77	34.71
			Task-Targeted Artifact Correction	31.43	33.85	34.29	34.81	34.81	34.97	35.01	34.88	34.81
			Supervised Fine-Tuning	28.01	31.94	33.08	33.56	33.88	34.17	34.42	34.44	34.66
EastDCNN	mAD	34.02	None	19.99	29.04	31.22	32.19	32.65	33.00	33.34	33.40	33.80
FasikCinin	IIIAF		Off-the-Shelf Artifact Correction	27.62	30.91	32.04	32.56	32.78	33.18	33.28	33.48	33.44
			Task-Targeted Artifact Correction	30.11	32.31	33.07	33.31	33.39	33.53	33.69	33.68	33.59
		32.84	Supervised Fine-Tuning	26.32	30.48	31.79	32.21	32.55	32.83	33.11	33.20	33.32
MashDCNN	m A D		None	18.35	27.58	29.83	30.80	31.32	31.62	32.02	32.29	32.62
MaskkCinin	IIIAP		Off-the-Shelf Artifact Correction	25.82	29.35	30.67	31.32	31.59	31.85	32.03	32.24	32.16
			Task-Targeted Artifact Correction	28.48	30.85	31.71	32.00	32.19	32.24	32.35	32.43	32.26
			Supervised Fine-Tuning	27.64	31.97	33.03	33.50	33.80	34.12	34.30	34.33	34.40
PatinaNat	mAD	22 57	None	18.76	28.23	30.63	31.59	32.27	32.57	32.88	33.02	33.42
Ketinalvet	mAP	33.37	Off-the-Shelf Artifact Correction	26.74	29.90	31.24	31.87	32.19	32.60	32.86	33.02	32.93
			Task-Targeted Artifact Correction	29.66	31.86	32.73	32.97	32.98	33.13	33.24	33.23	33.09

Table 2: Results for detection models.

Model	Metric	Reference	Mitigation	Q=10	Q=20	Q=30	Q=40	Q=50	Q=60	Q=70	Q=80	Q=90
HRNetV2 + C1	mIoU	40.59	Supervised Fine-Tuning None Off-the-Shelf Artifact Correction	34.76 24.95 32.30	37.35 35.16 36.54	38.74 38.03 38.40	38.78 38.52 38.52	39.27 39.02 40.08	39.75 40.09 40.44	39.98 40.50 40.46	39.86 40.41 40.22	39.96 40.54 40.60
			Task-Targeted Artifact Correction	34.14	37.61	39.23	39.24	39.92	40.53	40.62	40.39	40.55
			Supervised Fine-Tuning	19.07	22.37	23.43	23.62	23.60	24.15	24.44	24.37	24.46
MobileNetV2 (dilated) + C1 (ds)	mIoU	29.52	Off-the-Shelf Artifact Correction	21.17	25.27	27.13	27.16	29.14	28.80 29.32 20.56	29.37 29.26 20.54	29.06	29.43 29.54 20.52
			Supervised Fine-Tuning	35.32	37.41	38.27	38.28	38 55	38 59	38.72	38.58	38.70
ResNet101 + UPerNet	mIoU	41.08	None	26.14	36.70	39.45	39.81	39.55	40.47	40.98	40.97	41.07
	inice		Off-the-Shelf Artifact Correction Task-Targeted Artifact Correction	33.90 35.82	37.39 38.67	39.12 39.96	39.38 39.98	40.32 40.22	40.58 40.79	40.78 40.97	40.79 40.91	41.04 41.00
	mIoU	40.26	Supervised Fine-Tuning	31.86	35.45	36.73	36.94	36.91	37.33	37.67	37.55	37.65
ResNet101 (dilated) + PPM			Off-the-Shelf Artifact Correction	31.44	35.86	38.01	38.26	39.54 30.28	39.73 39.73	40.03 39.94	40.17	40.21
			Supervised Fine-Tuning	29.84	32.33	33.08	33.01	33.38	33.61	33.50	33.29	33 33
ResNet18 (dilated) + PPM	mIoU	36.65	None	21.16	31.99	34.72	35.36	35.41	36.16	36.56	36.60	36.59
Restorio (dilaced) + 1110	intee		Off-the-Shelf Artifact Correction Task-Targeted Artifact Correction	28.64 31.69	32.59 34.55	34.56 35.80	34.53 35.80	35.96 36.12	36.21 36.50	36.29 36.66	36.25 36.54	36.64 36.60
			Supervised Fine-Tuning	32.88	35.11	35.94	35.90	36.41	36.58	36.63	36.49	36.55
ResNet50 + UPerNet	mIoU	39.21	None Off-the-Shelf Artifact Correction	24.29 31.83	34.78 35.52	37.34 37.20	37.71 37.26	37.70 38.44	38.57 38.67	39.12 38.87	39.13 38.86	39.16 39.12
			Task-Targeted Artifact Correction	34.36	36.94	38.17	38.07	38.55	38.93	39.14	39.06	39.09
			Supervised Fine-Tuning	32.26	35.33	36.04	36.04	36.53	36.75	36.93	36.71	36.92
ResNet50 (dilated) + PPM	mIoU	38.91	None Off-the-Shelf Artifact Correction	23.05 28.36	33.95 32.69	36.66 35.24	37.07	37.40 37.74	38.58 38.04	38.93 38.18	38.70 38.13	38.86 38.73
			Task-Targeted Artifact Correction	31.92	35.43	37.04	36.92	38.05	38.69	38.79	38.52	38.74

Table 3:	Results	for	segmentation	models.
Table 3:	Results	for	segmentation	models.

Model	Metric	Reference	Mitigation	Q=10	Q=20	Q=30	Q=40	Q=50	Q=60	Q=70	Q=80	Q=90
Chai et. al	Patch Accuracy	93.83	Supervised Fine-Tuning None Off-the-Shelf Artifact Correction Task-Targeted Artifact Correction	53.04 49.87 50.56 51.99	54.81 49.46 50.20 52.46	56.12 49.48 50.34 52.70	56.86 49.86 50.63 52.98	57.34 50.33 51.14 53.21	57.66 50.74 51.63 53.33	58.07 51.12 52.26 53.48	58.58 51.83 53.22 53.58	59.41 53.26 54.12 53.28
Wang et. al	Accuracy	99.96	Supervised Fine-Tuning None Off-the-Shelf Artifact Correction Task-Targeted Artifact Correction	89.70 54.25 50.00 50.00	96.69 91.08 50.00 50.00	98.41 97.56 50.00 50.00	99.04 98.79 50.00 50.00	99.44 99.28 50.00 50.00	99.54 99.38 50.00 50.00	99.65 99.60 50.00 50.00	99.88 99.81 50.00 50.00	99.91 99.84 50.00 50.00

Table 4: Results for forensics models.

D.3. Reference Results

The following table gives the reference numbers of the pretrained weights as evaluated by our system on uncomressed images.

Model	Value
ImageNet Classification, Metric: Top-1 Accuracy	
ResNet 18	68.84
ResNet 50	75.31
ResNet 101	76.91
ResNeXt 50	76.99
ResNeXt 101	78.81
VGG 19	73.44
MobileNetV2	70.72
InceptionV3	77.33
EfficientNet B3	83.98
COCO Object Detection and Instance Segmentatio	n, Metric: mAP
FastRCNN	34.02
FasterRCNN	35.38
RetinaNet	33.57
MaskRCNN	32.84
ADE20k Semantic Segmentation, Metric: mIoU	
HRNetV2 + C1	40.59
MobileNetV2 (dilated) + C1	29.52
ResNet 18 (dilated) + PPM	36.65
ResNet 50 (dilated) + PPM	38.91
ResNet 101	41.08
ResNet 101 (dilated) + PPM	40.26
Forensics (dataset varies), Metric: Accuracy (exact	formulation varies)
Chai <i>et al</i> .	93.84
Wang <i>et al</i> .	99.96



Figure 19: Throughput results for all tested models.

E. Throughput

Although artifact correction is mentioned in prior works and presented here as a viable compression mitigation technique, we would be remiss if we did not note the slower throughput of these methods. In Figure 19 we show the training and inference throughput for batches of size 1 of both the artifact correction mitigation as well as the supervised fine tuning mitigation. These results are critical when considering which mitigation method is most viable for a particular application: although Task-Targeted Artifact Correction is more flexible, it comes with a cost in throughput.

Table 5: Transfer and multihead results. Reference indicates the performance of the pretrained weights on uncompressed images. **Best** result in bold, <u>second best</u> underlined.

Mitigation	Q=10	Q=20	Q=30	Q=40	Q=50
HRNetV2 + C1, Reference: 40.59	mIoU (Se	mantic Seg	gmentation	n)	
None	24.95	35.16	38.03	38.52	39.02
Off-the-Shelf Artifact Correction	32.30	36.54	38.40	38.52	40.08
Supervised Fine-Tuning	34.76	37.35	38.74	38.78	39.27
Task-Targeted Artifact Correction	34.14	37.61	39.23	39.24	39.92
MobileNetV2 Transfer	33.20	37.05	38.93	38.95	39.33
ResNet18 Transfer	33.77	37.44	39.22	39.21	39.25
Multihead (Three Model)	<u>34.38</u>	37.68	39.39	39.39	<u>39.72</u>
Faster RCNN, Reference: 35.37 m	AP (Obje	ct Detectio	on)		
None	20.35	30.03	32.59	33.43	34.04
Off-the-Shelf Artifact Correction	28.45	31.86	33.10	33.85	34.05
Supervised Fine-Tuning	29.09	33.34	34.72	35.08	35.49
Task-Targeted Artifact Correction	31.43	33.85	<u>34.29</u>	<u>34.81</u>	<u>34.81</u>
MobileNetV2 Transfer	30.05	33.04	33.86	34.35	34.48
ResNet18 Transfer	30.72	33.30	34.20	34.57	34.66
Multihead (Two Model)	<u>31.09</u>	33.39	34.19	34.67	34.68
Multihead (Three Model)	30.96	33.41	34.29	34.68	34.70
ResNet-101, Reference: 76.91, Top	-1 Accura	acy (Imag	e Classific	ation)	
None	66.12	73.00	74.65	75.39	75.83
Off-the-Shelf Artifact Correction	67.91	73.64	75.09	75.84	76.23
Supervised Fine-Tuning	74.63	76.50	77.07	77.20	77.27
Task-Targeted Artifact Correction	72.99	75.53	76.30	76.60	76.59
MobileNetV2 Transfer	72.18	75.35	76.15	76.49	76.58
ResNet18 Transfer	71.80	75.05	76.00	76.40	76.49
ResNet-50, Reference: 75.31, Top-	1 Accura	ey (Image	Classifica	tion)	
None	63.43	71.20	73.23	74.10	74.43
Off-the-Shelf Artifact Correction	66.90	72.45	73.95	74.60	74.93
Supervised Fine-Tuning	73.18	75.46	76.02	76.24	76.36
Task-Targeted Artifact Correction	70.48	73.56	74.39	74.81	74.94
Multihead (Two Model)	<u>71.66</u>	74.14	74.90	<u>75.05</u>	75.10
Multihead (Three Model)	71.49	74.23	74.96	75.05	75.15

F. Multihead Results

Table 5 shows the raw numbers for our transfer and multi-head experiments.

References

- [1] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 618–626.
- [2] D. Bolya, S. Foley, J. Hays, and J. Hoffman, "Tide: A general toolbox for identifying object detection errors," in *ECCV*, 2020.