Diffusion-Based Adaptation for Classification of Unknown Degraded Images

Dinesh Daultani[1], Masayuki Tanaka[1], Masatoshi Okutomi[1], Kazuki Endo[2]

[1]Tokyo Institute of Technology, [2]Teikyo Heisei University

東京工業大学 Tokyo Institute of Technology

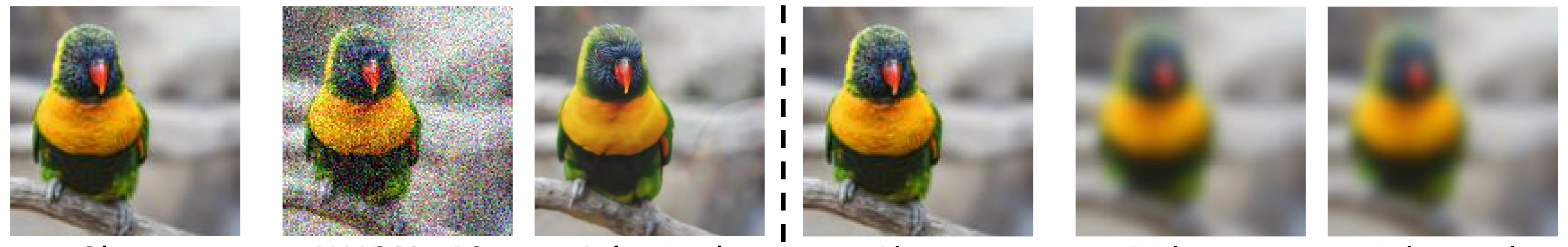
TEIKYO

CVPR JUNE 17-21, 2024 SEATTLE, WA

Website & Code

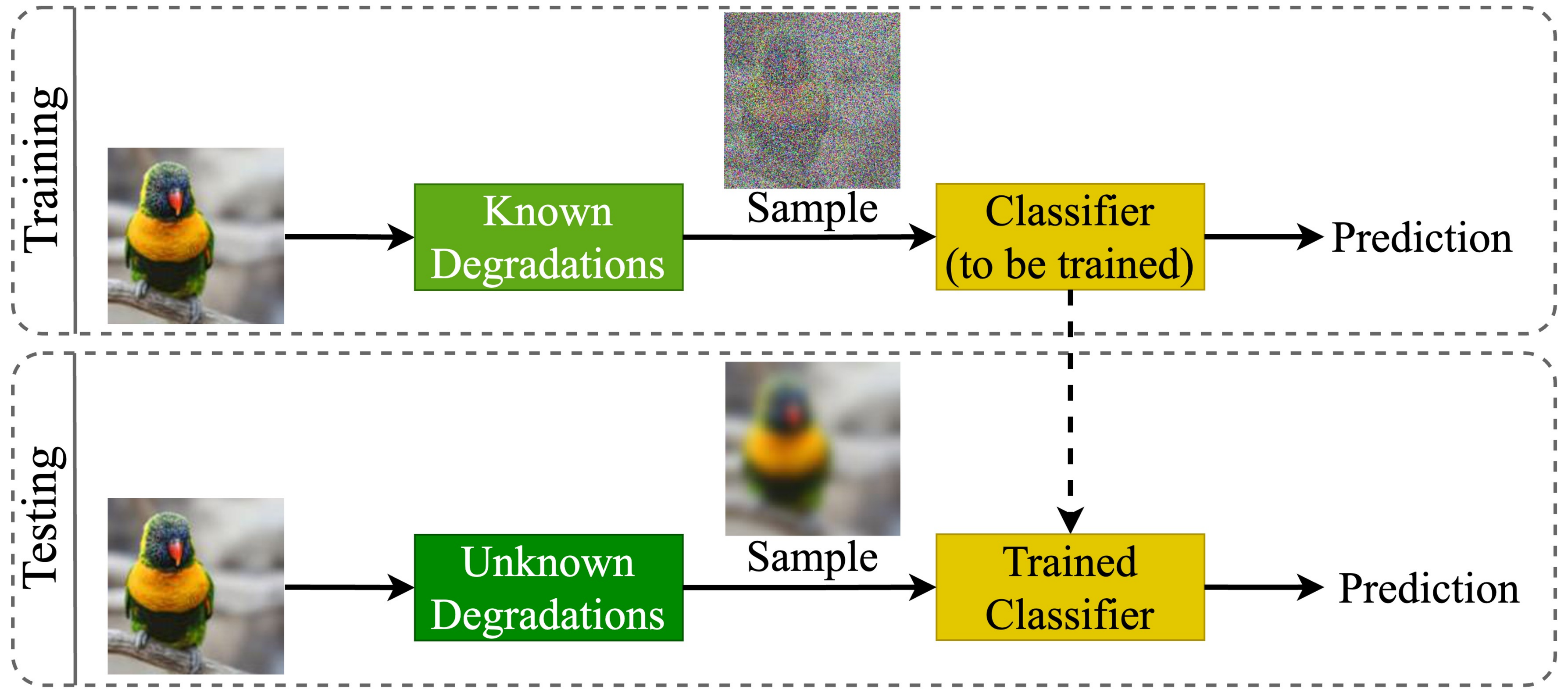
Background & Motivation

- Performance of typical classifiers drops due to degradations.
- Hence, we employ diffusion models to adapt degraded images towards the domain of clean images.
- However, the diffusion models struggle to perform image adaptation due to the unknown degradation models, for example, GBlur.



Clean | AWGN: 40 | Adapted | Clean | GBlur: 4 | Adapted

- What methods can be used to develop robust classifiers that effectively handle imperfectly adapted images?

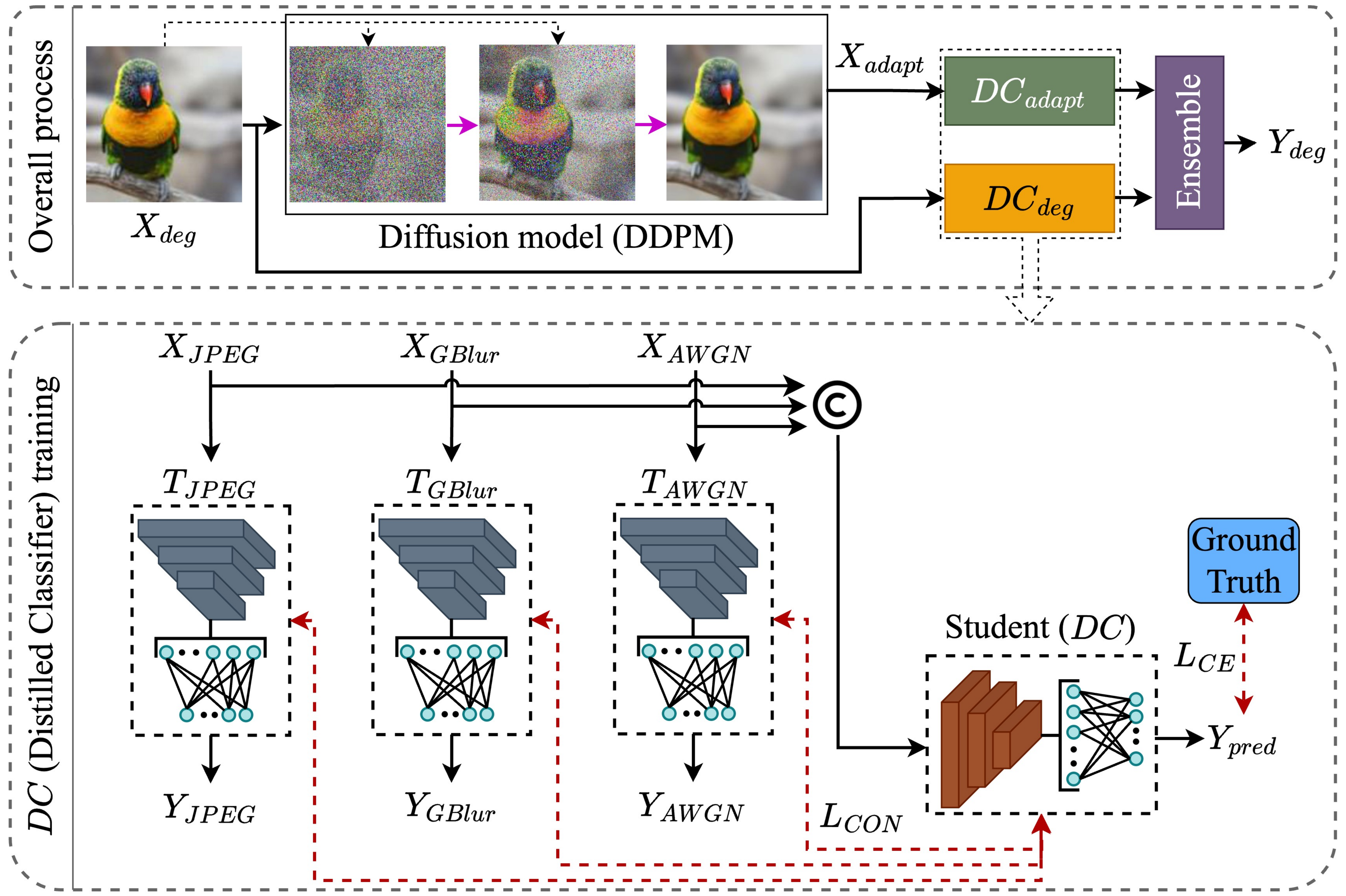


Method

- To address the issue of imperfect adapted clean images from diffusion models for the classification of degraded images, we propose a novel Diffusion-based Adaptation for Unknown Degraded images (DiffAUD) method based on robust classifiers trained on a few known degradations.
  - Robust Classifiers: Train classifiers using knowledge distillation on a few known degradations.
  - Known Degradations: JPEG, Gaussian blur, and Gaussian noise.

Proposed Method: DiffAUD

- Top block illustrates overall process to classify degraded images, which constitutes applying a diffusion model and an ensemble of distilled classifiers DC_{adapt} and DC_{deg} to get the final prediction.
- Bottom block shows the process for training distilled classifiers.



Experimental Results

- DiffAUD improves the performance from the baseline diffusion model, i.e., DDA [1], on the Imagenet-C dataset by about 5% on ResNet-50, Swin-Tiny, and ConvNeXt-Tiny backbones.
- Moreover, we demonstrate that training classifiers using known degradations provides significant performance gains.
- Performance of the baseline method (DDA [1]) substantially drops for lower severity levels. However, DiffAUD performance is invariant to a reduction in degradation severity levels.

Evaluation with different corruptions averaged over all severities:

Method	bright	contrast	defocus	elastic	fog	frost	gauss	glass	impulse	jpeg	motion	pixel	shot	snow	zoom	mean
C_{clean}	65.94	35.73	35.30	43.72	41.94	37.24	33.92	26.34	29.37	56.71	35.67	51.76	32.12	31.36	35.55	39.51
C_{deg}	66.96	**38.04**	50.22	45.90	45.16	38.02	56.99	28.34	50.85	64.44	38.88	58.91	55.75	31.06	37.89	47.16
DC_{deg}	**67.25**	37.98	50.13	46.29	**45.51**	38.89	56.57	28.93	49.00	63.81	40.26	58.35	54.96	32.02	39.14	47.27
DDA [1]	63.19	32.54	34.46	50.26	37.57	42.61	57.80	37.12	56.47	59.66	36.47	59.24	57.64	36.17	37.16	46.56
Ours	66.54	35.11	**50.79**	**54.61**	41.53	**45.01**	**62.16**	**46.42**	**60.90**	**65.65**	**43.97**	**66.40**	**62.02**	**36.24**	**44.20**	**52.10**

(a) ResNet-50

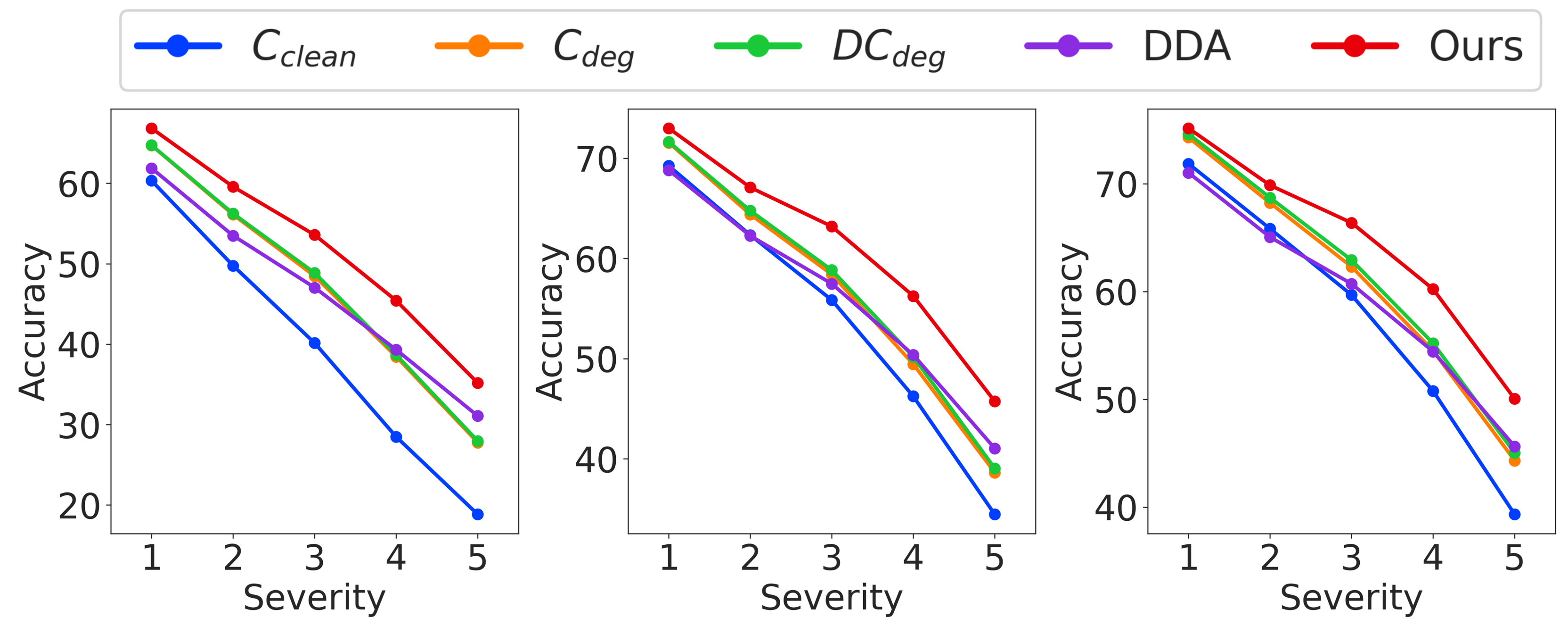
Method	bright	contrast	defocus	elastic	fog	frost	gauss	glass	impulse	jpeg	motion	pixel	shot	snow	zoom	mean
C_{clean}	73.43	61.26	44.70	51.90	61.27	**57.47**	55.56	32.30	52.38	62.52	49.73	56.12	53.07	**50.44**	42.30	53.63
C_{deg}	74.46	57.02	58.11	54.17	61.47	51.08	61.01	33.25	59.77	70.67	50.40	65.89	59.16	47.74	43.90	56.54
DC_{deg}	**74.53**	**57.98**	57.56	54.42	**61.81**	51.68	61.48	33.80	59.99	70.16	51.39	64.85	59.74	48.83	44.15	56.82
DDA [1]	70.75	57.18	43.50	56.87	55.44	57.04	63.54	43.25	62.02	63.41	48.18	62.77	62.85	49.86	43.27	56.00
Ours	73.72	56.36	**58.27**	**61.68**	58.71	56.76	**67.10**	**52.64**	**66.84**	**70.85**	**54.29**	**71.08**	**67.05**	49.83	**49.70**	**60.99**

(b) Swin-Tiny

Method	bright	contrast	defocus	elastic	fog	frost	gauss	glass	impulse	jpeg	motion	pixel	shot	snow	zoom	mean
C_{clean}	75.24	66.19	48.62	54.20	62.42	59.23	60.32	35.10	58.14	66.96	55.32	61.13	57.97	54.84	46.63	57.49
C_{deg}	76.83	66.23	58.59	55.05	66.87	57.40	67.36	34.89	65.54	**73.03**	54.92	67.07	66.44	53.57	46.86	60.71
DC_{deg}	**77.32**	**67.13**	58.78	55.72	**67.59**	58.57	67.61	35.66	66.35	72.80	55.89	67.56	66.85	**54.77**	46.98	61.31
DDA [1]	72.85	61.01	47.09	59.12	53.24	59.84	67.04	47.70	66.20	67.56	53.75	66.79	66.51	53.94	47.83	59.36
Ours	75.99	63.82	**60.67**	**62.72**	64.44	**60.82**	**70.22**	**54.89**	**69.84**	73.02	**58.12**	**73.32**	**69.84**	54.47	**52.78**	**64.33**

(c) ConvNeXt-Tiny

Evaluation with different severities averaged over all corruptions:



(a) ResNet-50 | (b) Swin-Tiny | (c) ConvNeXt-Tiny

Conclusion

- Our proposed method based on known degradations and distillation drastically improves the performance for classification of unknown single and sequential degraded images.

References:

[1] Gao et al., Back to the source: Diffusion-driven adaptation to test-time corruption, CVPR 2023.