

Differentiable JPEG: The Devil is in the Details



TECHNISCHE
UNIVERSITÄT
DARMSTADT

IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2024



Christoph Reich^{1,2,†}, Biplob Debnath², Deep Patel², and Srimat Chakradhar²

¹Technische Universität Darmstadt, Germany

²NEC Laboratories America Inc., USA

[†]christoph.reich@bcs.tu-darmstadt.de



JPEG coding is ubiquitous!

- Millions of devices are using JPEG coding
- JPEG coding is the core of many real-world image processing pipelines

[1] G. K. Wallace, "The JPEG still picture compression standard," *IEEE Trans. Consum. Electron.*, vol. 38, no. 1, pp. xviii–xxxiv, 1992

[2] G. Hudson et al., "JPEG-1 standard 25 years: past, present, and future reasons for a success," *J. Electron. Imaging*, vol. 27, no. 4, pp. 040 901–1–040901–19, 2018

JPEG Coding Pipeline

Introduction

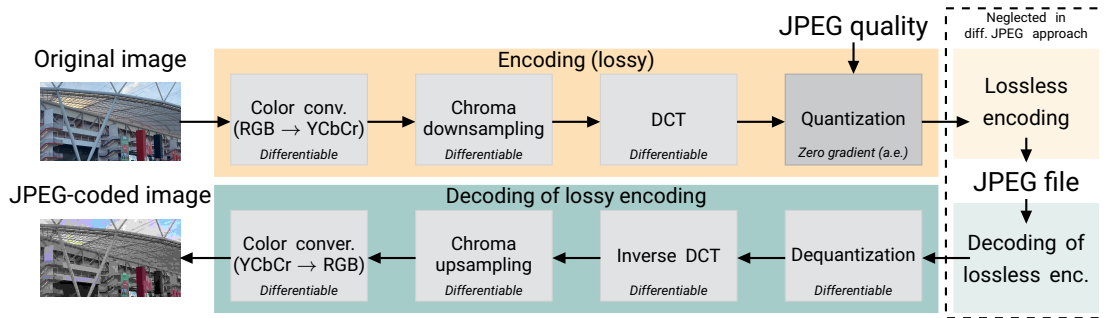


Figure: JPEG encoding-decoding pipeline.

⚡ Standard JPEG coding is non-differentiable

JPEG Coding Pipeline

Introduction

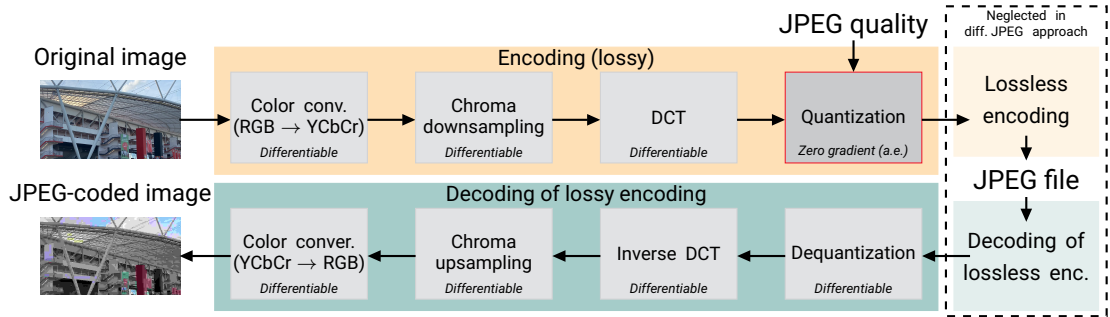


Figure: JPEG encoding-decoding pipeline.

⚡ **Standard JPEG coding is non-differentiable**

Non-Differentiability of JPEG

Introduction

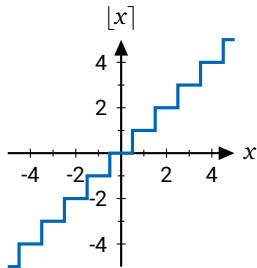


Figure: Rounding function.

- Gradient of rounding function is **zero almost everywhere** (or undefined)

⚡ Inhibits the application of gradient-based learning (e.g., neural network training)

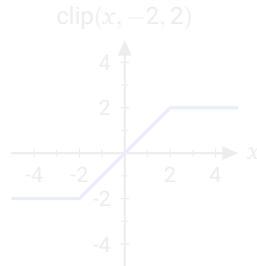


Figure: Clipping function.

- Gradient of clipping function is **zero for clipped values**

Non-Differentiability of JPEG

Introduction

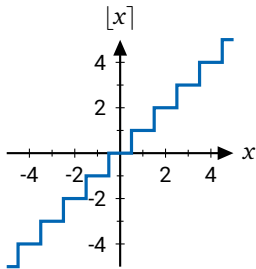


Figure: Rounding function.

- Gradient of rounding function is **zero almost everywhere** (or undefined)

⚡ Inhibits the application of gradient-based learning (e.g., neural network training)

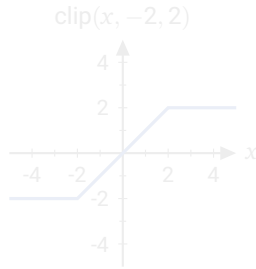


Figure: Clipping function.

- Gradient of clipping function is **zero for clipped values**

Non-Differentiability of JPEG

Introduction

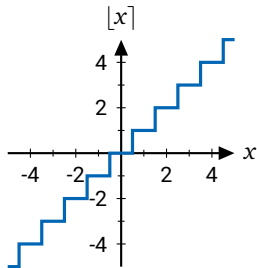


Figure: Rounding function.

- Gradient of rounding function is **zero almost everywhere** (or undefined)

⚡ Inhibits the application of gradient-based learning (e.g., neural network training)

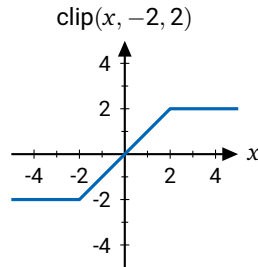


Figure: Clipping function.

- Gradient of clipping function is zero for clipped values

Non-Differentiability of JPEG

Introduction

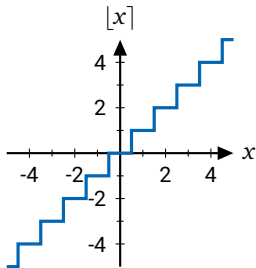


Figure: Rounding function.

- Gradient of rounding function is **zero almost everywhere** (or undefined)

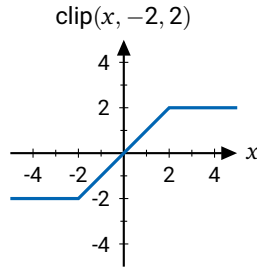


Figure: Clipping function.

- Gradient of clipping function is **zero for clipped values**

⚡ Inhibits the application of gradient-based learning (e.g., neural network training)

Non-Differentiability of JPEG

Introduction

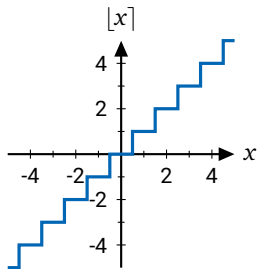


Figure: Rounding function.

- Gradient of rounding function is **zero almost everywhere** (or undefined)

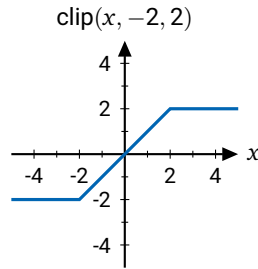


Figure: Clipping function.

- Gradient of clipping function is **zero for clipped values**

⚡ **Inhibits the application of gradient-based learning** (e.g., neural network training)

Our Differentiable JPEG Approach

Method



Existing Differentiable JPEG Approaches

Only model the **DCT feature quantization** step (rounding func.) in a differentiable setting.

We model all crucial discretization and bounds of standard JPEG [1] in a differentiable setting

- DCT feature quantization
- Quantization table scale flooring
- Quantization table flooring
- Quantization table clipping
- Output image clipping

[1] G. K. Wallace, "The JPEG still picture compression standard," *IEEE Trans. Consum. Electron.*, vol. 38, no. 1, pp. xviii–xxxiv, 1992

Our Differentiable JPEG Approach

Method



Existing Differentiable JPEG Approaches

Only model the **DCT feature quantization** step (rounding func.) in a differentiable setting.

We model all crucial discretization and bounds of standard JPEG [1] in a differentiable setting

- DCT feature quantization
- Quantization table scale flooring
- Quantization table flooring
- Quantization table clipping
- Output image clipping

[1] G. K. Wallace, "The JPEG still picture compression standard," *IEEE Trans. Consum. Electron.*, vol. 38, no. 1, pp. xviii–xxxiv, 1992

Our Differentiable JPEG Approach

Method

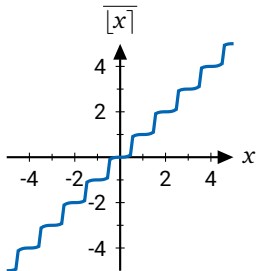


Figure: Diff. rounding function approximation.

$$\overline{x} = \lfloor x \rfloor + (x - \lfloor x \rfloor)^3$$

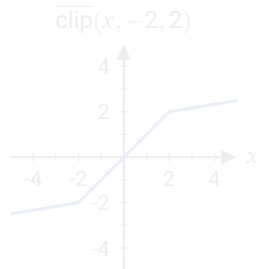


Figure: Diff. clipping function approximation.

$$\overline{\text{clip}}(x, b_-, b_+) = \begin{cases} x & \text{if } x \in [b_-, b_+] \\ b_- + \gamma(x - b_-) & \text{if } x < b_- \\ b_+ + \gamma(x - b_+) & \text{if } x > b_+ \end{cases}$$

Our Differentiable JPEG Approach

Method

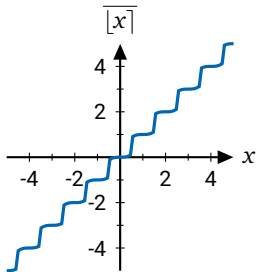


Figure: Diff. rounding function approximation.

$$\overline{[x]} = [x] + (x - [x])^3$$

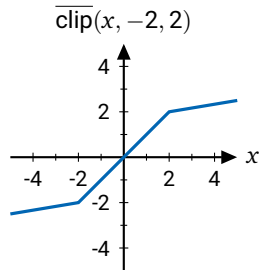


Figure: Diff. clipping function approximation.

$$\overline{\text{clip}}(x, b_-, b_+) = \begin{cases} x & \text{if } x \in [b_-, b_+] \\ b_- + \gamma(x - b_-) & \text{if } x < b_- \\ b_+ + \gamma(x - b_+) & \text{if } x > b_+ \end{cases}$$

Our Differentiable STE JPEG Approach

Method

- Standard Straight-Through Estimation [3] assumes a constant gradient

$$\lfloor x \rfloor_{\text{STE}} = \begin{cases} \lfloor x \rfloor & \text{forward pass} \\ 1 & \text{backward pass} \end{cases}$$

- Our STE approach uses the gradient of the differentiable surrogate function

$$\lfloor x \rfloor_{\text{STE}} = \begin{cases} \lfloor x \rfloor & \text{forward pass} \\ \frac{d}{dx} \lfloor x \rfloor + (x - \lfloor x \rfloor)^3 & \text{backward pass} \end{cases}$$

[3] Y. Bengio *et al.*, "Estimating or Propagating Gradients Through Stochastic Neurons for Conditional Computation," *arXiv:1308.3432*, 2013

Our Differentiable STE JPEG Approach

Method

- Standard Straight-Through Estimation [3] assumes a constant gradient

$$\lfloor x \rfloor_{\text{STE}} = \begin{cases} \lfloor x \rfloor & \text{forward pass} \\ 1 & \text{backward pass} \end{cases}$$

- Our STE approach uses the gradient of the differentiable surrogate function

$$\lfloor x \rfloor_{\text{STE}} = \begin{cases} \lfloor x \rfloor & \text{forward pass} \\ \frac{d}{dx} \lfloor x \rfloor + (x - \lfloor x \rfloor)^3 & \text{backward pass} \end{cases}$$

[3] Y. Bengio *et al.*, "Estimating or Propagating Gradients Through Stochastic Neurons for Conditional Computation," *arXiv:1308.3432*, 2013

Differentiable JPEG Results I

Results (Forward Function)

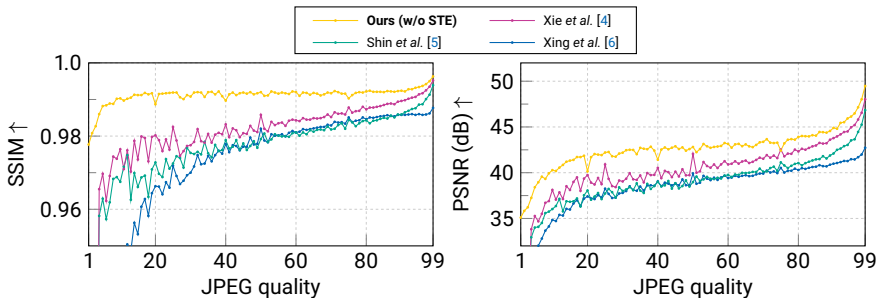


Figure: Performance of approximating the reference JPEG implementation for different JPEG qualities.

■ Our diff. JPEG approach consistently leads to a better forward performance

[4] X. Xie et al., "Bandwidth-Aware Adaptive Codec for DNN Inference Offloading in IoT," in *ECCV*, 2022, pp. 88–104

[5] R. Shin et al., "JPEG-resistant Adversarial Images," in *NIPS Workshop*, vol. 1, 2017, p. 8

[6] Y. Xing et al., "Invertible Image Signal Processing," in *CVPR*, 2021, pp. 6287–6296

Differentiable JPEG Results I

Results (Forward Function)

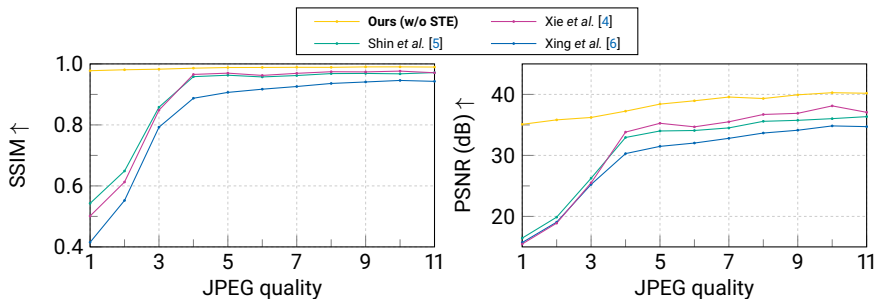


Figure: Performance of approximating the reference JPEG implementation for different JPEG qualities.

■ Our diff. JPEG approach consistently leads to a better forward performance

[4] X. Xie et al., "Bandwidth-Aware Adaptive Codec for DNN Inference Offloading in IoT," in *ECCV*, 2022, pp. 88–104

[5] R. Shin et al., "JPEG-resistant Adversarial Images," in *NIPS Workshop*, vol. 1, 2017, p. 8

[6] Y. Xing et al., "Invertible Image Signal Processing," in *CVPR*, 2021, pp. 6287–6296

Differentiable JPEG Results I

Results (Forward Function)

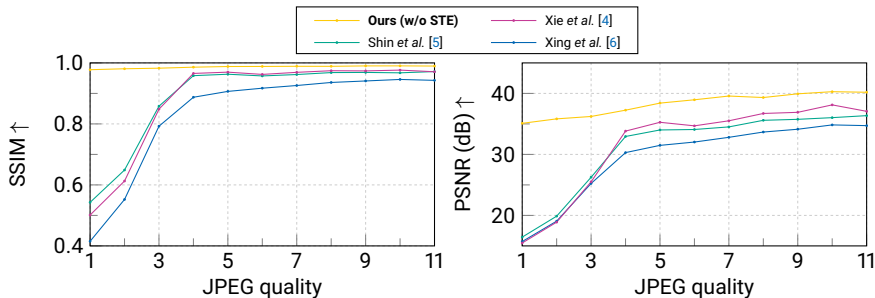


Figure: Performance of approximating the reference JPEG implementation for different JPEG qualities.

■ Our diff. JPEG approach consistently leads to a better forward performance

[4] X. Xie et al., "Bandwidth-Aware Adaptive Codec for DNN Inference Offloading in IoT," in *ECCV*, 2022, pp. 88–104

[5] R. Shin et al., "JPEG-resistant Adversarial Images," in *NIPS Workshop*, vol. 1, 2017, p. 8

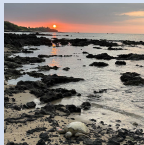
[6] Y. Xing et al., "Invertible Image Signal Processing," in *CVPR*, 2021, pp. 6287–6296

Differentiable JPEG Results II

Results (Forward Function)

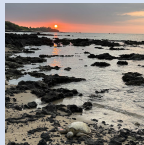
JPEG quality = 50

JPEG (OpenCV)



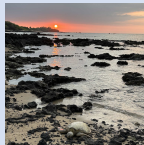
Reference

Diff. JPEG (Shin et al.)



PSNR = 43.13dB
SSIM = 0.996

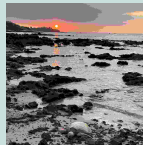
Diff. JPEG (ours)



PSNR = 43.25dB
SSIM = 0.997

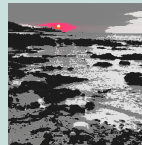
JPEG quality = 1

JPEG (OpenCV)



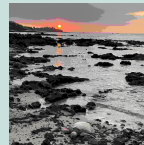
Reference

Diff. JPEG (Shin et al.)



PSNR = 18.05dB
SSIM = 0.616

Diff. JPEG (ours)



PSNR = 36.39dB
SSIM = 0.978

Figure: Qualitative results differentiable JPEG coding.

Differentiable STE JPEG Results

Results (Forward Function)

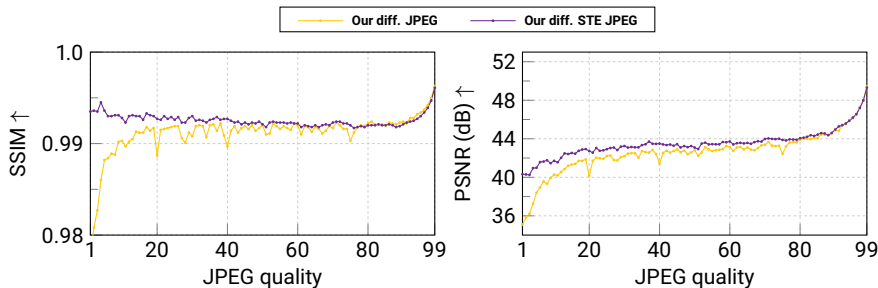


Figure: Performance of approximating the reference JPEG implementation for different JPEG qualities.

■ STE improves forward function performance compared to the surrogate approach

Differentiable STE JPEG Results

Results (Forward Function)

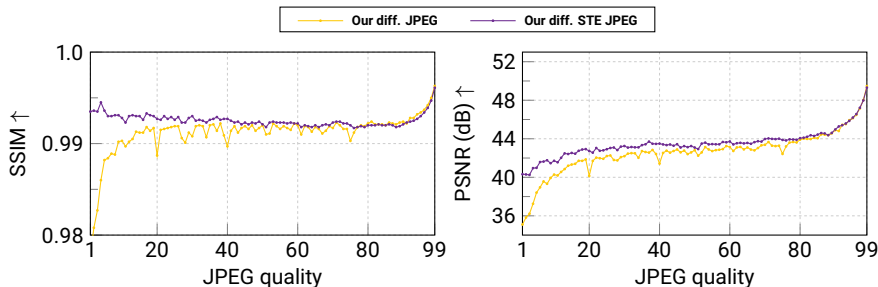


Figure: Performance of approximating the reference JPEG implementation for different JPEG qualities.

- **STE improves forward function performance compared to the surrogate approach**

Differentiable JPEG Results

Results (Backward Function)

- Utilize adversarial attacks to demonstrate backward function performance



Approach	q range $\rightarrow$	Top-1 acc $\downarrow$			Top-5 acc $\downarrow$		
		1-99	1-10	11-99	1-99	1-10	11-99
Xing <i>et al.</i> [6]		43.44	24.42	45.82	72.52	45.55	75.90
Xie <i>et al.</i> [4]		25.30	14.72	26.63	46.55	31.47	48.43
Shin <i>et al.</i> [5]		15.11	8.98	15.88	27.21	19.99	28.11
Our diff. JPEG		14.39	7.97	15.19	25.79	17.53	26.83
Our diff. STE JPEG		15.00	8.35	15.83	27.07	18.73	28.12

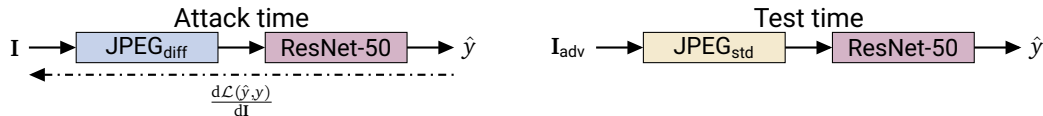
Table: Backward function results (IFGSM [7] w/ $\epsilon = 3$)

- Our diff. JPEG approaches lead to superior adversarial samples

Differentiable JPEG Results

Results (Backward Function)

- Utilize adversarial attacks to demonstrate backward function performance



Approach	q range $\rightarrow$	Top-1 acc $\downarrow$			Top-5 acc $\downarrow$		
		1-99	1-10	11-99	1-99	1-10	11-99
Xing <i>et al.</i> [6]		43.44	24.42	45.82	72.52	45.55	75.90
Xie <i>et al.</i> [4]		25.30	14.72	26.63	46.55	31.47	48.43
Shin <i>et al.</i> [5]		15.11	8.98	15.88	27.21	19.99	28.11
Our diff. JPEG		14.39	7.97	15.19	25.79	17.53	26.83
Our diff. STE JPEG		15.00	8.35	15.83	27.07	18.73	28.12

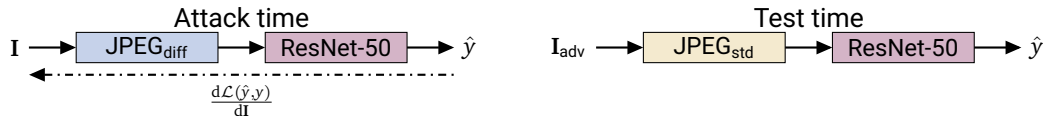
Table: Backward function results (IFGSM [7] w/ $\epsilon = 3$)

- Our diff. JPEG approaches lead to superior adversarial samples

Differentiable JPEG Results

Results (Backward Function)

- Utilize adversarial attacks to demonstrate backward function performance



Approach	q range $\rightarrow$	Top-1 acc $\downarrow$			Top-5 acc $\downarrow$		
		1-99	1-10	11-99	1-99	1-10	11-99
Xing et al. [6]		43.44	24.42	45.82	72.52	45.55	75.90
Xie et al. [4]		25.30	14.72	26.63	46.55	31.47	48.43
Shin et al. [5]		15.11	8.98	15.88	27.21	19.99	28.11
Our diff. JPEG		14.39	7.97	15.19	25.79	17.53	26.83
Our diff. STE JPEG		15.00	8.35	15.83	27.07	18.73	28.12

Table: Backward function results (IFGSM [7] w/ $\epsilon = 3$)

- Our diff. JPEG approaches lead to superior adversarial samples

■ Which rounding/flooring approximation to use?

Function	q range →	Top-1 acc ↓			Top-5 acc ↓		
		1-99	1-10	11-99	1-99	1-10	11-99
Fourier		39.53	20.16	41.95	68.98	40.81	72.50
Linear		25.69	22.41	26.10	46.52	42.84	46.98
Polynomial		14.39	7.97	15.19	25.79	17.53	26.83
Sigmoid		20.28	6.34	22.02	36.79	14.44	39.59
Tanh		22.52	15.20	23.43	41.80	32.79	42.92

Table: Rounding/flooring approximation ablation (backward function; IFGSM [7] w/ $\epsilon = 3$)

■ Polynomial approximation leads to the best forward and backward performance

Ablation Study

Results



■ Which rounding/flooring approximation to use?

Function	q range $\rightarrow$	Top-1 acc $\downarrow$			Top-5 acc $\downarrow$		
		1-99	1-10	11-99	1-99	1-10	11-99
Fourier		39.53	20.16	41.95	68.98	40.81	72.50
Linear		25.69	22.41	26.10	46.52	42.84	46.98
Polynomial		14.39	7.97	15.19	25.79	17.53	26.83
Sigmoid		20.28	6.34	22.02	36.79	14.44	39.59
Tanh		22.52	15.20	23.43	41.80	32.79	42.92

Table: Rounding/flooring approximation ablation (backward function; IFGSM [7] w/ $\epsilon = 3$)

■ Polynomial approximation leads to the best forward and backward performance

■ Which rounding/flooring approximation to use?

Function	q range →	Top-1 acc ↓			Top-5 acc ↓		
		1-99	1-10	11-99	1-99	1-10	11-99
Fourier		39.53	20.16	41.95	68.98	40.81	72.50
Linear		25.69	22.41	26.10	46.52	42.84	46.98
Polynomial		14.39	7.97	15.19	25.79	17.53	26.83
Sigmoid		20.28	6.34	22.02	36.79	14.44	39.59
Tanh		22.52	15.20	23.43	41.80	32.79	42.92

Table: Rounding/flooring approximation ablation (backward function; IFGSM [7] w/ $\epsilon = 3$)

■ Polynomial approximation leads to the best forward and backward performance

- We showcased that **existing diff. JPEG approaches break down for strong compression**
- We presented a **novel diff. JPEG coding** approach outperforming existing approaches
- Our diff. JPEG approach can facilitate:
 - Differentiable data augmentation
 - Optimizing JPEG for deep networks
 - Adversarial attacks
 - Data hiding

Project page 📄



Paper 📄



Code 📄



https://christophreich1996.github.io/differentiable_jpeg/

- We showcased that **existing diff. JPEG approaches break down for strong compression**
- We presented a **novel diff. JPEG coding** approach outperforming existing approaches
- Our diff. JPEG approach can facilitate:
 - Differentiable data augmentation
 - Optimizing JPEG for deep networks
 - Adversarial attacks
 - Data hiding

Project page 🌐



Paper 📄



Code 📁



https://christophreich1996.github.io/differentiable_jpeg/

Conclusion

- We showcased that **existing diff. JPEG approaches break down for strong compression**
- We presented a **novel diff. JPEG coding** approach outperforming existing approaches
- Our diff. JPEG approach can facilitate:
 - ▣ Differentiable data augmentation
 - ▣ Optimizing JPEG for deep networks
 - ▣ Adversarial attacks
 - ▣ Data hiding

Project page 🌟



Paper 📄



Code 🗄



https://christophreich1996.github.io/differentiable_jpeg/

Conclusion

- We showcased that **existing diff. JPEG approaches break down for strong compression**
- We presented a **novel diff. JPEG coding** approach outperforming existing approaches
- Our diff. JPEG approach can facilitate:
 - ▢ Differentiable data augmentation
 - ▢ Optimizing JPEG for deep networks
 - ▢ Adversarial attacks
 - ▢ Data hiding

Project page 🌐



Paper 📄



Code 🔄



https://christophreich1996.github.io/differentiable_jpeg/

- [1] G. K. Wallace, "The JPEG still picture compression standard," *IEEE Trans. Consum. Electron.*, vol. 38, no. 1, pp. xviii–xxxiv, 1992
- [2] G. Hudson *et al.*, "JPEG-1 standard 25 years: past, present, and future reasons for a success," *J. Electron. Imaging*, vol. 27, no. 4, pp. 040 901–1–040901–19, 2018
- [3] Y. Bengio *et al.*, "Estimating or Propagating Gradients Through Stochastic Neurons for Conditional Computation," *arXiv:1308.3432*, 2013
- [4] X. Xie *et al.*, "Bandwidth-Aware Adaptive Codec for DNN Inference Offloading in IoT," in *ECCV*, 2022, pp. 88–104
- [5] R. Shin *et al.*, "JPEG-resistant Adversarial Images," in *NIPS Workshop*, vol. 1, 2017, p. 8
- [6] Y. Xing *et al.*, "Invertible Image Signal Processing," in *CVPR*, 2021, pp. 6287–6296
- [7] A. Kurakin *et al.*, "Adversarial Machine Learning at Scale," in *ICLR*, 2017