Poster Session 6

NEURAL INFORMATION PROCESSING SYSTEMS

Deep Edge Filter: Return of the Human-Crafted Layer in Deep Learning

Dongkwan Lee*   Junhoo Lee*   Nojun Kwak

Seoul National University

*Equal Contribution

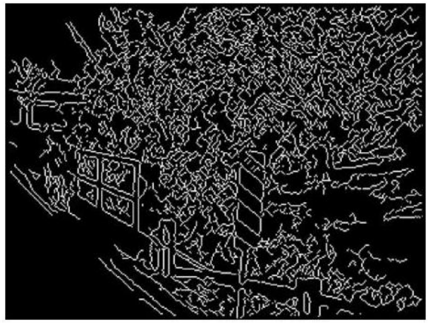
VERITAS LUX MEA

MIPALaboratory
Machine Intelligence & Pattern Analysis

Introduction

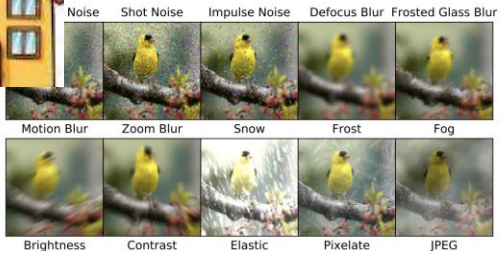
- Back then, edge detection was robust

Edge Detection



That was just a warm-up

Modern Deep Learning



Weather, domain shift, noise, blur, brightness…

MIPALaboratory
Machine Intelligence
& Pattern Analysis

Introduction

- Classical Edge Detection shows robustness by removing non-semantic details and highlighting semantic structure.

- What if we **apply edge-filtering logic** to **Deep Learning features** to enhance stability?

Deep Edge Filter

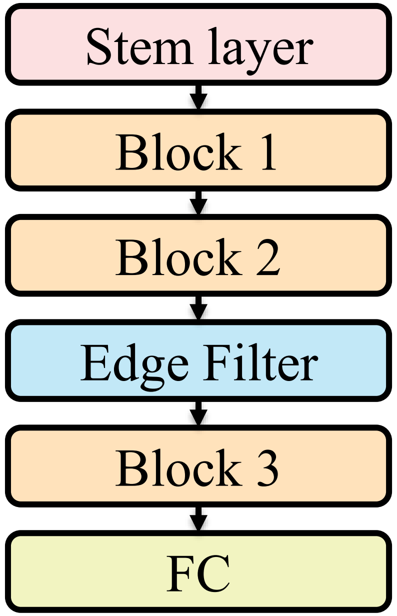
- Key hypothesis

  As in image edge detection, we hypothesize that the high-frequency components of deep features will carry meaningful information.

- Implementation

  High-Pass Filter(HPF) directly applied to deep feature by subtracting LPF value

$$F_{edge}(h) = h - LPF(h)$$



Observation

- When applying a frequency filter to the backbone output in CIFAR100 Linear Probing, the LPF reduced performance, whereas applying an Edge Filter (HPF) improved performance.

	w/o Filter	LPF	Edge Filter
Train acc	21.8	17.4	**22.4**
Val acc	21.0	17.4	**21.4**

Experiment Goal

- We aim to demonstrate that the phenomenon where edge filters remove superficial information while preserving semantic information, thereby increasing a model's generalizability, holds generally across deep learning in architecture-agnostic and modality-agnostic manner.

- To this end, we select and perform tasks where generalization is crucial across various modalities.

Experiment

Vision modality

- task : Test-Time Adaptation

Method	CIFAR-10C				CIFAR-100C				ImageNet200-C			
	Source	Direct	NORM	TENT	Source	Direct	NORM	TENT	Source	Direct	NORM	TENT
WRN28-10	91.9	49.6	73.8	74.6	69.8	26.2	46.6	48.0	58.8	2.0	23.9	21.6
$+F_{edge}$	90.8	57.7 **(+8.0)**	75.3 **(+1.5)**	75.8 **(+1.2)**	65.4	36.4 **(+10.2)**	49.7 **(+3.1)**	50.5 **(+2.5)**	55.5	2.1 **(+0.1)**	25.8 **(+1.9)**	23.2 **(+1.6)**
ViT-B/32	95.6	60.8	-	60.9	82.3	41.3	-	41.0	83.4	27.5	-	30.7
$+F_{edge}$	95.3	68.1 **(+7.3)**	-	69.4 **(+8.5)**	81.8	41.7 **(+0.4)**	-	42.0 **(+1.0)**	93.1	29.4 **(+1.9)**	-	32.6 **(+1.9)**

Experiment

Language modality

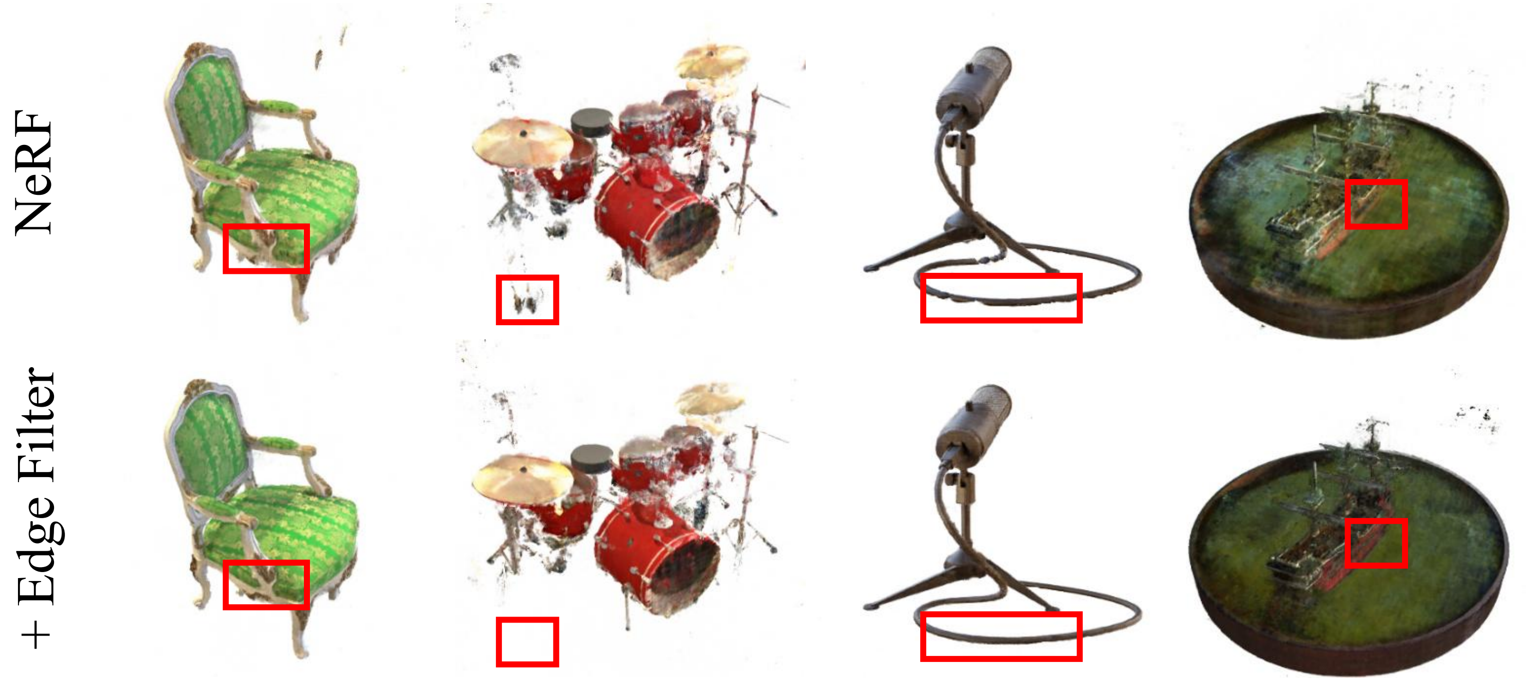
- task : Sentiment Analysis (GLUE benchmark)

	SST-2(Sentiment)	QQP(Paraphrase)	QNLI(Inference)
Transformer	79.36	83.42	62.40
$+F_{edge}$	80.85 **(+1.49)**	83.46 **(+0.04)**	63.30 **(+0.90)**

Experiment

3D modality

- task : Fewshot NeRF (8-view, Blender dataset)
- PSNR : **22.95** → **23.39** **(+0.44)**



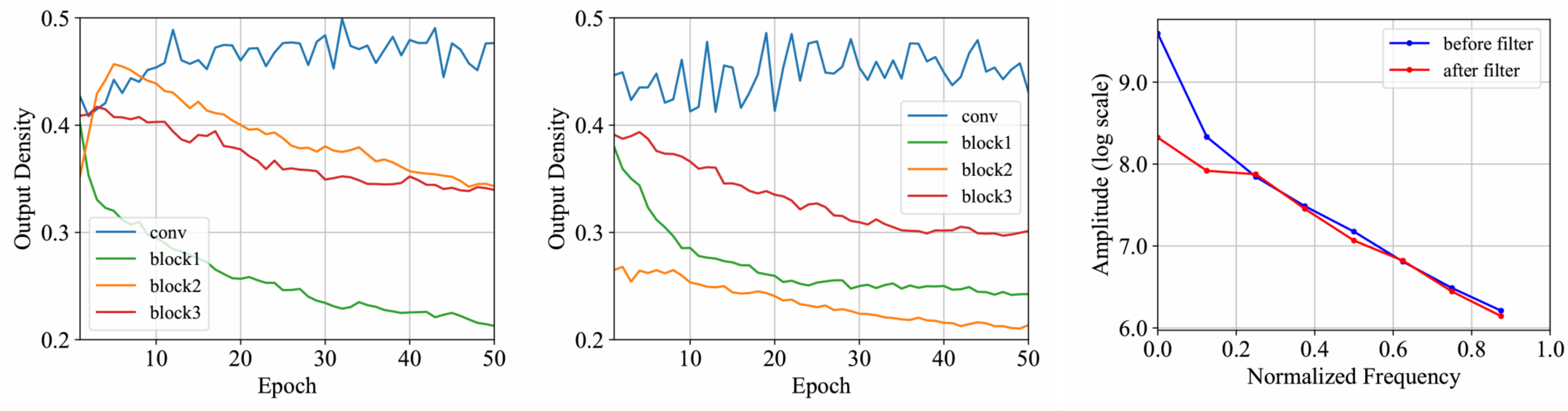
Experiment

Audio modality

- task : Sound classification (UrbanSound8k)
- Accuracy : **77.42% → 81.72%** **(+4.3%p)**

Analysis

- Analysis on block output density and FFT result



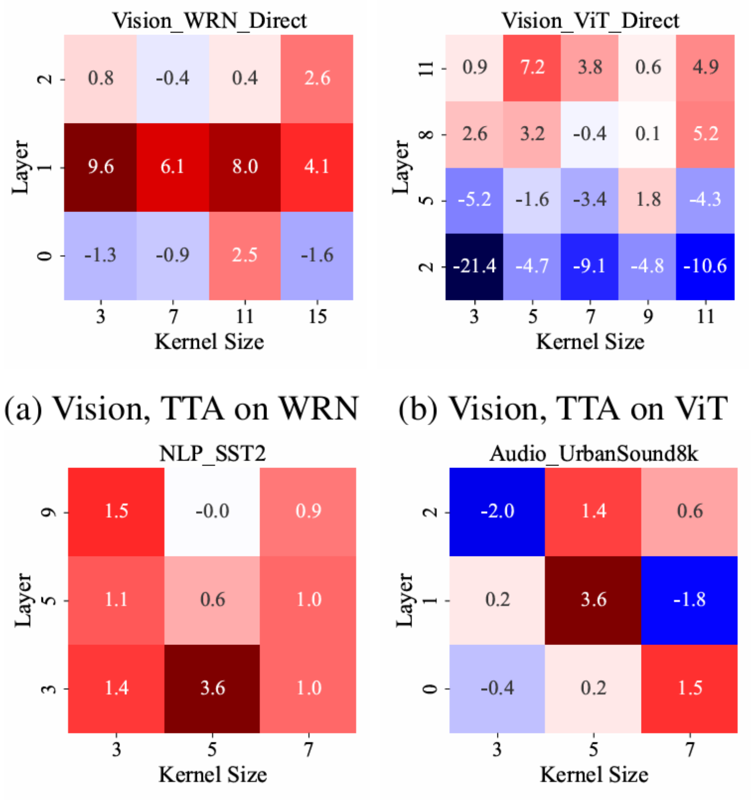
(a) Without Filter **(b) Edge Filter after block1** **(c) Deep Feature FFT**

Ablation studies

by Filter Type

Modality	Vision		Language
Benchmark	CIFAR-10C (Direct)		SST-2
Backbone	WRN	ViT	BERT
Baseline	49.64	60.84	79.36
+Mean $\mathcal{F}_{\text{edge}}$	57.65 **(+8.01)**	68.09 **(+7.25)**	80.85 **(+1.49)**
+Median $\mathcal{F}_{\text{edge}}$	59.69 **(+10.05)**	67.28 **(+6.44)**	81.60 **(+2.24)**
+Gaussian $\mathcal{F}_{\text{edge}}$	57.66 **(+8.02)**	56.37 **(-4.47)**	80.58 **(+1.12)**
+Mean LPF	39.69 **(-9.95)**	47.96 **(-26.68)**	57.56 **(-3.28)**

by Filter Position and Kernel Size



(a) Vision, TTA on WRN (b) Vision, TTA on ViT

(c) Language, SST2 (d) Audio, UrbanSound8k

Conclusion

- We brought the concept of **“Edge Filtering”** to deep neural networks, creating a modality-agnostic and architecture-agnostic approach.

- We reveal that deep learning models store semantic information in **“High Frequency Components”** of deep features.

- Therefore, the Edge Filter serves as a simple yet effective human-crafted layer that significantly improves the robustness of deep learning models