ICLR 2024

PAC-FNO: Parallel-Structured All-Component Fourier Neural Operators for Recognizing Low-Quality Images

Jinsung Jeon[1], Hyundong Jin[1], Jonghyun Choi[2], Sanghyun Hong[3], Dongeun Lee[4], Kookjin Lee[5], Noseong Park[6]

[1]Yonsei University, [2]Seoul National University, [3]Oregon State University,
[4]Texas A&M University-Commerce, [5]Arizona State University, [6]KAIST

Outlines

- **Motivation**
- **Related Works**
- **PAC-FNO**
- **Experiments**
- **Results**

Motivation

- When deep **learning models are deployed to the real world**, they are likely to **face low-quality inputs at inference** (e.g., low-resolution or natural input variations)
- The use of such **low-quality inputs significantly degrades the performance** of visual recognition models.

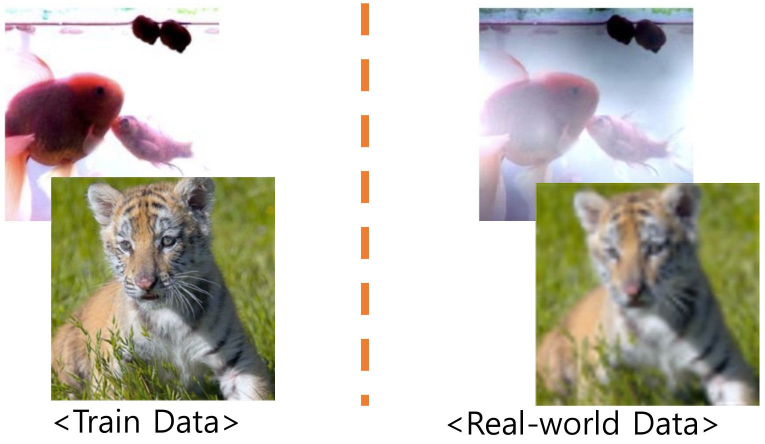


Figure 1: Comparing train data and real-world data.

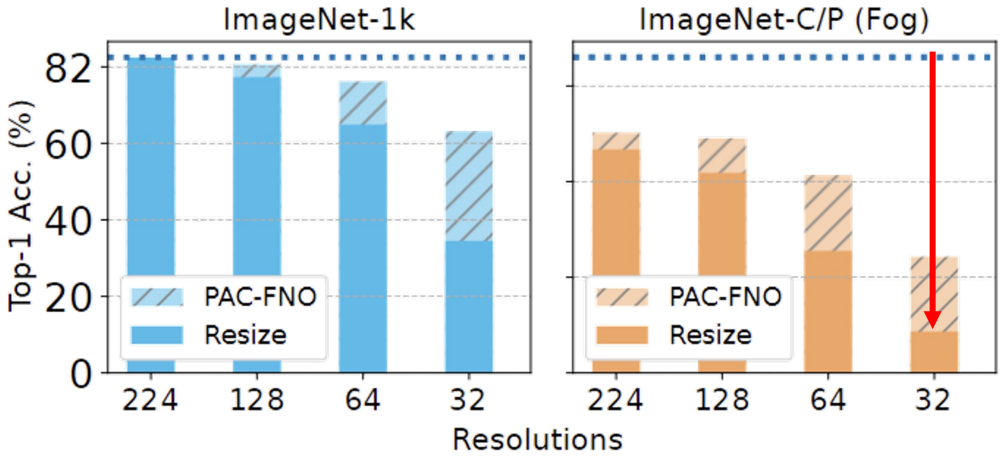


Figure 2: Performance degradation in ImageNet-1k and ImageNet-C/P (Fog).

Related Works

- **Fourier Neural Operators (FNOs)** which achieve remarkable performance in solving PDEs, are mathematically defined under the infinite-dimensional continuous space regime and for this reason, **they can process various resolutions of the continuous space without model change.**
- For efficiency, FNOs restrict the size of the learnable parameter R_θ by using an ideal low-pass filter ξ.

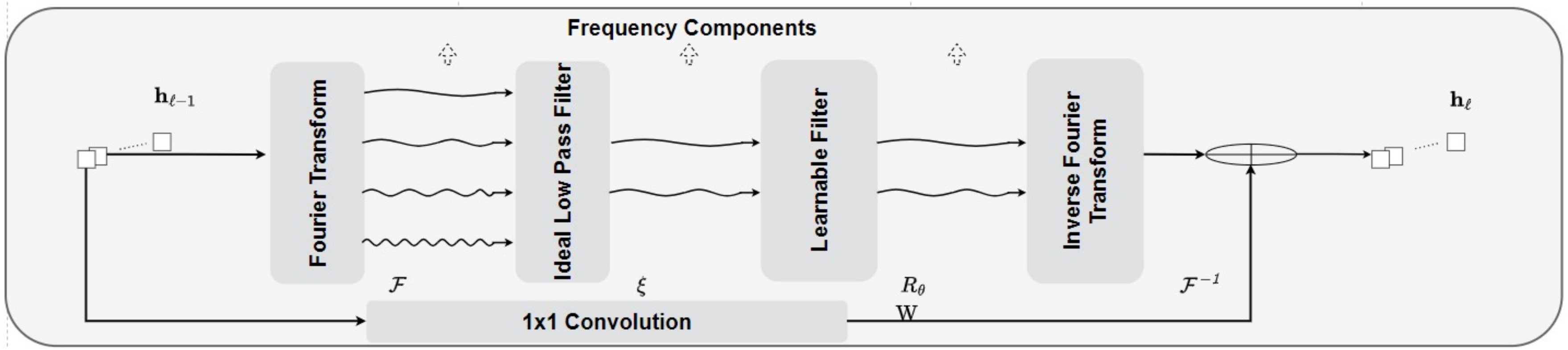
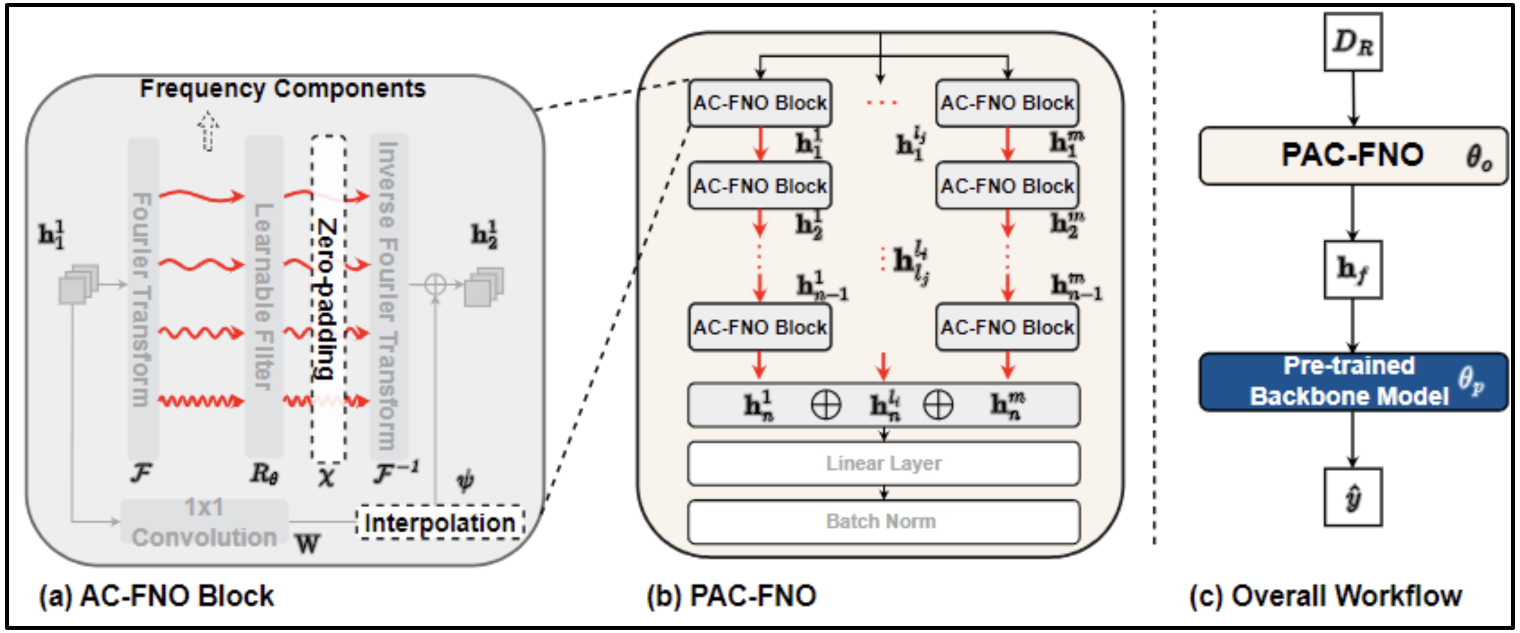


Figure 3: Fourier Neural Operators.

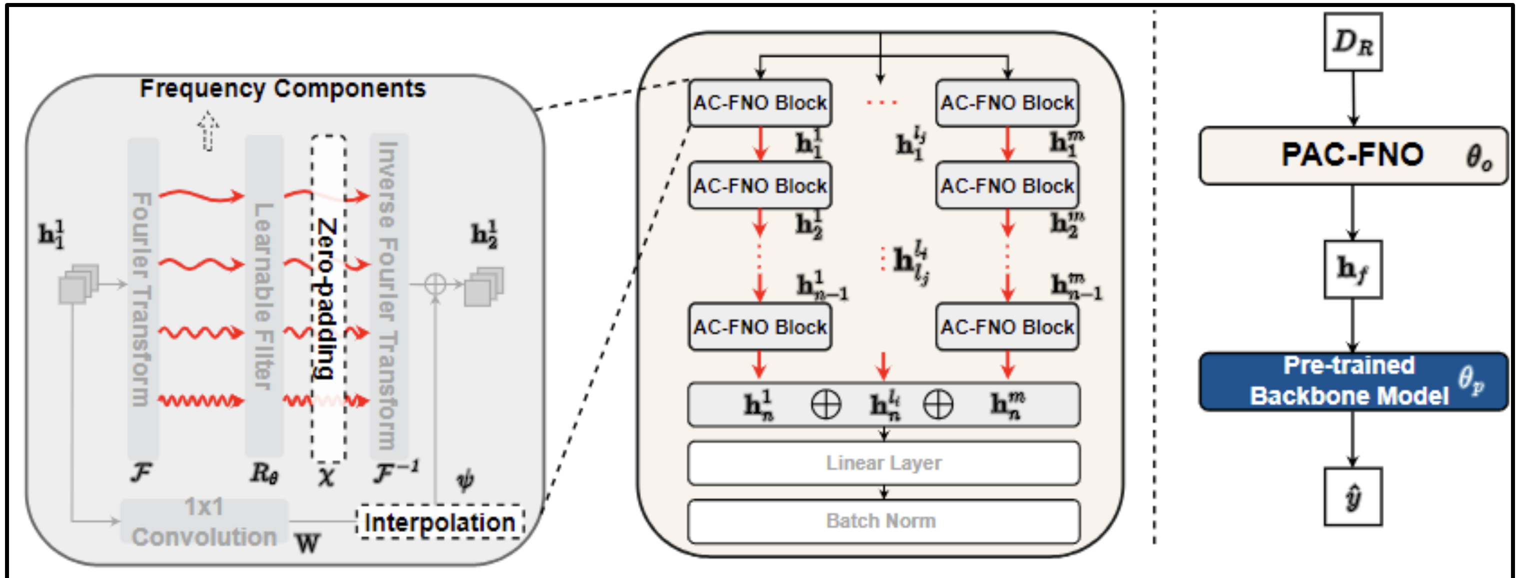
PAC-FNO: Parallel-Structured All-Component Fourier Neural Operators for Recognizing Low-Quality Images



(a) All-Component-FNO (AC-FNO) Block

- In the case of images, high-frequency information sometimes plays an important role in image classification, especially when detailed information is required (type of bird, type of car, etc). To this end, we propose an AC-FNO block without any band pass filters.
- AC-FNO blocks use all frequency components and rely on Zero-padding and Interpolation to construct the images for the target resolution.
- $\mathcal{F}$ is a Fourier transform that transforms the hidden vector into the frequency domain, χ a zero-padding, R_θ is a learnable filter, ψ is an interpolation and W is a 1x1 convolutional operation.

PAC-FNO: Parallel-Structured All-Component Fourier Neural Operators for Recognizing Low-Quality Images

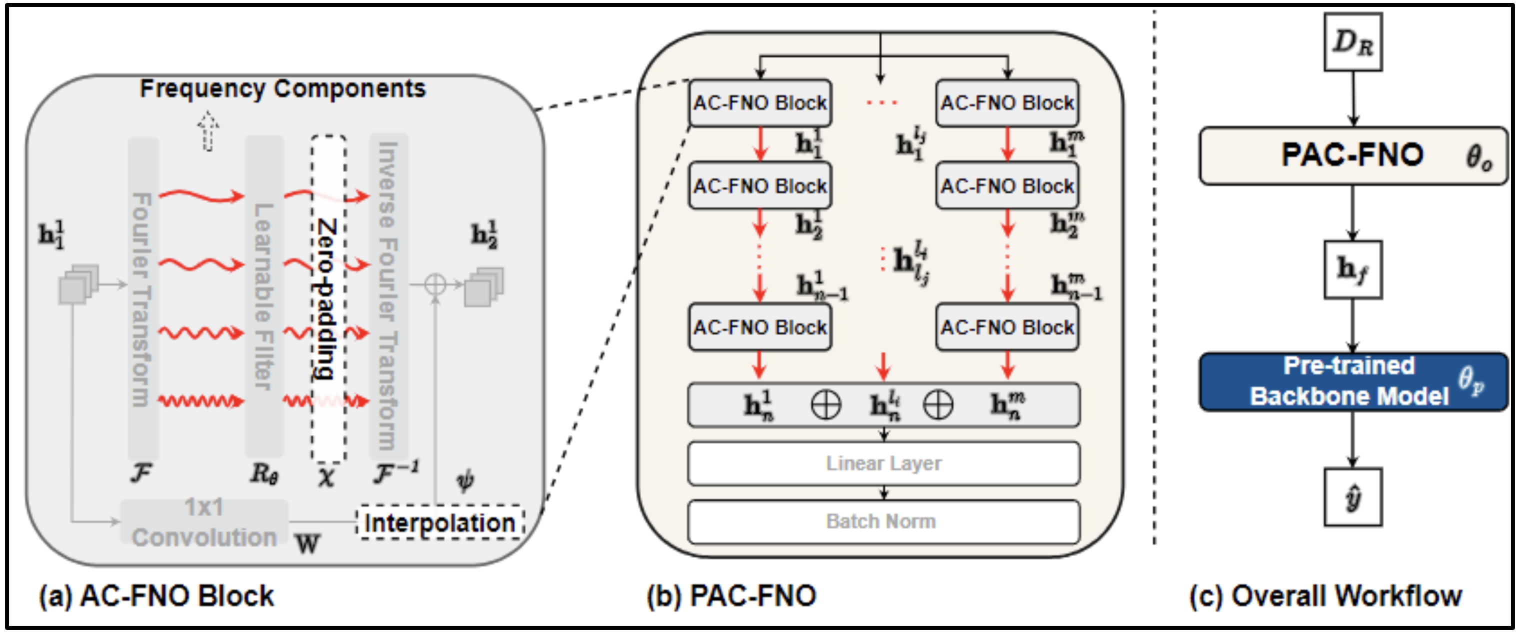


(a) AC-FNO Block (b) PAC-FNO (c) Overall Workflow

(b) Parallel Configuration of AC-FNO Blocks

- To increase the capacity to learn various types of input variations, we propose to configure AC-FNO blocks in a parallel structure.
- There are n x m AC-FNO blocks, where n is the number of stages and m is the number of parallel AC-FNO blocks in a stage.

PAC-FNO: Parallel-Structured All-Component Fourier Neural Operators for Recognizing Low-Quality Images



Two-stage Learning Algorithm

- **First stage: We jointly train the PAC-FNO and the pre-trained backbone model together using only the target resolution for which the backbone model was trained.**
- **Second stage: We fine-tune the well-harmonized model with images in low resolution to generate a unified hidden space for all resolutions. In this stage, the backbone model is frozen.**

Experiments

Table 1: Performance of PAC-FNO on the low-resolution tasks using ImageNet-1k

Model	Method	Resolution							
		28	32	56	64	112	128	**224**	**299**
ResNet-18	Resize	16.7	22.1	45.7	50.5	63.7	65.5	69.8	-
	Fine-tune	1.01	2.15	10.2	16.3	34.5	50.6	65.0	-
	DRLN	0.22	-	17.1	-	62.8	-	69.8	-
	DRPN	31.6	-	55.6	-	67.5	-	69.8	-
	FNO	40.2	45.2	59.1	61.5	67.6	68.5	70.1	-
	UNO	39.9	45.3	58.9	61.3	67.1	68.1	69.4	-
	A-FNO	**43.3**	**49.5**	60.2	62.5	66.9	67.5	68.5	-
	PAC-FNO	42.7	47.7	**60.5**	**62.8**	**68.3**	**69.1**	**70.2**	-
Inception-V3	Resize	16.7	22.0	48.6	53.9	69.5	71.8	-	77.3
	Fine-tune	39.6	47.2	63.7	69.8	72.9	73.3	-	77.5
	FNO	48.9	54.0	68.5	70.2	74.9	76.5	-	**78.4**
	UNO	42.2	48.1	65.4	68.0	74.7	75.5	-	77.3
	A-FNO	33.4	39.6	59.6	62.7	71.0	72.1	-	74.9
	PAC-FNO	**49.3**	**54.9**	**68.8**	**70.7**	**76.1**	**76.9**	-	**78.4**

Model	Method	Resolution							
		28	32	56	64	112	128	**224**	**299**
ViT-B16	Resize	41.6	47.4	66.0	68.9	76.8	78.1	**81.1**	-
	Fine-tune	39.8	48.5	66.6	69.6	77.3	78.4	81.1	-
	DRLN	3.75	-	42.0	-	77.0	-	81.1	-
	DRPN	52.2	-	**73.0**	-	**80.0**	-	81.1	-
	FNO	39.1	46.6	66.2	69.2	77.1	78.0	79.7	-
	UNO	43.7	50.6	67.9	70.3	77.9	78.8	80.2	-
	A-FNO	**54.9**	**59.5**	72.3	**74.1**	78.3	**78.9**	79.5	-
	PAC-FNO	46.6	52.1	69.0	71.6	77.7	78.6	79.6	-
ConvNeXt-Tiny	Resize	27.5	34.5	60.7	64.9	75.0	77.3	**82.5**	-
	Fine-tune	40.2	62.3	65.8	76.0	76.4	**80.7**	81.8	-
	DRLN	0.30	-	25.1	-	71.8	-	82.5	-
	DRPN	40.7	-	68.2	-	79.4	-	82.5	-
	FNO	42.8	48.9	67.4	69.8	76.5	77.4	79.2	-
	UNO	58.6	62.9	73.7	75.5	79.1	79.7	80.6	-
	A-FNO	48.3	53.5	68.1	70.5	75.9	76.6	78.1	-
	PAC-FNO	**58.9**	**63.2**	**74.5**	**76.2**	**80.2**	**80.7**	81.5	-

Table 2: Performance of PAC-FNO on low-resolution tasks using fine-grained datasets and input variation tasks

Dataset	Method	Resolution						
		28	32	56	64	112	128	**224**
Oxford-IIIT Pets	Resize	29.4	36.4	70.1	77.0	89.9	91.2	93.7
	Fine-tune	32.6	41.1	73.2	79.3	91.0	**91.5**	**93.8**
	DRLN	3.37	-	36.8	-	88.1	-	93.7
	DRPN	41.5	-	84.2	-	**92.6**	-	93.7
	FNO	19.1	54.0	60.6	70.2	86.7	90.4	91.4
	UNO	11.1	15.7	43.2	50.5	80.1	83.8	90.3
	A-FNO	27.0	33.3	63.5	70.2	85.2	86.9	89.1
	PAC-FNO	**73.4**	**77.3**	**86.0**	**87.6**	90.5	91.1	91.7
Flowers	Resize	39.5	47.6	75.3	80.3	92.1	93.7	95.9
	Fine-tune	44.3	51.1	78.8	81.9	92.5	94.0	95.9
	DRLN	9.92	-	53.4	-	91.3	-	95.9
	DRPN	64.5	-	89.0	-	**95.0**	-	95.9
	FNO	25.3	32.8	65.9	73.0	91.7	93.7	**96.6**
	UNO	23.2	30.2	64.5	71.7	90.8	92.6	96.1
	A-FNO	26.8	33.6	63.7	70.0	86.2	88.8	91.9
	PAC-FNO	**74.1**	**78.0**	**87.6**	**89.3**	93.6	**94.2**	94.4

Variation	Method	Resolution						
		24	32	54	64	112	128	224
Fog	Resize	8.21	10.8	27.2	31.9	48.5	52.3	58.4
	Fine-tune	23.2	28.2	47.5	51.4	**61.0**	**62.2**	**63.0**
	DRLN	0.16	-	6.54	-	33.5	-	58.4
	DRPN	0.67	-	0.99	-	1.32	-	58.4
	FNO	14.4	18.6	38.9	43.5	56.6	58.4	60.3
	UNO	24.6	29.7	46.8	49.5	59.2	60.4	59.0
	A-FNO	11.0	14.9	33.4	38.3	51.9	53.6	53.3
	PAC-FNO	**25.4**	**30.4**	**48.2**	**51.7**	60.1	61.4	62.8
Brightness	Resize	22.0	28.2	52.7	56.8	67.8	69.9	73.5
	Fine-tune	42.8	48.9	65.8	68.1	**73.5**	**74.4**	**75.7**
	DRLN	6.64	-	35.71	-	56.1	-	58.4
	DRPN	3.09	-	1.08	-	8.01	-	58.4
	FNO	34.9	40.7	59.4	62.6	70.0	71.3	73.1
	UNO	50.8	55.6	66.7	67.9	71.4	72.1	72.8
	A-FNO	37.6	43.2	59.4	62.5	68.8	69.8	70.0
	PAC-FNO	**51.9**	**56.5**	**68.2**	**69.8**	73.4	74.1	74.7

Resize: Resize is a method that directly feeds the resized images to a pre-trained classification model using interpolation
Fine-tune: Fine-tune is a method of fine-tuning the pre-trained classification model with the resized images.
DRLN and DRPN: These two models are representative models capable of up to 8 times super-resolution (SR). With SR models, low-resolution images are upscaled to the target resolution and fed into a pre-trained model.
FNO, UNO, A-FNO: These three models are Fourier neural operator models. These models are most similar to our PAC-FNO and are trained by our proposed training method for fair comparison.

Experiments

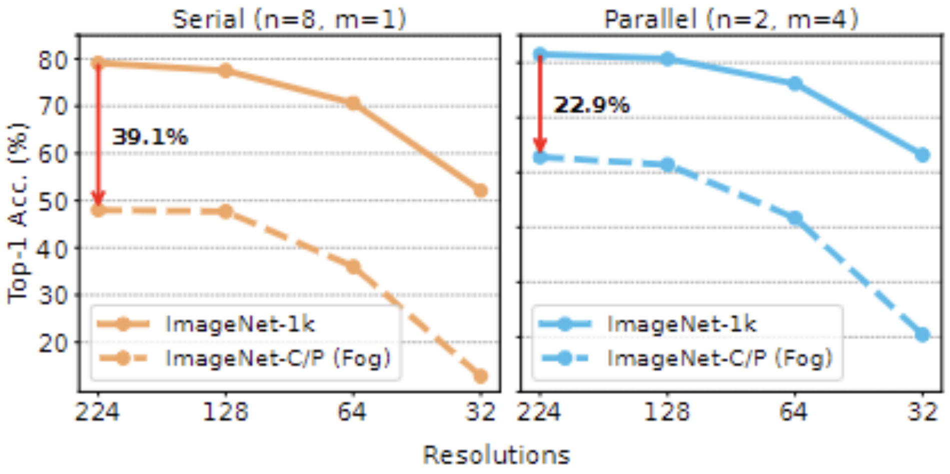


Figure 3: Benefit of parallel structure.

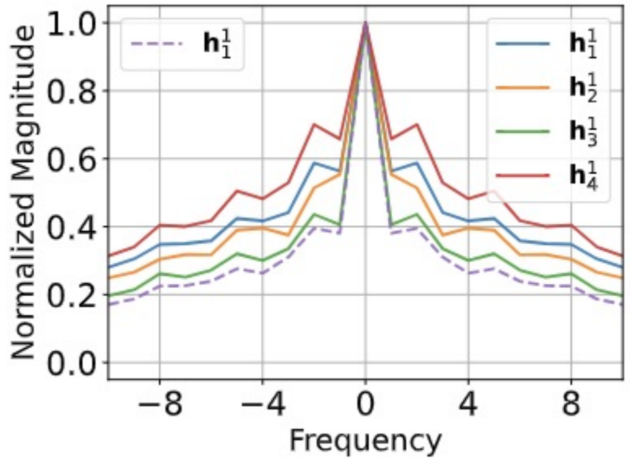


Figure 4: Comparison of spectral responses according to the configuration of the AC-FNO block

Results

- We proposed **parallel-structured and all-component Fourier neural operators** (PAC-FNOs) for visual recognition under low-quality images.
- To this end, we design i) an **AC-FNO** and ii) a **parallel** configuration of AC-FNO blocks and also propose a **two-stage training algorithm.**
- As a result, PAC-FNO provides **two advantages** over existing methods: (i) **It can handle both low-resolution and input variations** typically observed in low-quality images with a single model; (ii) **One can attach PAC-FNO to any visual recognition model** and fine-tune it.
- In the evaluation with four visual recognition models and seven datasets, we show that PAC-FNO achieves high accuracy for various resolutions and input variations.

ICLR 2024

PAC-FNO: Parallel-Structured All-Component Fourier Neural Operators for Recognizing Low-Quality Images

Presenter: Jinsung Jeon

Email: jjsjjs0902@yonsei.ac.kr, Homepage: https://sites.google.com/view/jsjeon

Tue 7 May 10:45 a.m. CEST — 12:45 p.m. CEST, Hall B

YONSEI UNIVERSITY
SEOUL NATIONAL UNIVERSITY
Oregon State University
TEXAS A&M UNIVERSITY COMMERCE
Arizona State University
KAIST