2025 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING (ICASSP 2025)
April 06 - 11, 2025 Hyderabad, India

Mobile-friendly Image de-noising: Hardware Conscious Optimization for Edge Application

Sowmya Vajrala#, Srinivas S Miriyala#*, Hitesh Kumar, Sravanth Kodavanti, and Vikram N Rajendiran
Samsung Research Institute Bangalore, India
{v.lahari, srinivas.m1, h1.kumar, k.sravanth, vikram.nr}@samsung.com | #Equal contribution & *Corresponding author

Background

- Image enhancement is a critical task in computer vision and Mobile photography that is often entangled with noise.
- This renders the traditional Image Signal Processing (ISP) ineffective compared to the advances in deep learning (DL).
- However, the success of such methods is increasingly associated with their ease of on-device (e.g. mobile) deployment.
- Image de-noising with an efficiently designed AI has become the need-of-the-hour in Mobile Image Signal Processing Pipeline.
- One-shot super-net based strategies remain the most effective methods for the design of deep neural networks, they have been less explored for Image De-noising in the literature.
- In summary, the reported strategies for differentiable NAS face the following issues when scaled for the task of Image De-noising:

Limited Search Space exploration
Differentiable NAS works efficiently in simpler search spaces but struggle in complex domains

Performance Collapse
To facilitate training, the discrete optimization is smoothened, resulting in a bias for suboptimal small networks

Indecisiveness
due to equal weightage on different subnets because of improper tuning of hyper-parameters (e.g. the temp. in Softmax).

Optimization instabilities
arising from different scales of architecture parameters and weights combined with Training instabilities

- Further, none of the existing NAS methods utilize the trained SoTA.

Methodology

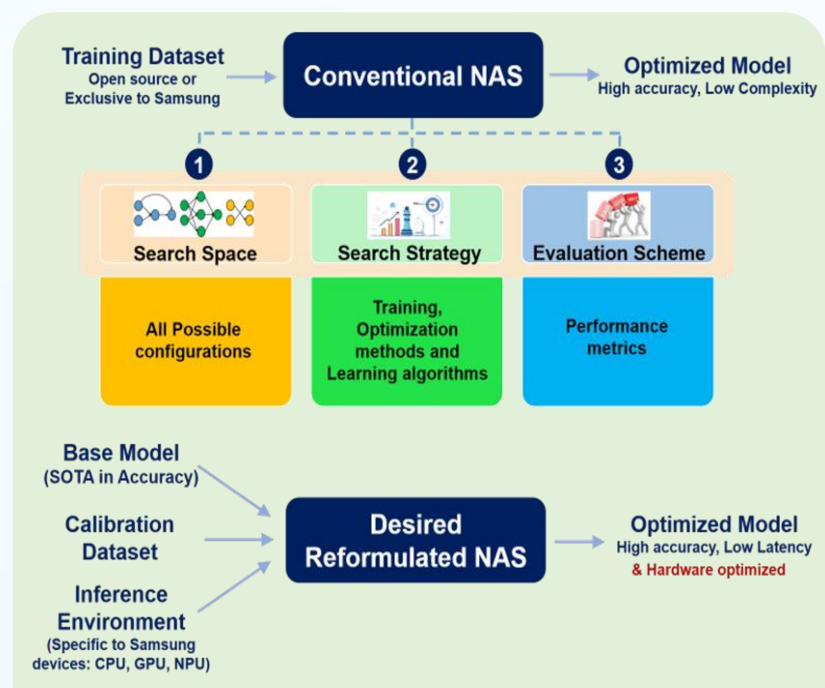

Figure 1. Need for Reformulated NAS

SoTA Pre-trained AI as the Base model
↓
Hardware aware Search Space design
Hardware aware Alternatives are associated with a penalty $\wp_i \; \forall \, i = 1 \text{ to } K$ that correlates with their on-device latency and/or GFLOPs
↓
Entropy Regularized Differentiable NAS
↓
Entropy Regularized Neural Network or **ERN-Net**

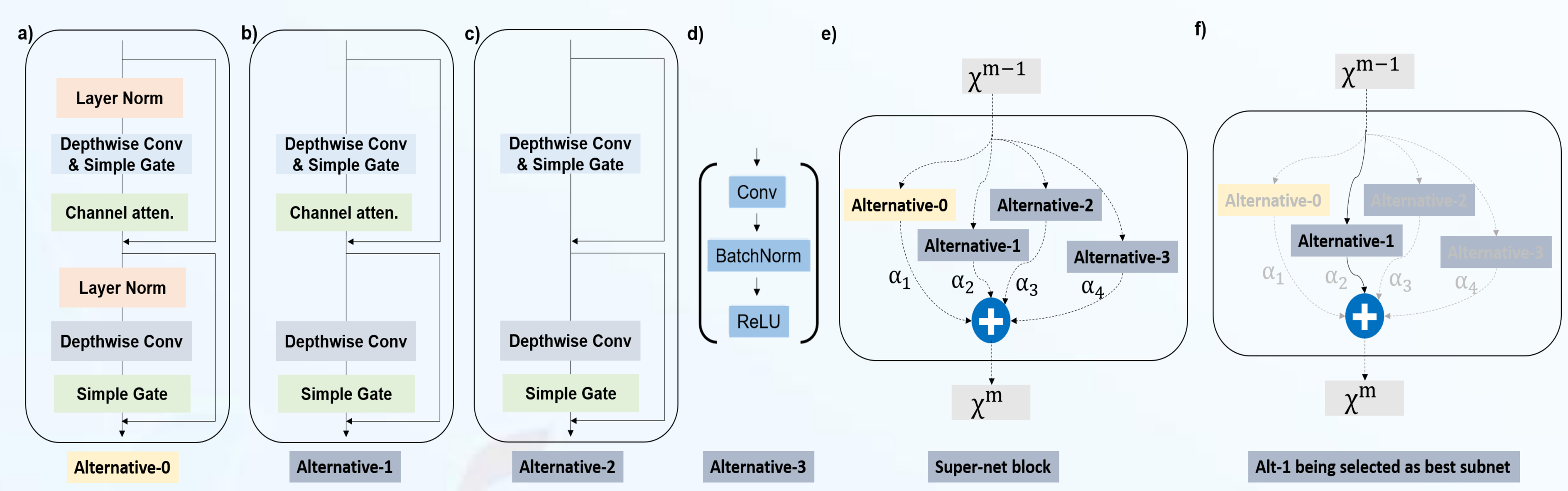

Figure 2. Hardware aware search space with a) base block, b) to d) hardware aware alternatives and e) the super-net block.

Output of m^{th} block → $\chi^m = \sum_{i=1}^{K} \alpha_i f(\chi^{m-1})_i$

Arch. Encodings → $\alpha_i = \mathrm{Softmax}(\varphi_i \forall\, i = 1 \text{ to } K)$

Training Loss → $L_T \leftarrow \mathcal{F}(\chi^{out}, \chi^{GT})$

$L_P = \sum_{m=1}^{M} \left[\sum_{i=1}^{K} \alpha_i [\wp_i]\right]_m$ ← Penalty for complexity

$L_{ER} = -\sum_{m=1}^{M} \left[\sum_{i=1}^{K} \alpha_i \log(\alpha_i)\right]_m$ ← ER Loss

$\{\boldsymbol{w}, \boldsymbol{\varphi}\}^* = \operatorname{argmin}(L = L_T + L_P + L_{ER})$

Results

- The search space for the Image De-noising was designed based on the 36-block NAFNet model [1]
- Given the hardware for deployment (e.g. the Neural Processing Unit on the Qualcomm's sm8650 chipset), various hardware-friendly alternatives to the NAF block (see Fig. 2) were identified.
- Super-net is trained on SIDD [2] (real de-noising) and tested on SenseNoise [3] and Gaussian De-noising.

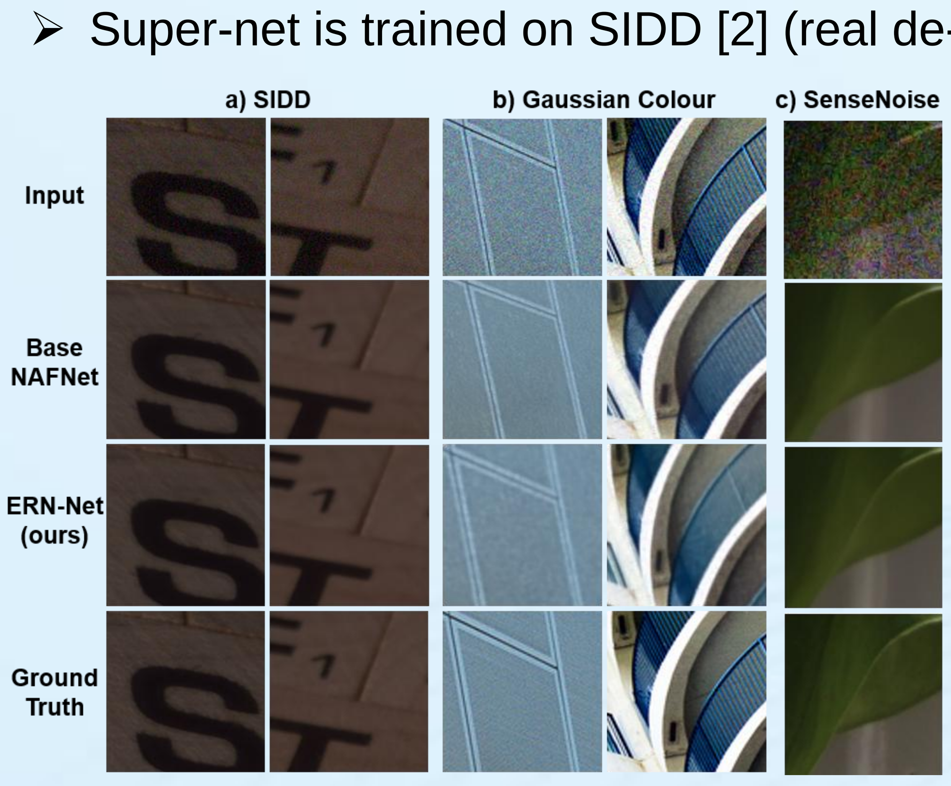

Figure 3. Visual comparison of outputs from ERN-Net and the base model for de-noising Gaussian and raw noise (SIDD & SenseNoise)

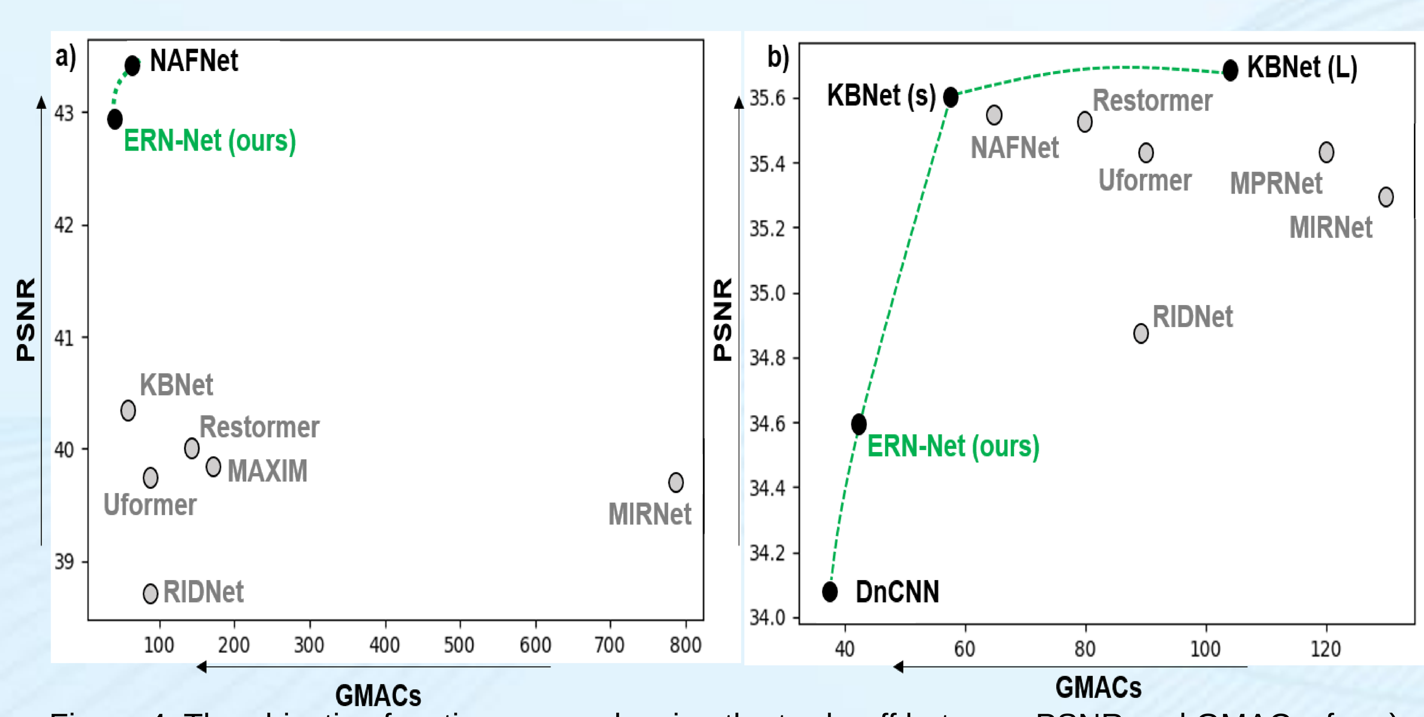

Figure 4. The objective function space showing the trade-off between PSNR and GMACs for a) SIDD and b) SenseNoise dataset. Green dashed curve is the non-dominated front.

		MIRNet	MAXIM	Restormer	RIDNet	Uformer	KBNet	NAFNet	ERN-Net
SIDD	PSNR	39.72	39.89	40.02	38.71	39.75	40.35	43.42	43.09
	SSIM	0.959	0.960	0.960	0.914	0.959	0.972	0.959	0.956
	GMACs	786	169.5	140	89	88.8	57.8	65	42
Sense Noise	PSNR	35.30	35.43	35.52	34.88	35.43	35.6	35.55	34.6
	SSIM	0.919	0.922	0.924	0.915	0.920	0.924	0.923	0.988
	GMACs	130	120	80	89	90	57.8	65	42

- 12% less parameters, ~2 fold improvement in on-device (Samsung GGS24 ultra) latency and 13% improvement in memory with 0.7% and 0.3% drop in PSNR and SSIM, respectively.

Conclusions

- Effort to optimize pre-trained AI for Image De-noising for edge deployment.
- Proposed framework offers network surgery, hardware-aware search space design and entropy regularized differentiable search strategy, that is first-of-its-kind.
- Development and deployment of complex AI models for image de-noising are both resolved successfully without downplaying one on the grounds of other.
- Future scope: the generic strategy can be applied for optimization and deployment of compute and memory-intensive models for Generative Artificial Intelligence.

References

[1] Chen, L., Chu, X., Zhang, X., & Sun, J. Simple baselines for image restoration. In European conference on computer vision, 17-33 (2022).
[2] Abdelhamed, A., Lin, S., & Brown, M. S. A high-quality denoising dataset for smartphone cameras. In Proceedings of the IEEE conference on CVPR, 1692-1700 (2018).
[3] Zhang, Y., Li, D., Law, K. L., Wang, X., Qin, H., & Li, H. (2022). Idr: Self-supervised image denoising via iterative data refinement. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 2098-2107).
[4] Liu, H., Simonyan, K., & Yang, Y. (2018). Darts: Differentiable architecture search. arXiv preprint arXiv:1806.09055.

Thank You!