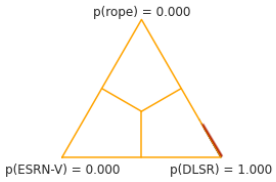
# Supplementary Material: An Experimental Protocol for Neural Architecture Search in Super-Resolution

Jesús Leopoldo Llano García Raúl Monroy Víctor Adrián Sosa Hernández
Tecnologico de Monterrey, School of Engineering and Sciences
Av. Lago de Guadalupe Km 3.5, Atizapán de Zaragoza, Edo. Mexico 52926, MEXICO
{A01748867, raulm, vsosa}@tec.mx

## 1. Visual presentation of the Bayesian analysis

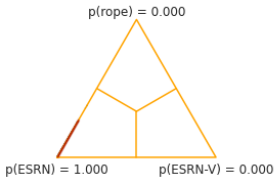
The figures presented below are representations of the posterior probabilities from performing a Bayesian signed-rank test on the results of our evaluation protocol. The experimental set considers 9 architectures: NAS-DIP [2], architectures A and C from HNAS [5], architectures A and B from FALSR [3], architecture A from MoreMNAS [4], ESRN and ESRN-V from [9], and DLSR from [6]. We measure architecture performance using their default parameters over three different resolutions of Set5 [1], Set14 [10], BSD100 [8] and Urban100 [7].



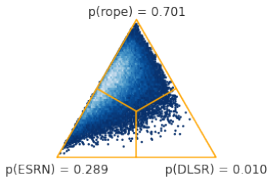

(a) ESRN-V VS DLSR

Figure 1: The probabilities of ESRN performing better than DLSR. Each vertex of the simplex describes numerically which algorithm has a larger probability of performing better. The blue dots represent the Monte Carlo samples computed for each comparison. A lack of visible blue dots on a graph signifies that all samples rest directly on the edge and vertex of a specific simplex region; this region is delimited with red.

Each simplex is generated plotting 1,500,000 Markov-Chain Monte Carlo samples, based on the probabilities of paired models, allocated in barycentric coordinates: each vector of probabilities is a point in a simplex having vertices (1,0,0),(0,1,0),(0,0,1). The vertices of the triangular structure seen on the images represent decisions that favor model $A$, equal performance, or model $B$, respectively.



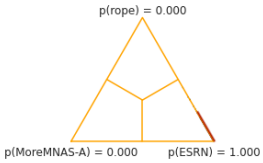

(a) ESRN VS ESRN-V



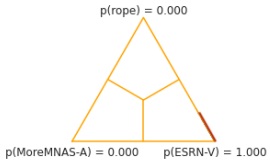

(b) ESRN VS DLSR

Figure 2: The probabilities of ESRN performing better than the other models. Each vertex of the simplex describes numerically which algorithm has a larger probability of performing better. The blue dots represent the Monte Carlo samples computed for each comparison. A lack of visible blue dots on a graph signifies that all samples rest directly on the edge and vertex of a specific simplex region; this region is delimited with red.

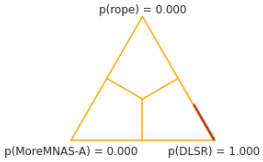
Some graphs seemingly present no blue dots, this indicates that the probability samples are located on the edge and vertex of a particular region. A situation that derives from the posterior probabilities determining a negligible chance for the other vertex to attract samples. We delimit with a red bold line the region where samples cluster on the edges of the simplex.




(a) MoreMNAS-A VS ESRN

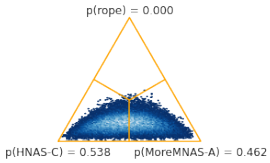



(b) MoreMNAS-A VS ESRN-V

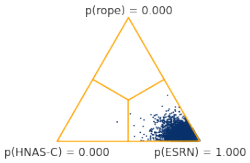



(c) MoreMNAS-A VS DLSR

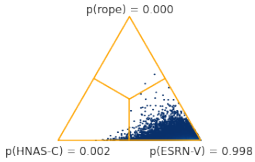
Figure 3: The probabilities of MoreMNAS-A performing better than the other models. Each vertex of the simplex describes numerically which algorithm has a larger probability of performing better. The blue dots represent the Monte Carlo samples computed for each comparison. A lack of visible blue dots on a graph signifies that all samples rest directly on the edge and vertex of a specific region of the simplex, this region is delimited with red.



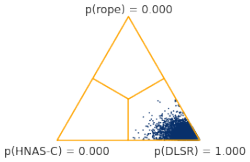

(a) HNAS-C VS MoreMNAS-A




(b) HNAS-C VS ESRN

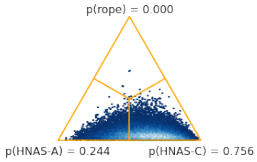



(c) HNAS-C VS ESRN-V

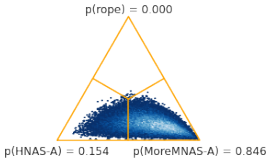



(d) HNAS-C VS DLSR

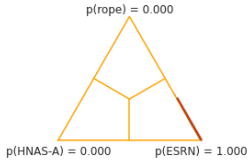
Figure 4: The probabilities of HNAS-C performing better than the other models. Each vertex of the simplex describes numerically which algorithm has a larger probability of performing better. The blue dots represent the Monte Carlo samples computed for each comparison. A lack of visible blue dots on a graph signifies that all samples rest directly on the edge and vertex of a specific simplex region; this region is delimited with red.



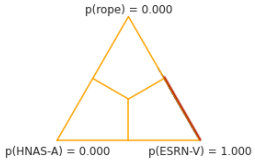

(a) HNAS-A VS HNAS-C



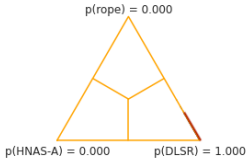

(b) HNAS-A VS MoreMNAS-A




(c) HNAS-A VS ESRN

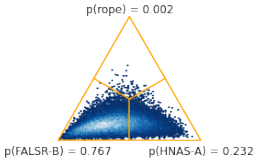



(d) HNAS-A VS ESRN-V

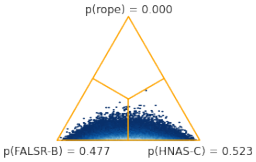



(e) HNAS-A VS DLSR

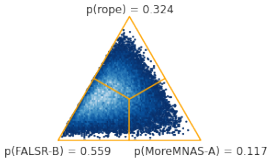
Figure 5: The probabilities of HNAS-A performing better than the other models. Each vertex of the simplex describes numerically which algorithm has a larger probability of performing better. The blue dots represent the Monte Carlo samples computed for each comparison. A lack of visible blue dots on a graph signifies that all samples rest directly on the edge and vertex of a specific simplex region; this region is delimited with red.



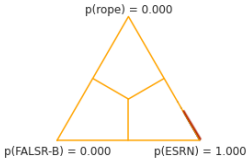

(a) FALSR-B VS HNAS-A



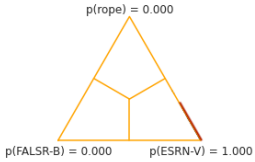

(b) FALSR-B VS HNAS-C



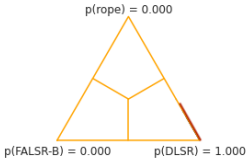

(c) FALSR-B VS MoreMNAS-A




(d) FALSR-B VS ESRN

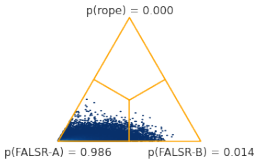



(e) FALSR-B VS ESRN-V

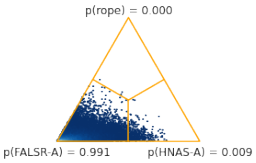



(f) FALSR-B VS DLSR

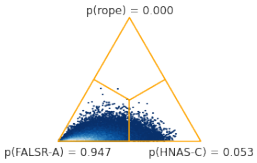
Figure 6: The probabilities of FALSR-B performing better than the other models. Each vertex of the simplex describes numerically which algorithm has a larger probability of performing better. The blue dots represent the Monte Carlo samples computed for each comparison. A lack of visible blue dots on a graph signifies that all samples rest directly on the edge and vertex of a specific simplex region; this region is delimited with red.



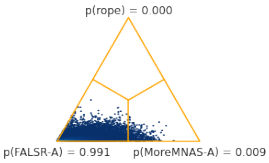

(a) FALSR-A VS FLSR-B



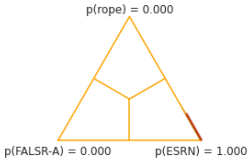

(b) FALSR-A VS HNAS-A



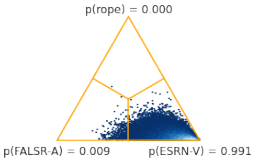

(c) FALSR-A VS HNAS-C



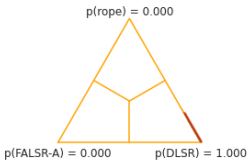

(d) FALSR-A VS MoreMNAS-A




(e) FALSR-A VS ESRN

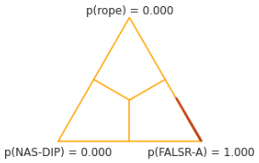



(f) FALSR-A VS ESRN-V

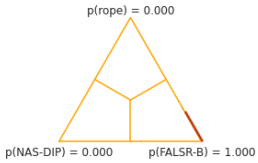



(g) FALSR-A VS DLSR

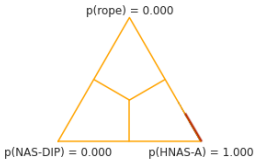
Figure 7: The probabilities of FALSR-A performing better than the other models. Each vertex of the simplex describes numerically which algorithm has a larger probability of performing better. The blue dots represent the Monte Carlo samples computed for each comparison. A lack of visible blue dots on a graph signifies that all samples rest directly on the edge and vertex of a specific simplex region; this region is delimited with red.



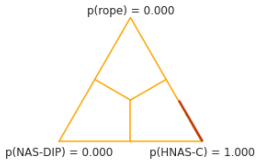

(a) NAS-DIP VS FALSR-A



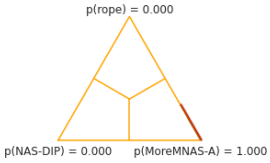

(b) NAS-DIP VS FALSR-B



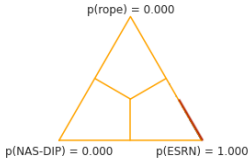

(c) NAS-DIP VS HNAS-A



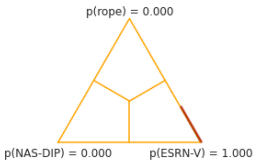

(d) NAS-DIP VS HNAS-C



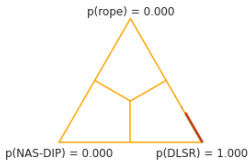

(e) NAS-DIP VS MoreMNAS-A




(f) NAS-DIP VS ESRN




(g) NAS-DIP VS ESRN-V




(h) NAS-DIP VS DLSR

Figure 8: The probabilities of NAS-DIP performing better than the other models. Each vertex of the simplex describes numerically which algorithm has a larger probability of performing better. The blue dots represent the Monte Carlo samples computed for each comparison. A lack of visible blue dots on a graph signifies that all samples rest directly on the edge and vertex of a specific simplex region; this region is delimited with red.

## References


[1] Marco Bevilacqua, Aline Roumy, Christine Guillemot, and Marie Line Alberi-Morel. Low-complexity single-image super-resolution based on nonnegative neighbor embedding. In *23rd British Machine Vision Conference (BMVC)*, pages 135.1–135.10. BMVA press, 2012.

[2] Yun-Chun Chen, Chen Gao, Esther Robb, and Jia-Bin Huang. Nas-dip: Learning deep image prior with neural architecture search. In Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm, editors, *Computer Vision – ECCV 2020*, pages 442–459, Cham, 2020. Springer International Publishing.

[3] Xiangxiang Chu, Bo Zhang, Hailong Ma, Ruijun Xu, and Qingyuan Li. Fast, accurate and lightweight super-resolution with neural architecture search, 2020.

[4] Xiangxiang Chu, Bo Zhang, Ruijun Xu, and Hailong Ma. Multi-objective reinforced evolution in mobile neural architecture search, 2019.

[5] Yong Guo, Yongsheng Luo, Zhenhao He, Jin Huang, and Jian Chen. Hierarchical neural architecture search for single image super-resolution. *IEEE Signal Processing Letters*, 27:1255–1259, 2020.

[6] Han Huang, Li Shen, Chaoyang He, Weisheng Dong, Haozhi Huang, and Guangming Shi. Lightweight image super-resolution with hierarchical and differentiable neural architecture search, 2021.

[7] Jia-Bin Huang, Abhishek Singh, and Narendra Ahuja. Single image super-resolution from transformed self-exemplars. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5197–5206, 2015.

[8] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, volume 2, pages 416–423 vol.2, 2001.

[9] Dehua Song, Chang Xu, Xu Jia, Yiyi Chen, Chunjing Xu, and Yunhe Wang. Efficient residual dense block search for image super-resolution, 2019.

[10] Roman Zeyde, Michael Elad, and Matan Protter. On single image scale-up using sparse-representations. In Jean-Daniel Boissonnat, Patrick Chenin, Albert Cohen, Christian Gout, Tom Lyche, Marie-Laurence Mazure, and Larry Schumaker, editors, *Curves and Surfaces*, pages 711–730, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.