Neural Architecture Search for Efficient Uncalibrated Deep Photometric Stereo

Francesco Sarno[1], Suryansh Kumar[1*], Berk Kaya[1], Zhiwu Huang[1], Vittorio Ferrari[2], Luc Van Gool[1,3]
Computer Vision Lab, ETH Zurich[1], Google Research[2], KU Leuven[3]

WACV
WAIKOLOA, HAWAII
JANUARY 4-8, 2022

Problem Definition and Contributions

Motivation
- Calibrated settings are challenging to reproduce**.**
- Existing deep learning methods are suboptimal.

Goal
Explore applicability of NAS to deep uncalibrated photometric stereo.

Contributions
- Automatic network design for deep uncalibrated photometric stereo.
- Comparable or better performances than existing methods.
- Dramatically lower number of required parameters.

Proposed Method

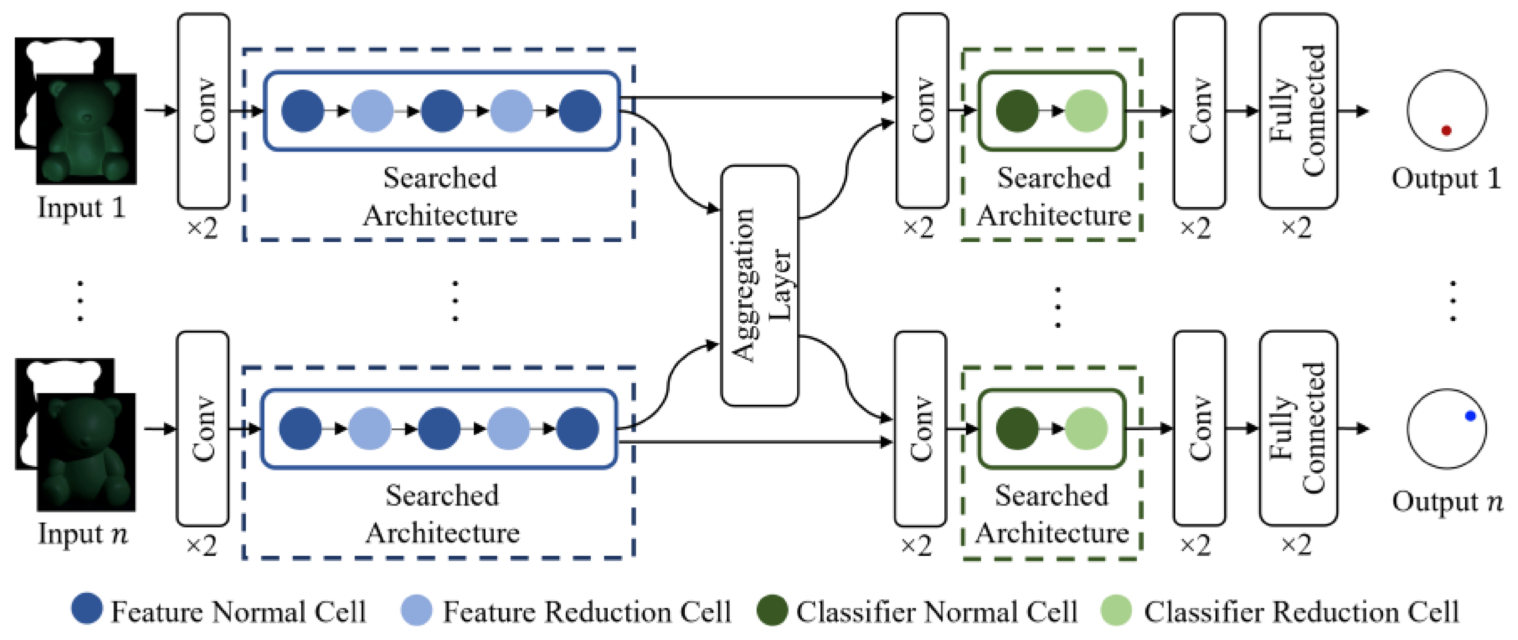
Deep Uncalibrated Photometric Stereo
- Siamese architecture → image specific and global information.
- Two-stage approach → avoid GBR ambiguity.

Neural Architecture Search
- Algorithm: based on DARTS [1].
- Cell structure: 2 input nodes, 1 output node and 4 intermediate nodes.
- Search space: $1\times1,\ \ 3\times3,\ \ 5\times5$ separable convolutions, zero, identity.
- Cell typologies: normal ($stride\ =1$) & reduction ($stride\ =2$).

Pipeline
For both the networks independently:
- Fix modules to ensure physical constraints and feature dimensions (normalization layer, aggregation operation, fully connected layer).
- Choose where, the number and the typology of cells to stack.
- Search for optimal cells.
- Fix optimal architecture and train from scratch.

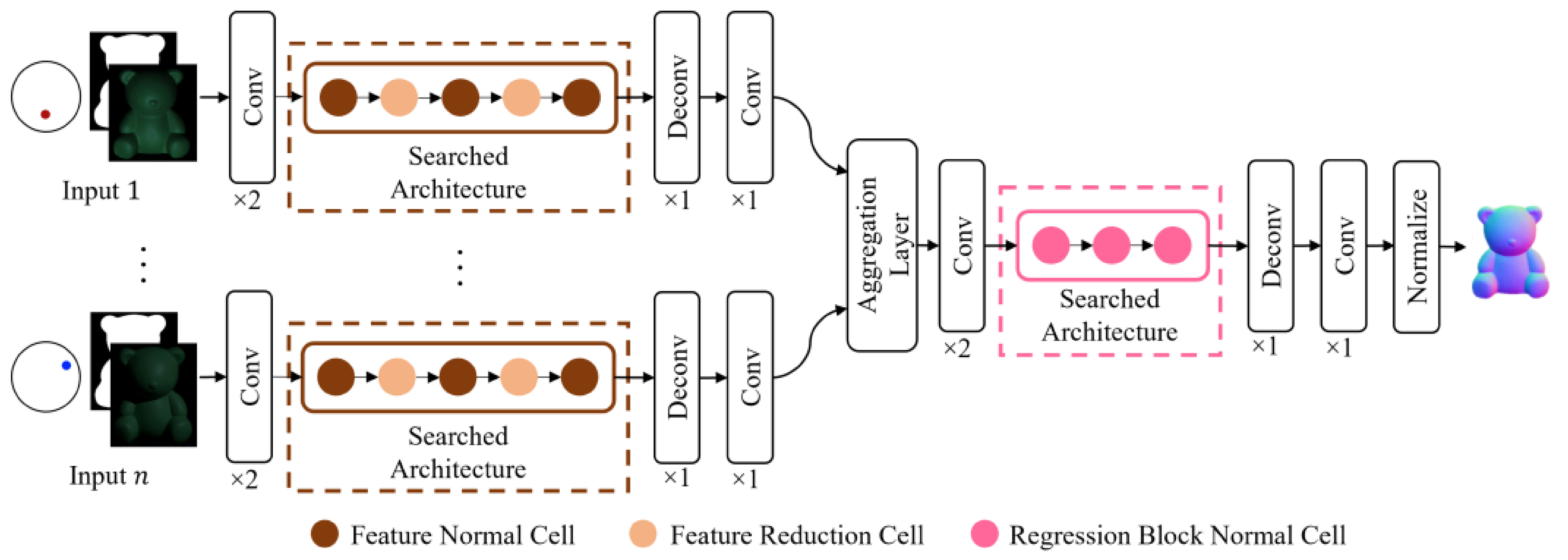
Network Architectures

Light Calibration Network



- Estimation of light direction and intensity for each light source.
- $L_{light} = L_{l_a} + L_{l_e} + L_e$

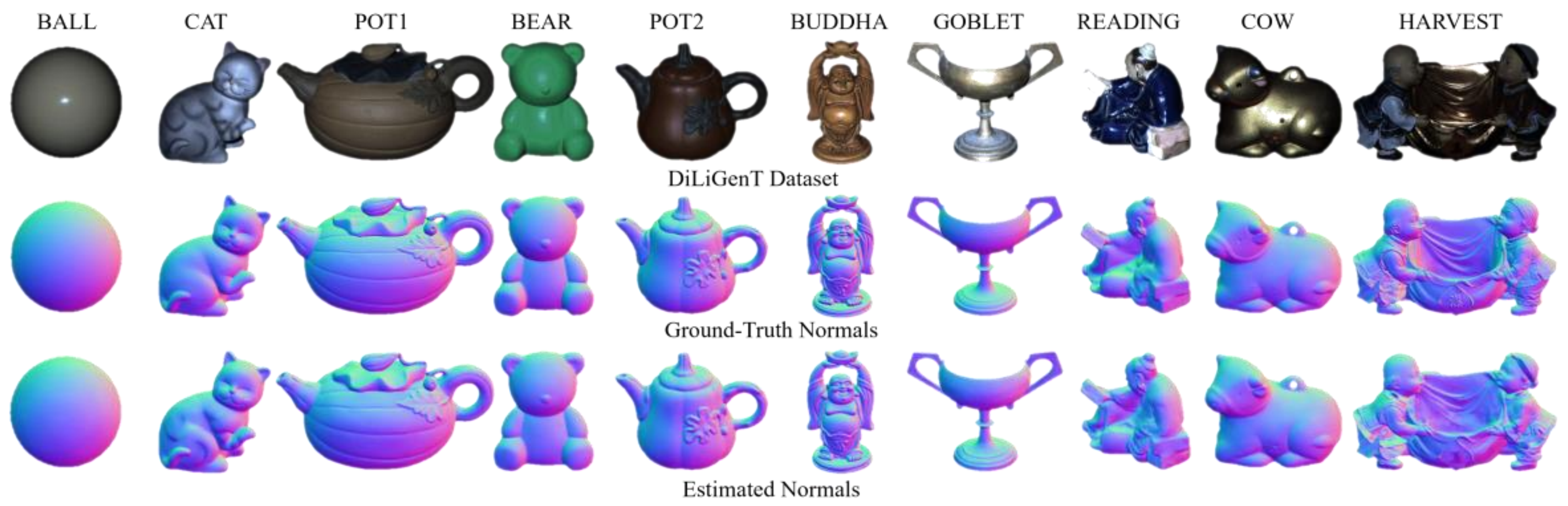
Normal Estimation Network



- Estimation of surface normals.
- $L_n = \frac{1}{hw}\sum_i^{hw}(1 - \tilde{\boldsymbol{n}}_i^T \boldsymbol{n}_i)$

Experiments and Results

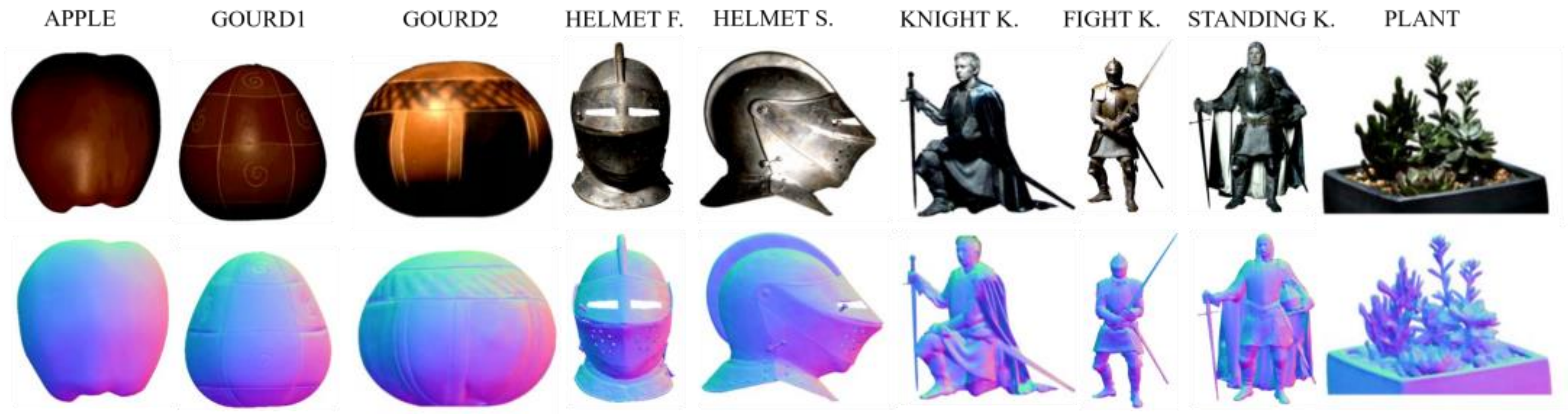
Visual Results on DiLiGenT Dataset (Normals) [2]



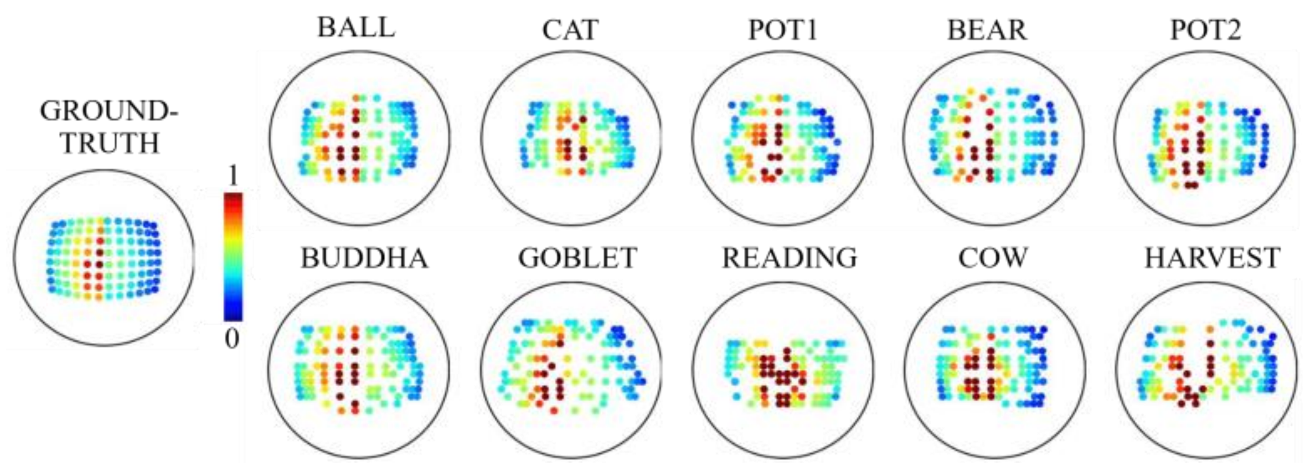
Quantitative Results on DiLiGenT Dataset (Normals) [2]

Method	Params[M]	Ball	Cat	Pot1	Bear	Pot2	Buddha	Goblet	Reading	Cow	Harvest	Average
UPS-FCN [3]	6,1	3,96	12,16	11,13	7,19	11,11	13,06	18,07	20,46	11,84	27,22	13,62
SDPS-Net [4]	6,6	2,77	8,06	8,14	6,89	7,50	8,97	11,91	**14,90**	8,48	17,43	9,51
GCNet+PS-FCN [15]	6,8	**2,50**	**7,90**	**7,20**	5,60	**7,10**	**8,60**	**9,60**	**14,90**	7,80	**16,20**	**8,70**
UNIR [6]	8,1	3,78	7,91	8,75	5,96	10,17	13,14	11,94	18,22	10,85	25,49	11,62
OURS	**4,4**	3,46	8,94	7,76	**5,48**	**7,10**	10,00	9,78	15,02	**6,04**	17,97	9,15

Visual Results on Gourd&Apple Dataset [7] and Light Stage Data Gallery [8]



Visual Results on DiLiGenT Dataset (Light) [2]



Key References:

[1] DARTS [Liu *et al.*, ICLR19]
[2] DiLiGenT [Shi *et al.*, TPAMI16]
[3] UPS-FCN [Chen *et al.*, ECCV18]
[4] SDPS-Net [Chen *et al.*, CVPR19]
[5] GCNet+PS-FCN [Chen *et al.*, ECCV20]
[6] UNIR [Kaya *et al.*, CVPR21]
[7] Gourd&Apple [Alldrin *et al.*, CVPR08]
[8] Light Stage [Einarsson *et al.*, EGSR06]