ISP4ML: The Role of Image Signal Processing In Efficient Deep Learning Vision Systems

Patrick Hansen, Alexey Vilkin, Yury Krustalev, James Imber, Dumidu Talagala,
David Hanwell, Matthew Mattina, and Paul N. Whatmough

Overview

Understand the impact of image signal processing on deep learning vision systems

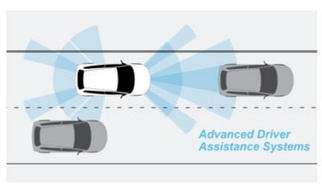
Design an ISP with traditional components for efficient system performance

Application Drivers for Efficient Computer Perception

Augmented Reality

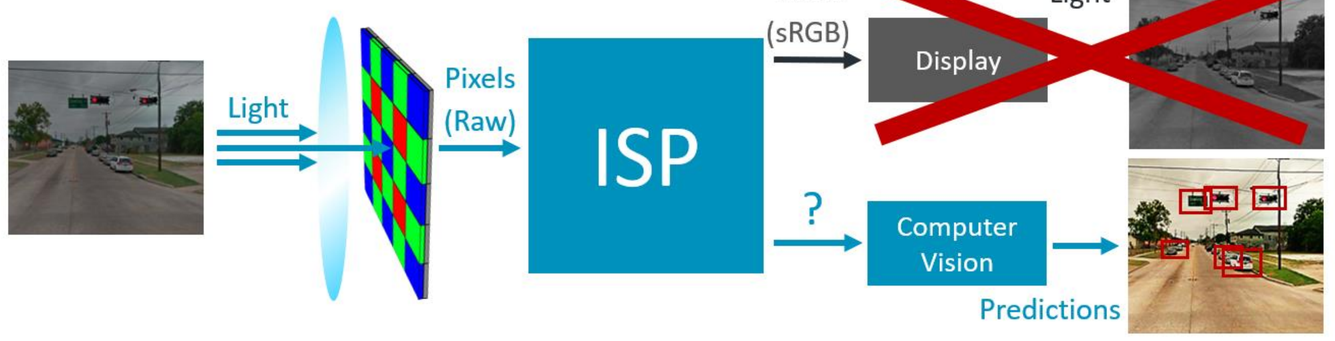
Virtual Reality

Autonomous drones



ADAS

Image Signal Processor Landscape

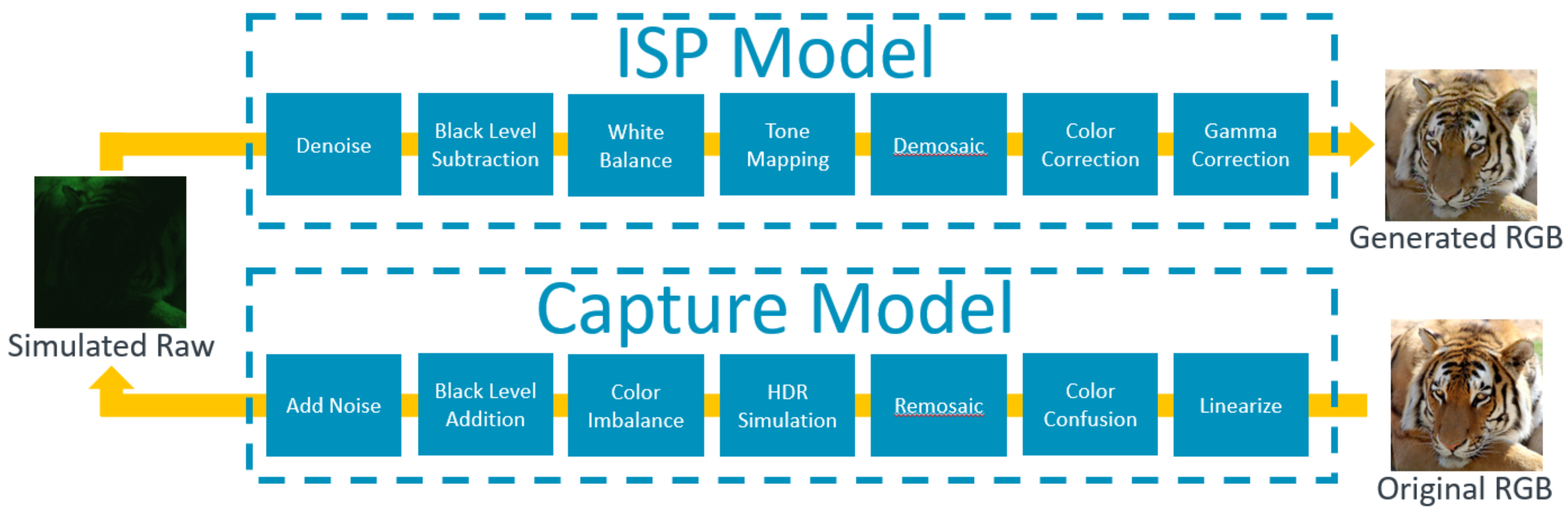


As computer vision algorithms improve, applications will no longer demand image displays

Without the constraint to produce human-viewable images, what should an ISP look like?

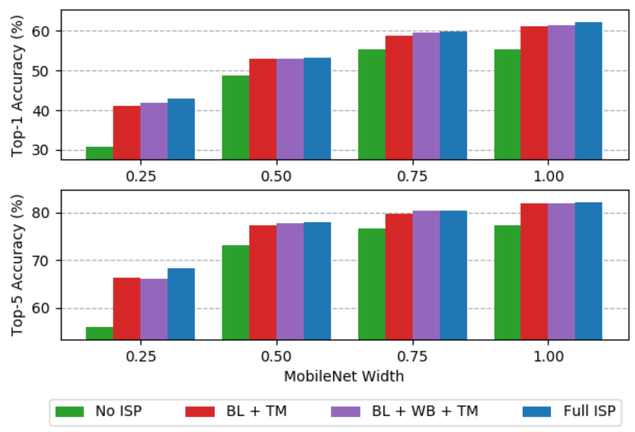
Methodology: Generating Data

- Neural networks dominate computer perception tasks
- No public large-scale raw image dataset useful for training classification or detection networks
- Solution: simulate raw images from standard datasets (e.g. ImageNet)
- Capture model simulates a camera sensor and lens to approximate a raw image from RGB
- ISP software model allows stages to be selectively enabled/disabled to test different ISP configurations

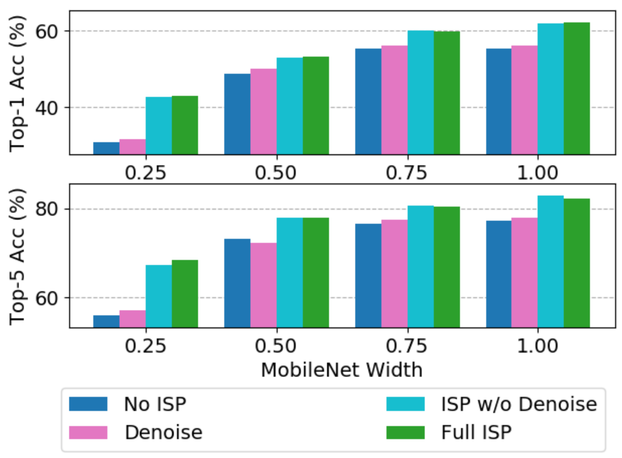


ISP Evaluation

- Trained MobileNets on ImageNet
- Experiments on ResNets support these results
- Tested models on lab-captured test set to validate findings

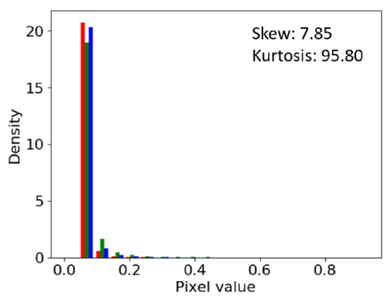


Ablation experiment demonstrates substantial benefit of the ISP. Most significant benefit provided by tone mapping.

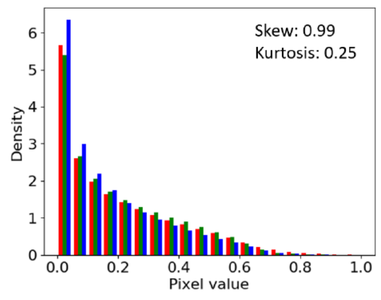


Denoising provides negligible benefit, although it takes up a large percentage of ISP area on silicon

Tone Mapping Investigation



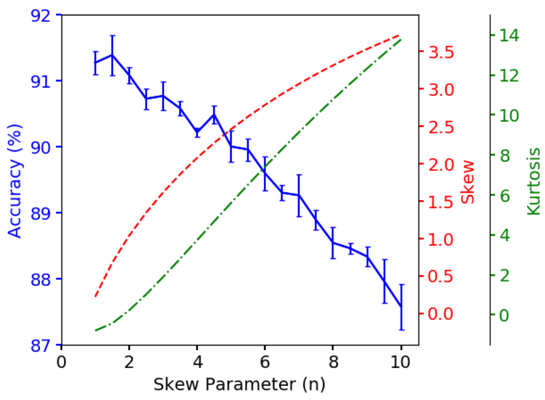
Raw image distributions heavily skewed



Raw image distributions heavily skewed

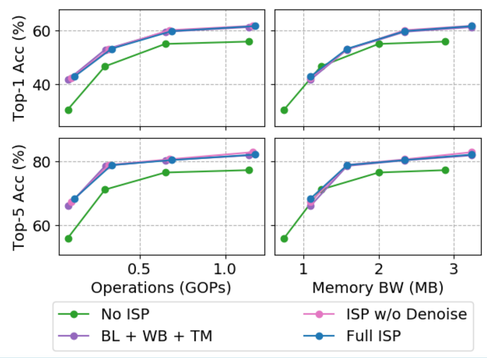
- Trained ResNets on CIFAR-10 with a per-pixel non-uniform mapping:
  $f(x,n)=x^n$
- Produces similar distributions as simulated raw images
- Conclusion: Networks fair to learn effectively from heavily non-uniform data distributions
- Histogram normalization highly valuable

Hypothesis: The disproportionate concentration of raw image pixel values hinders CNN performance. Tone mapping corrects this, resulting in better prediction accuracy



System Level Impact of ISP

- Improvements in accuracy provided by ISP worthwhile because relative cost to CNN inference is miniscule
- ISPs also enable hardware-agnostic inference by enforcing a standard image representation
- Many models are fundamentally incompatible with raw images



Machine Learning Research

ARM Research