AI-ISP:

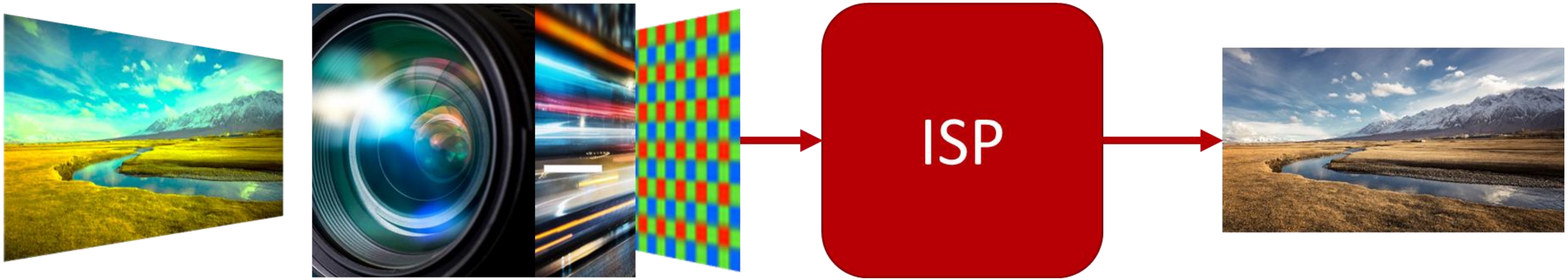
Adding Real–Time AI Functionality to Image Signal Processing with Reduced Memory Footprint and Processing Latency

Mankit Lo

Chief Architect, NPU IP Development

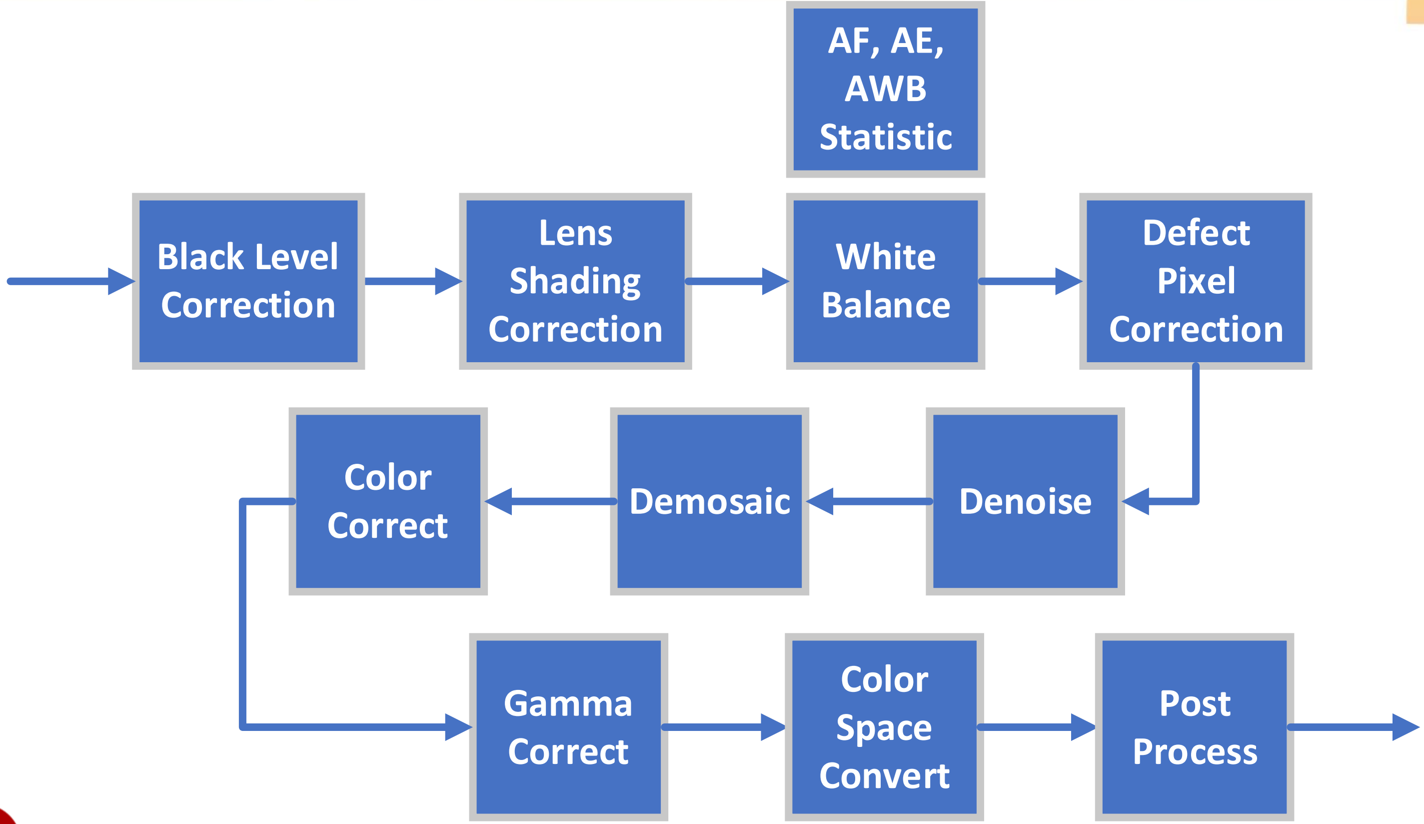
VeriSilicon Inc.

What Is an ISP (Image Signal Processor)?



- Image Signal Processor is responsible for processing the raw image data captured by a camera’s sensor and turning it into a usable image

ISP Traditional Hardware Pipeline



The Growing Body of Research on AI for ISP

- "Learning to See in the Dark": low-light image enhancement

(a) Camera output with ISO 8,000 (b) Camera output with ISO 409,600 (c) Our result from the raw data of (a)

The Growing Body of Research on AI for ISP (cont 2/3)

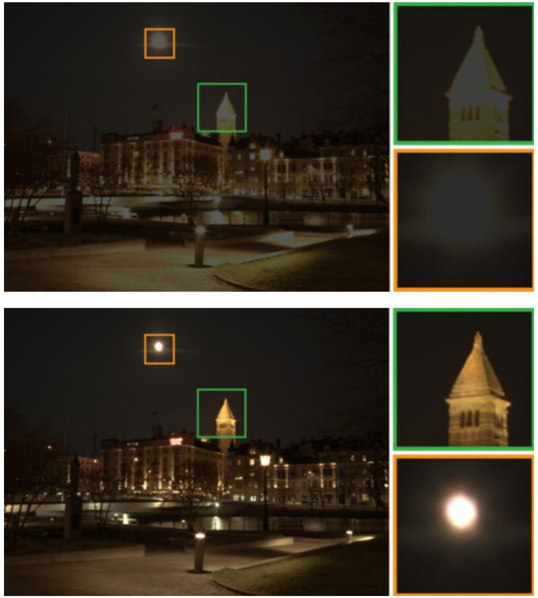
- "Deep Joint Demosaicing and Denoising"

- "Learning Deep Priors for Image Dehazing"

The Growing Body of Research on AI for ISP (cont 3/3)

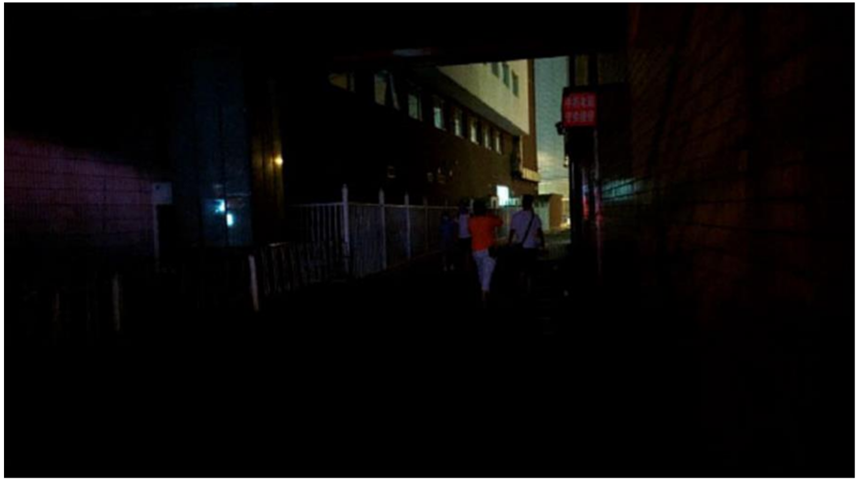
- "HDR image reconstruction from a single exposure using deep CNNs"

- "Merging-ISP: Multi-Exposure High Dynamic Range Image Signal Processing"



6

Noise Reduction with AI-ISP



Without AI-ISP

With AI-ISP

LUX 0.2 lux

VeriSilicon

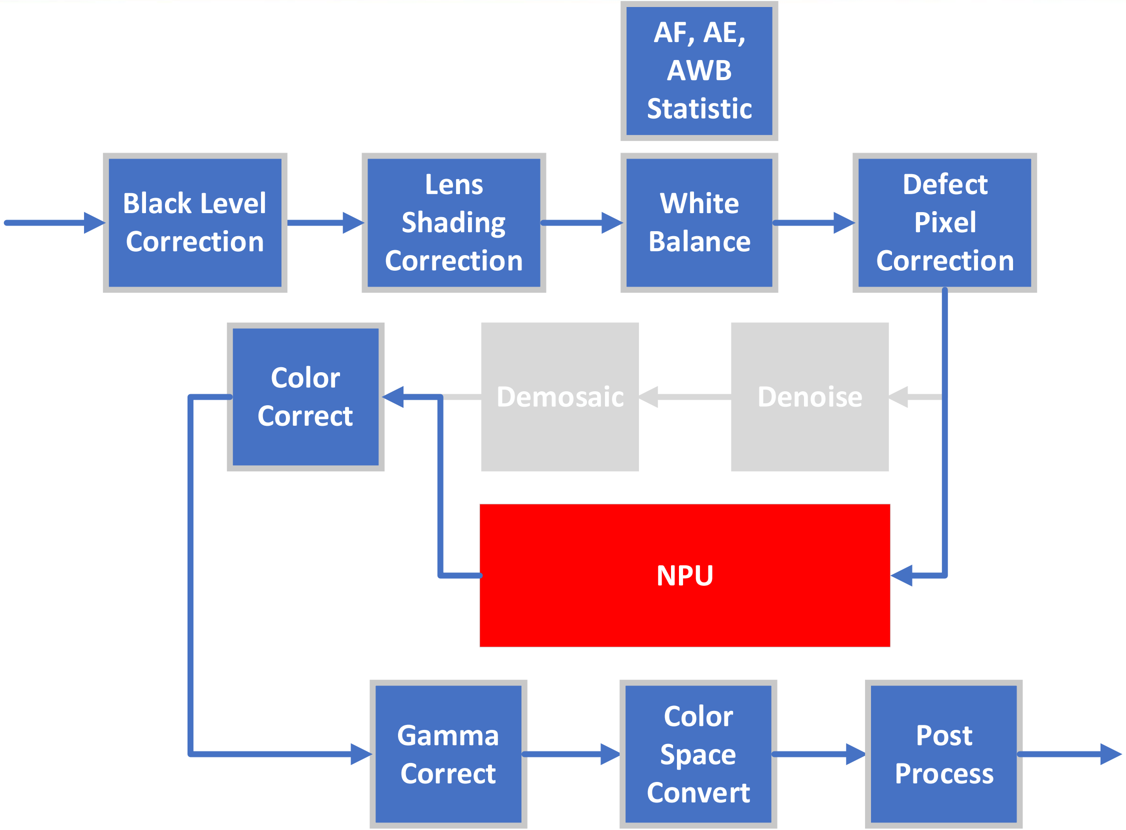
Low Light Image Enhancement for Automotive with AI-ISP

Without AI-ISP

With AI-ISP

Not to Fall Behind

How to Implement an AI ISP?

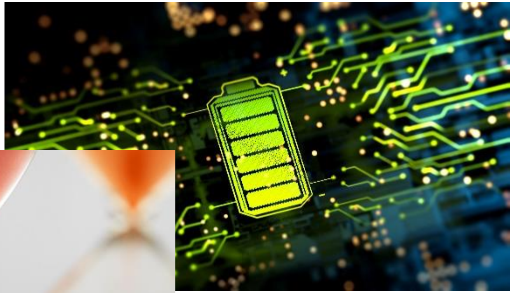


Real-time AI-ISP Challenges

- Programmable: Ever improving algorithms and changing scenarios demand a programmable solution.
- Computation: High resolution sensors and powerful algorithm require high computation power

Real-time AI-ISP Challenges (cont 2/3)

- Memory usage: Conventional image signal processing techniques often require the whole image frame to be stored in memory before processing can begin, which result in high memory usage.
- DDR-SDRAM bandwidth: Requiring the whole image stored means the need of using DDR-SDRAM.
- Power: Accessing DDR-SDRAM requires more power
- Latency: Latency is measured in frames



Real-time AI-ISP Challenges (cont 3/3)

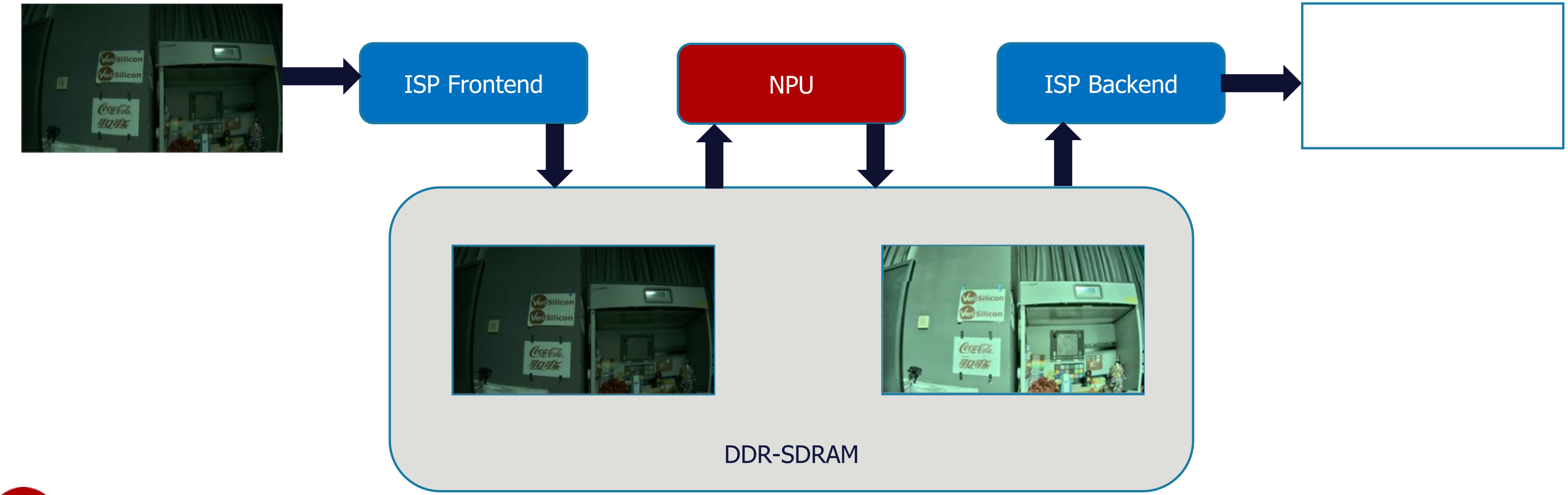
- Task partitioning: The need to distribute tasks to multiple computation units efficiently
- Data sharing: Among multiple computation units
- Synchronization: How to synchronize among multiple computation units

Without AI Processing

- ISP can typically process one raster line at a time

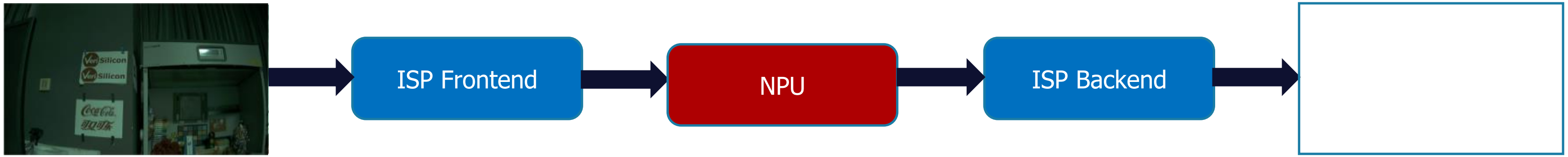
With AI: Loosely Coupled Frame-based

- ISP is forced to store/read whole frame in/from memory



With AI: Tightly Coupled Line-Based

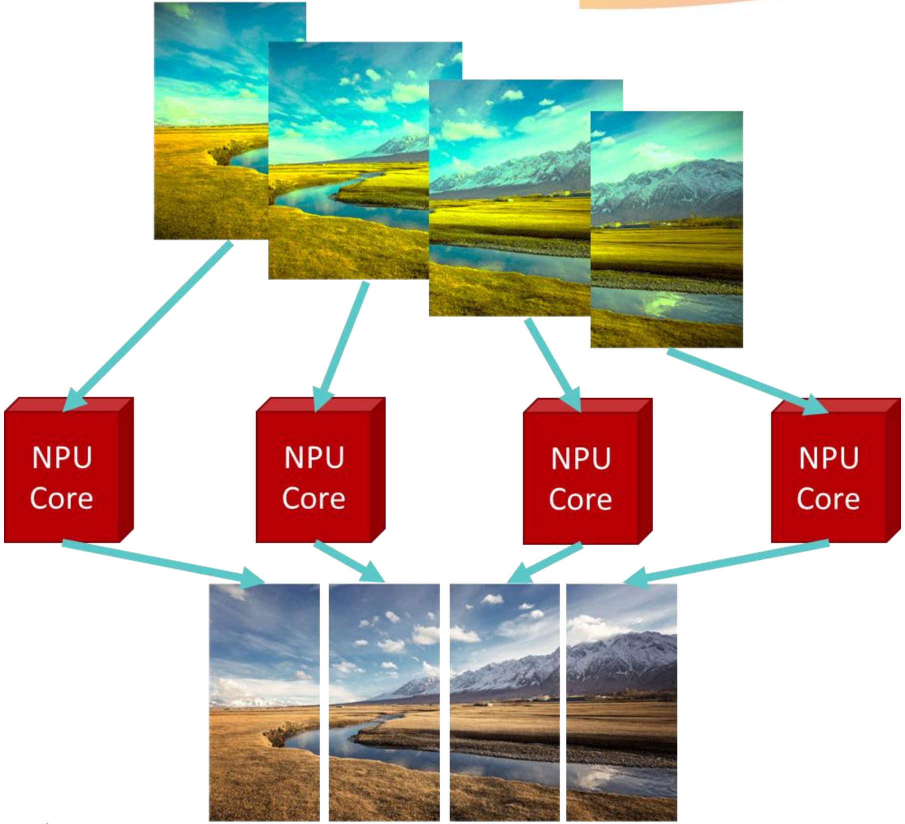
- Need NPU that can do processing line-by-line
- Need ISP and NPU that work closely together



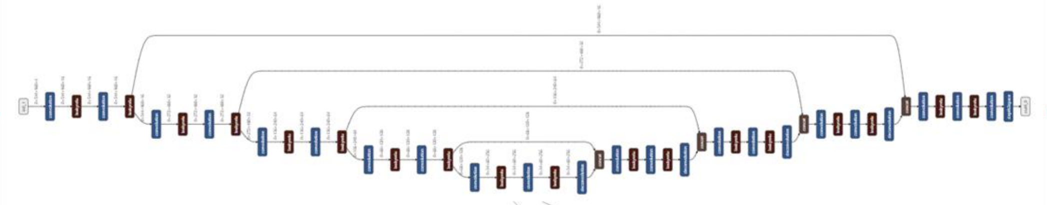
- DDR-less
- Low latency
- Low power

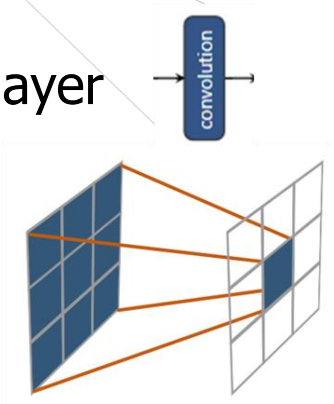
Task Splitting

- High computation and high internal bandwidth requires multiple NPU cores
- Split image in the horizontal dimension



Large Network Receptive Field

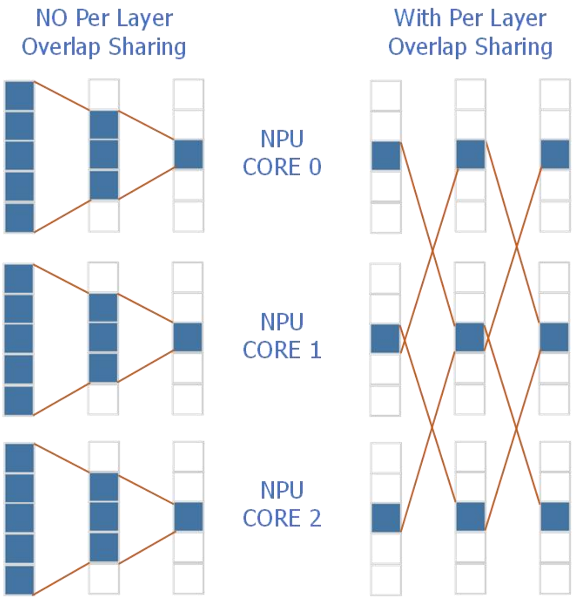


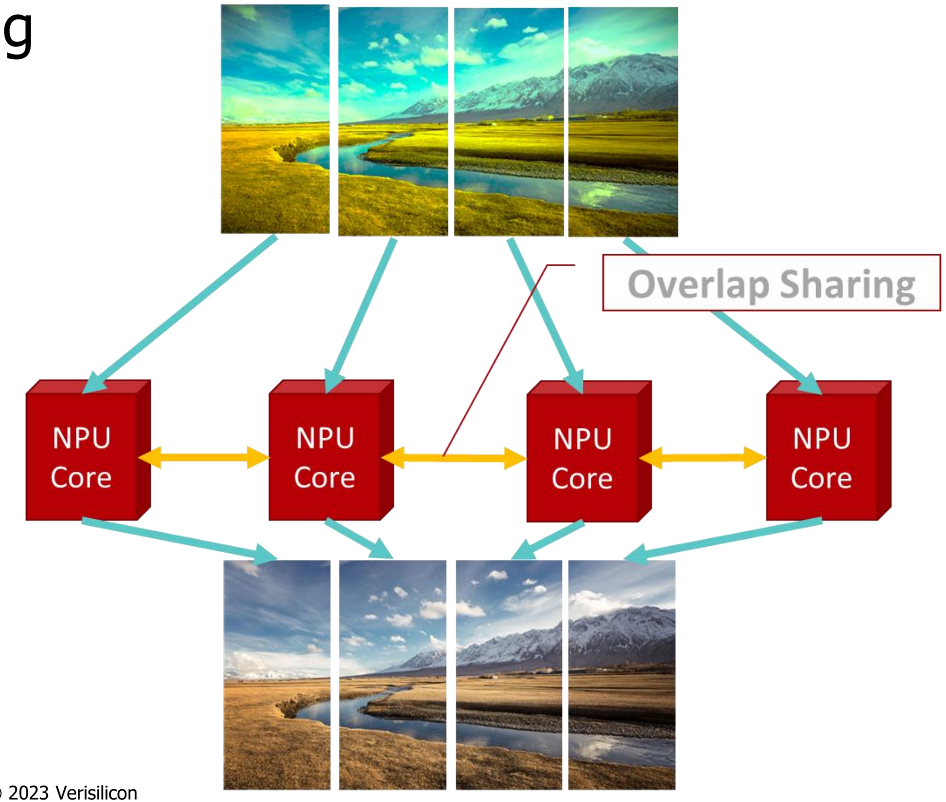


- Even the receptive field is small in each layer
- Network receptive field is very large
- Waste processing Power

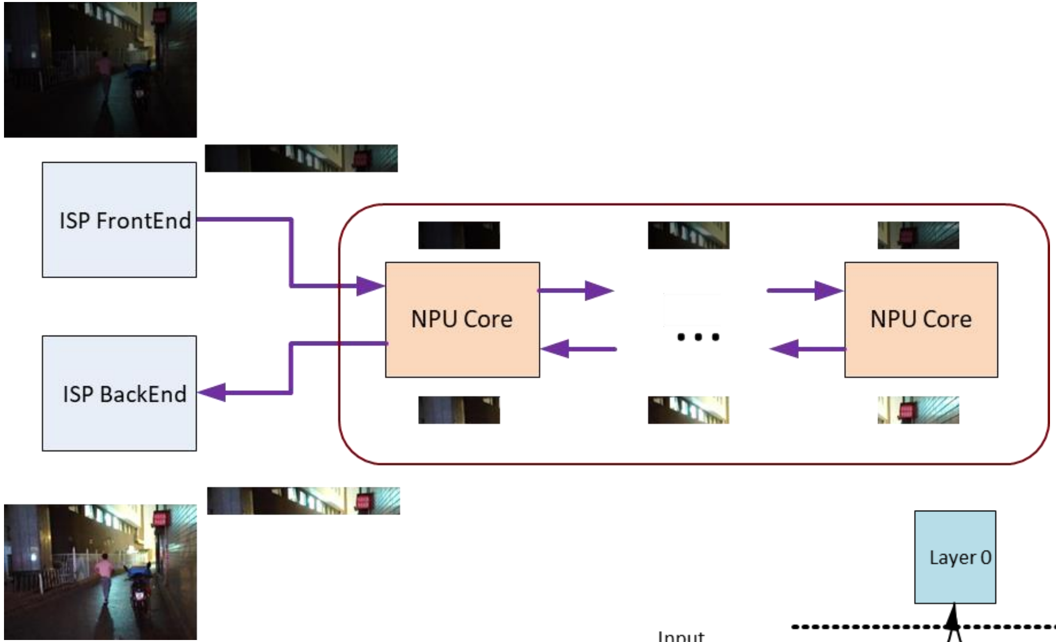
Per Layer Overlap Sharing

- Overlapped data shared among cores in every layer

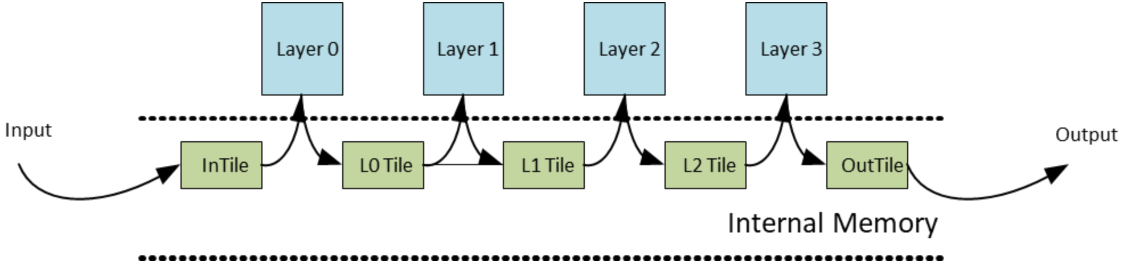




AI-ISP: Pixel Streaming

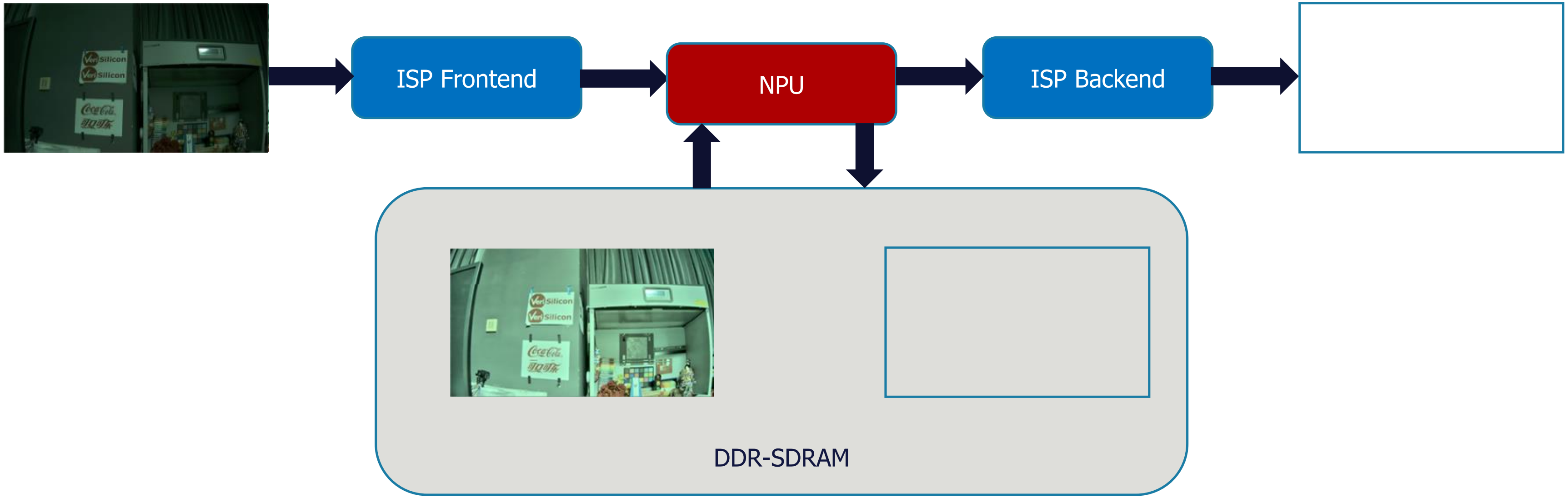


- Line-based latency
- "SplitX" job partition
- Layer fusion for DDR-less



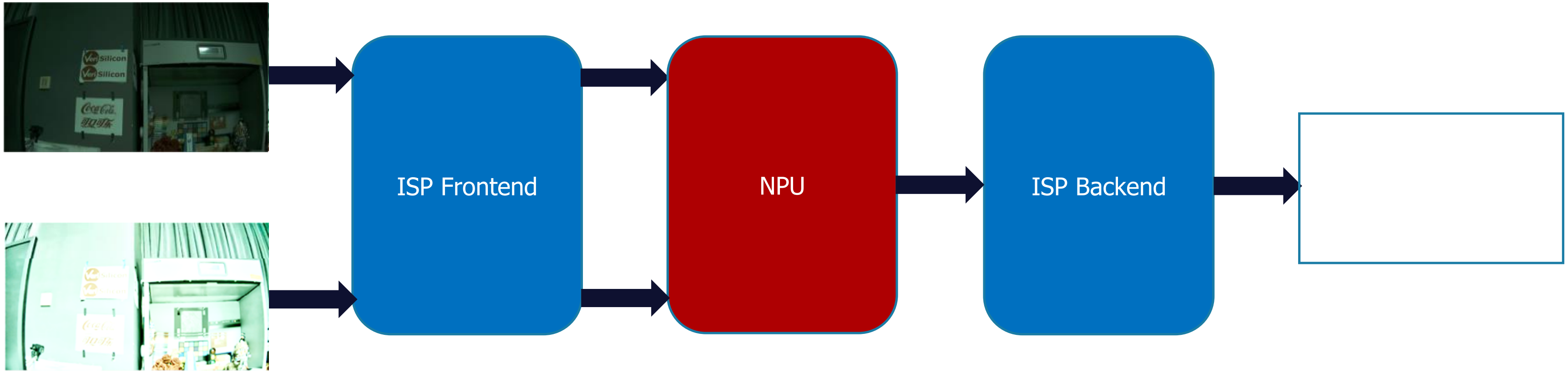
Use Case: AI-3DNR

- Current frame data from ISP streaming through NPU while previous frame reference is fetched from DDR-SDRAM

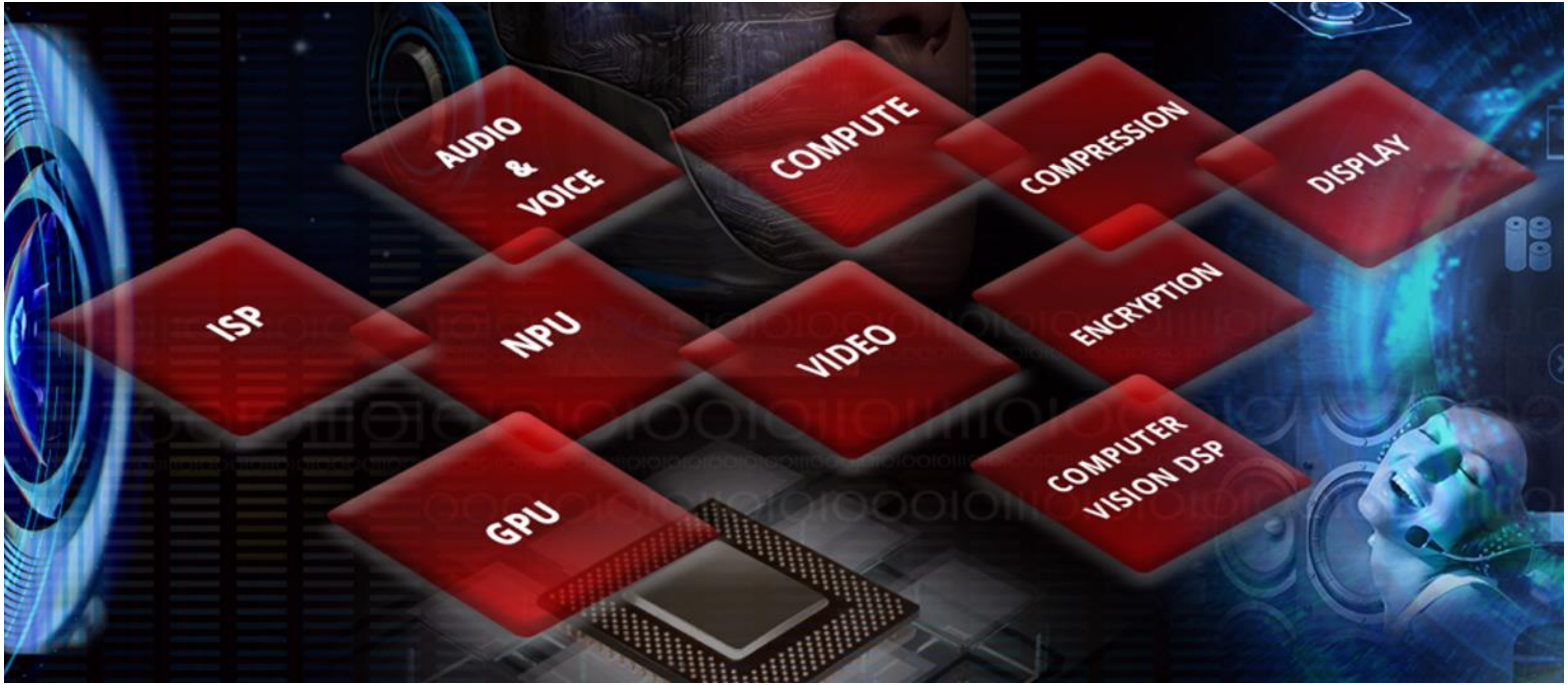


Use Case: AI-HDR

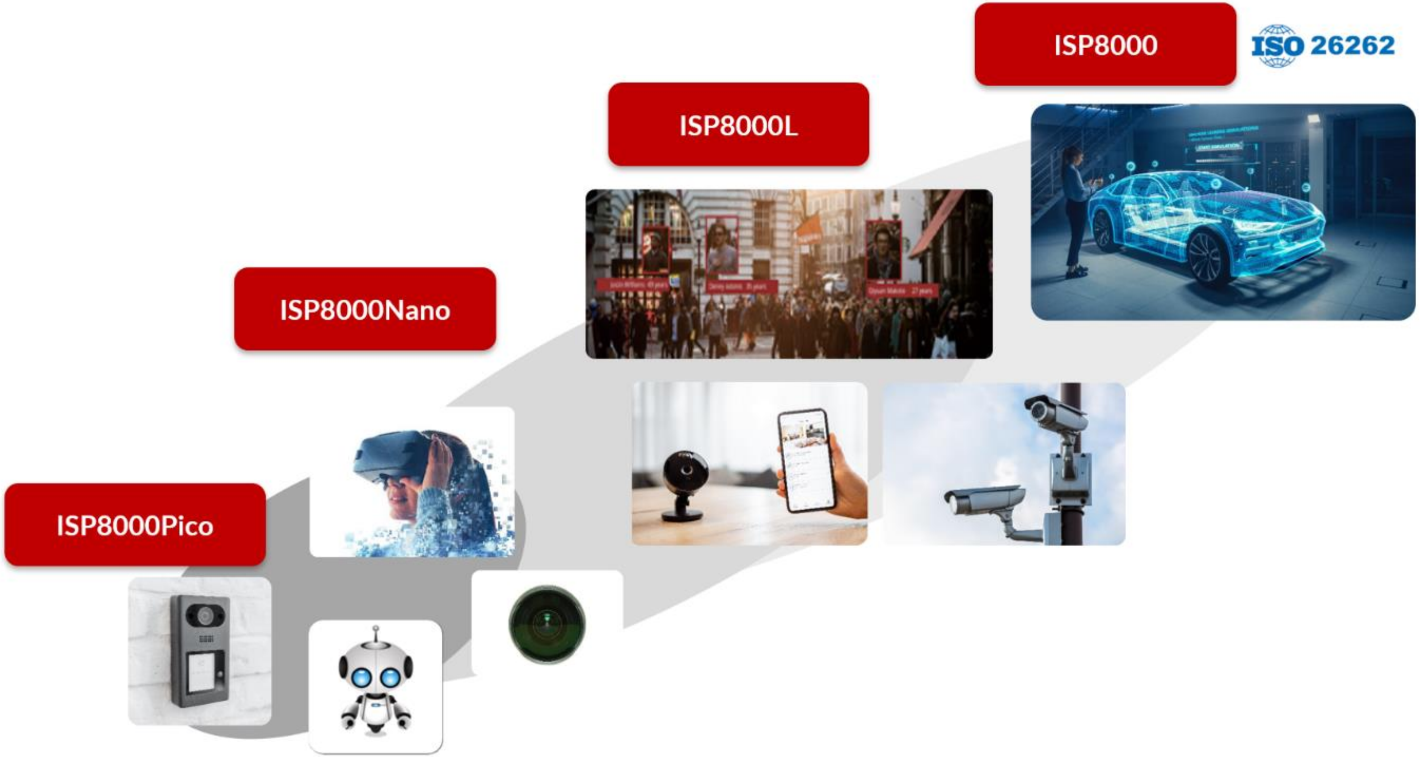
- Multiple sensor inputs stream through NPU



VeriSilicon IPs



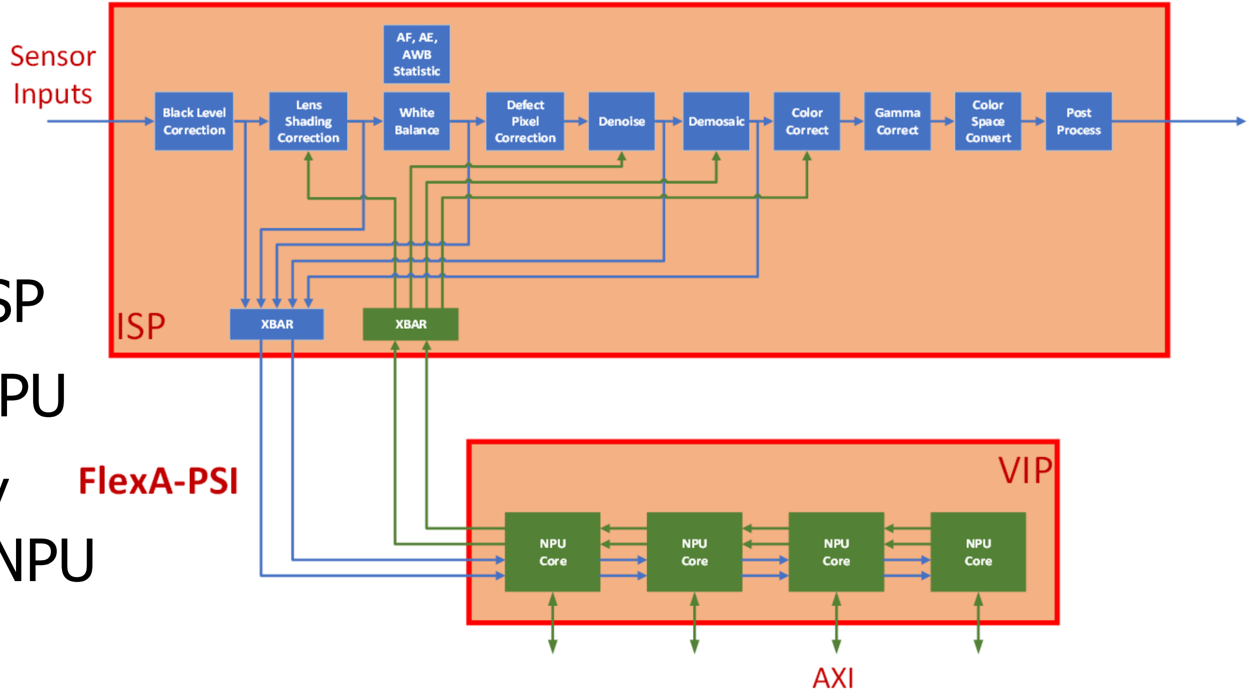
VeriSilicon Vivante® ISP IP



VeriSilicon Vivante® NPU IP

VeriSilicon AI-ISP solution

- VeriSilicon Vivante ISP
- VeriSilicon Vivante NPU
- FlexA-PSI seamlessly connecting ISP and NPU
- AI algorithms



AI Contents

AI-ISP

Rich Input / Output configuration

- Support inserting AI-NR module
- Support inserting AI-Demosaic module
- ……

Face Scenarios

- Supporting Skin detection
- Skin area different process
- ……

Rich ROI configuration

- Support inserting AI-Detect
- Traffic light, Body, Vehicle, bad Weather
- AWB, AE optimization for objects
- ……

AI inside

Scene Scenarios

- Different Scene Color Scenario
- Different Scene Sharpness Scenario
- ……

Unified Software Architecture, User API, Test Case and 3rd Algorithm Support……

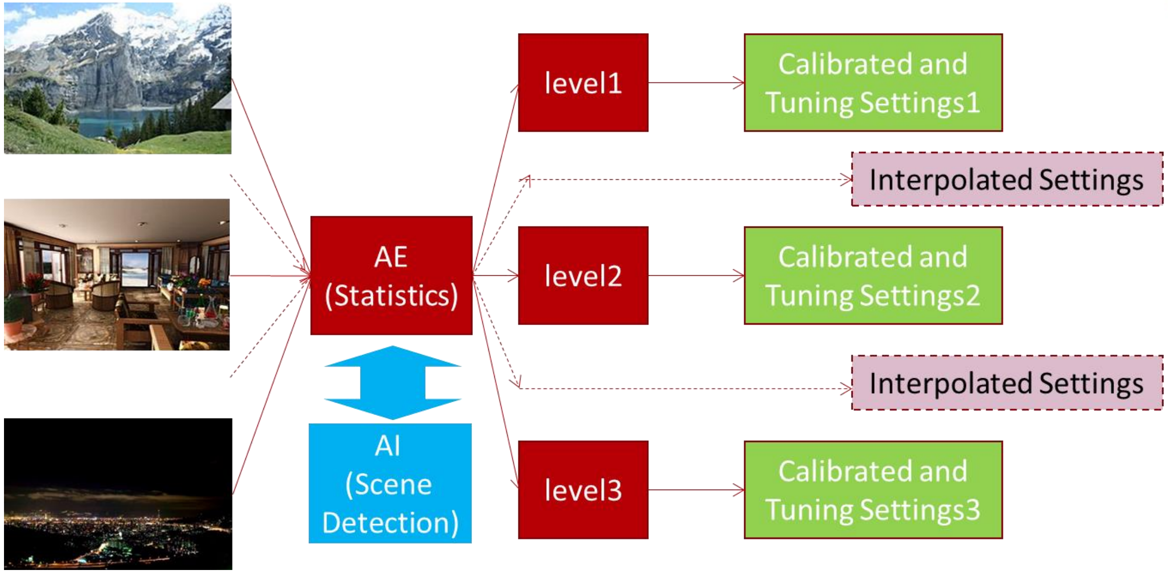
AI-Face Detect – 3A

- Auto exposure
  - based on face detection
- Skin tone style fine tune
  - based on face detection

Gain = 8.6

Gain = 16
Higher Exposure based

AI-Scene Categorization



- AI to distinguish all scenes into different categories with higher accuracy
- Interpolate between different scene settings based on scene probability

AI-Detect (Traffice Light, Bad Visibility)

- AI traffic object detection
  - Traffic light, vehicle
  - Pedestrian, bad weather
- Identify region of interest
- Adjust image
  - 3A
  - Global tone mapping
  - Wide dynamic range

Without ROI

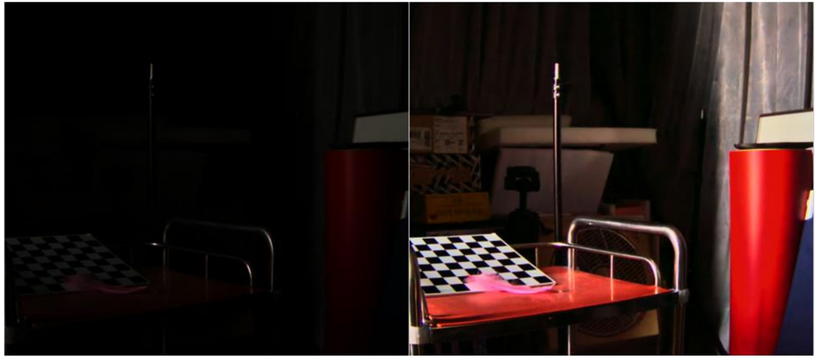
With ROI

Clear Shape of light

Bad weather visibility

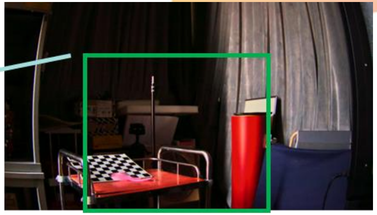
AI-BR 1 Lux – Low-light Condition

Max maintenance of details

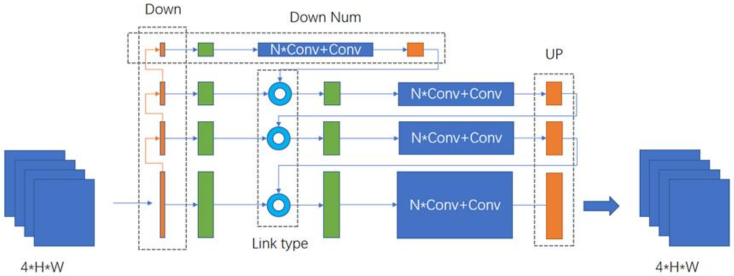


(Left) : before Brighten

(Right) : after Brighten



- Lightweight models
- Raw In / Raw Out



Conclusion

- **ISP**
  - Flexible in-pipeline replacement
- **NPU**
  - Programmability
  - Performance
  - Line-based processing
  - Layer level overlap sharing
- **ISP/NPU communications**
  - Synchronization
  - Line-based data transfer
  - Multiple channels
- **DDR-less or low memory footprint**
- **Low latency**
- **Low power**

Resources

VeriSilicon Vivante ISP IP

https://www.verisilicon.com/en/IPPortfolio/VivanteISPIP

VeriSilicon Vivante NPU IP

https://www.verisilicon.com/en/IPPortfolio/VivanteNPUIP

2023 Embedded Vision Summit

- Visit VeriSilicon's booth to speak with technology experts and watching exciting demos

References (1/2)

- C. Chen, Q. Chen, J. Xu, and V. Koltun (2018). Learning to See in the Dark. https://doi.org/10.48550/arXiv.1805.01934
- M. Charbi, C. Chaurasia, S. Paris, and F. Durand (2016). Deep Joint Demosaicking and Denoising. http://dx.doi.org/10.1145/2980179.2982399
- Y. Liu, J. Pan, J. Ren, and Z. Su (2019). Learning Deep Priors for Image Dehazing. https://doi.org/10.1109/ICCV.2019.00258

References (2/2)

- G. EILERTSEN, J. KRONANDER, G. DENES, R. K. MANTIUK, and J. UNGER (2017). HDR image reconstruction from a single exposure using deep CNNs. https://doi.org/10.48550/arXiv.1710.07480
- P. Chaudhari, F. Schirrmacher, A. Maier, C. Riess, and T. Köhler (2021). Merging-ISP: Multi-Exposure High Dynamic Range Image Signal Processing. https://doi.org/10.48550/arXiv.1911.04762