Selecting Image Sensors for Embedded Vision Applications: Three Case Studies

Monica Houston

Manager, AI / ML Team

Avnet

What does an image sensor do?

1. Camera shutter opens

2. Photon wells ("pixels") store photons

3. Analog front-end converts analog to digital

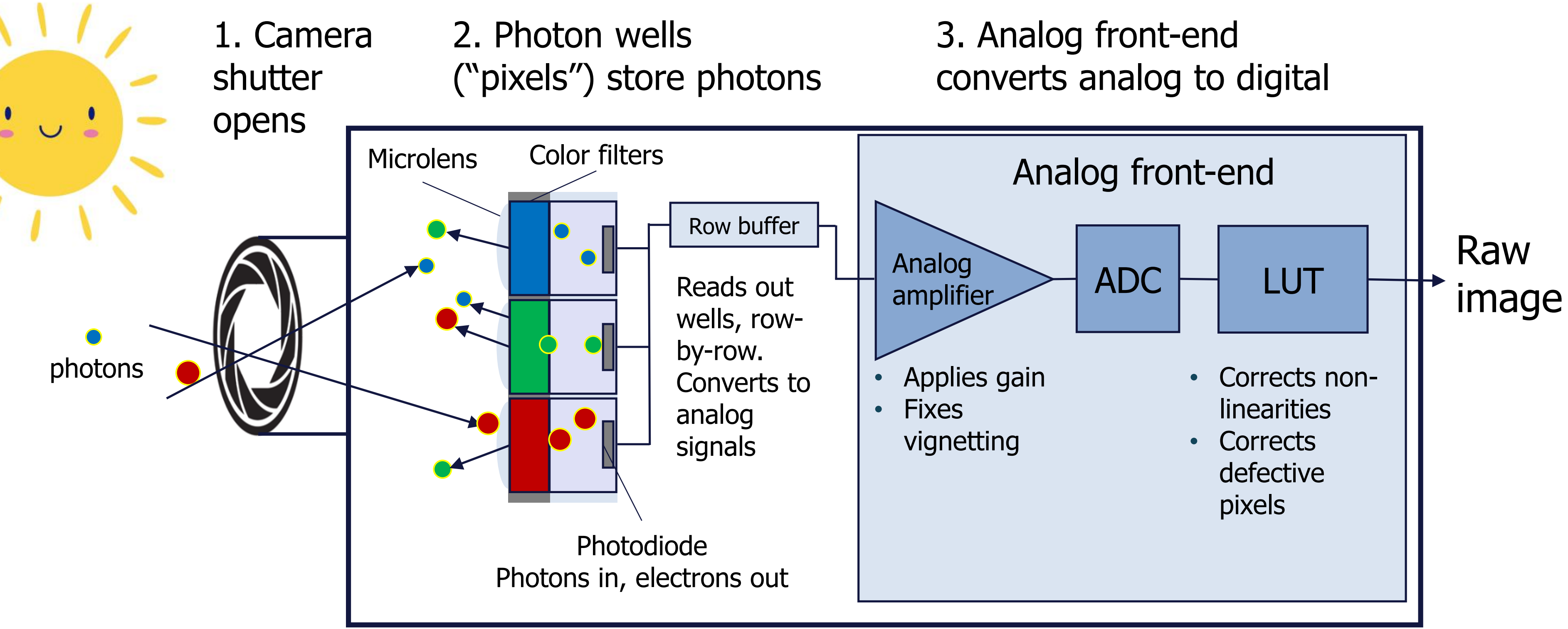
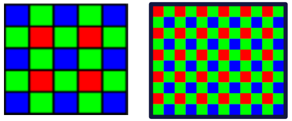
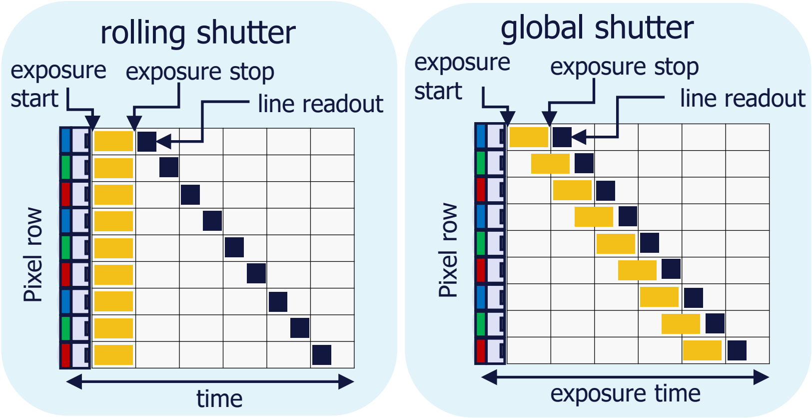


Image sensor specifications

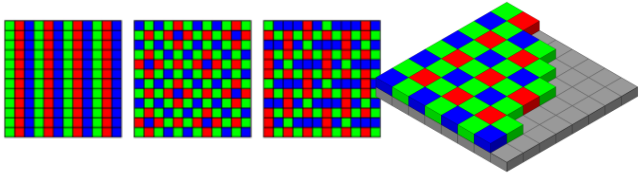
Resolution



Shutter type and speed



Color Filter Arrays (CFAs)



Signal to Noise Ratio (SNR) & Dynamic Range

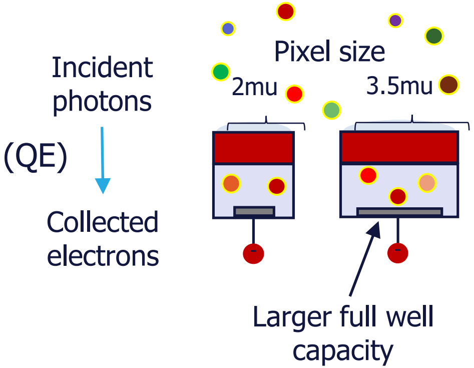
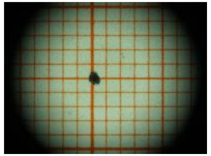


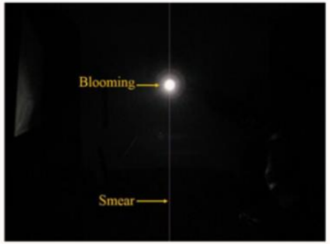
Image sensor specifications (continued)

Other specs:

- Near Infrared Optimized (NIR)
- Chief Ray Angle (CRA)
- Field of View (FOV)
- Defective Pixels
- Pixel size / full well

vignetting

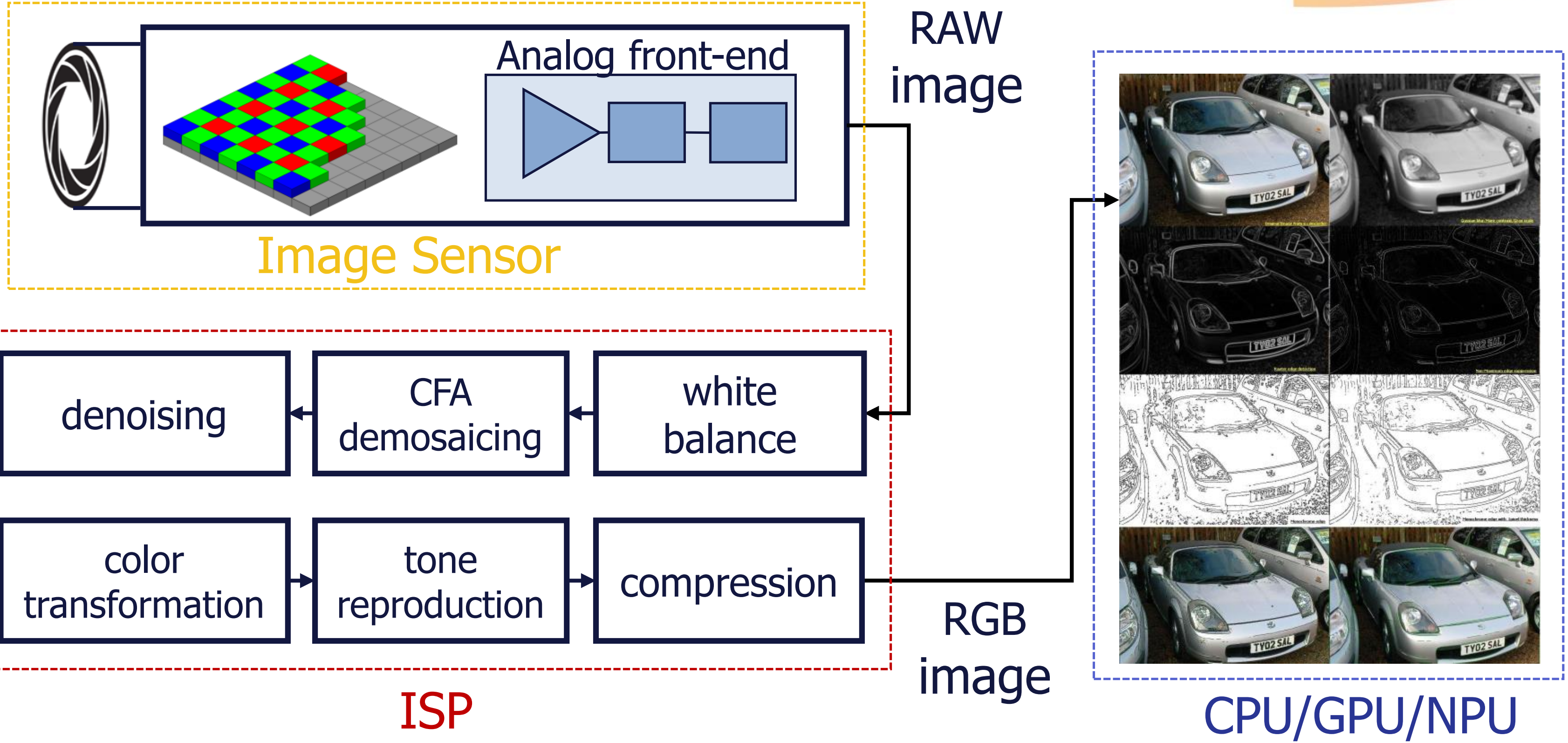




Other things to consider:

- Image Signal Processor (ISP)
- Interfaces
- Compression
- Sensor size
- Power consumption (heat)
- Lens
- Price
- Lead time
- Support

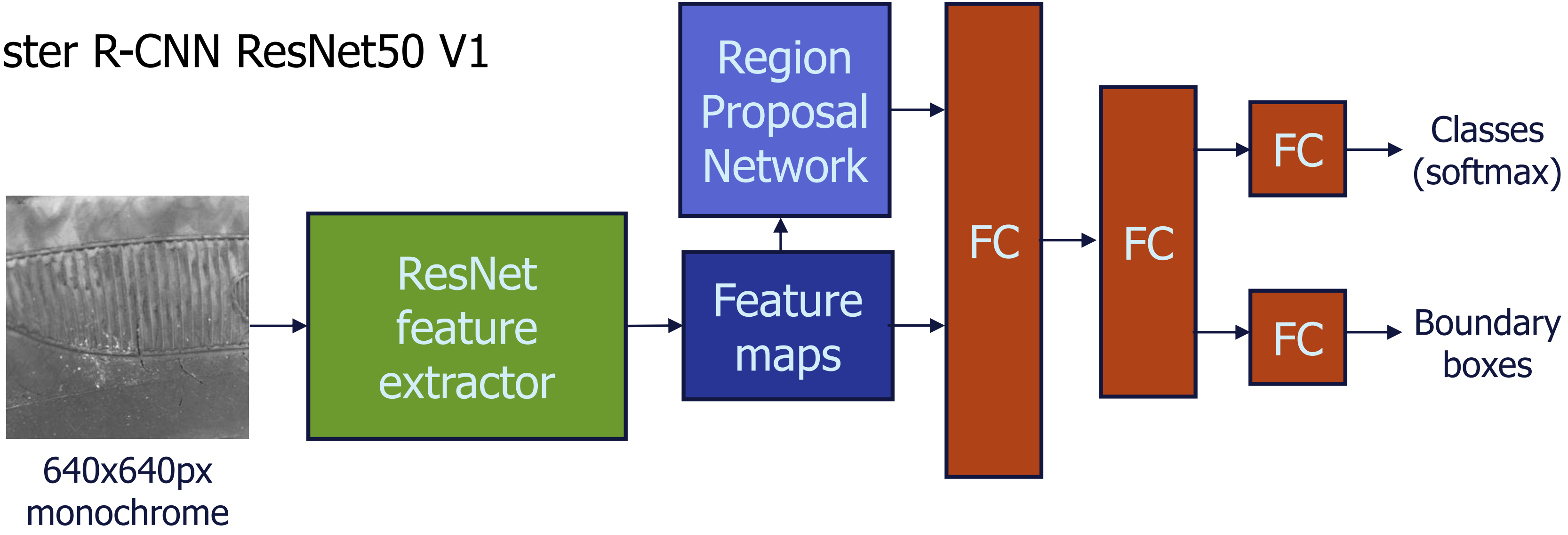
Why choosing the right image sensor is important



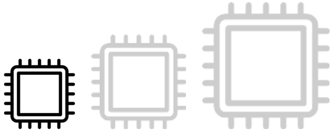
Case Study #1:
Tire Defect Detection

The model

Faster R-CNN ResNet50 V1



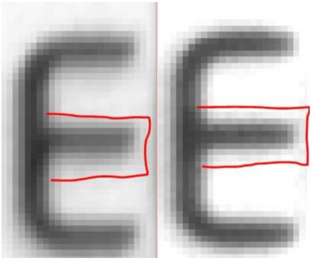
Computational requirement:



Monochrome sensors vs color sensors

Monochrome pros:

- Higher quantum efficiency
- Better features:
  - Reduced noise
  - Improved contrast
- Increased speed
- Lower cost



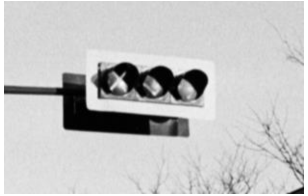
color image | monochrome image

Can you just use a color image sensor and convert to greyscale?

Some models do better with color

oxidation detection

clock or stove?



navigation

tiger or snow leopard?

Case study #1 summary

embedded VISION SUMMIT

When to consider using a monochrome image sensor:

- Model accuracy doesn't depend on color information
- Low-light conditions
- Intensity-based features such as texture and edges are important
  - Bar code scanning, OCR, defect detection
- Low cost requirements
- Limited computational ability

Assess:

- Is your model architecture designed for monochrome input?
- Is color important at any step in the application?
  - E.G. color might not be important for your object detection but is relevant for your classification

Case Study #2:
License Plate Recognition

The models

1920×1080px

YOLOv2

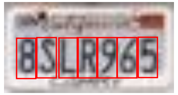
Vehicle detection

YOLOv3

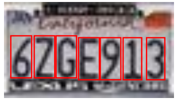
Plate detection

YOLO-Net & Tessaract OCR

Text recognition

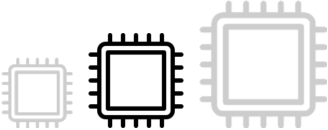


8SLR965



6ZGE913

Computational requirement:



Shutter type and speed

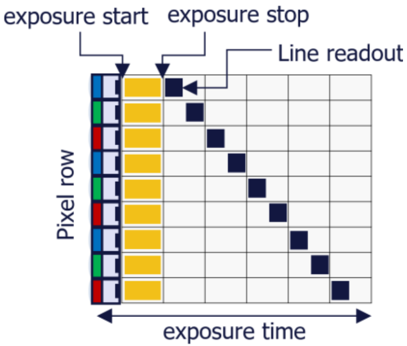
Rolling shutter pros:

- Lower noise
- Lower cost
- Typically lower power

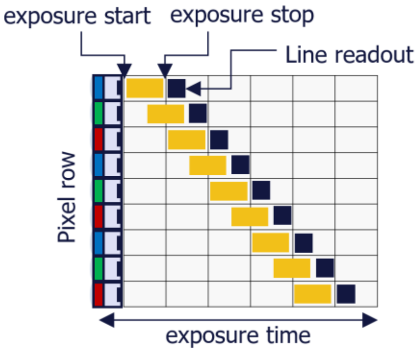
Global shutter pros:

- Higher frame rates
- No motion distortion
- Better low-light quality
- Larger FOV

Global shutter

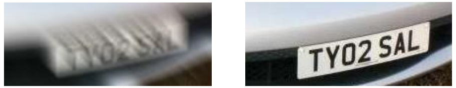


Rolling shutter



High frame rates

- 1/60, 1/120, 1/240…
- Increased accuracy
  - Enable more detections
  - More likely to capture at least one clear image
- Reduced motion blur

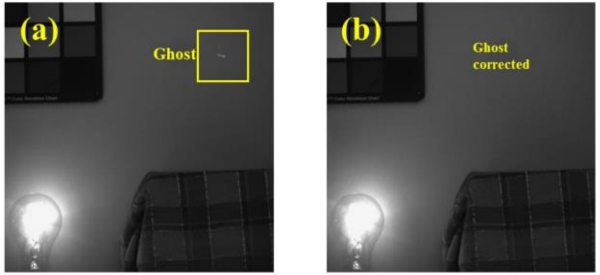


How fast must frame rate be to capture a car going 100 mph?

Variable light conditions

Measurement	Definition	Good range
High dynamic range (HDR)	Includes temporal dark noise	120 dB – 140 dB
Signal to noise ratio (SNR)	Includes temporal dark noise, shot noise	40+ dB
ADC resolution (luminance)	Intensity of light captured	12-bit+
Near-infrared sensitivity (NIR)	Sensitivity to NIR wavelengths	650 nm to 2500 nm
Shutter efficiency ratio	Time all pixels exposed to light / shutter speed	90%+

Higher **shutter efficiency ratio** =
less **ghosting** (light or motion artifacts)
Rejects undesired light (outdoors)



Case study #2 summary

For moving objects, consider:

- Global shutter
- Frame rate

For varied lighting conditions, including glare, overcast, strobing, and night, consider:

- HDR
- SNR
- ADC resolution
- NIR

Assess:

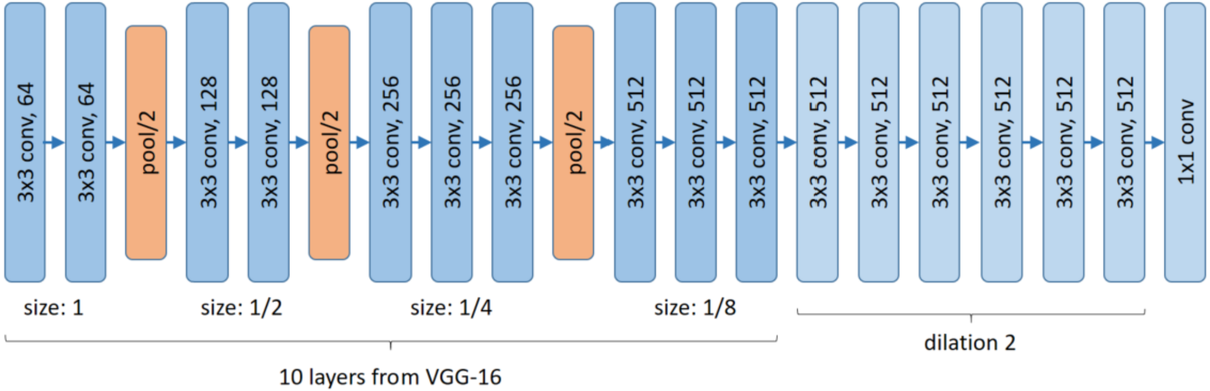
- Is shutter speed fast enough to prevent motion blur but slow enough to let in enough light?
- Power requirements for HDR and higher frame rates?
- Color or monochrome sensor?

Case Study #3: Crowd Counting

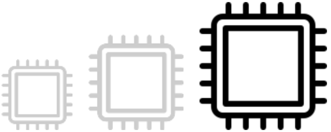
The model

2048×1080 px

CSRNet



Computational requirement:



High resolution images

How much resolution do you need?

Spatial resolution =
(feature size) / (minimum resolution to find an edge)

15cm* / 4px* = 3.75cm px

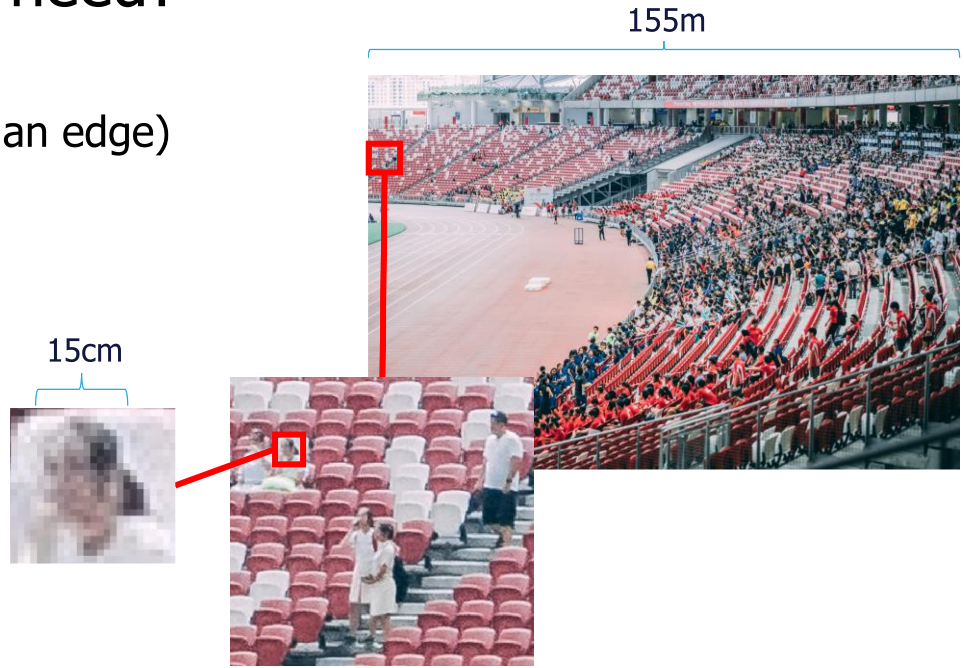
*average diameter of a human head is 15cm

*minimum of 3-4px in order to find an edge

Image sensor resolution =
(FOV) / (spatial resolution)

155m / 3.75cm px = 4,133px* wide

*This image is 5758px wide



AVNET

Case study #3 summary

When to consider high resolution image sensors:

- Large field of view is needed
  - Multiple cameras are not feasible
- Image pipeline requires cropping, resizing, or zooming
- Computation power, memory, and data interfaces are adequate

Assess:

- How much resolution can camera interfaces handle at required speed?
- Are you able to collect or find a dataset with the desired resolution?

Tips and Tricks for high resolution input:

- Downsampling
- Patch-based training
- Try monochrome input
- Trade-off with frame rate
- Trade-off with pixel size

Conclusion

1. Different embedded vision applications require different image sensors
2. Choice of image sensor has a significant impact on the accuracy, power, and computational cost of your vision system
3. Choosing the right image sensor will require trade-offs

Consider:

- What features of the input image are most important?
    - Edges, textures, hue, etc.
- Lighting conditions
- Motion and speed (of subject or camera)
- Size of required field of view
- Size of ROI or detail

Image sensor resources

Specifications

- Understanding the In-Camera Image Processing Pipeline for Computer Vision
- Operating principle and features of CMOS sensors
- Image Sensor Terminology

Monochrome vs Color

- How much resolution do I lose using a color industrial camera in a mono mode?
- Does Colour Really Matter? Evaluation via Object Classification

Shutter Type and Speed

- Real-Time Camera Tracking: When is High Frame-Rate Best?
- Global Shutter Efficiency Improvement to >100dB in Advanced Global Shutter Imager

Resolution

- Imaging Electronics 101: Camera Resolution for Improved Imaging System Performance
- Efficient High-Resolution Deep Learning: A Survey

More information on models and techniques

Defect Detection

- Faster R-CNN
- Tire Defect Detection Using Fully Convolutional Network

License Plate Recognition

- License Plate Recognition in Urban Road Based on Vehicle Tracking and Result Integration
- License Plate Detection and Recognition in Unconstrained Scenarios

Crowd Counting

- CSRNet
- NWPU-Crowd: A Large-Scale Benchmark for Crowd Counting and Localization
- To choose or to fuse? scale selection for crowd counting
- Efficient High-Resolution Deep Learning: A Survey