2020 embedded VISION summit®

Using an ISP for Real-time Data Augmentation

Timofey Uvarov
Pony.ai
Sept 2020

Contents

Introduction to Pony.ai

Exposure in photography and computer vision

Experiments with exposure control

Conclusions and problem formulation

Proposed solution with automotive ISP

Rererences

Pony.ai

Pony.ai – global autonomous technology leader

HIGHLIGHTS

100+ vehicles

500km²+ operational coverage

3.0+ million autonomous kilometers driven

FREMONT

coverage ~100 km²

R&D hub with commuter pilot with City of Fremont

IRVINE

coverage ~50 km²

Robotaxi pilot serving the general public with Hyundai

GUANGZHOU

coverage ~200 km²

Robotaxi pilot launched Q4 2018

BEIJING

coverage ~100 km²

Actively testing in two urban locations under T3 permit

SHANGHAI

coverage ~30 km²

To launch large-scale Robotaxi fleet operations

Existing sites

New launches in 2020

Core technology: vehicle-agnostic “virtual driver”

In-depth partnerships *with camera industry leaders to build next generation* ***camera for machine vision***

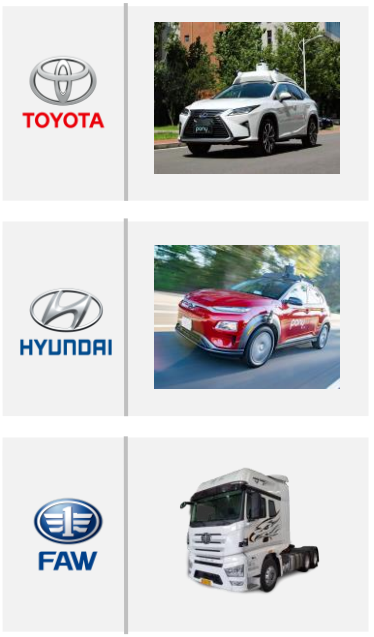
Product layer — End Customer Interface

Sensors & Hardware
- LiDAR
- Camera
- Radar
- GNSS/ IMU
- Compute

AD Software Stack
- Sensor fusion
- Perception
- Prediction
- Planning
- Control
- HD Map Integration & Localization
- Data & Infrastructure
- Vehicle Interface

Vehicle

Select Vehicle Platforms



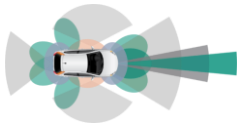
Pony.ai	Developed in-house by Pony.ai
Partner	Developed by partner selected by Pony.ai

Collaboration between Pony.ai and ON Semiconductor

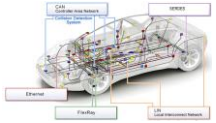
About ON Semiconductor

- Global semiconductor company with ~35,000 employees
- Broad automotive portfolio, segment contributing 33% of total revenue
  - #1 automotive image sensor provider globally

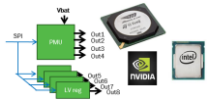
ONSemi Products Enabling AD



Active Safety Solutions



In-Vehicle Networking



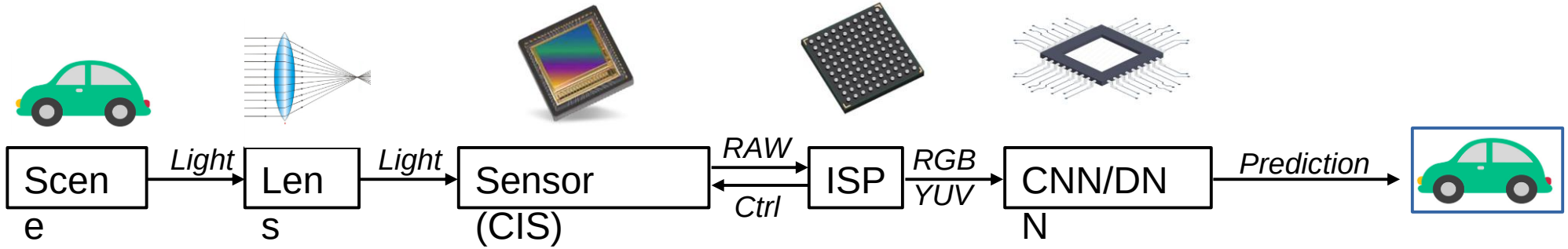
ASIL Compliant Power Management

Scope of Collaboration

- Collaborate on developing next-generation image sensing and processing technologies for machine vision and AI
- Today's presentation covers one of the collaboration topics *"Using an ISP for real-time data augmentation"*

Exposure

Perceptual CV pipeline architecture



Exposure in photography

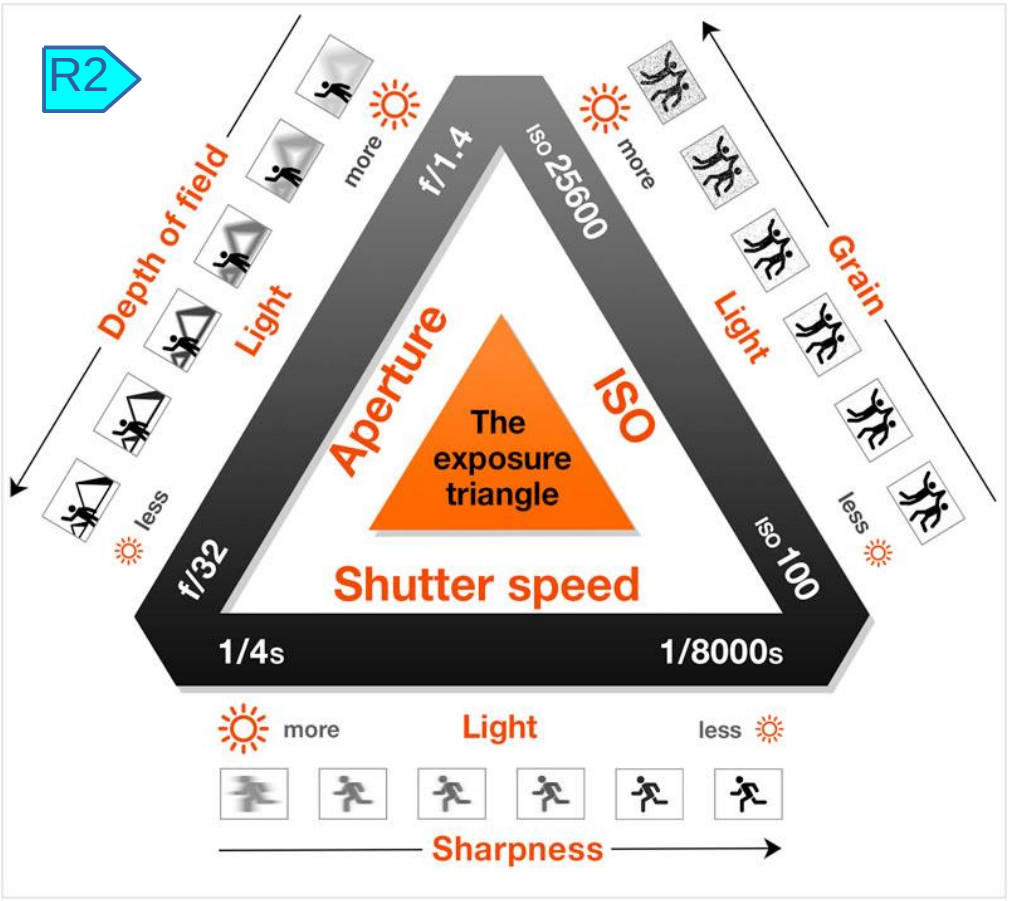
R1 Exposure is the amount of light per unit area reaching a frame of photographic film or the surface of an electronic image sensor

Exposure is determined by shutter speed, lens aperture, and scene luminance.

Exposure control is fundamental in photography.

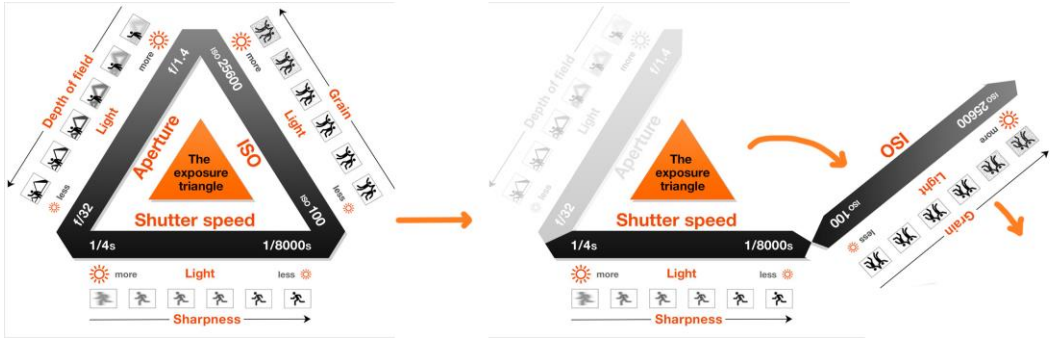
"Correct" exposure (in photography) may be defined as an exposure that achieves the effect the photographer intended.

R2

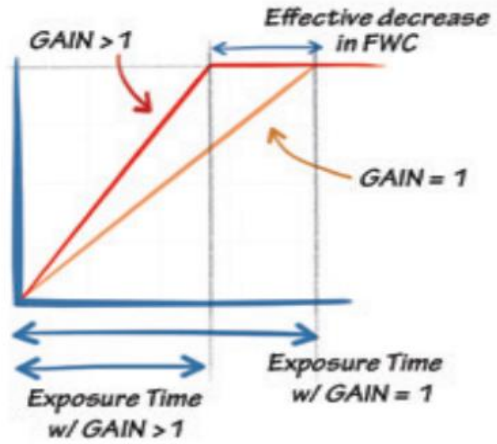


Exposure in CV and automotive

Lens aperture is fixed (no moving parts)



R1



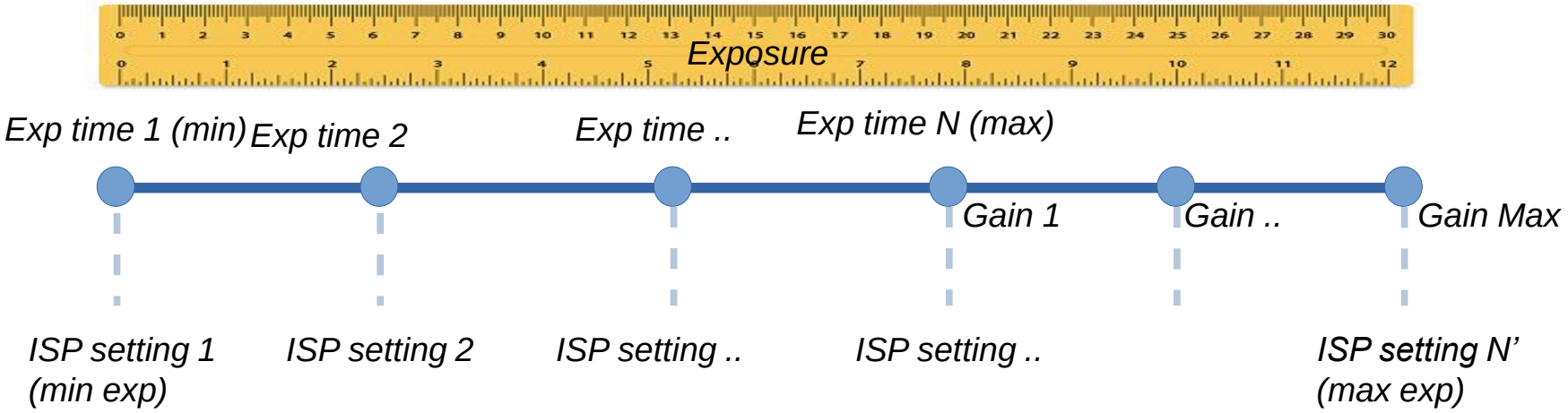
ISP (image signal processor) controls:

- Exposure time ET (integration time/shutter speed)
- Gain (ISO / film sensitivity)

based on scene light metering and history (statistics)

Exposure time is first priority. Gain compensates for dark image with short ET

Exposure in automotive and CV



Exposure time is first priority. Gain compensates for dark image with short ET

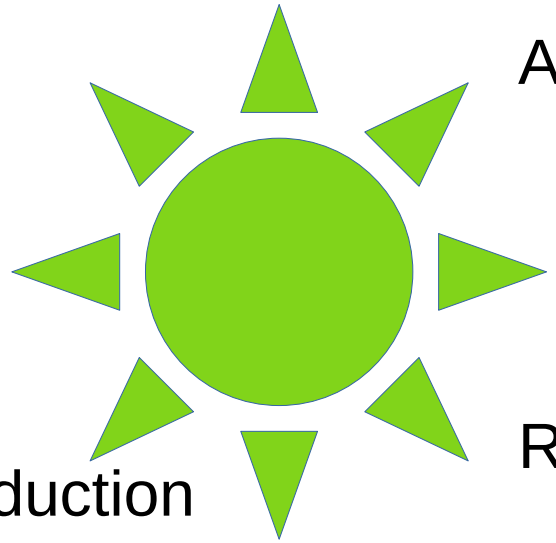
ISP usually has different setings for highest and lowest exposure points on the spectrum above (sometimes in between)

Exposure

Affect Dynamic Range (DR), tone reproduction

Noise/SNR

Affects sharpness, motion blur,



controls image sensor

Quantization

Color accuracy and reproduction

Responsible for under/over saturation

Drives ISP setting

Exposure control is crucial for object detection and safety of AV

Lab experiment 1: Automatic Exposure

Setup

Camera:
- Sensor in single exposure mode
- ISP: OnSemi AP200
- Lens: 180D ImmerVision

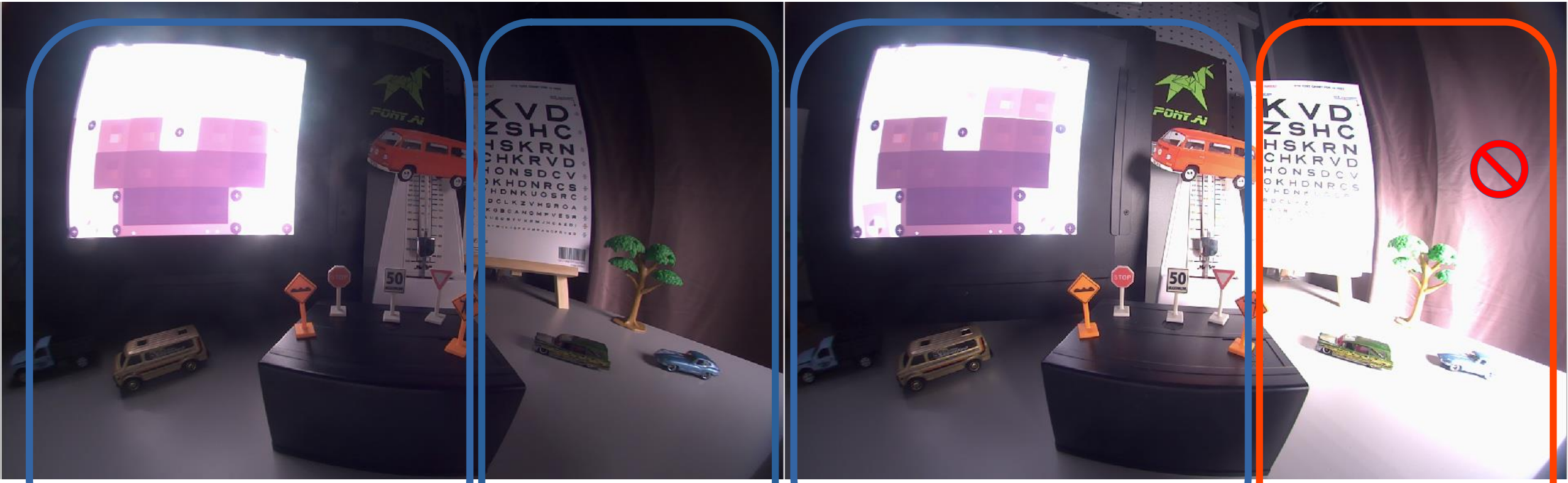
Light:
- Cool-lux LED (1/25/50/75 strength)
- Imatest Light-box

Scene:
- Imatest Contrast resolution chart
- Small car models (4x)
- Human Vision test chart
- Moving object

Test 1: AE + changed light

Weak light (<100lux)

Strong light (~500lux)



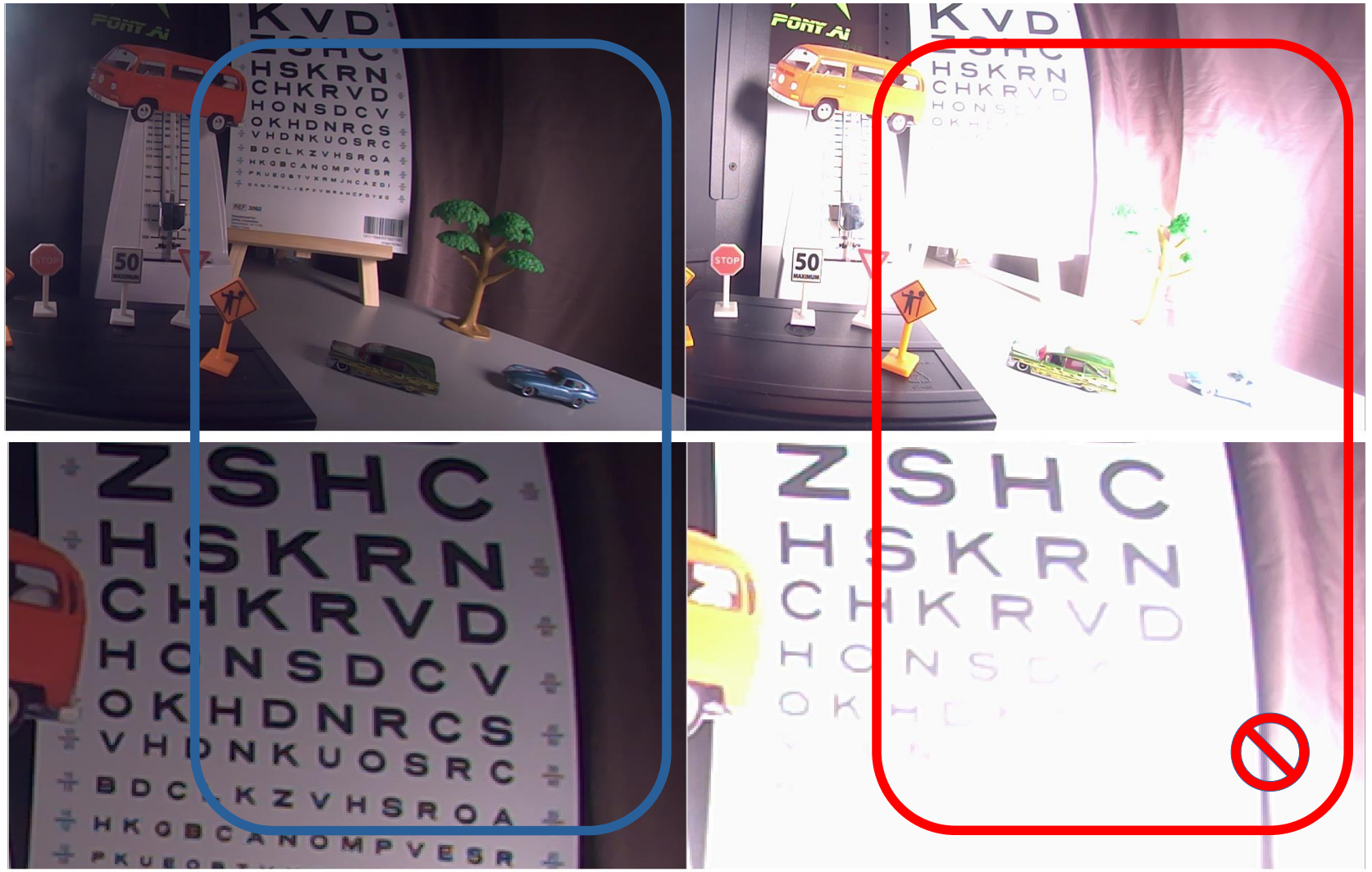
Within saturation

Within saturation

Within saturation

Over - saturation

Test 1: Automatic Exposure / Changed light



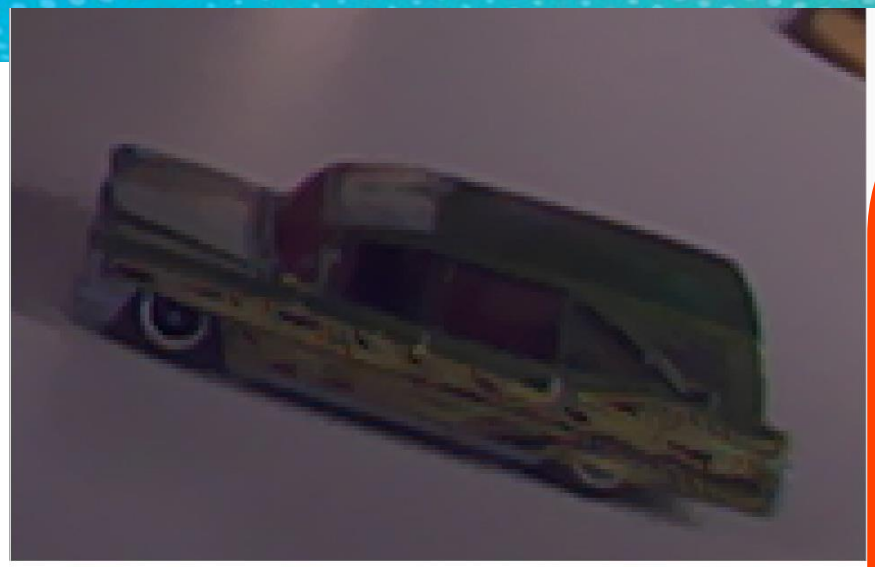
Lost visibility

Lost visibility

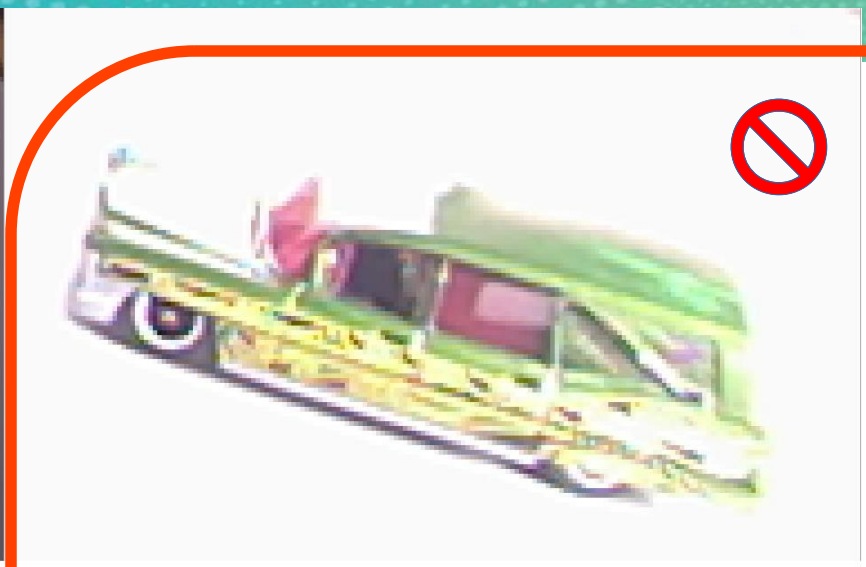
Automatic Exposure (dark light, 10)

Automatic Exposure (bright light, 250)

Test 1: AE / Changed light



Automatic Exposure (dark light, 10)



Automatic Exposure (bright light, 250)

Same Image

Automatic Exposure (dark light, 10)

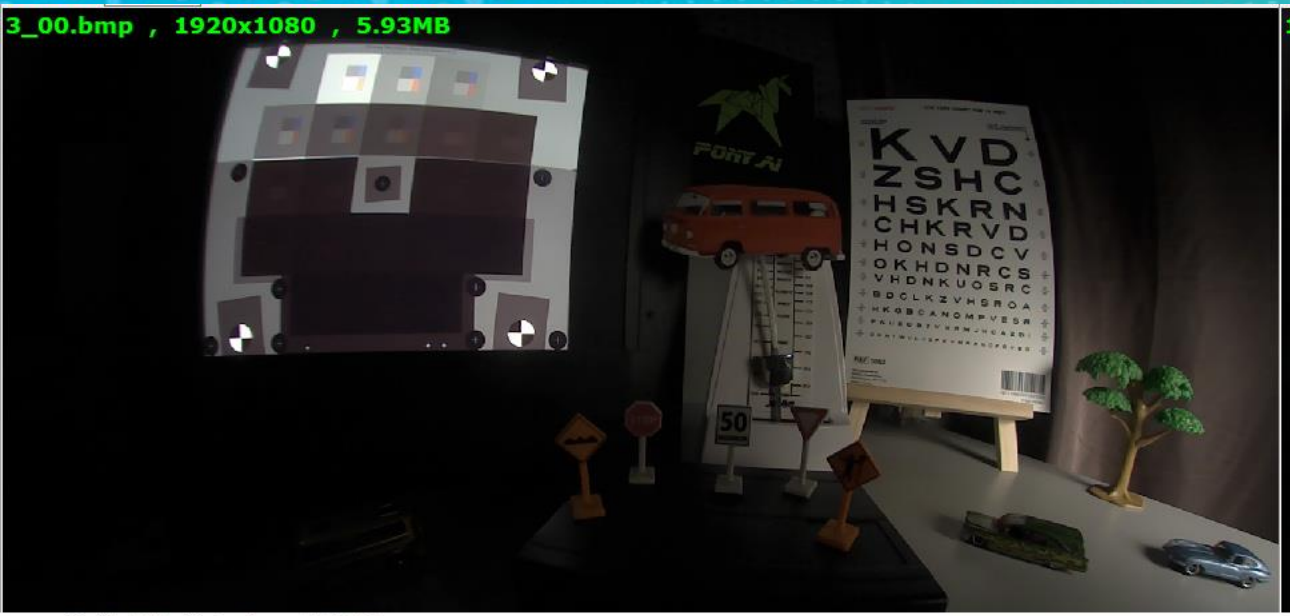
Too dark

Automatic Exposure (bright light, 250)

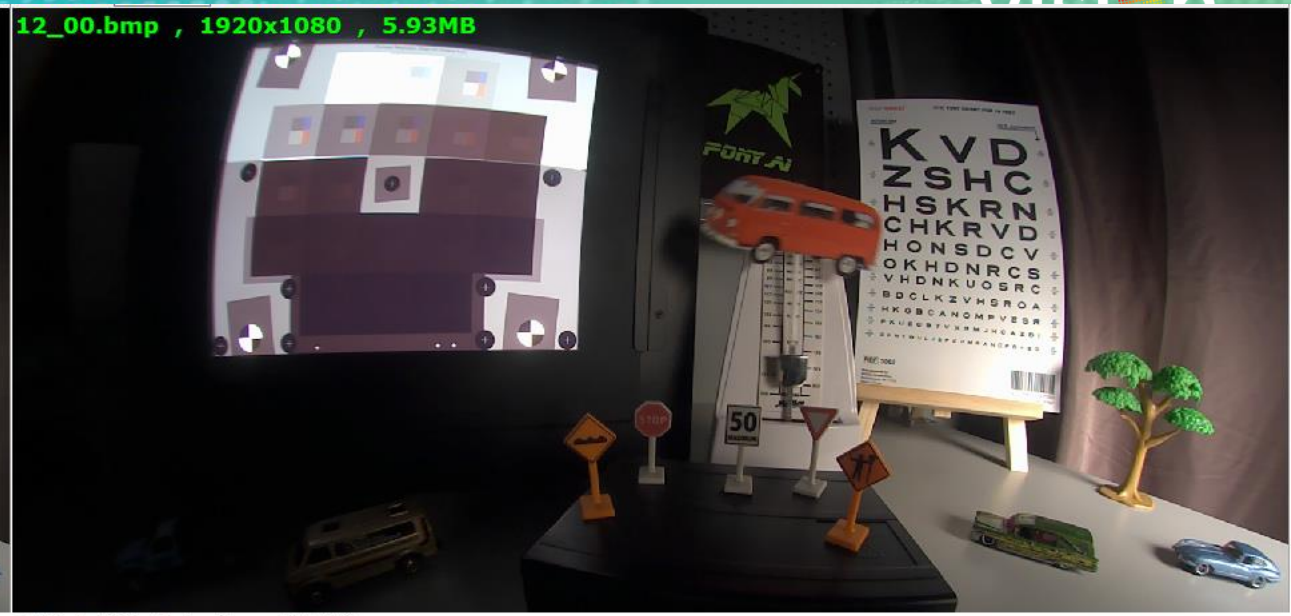
Can't Be Properly exposed

Lab Experiment 2: Manual Exposure

Manual Exposure : Full Scene

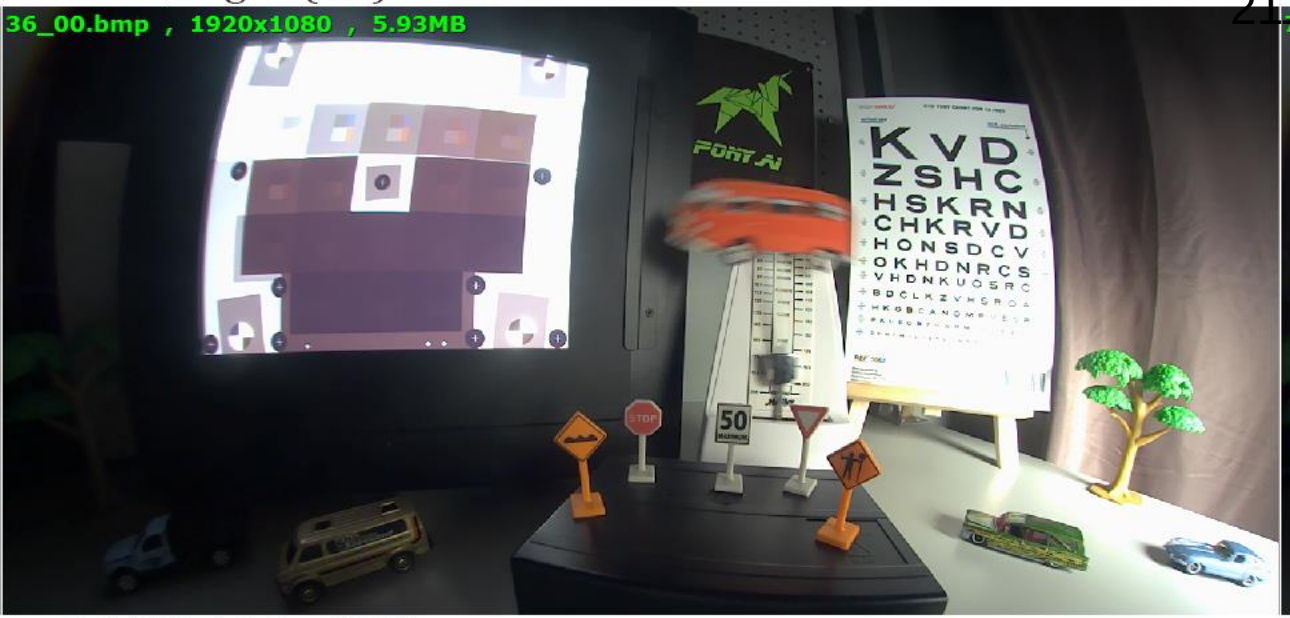


3 MS, Light (99)

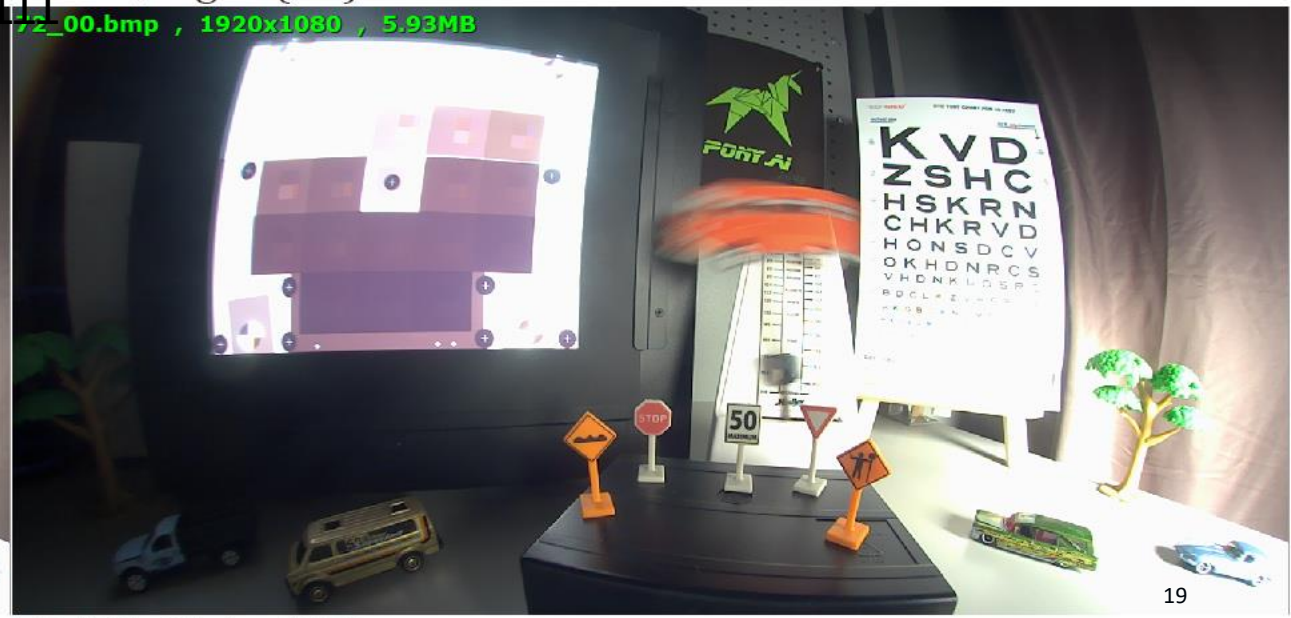


12 MS, Light (99)

2111

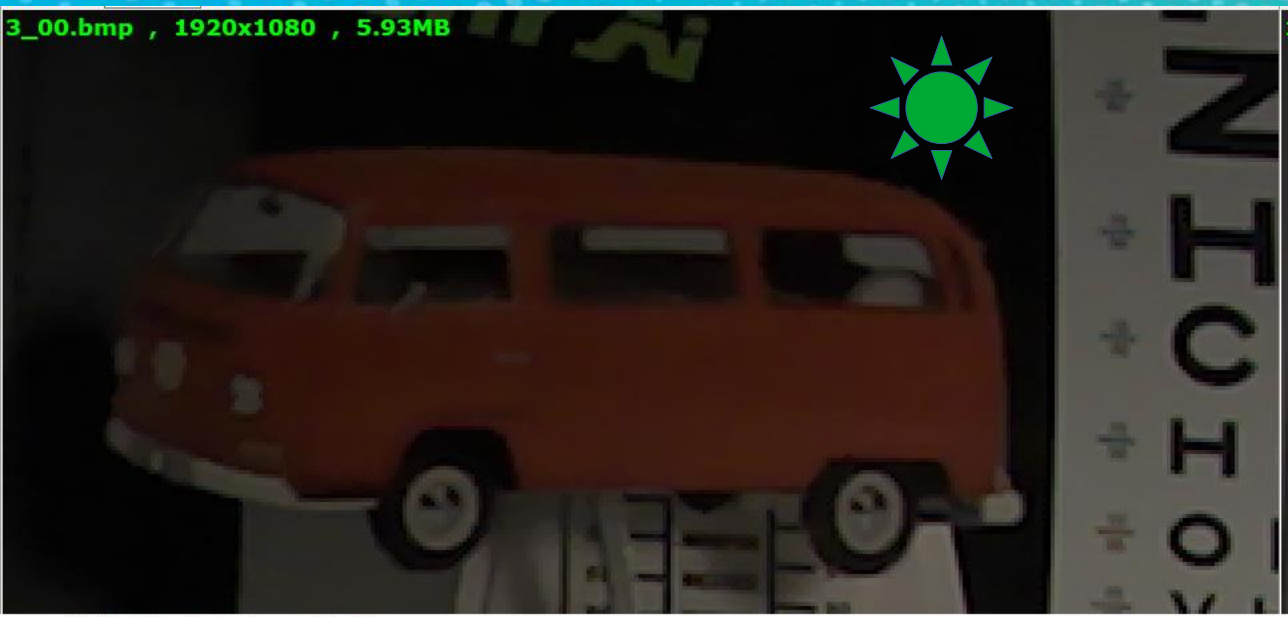


36 MS, Light (99)



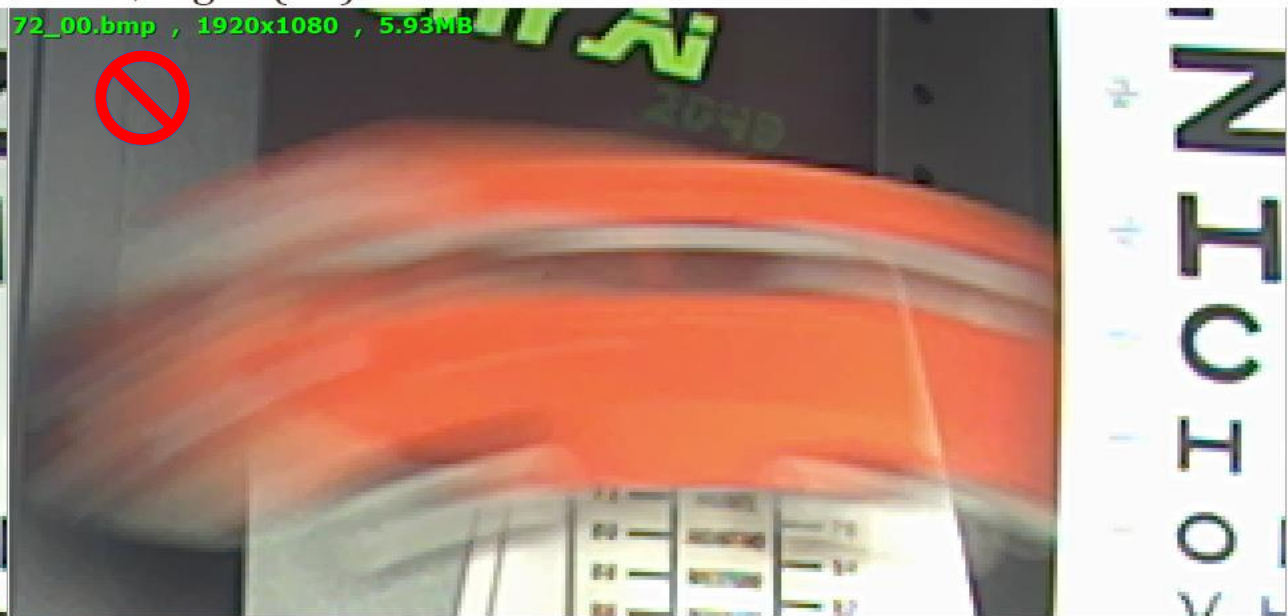
72 MS, Light (99)

Manual Exposure: Moving object

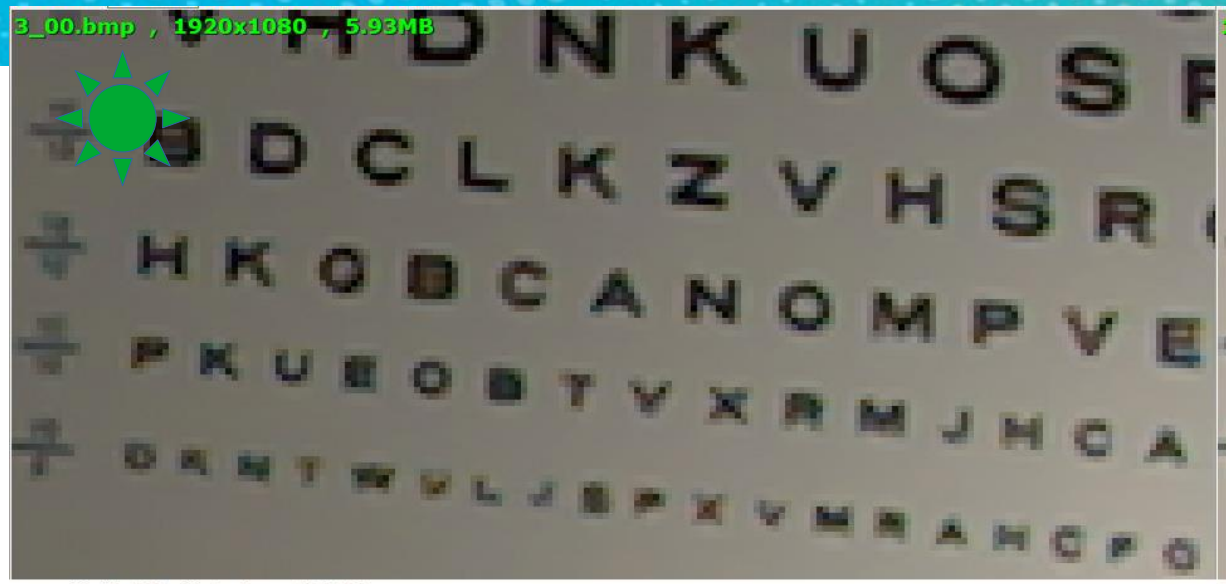


3 MS, Light (99)

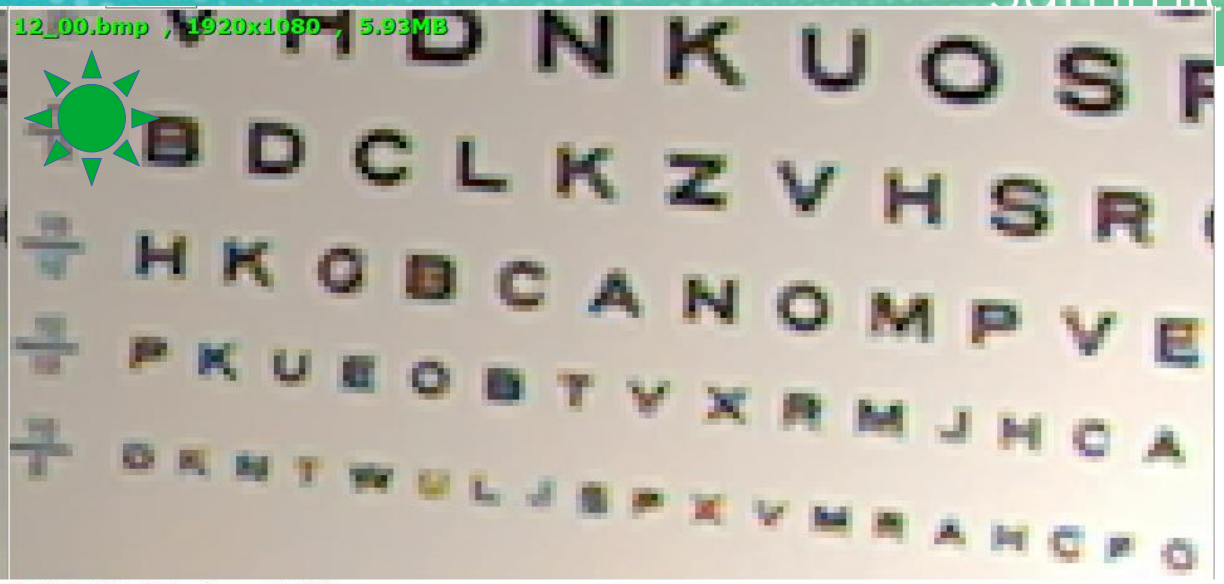
12 MS, Light (99)



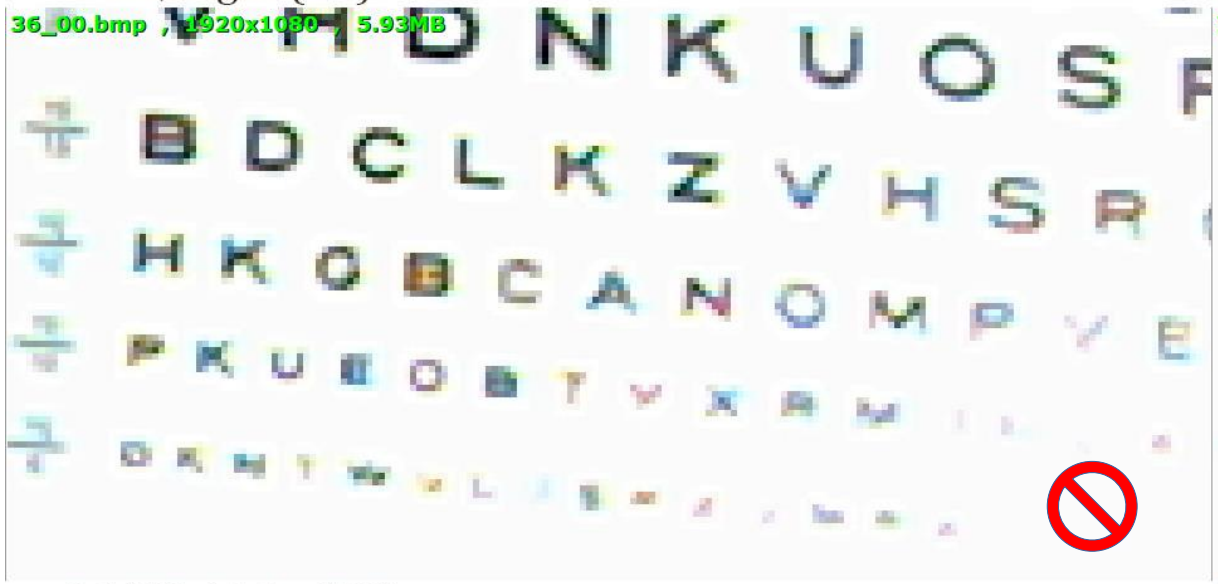
Manual Exposure: Human Vision Chart



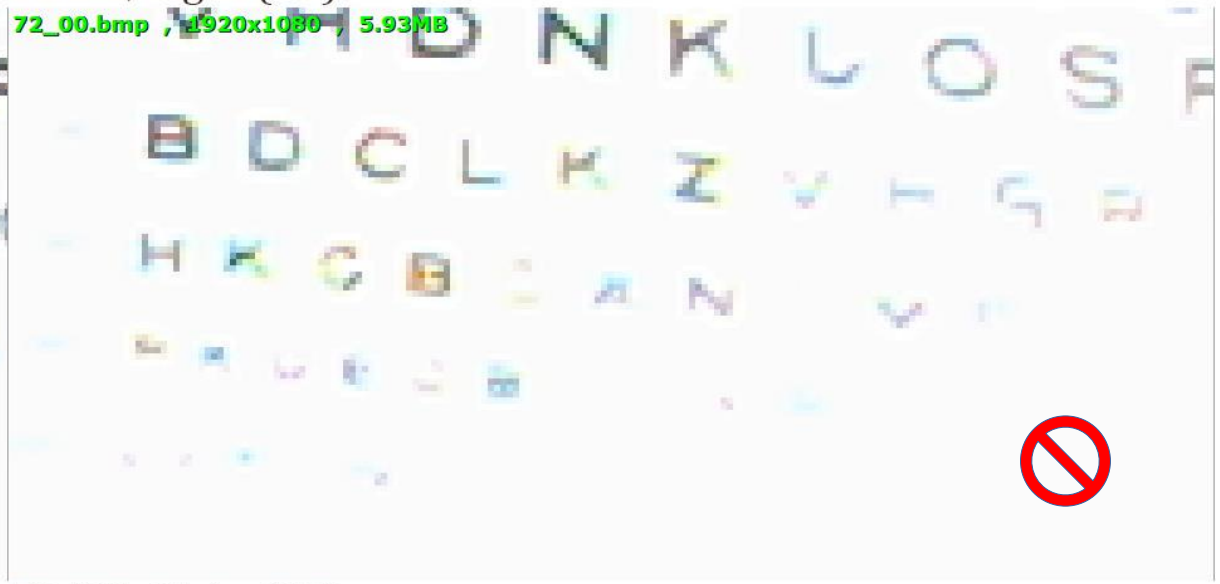
3 MS, Light (99)



12 MS, Light (99)



36 MS, Light (99)

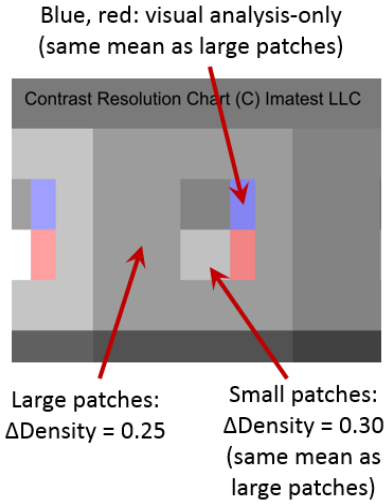


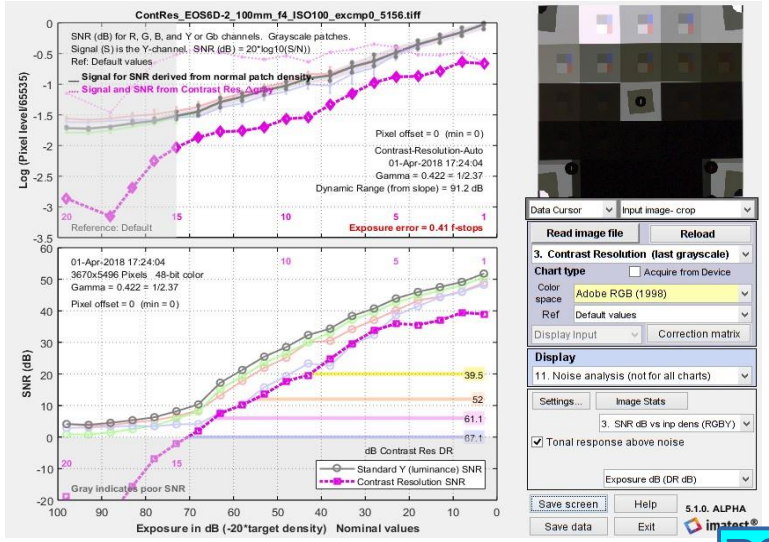
72 MS, Light (99)

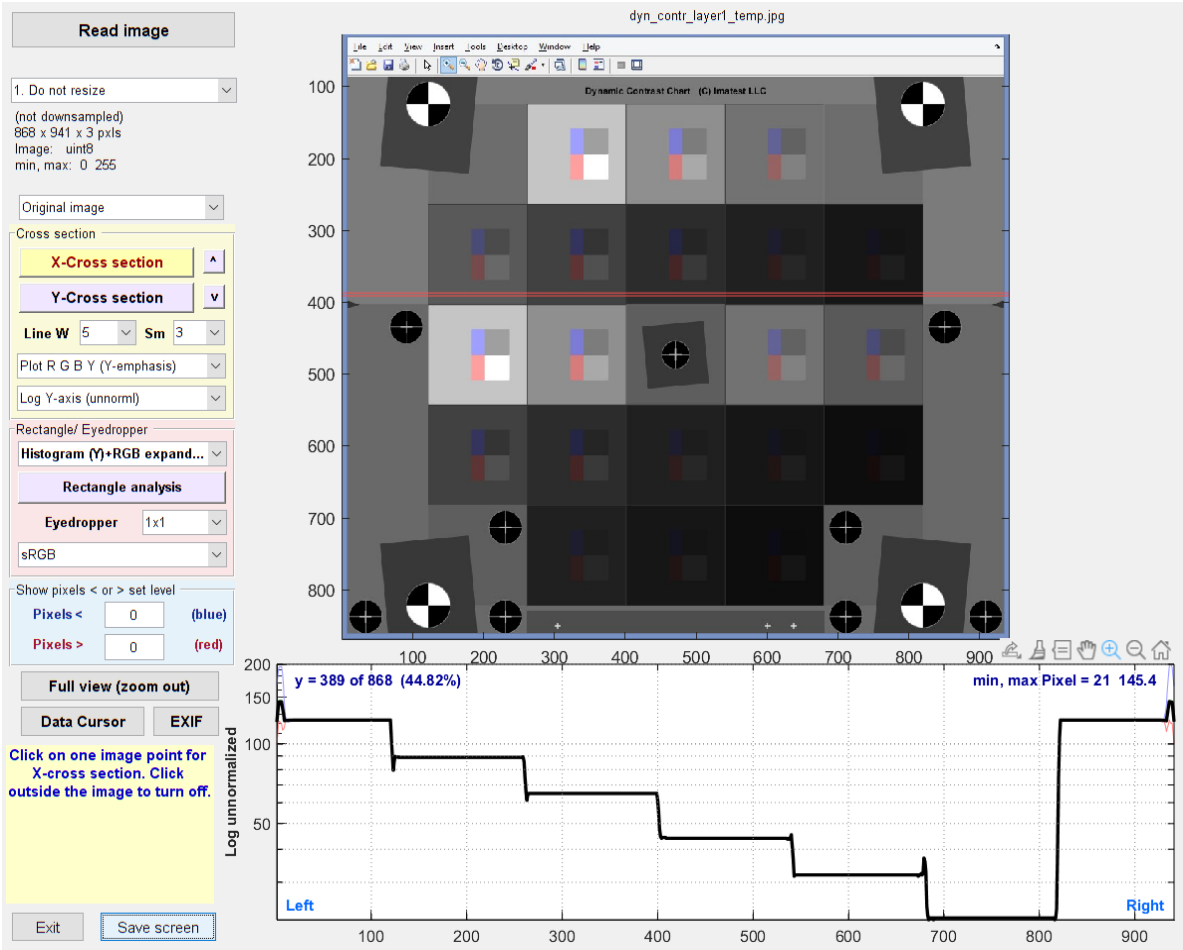
Imatest Contrast Resolution

Measures the visibility of low contrast objects in larger fields over a wide range of brightness

The range of tones a camera responds to with good contrast and Signal-to-Noise Ratio (SNR))
is a function of the sensor and lens (and to some extent, the signal processing)

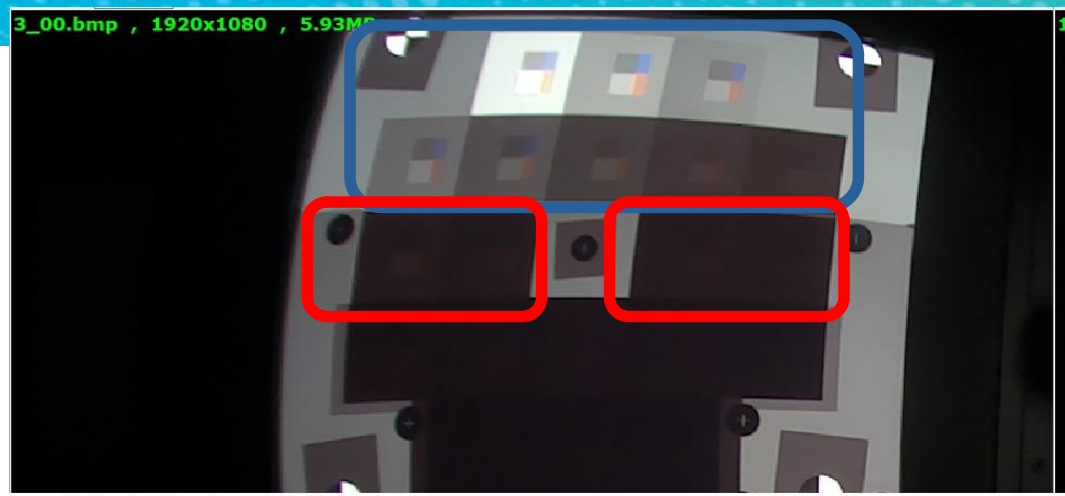




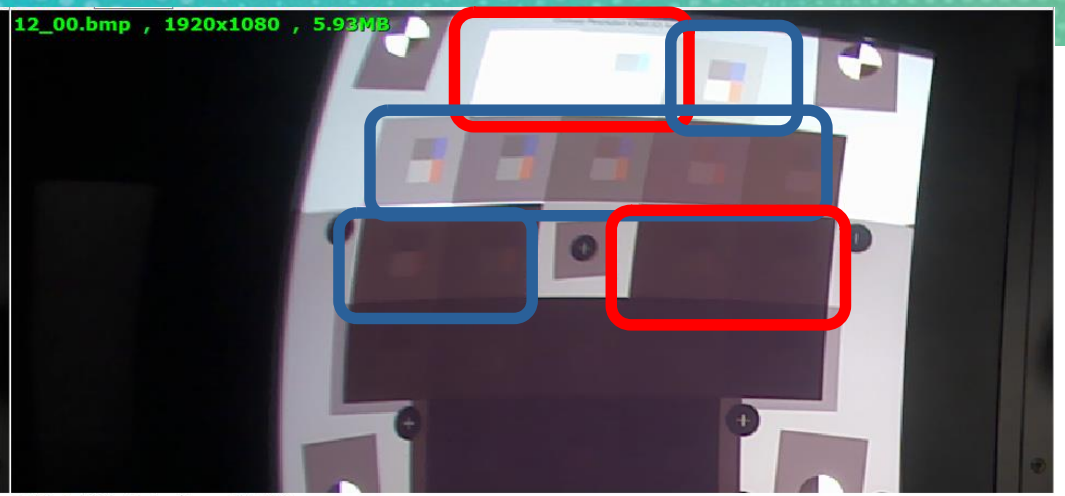


R3

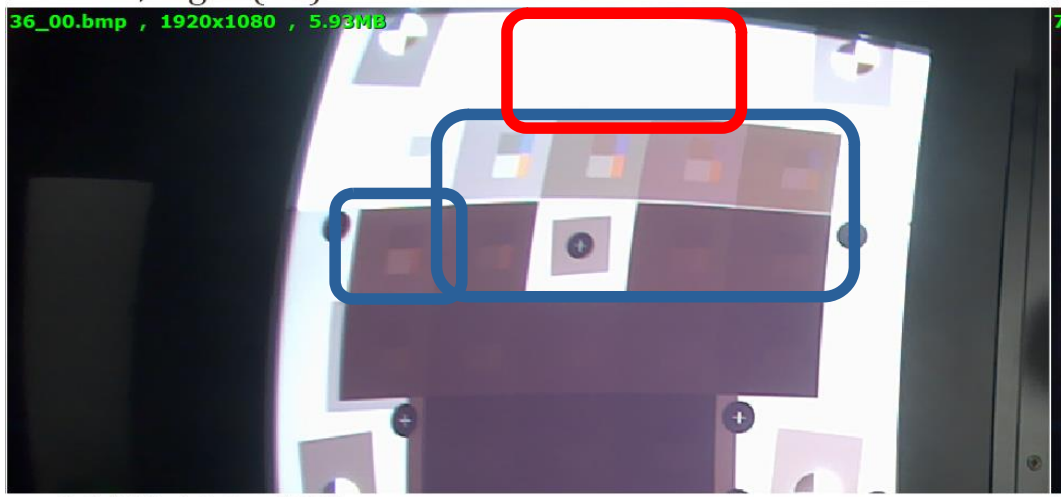
Manual Exposure: Imatest Contrast Resolution



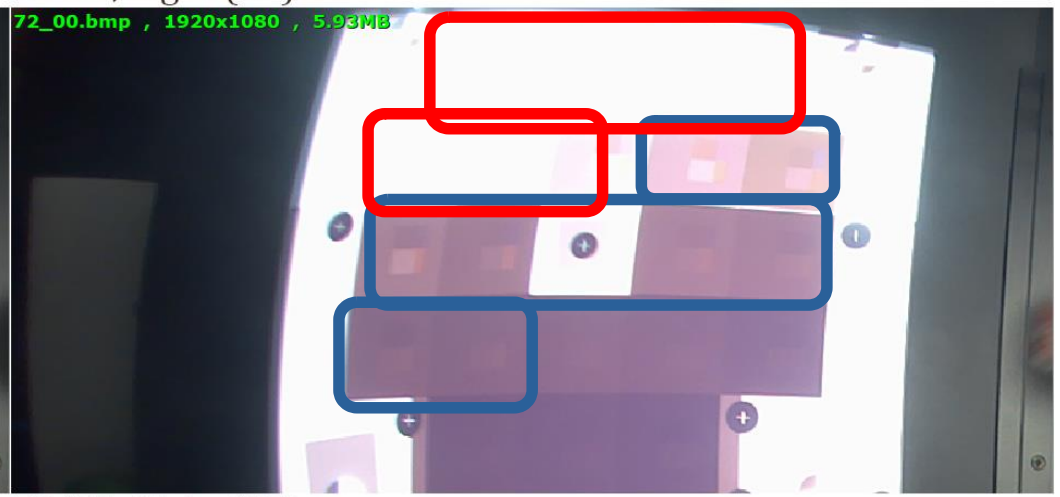
3 MS, Light (99)



12 MS, Light (99)



36 MS, Light (99)



72 MS, Light (99)

Manual Exposure: Shadows

3 MS, Light (99)

12 MS, Light (99)

36 MS, Light (99)

72 MS, Light (99)

Manual Exposure: Highlights

3 MS, Light (99)

12 MS, Light (99)

36 MS, Light (99)

72 MS, Light (99)

Field experiment, conclusions

Exposure time difference

mAP depends:
- ISP settings
- Training / Test set / Object sizes
- Model architecture

Chaning one of the the above can change detection accuracy?

Common question for ISP tuning: **A or B**?

Answer: **A and B.**

A: Long exposure time: motion blur, less noise (low gain)

B: Short exposure time: sharp image, high noise (high gain)

Conclusions

Its is impossible to choose the exposure parameters which will be optimal for all objects in the scene.

Same object appears differently depending on light, exposure control and ISP parameters.

Training set never has all possible variations of lighting conditions for any object.

.

Augmentation

Data improvement

Three ways to improve data

R7

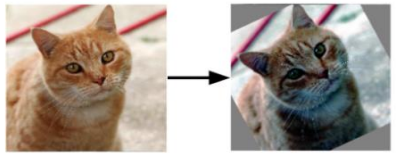
1 - Collect more

- expensive
- requires manual labor

2 - Synthesize

- complicated
- might not truly represent the real data

3 - Augment



- simple
- but finding a good augmentation strategy takes lots of trial & error (**=time of AI engineers**)

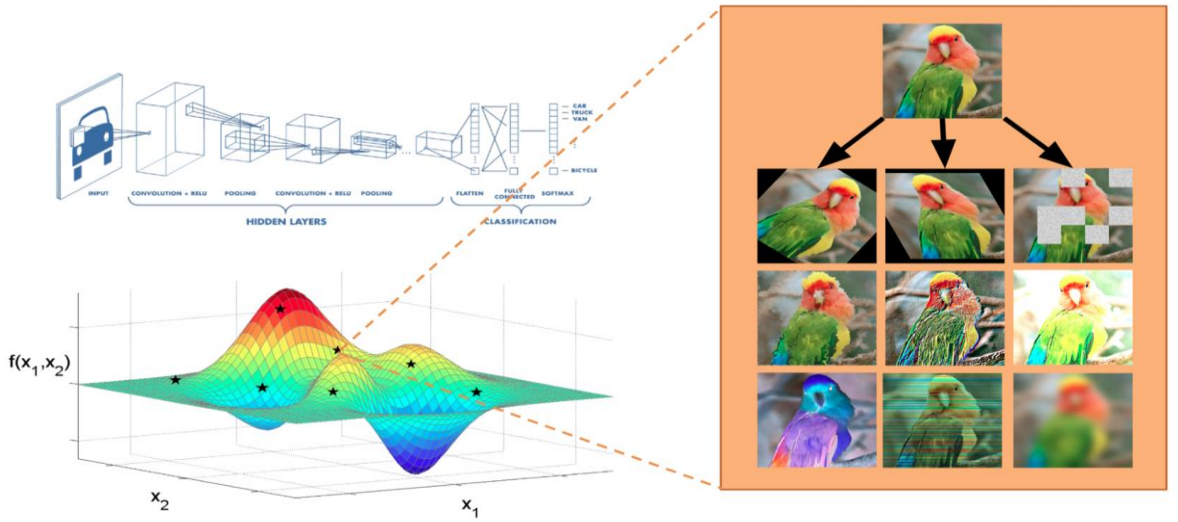
4

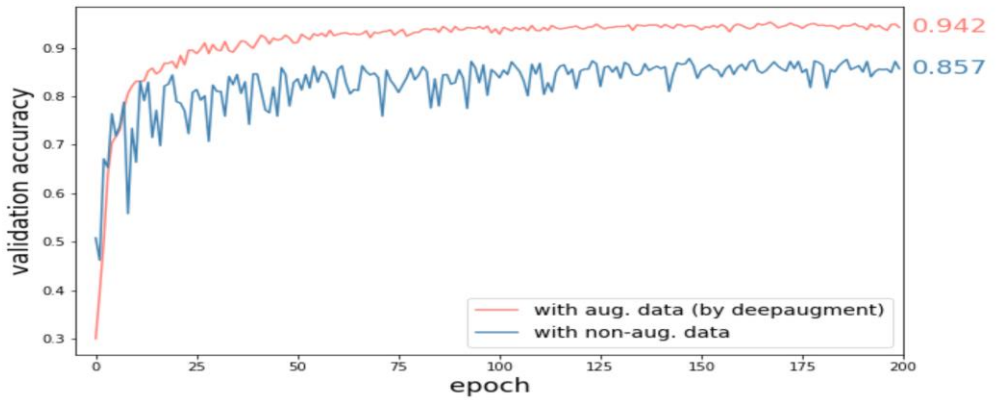
Augmentation (Auto ML)

R7

Data is the most critical piece of AI applications. Not having enough labeled data often leads to overfitting, which means the model will not be able to generalize to unseen examples.

This can be mitigated by data augmentation, which effectively increases the amount and diversity of data seen by the network. It is done by artificially producing new data by applying transformations on an original dataset such as rotation, cropping, occlusion, etc.





Comparison of validation accuracies of WideResNet-28-10 CNN model with CIFAR10 images when they are augmented by policies found by DeepAugment, and when they are not augmented. Validation accuracy is increased by 8.5%, equivalent to 60% reduction in error.

Learning Data Augmentation Strategies for Object Detection

2020 embedded vision summit

Learned augmentation policy Systematically improves object detection

Learned data augmentation improves model regularization

Sub-policy 1. (Color, 0.2, 8), (Rotate, 0.8, 10)
Sub-policy 2. (BBox_Only_ShearY, 0.8, 5)
Sub-policy 3. (SolarizeAdd, 0.6, 8), (Brightness, 0.8, 10)
Sub-policy 4. (ShearY, 0.6, 10), (BBox_Only_Equalize,0.6, 8)
Sub-policy 5. (Equalize, 0.6, 10), (TranslateX, 0.2, 2)

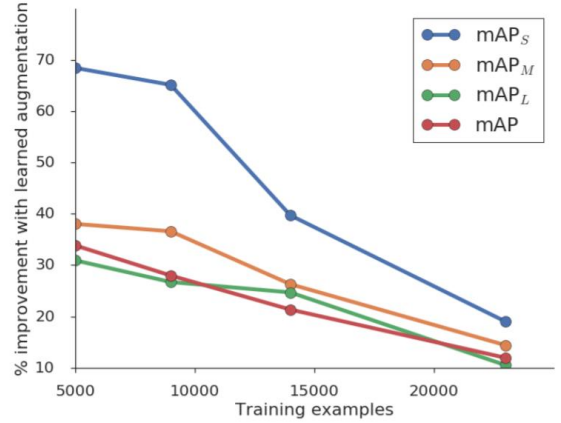
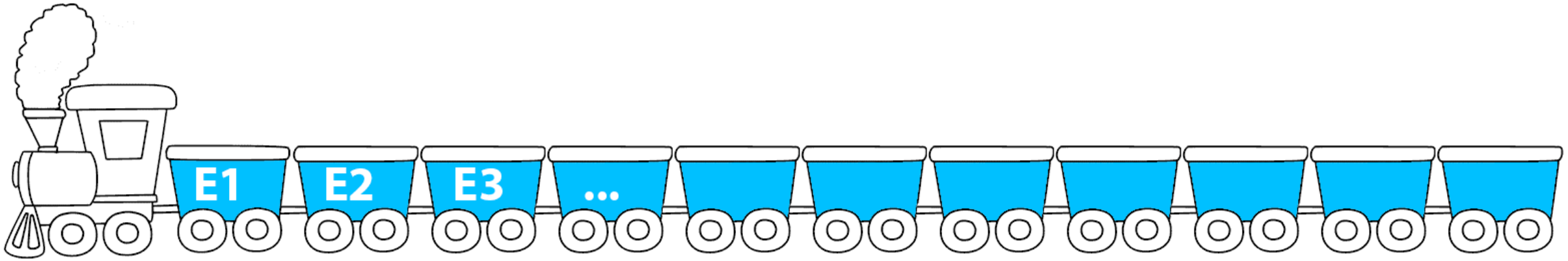


Figure 3: Percentage improvement in mAP for objects of different sizes due to the learned augmentation policy.

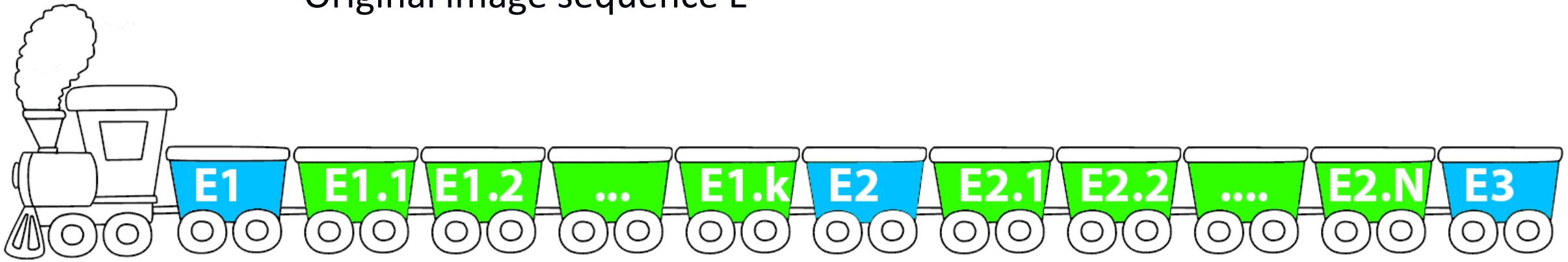
training set size	Baseline mAP$_S$	mAP$_M$	mAP$_L$	mAP	Our results mAP$_S$	mAP$_M$	mAP$_L$	mAP
5000	1.9	7.1	9.7	6.5	3.2	9.8	12.7	8.7
9000	4.3	12.3	17.6	11.8	7.1	16.8	22.3	15.1
14000	6.8	17.5	23.9	16.4	9.5	22.1	29.8	19.9
23000	10.0	24.3	33.3	22.6	11.9	27.8	36.8	25.3

Resnet-101/COCO (Small,Med, Large)

Augmented train

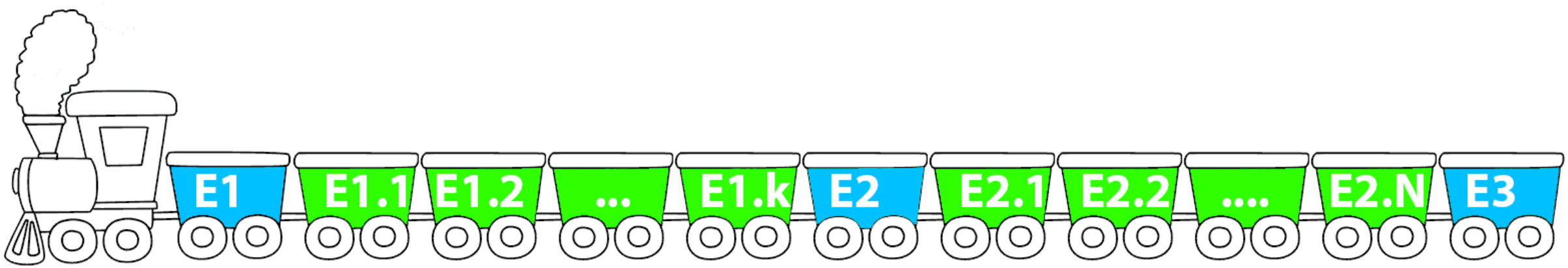


- Original image sequence E



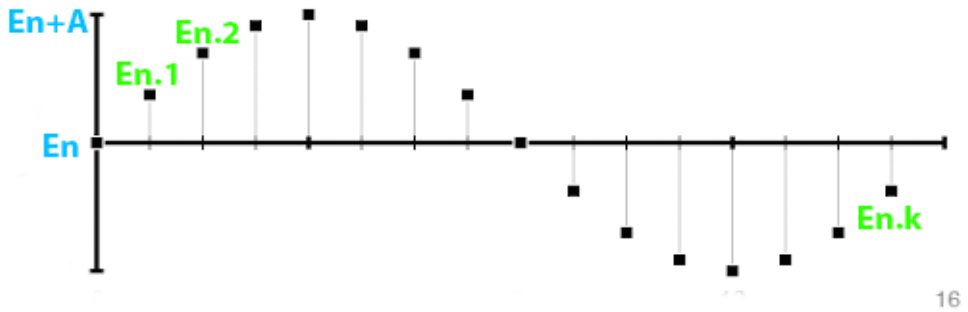
- Augmented sequence: E – anchor frames,
- k augmented frames between each two anchors

Augmented train for training networks

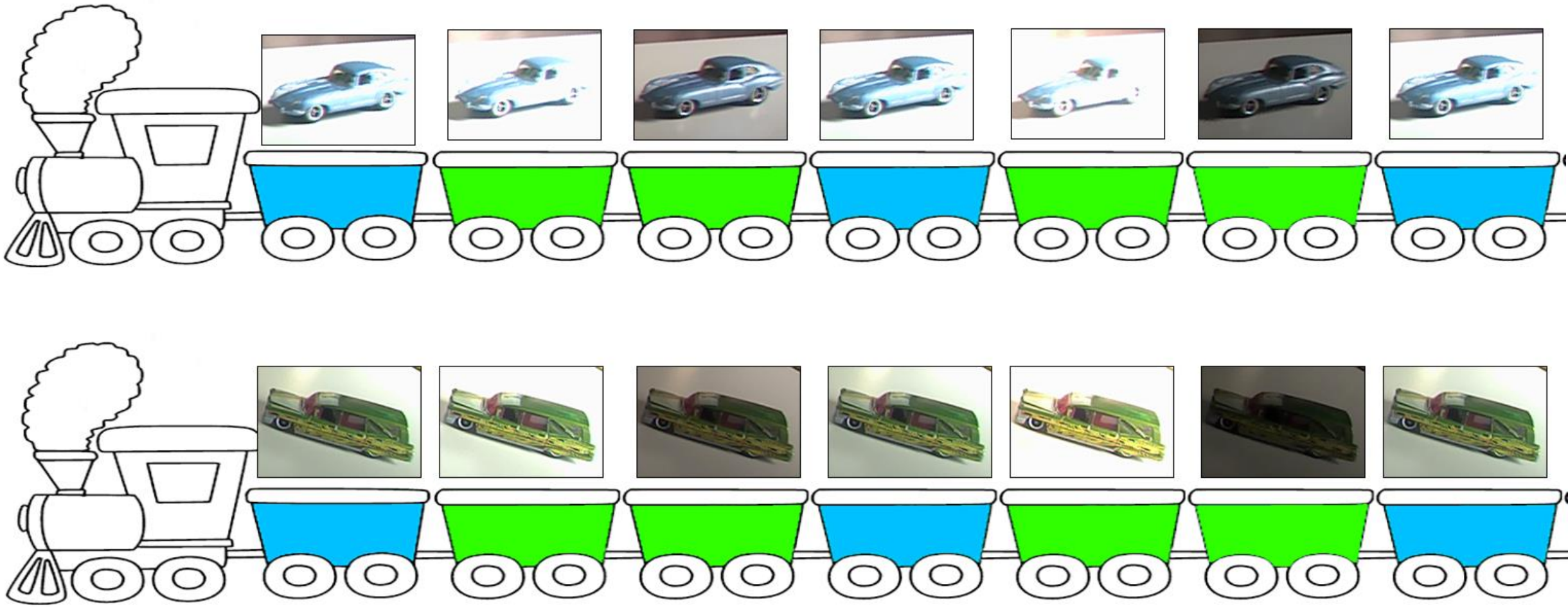


- (En)’ – exposure time of En
- (En.i)’ = (En)’ + A*sin(i*2pi/(k+1)),
- A – bracketing / augmentation strength

If (k == 1), (En.i)’ = rand(A)*(-1)^n



Example of augmented exposure trains



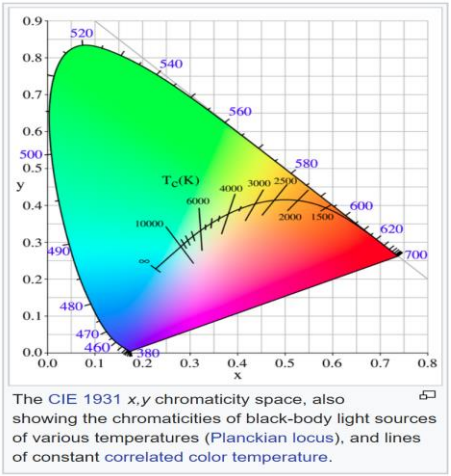
White Balance (WB) augmentation

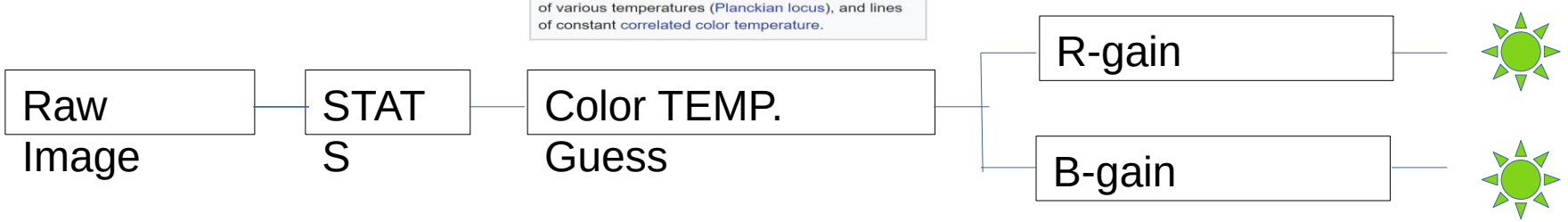
White balance (WB) is the process of removing unrealistic color casts, so that objects which appear white in person are rendered white in your photo

R6

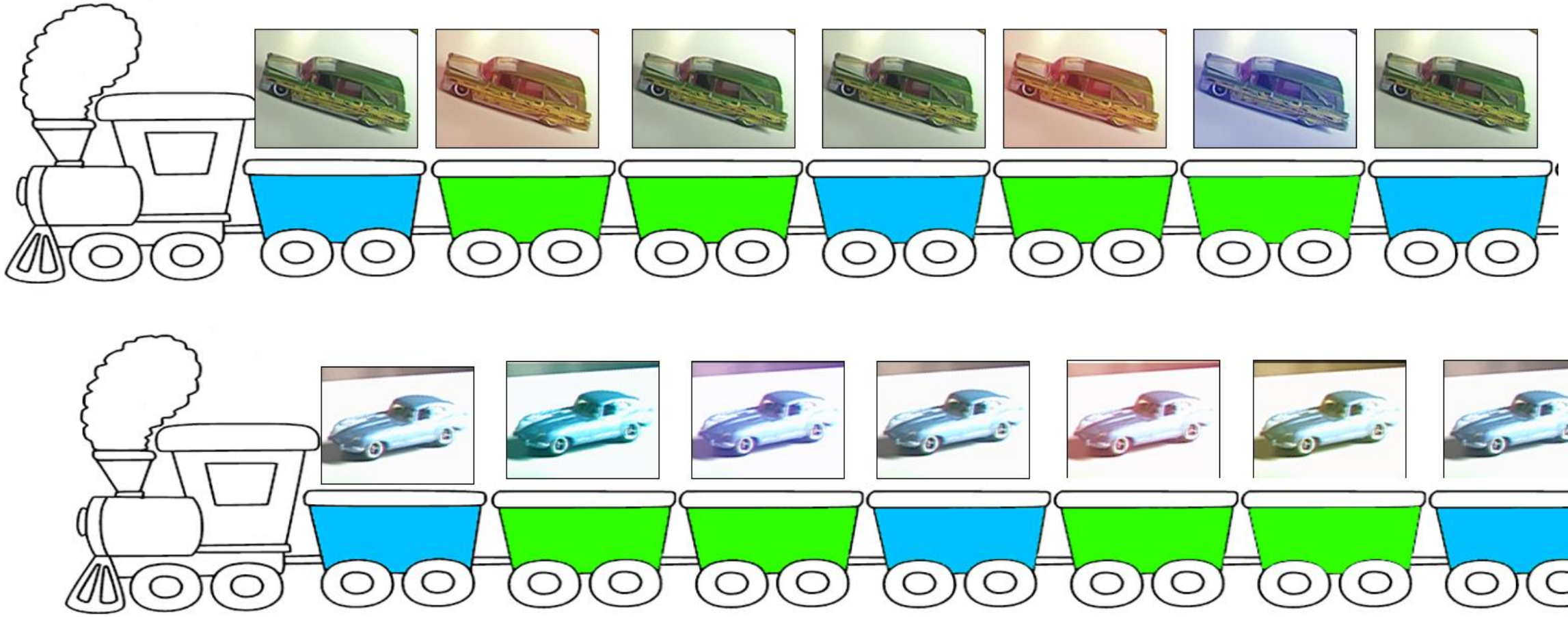
Proper camera white balance has to take into account the "color temperature" of a light source, which refers to the relative warmth or coolness of white light

R5

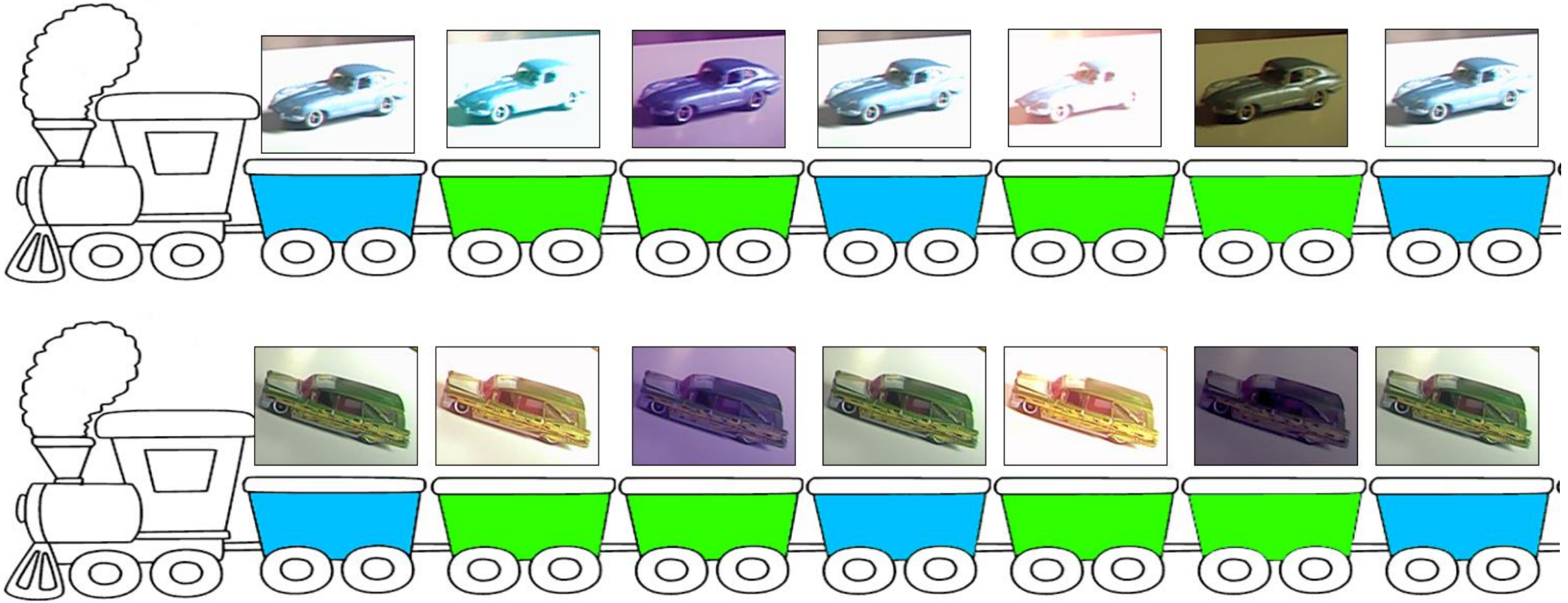




WB augmentation (Auto-Exposure)



WB + Exposure augmentation



Augmented demo

Prototype:
- Single exposure mode CIS
- ISP: OnSemi AP0200

Target Platform:
OnSemi AP0300
CEVA XM-4 dsp

ISP augmentation vs. Post processing

	ISP augmentation	Post ISP Pre-processing
Baseline	Augments capturing of the object on physical level (different physical level of noise, readout, motion blur)	Only changes image data after ISP processing
Simulation ISP parameters	Physically accurated simulation of oject representation at different lighting condition	Trying to mimic contrast, gamma, etc on existing processed image.
Data	RAW 12-16bit (pre-ISP)	8 bit RGB/YUB (post ISP)
ISP parameters	Augmented based on exposure change in realistic	ISP parameters are not change/augmented
Compute	Real-time, hw-accelerated	Uses software, GPU

ISP augmentation for imaging, post processing for geometry

Resource

1. Exposure in photography (Wikipedia)
https://en.wikipedia.org/wiki/Exposure_(photography)

2.The Exposure Triangle: Making Sense of Aperture, Shutter Speed, and ISO (Petapixel)
https://petapixel.com/2017/03/25/exposure-triangle-making-sense-aperture-shutter-speed-iso

3. Boosting sensitivity (Vision research, Phantom):
https://phantomhighspeed-service.force.com/servlet/servlet.FileDownload?file=00P1N00000dSuGDUA0

4. Contrast Resolution chart and analysis (Imatest)
https://www.imatest.com/docs/contrast-resolution/

5. Color Temperature (Wikipedia)
https://en.wikipedia.org/wiki/Color_temperature

6. Understanding White Balance (Cambridgecolor)
https://www.cambridgeincolour.com/tutorials/white-balance.htm

7. AutoML for Data Augmentation (Baris Ozmen)
https://blog.insightdatascience.com/automl-for-data-augmentation-e87cf692c366

8. Learning Data Augmentation Strategies for Object Detection (Google Research):
https://arxiv.org/abs/1906.11172

Thank you!