Mastering Automated Visual Inspection

Future Trends in Automated Inspection

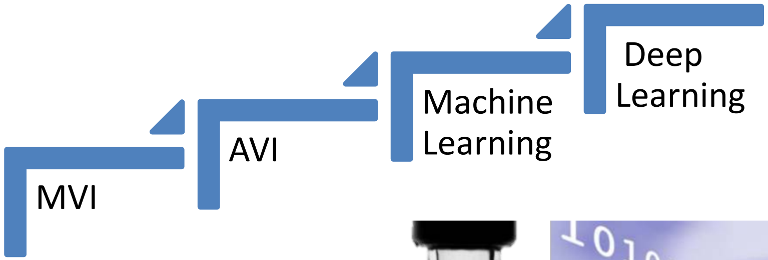
- What is deep Learning ?
- How deep Learning will transform VI ?
- Key Milestones last decade
- 1st proof of concept with cracks and particles

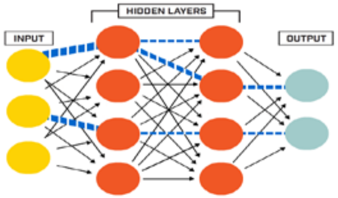
How Deep Learning will transform AVI?

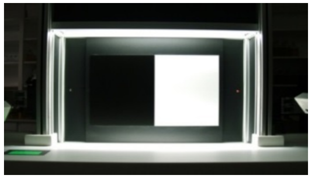
Objective: Presentation of current trends in AVI and computer vision

1. Main function blocks of AVI
2. What can « see » a machine?
3. Historic milestones
4. Comparaison Man/machine
5. How is working « deep learning » ?
6. Some practical demos

- From Manual Visual Inspection to Deep Learning



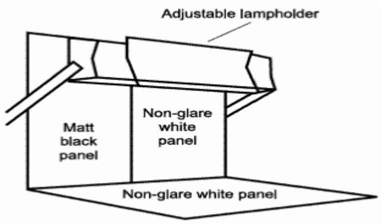


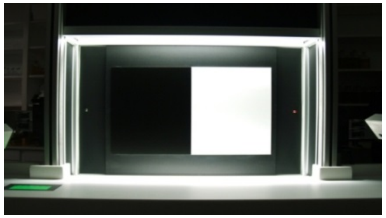


The content, material, perspectives and opinions expressed in this presentation are solely those of the presenter, and do not reflect the views or opinions of the presenters employers or PDA. Conflict of interest: Romain Veillon is an employee of the GSK group of companies. This work was sponsored by GlaxoSmithKline Biologicals SA.

• MVI is a Baseline

- Pharmacopeia test



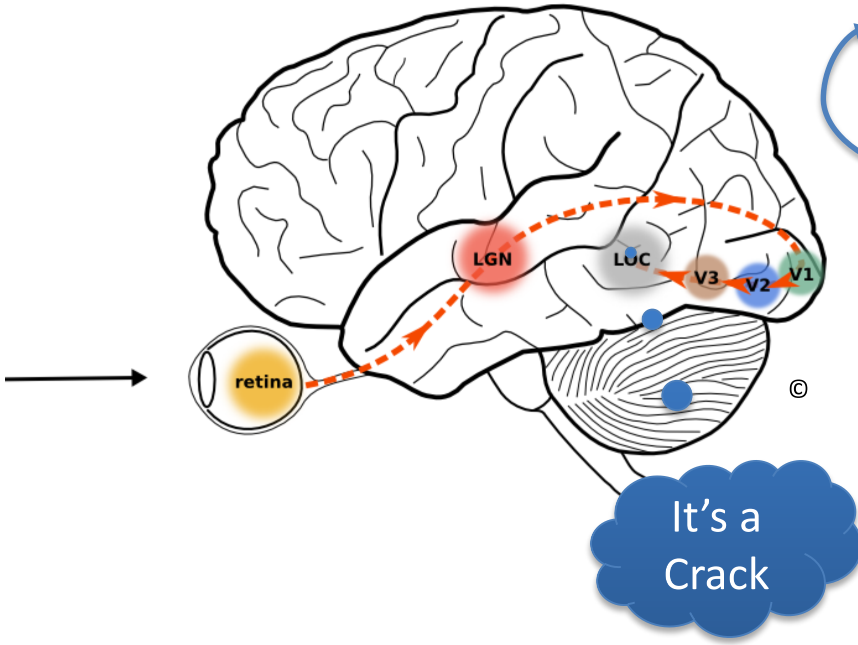


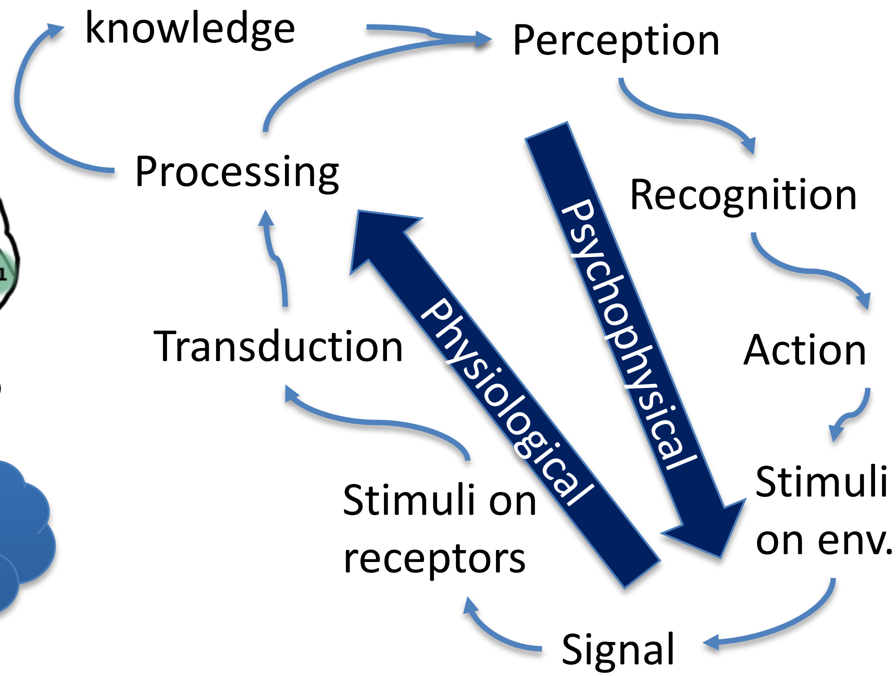
Standard work defined

- Illumination intensity lower limit 2000-3750Lux
- Black & White Background contrast
- 5 sec per each background for each container
- Maximum inspection time must be defined
- Break time must be defined
- Standard work defined
- Certification of operators

Key Take Away:

- MVI method is described in Pharmacopeias
- It remains the Baseline

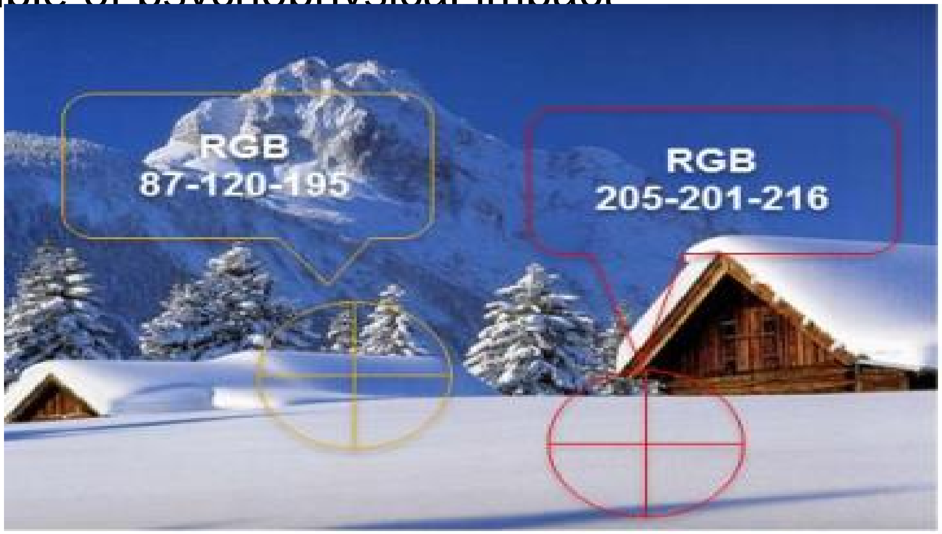




Key Take Away:

- MVI operator can detect + classify defects
- MVI operators are subject to psychophysical interactions, lead to variability in performance

- Example of psychophysical impact
-



Chromatic continuity:
We see snow even when color changes drastically (RGB)

Key Take Away:

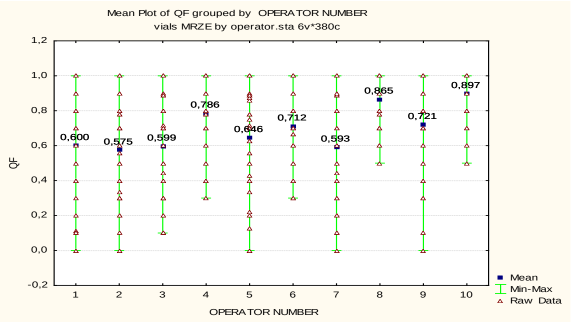
- Example of chromatic continuity shows how our brain influence decision making

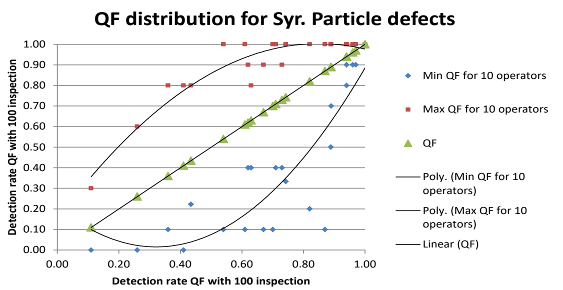
• What are MVI Baseline Performance ?

✓ Average Probability Of Detection (PoD) : [70-99%] on true defects

✓ Inter operator Variability: *

Less variability in true defect Zone: **





✓ False reject rate in MVI : 1 to 2%

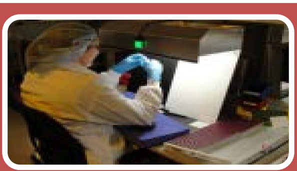
✓ Throughput +/- 1800 units / shift / operator including breaks

Key Take Away:

- MVI slow speed
- MVI variability
- Low False rejects
- Classification

*Romain veillon PDA Bethesda 2011 / ** Romain Veillon PDA Bethesda 2013

Comparison MVI to AVI



Manual Visual Inspection (MVI)

- ✓ Adaptation
- ✓ Brain
- ✓ Flexible
- ✓ Decision capable
- ✓ Classification of defects

- ✓ Inconsistent (fatigue effect, emotional)
- ✓ Not highly reproducible
- ✓ Susceptible to influence
- ✓ Slow
- ✓ Monotonous repeated work

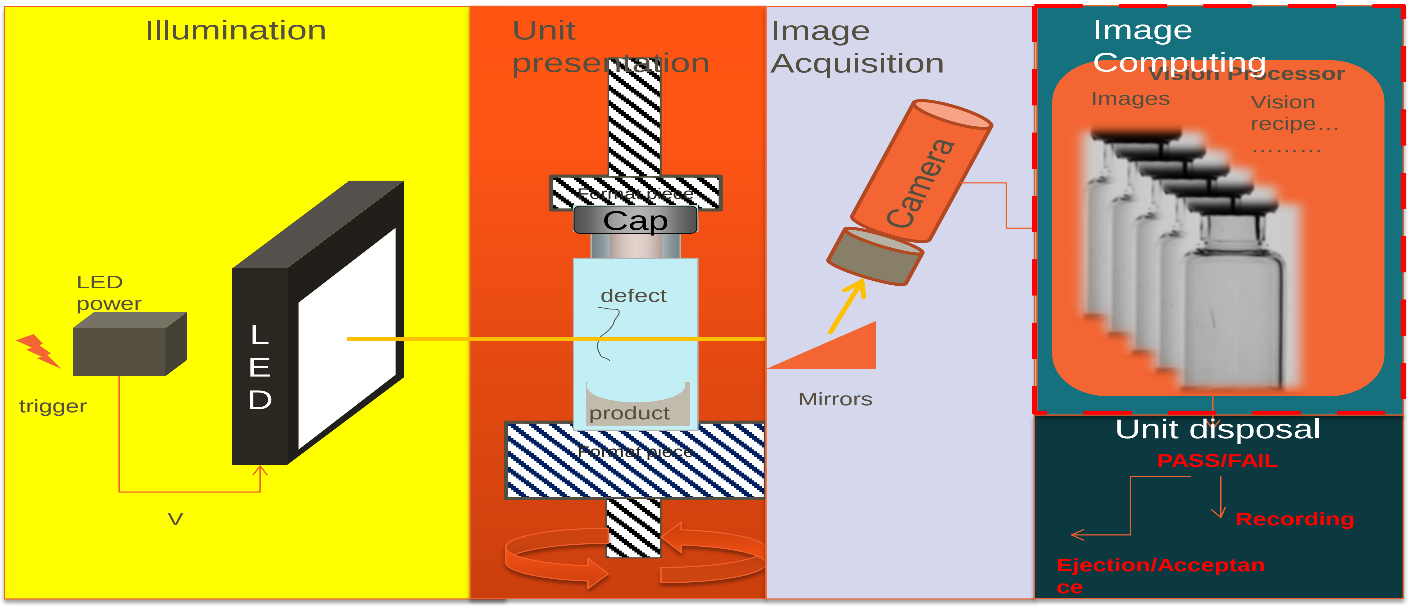
Automated Visual Inspection (AVI)

- ✓ High speed and high capability
- ✓ Highly reproducible
- ✓ Consistent (no fatigue effect)
- ✓ Defects presentation

- ✓ High initial investment
- ✓ Works within strict condition (validated upstream process)
- ✓ Indiscriminative (i.e.: fiber and cracks are seen the same way)
- ✓ Some uncovered area
- ✓ Higher false reject rate

Key Take Away: AVI and MVI are complementary

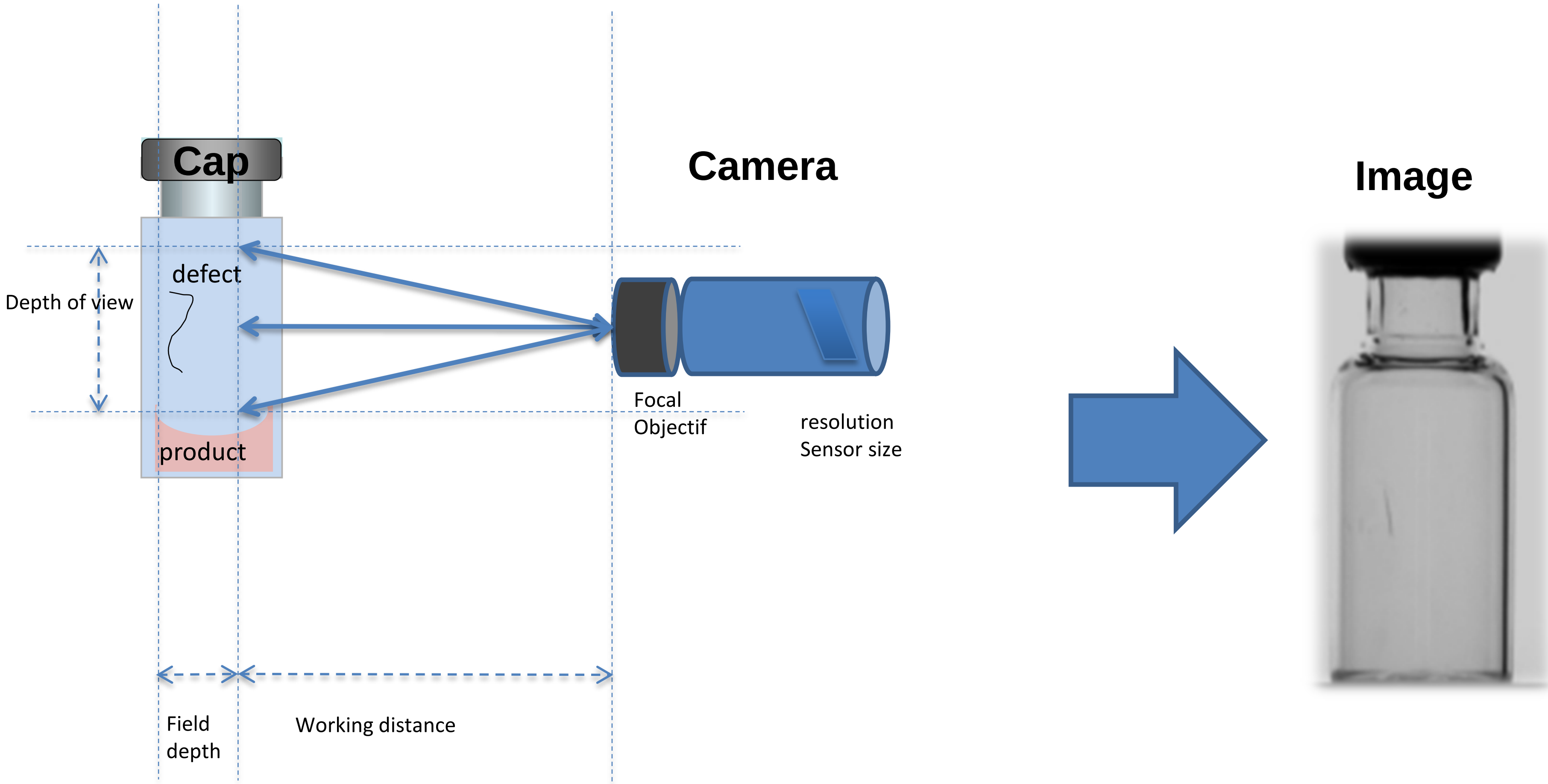
What are main function of AVI equipment ?



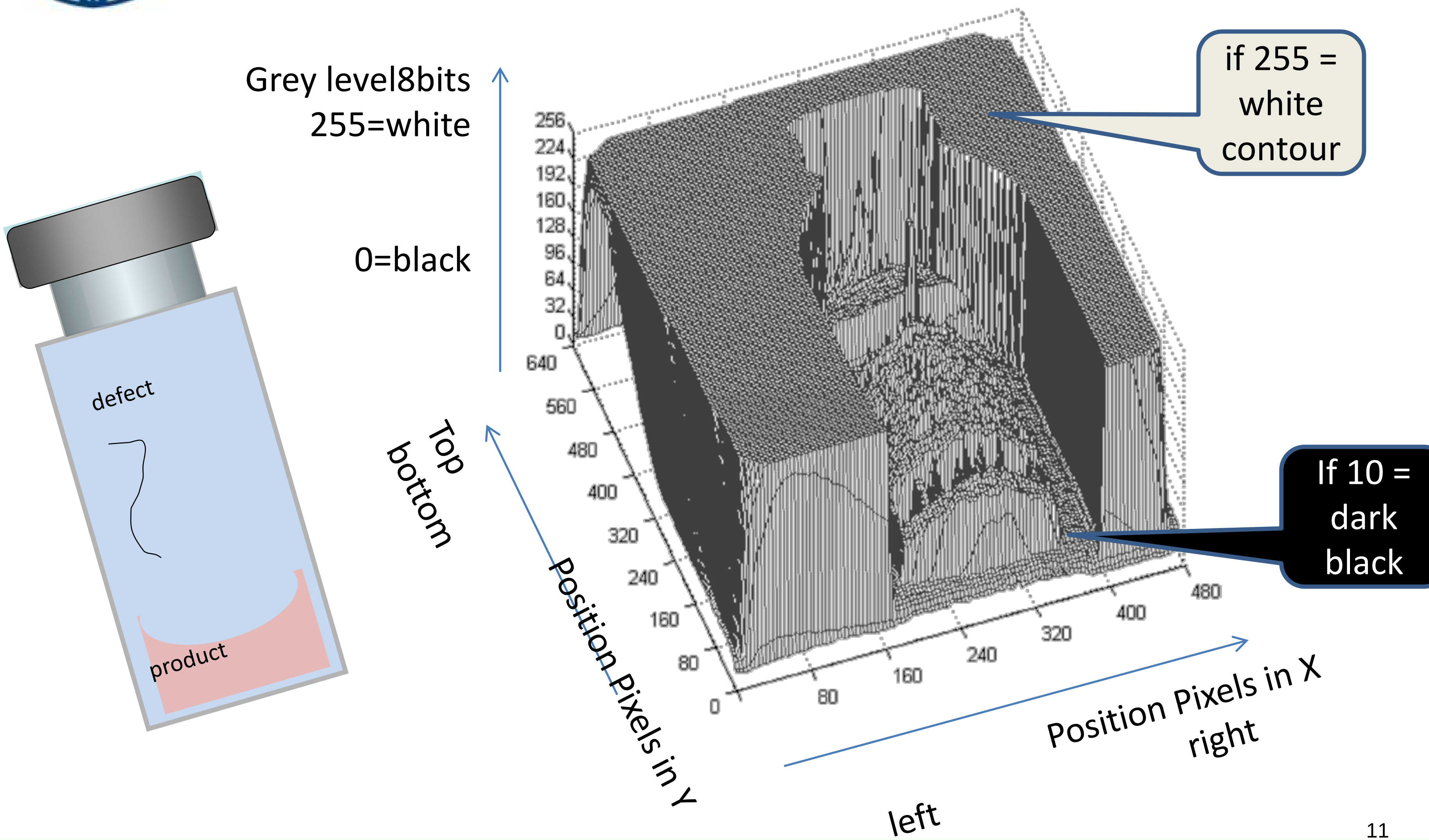
Key Take Away:

- AVI machines very complex equipment to combine
- Fast accurate mechanical handling (swiss watch like)
- Light illumination fast synchro / optical path with fast moving mirrors / vision processors to compute images in less than 15ms.
- Complex Automation to synchronize and store with data integrity

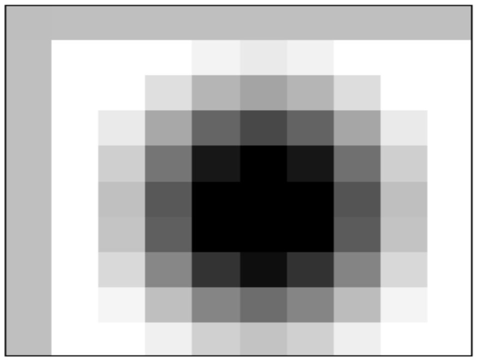
Image Acquisition principle will remain the same



What a machine really sees, what is a Digital image?



What is a digital image ?



1 particle image

0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75
0.75	1.00	1.00	1.00	0.95	0.92	0.95	1.00	1.00	1.00
0.75	1.00	1.00	0.87	0.71	0.64	0.71	0.87	1.00	1.00
0.75	1.00	0.92	0.66	0.40	0.28	0.39	0.65	0.92	1.00
0.75	1.00	0.81	0.46	0.09	0.00	0.09	0.44	0.81	1.00
0.75	1.00	0.75	0.35	0.00	0.00	0.00	0.33	0.75	1.00
0.75	1.00	0.77	0.37	0.00	0.00	0.00	0.36	0.76	1.00
0.75	1.00	0.85	0.53	0.20	0.05	0.20	0.52	0.85	1.00
0.75	1.00	0.96	0.75	0.52	0.43	0.52	0.74	0.96	1.00
0.75	1.00	1.00	0.94	0.82	0.76	0.82	0.94	1.00	1.00

Image with grey levels…

0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75
0.75	1.00	1.00	1.00	0.95	0.92	0.95	1.00	1.00	1.00
0.75	1.00	1.00	0.87	0.71	0.64	0.71	0.87	1.00	1.00
0.75	1.00	0.92	0.66	0.40	0.28	0.39	0.65	0.92	1.00
0.75	1.00	0.81	0.46	0.09	0.00	0.09	0.44	0.81	1.00
0.75	1.00	0.75	0.35	0.00	0.00	0.00	0.33	0.75	1.00
0.75	1.00	0.77	0.37	0.00	0.00	0.00	0.36	0.76	1.00
0.75	1.00	0.85	0.53	0.20	0.05	0.20	0.52	0.85	1.00
0.75	1.00	0.96	0.75	0.52	0.43	0.52	0.74	0.96	1.00
0.75	1.00	1.00	0.94	0.82	0.76	0.82	0.94	1.00	1.00

Digital Image = matrix grid of figures in X and Y

In computer vision language (python/C++) it is a matrix object:

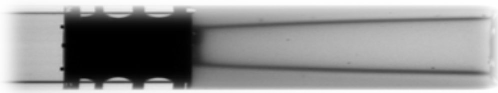
```
np.zeros(img.shape, dtype=img.dtype)
```

Key Take Away:

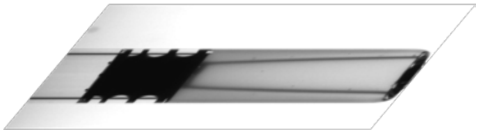
- Computer vision see only a matrix
- That represent spatial distribution of grey levels
- Neural Network will work with image matrix

- AVI: From image acquisition
- to feature extraction and decision

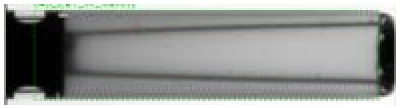
1 3D Object presented



2 2D image

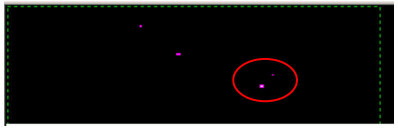


3 Area Of Interest (=AOI)



4 Binarization

5 Object detection



6

Image feature rules

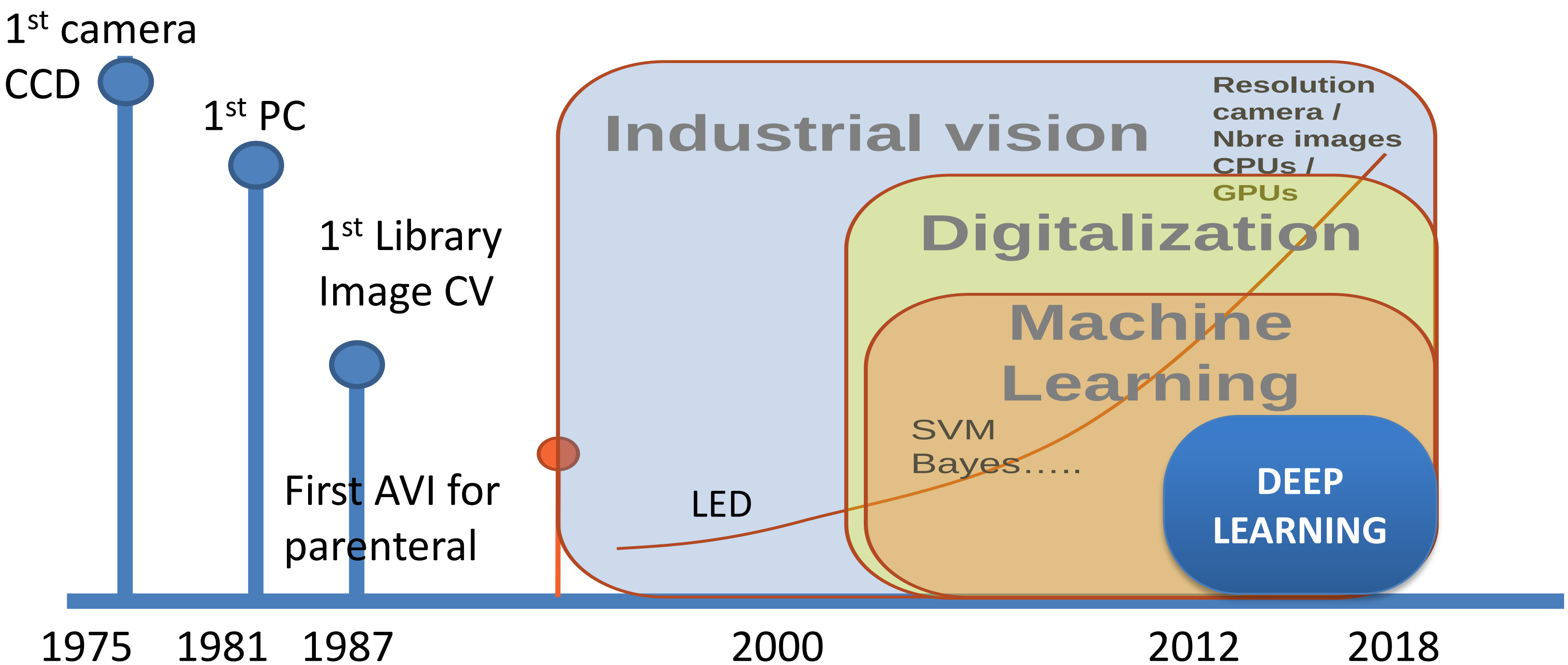
Feature Name	Current Value
Box X Min	227.0
Box Y Max	72.0
Box Y Min	70.0
Center of Gravity X	228.2
Center of Gravity Y	71.0
Convex Perimeter	12.0
Elongation	1.4
Feret Elongation	1.4
Feret Max Diameter	4.2
Moment Central X1 Y1	0.0

7 Pass / Fail

Key take away:

- **this engineering step of vision recipe development is done to reduce information of images and to enhance specificity of decision….in less than 10ms**
- **It remains low level vision**

- AVI is a fast evolving technology
- with LED / Digitalization / Deep Learning

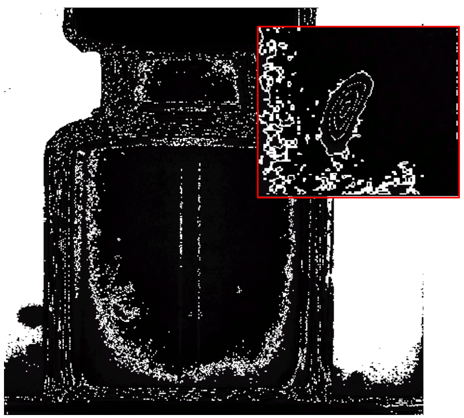


Key Take Away:

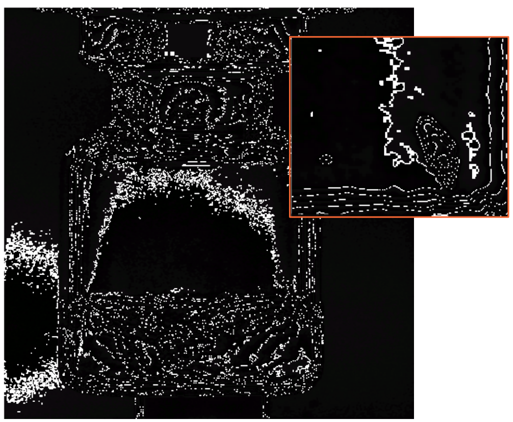
- AVI is a young, maturing technology
- Many changes over the last 30 years, next one is deep learning

• What is benefit of digitalization ?

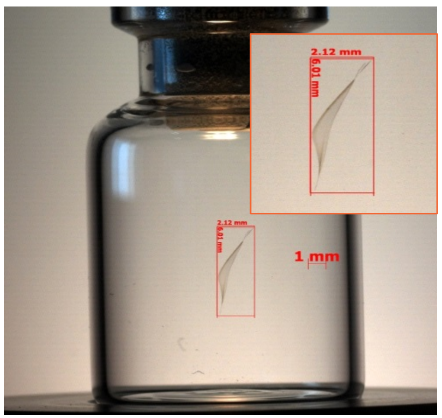
VGA analogic



2M Pixels digital



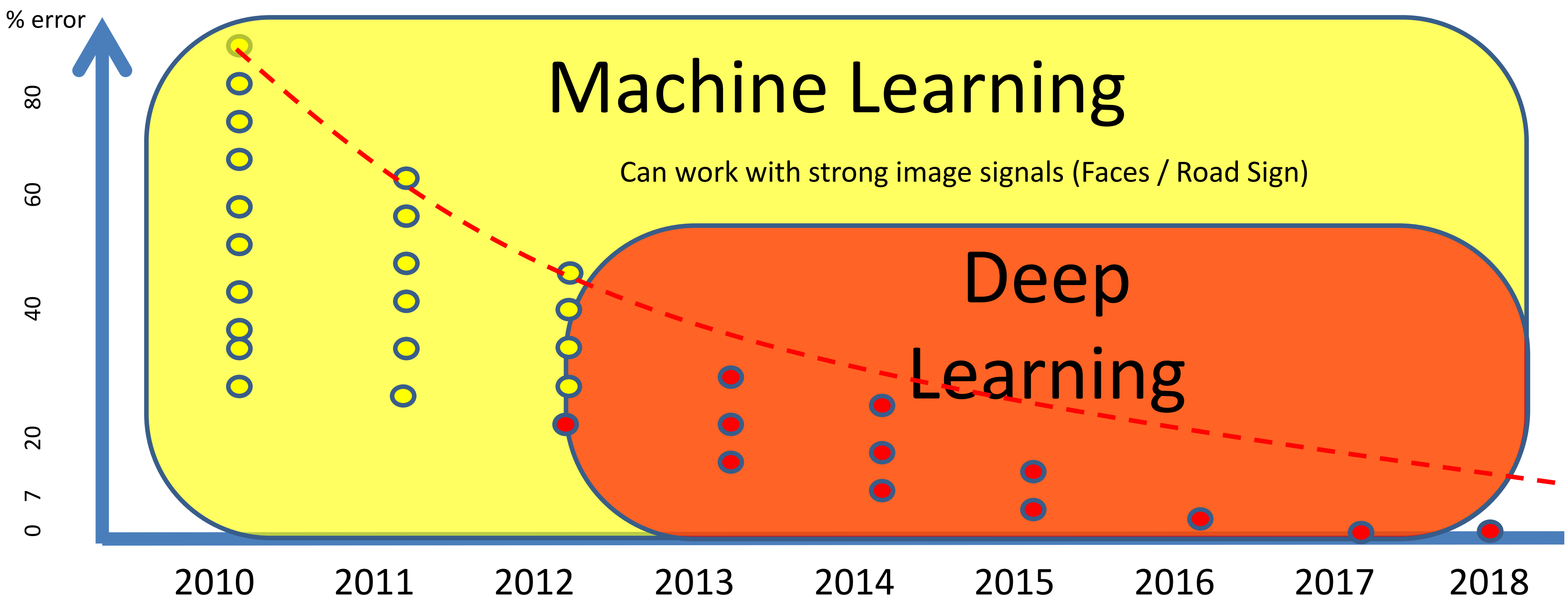
12MPixels digital



Key Take Away:

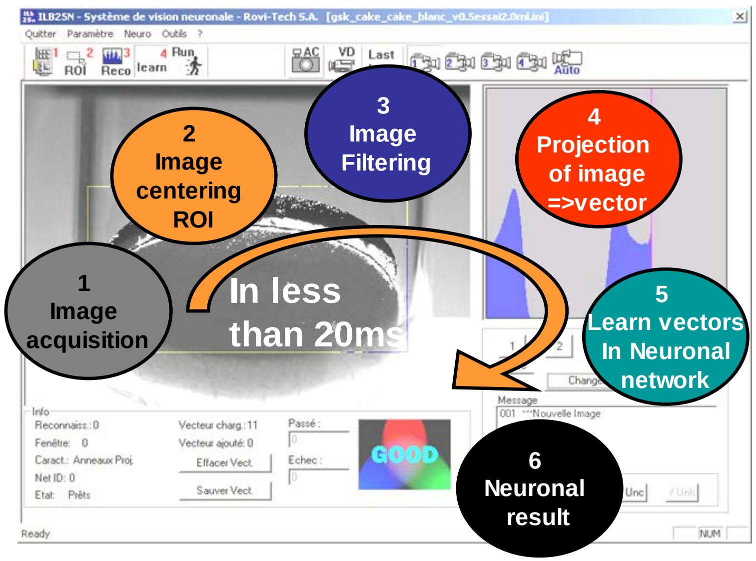
- AVI machines with Digital Cameras pave the way to very high resolutions, higher Detection and lower False rejects
- Higher resolutions will pave the road for deep learning model fine tuning

Machine Learning versus Deep Learning ?



Key Take Away: Machine Learning (SVM) never achieved promising results with parenteral

In 2008, first PDA presentation with Neural Network for Lyo cakes (R. Veillon)

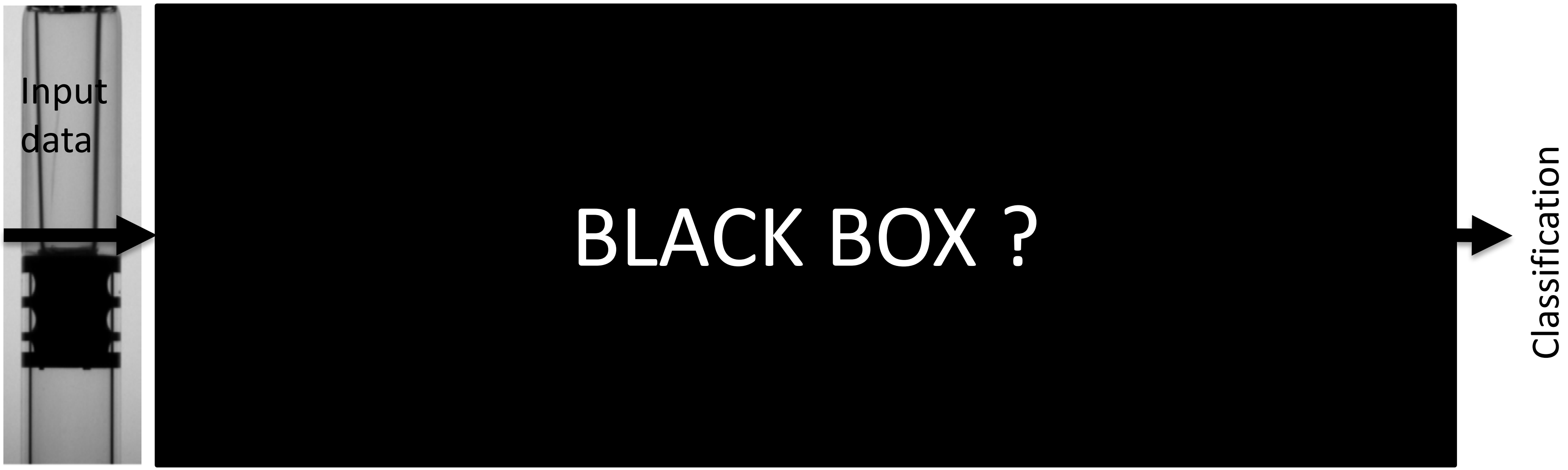


- First concept of Neural network were hybrid between classical vision and NN use.
- They were based on feature extraction in image. (8bit vector projection)
- This gave too much information loss.
- Worked fine with obvious defects but not ability to detect particles and cracks.

Key Take Away:

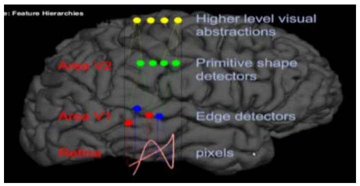
- Machine Learning (SVM) never achieved promising results with parenteral
- Neural network with vectorization gave limited results

• Basic concept of neural network



<u>Key take away :</u> Neural Network can be compared to multiple layers of wired nodes with adjustable weight (potentiometers) in order to link an input signal (image) to an output response (classification)

Comparison MVI and Convolutional Neural Network



- ✓ 1 real object that was trained to the operator = supervised learning
- ✓ Observation; concentration / light/ fatigue….
- ✓ Cones in retina activated
- ✓ Image projected in V1 area of brain for detection angles/ edges /contours
- ✓ Area V2 of brain to detect gross forms/shapes
- ✓ Area V4 -V5 for forms more abstract
- ✓ Activation memory area

➔ **Object Identification + classification**

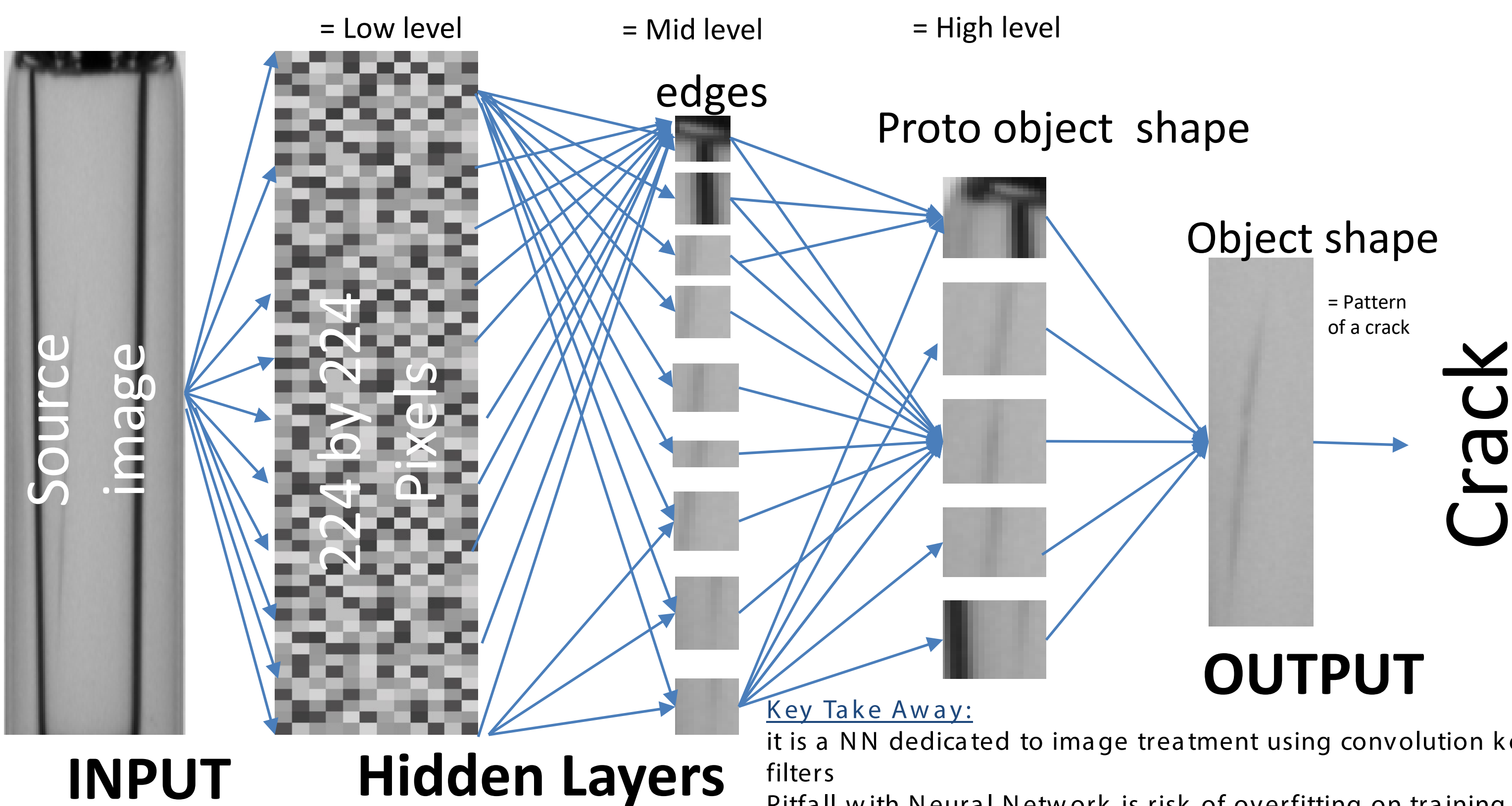
- ✓ 1 real object that was trained to the machine = supervised learning
- ✓ Image camera capture
- ✓ Presentation of image in 1st layer of neuron = number of pixels
- ✓ Each part of image is sent to other layers of neuron that are interconnected, adjustment of coefficients to match elements as best (weight+bias)
- ✓ Last neuron layer to classify object

➔ **Object Identification + classification**

Key Take Away:

- CNN deep learning mimic human brain with better performance on repetitive task
- But it is Not Artificial 'Intelligence'

• What is a Convolution Neural Network (CNN) ?

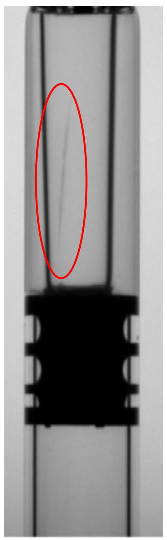


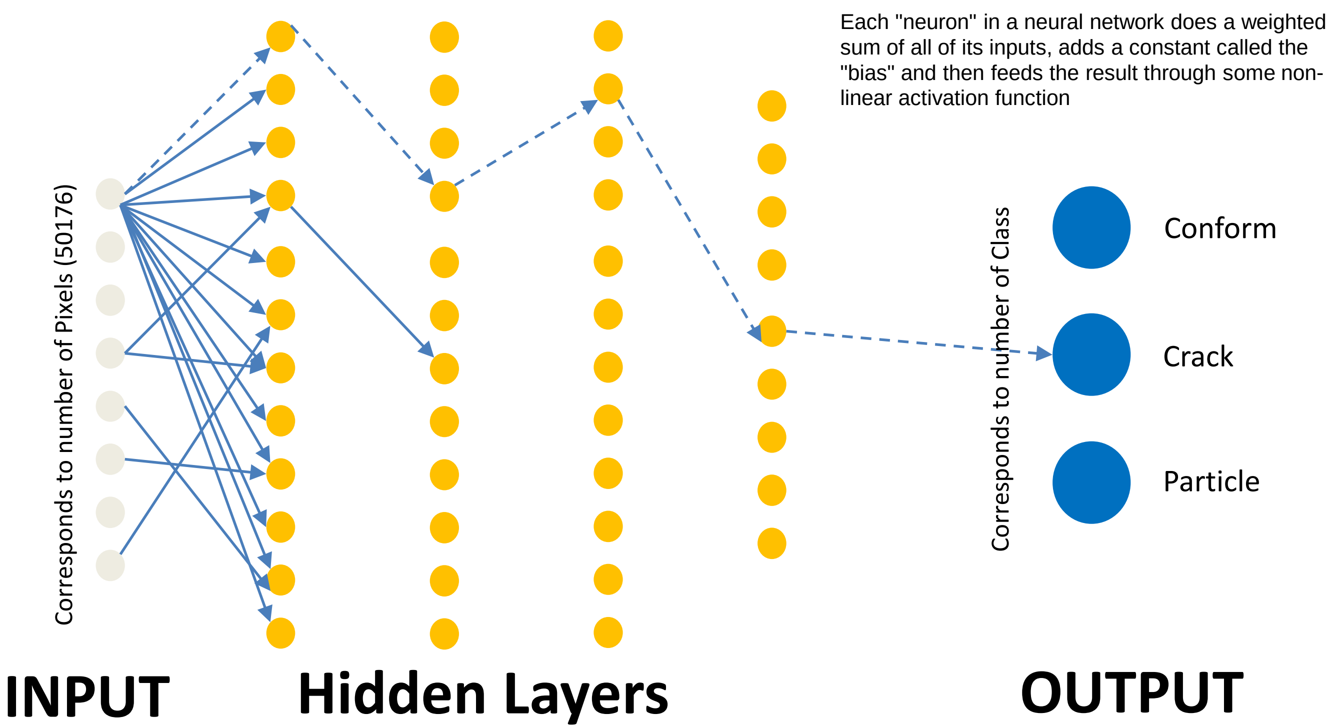
Key Take Away:

it is a NN dedicated to image treatment using convolution kernel filters

Pitfall with Neural Network is risk of overfitting on training images

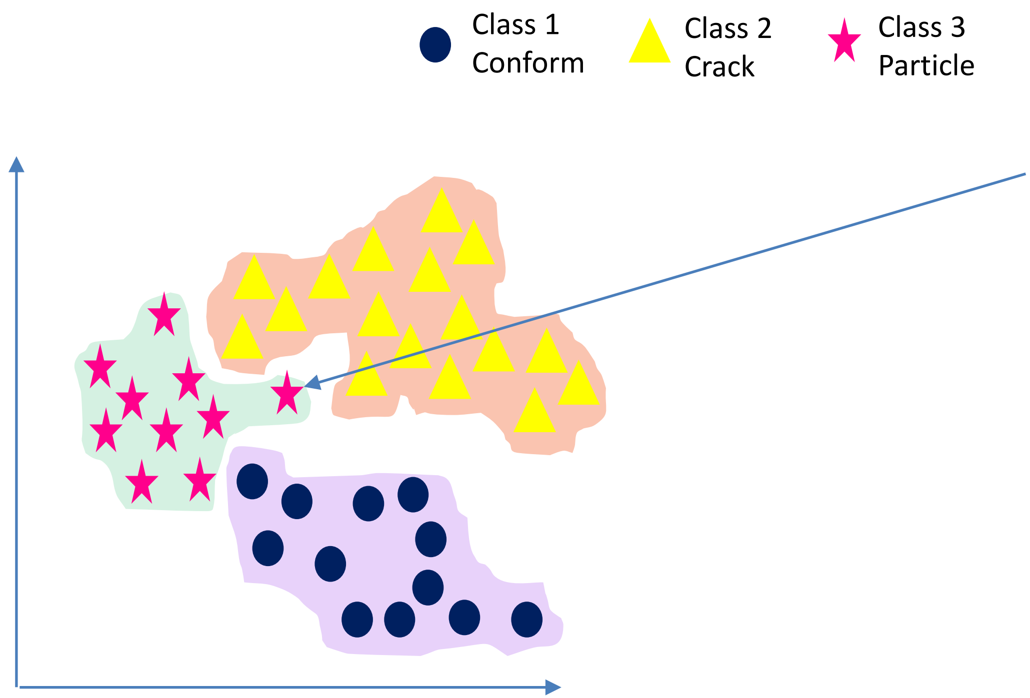
• How Convolution Neural Network (CNN) work?





Key Take Away: Goal: To learn a classification model from the data that can be used to predict the classes of new (future, or test) cases/instances.

• Risk of overfitting with training data



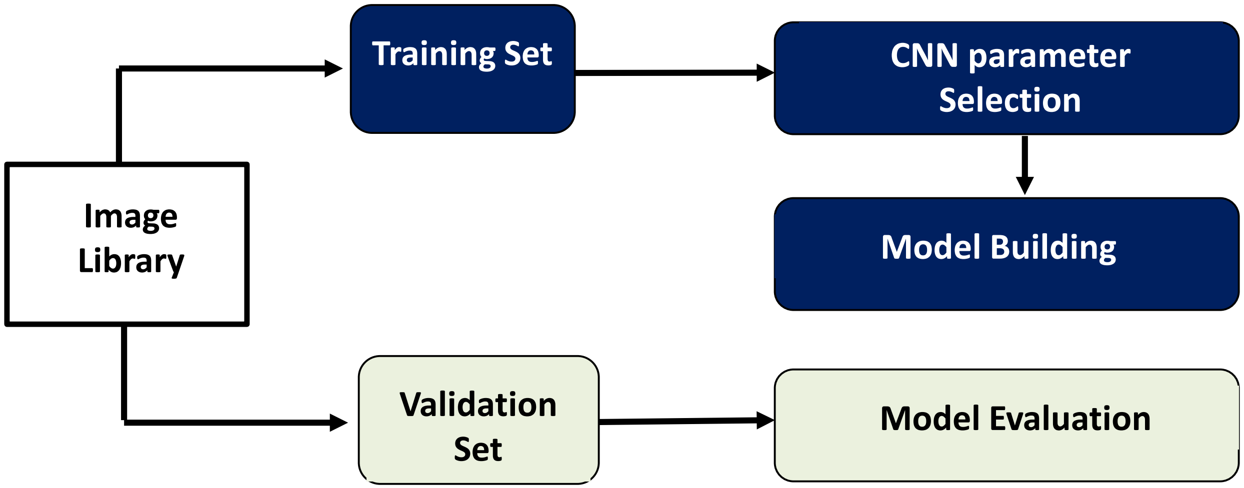
- If this image in Class 3 is an outlier (atypical to this particle class) then the model tries to adapt and this creates overfitting

- This schematic representation is a 2D representation of 3 class
- A Neural Network is a N dimension vectorization of the 3 classes so it is prone to more overfitting

- Need to develop special cross validation strategy (cross entropy) to control any overfitting

Key Take Away: to avoid overfitting there must be cross validation strategies and use of independent prediction set

• Validation strategy to avoid overfitting



- Supervised learning: classification is seen as supervised learning.
- Supervision: The data (images) are labeled with pre-defined classes (defect categories). Supervised because it is like a teacher.

Key Take Away:

- Supervised learning is a frozen model that stops learning and never evolves
- With multiple independent image sets and use of cross validation => avoid overfitting

Supervised learning WOW

1 images library 3 class: 600 images of syringes on 1 Camera
(200 Conform / 200 Crack / 200 Particle)
+ all images checked visually
+ corresponding labels

1 training set 400 images random selection
(Conform / Crack / Particle)
+ corresponding labels

1 Set validation
random and never
shown to the CNN.
100 images
Without labels

1 Set prediction
random
all 600 images
Without labels

Key Take Away:

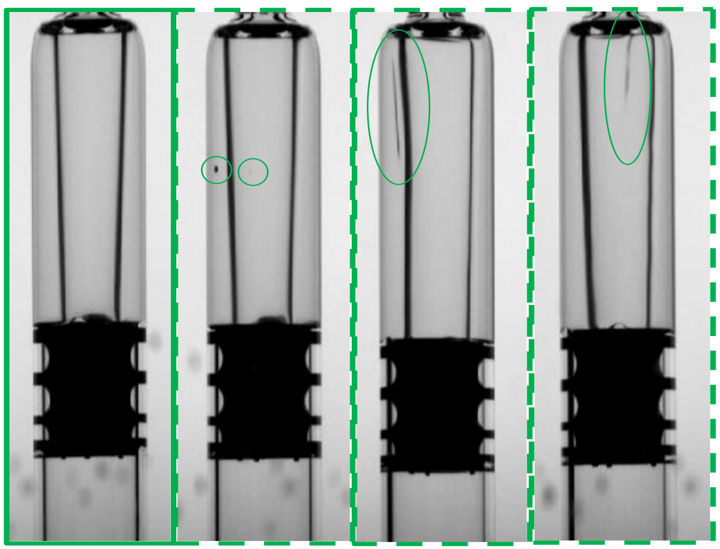
- Library preparation is key
- Need to concentrate effort on image library selection and review

- Problem statement for Deep Learning Proof of Concept
- Detect and classify crack and particle defects

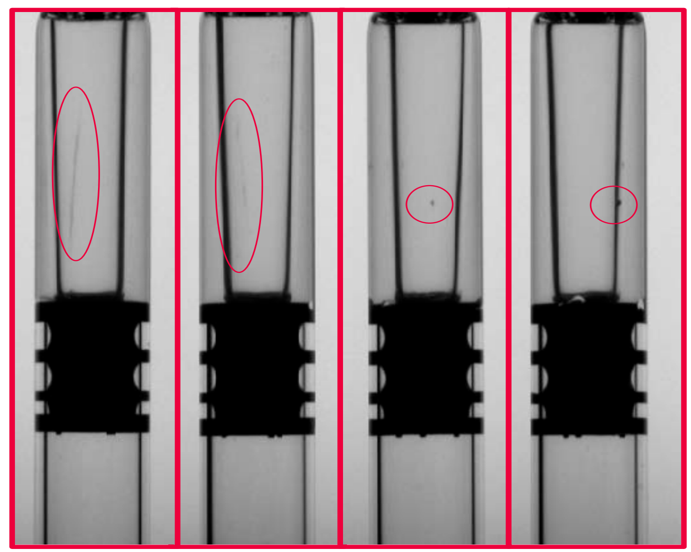
✓ Detect real Defect / conform units with artifacts very borderline

✓ Classify units in 3 class: Conform / Crack / Particle

✓ Do it in less than 15ms



Choice of most boarder line conforms



Choice of most difficult defects

Wide Design space

Wide Design space

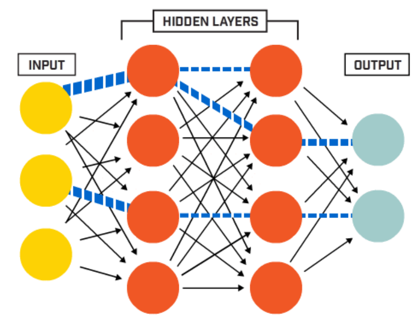
Key Take Away:

- Choice of most difficult images very barderline to this first Proof of Concept
- Conform have many artifacts / Defect are very small and similar to conform

• First Proof of Concept
• with real AVI Defects images

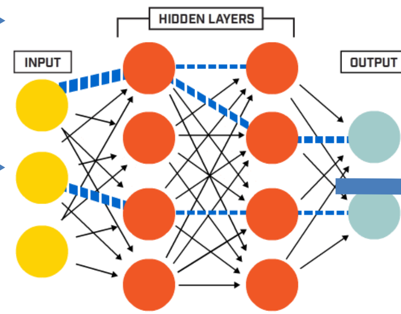
1 New images library 3 class:
1224 images of syringes on 1 Camera
(645Conform / 234Crack / 345Particle)
+ corresponding labels

1 training set 400 images
(Conform / Crack / Particle)
+ corresponding labels



pre-trained CNN on 1000 image class
(ImageNet.org)

1 New Set validation
never shown to the
CNN. 122 images
Without labels



>99% defect detection
<1% False reject
95.1% of labels found
4.9% wrong classification
Cross entropy very low <0.05

Key Take Away:
- With Pre trained CNN on large number of image class, it is possible to re-train on new images.
- Performance is very high >99% and units are classified in 3 class.

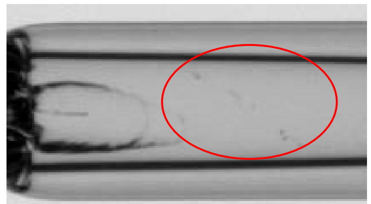
Investigation for miss classification errors

100% defect detection
0% False Reject

95.1% of labels found accurately

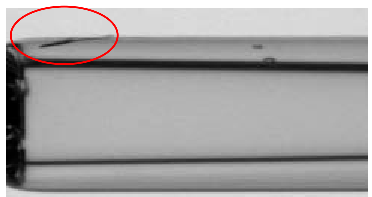
6 defect images seen as defect but with wrong classification
Crack ←→ Particle

Retrained without 6 images = 97.9% accuracy

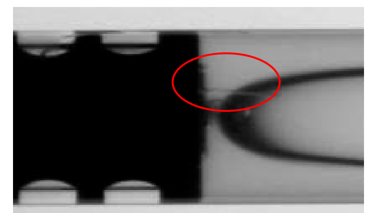


Crack 5 classified as particle

Particle 12 classified as crack

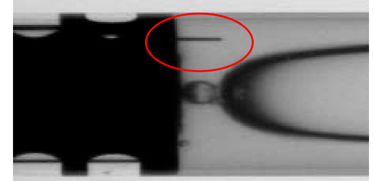


Crack 26 classified as particle



Particle 363 classified as crack

Crack 31 classified as particle



Particle 367 classified as crack

8 images per syringes

min 2 images well classified can balance this

Key Take Away: quality of training set is key, avoid any misleading images. For example when some angle of view can show only one part of crack => considered as particle

Challenge deep learning model with stress test

Deep Learning parameters

- Training Day1 / 2 / 3 Same results 97.9% → repeatability
- Drop in accuracy if less than 100 images per class, more is better 97.9% with 1600 images → Number images
- Change Neural network had significant impact on accuracy → Neural Network
- Source image must be minimum 224 by 224 pixels → Resolution

Deep Learning Model

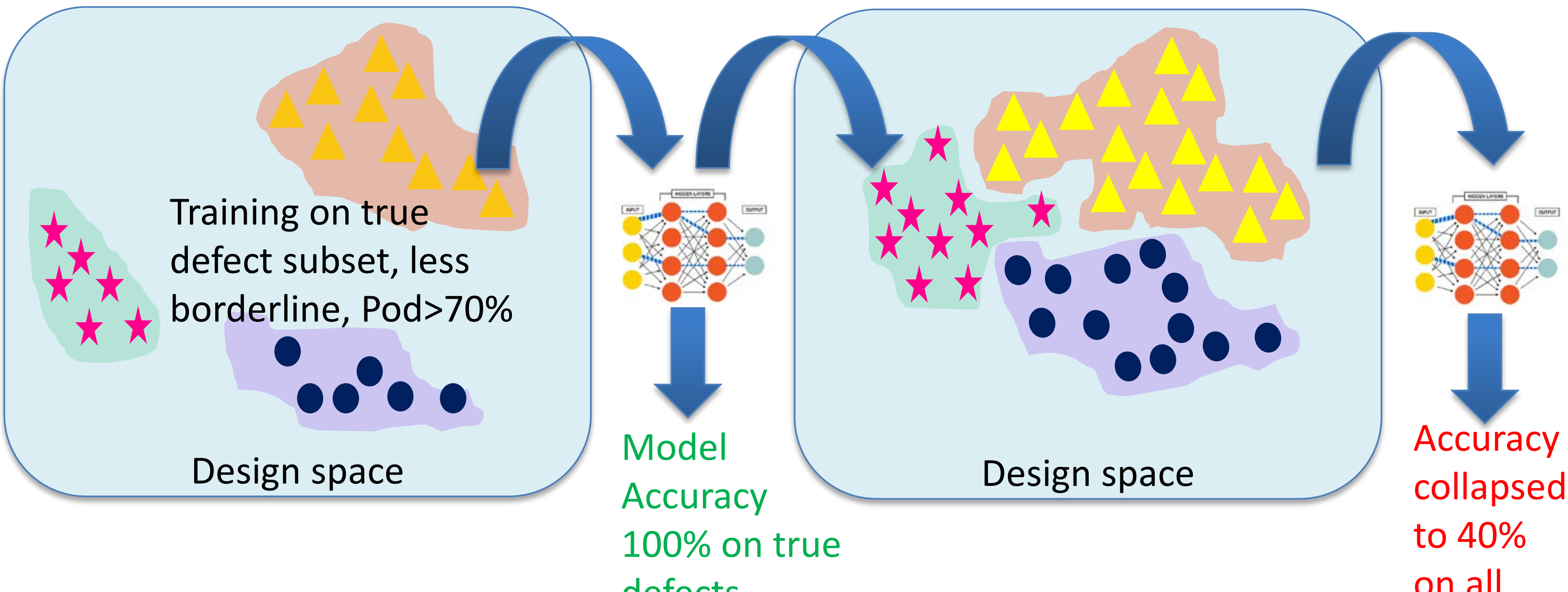
AVI vision parameters

- New image set → New image set (different AVI eq.) had limited effect on accuracy: 96.61% vs 97.9%
- Flip image → Flipping image = not significant impact
- Image moving → Image moving +/- 15 pixels = not significant impact
- Image luminance → Image luminance has significant impact on detection

Key Take Away:

- Stress test are key to evaluate if model can be robust on industrial machines
- As initial source image were taken on industrial AVI machine at fast speed, vision parameters can be controlled
- Parameters of model design are critical and must be defined in project definition

• Coming back to J. Knapp
• Training test with only true defects zone (MVI PoD > 70%)



Key Take Away:

- If training is done only on true defect, deep learning model will detect all true defects with perfect accuracy
- However, if we submit all defect to this model = poor ability (40%) to detect gray zone defect (PoD <70%)
- This shows that Human and Neural Network will have to share same training practices

How Deep Learning will transform AVI ?

Current Industrial vision

- ✓ Ajustement light + optic + image
- ✓ Image Capture

  Images conforming units (kits)

  + defect images (kits)

  + identification defects (logbooks)
- ✓ Preparation of image treatment

  for each camera

  - click and drag software

  - Optimization processing time 1 to 3weeks
- ✓ Ajustment on images / auto ajust.
- ✓ Evaluation on machine
- ✓ Validation (PQ)
- ✓ Go Live to production

More effort in computer vision tools

Deep Learning Proof of Concept

- ✓ Need AVI Machine
- ✓ Ajustment light

  + optic + image
- ✓ Image Capture

  Images conforming units (kits)

  + defect images (kits)

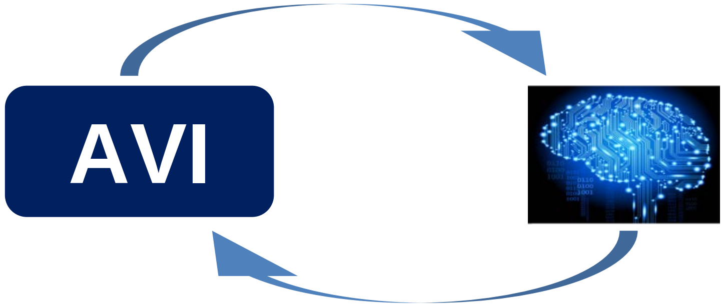
  + identification defects (logbooks)
- ✓ Construction of 2 data bases of image: 1 week
- ✓ learning (training_set) 20 min
- ✓ Evaluation (Prediction_set) 5 min
- ✓ Programmation neural network 1 hour
- ✓ Evaluation & adjustments 1 day
- ✓ Optimization processing time
- ✓ Validation (PQ)
- ✓ Go live to production

More effort in image library

Key Take Away:

- Deep learning will benefit to AVI
- Need to develop image libraries, opportunities for PDA to launch initiatives

- How Deep Learning will transform AVI ?
- Future Perspectives



Low false rejects

Machine Learning

Digitalization

Prior Art Deep Learning Parenteral defects

virtual machines

Multispectral vision

Defect classification

Real time Libraries

images libraries

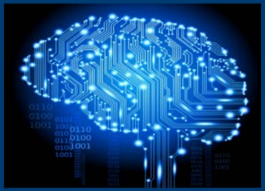
Higher resolution

Conclusion:

- Accurate parenteral defect deep learning is possible.
- Future AVI equipment will encompass digitalization combined with Deep Learning and virtual machines.
- Need for high quality images on robust AVI equipment
- For Validation purpose need to restrict to supervised learning concept.
- Will human reference remain ?
- Need to work on image libraries for parenteral, PDA could lead this initiative.

- Thank you for your attention

- Contact: romain.veillon@gsk.com



- Many thanks to Fernand Koert for providing image library

- ….is there any artificial „intelligence“ behind this ?
  - “…and we call invisible, either what is absolutely – as we consider impossible in other cases -,
  - Or what is visible by its inherent nature, but in fact it may only be hardly visible or invisible »
  - Aristotle, De Anima, Book 2, 10

What do you need ?

- ✓ 1 PC Linux
- ✓ Python + OpenCV
- ✓ Scikit-learn
- ✓ TensorFlow + Keras
- ✓ Some tagged images(many!!)

OpenCV python™

Deep Learning

K + TensorFlow