Mastering AVI

Part9: Future trends in AVI

Instructor Lead: Romain Veillon / Fernand Koert / Sébastien Koch

The Team

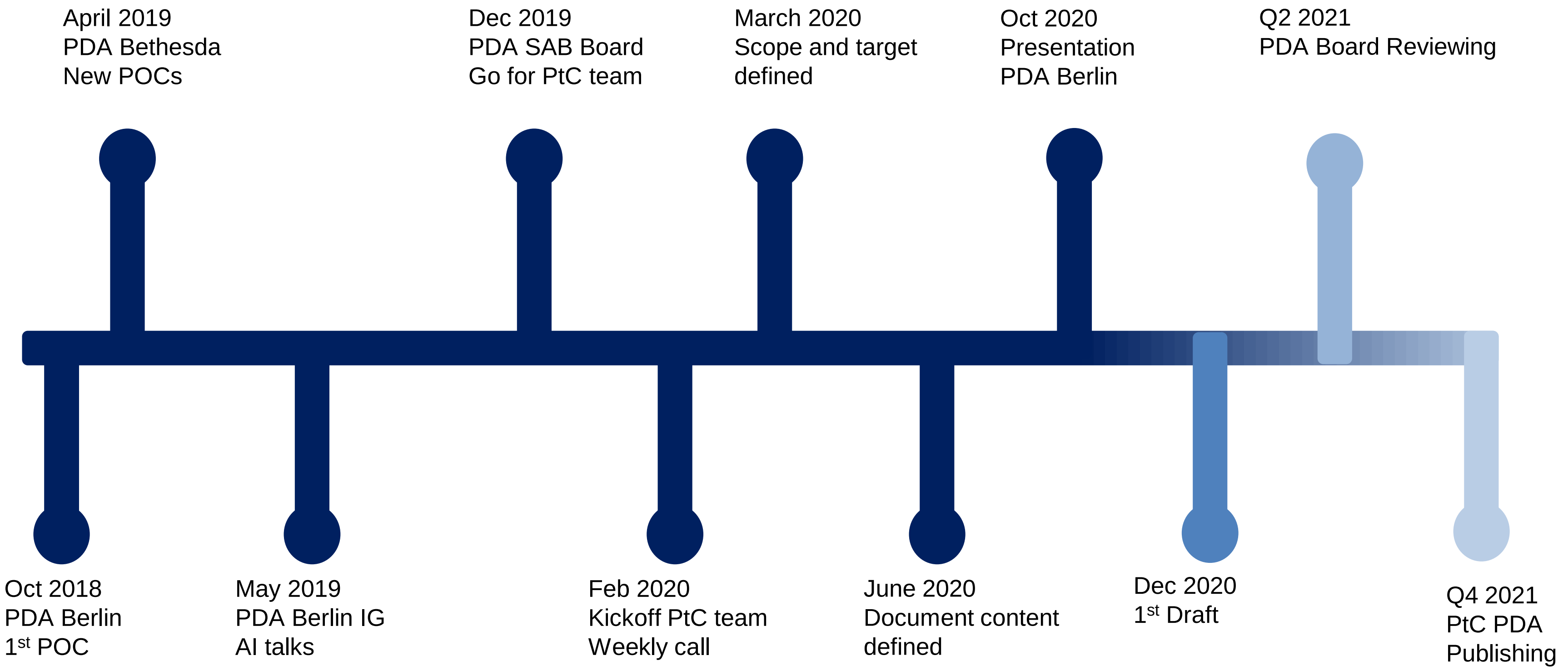
PTC task Force Lead:

- Romain Veillon, GSK Vaccines
- John Shabushnig, Insight Pharma Consulting

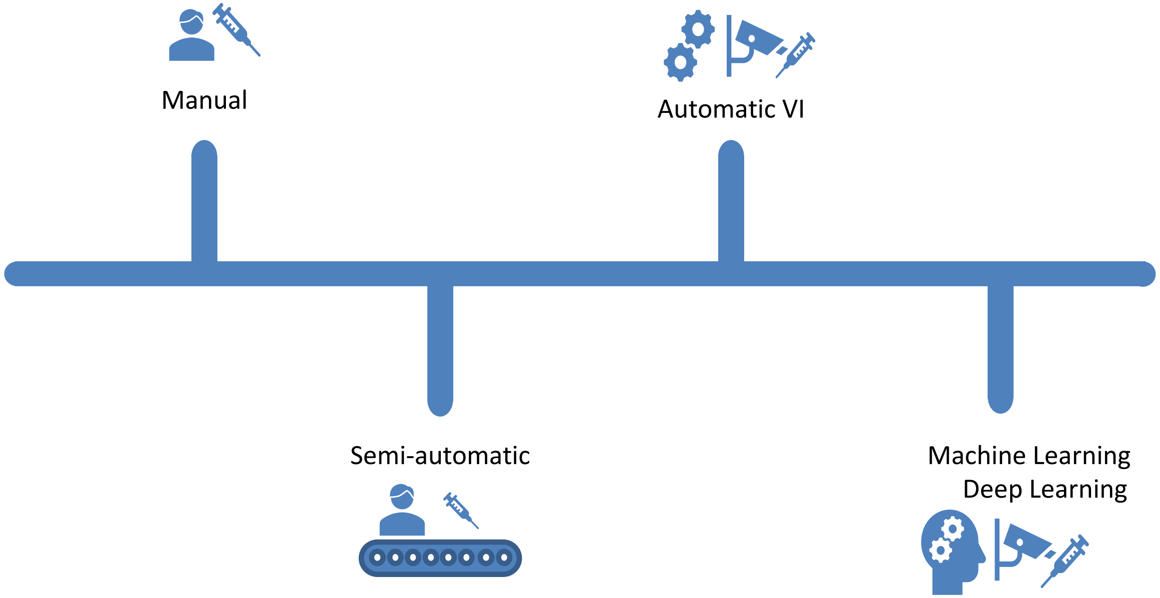
PTC Team Members

- Janie Miller, PDA
- Matthieu Duvinage, Data scientist GSK
- Soto, Manuel, Data Scientist Amgen
- Brian Turnquist, Data scientist Boon Logic
- Christian Eckstein, Data scientist MVTec
- Andrea Sardella, Stevanato Group
- Lars Aabye-Hansen, Novo Nordisk A/S
- Zheng Li, Genentech
- Jorge Delgado Torres, Amgen
- Chady Elahmad, MVTec

Timeline

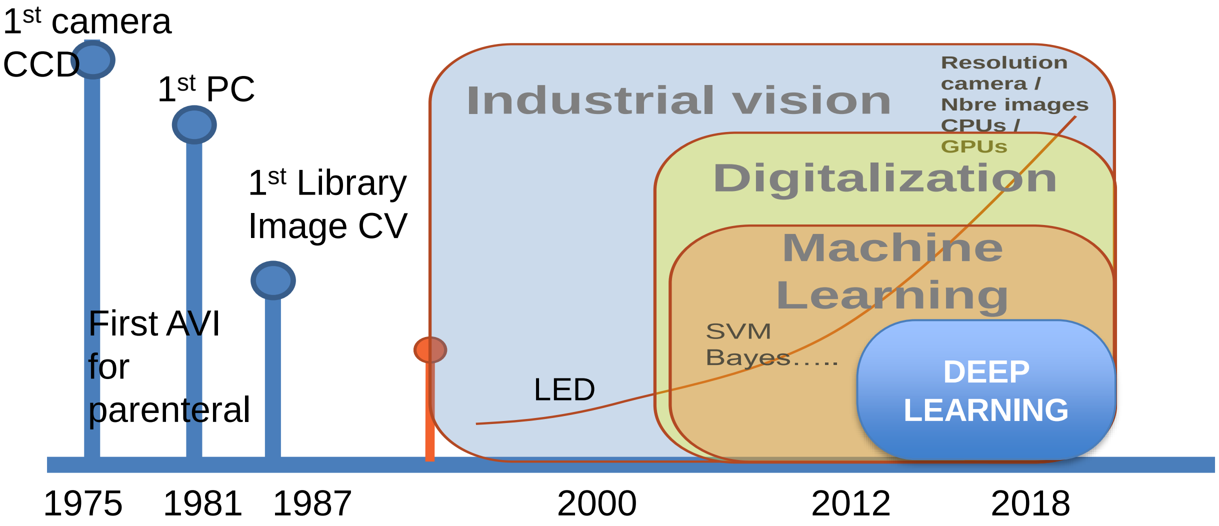

Oct 2018
PDA Berlin
1st POC
April 2019
PDA Bethesda
New POCs
May 2019
PDA Berlin IG
AI talks
Dec 2019
PDA SAB Board
Go for PtC team
Feb 2020
Kickoff PtC team
Weekly call
March 2020
Scope and target
defined
June 2020
Document content
defined
Oct 2020
Presentation
PDA Berlin
Dec 2020
1st Draft
Q2 2021
PDA Board Reviewing
Q4 2021
PtC PDA
Publishing

Where do we come from ?



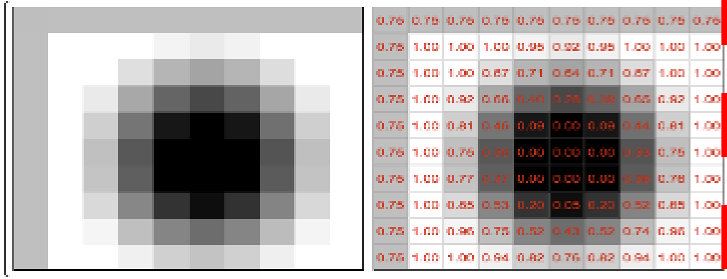
- **AVI is a young, but maturing technology**
- **Many step changes over the last 30 years,**
- **next step change is AI**

AVI is a fast-evolving technology



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

• What is a digital 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	0.82	0.66	0.40	0.28	0.38	0.65	0.82	1.00
0.75	1.00	0.81	0.46	0.08	0.00	0.08	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	0.78	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	0.86	0.75	0.52	0.43	0.52	0.74	0.86	1.00
0.75	1.00	1.00	0.94	0.82	0.76	0.82	0.94	1.00	1.00

1 particle image **Image with grey levels…** **Digital Image = matrix grid of figures in X and Y**

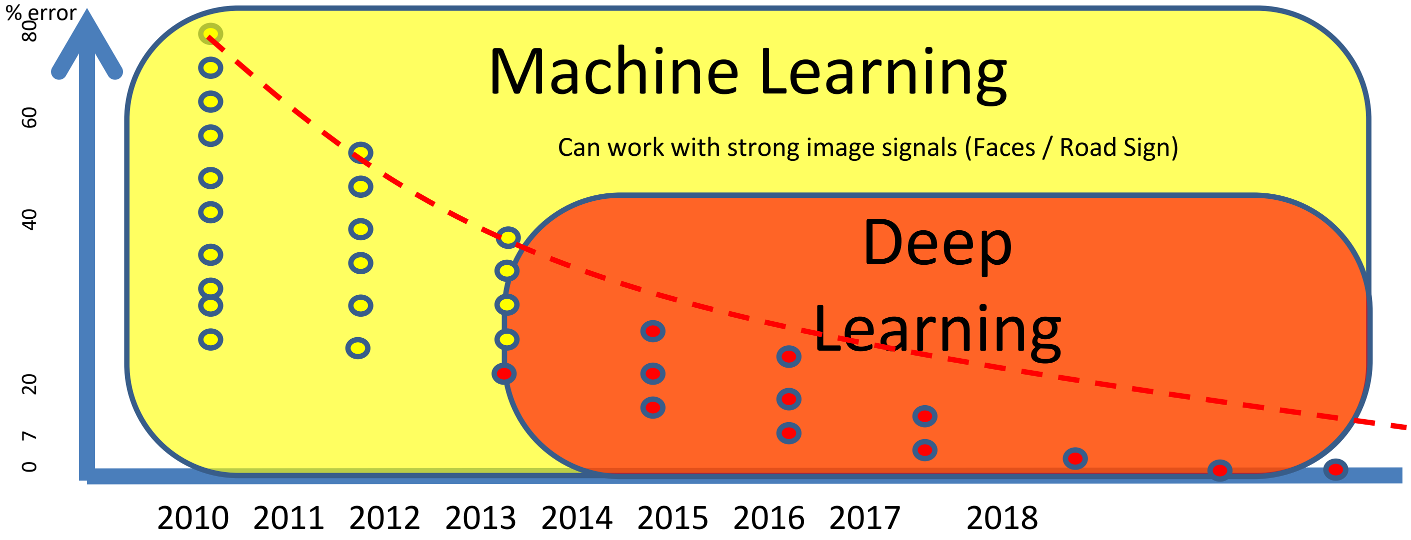
In computer vision language (python/C++) it is a matrix object:

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

Key Take Away:

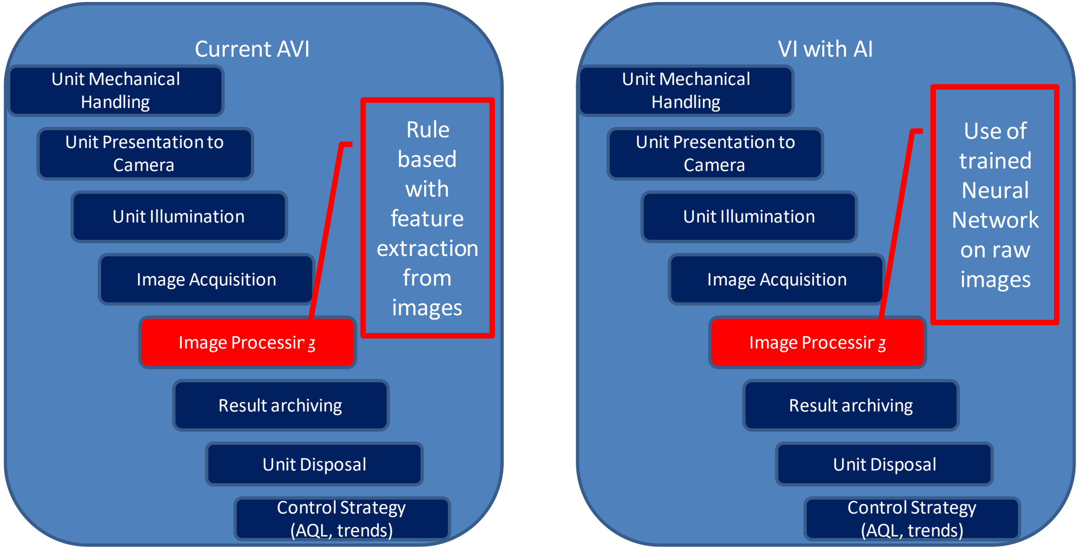
- **Computer vision see only a matrix**
- **That represent spatial distribution of grey levels**
- **Neural Network will work with image matrix**

Machine Learning versus Deep Learning ?



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

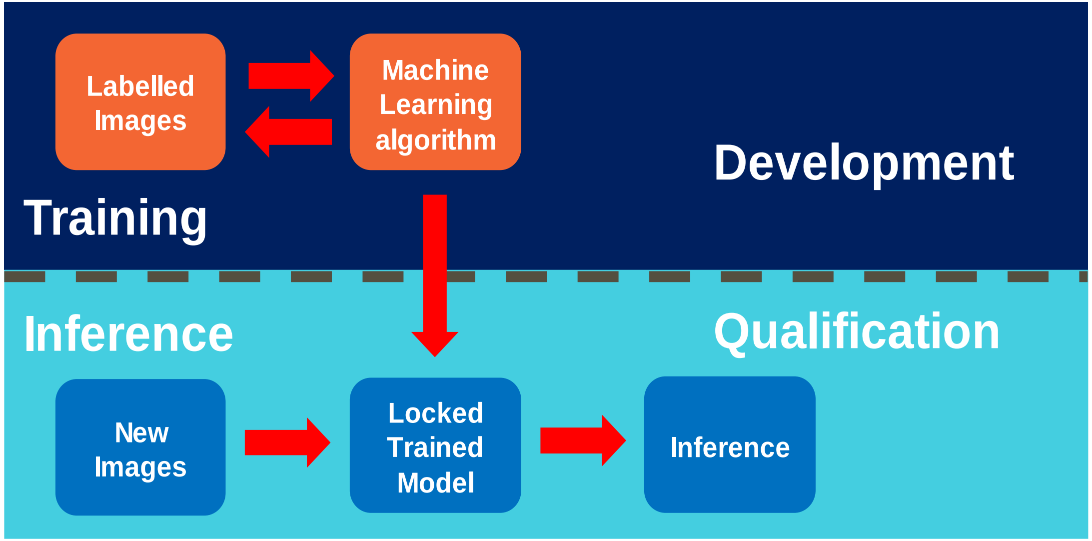
Current AVI versus machine Learning



Point to consider:

- ❑ Scope of change with AI deployment is limited to image processing, all other crucial element remain the same

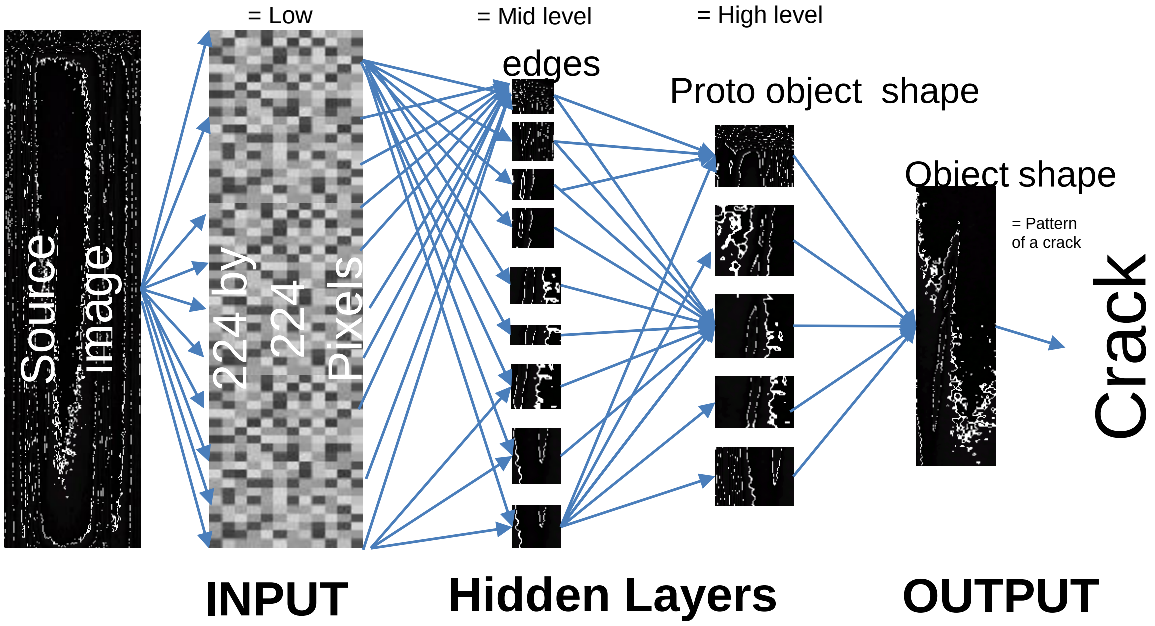
Principle of Deep Learning



Point to consider:

- with Supervised Deep Learning the vision set-up can be frozen and versioned before qualification and later use, it will not evolve, need for versioning control and audit trails

What is a Convolution Neural Network (DNN) ?

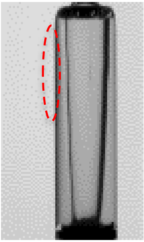


Key Take Away:
it is a NN dedicated to image treatment using convolution kernel filters
Pitfall with Neural Network is risk of overfitting on training images

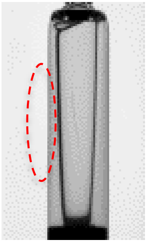
Image Labelling

Example of a binary detection between 2 classes: conform and crack

Conform (is a scratch conform?)

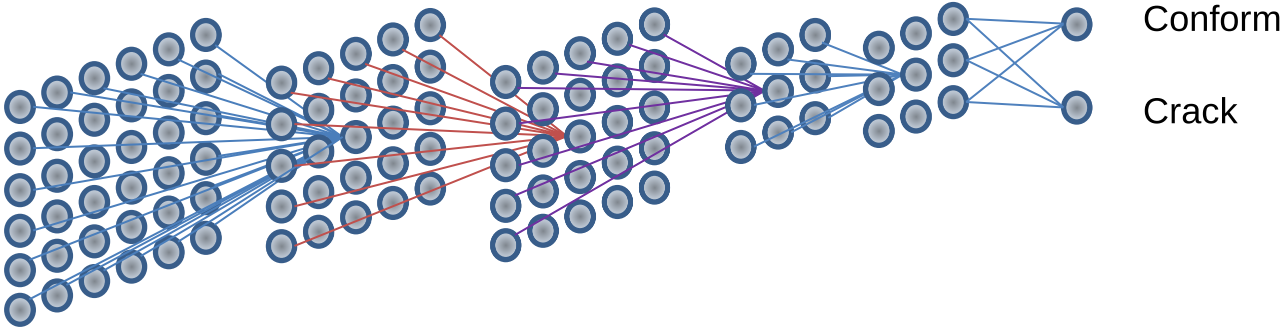


Crack

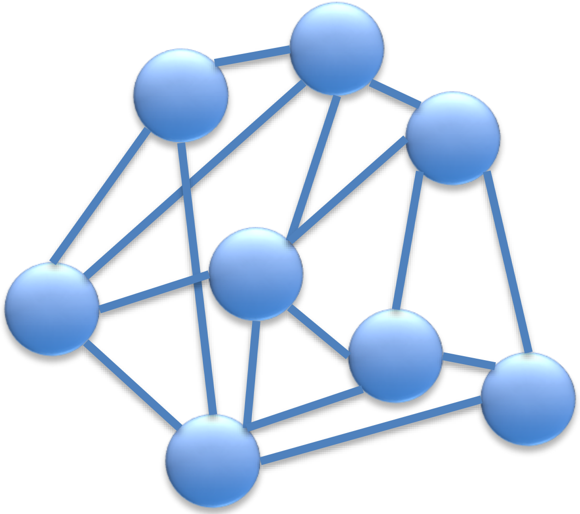


Point to consider:

- ❑ **Labelling defect per class is also very critical.**
- ❑ **Who can label an image ?**
- ❑ **How to document labelling ?**
- ❑ **What are boundaries of conforming class?**



What are main points to consider to explore when moving to AI?



AVI with AI:

- Defect kit design space, explore grey zone
- Design space to the limit of unknown
- Image libraries for conforming unit class
- Defect labelling is a critical steps
- Vision engineers skills will remain
- Data science is new capabilities to develop
- Solid backend GMP IT infrastructure

Labelling process

Defined labelling scenario

Definition of teams:

TEAM 1	TEAM 2	TEAM 3	TEAM 4
Lead Labeler: Mr. W	Lead Labeler: Mr. X	Lead Labeler: Mr. y	Lead Labeler: Mr. Z
Support Labeler: Mr. Wa	Support Labeler: Ms. Xa	Support Labeler: Mr. Ya	Support Labeler: Ms. Za
Mr. Wb	Mr. Xb	Mr. Yb	Mr. Zb
Label job I: C61 crack	Label job I: C61 crack	Label job I: C61 particle	Label job I: C61 particle
Label job II: C61 cake	Label job II: C61 cake	Label job II: C61 closure	Label job II: C61 closure

Note

All members had to plan this labelling within their own workload. 20000 images had to be labeled for just one camera system!

Labelling process, interim result

defect kit	Kit ID	Folder	criticality	Station	defect definition	kit used for vision sations	samples	runs	C131	C132	C133	C42	C72	C61	C62
Comment									Heel view	Heel view	Heel view	Bottom Crimping and neck	cake surface, heel, cake side, sidewall	sidewall, cake side, neck	sidewall, cake side, neck
Number of image per									1	1	1	8	8	8	8
defect kit	Kit ID	Folde	critical	Station	defect description	kit used for vision sations	sampl	runs	C131	C132	C133	C42	C72	C61	C62
crack	CLY01	210720	critical	crack neck	small vertical	C41, C42	3					392		336	no images
crack	CLY02	210720	critical	crack neck	small horizontal	C41, C42	3					392		336	no images
crack	CLY03	210720	critical	crack heel	base small vertical	C61, C62, C71, C72, C121, C13x	3		49	49	49		392	728	392
crack	CLY04	210720	critical	crack sidewall	body big vertical	C51, C52, C61, C62; C72	3						no images	672	336
crack	CLY05	210720	critical	crack shoulder	chip shoulder	C51, C52, C61, C62; C72	3					no images	no images	672	336
crack	CLY07	210720	critical	crack heel	base whole circle	C61; C62, C72, C121, C13x	3		49	49	49		no images	728	392
crack	CLY08	210720	critical	crack sidewall	body small horizontal	C51, C52, C61, C62; C72	3						no images	728	392
crack	CLY09	210720	critical	crack bottom	bottom	C121, C13x	3		49	49	49		no images	336	
crack	CLY10	210720	critical	crack shoulder	bumpcheck shoulder	C51, C52, C61, C62; C72	3					no images	no images	784	448
particle	PLY04		major	moving particle	transparent on cake	C41, C42, C61, C62	8						500	448	no images
particle	PLY02		major	particle top cake	transparent big	C41, C42	8						448	448	
particle	PLY03		major	particle top cake	small black	C41, C42	8						448	448	
particle	PLY01		major	particle below stopper	white particle	C41, C42	8					384		448	
closure	CSLY01		critical		Sidewall	C61								120	
closure	CSLY02		critical		Sidewall	C61								120	
closure	CSLY03		critical		Sidewall	C61								120	
closure	CSLY04		critical		Sidewall	C61								120	
closure	CSLY05		critical		Sidewall	C61								120	
closure	CSLY06		critical		Sidewall	C61								120	
closure	CSLY07		critical		Sidewall	C61								120	
closure	CSLY08		critical		Sidewall	C61								120	
closure	CSLY09		critical		Sidewall	C61								120	
closure	CSLY10		critical		Sidewall	C61								120	
cake	FLY01		critical	melted cake	Sidewall	C61								312	
cake	FLY02		critical	liquid	Sidewall	C61								312	
cake	FLY03		major	half moon	Sidewall	C61								312	
cake	FLY04		major	peaked	Sidewall	C61								312	
cake	FLY05		major	expanded / inflated	Sidewall	C61								312	
cake	FLY06		major	retracted	Sidewall	C61								312	
cake	FLY07		major	product in the neck	Sidewall	C61								312	
cake	FLY08		major		Sidewall	C61								312	

main view
contribute
not seen

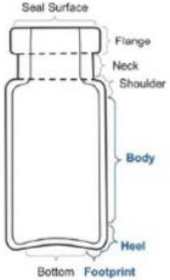
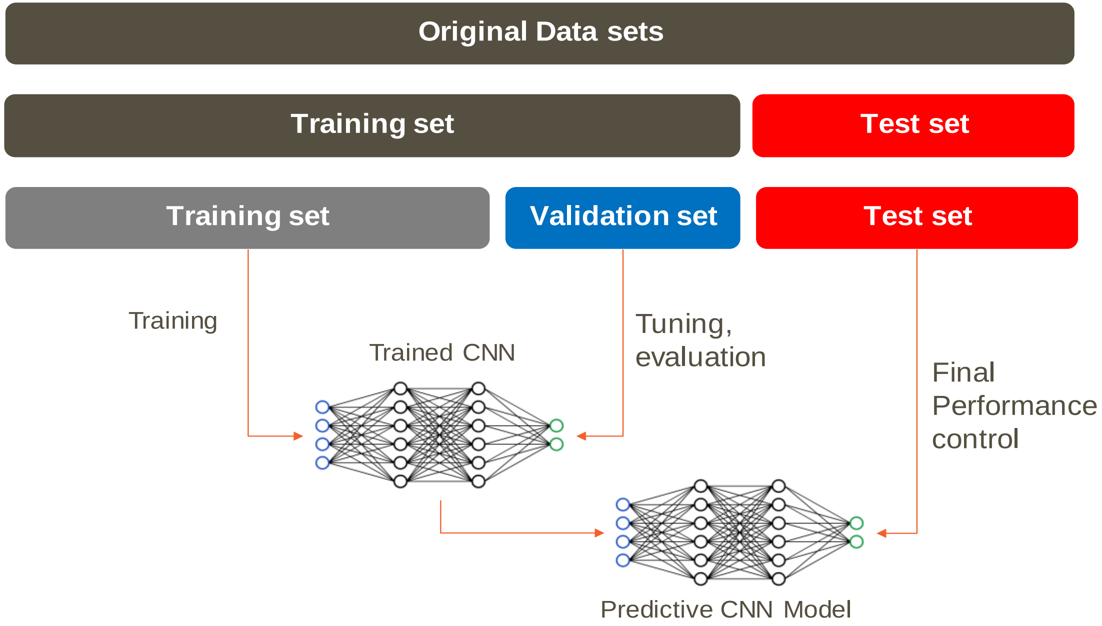


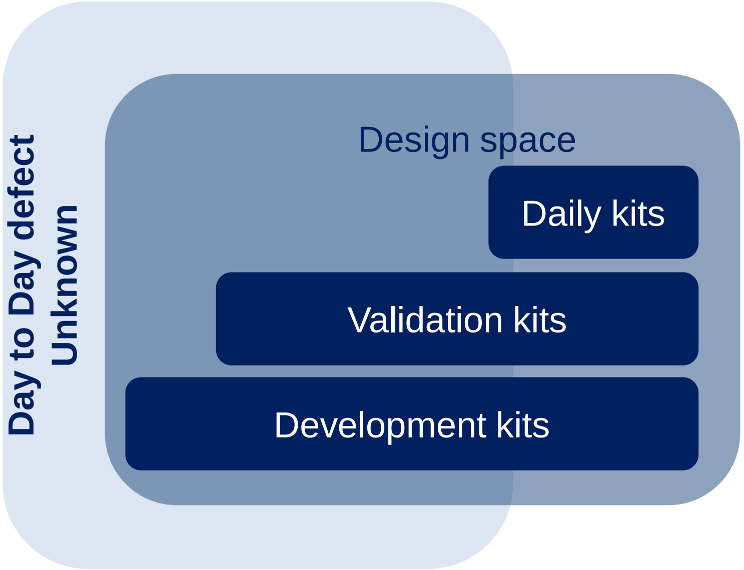
Image test sets



Point to Consider:

- ❑ Number of image per class is critical,
- ❑ the independence of training, validation and test set is critical,
- ❑ audit trail could prevent data leakage
- ❑ Proportion of each
- ❑ Number of images per class
- ❑ Augment image conditions
- ❑ Try to generalize from input images

Defect Design space

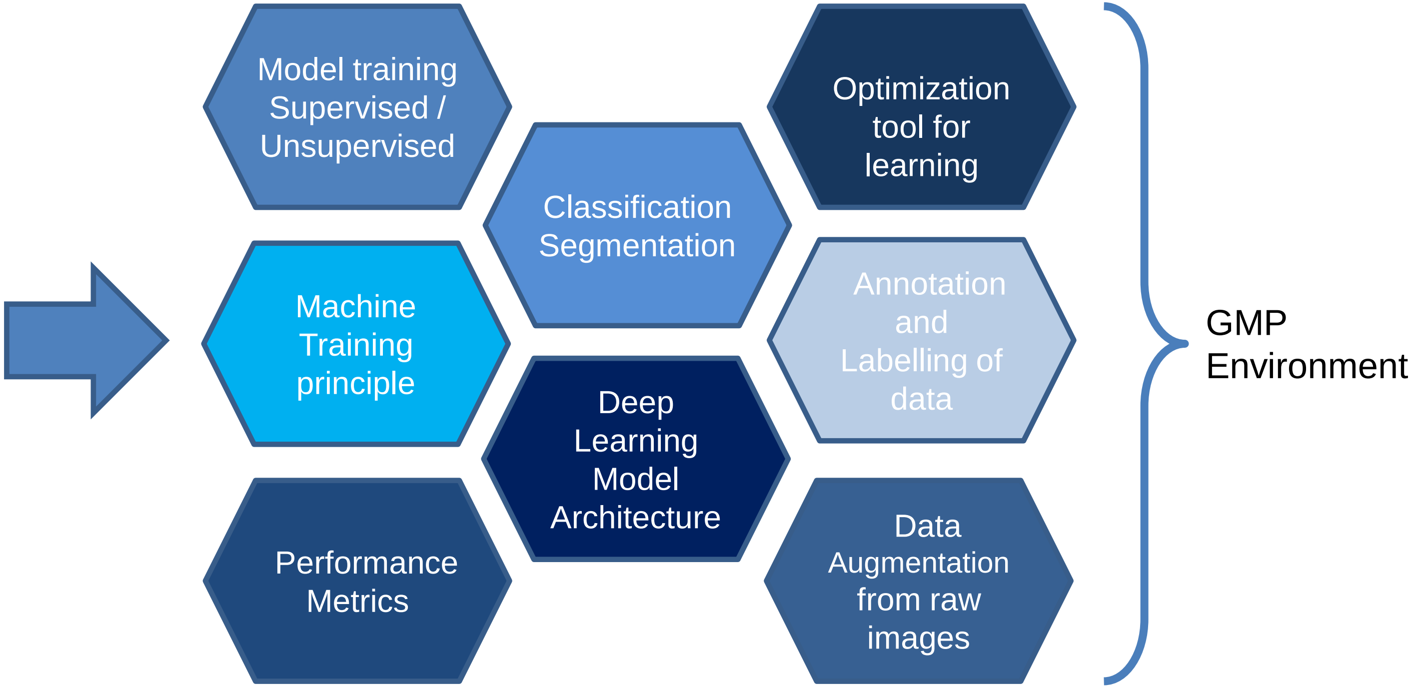


Point to Consider:

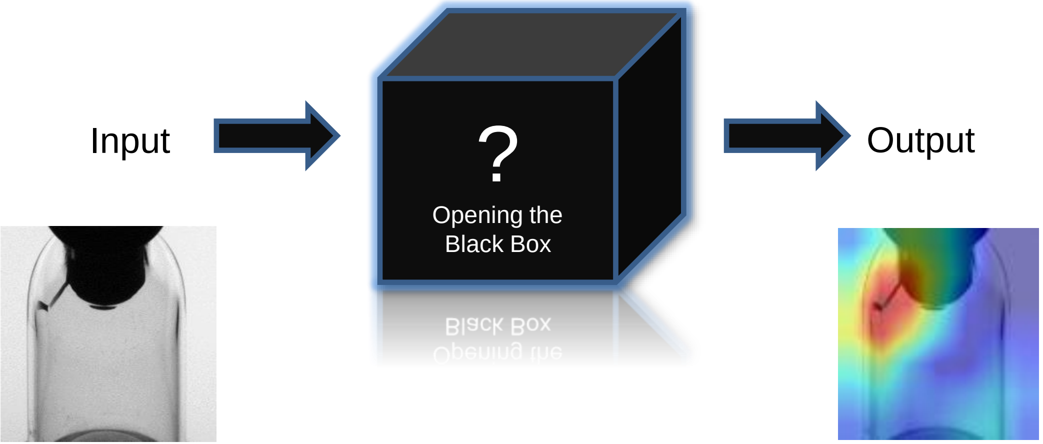
- ❑ **Risk to overfit on specific defect types and to have poor ability to the unknown**
- ❑ **The Design space should be extended to the maximal polymorphism of defect,**
- ❑ **The true defect zone may be too restrictive for AI development and training**
- ❑ **Need to explore to limit of the unknown, need more development kits to feed digital libraries**

Key Data Science element to cover

Data Scientist Team
For solid scientific advisory



Why Visualization of results is so critical ?



Point to consider:

- ❑ it is key to report the results of AI with some visualization tools like heatmap, bounding box or segmentation to well document and give transparency
- ❑ Segmentation (SSD) can show where Deep Learning found a defects

Validation of AI applied to VI

Desktop Qualification of ML

- Desktop performance qualification
- Strong metrics documentation

On line performance Qualification

- Digital Image testing
- Real defect kit testing

Life cycle Continuous improvement

- Data gathering for improvement
- Continuous improvement with new validation of performance

Some References

- **Compendia**
  - ✓ USP<790> visual inspection, Jan 2014
  - ✓ USP<1790>, visual inspection companion chapter (guidance), Aug 2017
  - ✓ Ph. Eur., JP Visual inspection
- **Articles:**
  - ✓ PDA Journal all Knapp Articles from [1980-1992]
  - ✓ J. Shabushnig, PDA 2014, PDA survey visual inspection;
- **Books:**
  - ✓ Computer Vision: Algorithms and applications Richard Szeliski 2011
  - ✓ Computer Vision: Detection, Recognition and reconstruction, Roberto Cipolla 2010
  - ✓ Particle for Parenteral, J. Shabushnig, R. Cherris, PDA 2016,
- **Lectures/Web-resources:**
  - ✓ Standford Univ. CA: Bernd Girod, Digital Image Processing
  - ✓ Python OpenCV documentation

- Acknowledgements
- Fernand Koert / Romain Veillon / Aurélien/Sebastien Koch
- PDA Europe
- PDA visual inspection committee