Federated Conference on Computer Science and Information Systems
FedCSIS 2019
September 1-4
Leipzig, Germany

Deep Data and Big Learning: More quality data for better knowledge

Francisco Herrera

Data Science and Computational Intelligence Andalusian Research Institute

Dept. of Computer Science and A.I.

University of Granada, Spain

Email: herrera@decsai.ugr.es

http://sci2s.ugr.es

UNIVERSIDAD DE GRANADA

DaSCI

Deep Data and Big Learning:
More quality data for better knowledge

Outline

1. **Deep Data:** Towards quality data
2. **Big Learning**: CNNs with quality data
3. Case of study 1: **MNIST**
4. Case of study 2: **Whale detection**
5. Case of study 3: **Knife detection**
6. Concluding Remarks: **More quality data for better knowledge**

Deep Data and Big Learning:
More quality data for better knowledge

Outline

1. **Deep Data:** Towards quality data
2. **Big Learning**: CNNs with quality data
3. Case of study 1: **MNIST**
4. Case of study 2: **Whale detection**
5. Case of study 3: **Knife detection**
6. Concluding Remarks: **More quality data for better knowledge**

Towards quality data

Quality decisions

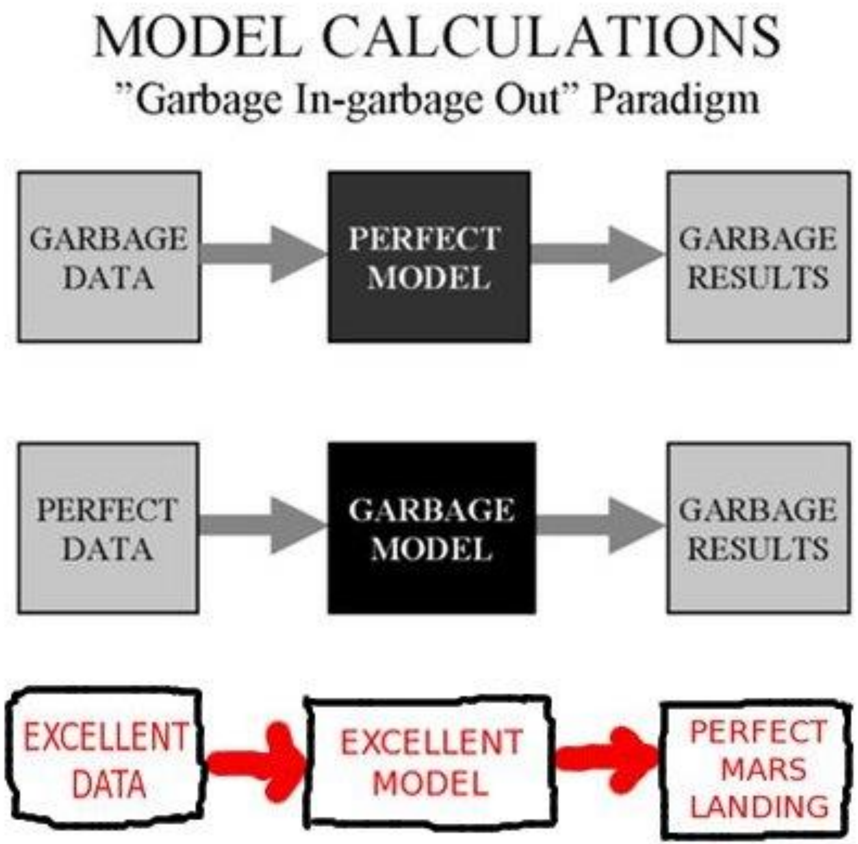
("quality models/patterns/rules")

are based on

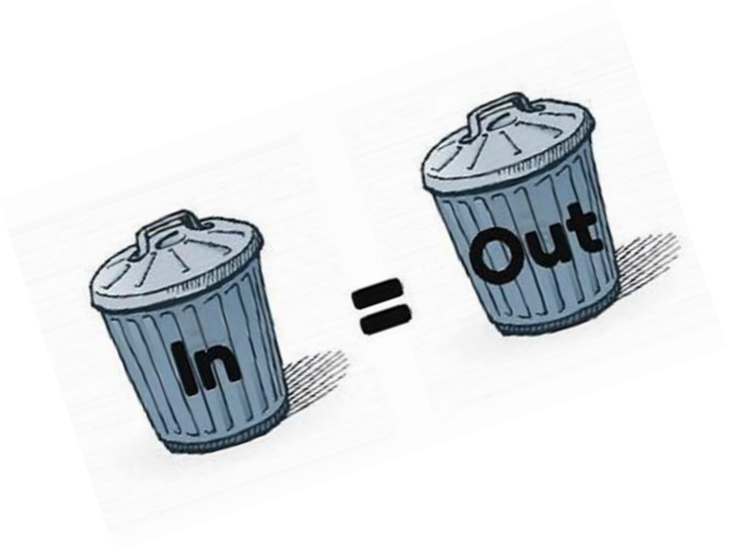
Quality Data!

Towards quality data

Quality decisions
(“quality patterns/rules”)
are based on Quality Data!



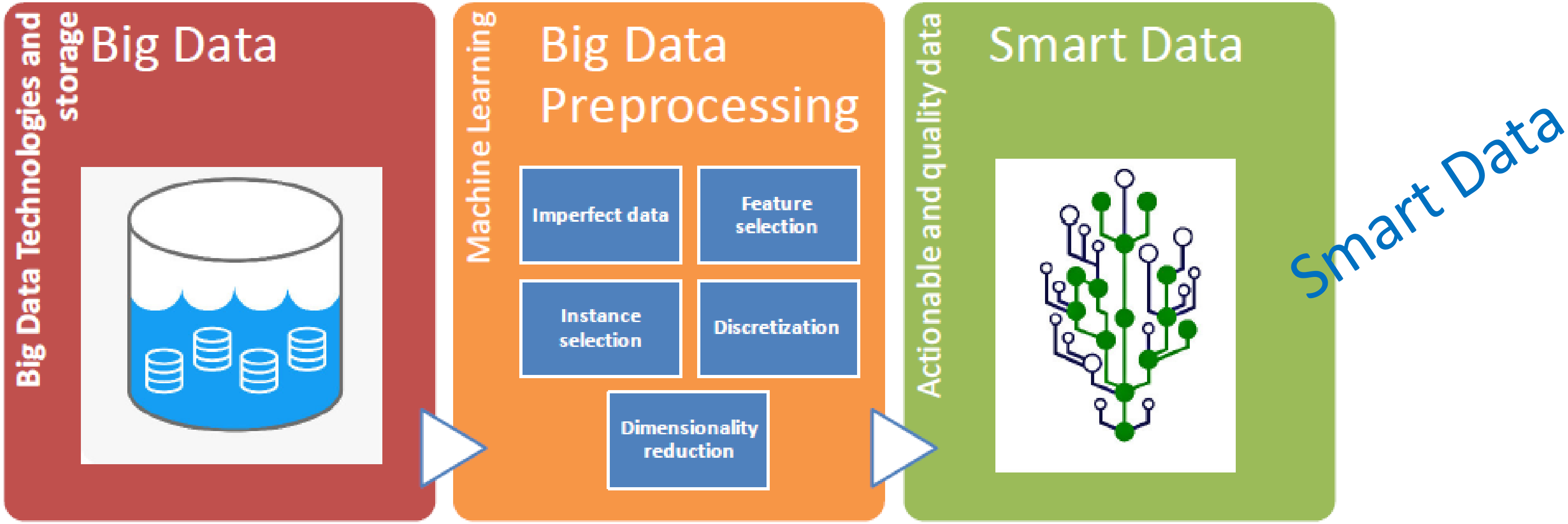
More quality data for better knowledge



Towards quality data

More quality data for better knowledge

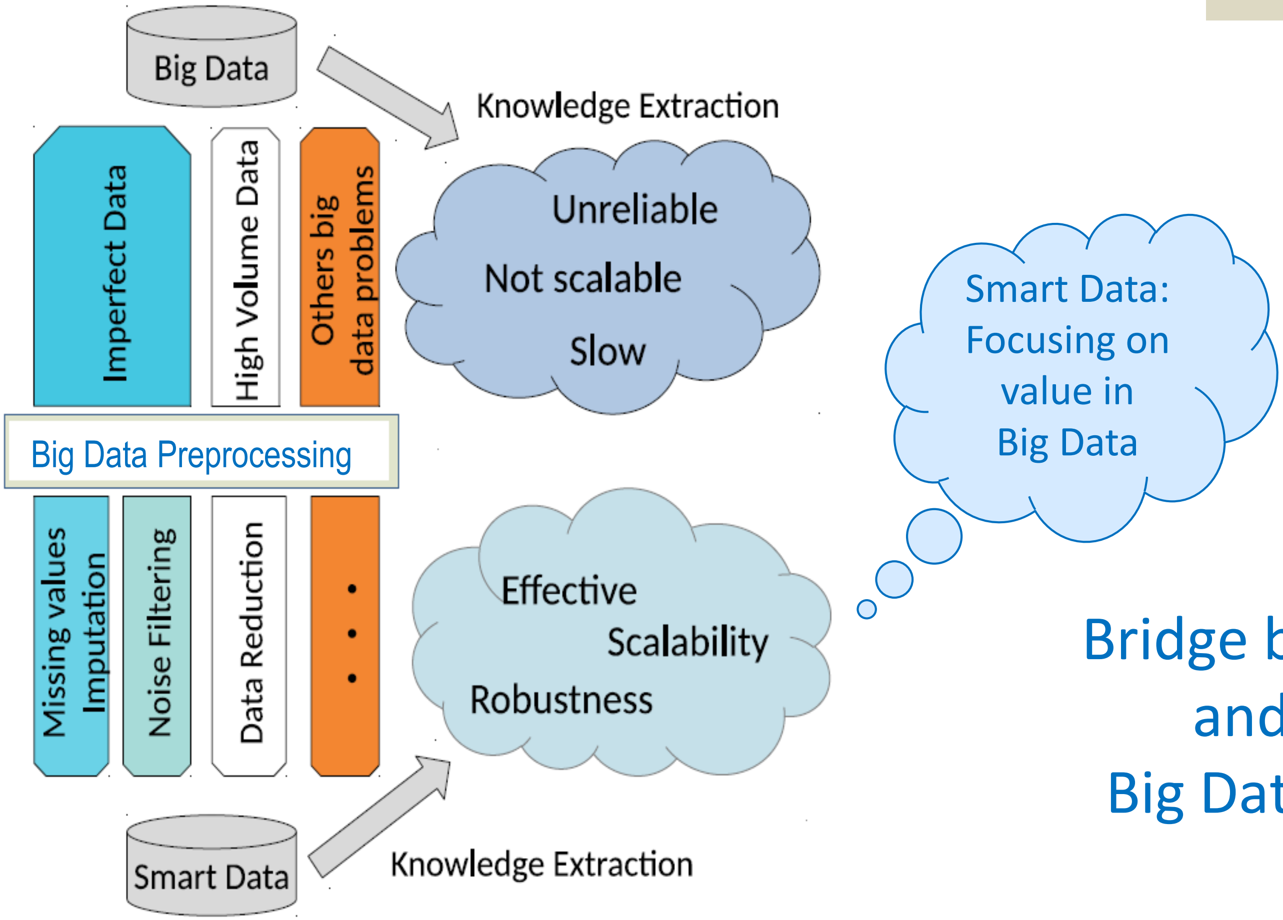
Big data preprocessing is the key to transform raw big data into quality and smart data.



Transforming big data into Smart data: An insight on the use of k-Nearest Neighbours algorithm to obtain quality data
I. Triguero, J. Maillo, D. García, S. Garcia, F. Herrera
Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2019. Open access

Towards quality data

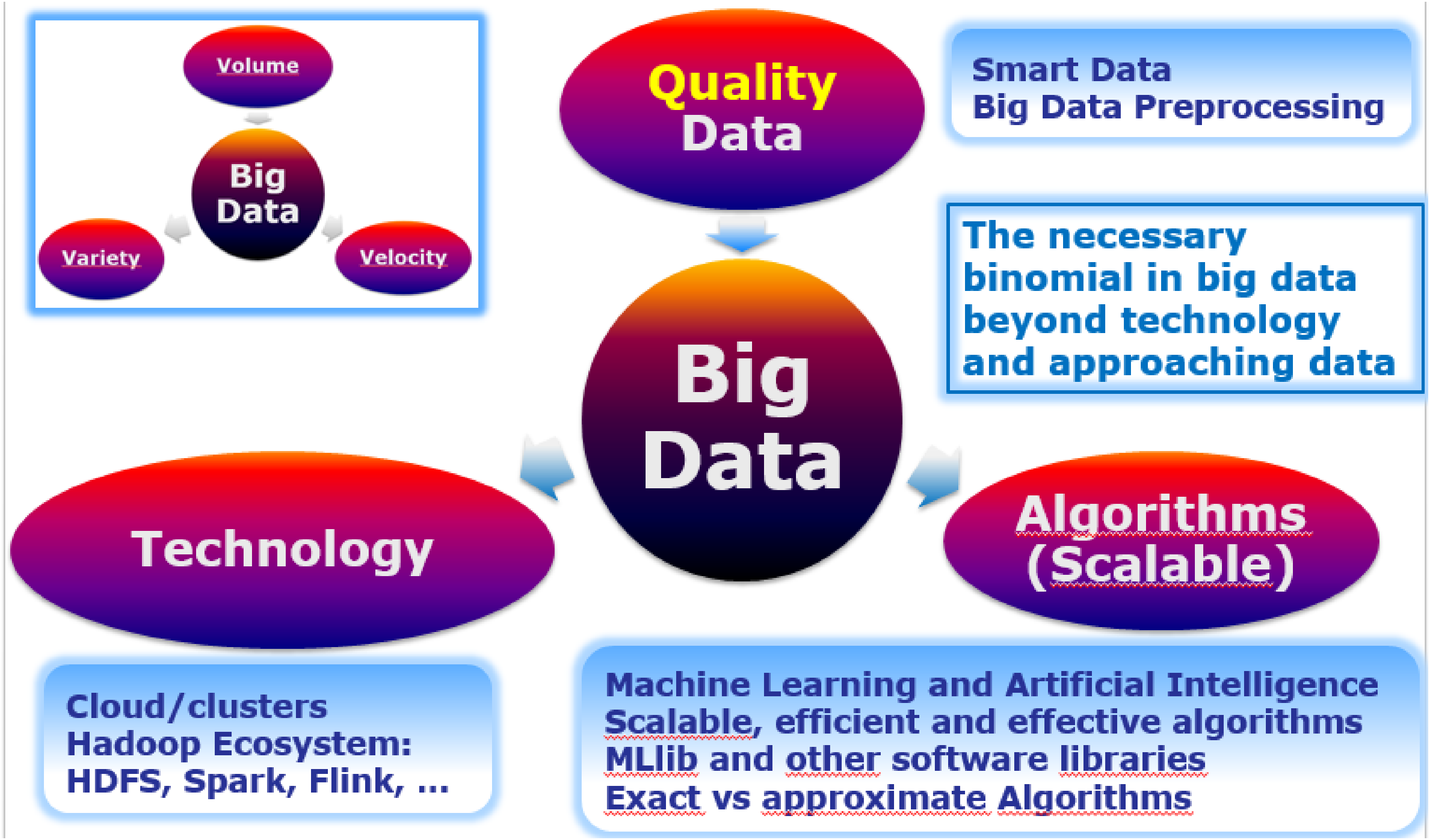
More quality data for better knowledge



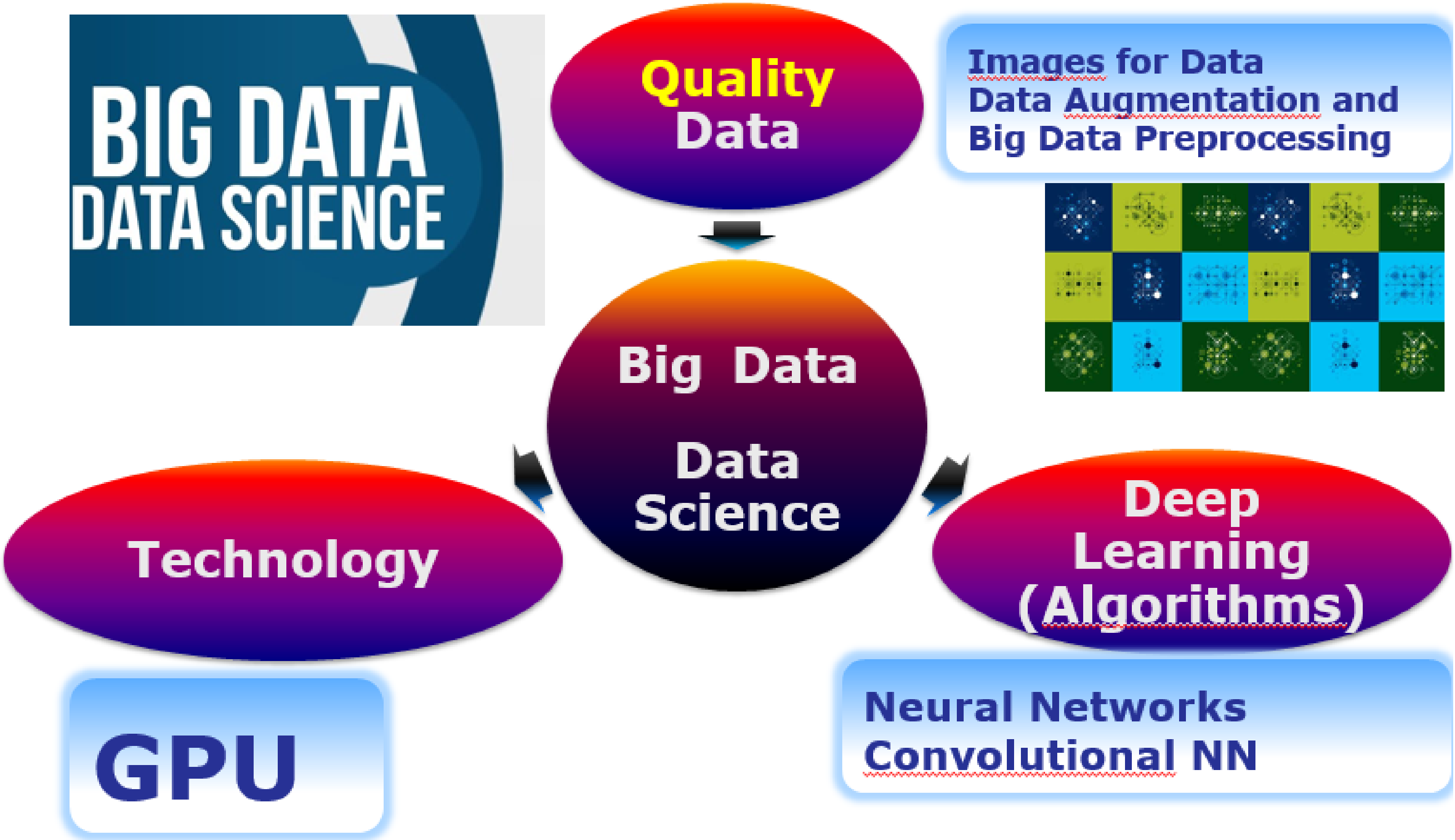
Bridge between Big Data and Smart Data:
Big Data Preprocessing

Towards quality data

Towards quality data

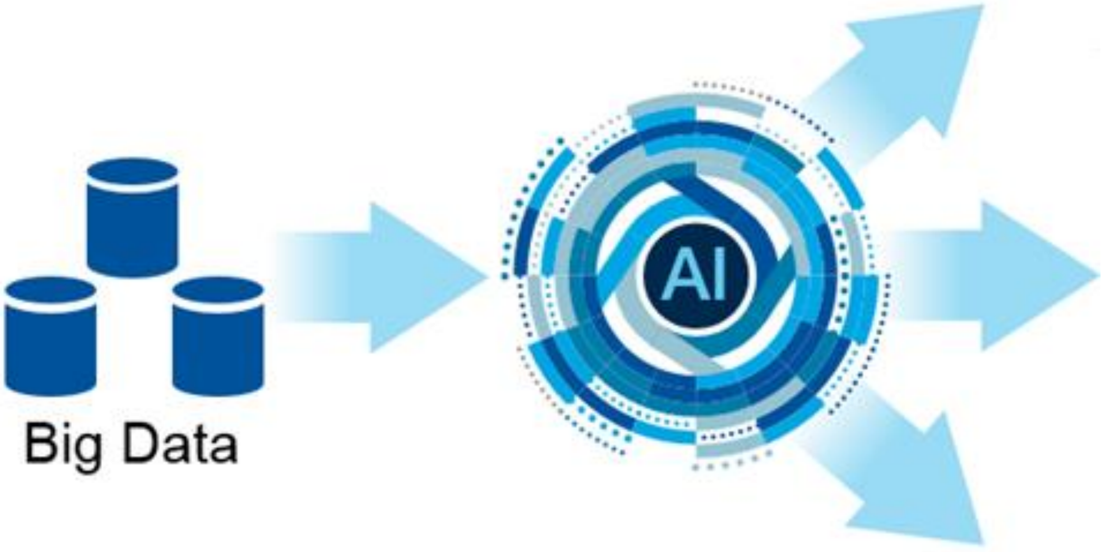


Towards quality data



Towards quality data

Smart Data

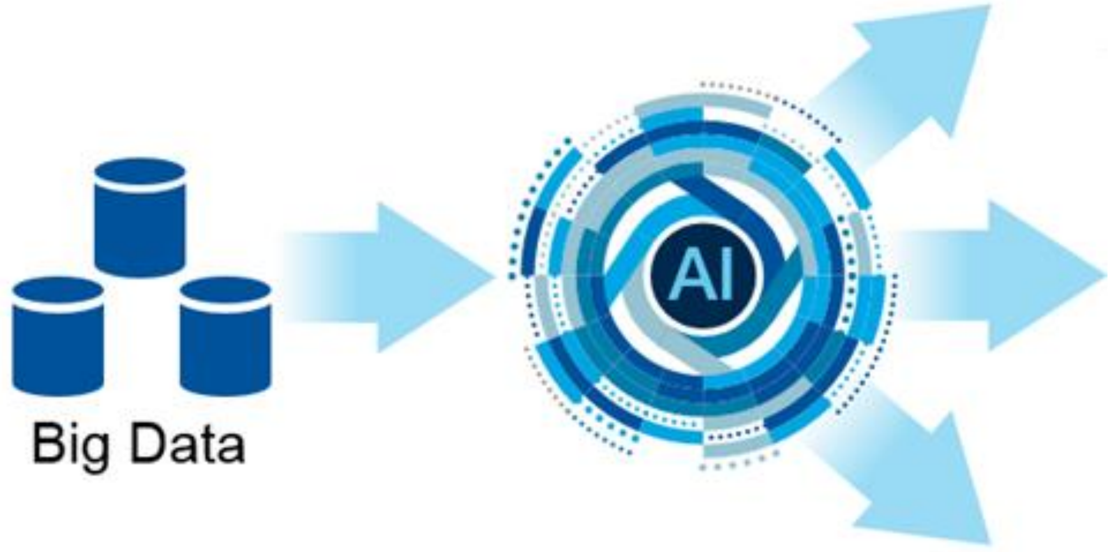


Knowledge

More quality data for better knowledge

Towards quality data

Data Preprocessing → Smart Data



Applications

Smart Data +
Machine Learning (AI) → Knowledge

More quality data for better knowledge

Deep Data and Big Learning:
More quality data for better knowledge

FedCSIS
Federated Conference on Computer Science and Information Systems

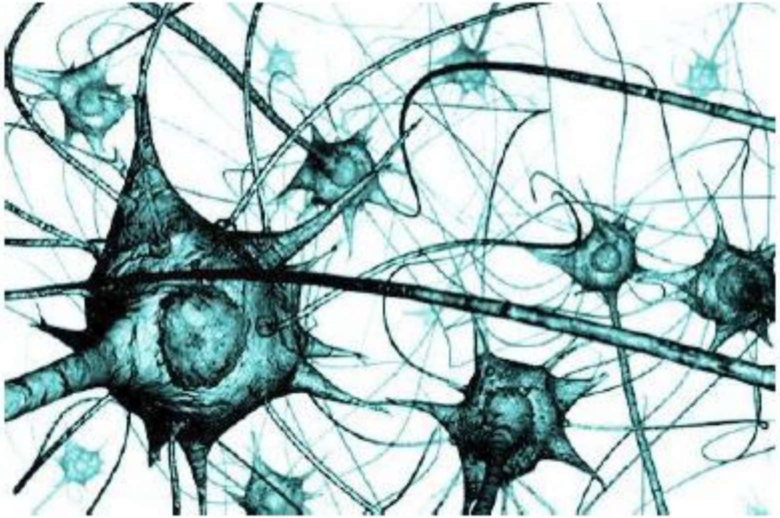
Outline

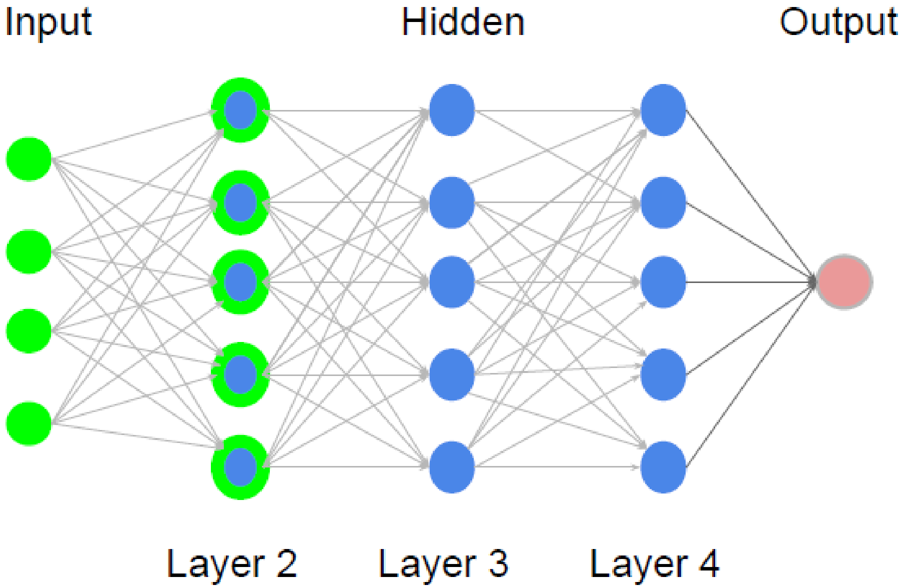
1. **Deep Data:** Towards quality data
2. **Big Learning:** CNNs with quality data
3. Case of study 1: **MNIST**
4. Case of study 2: **Whale detection**
5. Case of study 3: **Knife detection**
6. Concluding Remarks: **More quality data for better knowledge**

CNNs and quality data

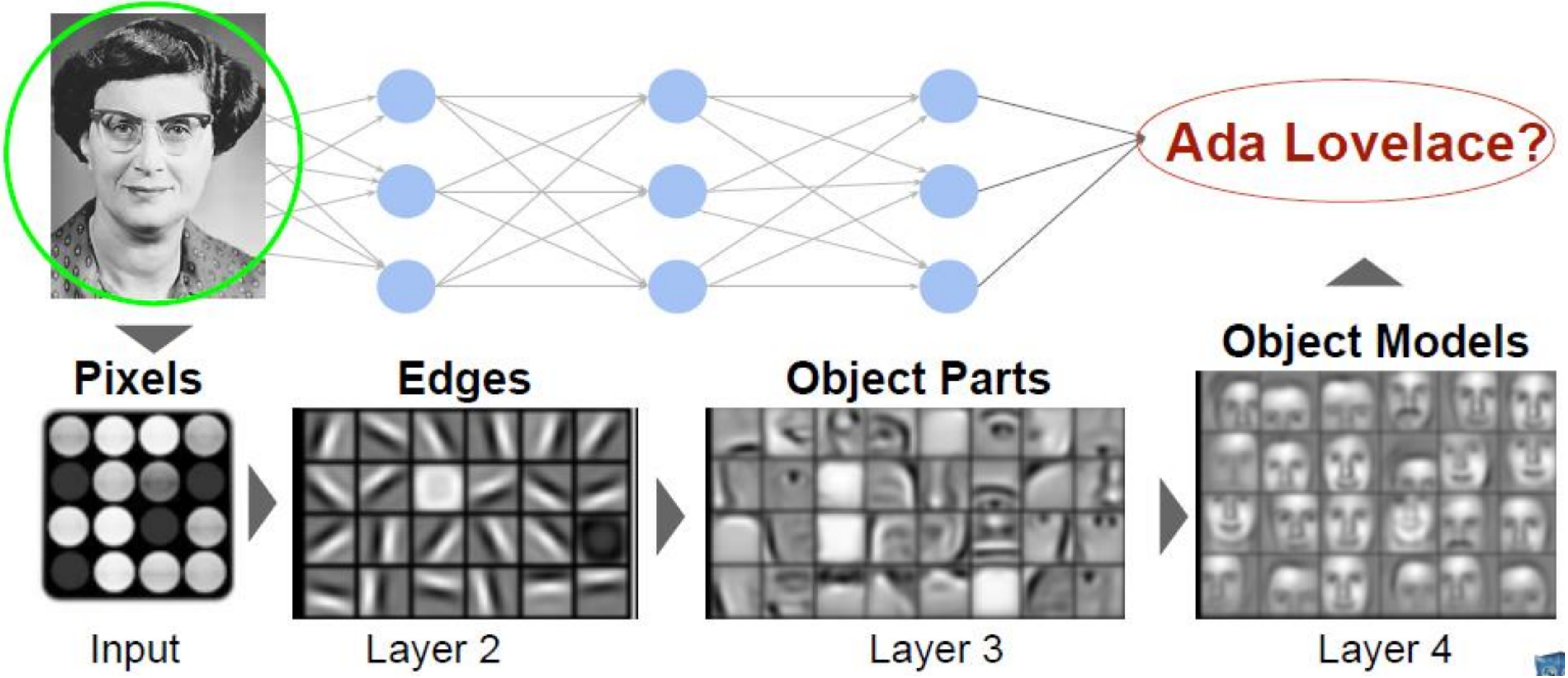
Artificial Neural Networks

Learn and predict on data



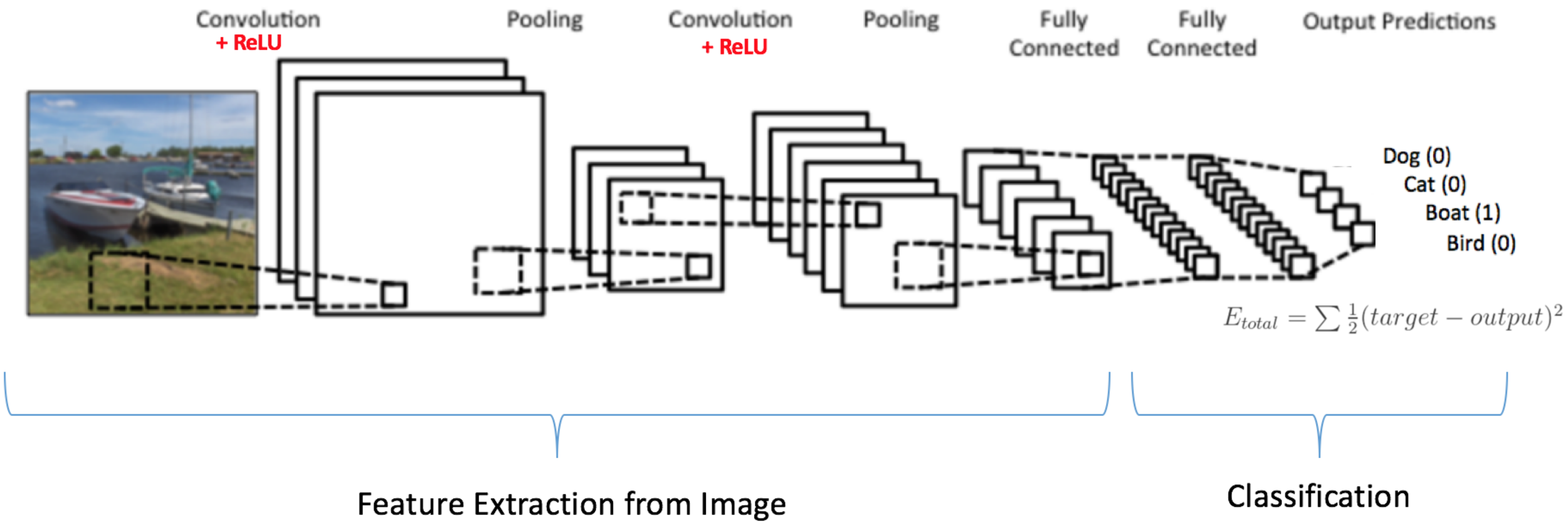


Convolutional Neural Networks



Convolutional Neural Networks

CNNs architecture



A CNN automatically learns the values of its filters based on the task you want to perform.

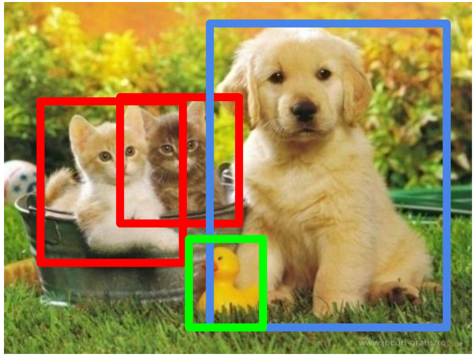
By the way, what is image classification?

Classification

CAT

Single object

Object Detection



CAT, DOG, DUCK

Instance Segmentation

CAT, DOG, DUCK

Multiple objects

Limitations: CNNs with quality data

CNNs require large amount of data to get better accuracies

Practical solutions: Smart data (deep data) as quality artificial data and quality original data together with Transfer learning

Quality artificial data: Data preprocessing as data augmentation

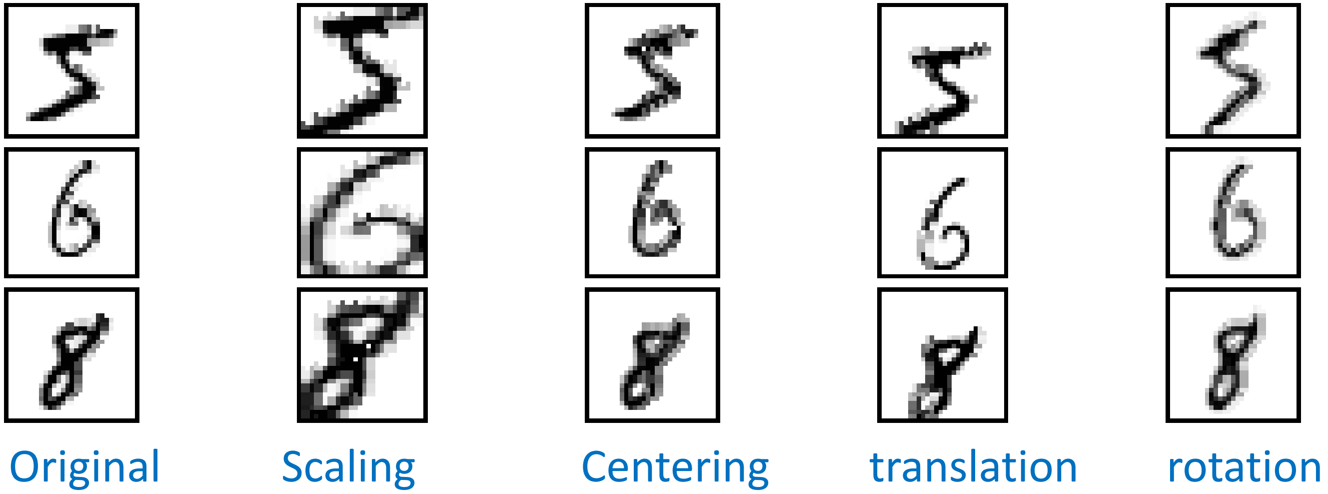
Data augmentation replicates the instances of the training set by introducing various types of transformations, e.g., translation, rotation, several types of symmetries, etc. Such techniques decrease the sensitivity of the training to noise and overfitting.

Deep Data: Smart data, Quality Data

(original and artificial data for Deep Learning)

Limitations: CNNs with quality data

Data augmentation replicates the instances of the training set by introducing various types of transformations, e.g., translation, rotation, several types of symmetries, etc. Such techniques decrease the sensitivity of the training to noise and overfitting.



Data preprocessing is very important to create quality artificial data

Limitations: CNNs with quality data

CNNs require large amount of data to get better accuracies

Practical solutions: Smart data **(deep data)** as quality artificial data and quality original data together with **Transfer learning**

Transfer learning is a machine **learning** technique where a model trained on one task is re-purposed on a second related task, applying a **fine tuning** process in deep learning.

Fine tuning is a process to take a network model that has already been trained for a given task, and make it perform a second similar task.

Big Learning: Deep Learning + Transfer Learning

Limitations: CNNs with quality data

Deep Data: Smart data, Quality Data

(original and artificial quality data for Deep Learning)

Big Learning: Deep Learning + Transfer Learning

Fundamental Idea

Deep Data and Big Learning:

More quality data for better knowledge

Deep Data and Big Learning:

More quality data for better knowledge

Outline

1. **Deep Data:** Towards quality data
2. **Big Learning**: CNNs with quality data
3. Case of study 1: **MNIST**
4. Case of study 2: **Whale detection**
5. Case of study 3: **Knife detection**
6. Concluding Remarks: **More quality data for better knowledge**

Case of study: MNIST

Handwritting recognition (60.000 training, 10.000 test)
Assign a digit from 0 to 9.

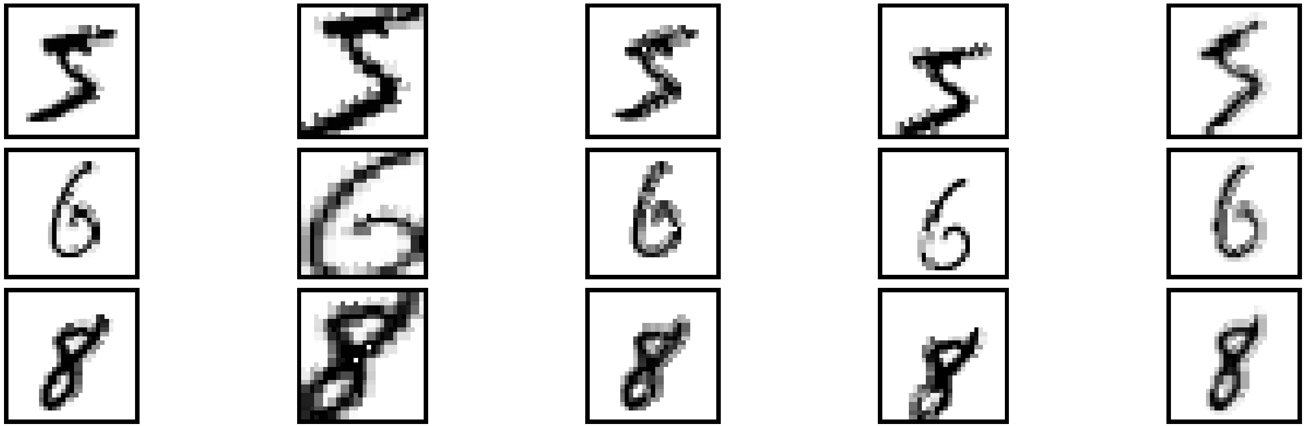
Case study: Data-augumentation for CNNs using MNIST

- Objective: Analyze the benefit of data-augmentation and ensembles on CNNs
- Methodology:
  - MNIST (60.000 train + 10.000 test) & 10 classes
  - Three CNNs: LeNet, Network3, Dropconnect
- Results

[A snapshot of image pre-processing for convolutional neural networks: case study of MNIST] S Tabik, D Peralta, A Herrera-Poyatos, F Herrera. International Journal of Computational Intelligence Systems 10, 555–568, 2017

Data augmentation

- Increase the training dataset volume artificially using transformations (tackling the large amount of data limitations)
- Objective: Improve model robustness



Original Scaling Centering translation rotation .... etc

Case study: Data-augmentation for CNNs using MNIST

Lenet-5 like CNNs:
LeNet,
Network3, DropConnect

Dataset	Combination	# of training instances
1	Original	60,000
2	Centering	60,000
3	Elastic	300,000
4	Translation	300,000
5	Rotation	300,000
6	Elastic-centering	300,000
7	Rotation-centering	300,000
8	Translation-elastic	1,500,000
9	Translation-rotation	1,500,000
10	Rotation-elastic	1,500,000
11	Rotation-elastic-centering	1,500,000
12	Elastic-elastic	1,500,000

Data-augmentation techniques

Case study: Data-augmentation for CNNs using MNIST

Test-set accuracies

	LeNet (10,000 iter.)				LeNet (50,000 iter.)			
Dataset	**Average**	**Best**	**Epochs**	**Time(s)**	**Average**	**Best**	**Epochs**	**Time(s)**
Original	99.08%	99.18%	10.67	267.91	99.05%	99.21%	213.33	1070.29
Centered	98.85%	99.06%	10.67	203.52	98.95 %	98.09%	213.33	926.38
Elastic	99.09%	99.19%	2.13	232.75	99.36%	99.44%	42.67	1065.38
Translation	99.09%	99.32%	2.13	268.75	99.30%	99.41%	42.67	1065.38
Rotation	99.05%	99.10%	2.13	268.03	99.25%	99.37%	42.67	1065.38
Elastic-centered	[5]99.17%	99.26%	2.13	267.20	99.27%	99.36%	42.67	925.51
Rotation-centered	98.90%	99.07%	2.13	232.73	99.19%	99.33%	42.67	950.38
Translation-elastic	[4]99.18%	99.32%	0.43	267.43	[5]99.39%	99.54%	8.53	1050.38
Translation-rotation	99.16%	99.40%	0.43	267.41	[3]99.40%	97.55%	8.53	1045.38
Rotation-elastic	[1]**99.31%**	99.39%	0.43	268.14	[1]**99.47%**	**99.57%**	8.53	1046.25
Rotation-elastic-centered	[3]99.19%	99.24%	0.43	232.30	[2]99.43%	99.52%	8.53	925.68
Elastic-elastic	[2]99.27%	**99.45%**	0.43	268.10	[4]99.40%	99.50%	8.53	1047.64

Case study: Data-augmentation for CNNs using MNIST

Test-set accuracies

	Network3(10 epochs)			**Network3(20 epochs)**		
Dataset	**Average**	**Best**	**Time(s)**	**Average**	**Best**	**Time(s)**
Original	99.01%	99.07%	124.45	99.25%	99.25%	205,21
Centered	98.73%	98.80%	118.32	98.97%	99.01%	196.92
Elastic	99.49%	99.54%	656,85	[3]99.61%	99.67%	1200,33
Translation	[5]99.49%	99.55%	631.53	[4]99.59%	99.63%	1228,71
Rotation	99.44%	99.50%	636.25	99.44%	99.50%	1256,95
Elastic-centered	99.32%	99.39%	566.44	99.57%	99.60%	1109,43
Rotation-centered	98.88%	98.94%	569.04	99.31%	99.32%	1167,32
Translation-elastic	[4]99.54%	99.57%	3647.78	[5]99.58%	99.63%	7111,65
Translation-rotation	[3]99.57%	99.61%	3650.66	99.58%	99.60%	7149,25
Rotation-elastic	[2]99.62%	**99.67 %**	3642,85	[2]99.67%	99.69%	6996,23
Rotation-elastic-centered	99.43%	99.51%	3054,43	99.51%	99.52%	6908,70
Elastic-elastic	[1]**99.65 %**	99.66%	3607.32	[1]**99.67 %**	**99.70 %**	7189,16

Case study: Data-augmentation for CNNs using MNIST

Test-set accuracies

	DropConnet(100 epochs)			DropConnet(200 epochs)		
Dataset	**Average**	**Best**	**Time(s)**	**Average**	**Best**	**Time(s)**
Original	98,32%	98,83%	7803.43	98.98%	98,99%	18748.53
Centered	95.35%	94,46%	6659.31	95.13%	98,85%	18635.54
Elastic	99.33 %	99,35%	7512.25	99.36%	99,36%	18606.15
Translation	[5]99.43%	99,46%	7736.41	[5]99.47%	99,47%	18710.45
Rotation	99.18%	99,29%	7151.73	99.37%	99,47%	18729.29
Elastic-centered	96.58%	96,69%	6969.89	97.08%	97,09%	18661.80
Rotation-centered	98.30%	98,41%	6974.23	98.55%	98,63%	18668.05
Translation-elastic	99.40%	99,57%	7162.37	[3]99.58%	99,58%	18745.93
Translation-rotation	[2]99.57%	99,59%	7410.32	[1]**99.69%**	**99,69%**	18772.40
Rotation-elastic	[3]99.54%	**99,60%**	7397.40	[4]99.56%	99,56%	18724.38
Rotation-elastic-centered	[4]99.47%	99,49%	7803.73	99,44%	99,46%	18220.50
Elastic-elastic	[1]**99,58%**	99,59%	7911.30	[2]99,59%	99,61%	18712.22

Case study: Data-augmentation for CNNs using MNIST

Results

	LeNet(500 neurons)		**Network3**		**DropConnect**	
	10,000 iter	**50,000 iter**	**10 epochs**	**20 epochs**	**100 epochs**	**200 epochs**
Ensemble-5	99,55%	**99,57%**	**99,72%**	**99,69%**	**99,72%**	99,66%
Ensemble-3	99,43%	**99,54%**	99,69%	99,67%	99,69%	**99,68%**

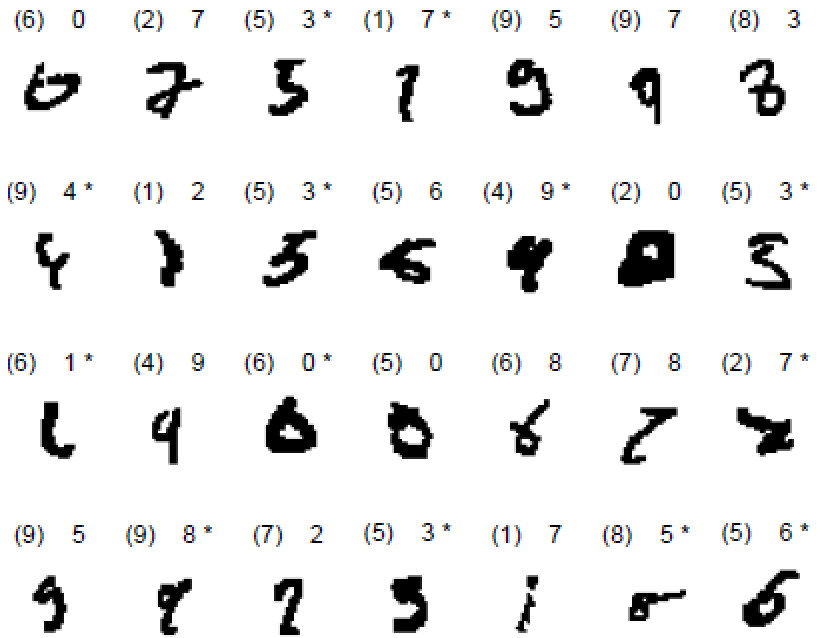
Error : 0.28% versus state of the art ensemble 0.16%

Case study: Data-augmentation for CNNs using MNIST

<u>Results</u>

The 28 misclassified characters (15 different)

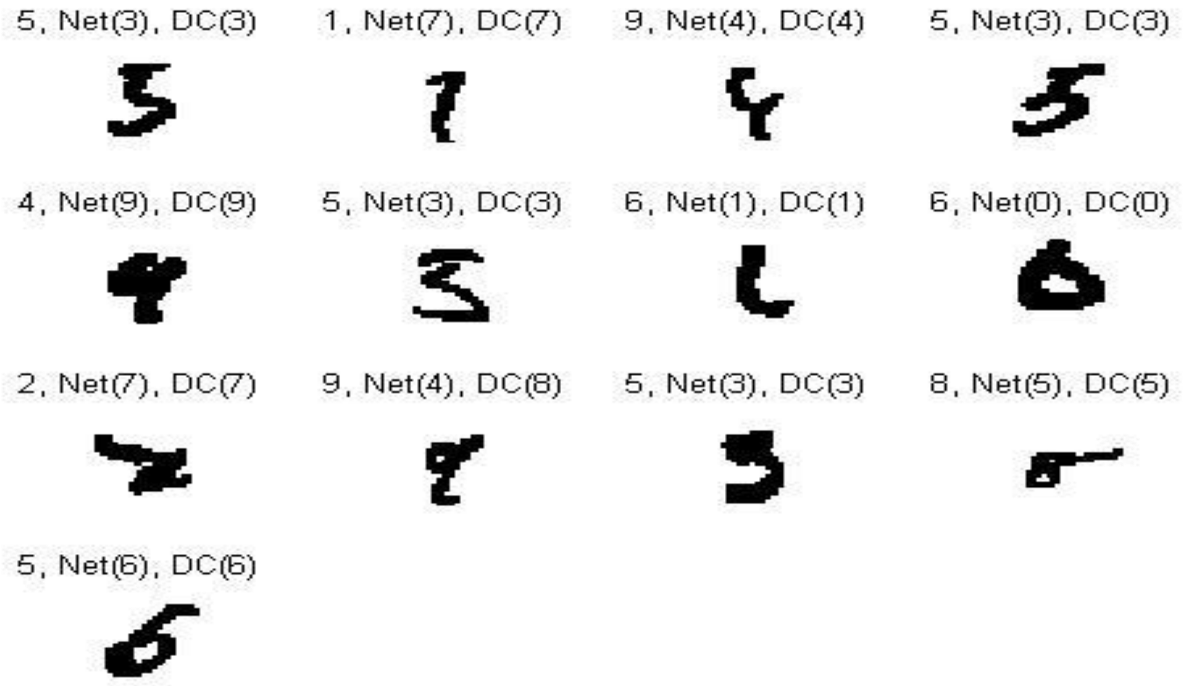
ensemble-5 (Network3)



ensemble-5 (DropConnect)

Case study: Data-augmentation for CNNs using MNIST

The 13 handwritten digits misclassified by ensemble-5 of DropConnet and Network3



Case study: Data-augmentation for CNNs using MNIST

Deep Learning: MNIST data (10.000 test)
Ensemble with different top CNN models

The digit between () represents the correct class.

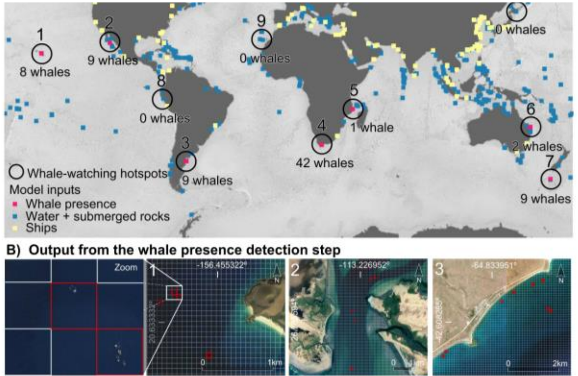
May 2019, Granada team (12 errors). World RECORD

Outline

1. **Deep Data:** Towards quality data
2. **Big Learning**: CNNs with quality data
3. Case of study 1: **MNIST**
4. Case of study 2: **Whale detection**
5. Case of study 3: **Knife detection**
6. Concluding Remarks: **More quality data for better knowledge**

- Training dataset: 700 (with 976 whales) **aereal** images extracted from: Google Earth, free Arkive, NOAA Photo Library y NWPU-RESISC45 dataset.

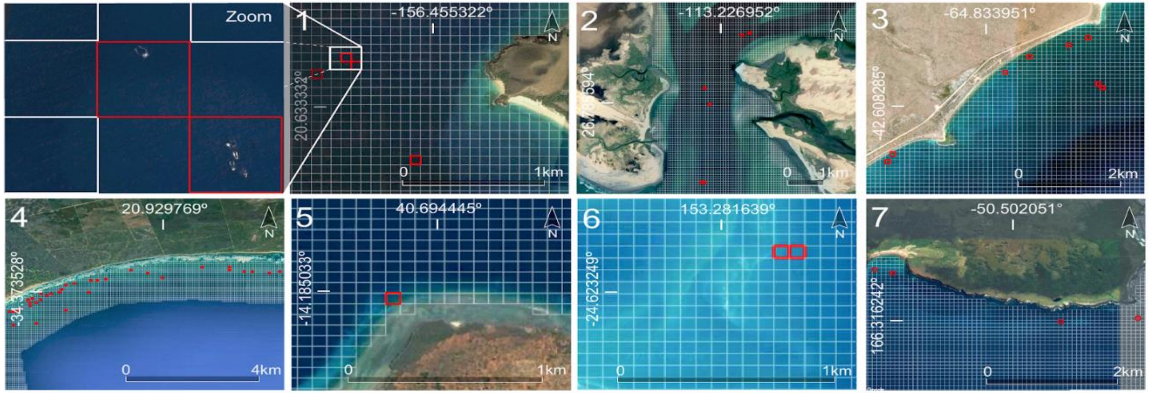
- Test dataset: **Satellite** images of ten whale watching hotsposts



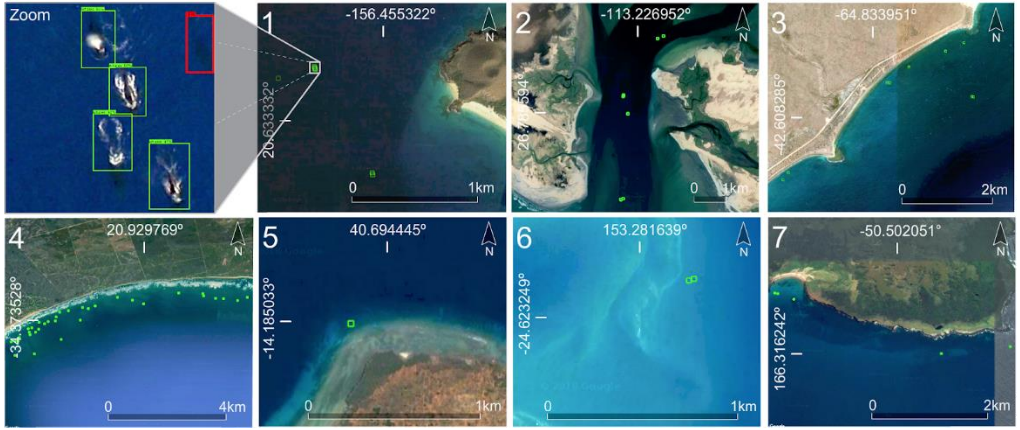
Problem	Database	Model	**Results**

The first step detects whale presence with an F1_score of 84%

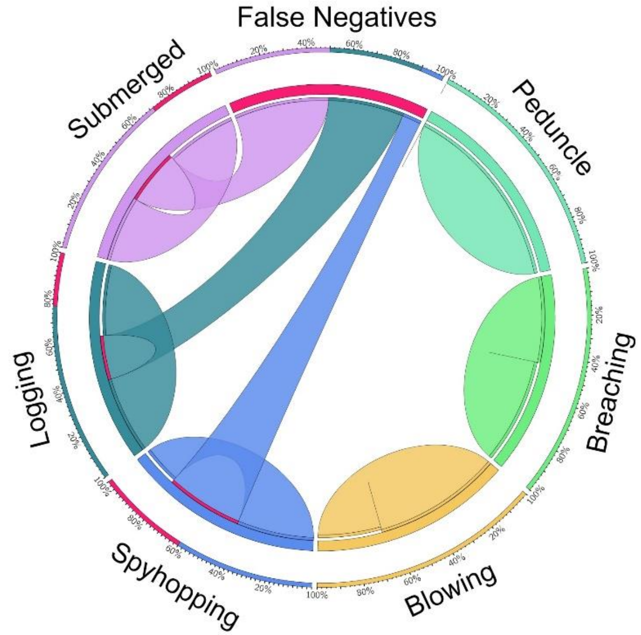
- Trimming
- Scaling up and down -> the key for generalizing from higher to lower resolution
- Rotations and translations
- Different illumination conditions

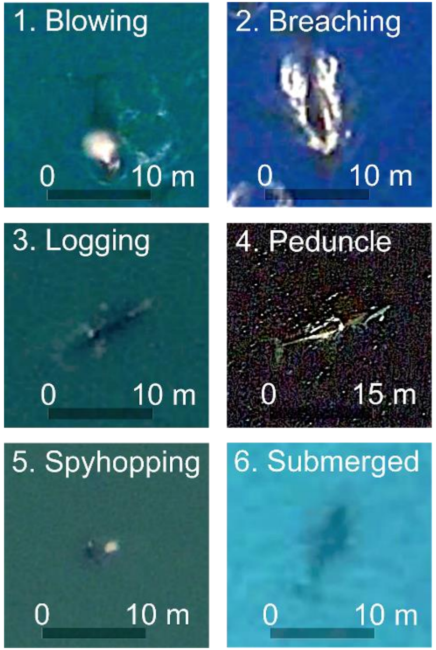


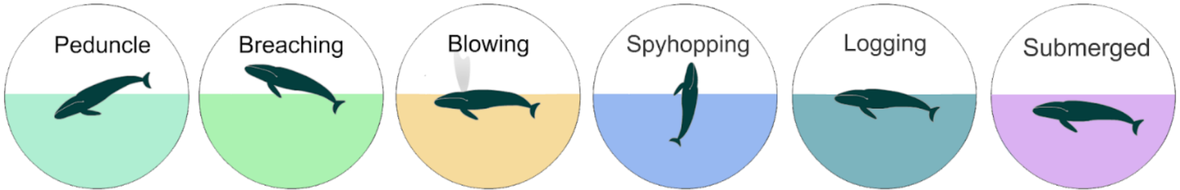
The second step counts whales with F1_score 97%



Impact of whale postures and movements on the performance of the model







Deep Data and Big Learning:
More quality data for better knowledge

Outline

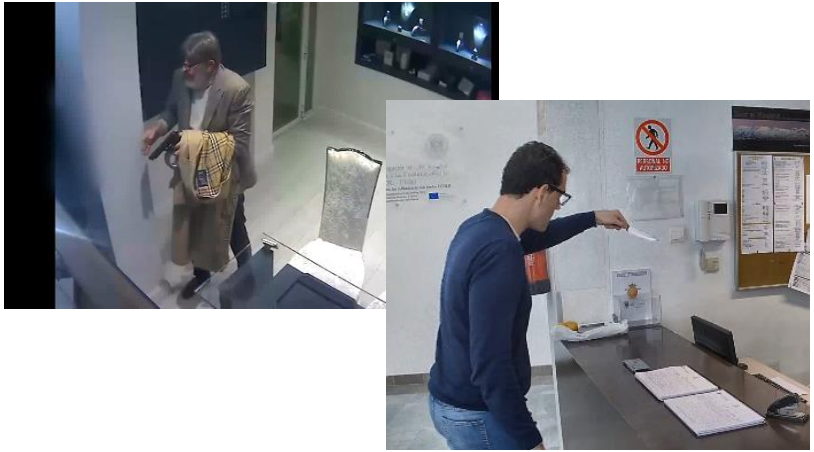
1. **Deep Data:** Towards quality data
2. **Big Learning**: CNNs with quality data
3. Case of study 1: **MNIST**
4. Case of study 2: **Whale detection**
5. Case of study 3: **Knife detection**
6. Concluding Remarks: **More quality data for better knowledge**

Project: Weapon Detection Alarm in Video Surveillance

The future of smart security

One of the ways to reduce the threat of violence generated by weapons is the early detection of their presence with enough time for agents or watchmen to act.

A novel solution could integrate an automatic weapon detection system with video surveillance system.



Project: Weapon Detection Alarm in Video Surveillance

No published work, patent, or commercial product addresses the problem of gun detection in real-time video using Deep Learning.

Our publication (February 2017) was the first work that use Deep Learning to detect weapons in video surveillance.

Automatic Handgun Detection Alarm in Videos Using Deep Learning

Roberto Olmos[1], Siham Tabik[1], and Francisco Herrera[1,2]

https://arxiv.org/abs/1702.05147

MIT Technology Review

Connectivity

The Best of the Physics arXiv (week ending March 4, 2017)

This week's most thought-provoking papers from the Physics arXiv.

by Emerging Technology from the arXiv March 4, 2017

A roundup of the most interesting papers from the arXiv:

Automatic Handgun Detection Alarm in Videos Using Deep Learning

R Olmos,, S Tabik, F Herrera
Automatic handgun detection alarm in videos using deep learning
Neurocomputing 275, 67-72, 2018

Case of study: Knife detection

Project: Weapon Detection Alarm in Video Surveillance

- Objective: Develop a fast and accurate arms detection model in videos
- Methodology: To create a database (knife / no knife) + To develop a Deep learning model

(a) (b) (c) (d)

Case of study: Knife detection

VGG16 ✗ Fine-tuning

CNN learns from scratch fitting weights through Backpropagation
Test: 178 knifes and 138 not knifes

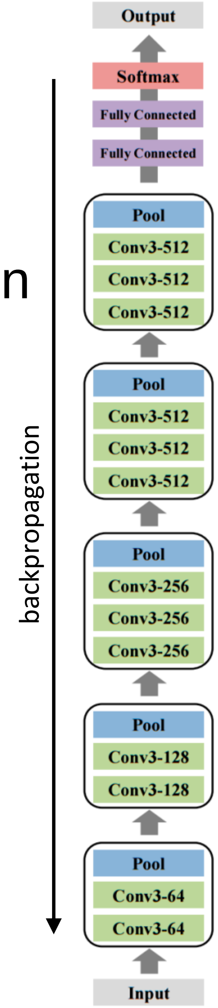
	Knife	no knife
Knife 178	147	31
No knife 138	37	101

Accuracy = 0.799
Recall = 0.826
F1 score = 0.812

FP




Knife	49,6%	69,8%	73,7%	50,1%



Case of study: Knife detection

VGG16 Previous FP ✓ Fine-tuning

- Fine-tuning improve classification because of pre-training

Class	Probability
No-Knife	99,99%
Knife	0%

Class	Probability
No-knife	98,77%
Knife	0%

Class	Probability
No-knife	99,99%
Knife	0%

Big Learning

Case of study: Knife detection

VGG16 ✓ Fine-tuning

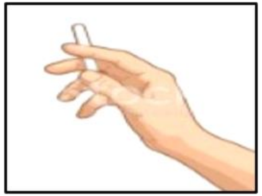
- CNN knows to extract key features and learns to classify
- Backpropagation fit fully connected layers only
- Test: 178 knifes and 138 not knifes

	Knife	no knife
Knife 178	168	10
No knife 142	7	135

Accuracy = 0.96 **(+ 0.16)**
Recall = 0.944
F1 score = 0.952

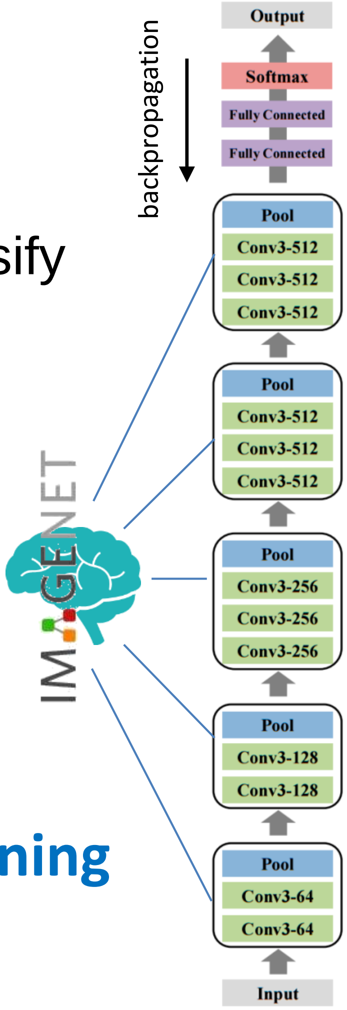
FP





Knife	100%	99,9%	100%

Big Learning



Case of study: Knife detection

Challenge: Brightness conditions may deteriorate image quality

Case of study: Knife detection

Challenge: Brightness conditions may deteriorate image quality

High brightness Low brightness

Fig. 7. An example of the detection results in two similar situations with different brightness conditions.

Case of study: Knife detection

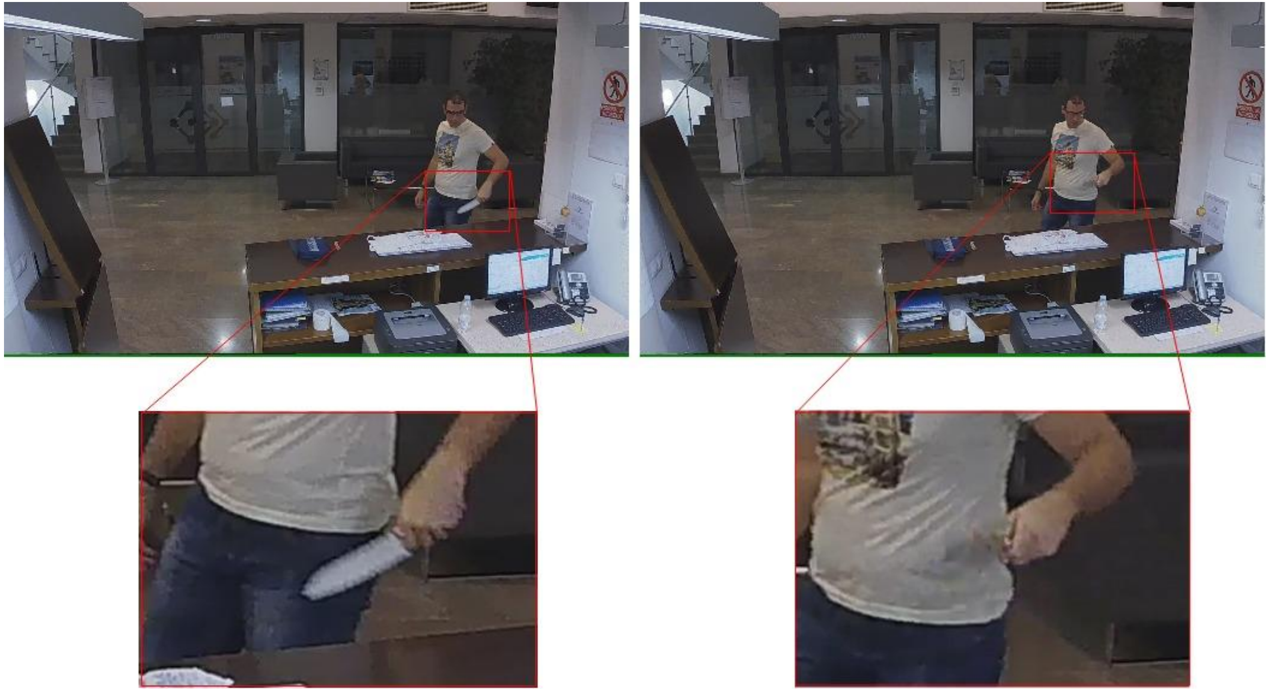
Challenge: Brightness conditions may deteriorate image quality

Table 6
Detection performance obtained on videos recorded in different brightness conditions.

Brightness	Knife size	#frames	#GT_P	#TP	#FP	Precision	Recall	F1
High	Large	121	112	78	0	100%	69.64%	82.1%
	Medium	107	90	44	0	100%	48.89%	65.67%
	Small	137	103	53	0	100%	51.46%	67.95%
			Average			100%	56.66%	71.91%
Medium	Large	109	98	85	0	100%	86.73%	92.89%
	Medium	116	98	73	0	100%	74.49%	85.38%
	Small	138	110	64	0	100%	58.18%	73.56%
			Average			100%	73.13%	83.94%
Low	Large	126	114	104	1	99.05%	92.04%	95.41%
	Medium	114	100	70	0	100%	70%	82.35%
	Small	138	101	74	0	100%	73.27%	84.57%
			Average			99.68%	78.44%	87.44%
Artificial	Large	119	110	95	0	100%	86.36%	92.68%
	Medium	113	99	75	3	96.15%	78.13%	86.21%
	Small	96	90	65	4	94.2%	75.58%	83.87%
			Average			96.78%	80.02%	87.59%

Case of study: Knife detection

Movement also generates noise and distortion



Case of study: Knife detection

DaCOLT: Darkening and Contrast at Learning and Test stages

Enhance robustness through data-augmentation and Preprocessing

Tackling brightness via darkening and contrast.

SSD(InceptionV2)
R-FCN(ResNet101)
Faster R-CNN (Inception-ResNet-V2, ResNet50, ResNet101, and Inception\

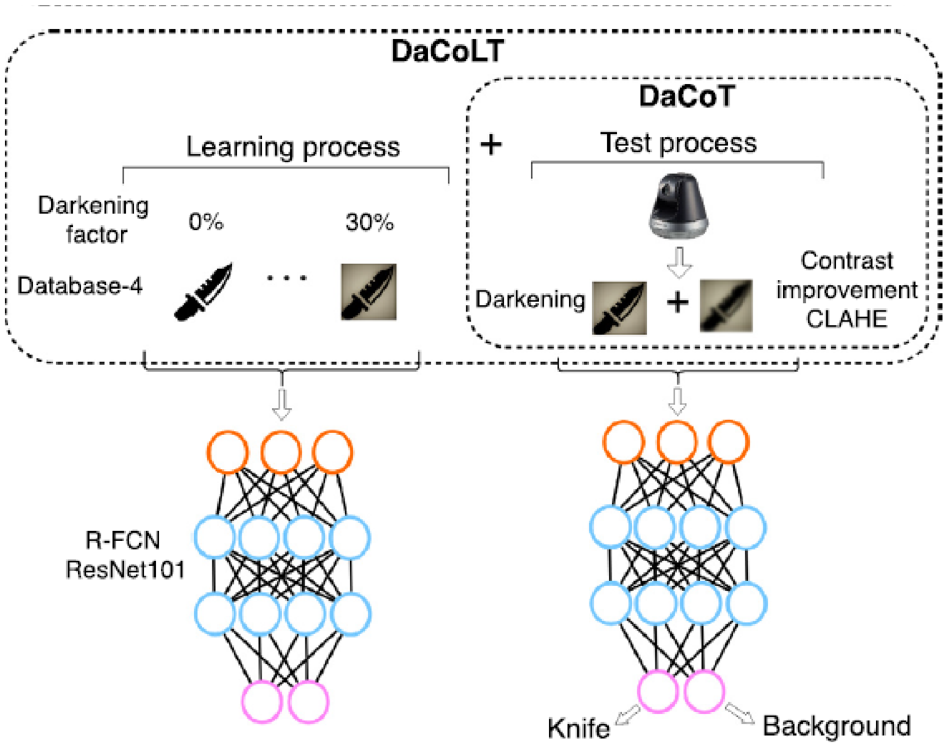


Fig. 8. An illustration of DaCoLT procedure applied at both, learning and test time.

Castillo, S Tabik, F Pérez, R Olmos, F Herrera
Brightness guided preprocessing for automatic cold steel weapon detection in surveillance videos with deep learning
Neurocomputing 330, 151-161, 2019

Case of study: Knife detection

Darkening and Contrast at Learning and Test stages

DaCoLT procedure consists of two stages:

- Training the detection model on a selected range of brightness conditions using data-augmentation
- Achieving the ideal brightness condition by adjusting the darkening of the frames and improving their visual quality using a preprocessing approach (CLAHE) before analyzing them with the detection model.

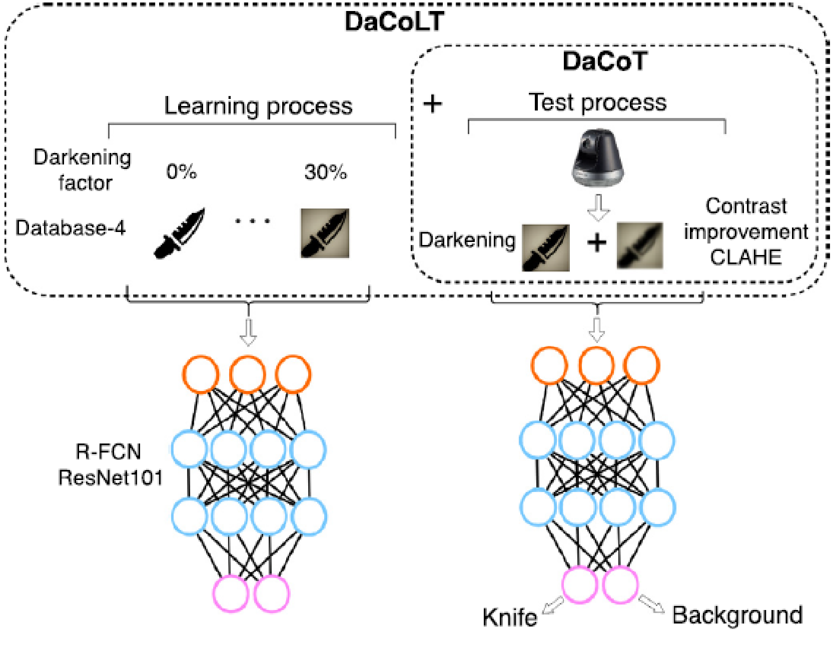
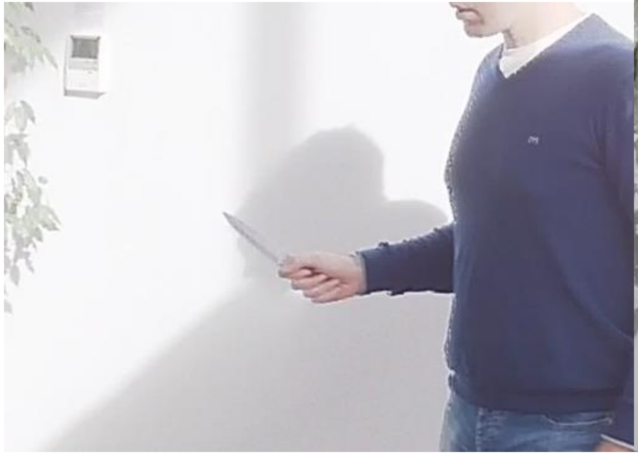


Fig. 8. An illustration of DaCoLT procedure applied at both, learning and test time.

Case of study: Knife detection

DaCoLT image sample



Original

No detection

Preprocessing

Detection: 71%

Preprocessing + data aug.

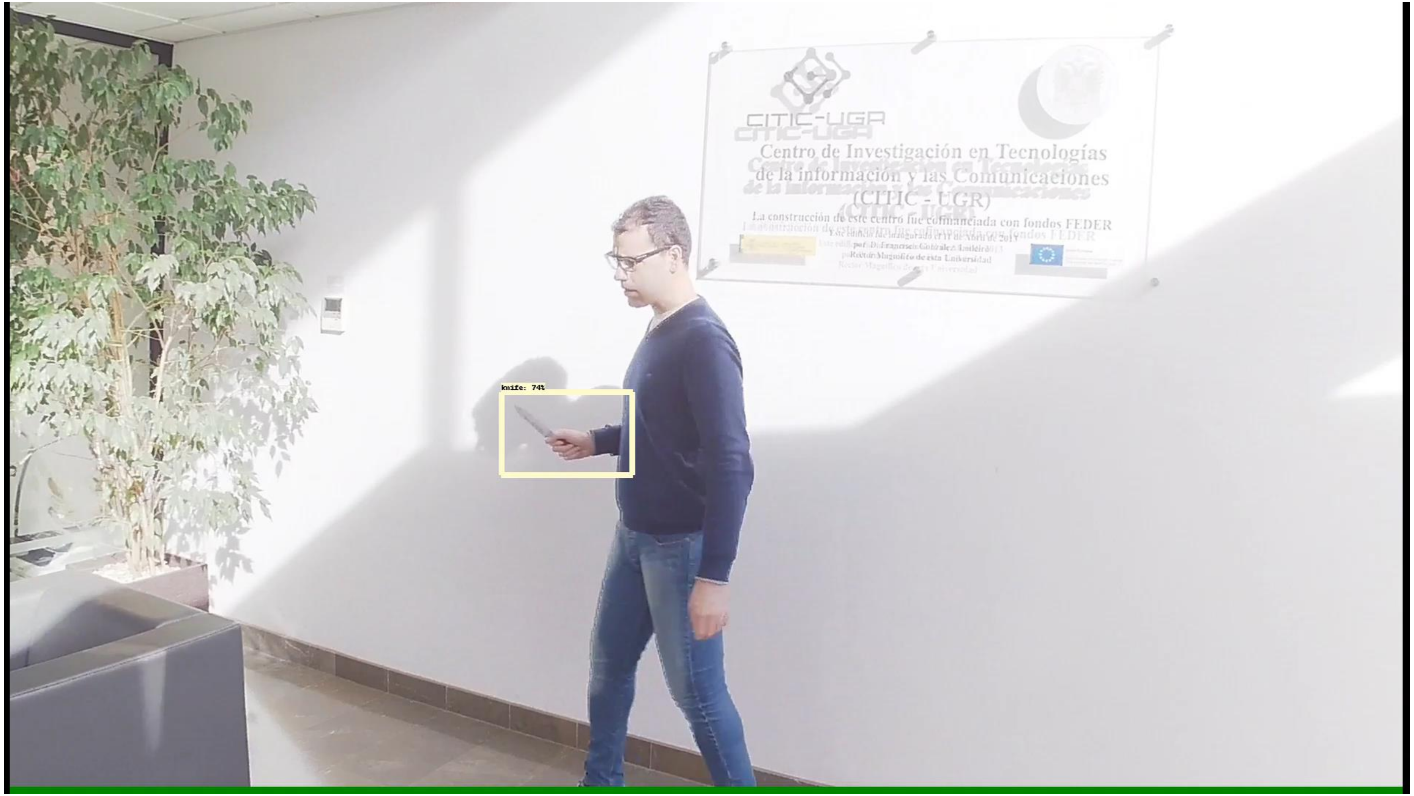
Detection: 99%

Case of study: Knife detection

Detection in original condition

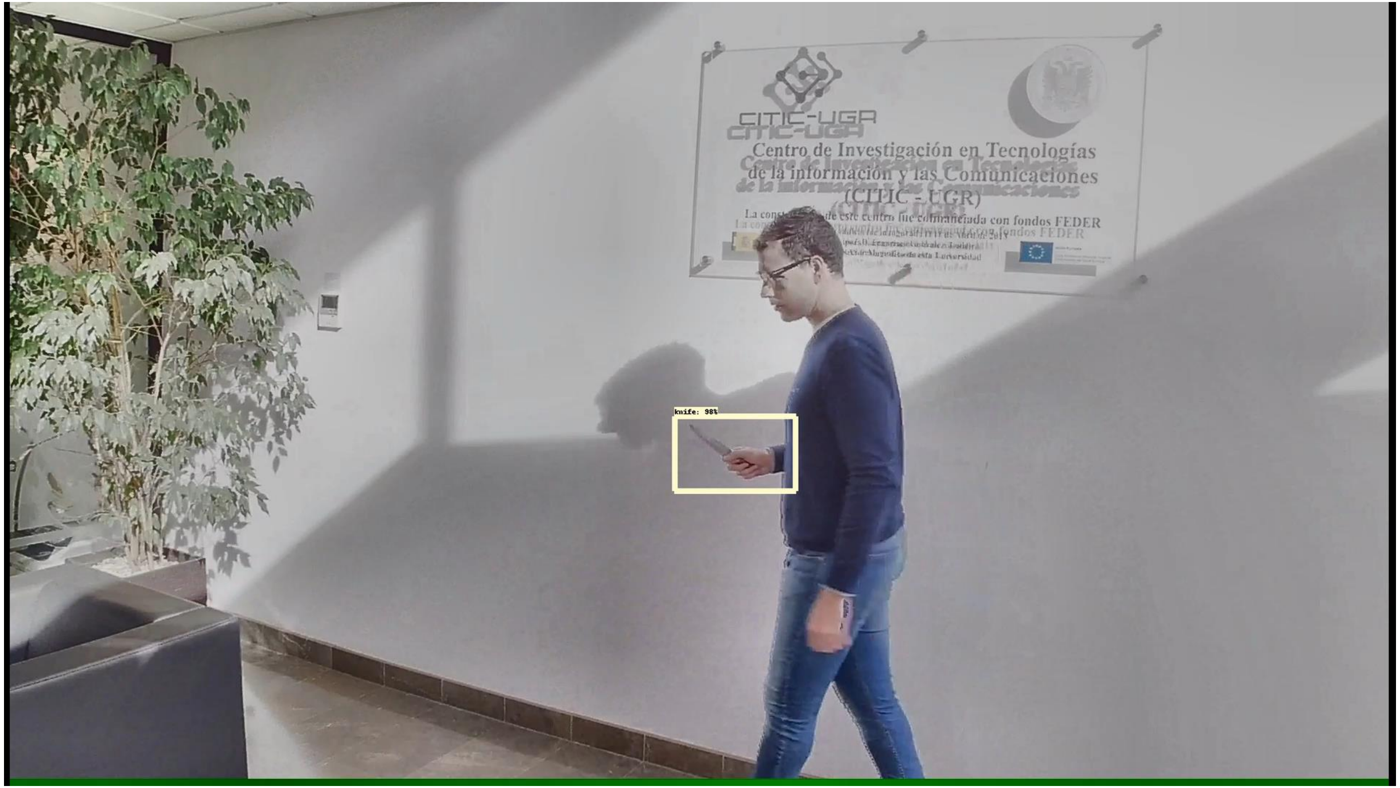
- The knife Surface reflects and make difficult the detection

- Sometimes the knife even dissapear



Case of study: Knife detection
Detection applying DaCoT

- Preprocessing applied at inference stage
- Improve the reflectance in brighter areas
- Slight True Positives rise

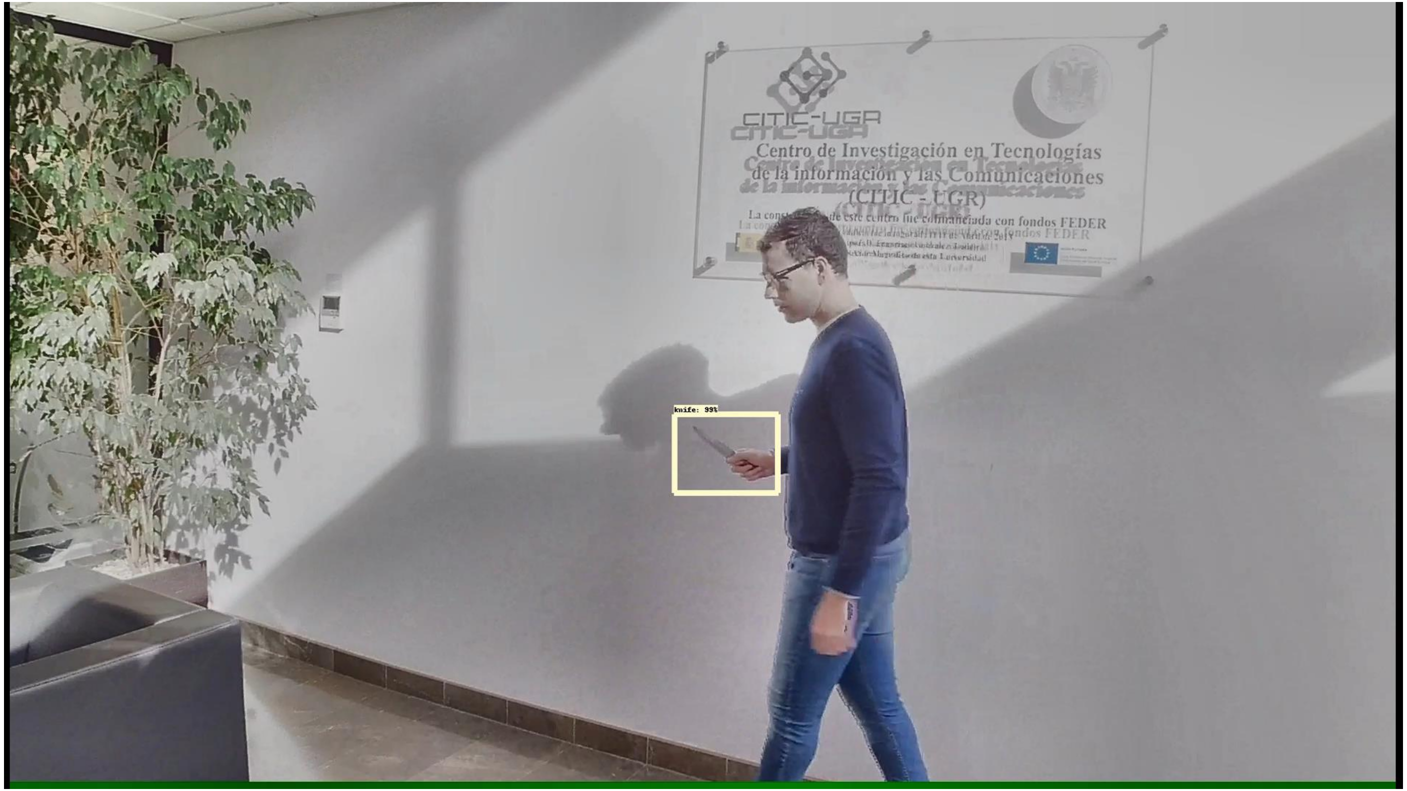


Case of study: Knife detection

Detection applying DaCoLT

- The preprocessing technique is applied for data-augmentation

- High True Positives rate rises



Case of study: Knife detection

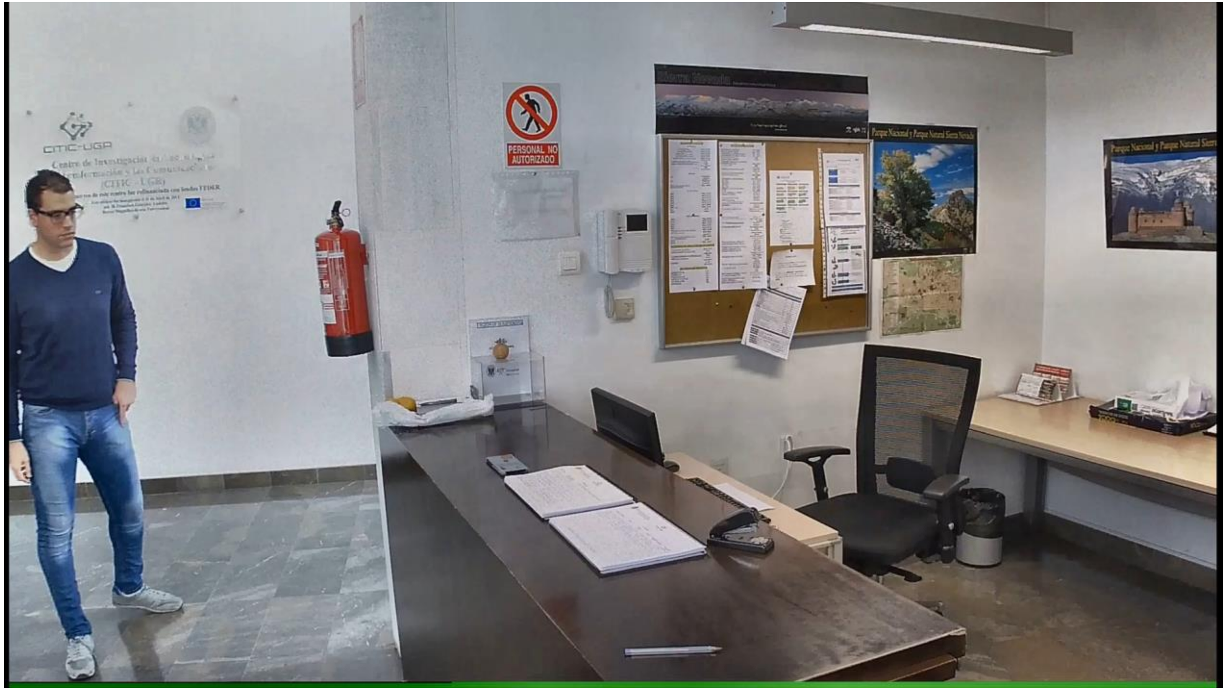
DaCoLT results

High brightness conditions with different knife sizes

	Knife size	#frames	#GT_P	#TP	#FP	Precision	Recall	F1
Original	Large	121	112	78	0	100%	69.64%	82.11%
High	Medium	107	90	44	0	100%	48.89%	65.67%
Brightness	Small	137	103	53	0	100%	51.46%	67.95%
			Average			100%	56.66%	71.91%
Guided brightness	Large	121	112	85	0	100%	75.89%	86.29%
DaCoT	Medium	107	90	56	0	100%	62.22%	76.71%
(Test time)	Small	137	103	53	0	100%	51.46%	67.95%
			Average			100%	63.19%	76.98%
Guided brightness	Large	121	112	84	0	100%	75%	85.71%
DaCoLT	Medium	107	90	64	0	100%	71.11%	83.12%
(Learning+Test)	Small	137	103	74	0	100%	71.84%	83.61%
			Average			100%	**72.65%**	**84.15%**

Case of study: Knife detection

Simulation applying DaCoLT: **Brightness guided preprocessing for knife detection**



Deep Data and Big Learning:

More quality data for better knowledge

Outline

1. **Deep Data:** Towards quality data
2. **Big Learning**: CNNs with quality data
3. Case of study 1: **MNIST**
4. Case of study 2: **Whale detection**
5. Case of study 3: **Knife detection**
6. Concluding Remarks: **More quality data for better knowledge**

Concluding Remarks

In contrast to the classical classification models, the high abstraction capacity of CNNs allows them to work on the original high dimensional space, which reduces the need for manually preparing the input.

However, a suitable preprocessing is still important to improve the quality of the result (including image preprocessing, data augmentation, …)

The supervised Deep Learning depends a lot on that phase of human annotation/labeling/selection.

Concluding Remarks

- Quality data preprocessing techniques adapt the data to fulfill the input demands of each data algorithm.
- Quality data preprocessing is an essential part of any automatic learning process.
- Quality data preprocessing techniques (including Data Augmentation) are very important for Deep Learning

Central idea. Deep Data and Big Learning:

More quality data for better knowledge

Concluding Remarks

Limitations and reflection

Focused in image analysis, the creation of the "Smart data" level databases in the context of supervised deep learning always goes through **the manual revision of the expert notebook.**

There are still no automatic methods that create "Smart data" for Deep Learning without the help of the human annotation".

Concluding Remarks

Limitations and reflection

Remember: There are still no automatic methods that create "Smart data" for the Deep Learning without the help of the human annotation.

There are open research studies towards quality data

Imbalanced classification: It needs preprocessing for the minority class

What is the meaning of noise data in deep learning?

Difficult instances for classification, selection and filtering

What is the correspondence with data reduction for getting quality small data for deep learning?

Concluding Remarks

Ending as we began

Quality decisions
(“quality models/patterns/rules”)
are based on
Quality Data!

More quality data for better knowledge

Quality Data to drive Deep Learning Applications

Deep Data and Big Learning:

More quality data for better knowledge