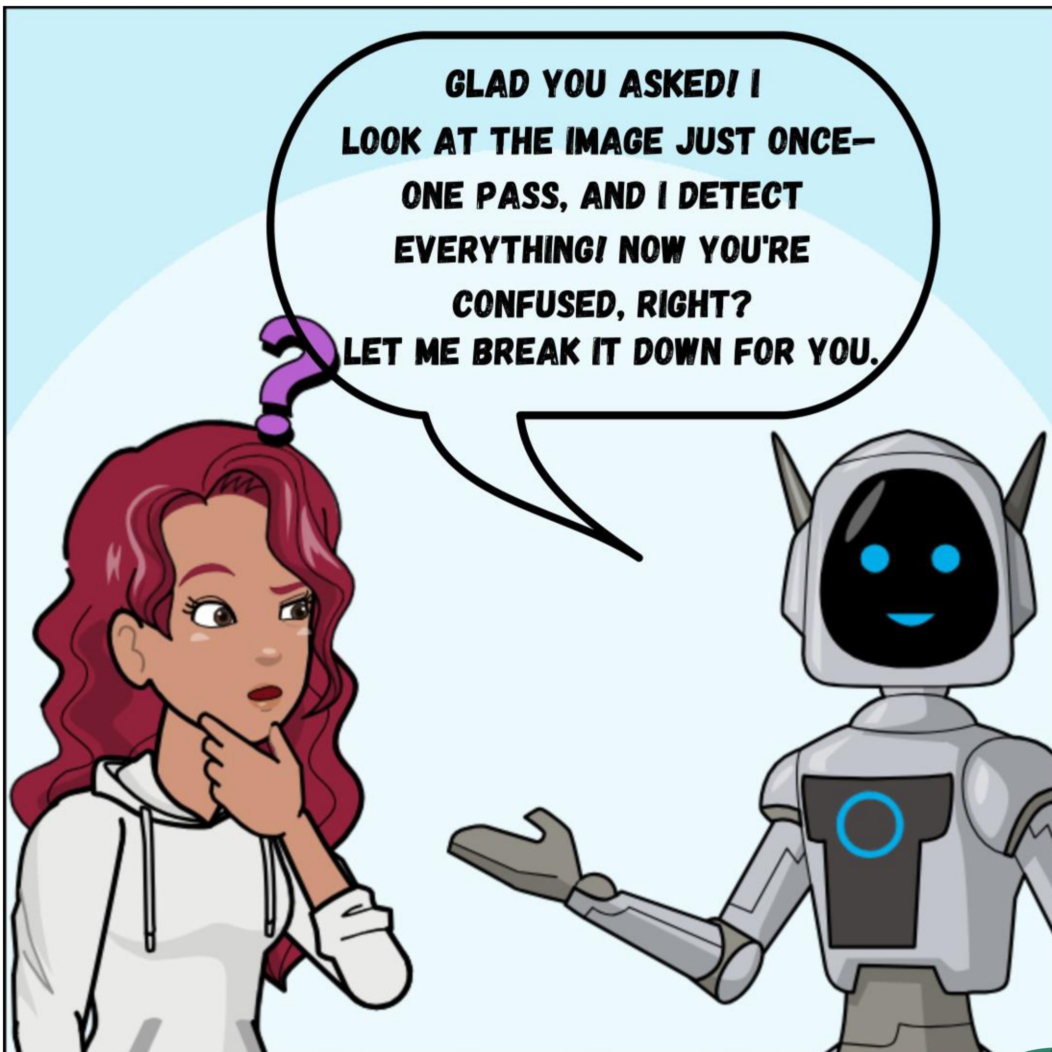
ISSN 2830-9820

# VisualComputing MAGAZiNE

Vol. 2, Issue 2
2024




## CHRONICLE OF COMPUTER VISION

## The Foreword

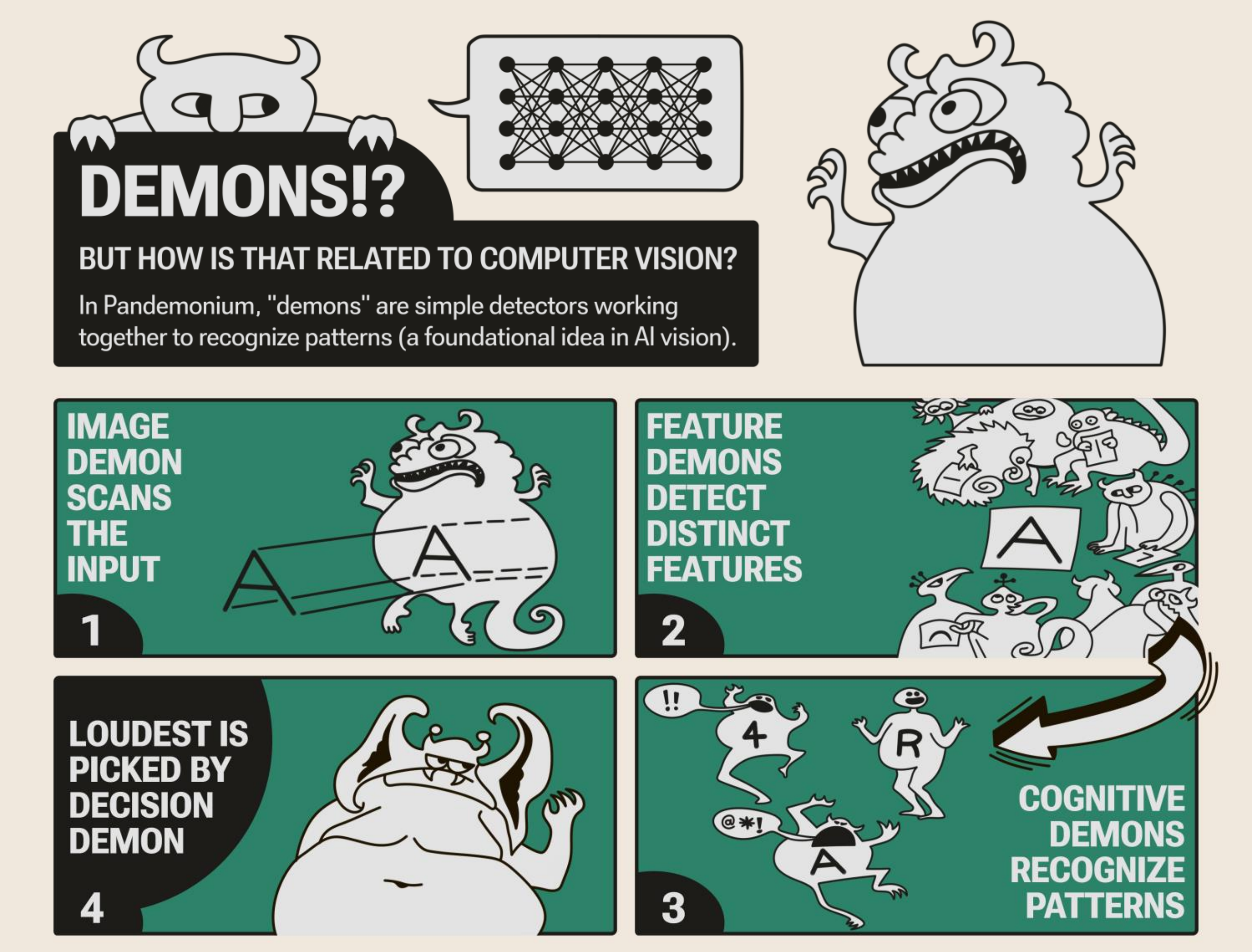
In this issue, we feature a series of posters made by Visual Computing Master students of USTHB University, each highlighting specific moments that redefined computer vision.
We begin with the concept of “Pandemonium” (1958), which laid the foundation for the perception of visual information through hierarchical processes. We then look at extracting 3D information from 2D images (1963) and highlight the landmark Summer Vision project (1966), which was an ambitious attempt to bring perception to machines at MIT.
Key breakthroughs followed: HOUGH's transformation (1972), optical flow and motion perception in 1981, and David Marr's paradigm of visual representation (1982). These developments were followed by tools such as the Canny Edge detector (1986) and object detection including the Viola-Jones detector (2001). The new millennium ushered in advances such as SIFT (2004) and SURF (2006), essential for primitive detection, and classifiers such as support vector machines (SVM, 2008).
The introduction of AlexNet in 2012 marked the beginning of the deep learning era, which saw models like GANs (2014), GoogleNet (2014) and ResNets (2015) bring remarkable advancements to the discipline . Recent innovations, including Vision Transformers (2021) and DALL-E 3 (2023), are pushing the boundaries of machine understanding and creative generation.
Each poster provides a window into these achievements, capturing the essence of the ideas and advances that have advanced computer vision over the decades.

**Chief Editor, Prof. Slimane LARABI**
**Computer Science Faculty**
**USTHB University, Algeria**


***Visual Computing: A quarterly magazine. Volume 2, Issue 2, 2024***
***ISSN:* 2830-9820**
***Chief Editor:   Prof. Slimane LARABI***

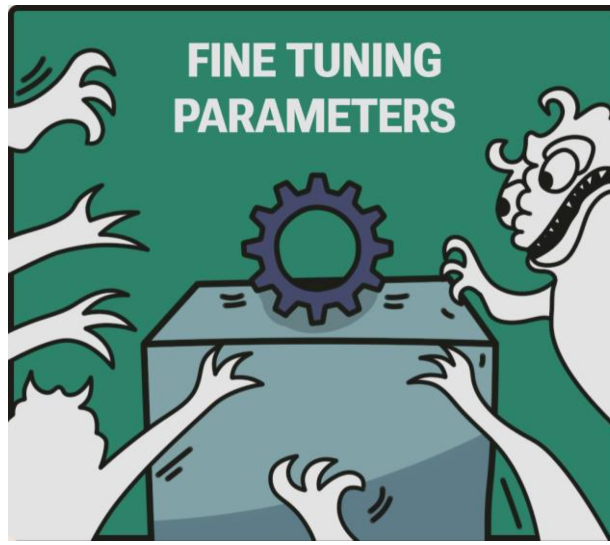


# CHRONICLE OF COMPUTER VISION
## EVENT: CONCEPT OF A "PANDEMONIUM", 1958
I. TIMSILINE, A. ABDENNOUZ, MASTER 2 VISUAL COMPUTING, USTHB



# CHRONICLE OF COMPUTER VISION
## EVENT: CONCEPT OF A "PANDEMONIUM", 1958

I. TIMSILINE, A. ABDENNOUZ, MASTER 2 VISUAL COMPUTING, USTHB


**PANDEMONIUM INFLUENCE**

- DEEP LEARNING MODELS
- SPEECH RECOGNITION
- OCR SYSTEMS
- NLP CONTRIBUTION

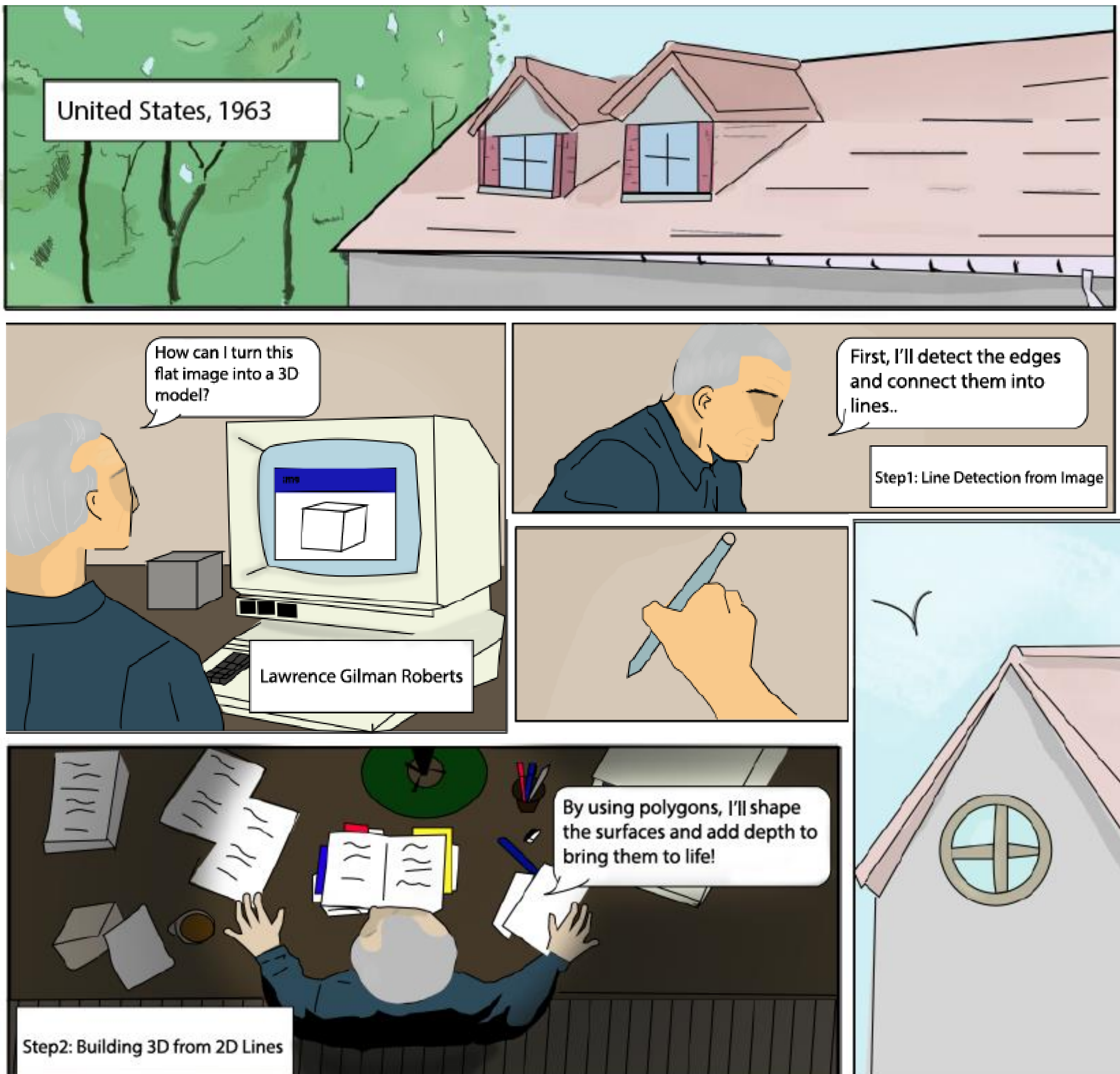
**REFERENCES**


- O. G. Selfridge, "Pandemonium: a paradigm for learning," Mechanisation of Thought Processes: Proceedings of a Symposium Held at the National Physical Laboratory, November 1958
- Hinton, L. (1977). Illustration. In D. A. Lindsey & N. A. Norman (Eds.), Human Information Processing (p. 116). Academic Press, Inc.

# CHRONICLE OF COMPUTER VISION

## EVENT: 3D information from 2D photographs , 1963

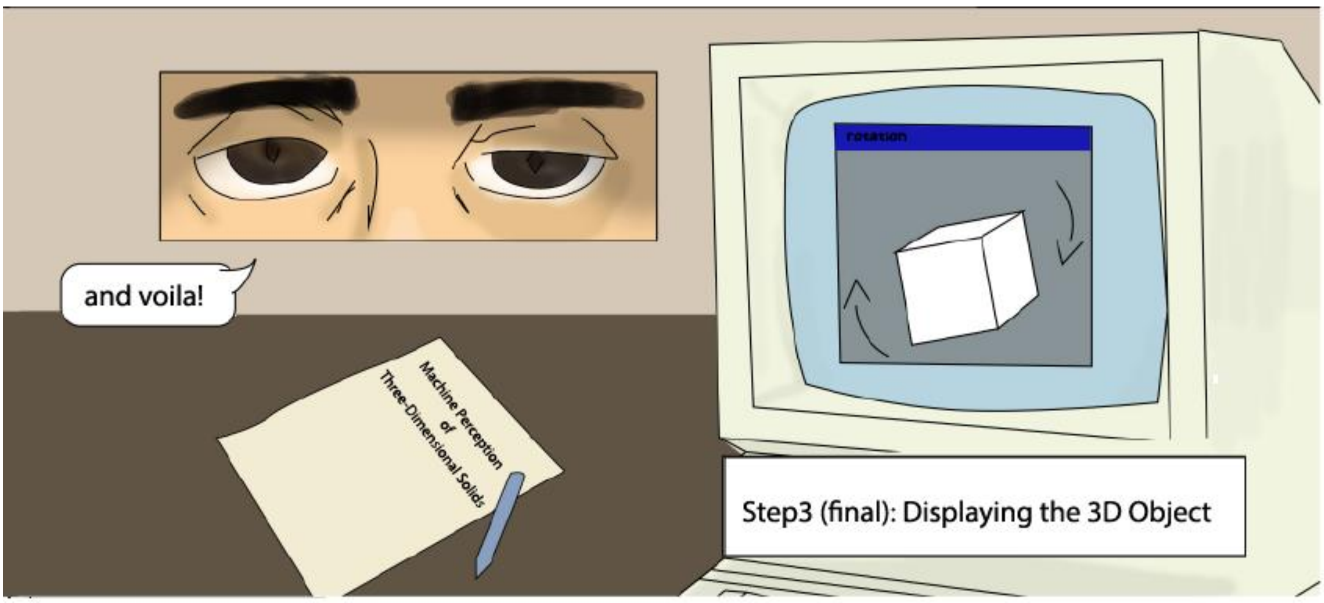
Y. ABOURA, M. LATIF, MASTER 2 VISUAL COMPUTING, USTHB



Visual Computing Magazine

## CHRONICLE OF COMPUTER VISION
## EVENT: 3D information from 2D photographs , 1963

Y. ABOURA, M. LATIF, MASTER 2 VISUAL COMPUTING, USTHB



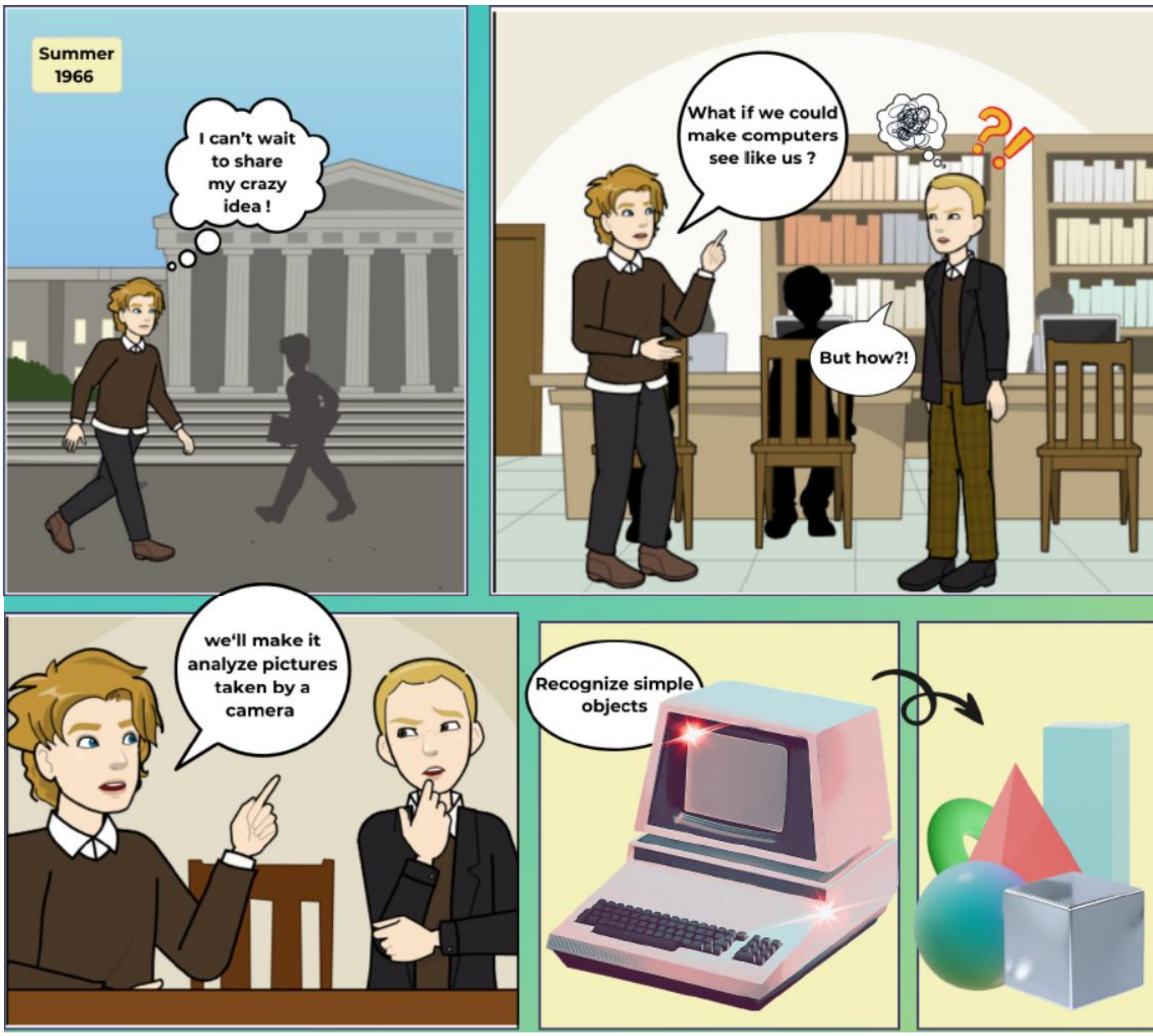

**Reference:**

**Roberts, Lawrence G., Machine perception of three-dimensional solids. Thesis (Ph. D.). Massachusetts Institute of Technology, Dept. of Electrical Engineering, 1963.**

# CHRONICLE OF COMPUTER VISION
## EVENT: Summer Vision Project, 1966
H. KADIRI, MASTER 2 VISUAL COMPUTING, USTHB

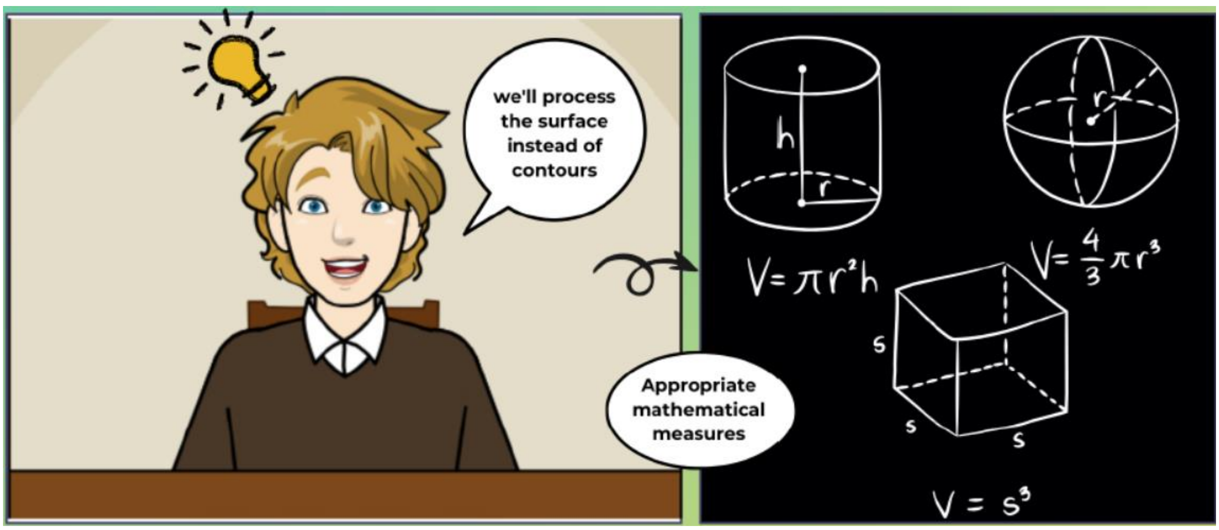
# CHRONICLE OF COMPUTER VISION
## EVENT: Summer Vision Project, 1966

H. KADIRI, MASTER 2 VISUAL COMPUTING, USTHB



Visual Computing Magazine

# CHRONICLE OF COMPUTER VISION
## EVENT: Summer Vision Project, 1966
H. KADIRI, MASTER 2 VISUAL COMPUTING, USTHB

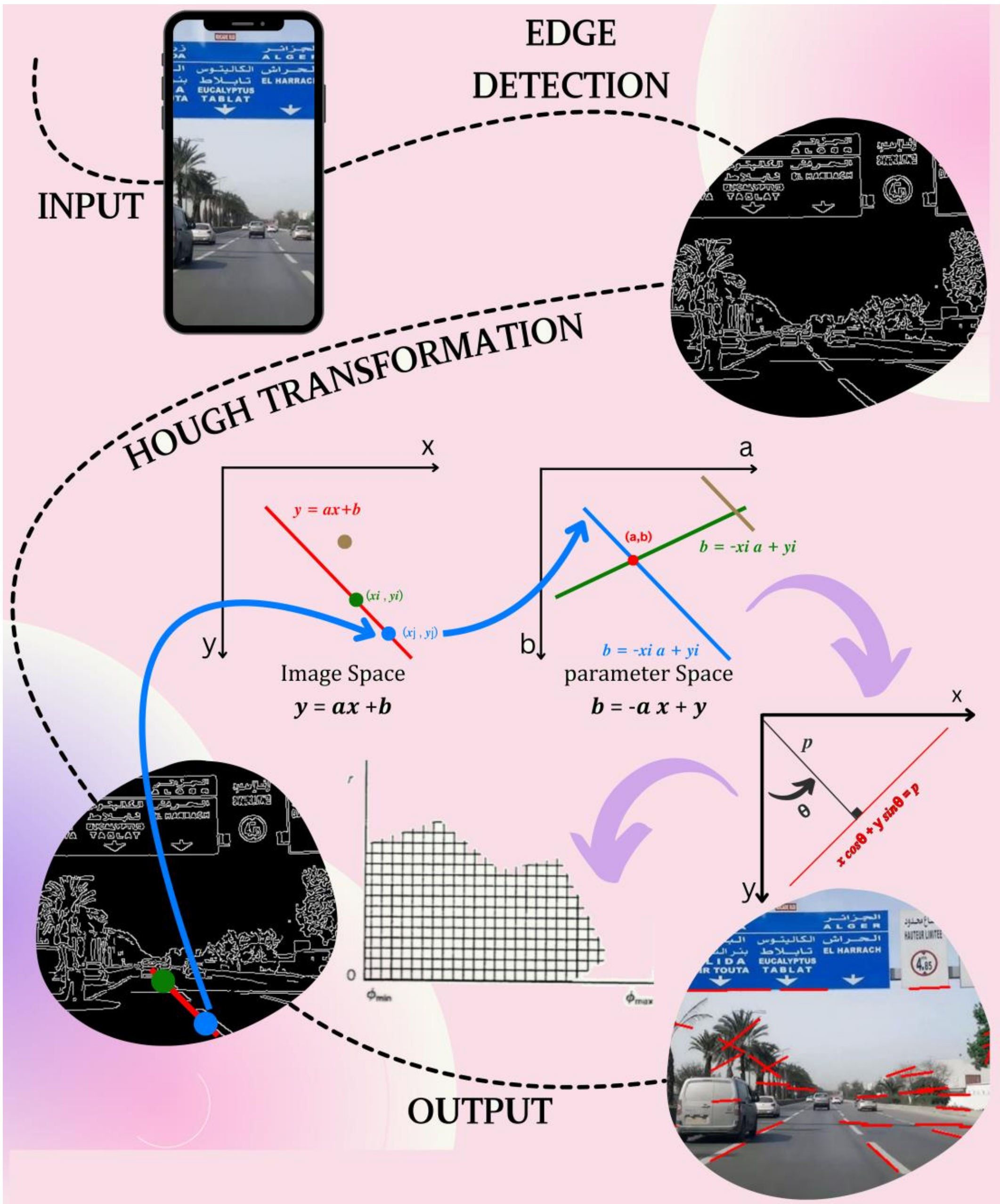




**Reference**:
Seymour Papert, The summer vision project, MIT, 1966
https://people.csail.mit.edu/brooks/idocs/AIM-100.pdf

Visual Computing Magazine

# CHRONICLE OF COMPUTER VISION
## EVENT: HOUGH Transform, 1972
A. BOUCENNA, A. OUHIB, MASTER 2 VISUAL COMPUTING, USTHB




**Reference**: Richard O. Duda, Peter E. Hart. Use of the Hough transformation to detect lines and curves in pictures. Communications of the ACM, Volume 15, Issue 2, 1972

**CHRONICLE OF COMPUTER VISION**

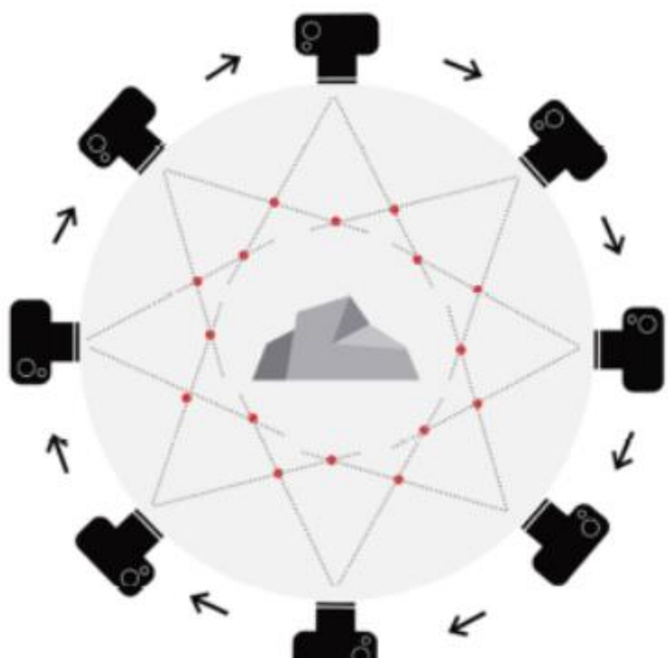
# EVENT: SFM Structure From Motion, 1981

H. SAAD, M. MEZIANI, MASTER 2 VISUAL COMPUTING, USTHB

*Structure from Motion is an algorithm that allows to recreate 3D objects into a cloud of points by simply taking a few images form different angles*


## Image Acquisition

*Capture multiple 2D images of an object or scene from different angles using a regular camera*



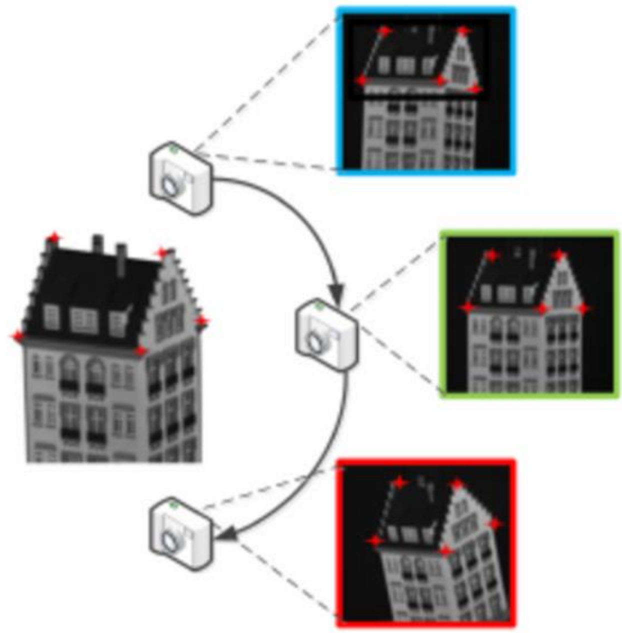
# CHRONICLE OF COMPUTER VISION

# EVENT: SFM Structure From Motion, 1981

**H. SAAD, M. MEZIANI, MASTER 2 VISUAL COMPUTING, USTHB**

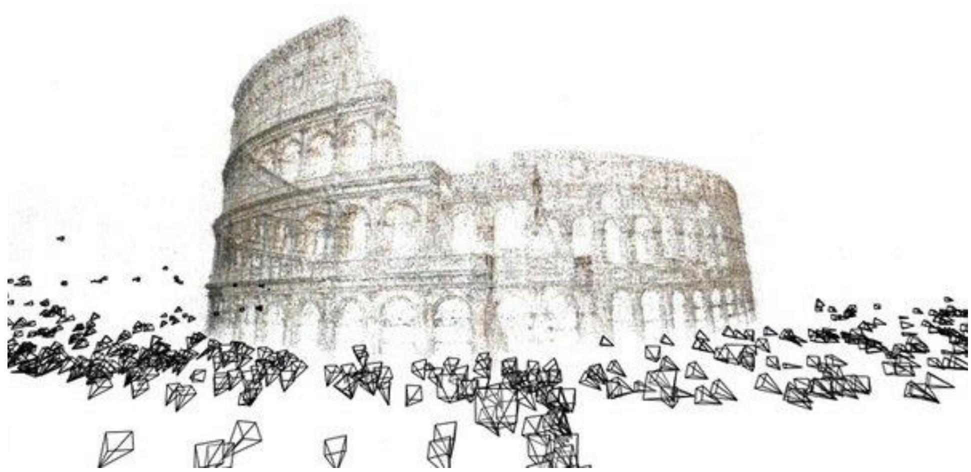
## Feature Detection & Matching

*The system identifies key features across the images and matches them to find overlapping parts between different pictures.*



## Sparse 3D Reconstruction

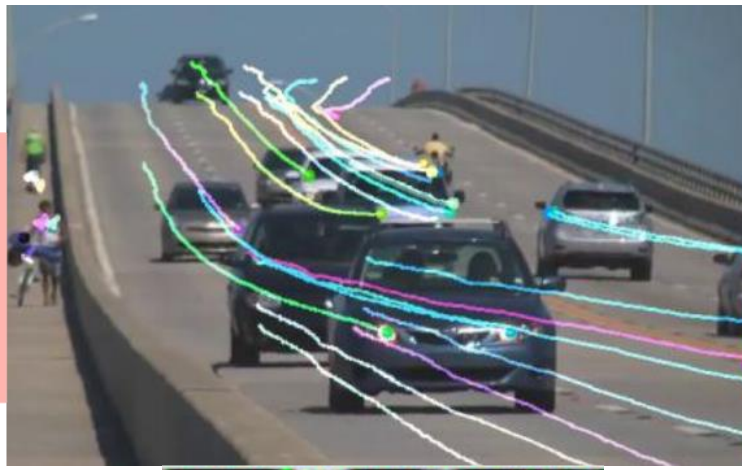
*With camera poses known, the algorithm builds a basic 3D model by plotting key points in 3D space.*



**Reference**: H.C. Longuest-Higgins. A computer algorithm for reconstructing a scene from two projections. Nature, vol. 293, pp. 133-135, 1981

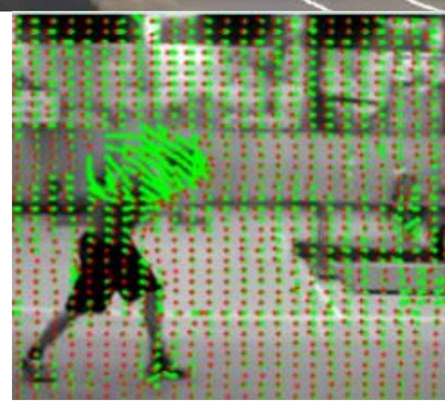
## CHRONICLE OF COMPUTER VISION
## EVENT: Optical Flow, 1981
K. AMROUNI, R. KADOUM, MASTER 2 VISUAL COMPUTING, USTHB

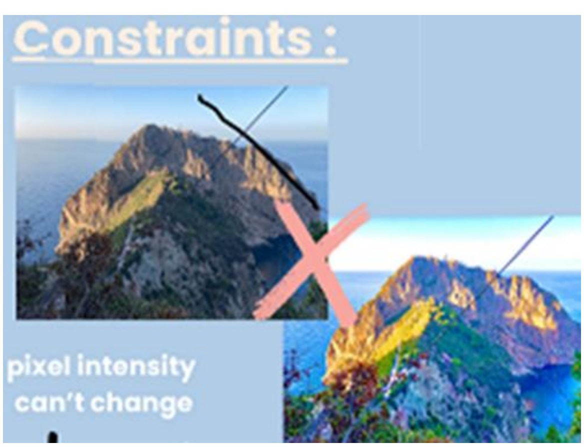
Optical flow is used in computer vision to estimate the motion of objects between consecutive frames in a video.
It looks at how pixels in an image move from one frame to the next and creates motion maps.



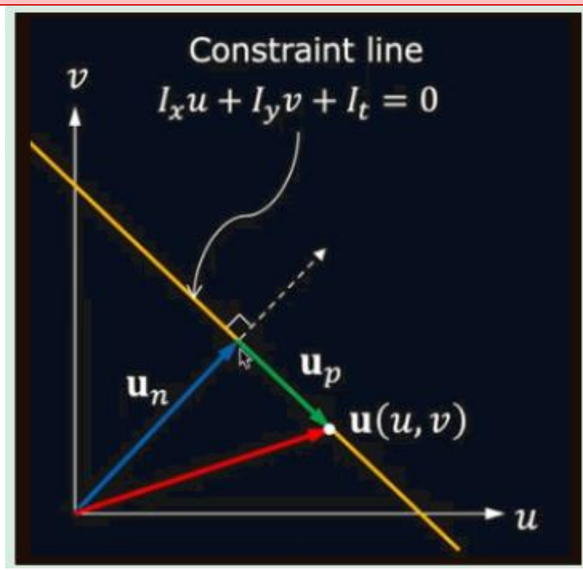
Motion flow can be used for :

- object tracking
- Robotics
- VR, to make a more interactive world

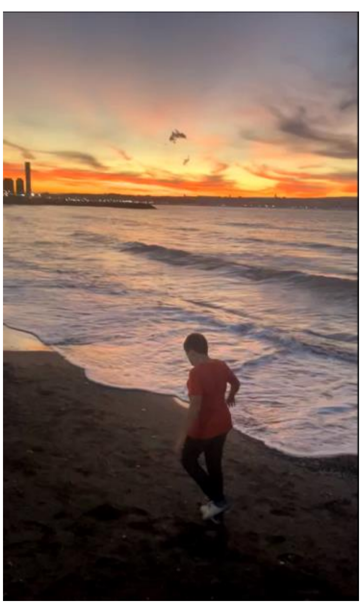


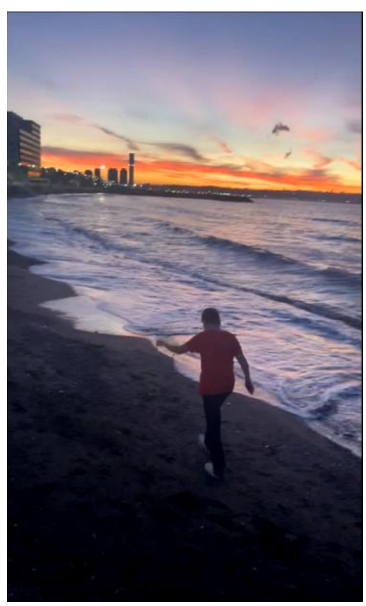


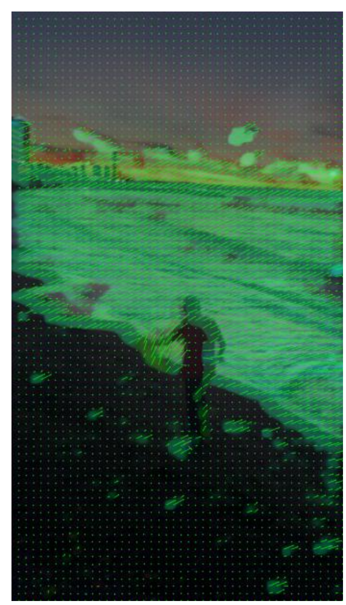

## CHRONICLE OF COMPUTER VISION
## EVENT: Optical Flow, 1981
K. AMROUNI, R. KADOUM, MASTER 2 VISUAL COMPUTING, USTHB




The value of the optical flow is on the constraint line.


Frame 1


Frame 2

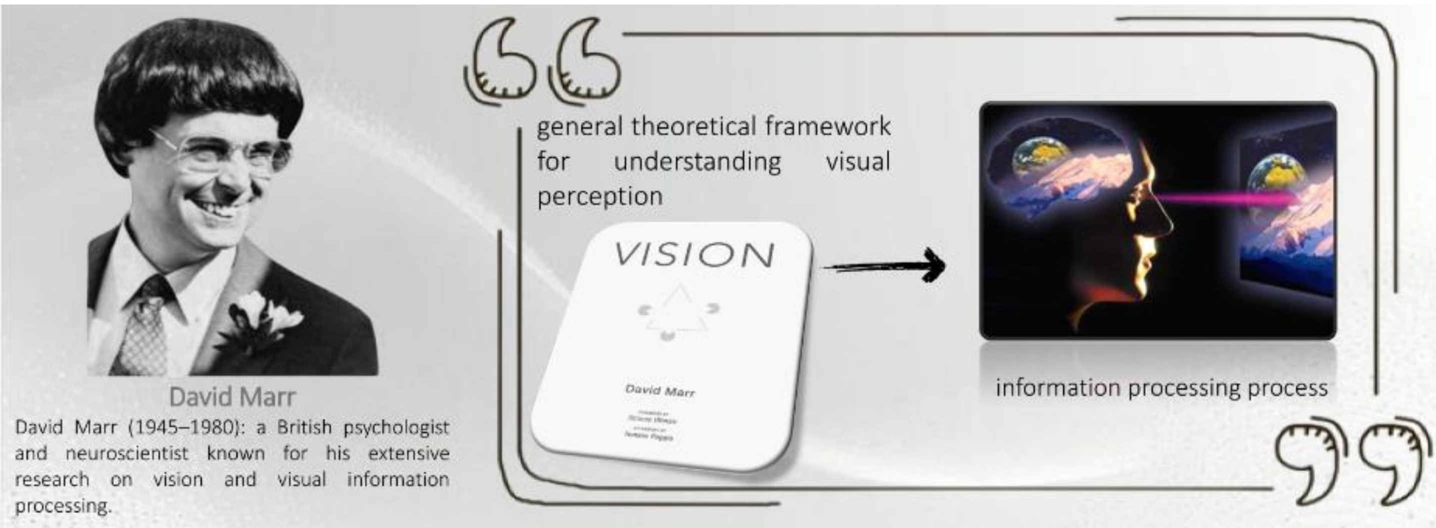

Optical flow

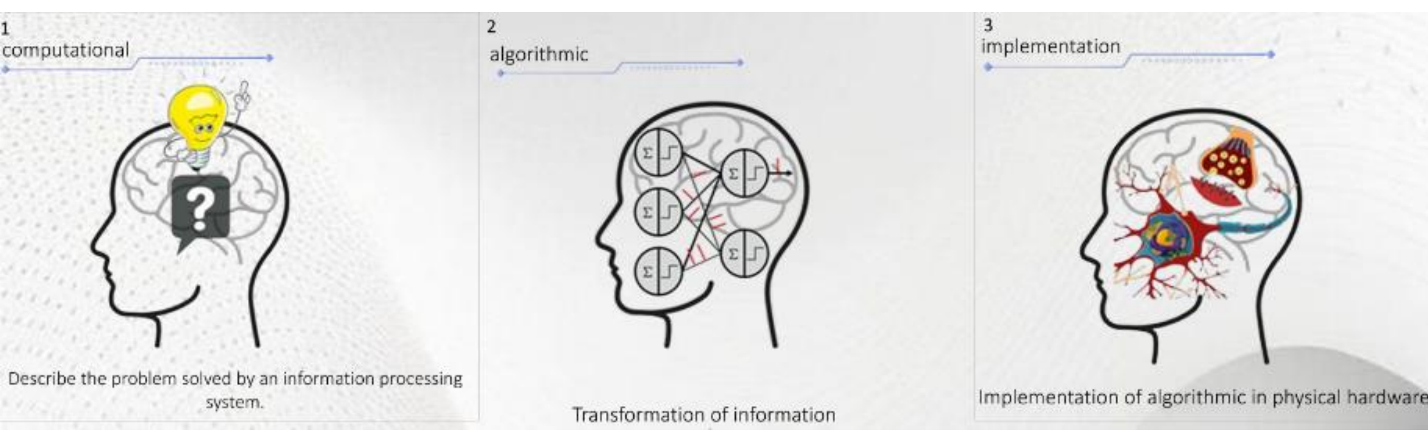
**Reference**: Berthold K.P. Horn, Brian G. Schunck. Determining optical flow. Artificial Intelligence. Volume 17, Issues 1–3, August 1981, Pages 185-203

Visual Computing Magazine

# CHRONICLE OF COMPUTER VISION
## EVENT: David Marr's Paradigm, 1982

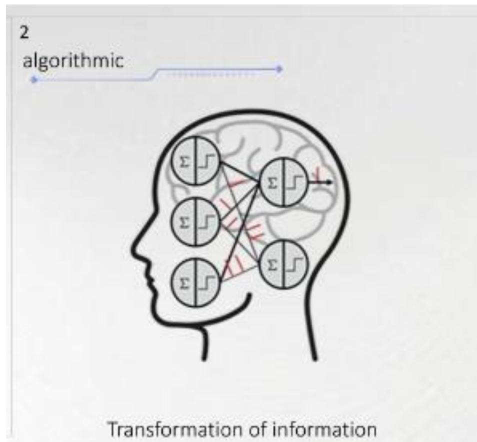
H. KACIOUSSALEH, Z. BISKRI, MASTER 2 VISUAL COMPUTING, USTHB

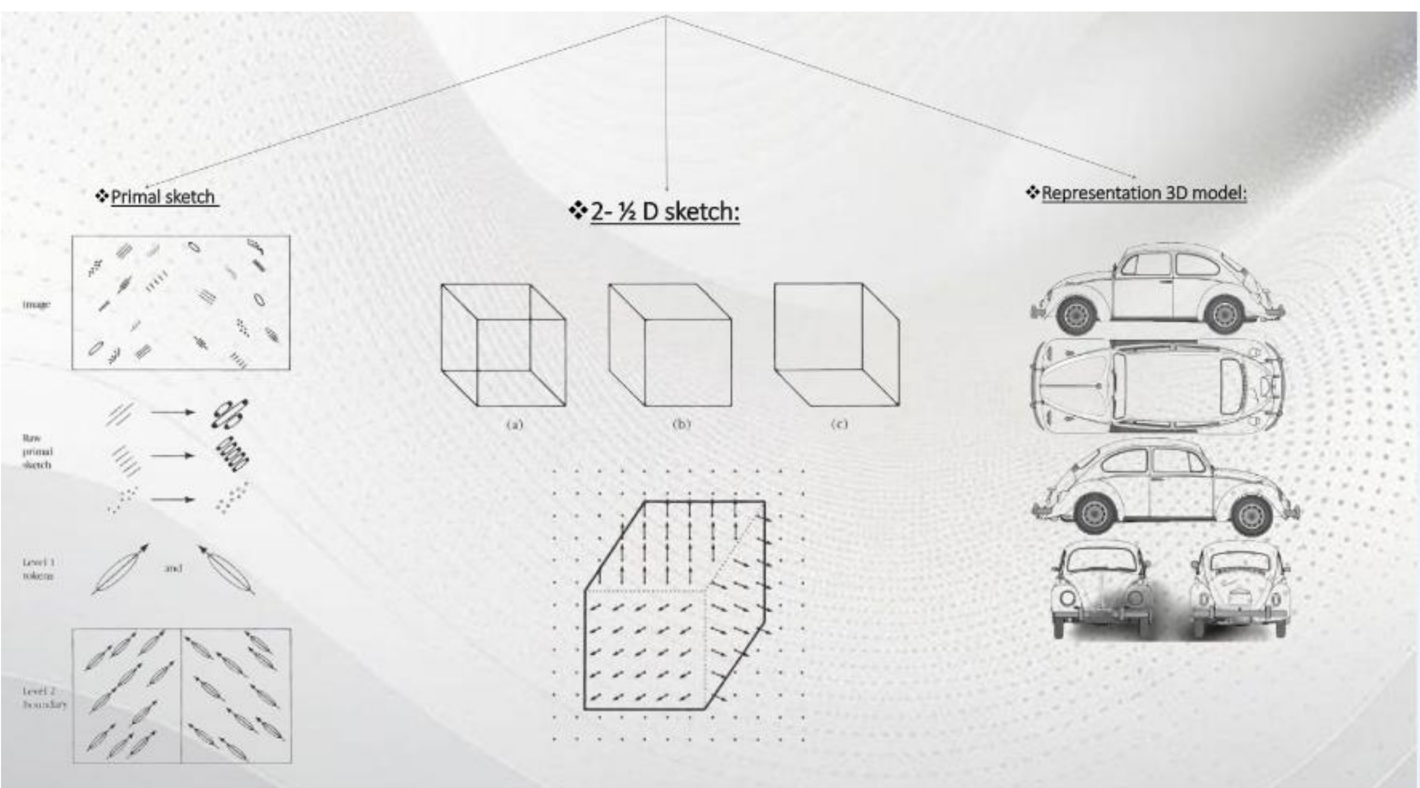



REFERENCE :David Marr. Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. 1982.

# CHRONICLE OF COMPUTER VISION
## EVENT: David Marr's Paradigm, 1982
H. KACIOUSSALEH, Z. BISKRI, MASTER 2 VISUAL COMPUTING, USTHB



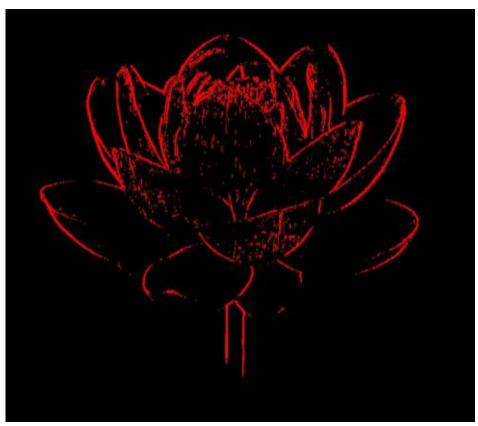
## CHRONICLE OF COMPUTER VISION
## EVENT: Canny Edge Detector, 1986

A. R. BADRI, A. MEKKAOUI, MASTER 2 VISUAL COMPUTING, USTHB

Original image

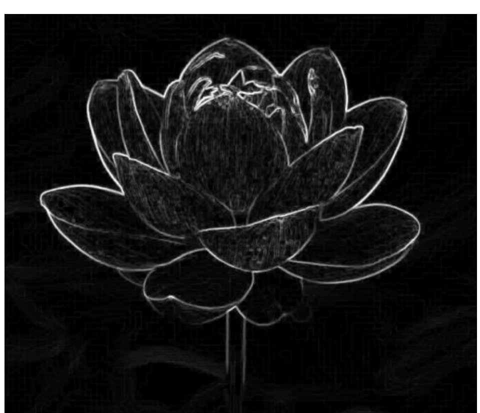
Conversion to Grayscale

Blurring with a Gaussian



X-Gradient

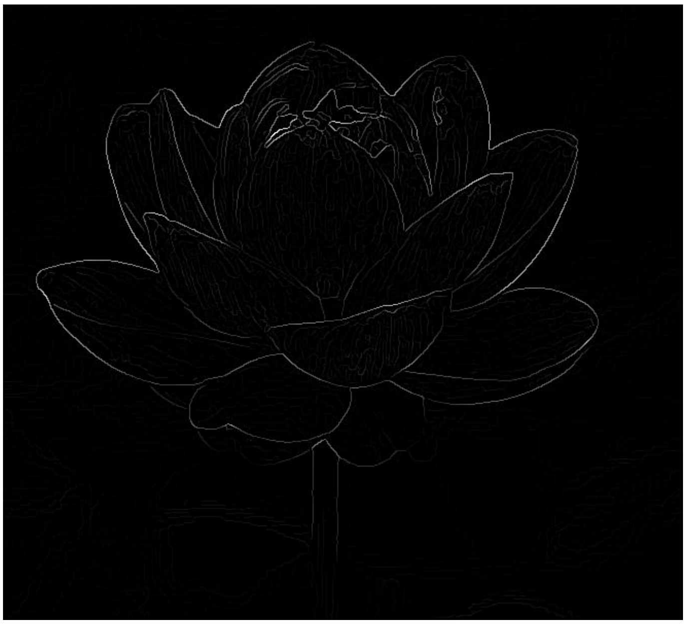
Y-Gradient



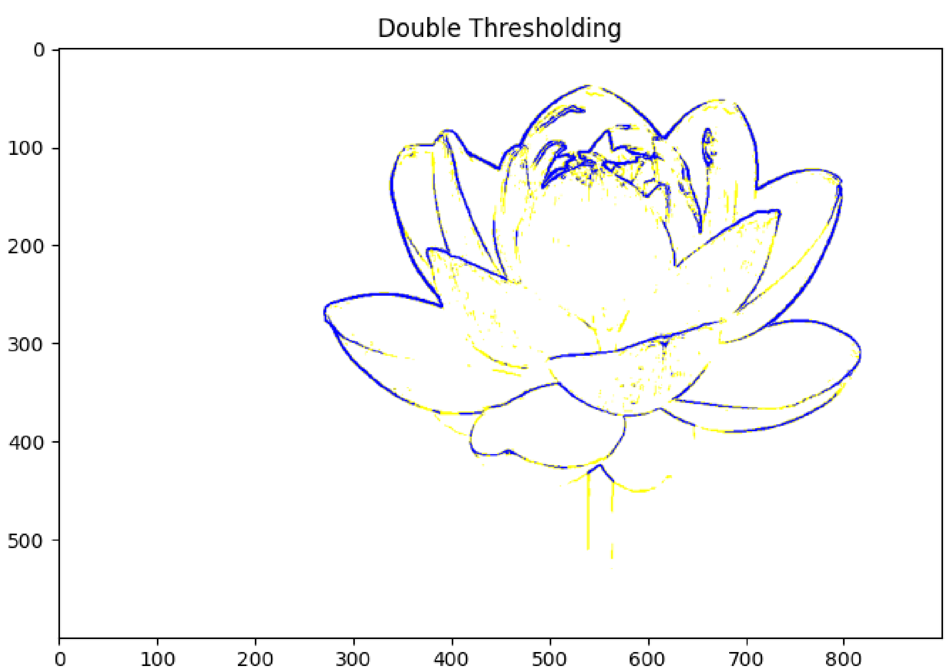
Sobel result magnitude

## CHRONICLE OF COMPUTER VISION
## EVENT: Canny Edge Detector, 1986

A. R. BADRI, A. MEKKAOUI, MASTER 2 VISUAL COMPUTING, USTHB

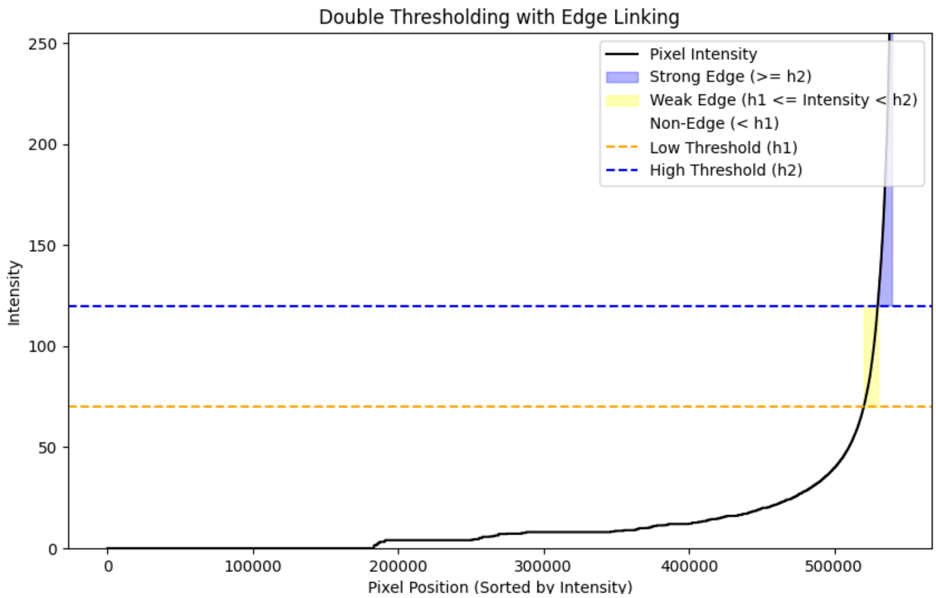


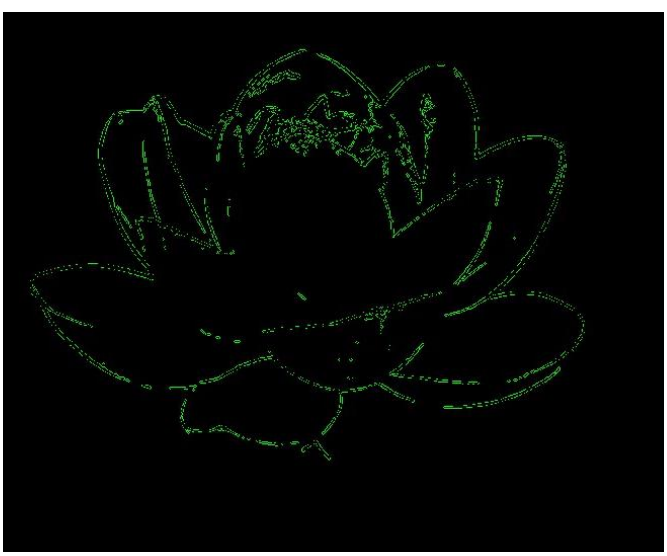
**Non Maxima suppression**



# CHRONICLE OF COMPUTER VISION
## EVENT: Canny Edge Detector, 1986

A. R. BADRI, A. MEKKAOUI, MASTER 2 VISUAL COMPUTING, USTHB



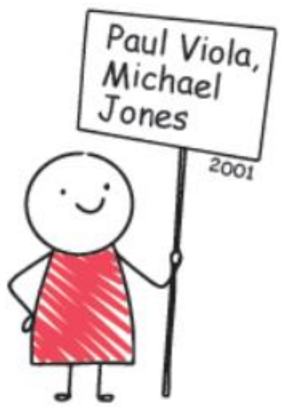
Edges linked only

Edge tracking by hysteresis

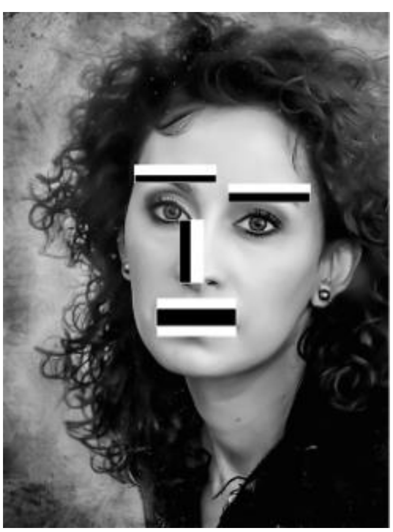
**Reference**: J. Canny, "A Computational Approach to Edge Detection," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679-698, Nov. 1986

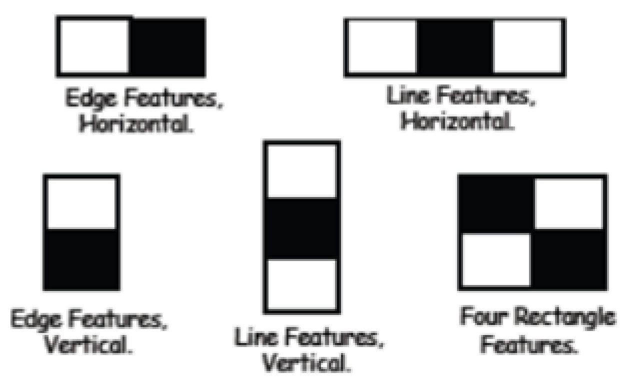
# CHRONICLE OF COMPUTER VISION
## EVENT: Viola-John Detector, 2001

Y. BELAIDI, K. MAZROU, MASTER 2 VISUAL COMPUTING, USTHB










Extraction of Haar features

Visual Computing Magazine

# CHRONICLE OF COMPUTER VISION

## EVENT: Viola-John Detector, 2001

Y. BELAIDI, K. MAZROU, MASTER 2 VISUAL COMPUTING, USTHB

Source image

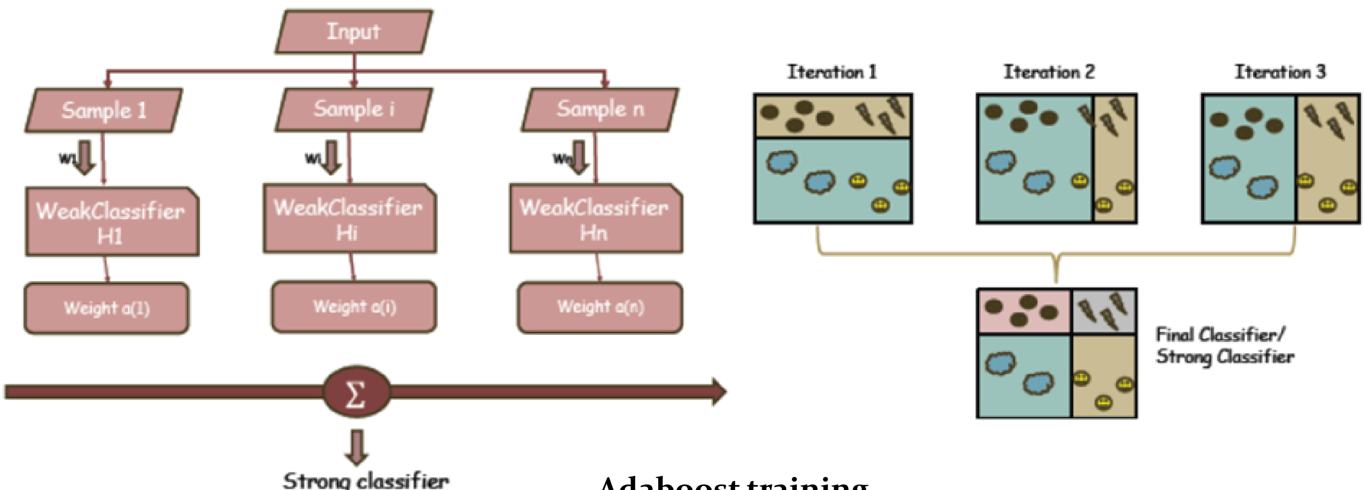
| | | | |
|---|---|---|---|
| 0 | 1 | 1 | 1 |
| 1 | 2 | 2 | 3 |
| 1 | 2 | 1 | 1 |
| 1 | 3 | 1 | 0 |

Integral image

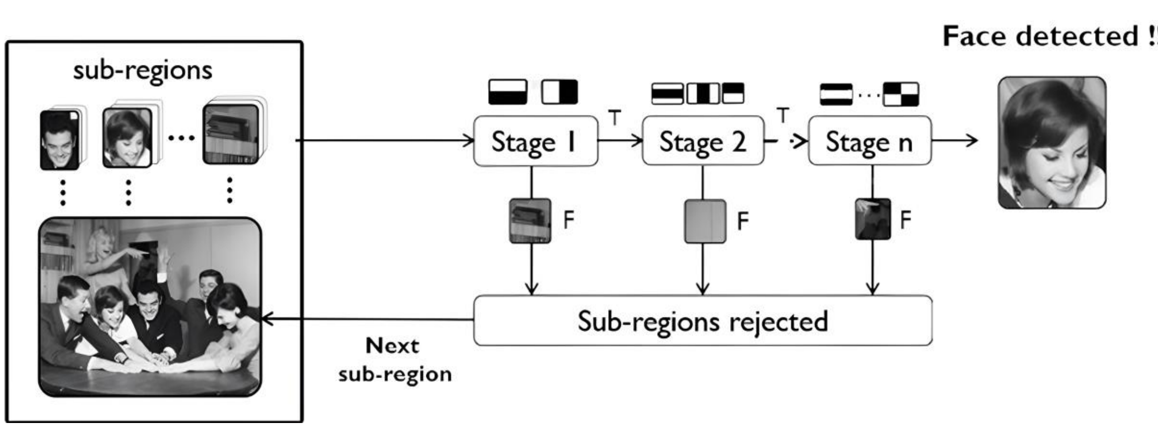
| | | | |
|---|---|---|---|
| 0 | 1 | 2 | 3 |
| 1 | 4 | 7 | 11 |
| 2 | 7 | 11 | 16 |
| 3 | 11 | 16 | 21 |

0+1+1+1+2+2=7

Compute the Integral of image




Adaboost training

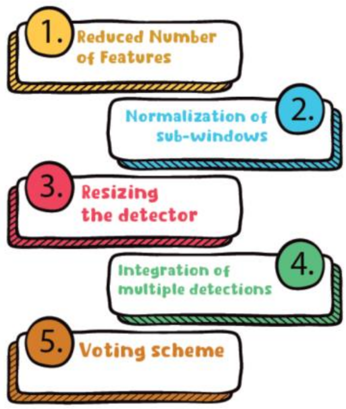


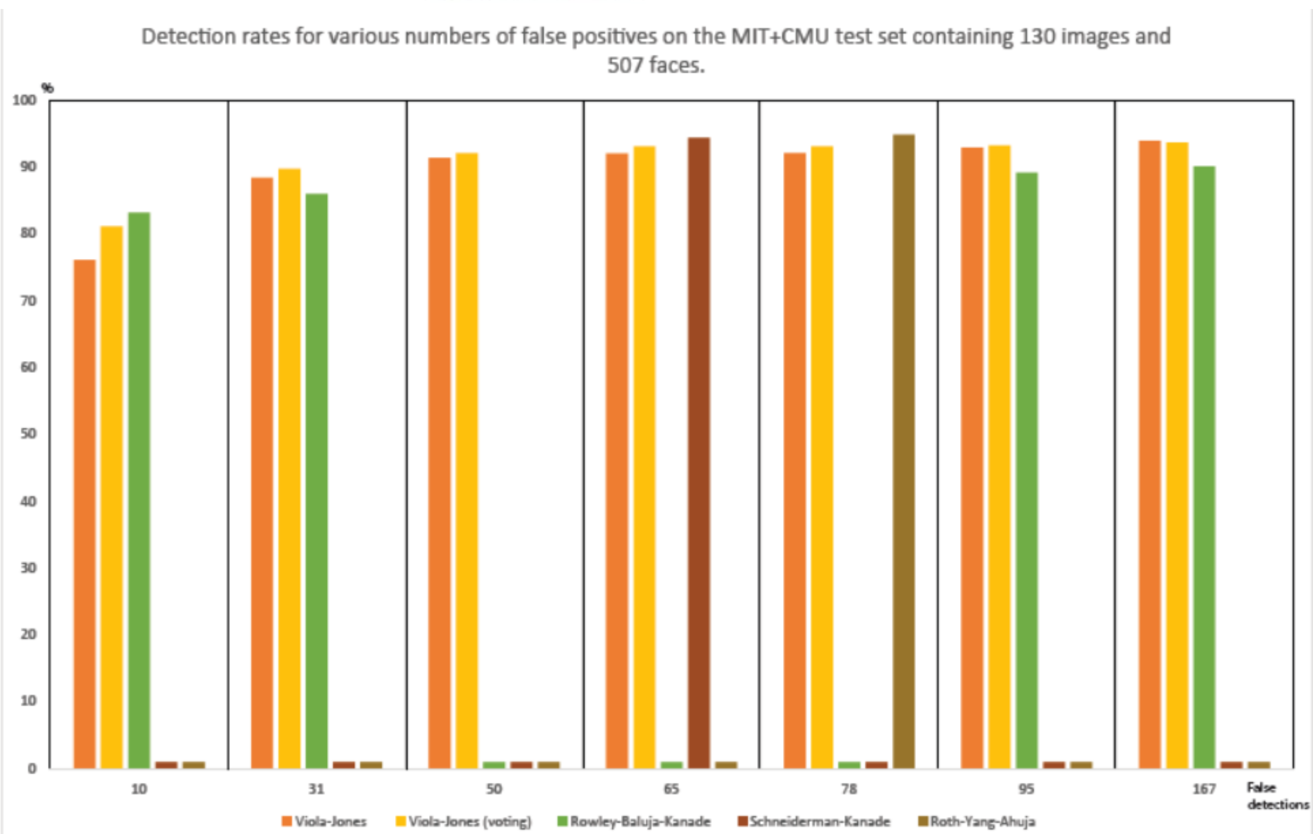

Cascade of classifiers

# CHRONICLE OF COMPUTER VISION
## EVENT: Viola-John Detector, 2001

Y. BELAIDI, K. MAZROU, MASTER 2 VISUAL COMPUTING, USTHB

**Improvements**

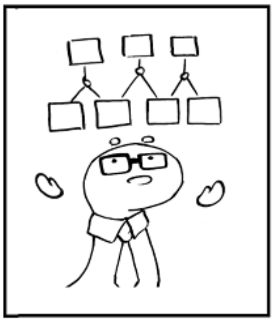


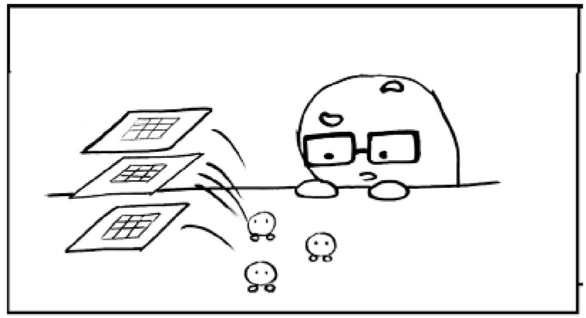

**Reference:** Paul Viola, Michael Jones. Rapid Object Detection using a Boosted Cascade of Simple Features, CVPR 2001

Visual Computing Magazine

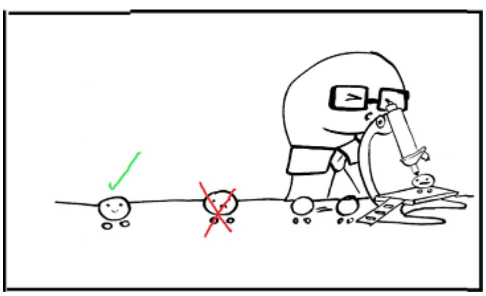
# CHRONICLE OF COMPUTER VISION
## EVENT: SIFT, 2004
L.R. BOUGUERRA, MASTER 2 VISUAL COMPUTING, USTHB

First we create a Scale-Space. We apply the gaussian with value Segma and up scale it. then for each two pairs apply the DoG.

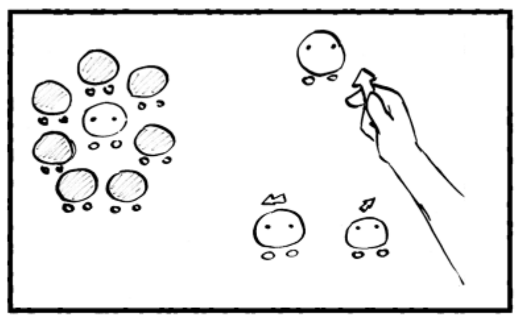


we detect the "Key ponits". using a 3*3 block and choosing the minima or maxima .

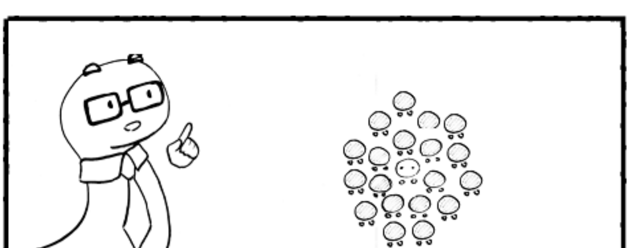


we have to get rid of unstable points:

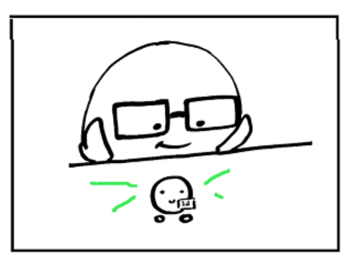
- the series of Tylor with the expansion of DoG for low-contrast points
-Hessian matrix to eliminait edges.



For the orientation, we select it neighboorhood and compute the gradient magnitude and direction, then create a direction histogram, it peak will be the orientation.



We'll repeat the same step except with a bigger neighborhood. and the collection of histograms is the ID.



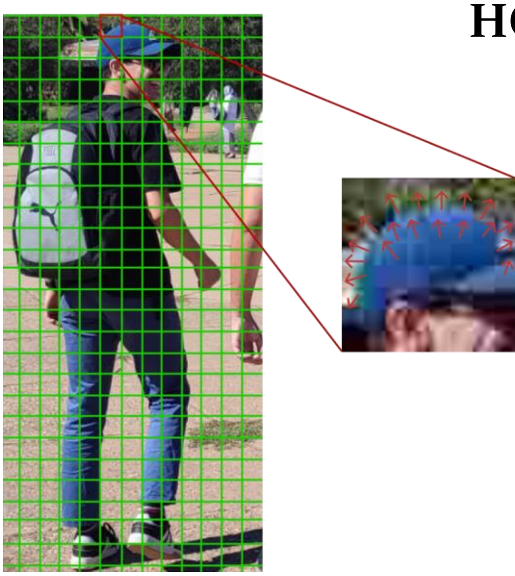
We add it in the data base.

**Reference:** Lowe, D.G. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision 60, 91–110 (2004).

# CHRONICLE OF COMPUTER VISION
## EVENT: HOG: Histogram of Oriented Gradients, 2005

A. AOUDJ, Y. CHELBI, MASTER 2 VISUAL COMPUTING, USTHB

## HOG Computation Process



| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 2 | 3 | 4 | 4 | 3 | 4 | 2 | 2 |
| 5 | 11 | 17 | 13 | 7 | 9 | 3 | 4 |
| 11 | 21 | 23 | 27 | 22 | 17 | 4 | 6 |
| 23 | 99 | 165 | 135 | 85 | 31 | 26 | 2 |
| 91 | 155 | 133 | 136 | 144 | 152 | 57 | 28 |
| 98 | 196 | 76 | 38 | 26 | 60 | 170 | 51 |
| 165 | 60 | 60 | 27 | 77 | 85 | 43 | 136 |
| 71 | 13 | 34 | 23 | 108 | 27 | 48 | 110 |

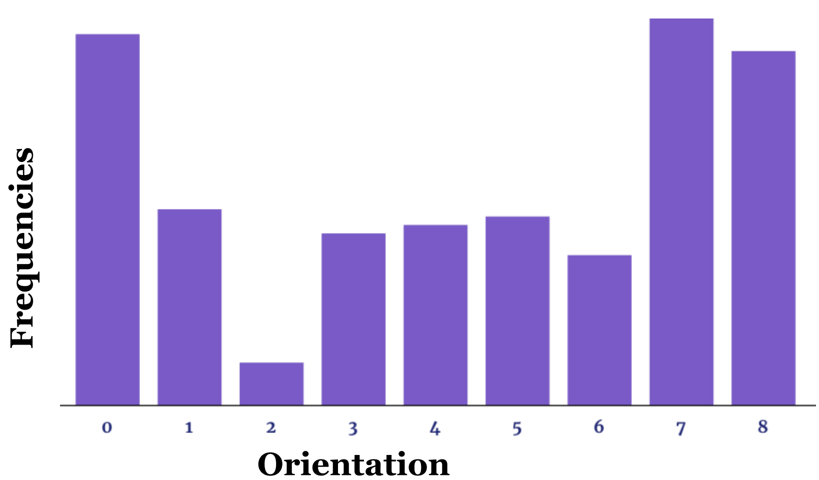
**Magnitude**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 80 | 36 | 5 | 10 | 0 | 64 | 90 | 73 |
| 37 | 9 | 9 | 179 | 78 | 27 | 169 | 166 |
| 87 | 136 | 173 | 39 | 102 | 163 | 152 | 176 |
| 76 | 13 | 1 | 168 | 159 | 22 | 125 | 143 |
| 120 | 70 | 14 | 150 | 145 | 144 | 145 | 143 |
| 58 | 86 | 119 | 98 | 100 | 101 | 133 | 113 |
| 30 | 65 | 157 | 75 | 78 | 165 | 145 | 124 |
| 11 | 179 | 91 | 4 | 110 | 17 | 133 | 110 |

**Orientation**

| 0 | 20 | 40 | 60 | 80 | 100 | 120 | 140 | 160 |
|---|---|---|---|---|---|---|---|---|
| 2 | 2 | | | 2 | | | | |

**Histogram of gradients**



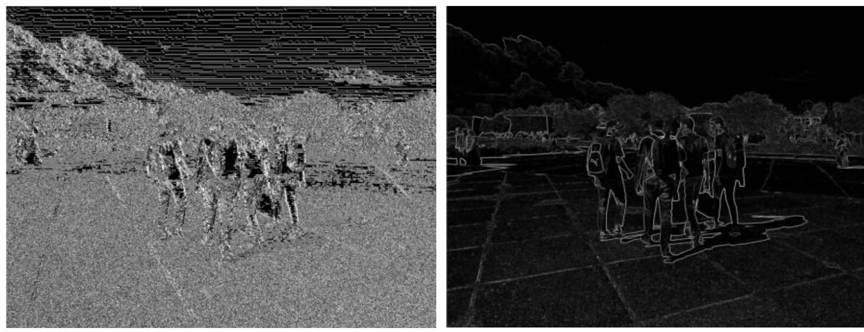
Visual Computing Magazine

# CHRONICLE OF COMPUTER VISION
## EVENT: HOG: Histogram of Oriented Gradients, 2005

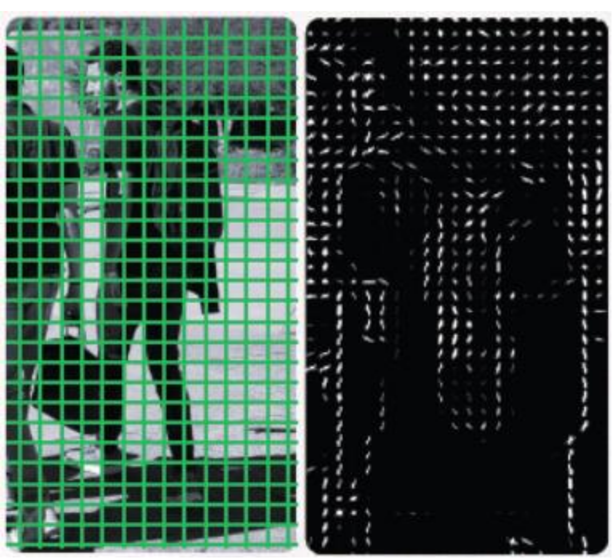
A. AOUDJ, Y. CHELBI, MASTER 2 VISUAL COMPUTING, USTHB

## HOG in Action: Real Images Analysis

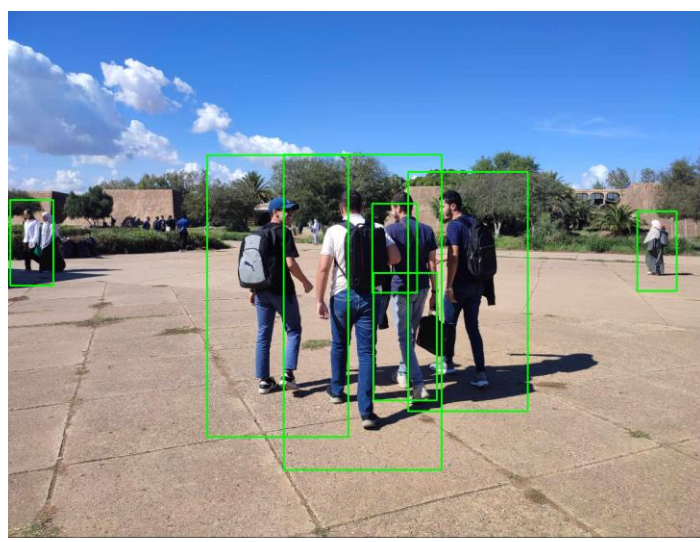
Grayscale conversion



Computation of Gradient (Magnitude, Orientation)



Grouping cells into blocks and computing histogram for each block.

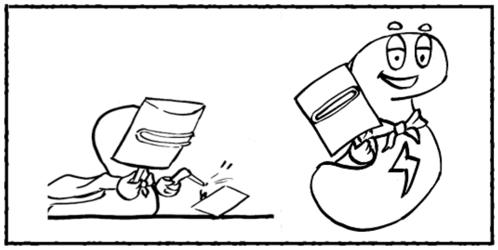


Block normalization, classification using SVM to detect humans.

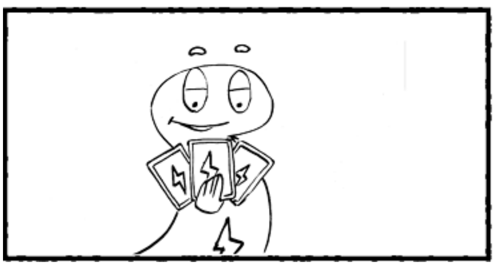
**Reference:** Dalal, N., & Triggs, B. (2005). Histograms of Oriented Gradients for Human Detection. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Vol. 1, pp. 886-893.

# CHRONICLE OF COMPUTER VISION
## EVENT: SURF, 2006

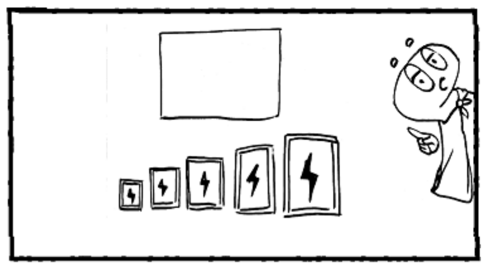
L.R. BOUGUERRA, MASTER 2 VISUAL COMPUTING, USTHB

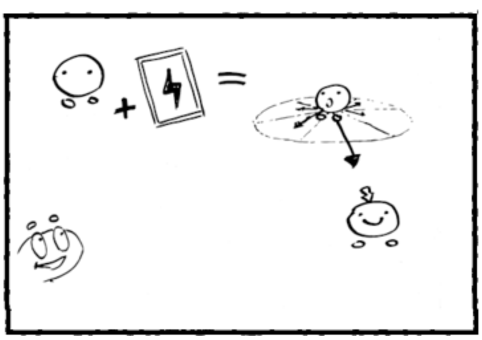
The first thing is to create an integral image from the original image.



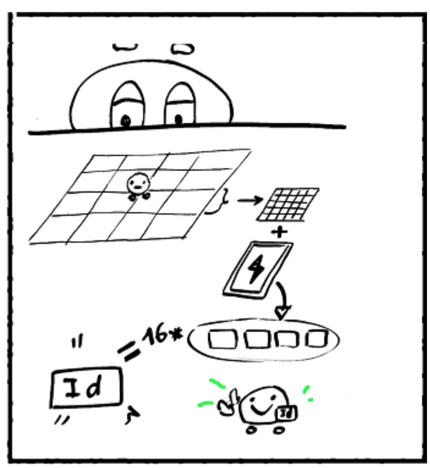
To detect key points we'll manipulate the integral image by using *Magic Cards*.



No more changes will occure on the image instead the magic cards will upscale.



we get the orientation by applying a magic card "Haar-wavelet", The point will follow it strongest direction.



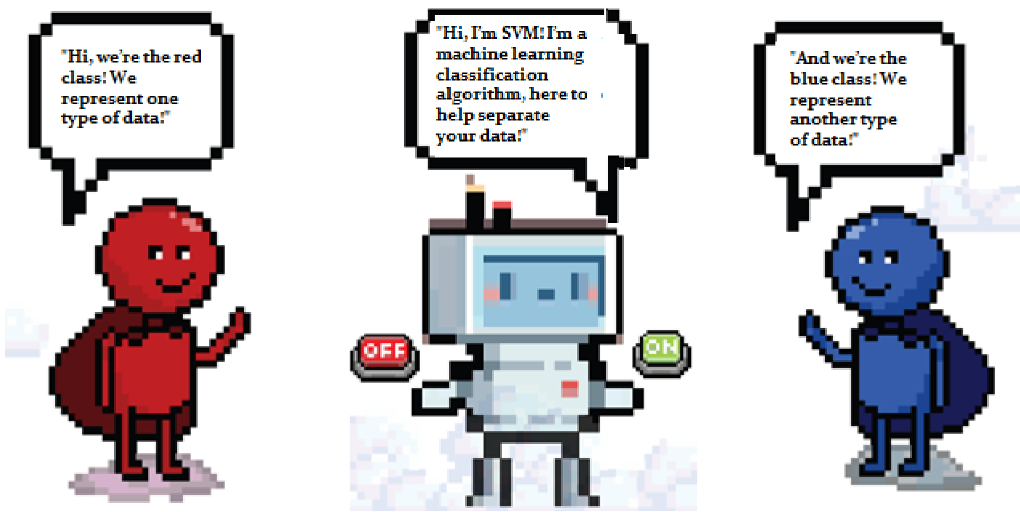

We'll repeat the earlier step with a wider neighborhood. The collection of all the values across all regions will be our ID.

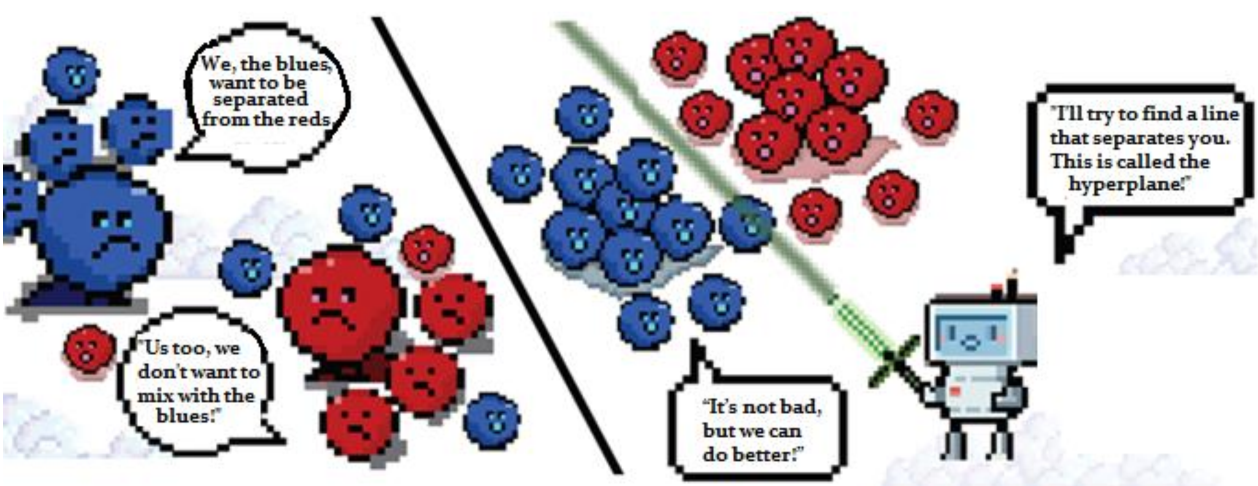
**Reference:** Bay, H., Tuytelaars, T., Van Gool, L. (2006). SURF: Speeded Up Robust Features. In: Leonardis, A., Bischof, H., Pinz, A. (eds) Computer Vision – ECCV 2006. ECCV 2006.

Visual Computing Magazine

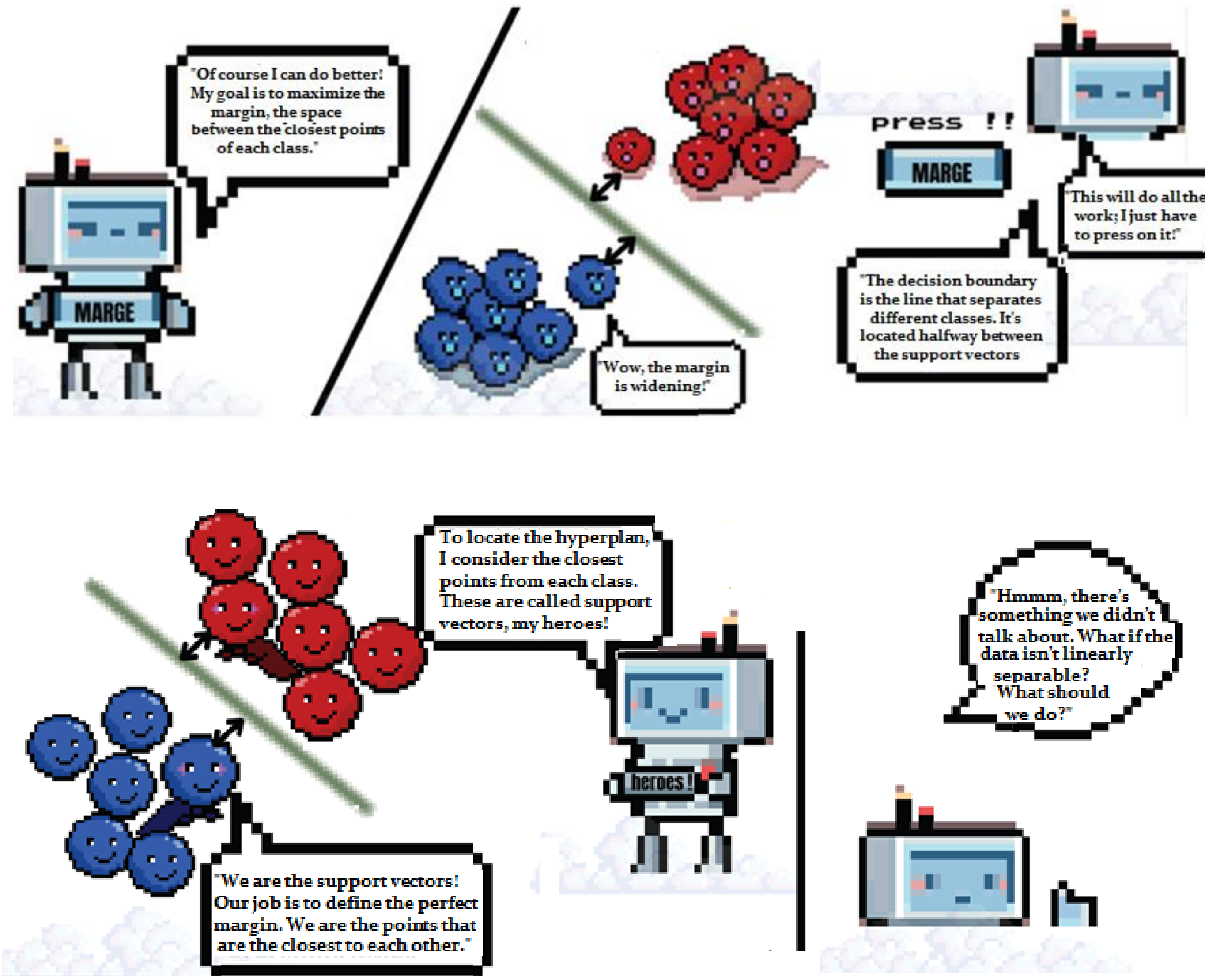
# CHRONICLE OF COMPUTER VISION
## EVENT: SVM, 2008

A. TAREB, A. KHOUAS, MASTER 2 VISUAL COMPUTING, USTHB



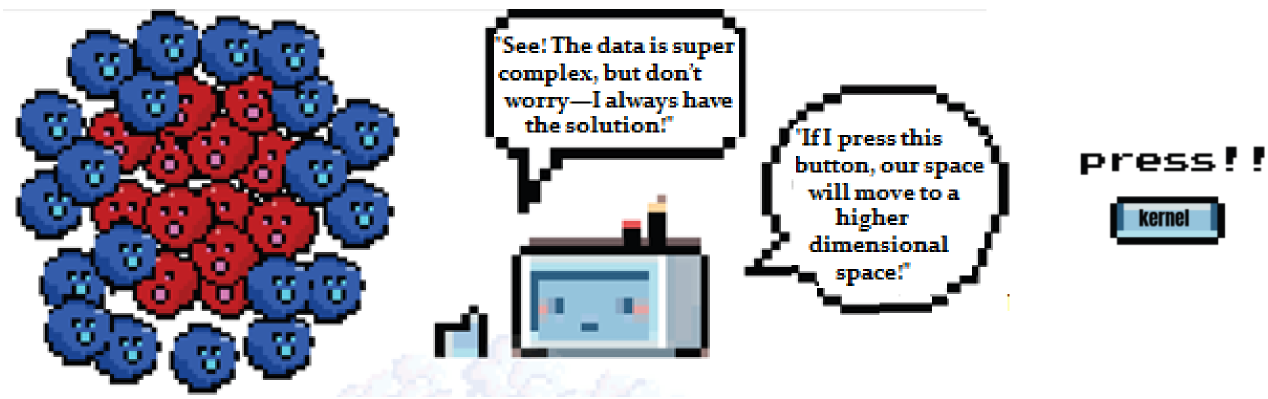



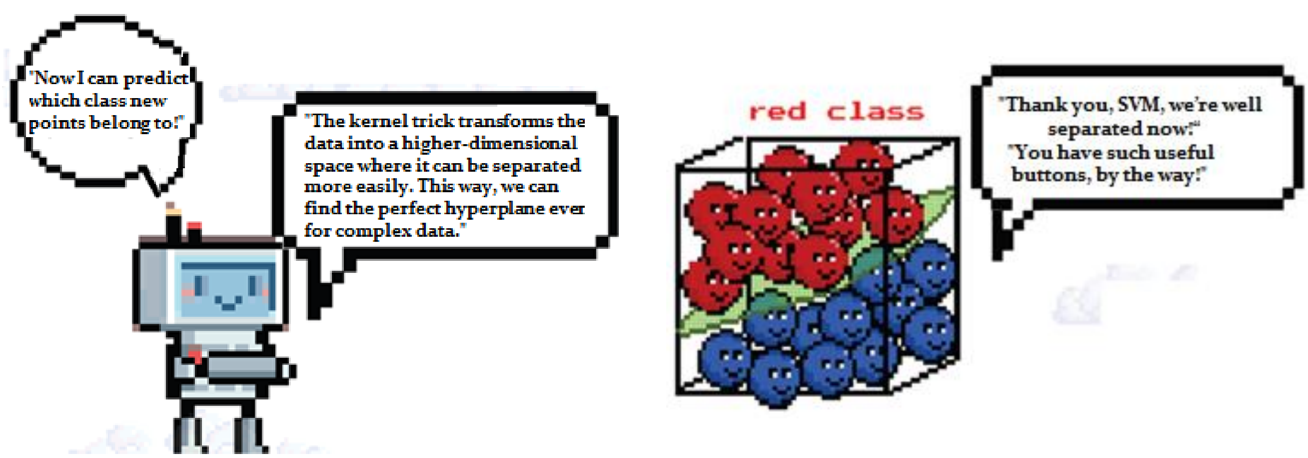
# CHRONICLE OF COMPUTER VISION
## EVENT: SVM, 2008
A. TAREB, A. KHOUAS, MASTER 2 VISUAL COMPUTING, USTHB



Visual Computing Magazine

# CHRONICLE OF COMPUTER VISION
## EVENT: SVM, 2008
A. TAREB, A. KHOUAS, MASTER 2 VISUAL COMPUTING, USTHB




**Reference:** Tzotsos,A. Argialas, D. (2008). Support Vector Machine Classification for Object-Based Image Analysis. Object-Based Image Analysis. Lecture Notes in Geoinformation and Cartography. https://doi.org/10.1007/978-3-540-77058-9_36

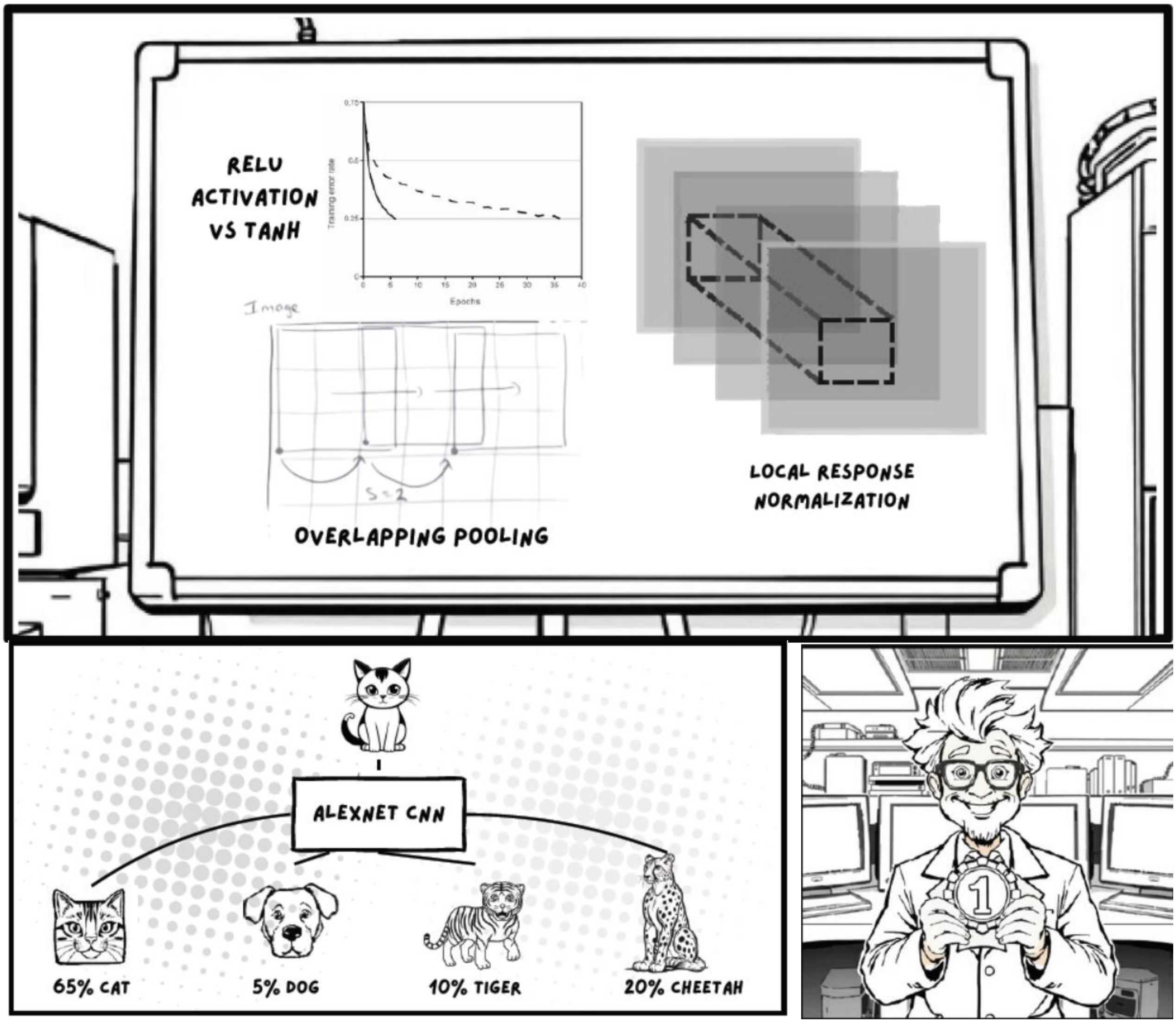
Visual Computing Magazine

# CHRONICLE OF COMPUTER VISION
## EVENT: AlexNet CNN, 2012
**A. ADLAOUI, A. BELLOULA**, MASTER 2 VISUAL COMPUTING, USTHB

Visual Computing Magazine

# CHRONICLE OF COMPUTER VISION
## EVENT: AlexNet CNN, 2012

**A. ADLAOUI, A. BELLOULA, MASTER 2 VISUAL COMPUTING, USTHB**




**Reference**: Authors: Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, ImageNet classification with deep convolutional neural networks », NIPS'12: Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1 Pages 1097 - 1105

Visual Computing Magazine

# CHRONICLE OF COMPUTER VISION

## EVENT: Generative Adversarial Networks , 2014.

A. KACI AISSA, R. NINE, MASTER 2 VISUAL COMPUTING, USTHB

GENERATOR

**Objective :**

Create data that mimics the real dataset.

**Purpose of training:**

Deceiving the discriminator.

**Type of loss :**

Minimize the success rate.

DESCRIMINATOR

**Objective :**

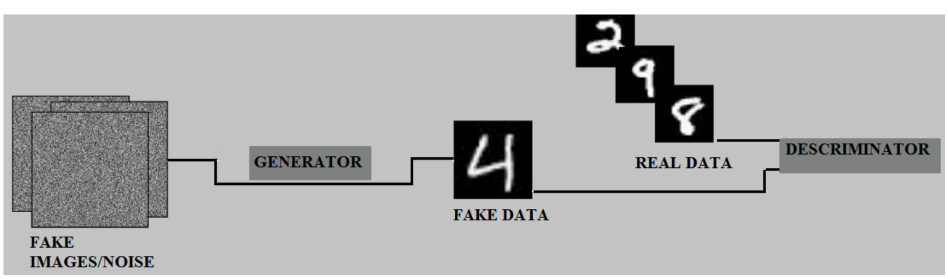
Distinguish between real data from generated data.

**Purpose of training:**

Correctly classify real and fake data..

**Type of loss :**

Maximize classification accuracy.




Reference: Goodfellow, Ian & Pouget-Abadie, Jean & Mirza, Mehdi & Xu, Bing & Warde-Farley, David & Ozair, Sherjil & Courville, Aaron & Bengio, Y.. (2014). Generative Adversarial Networks. Advances in Neural Information Processing Systems.

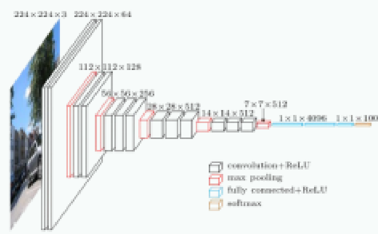
Visual Computing Magazine

# CHRONICLE OF COMPUTER VISION
# EVENT: GoogleNet, 2014, VGGNet, 2015.

S. BELKACEMI, L. ROUIBAH, MASTER 2 VISUAL COMPUTING, USTHB

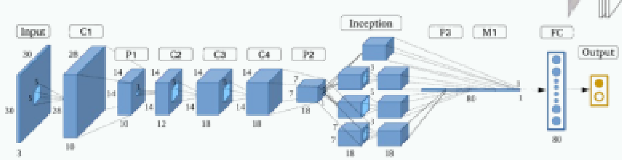
**what CNN should i use for image processing?**

I think **Visual Geometry Group** :
- VGG-16 & VGG-19
- very high precision





No no **Googlenet** :
- 22 layer
- Concept of Inception





## Performance

**Googlenet** achieves a top-5 test accuracy of nearly 93.3% in ImageNet with 7 million parammters only

**VGG** achieves a top-5 test accuracy of nearly 92.7% in ImageNet with 138 million parammters

## Applications

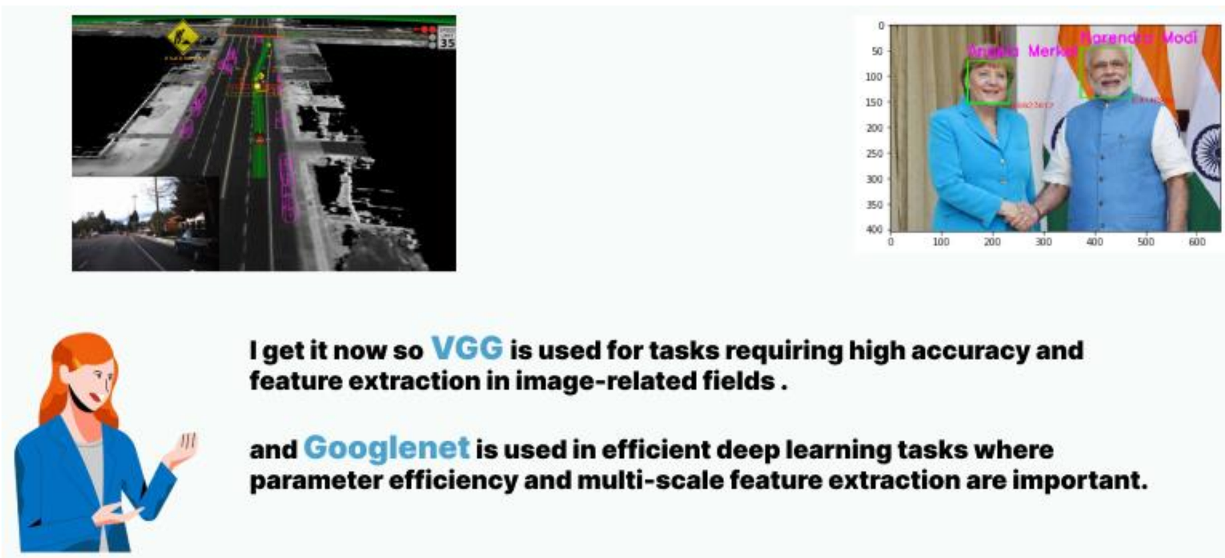
- Analyse vidéo
- Autonomous driving
- Satellite image analysis
- Image classification
- Object detection
- Medical imaging

- Image classification
- Feature extraction
- Facial recognition
- Medical image analysis
- Object detection and recognition

Visual Computing Magazine

# CHRONICLE OF COMPUTER VISION
## EVENT: GoogleNet, 2014, VGGNet, 2015.

**S. BELKACEMI, L. ROUIBAH, MASTER 2 VISUAL COMPUTING, USTHB**




**References:**

1- Karen Simonyan and Andrew Zisserman.Very Deep Convolutional Networks for Large-Scale Image Recognition,2015, https://arxiv.org/abs/1409.1556.GoogleNet:

2- Christian Szegedy,Wei Liu,Yangqing Jia,Pierre Sermanet,Scott Reed,Dragomir Anguelov,Dumitru Erhan,Vincent Vanhoucke,Andrew Rabinovich.
Going Deeper with Convolutions 2014, https://arxiv.org/abs/1409.4842

# CHRONICLE OF COMPUTER VISION
## EVENT: Deep Residual Networks, 2015
**W. KESBI, M. SAADI**, MASTER 2 VISUAL COMPUTING, USTHB

Hi NeoBot, i have a problem, my model isn't performing well, what should i do ?

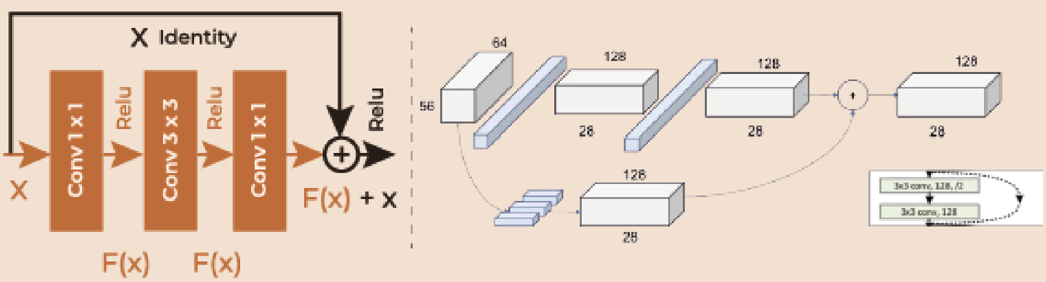
Have you tried adding more layers ?

Its even worst now...

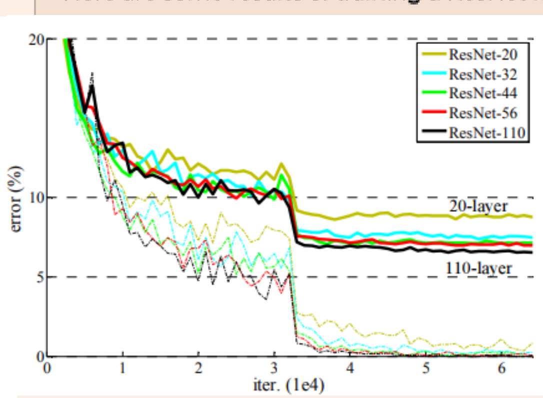
I think you should try ResNet

Huh, what's that ?

This architecture allows ResNet to perform complex tasks such as image recognition, object detection and segmentation.




Here are some results of training a ResNet model:




| | Model | Train acc | Validation acc | Time |
|---|---|---|---|---|
| 1 | ResNet18 | 0,83 | 0,87 | 2701s |
| 2 | ResNet34 | 0,8651 | 0,8519 | 4800s |
| 3 | ResNet50 | 0,8662 | 0,8095 | 5580s |
| 4 | ResNet101 | 0,8594 | 0,7884 | 6112s |
| 5 | ResNet152 | 0,8798 | 0,8836 | 9248s |

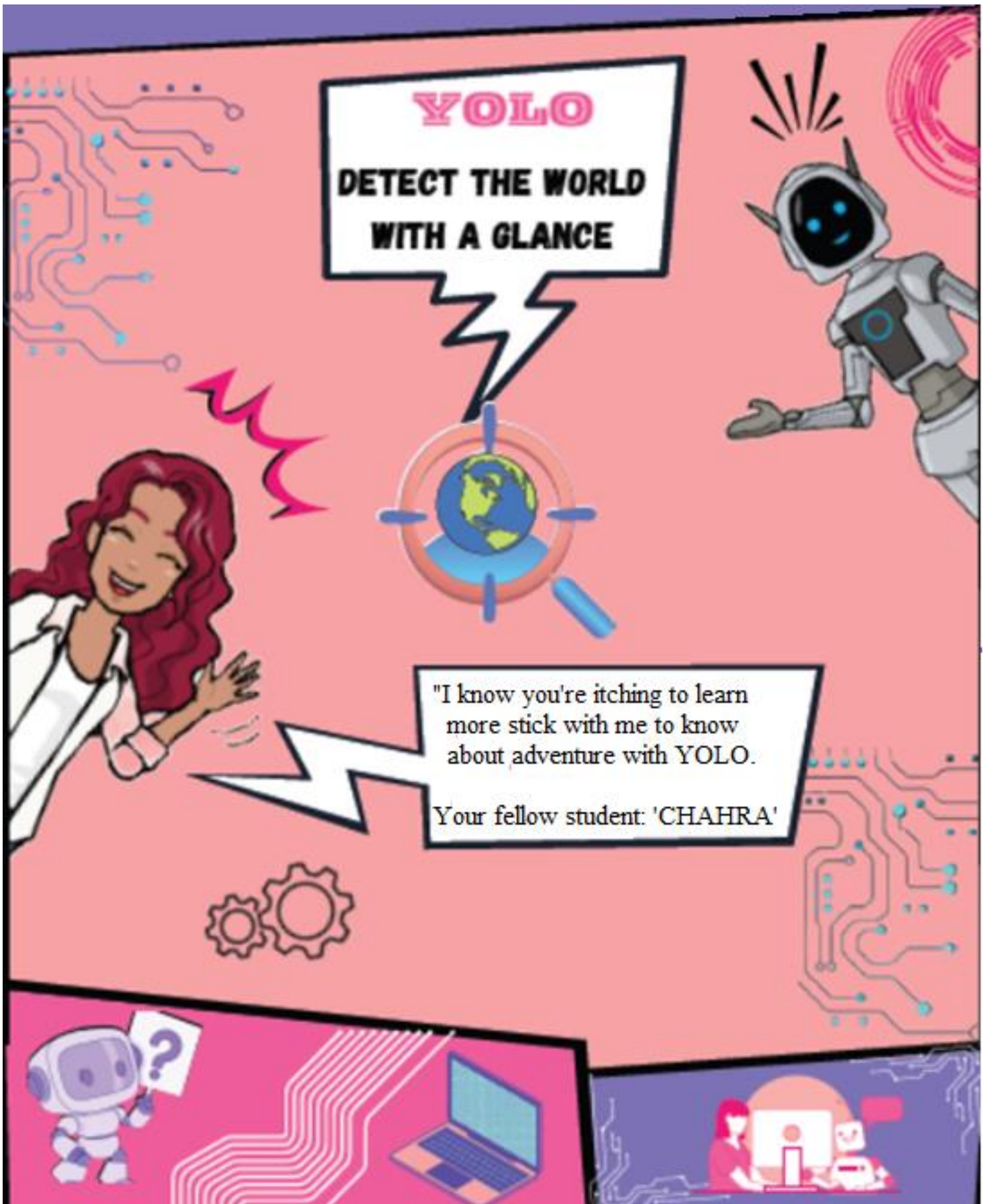
Intresting!! i will try that

Reference: Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Deep Residual Learning for Image Recognition, Microsoft Research, 10 Dec 2015.

Visual Computing Magazine

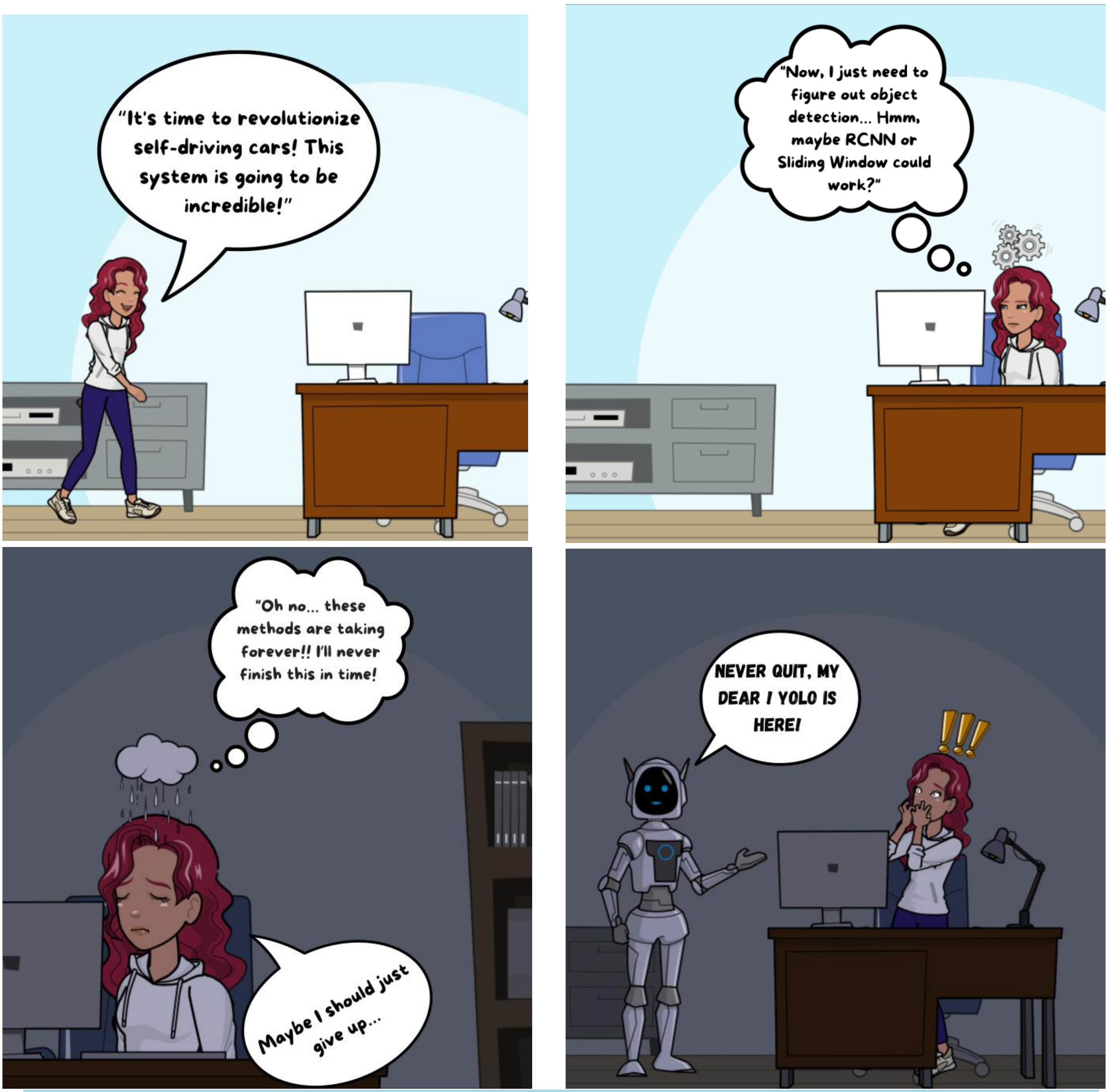
# CHRONICLE OF COMPUTER VISION
## EVENT: YOLO, 2016

**Y. SAADOUNE, W. Abboud**, MASTER 2 VISUAL COMPUTING, USTHB



## CHRONICLE OF COMPUTER VISION
## EVENT: YOLO, 2016

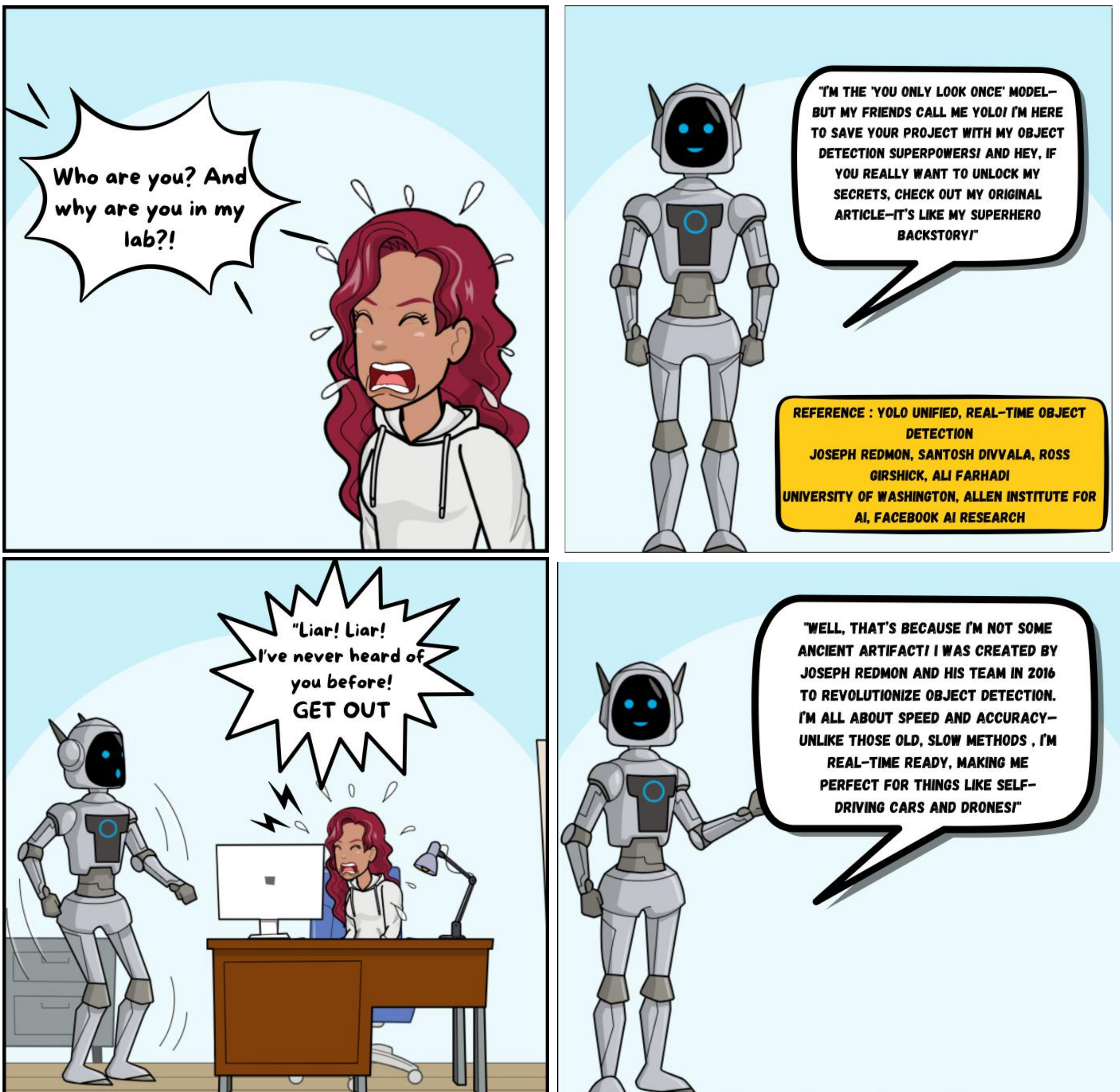
**Y. SAADOUNE, W. Abboud**, MASTER 2 VISUAL COMPUTING, USTHB



Visual Computing Magazine

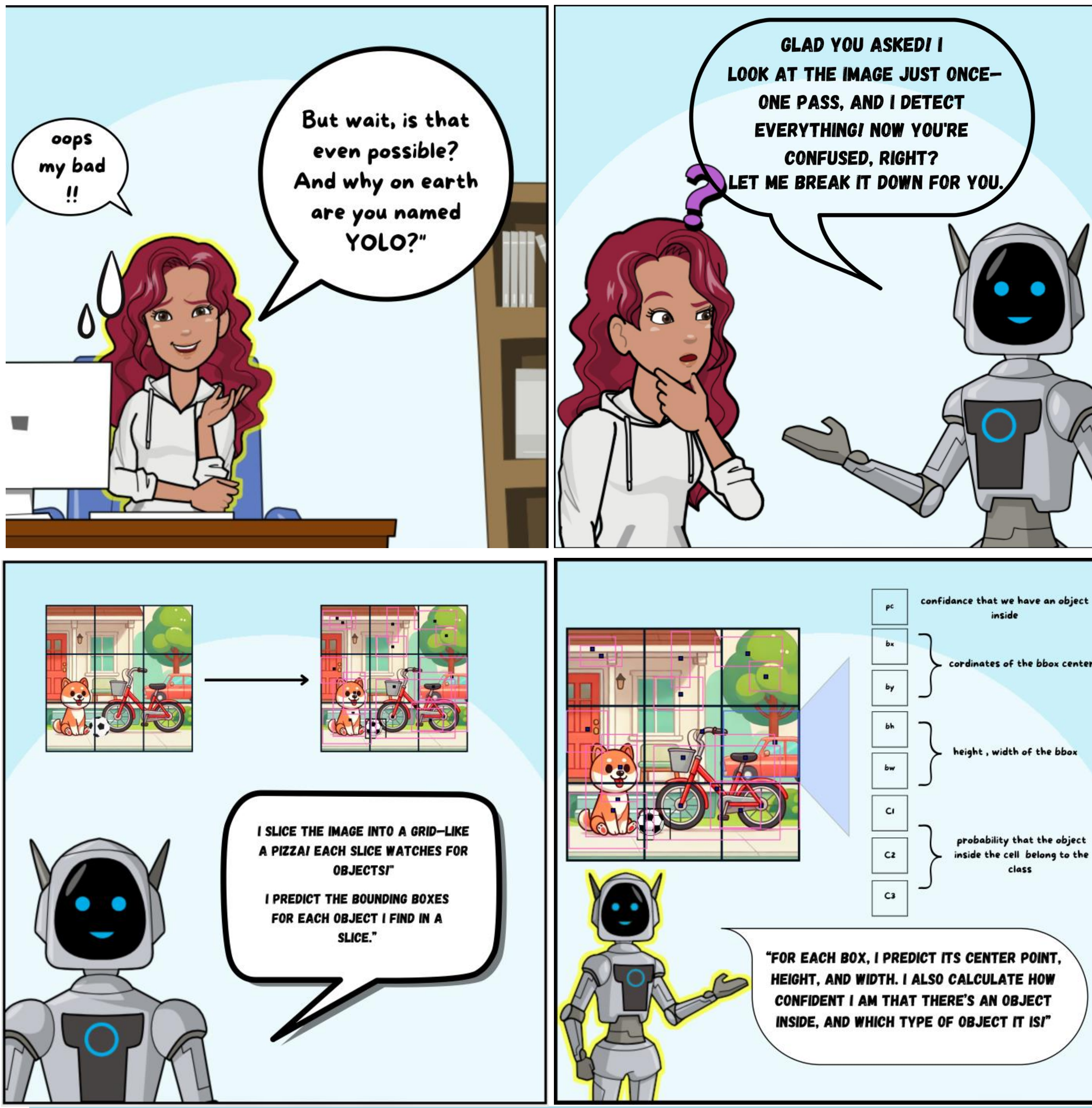
# CHRONICLE OF COMPUTER VISION
## EVENT: YOLO, 2016

**Y. SAADOUNE, W. Abboud**, MASTER 2 VISUAL COMPUTING, USTHB



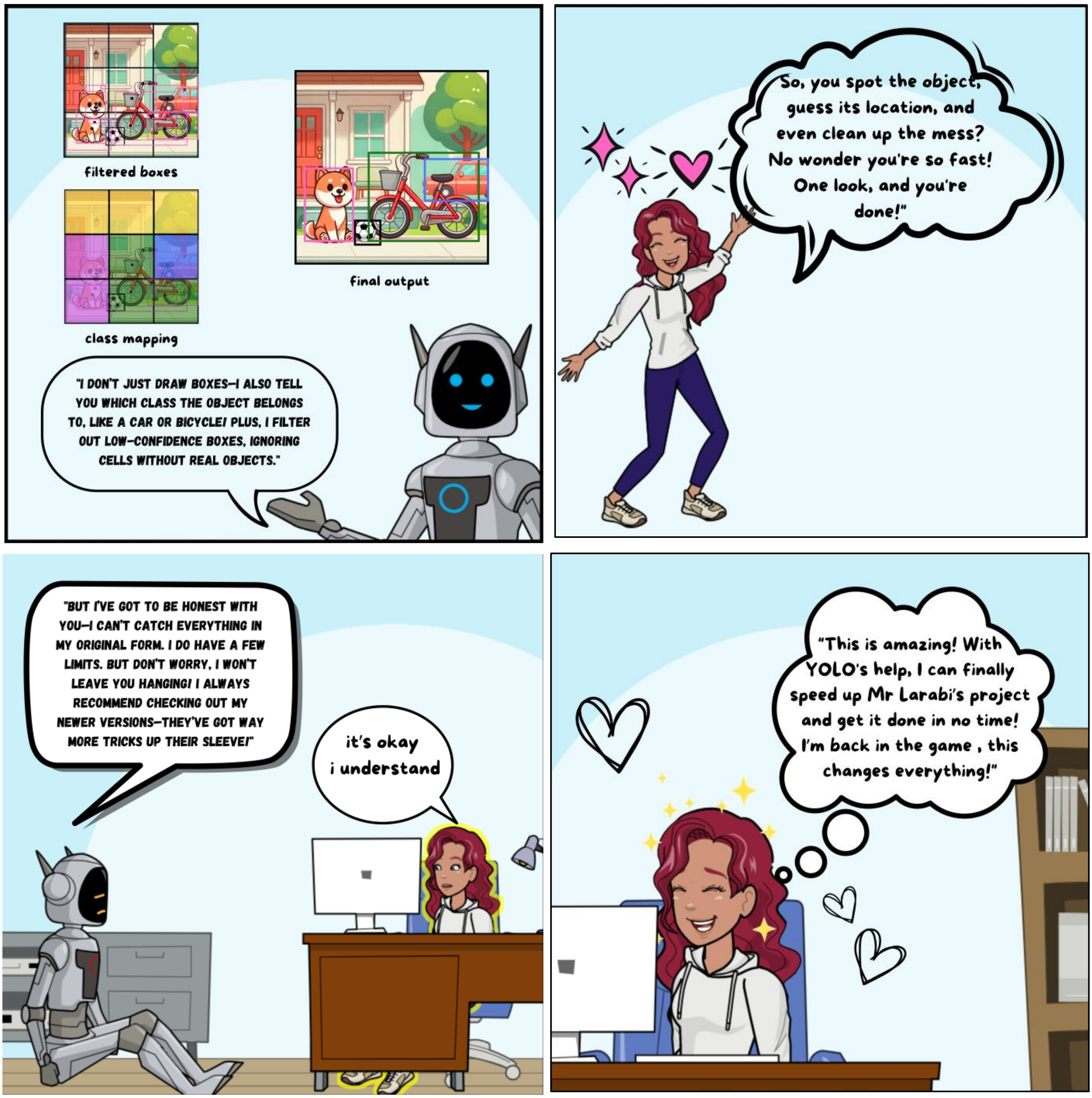
# CHRONICLE OF COMPUTER VISION
## EVENT: YOLO, 2016

Y. SAADOUNE, W. Abboud, MASTER 2 VISUAL COMPUTING, USTHB



# CHRONICLE OF COMPUTER VISION
# EVENT: YOLO, 2016

Y. SAADOUNE, W. Abboud, MASTER 2 VISUAL COMPUTING, USTHB



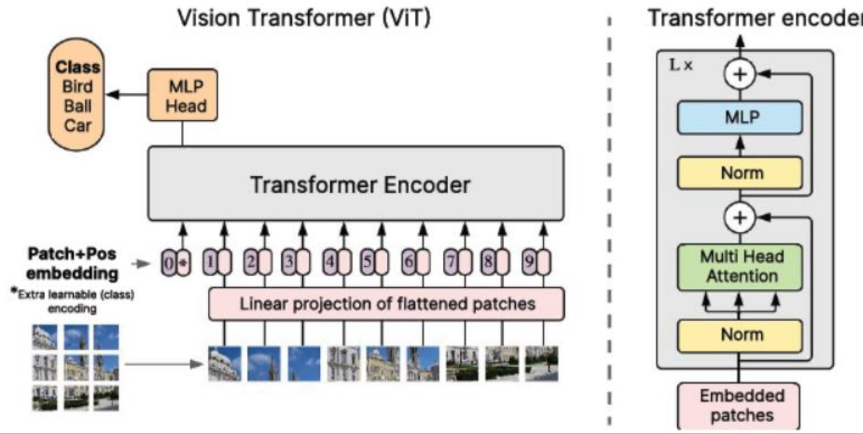
Visual Computing Magazine

# CHRONICLE OF COMPUTER VISION

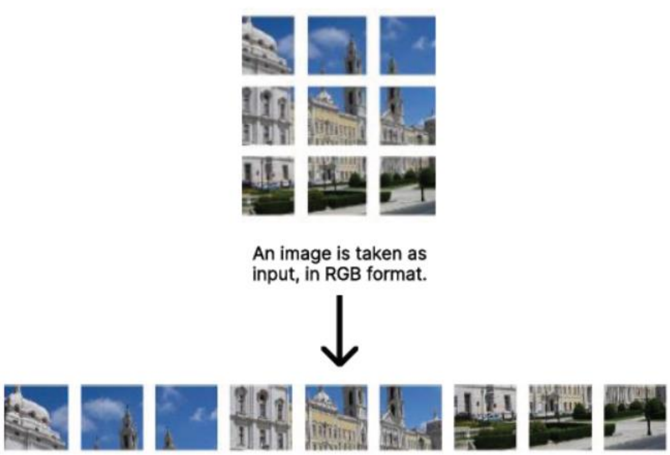
## EVENT: Vision Transformers, 2021

B. ABADLI, R. KEMMOUN, MASTER 2 VISUAL COMPUTING, USTHB

The vision transformer (VIT)'s architecture

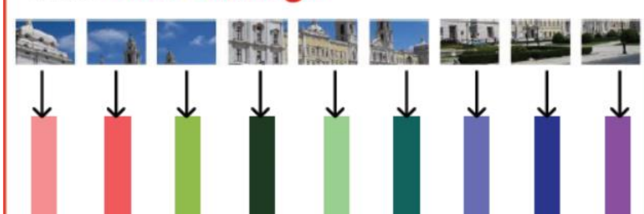



### First we need to prepare the **patches!**



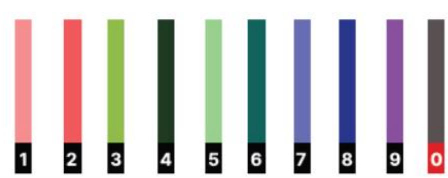

The image is divided into smaller, non-overlapping patches

### Then we reshape them into **embeddings**

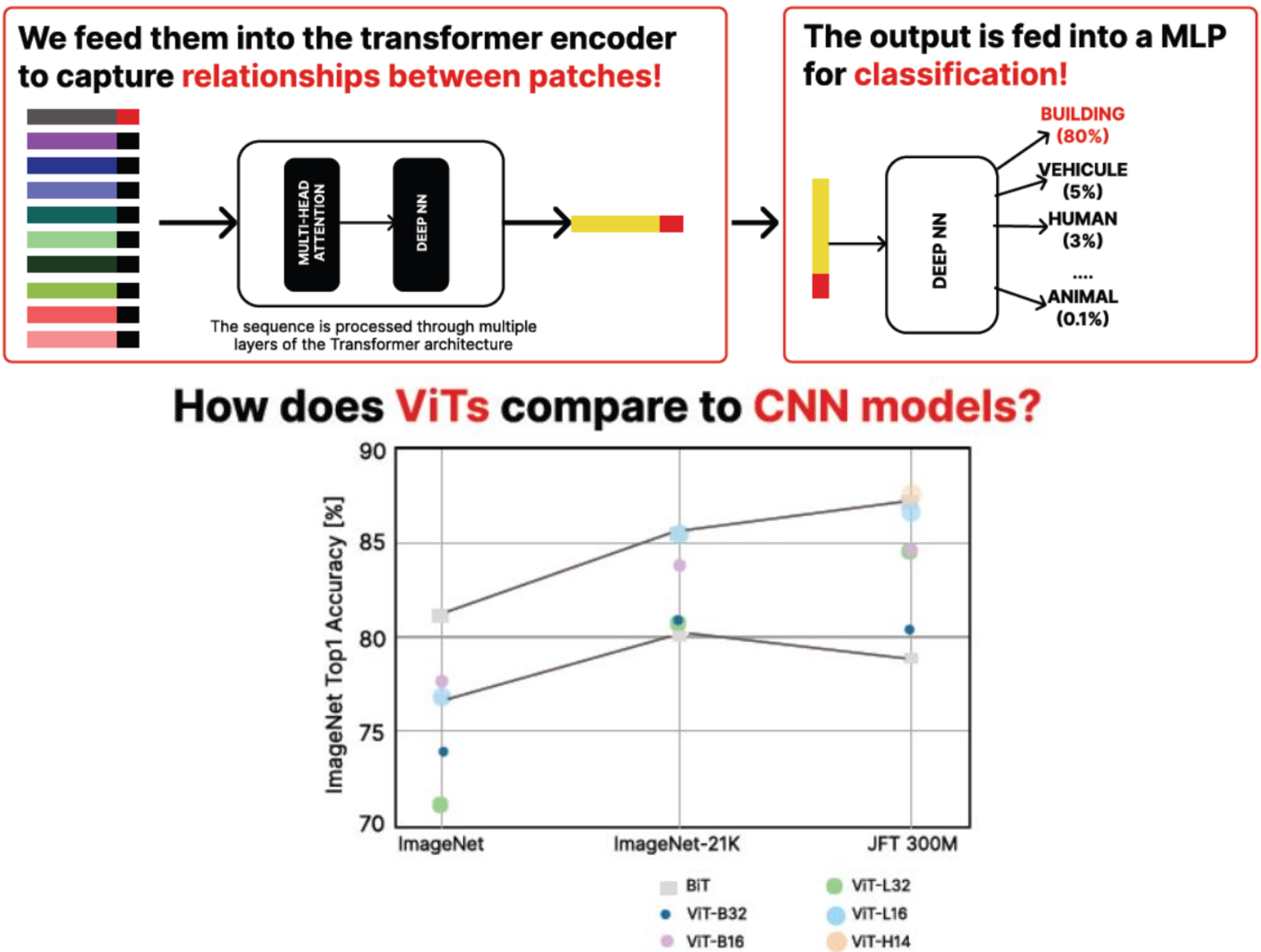


We convert the patches from RGB to numerical vectors

### We then add the **positions** and the **class token!**



# CHRONICLE OF COMPUTER VISION
## EVENT: Vision Transformers, 2021
**B. ABADLI, R. KEMMOUN**, MASTER 2 VISUAL COMPUTING, USTHB




Reference: Alexey Dosovitskiy , Lucas Beyer , Alexander Kolesnikov , Dirk Weissenborn , Xiaohua Zhai , Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, Neil Houlsby, AN IMAGE IS WORTH 16X16 WORDS:TRANSFORMERS FOR IMAGE RECOGNITION AT SCALE, ICLR 2021

# CHRONICLE OF COMPUTER VISION
# EVENT: DALL-E 3, 2023

R. BOUSSIS, W. WASSIM, MASTER 2 VISUAL COMPUTING, USTHB

Visual Computing Magazine

**CHRONICLE OF COMPUTER VISION**
**EVENT: DALL-E 3, 2023**

R. BOUSSIS, W. WASSIM, MASTER 2 VISUAL COMPUTING, USTHB

2

## CHRONICLE OF COMPUTER VISION
## EVENT: DALL-E 3, 2023

R. BOUSSIS, W. WASSIM, MASTER 2 VISUAL COMPUTING, USTHB

Visual Computing Magazine

## CHRONICLE OF COMPUTER VISION
## EVENT: DALL-E 3, 2023

R. BOUSSIS, W. WASSIM, MASTER 2 VISUAL COMPUTING, USTHB

## CHRONICLE OF COMPUTER VISION
## EVENT: DALL-E 3, 2023

R. BOUSSIS, W. WASSIM, MASTER 2 VISUAL COMPUTING, USTHB

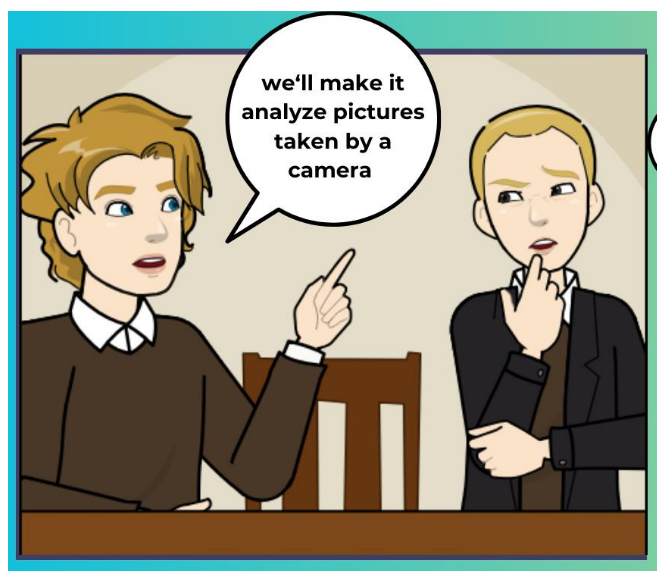

Reference: https://openai.com/index/dall-e-3/

# Visual Computing Magazine

## *Content*

Vol 2, Issue 2, 2024




Visual Computing Magazine