The new wave
Deep Learning algorithms
and
modern VCA in IP VSS

Vlado Damjanovski B.E.(electronics&TV)
ViDi Labs Pty Ltd
www.vidilabs.com
vlado@vidilabs.com

ViDi Labs
w w w . v i d i l a b s . c o m
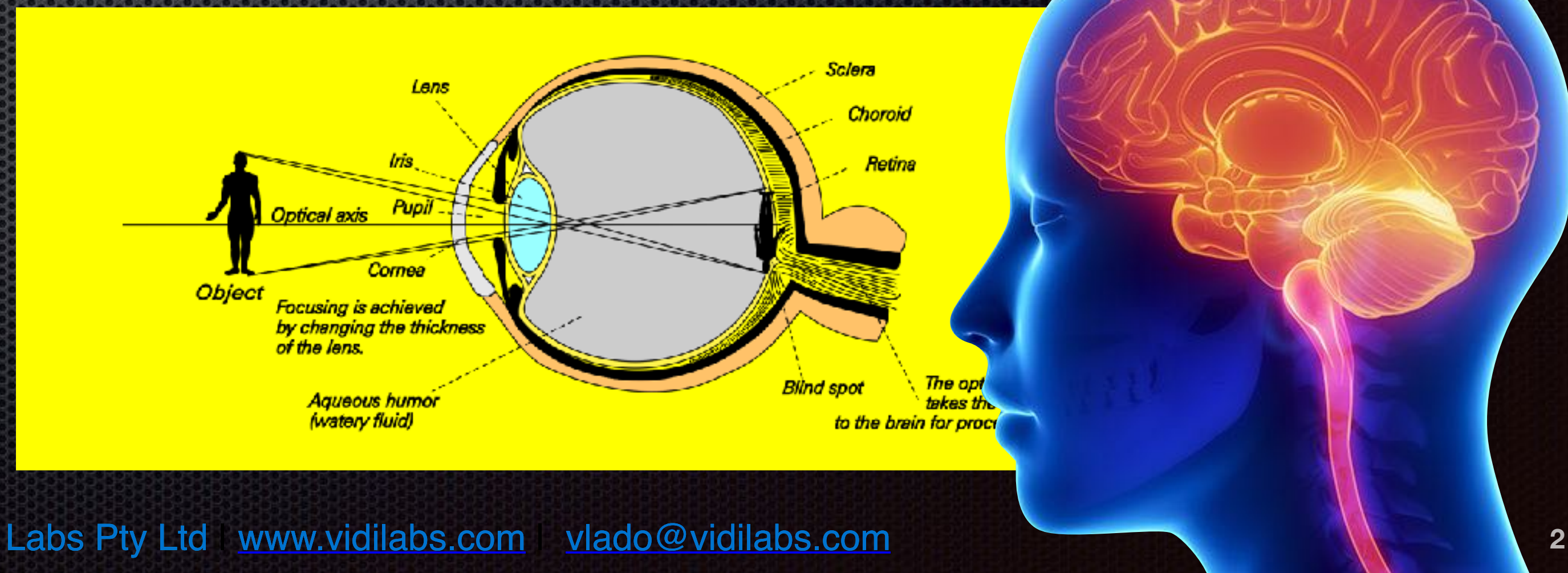
How we learn about the world around us

Sensory chart
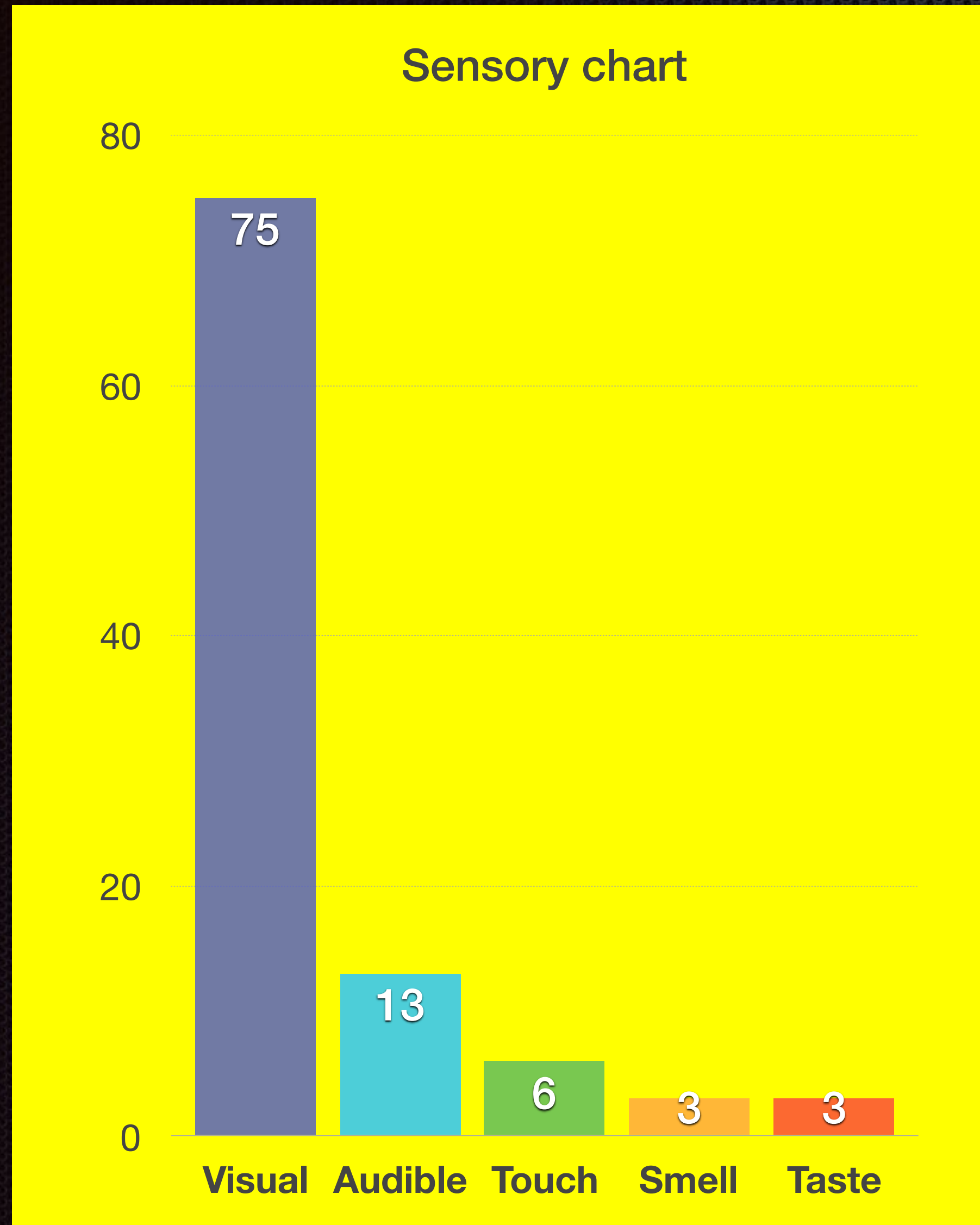
The visual perception is the most informative!
- 75 % visual sensors (eyes)
- 13% audible (ears)
- 6% touch (fingers, skin)
- 3% smell (nose)
- 3% taste (tongue)

"A picture is worth a thousand words"

CCTV as a protector

Ideology, religion and greed for money and profit practiced by ill-developed entities give birth to illegal activities, crime, extremism and terrorism.

While in the past, CCTV in public places was seen as **intrusion** of people's privacy, today CCTV is seeing rather as a **protector** of people's freedom and safety.

But, CCTV is no longer used just for safety…

CCTV problem:
finding the needle in the haystack...

Today, there is no problem recording very high quality digital video, HD or 4k, of **hundreds** of cameras, for **weeks, months or even years**.

The real major problem today is - **how do you find a footage of an event or incident that happened not long ago**, unless the date/time and camera covering the incident is known?
Playing back weeks or months of hundreds of cameras recorded footage or monitoring tens of live video screens is not practical nor efficient.

There is simply too much video data.

We need some intelligent help!

First, it was the VMD…

Many years ago, one of the most useful products in CCTV was the Video Motion Detector (VMD).

Initially, this was based on luminance changes.

Common problems: camera noise, flickering screens, trees, pets, rain,…

Then, a more advanced (outdoor) VMD came

A more advanced (outdoor) VMD device came which could discriminate motion of a person in an outdoor environment, ignoring rain, snow or wind and even taking perspective into account.

When recording got digitised, VMD was used to extend the recording capacity.

- % of Motion in a system…
- pre-VMD and post-VMD concepts.
- Objects classified as bit-maps…
- 3D-noise reduction got introduced,

(Worked only with fixed cameras!)

…then, FI and LPR followed…

Some years later, when CCTV recording became digital, more complex algorithms were developed, such as Face Identification (FI) and Licence Plates Recognition (LPR).

Initially, these algorithms **were exciting, but not really as successful** as originally intended.

The initial "intelligent" analytic systems **disappointed** and many took the stand that CCTV analytics was only **a waste of time and money…**

This was the time of computers becoming more powerful…

Patch Model
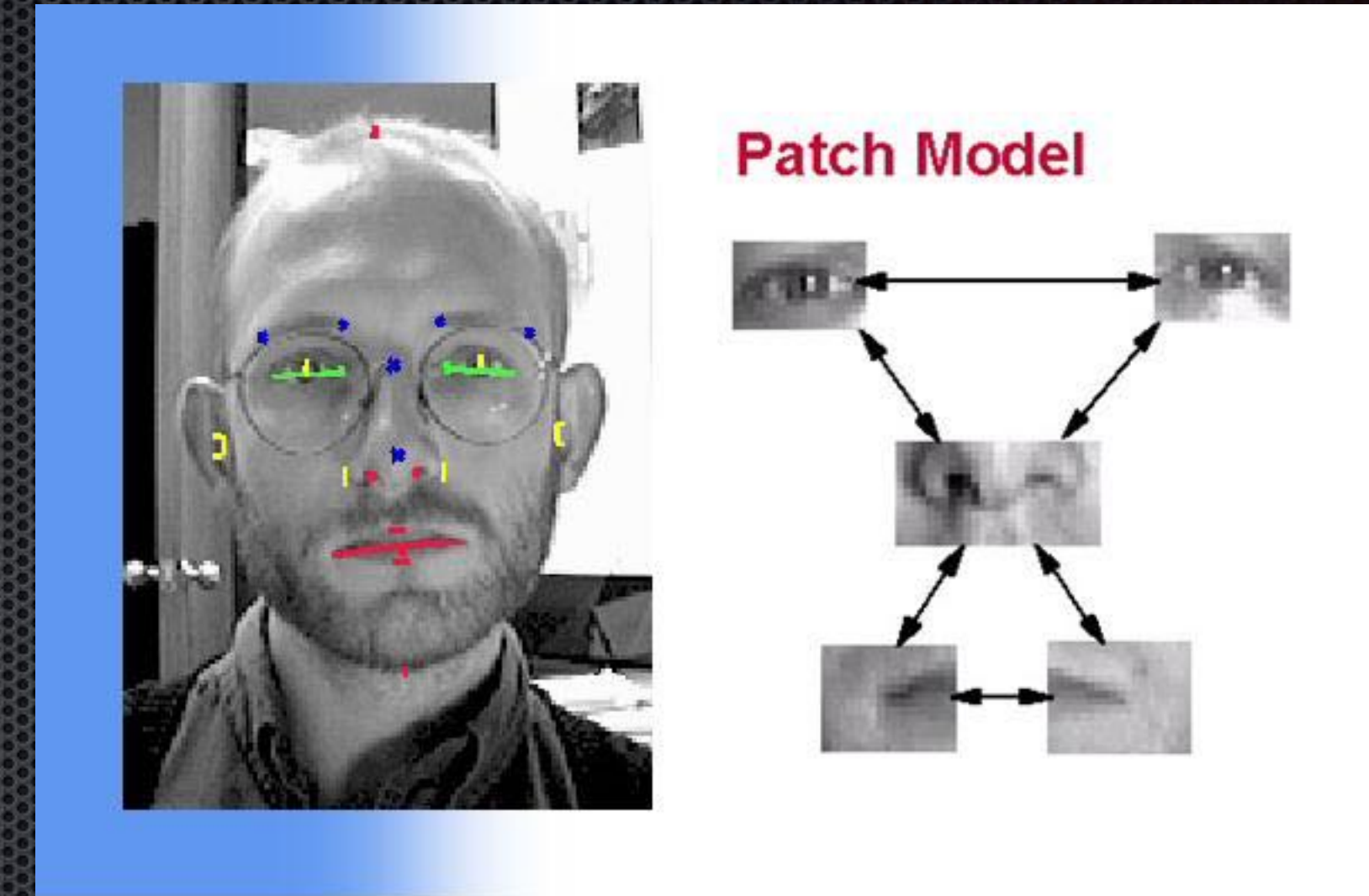
High expectations by the users ?

Setting up any analytic system in the past, whether it be outdoor VMD, FI or LPR, **was extremely difficult and required setting up of many variables**.

It included camera quality, light, lens, position, focus, noise, distance to object, computer algorithm, computer speed, etc…

Each different project required different settings and adjustments, being "walk tested" each and every time with different outcomes.

Slight variations in the captured images would yield false results, which frustrated many users, especially the ones expecting 100% accuracy.

For about 10 years not much has really changed…

Gary Kasparov vs IBM Deep Blue (1997)

A World Chess Master beaten by a computer!

Plenty of computer calculations, although not really intelligence.
Just a well programmed and exhaustive calculus algorithm (over 9M possibilities after the 3rd and 280B possibilities after the 4th move...)
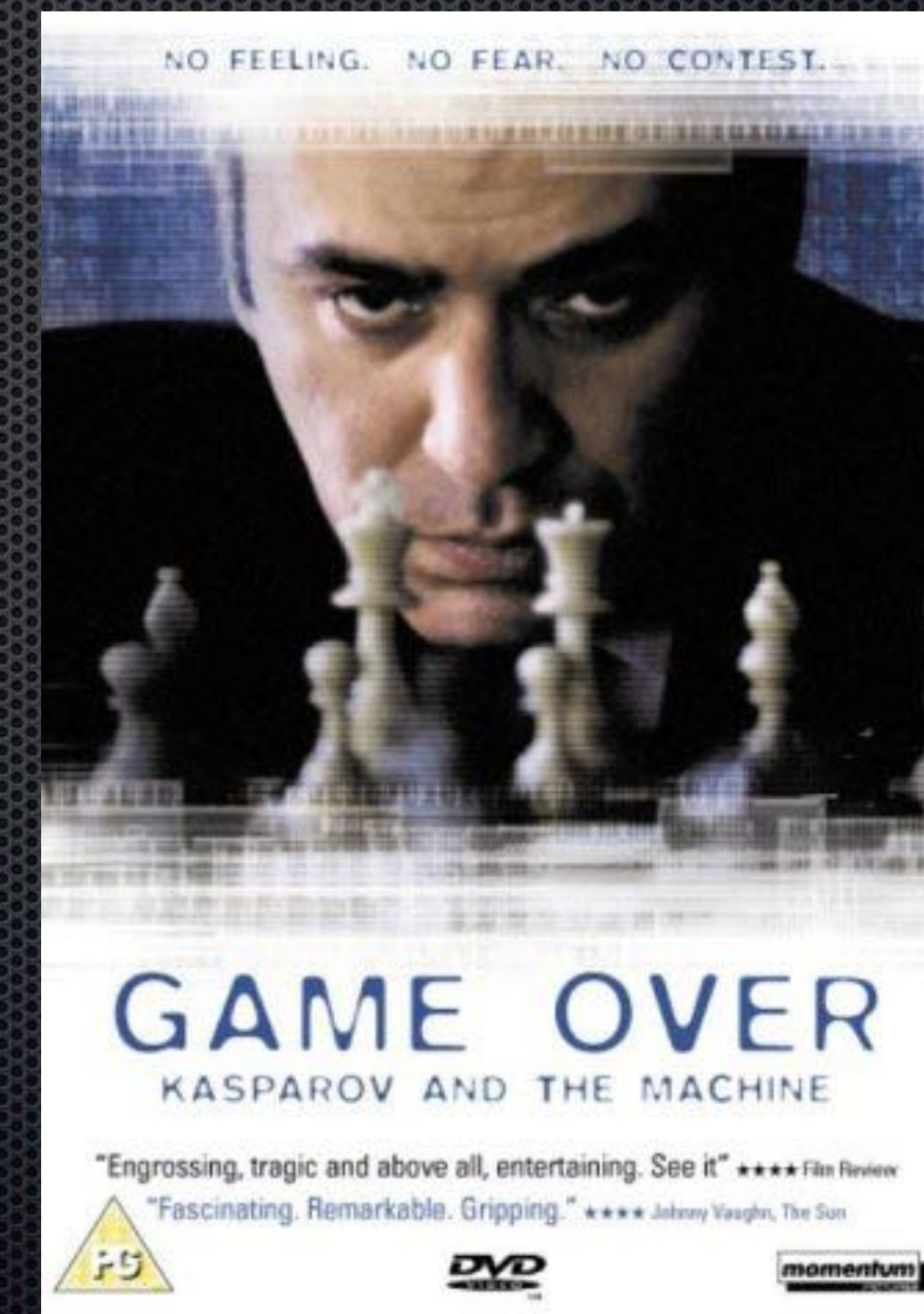
**Calculations is not the same as Intelligence!**

The proliferation of Digital imaging

The **Digital revolution** in consumer electronics (2000~2010) created a new and powerful momentum for CCTV too.

Digital music (MP3), Digital cameras, Smart phones, HD TV…

New market opportunities for all.

- Better resolution video
- Better picture quality
- Faster computers
- Smarter algorithms
- Smarter devices…
- and … VCA in CCTV

The Tidal Wave of Digital Photography

ViDi Labs
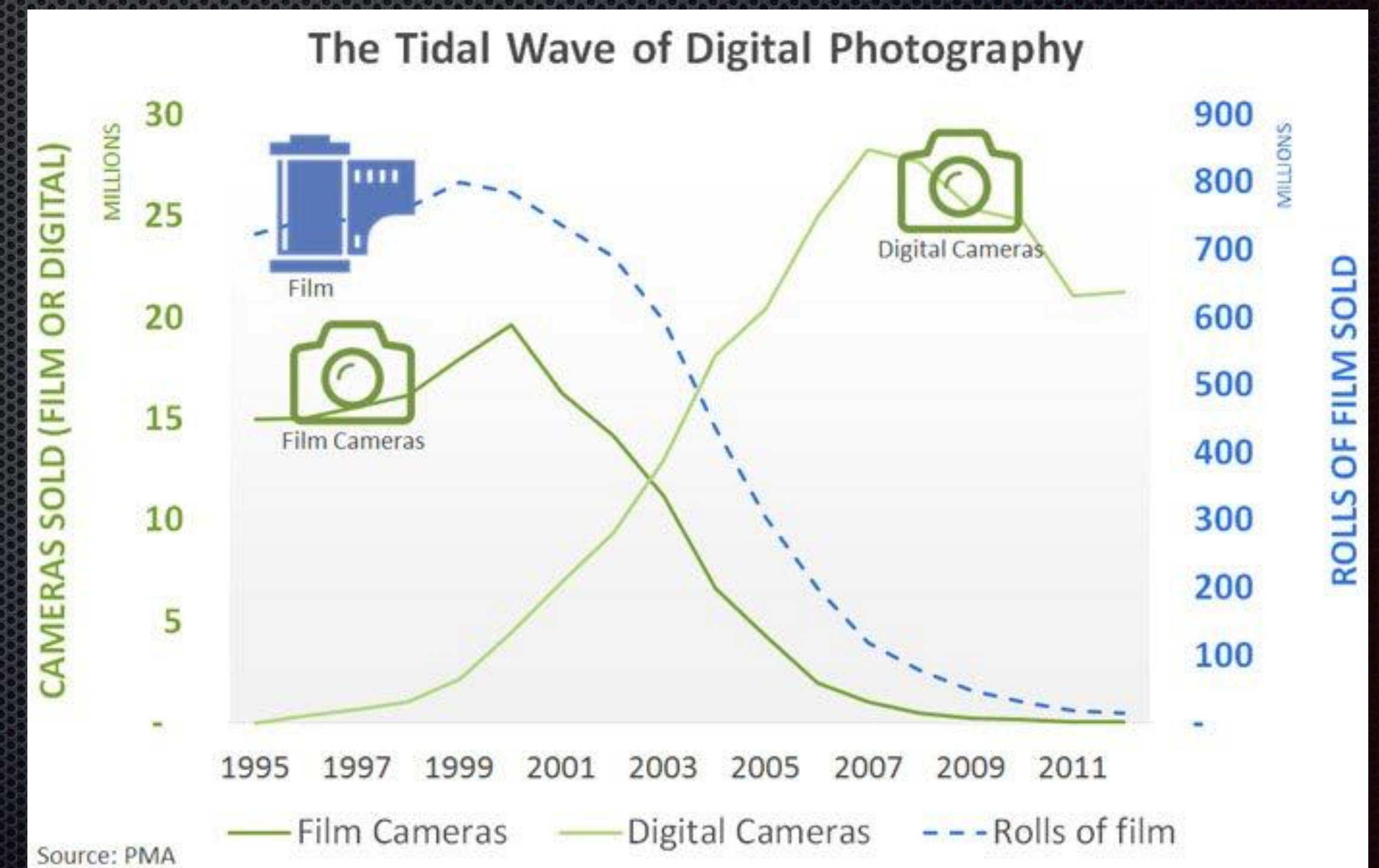
A programming breakthrough

A computer can only be as good as its programs.

Not long ago, a new approach in programming algorithms (combined with faster processing) based on brains' **neural networks** learning principles have yield **significant progress**.

These developments are in the early adoptation stage, but CCTV has already started seeing some very impressive advancements.

Users started **believing** the video analytics.

2011/2012 A Revolution in CNN and DNN

A developers competition *ImageNet Large Scale Visual Recognition Challenge* (ILSVRC) produced some **exceptional** results.

This is the time of maturation of HD digital video and fast intelligent video content processing and speech recognition.

Error rates dropped down from 25% to 16% (Today they are <5%!)

Adding GPU processing sped up the process of neural network programmed loops from days to minutes.