Introduction to

1. Computer Vision: From Theory to Application

2. OpenCV: Introduction to Image Processing, Embedded Robot Vision and Human-Computer Interaction

Jenn-Jier James Lien (連 震 杰)

Professor

Computer Science and Information Engineering

National Cheng Kung University

(O) (06) 2757575 ext. 62540

jjlien@csie.ncku.edu.tw

http://robotics.csie.ncku.edu.tw

I. Course Syllabus

Policy

- **No Popcorn, Please!!**
- **Cell phone off, please!!**

Syllabus

Basic Courses and Books

- **Undergraduate:**
  - **Linear Algebra**
  - **Introduction to Probability and Stochastic**
  - **Signals and Systems**
    - Alan V. Oppenheim and Alan S. Willsky, *Signals and Systems*, 2nd, 1996.
- **Graduate**
  - **Probability and Stochastic**
  - **Digital Signal Processing**
    -A.V. Oppenheim, R.W. Schafer, and J.R. Buck, *Discrete-Time Signal Processing*, Prentice Hall, 2nd, 1999. ISBN: 0137549202.
  - **Digital Image Processing**
    -R.C. Gonzalez and R.E. Woods, *Digital Image Processing*, Prentice Hall, 3rd, 2007. ISBN: 013168728X.
  - **OpenCV**
    -G. Bradski and A. Kaebler, *Learning OpenCV, Computer Vision with the OpenCV Library*, O'Reilly, 2008. ISBN-10: 0596516134 or ISBN-13: 978-0596516130.
  - **Computer Vision**
    -R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer, 2010. ISBN-10: 1848829345 or ISBN-13: 978-1848829343.
    -R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2nd, 2004. ISBN: 0521540518.
  - **Machine Learning and Pattern Recognition**
    -C.M. Bishop, *Machine Learning and Pattern Recognition*, Springer, 2007. ISBN: 0387310738.
  - **Deep Learning**
    -I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*, MIT, 2016. ISBN: 0262035618

Relative Conferences and Journals

- **Conference – CV and Deep Learning:**
  - CVPR: Computer Vision and Pattern Recognition
  - ICCV: International Conference on Computer Visio
  - ECCV: European Conference on Computer Vision
  - ACCV: Asian Conference on Computer Vision
  - ICML: International Conference on Machine Learning
  - NIPS: Advances in Neural Information Processing Systems
  - arXiv
  - ??
  - FG: International Conference on Automatic Face and Gesture Recognition
  - ICME:
  - ICPR:
  - ICIP:
  - eCVW: Embedded Computer Vision Workshop
- **Conferences - Robot:**
  - IEEE International Conference on Intelligent Robotics and Systems
  - IEEE International Conference on Systems, Man, and Cybernetics
- **Journal – CV and Deep Learning:**
  - IJCV: International Journal on Computer Vision
  - IEEE tPAMI
  - Image and Vision Computing
  - IEEE tVisualization and Computer Graphics
  - SIGGRAPH:
  - Deep Learning??
  - Pattern Recognition
  - IEEE tIP
  - IEEE tCSVT
  - IEEE tMM
- **Journal – Robot:**
  - IEEE tRobotics and Automation

Computer Vision: FTP Site and Website

- **FTP site to download and upload:**
  - **IP: 140.116.154.1**
  - **Port:**
  - **ID: cv2018**
  - **Password: cv2018**

- **Download lectures: After 17:00 every Wednesday**

- **Website:**
  - **http://robotics.csie.ncku.edu.tw/course.html**

- **Office Hour: CSIE 9F Robotics Lab**
  - **Monday – 15:00~17:00**
  - **Tuesday – 20:00~22:00**

OpenCV: FTP Site and Website

- **FTP site to download and upload:**
  - **IP: 140.116.154.1**
  - **Port:**
  - **ID: opencv2018**
  - **Password: opencv2018**

- **Download lectures: After 17:00 every Thursday**

- **Website:**
  - **http://robotics.csie.ncku.edu.tw/course.html**

- **Office Hour: CSIE 9F Robotics Lab**
  - **Monday – 15:00~17:00**
  - **Tuesday – 20:00~22:00**

II. Briefly:
James Lien / 連震杰

Biography in USA (1/2)

- **1993/08~1998/04:** Ph.D. ECE, U. of Pittsburgh
  - Position: Research Assistant, Robotics Institute (RI), School of Computer Science (SCS), Carnegie Mellon U. (CMU)
  - Advisor1: Prof. Takeo Kanade, Director of RI, SCS, CMU;
    - » Member of National Academy of Engineering
    - » Fellow of the American Academy of Arts and Sciences, USA;
    - » IEEE Fellow, ACM Fellow, AAAI Fellow

    Advisor2: Prof. Ching-Chung Li, ECE, U. of Pittsburgh, IEEE Fellow

    Advisor3: Prof. Jeffrey F. Cohn, Dept. of Psychology and Psychiatry, U of Pitt
  - Dissertation: Automatic Recognition of Facial Expression Using Hidden Markov Models and Estimation of Expression Intensity

- **1998/05~1998/12:** Robotics Institute (RI), School of Computer Science (SCS), CMU
  - Position: Visiting Research Scientist

- **1999/01~2002/07:** Identix (Visionics) - Biometric/Face Recognition/Surveillance
  - Position: Senior Research Scientist/DARPA HID Project Lead
  - Award:
    - 2000 USA FERET Face Recognition Competition: 第一名 at Visionics (第二名 MIT Media Lab)
    - 2000 DARPA Surveillance Project: Human Identification at a Distance - Multiple Measurements in Space and Time for Face Identification (BAA00-29), US$ 4 Millions
    - 2002 DARPA Surveillance Project: Human Identification at a Distance - Segmentation and Fused Face from Multiple Views (BAA00-29), US$ 1 Million

Biography in Taiwan (2/2)

- 連震杰教授現任成功大學資訊工程學系教授兼任資訊系副系主任及製造資訊與系統研究所所長、自動化科技學會副秘書長。
  - 在2002年時加入了國立成功大學資訊工程系，
  - 於2004年－AOI: 他的學生團隊及東台精機與友達、奇美（現為群創）TFT-LCD面板公司及茂迪、益通太陽能面板公司合作，致力於自動光學檢測(AOI)智慧機台之開發。
  - 於2009年－ADAS and Surveillance: 與美國德州儀器公司及研華科技合作研究開發嵌入式電腦視覺系統於先進駕駛輔助系統（ADAS）、監控(Surveillance)和人機互動市場。
  - 自2014年- ：他的團隊與東台精機及上銀科技合作開發DLP 3D檢測和重建及以電腦視覺與深度學習為基礎之智慧型機器手臂抓取與工業4.0的刀具磨耗監測和壽命預測等技術。

- **Academic Award:**
  - 博班學生涂瀞珽獲得「亞洲微軟學者獎」，並在北京亞洲微軟實習一年。
  - 2007 2篇國際會議最佳論文獎（IMECS 2007 、PSIVT 2007），另1篇被推薦競爭ACCV 2007最佳論文獎，推薦率約 1%（8篇/640篇）。
  - 2007 2篇CVGIP會議最佳論文獎（CVGIP 2004、 CVGIP 2007）。

- **Academic Performance: Training students**
  1) Contribute to this society
  2) Find good job with high salary.

Media: Newspaper about Human-Computer Interaction

- **1998/08/03: *Washington Post, Washington Times*, and *Pittsburgh Post-Gazette* –**

  "Look Closely: Computer program reads deep into our true feelings by analyzing our facial expressions.“

- **2001: ABC, NBC, and CBS TV Channels and Discovery Channel in USA: “Surveillance – Security via CCTV”**

Lie detector

Is this woman truly surprised? A computer can tell whether facial expressions are genuine.

Science, Page A-6

BIG UNIT MENACES BUCS/SPORTS, C-1

Pittsburgh Post-Gazette

ONE OF AMERICA'S GREAT NEWSPAPERS

MONDAY, AUGUST 3, 1998 VOL. 72, NO. 3 8/3/98

LOOK CLOSELY

Computer program reads deep into our true feelings by analyzing our facial expressions

By Sharon Voas
Post-Gazette Staff Writer

If Detective Sipowicz of "NYPD Blue" had this computer program, he wouldn't have to threaten so many suspects with bad things involving big men in prison.

He'd *know* if they were telling the truth most of the time.

A computer program just developed here can tell the difference between a spontaneous, or genuine expression, and one that is deliberate, but not necessarily false. In another year of study it's expected the program will be able to detect a true statement from a lie.

When computer analyses of suspects' expressions are used along with lie detector tests, which are accurate about 90 percent of the time, a detective would know most of the time whether a suspect is telling it straight.

This new computer program, developed by researchers at the University of Pittsburgh and Carnegie Mellon University, can discern thousands of facial expressions. The prototype is so precise that it deciphers the order and speed different facial muscles move in to create a genuine expression. That would be virtually impossible for someone to fake.

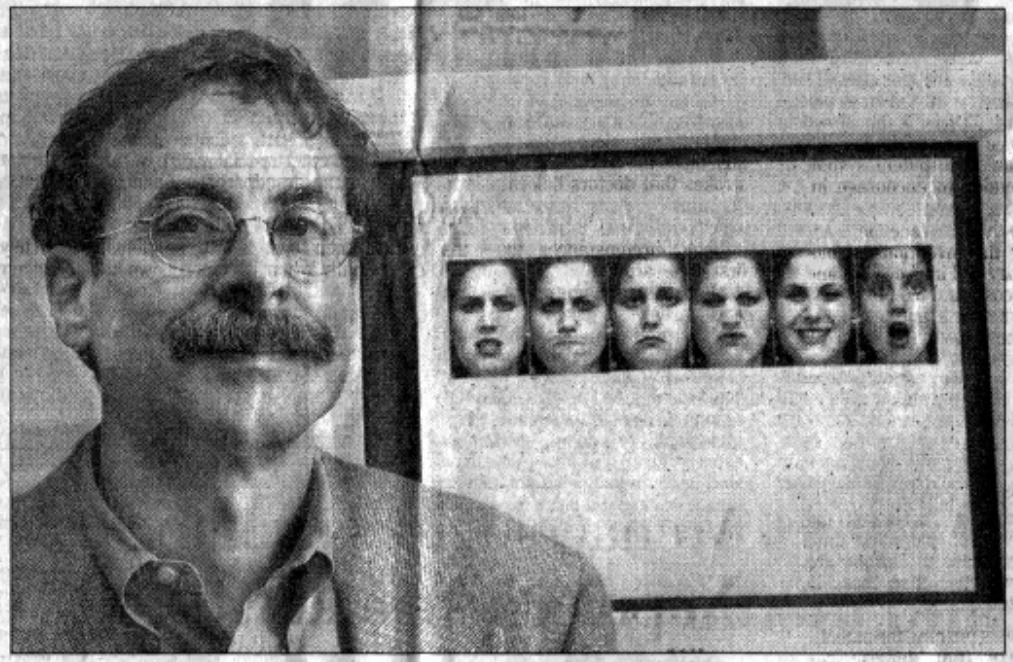
The new program is so far superior to existing methods of analyzing facial expressions that Pitt researcher Jeff Cohn says it's like inventing a new microscope.

"You can see so much more, and you can see things you couldn't see before," Cohn said.

The program has potential in mental health research as well, by revealing true feelings and helping doctors learn if a patient is really suicidal.

The program may make a big step toward developing computers users can actually talk to, as if talking to another human, instead of using keyboards and mice. But to have a conversation like two humans would, the computer would have to read the user's expressions. Also, the program may help movie animators draw more realistic expressions.

Automated Face Analysis, the official name of the program, was developed by Cohn, an associate professor of psychology and psychiatry at Pitt, Takeo Kanade, director of The Robotics Institute at Carnegie Mellon, James Lein, who recently received his doctorate in electrical engineering from Pitt, and Adena Zlochower, a doctoral candidate in psychology at Pitt.



Tony Tye/Post-Gazette

Jeff Cohn, associate professor of psychology and psychiatry at Pitt, says of the new program: "You can see so much more, and you can see things you couldn't see before."

Only one other group of researchers from the Salk Institute and the University of California at San Francisco is using a similar program of computer vision and facial analysis. Marian Stewart Bartlett, of the Salk Institute, said these programs will be especially useful for psychological research because researchers will no longer have to rely on inferring what the patient might feel. They will know.

"Facial expression is an important mode of human expression," Cohn said. "It regulates social behavior. It's important to people's self-presentation and to how they are perceived."

Facial expressions tell so much that we sometimes have "gut feelings" that something just doesn't feel right, even though we can't explain why.

For instance, President Clinton's chin boss is giving him trouble these days.

When a person smiles a genuine smile, his whole face moves, with the muscles around the eyes contract-

SEE **FACES,** PAGE A-7

SCIENCE & ENVIRONMENT

PITTSBURGH POST-GAZETTE ■ MONDAY, AUGUST 3 , 1998

Jeff Cohn/University of Pittsburgh

Adena Zlochower, a Pitt researcher working on a computer analysis of facial expressions, shows how muscles move in some basic expressions.
ABOVE LEFT: In joy, the muscles around her eyes tighten, making her cheeks rise and pulling up the corners of her lips and opening her mouth.
ABOVE CENTER: In anger, her lips press together and pull toward the middle, making a ball-shaped area called the "chin boss" push up. Eyes widen, the lower eyelid tightens and upper eyelids rise. Eyebrows draw together sharply.
ABOVE RIGHT: In surprise, the jaw drops, stretching her mouth down, while her eyes widen and her brows rise.

SCIENCE & ENVIRONMENT

Computer deciphers genuine, deliberate expressions

FACES FROM PAGE A-6

ing and the cheeks raising the mouth. The top of Clinton's face smiles, Cohn says, but Clinton pushes his lower lip up into what could be taken for a pout or some expression of sadness. Pulling that lip up tight creates a tennis-ball shaped area on the chin called the chin boss.

"That gives people a mixed message," Cohn said. "They're not quite sure about his smile."

Or take Susan Smith, the South Carolina woman who told police an attacker had kidnapped her 1- and 3-year-old sons, when she had actually belted them into her car and rolled it into a lake where they drowned.

"I knew right off she was lying," Cohn said. "When she was interviewed, she feigned sadness. In sadness, a characteristic is that the inside corners of the brows are pulled together and raised in a triangle, and the lip corners pull down. Very few people can do that voluntarily."

Cohn, 51, leans back in his office chair near a computer screen full of emoting faces. He's slender and wears the moss green cable knit sweater, khaki shorts and hiking shoes of an intellectual who spends a lot of time hiking or biking in the woods. A slender nose leads from his large blue eyes — that look larger because of thick-lensed, wire-rimmed glasses — to his thick brown and gray thatch of a mustache.

He swivels and sets the faces on the computer screen in motion, or, one could say, in emotion.

The program tracks each and every pixel — the tiniest unit of light on a computer or TV screen — of colleague Zlochower's face all the way through an expression. Little tracker lines on her face show not only which muscles move on different parts of her face, but in what order, with what intensity and how long each part lasts.

Those things weren't known before the development of this program because it uses video.

"Imagine Susan Smith has done some homework and was able to display sadness," Cohn said. "She would still have trouble performing that action in a realistic way because she'd have to pull her eyebrows up and her lip corners down in the right order and fast enough."

Most other methods of facial expression analysis focus on prototypes of a few basic expressions such as anger, joy, disgust, fear, sadness, surprise and so forth.

"But people very rarely show these peak expressions," Cohn said. "There are probably thousands of common expressions. So Automated Face Analysis examines components of expressions — a furrowed brow, widened eyes, tightened lips. We can [recognize] extreme expressions and subtle expressions."

Automated Face Analysis breaks components of facial analysis down into AUs, or action units, the smallest unit of change in an expression. How big is an AU?

"Think of Clint Eastwood giving someone a dirty look by narrowing his eyes by tightening his lower eyelids," Cohn said. "That's an AU-7."

AUs give researchers a much larger vocabulary to describe expressions.

"If we only have the terms for basic expressions, then [researchers] have an inadequate vocabulary to describe expression."

So, instead of saying someone looks happy, which may be deceiving, researchers might say she is displaying an AU-6 + 12 or an AU-6 + 12 + 20, types of smiles.

Other methods of facial analysis also use photos of people looking straight ahead and not moving. But in reality, you may close your eyes and shake your head.

Cohn turns back to the computer and calls up a series of photos of an apple-cheeked baby that start with the glimmer of a smile and move into a full-blown beaming smile that would elicit "oooohs."

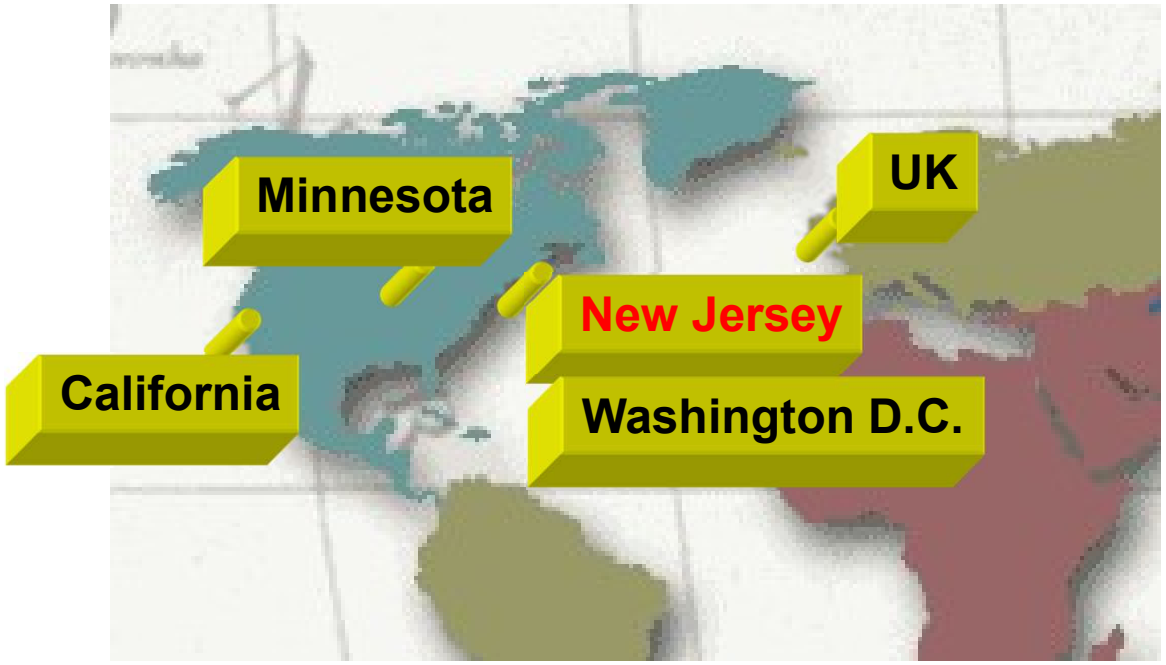
As the baby's smile progresses, he throws his head back and to the right. This computer program can account for that as part of the expression.

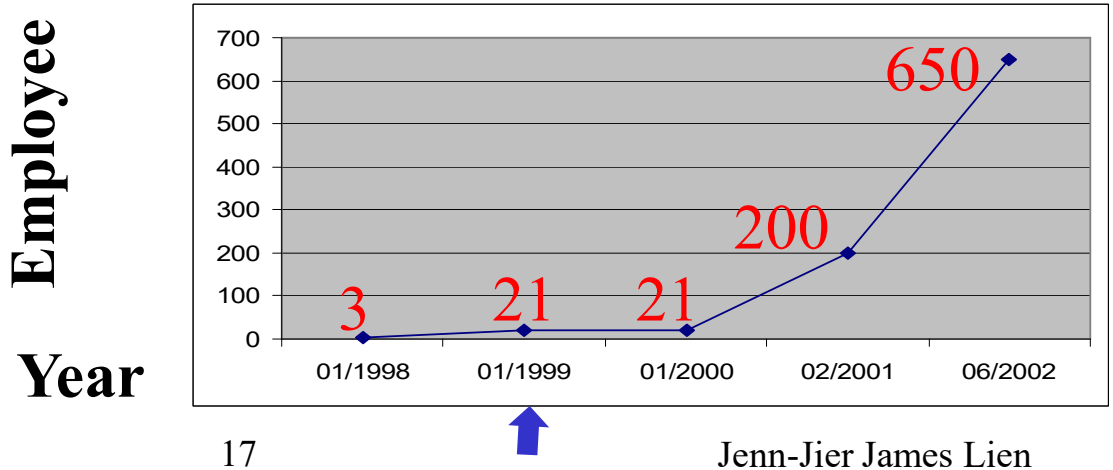
And babies' expressions are always spontaneous.

Jeffrey Cohn

These images from the Automated Face Analysis Program show how the program tracks the movement of muscles around the mouth and eyes when the subject is registering sadness, fear and disgust.

Visionics/Identix (IDNX)





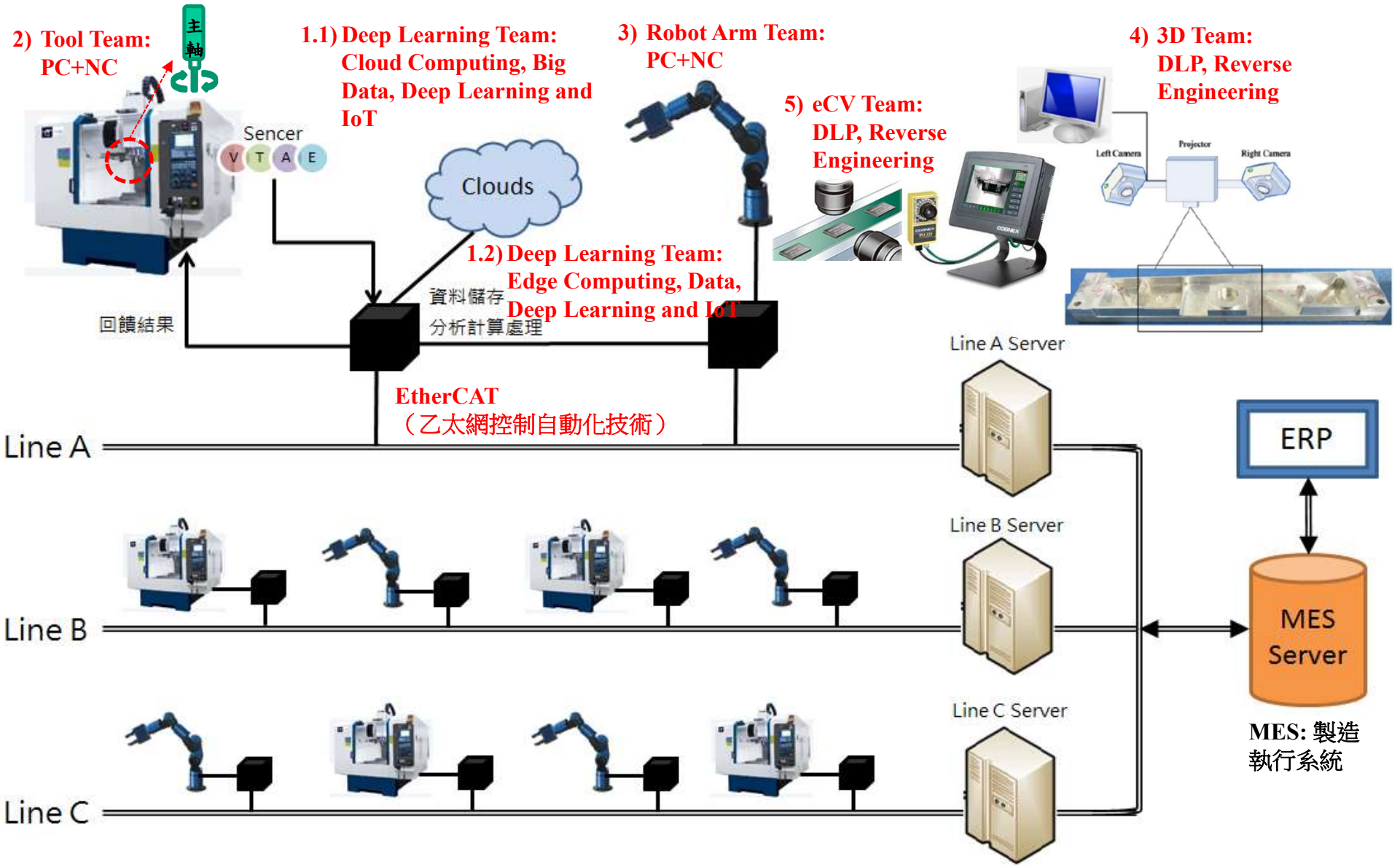
III. Computer/Robot Vision at Robotics Lab.

http://robotics.csie.ncku.edu.tw

RL Research Teams: 5 Teams

1. 深度學習組 (Deep Learning Team)
2. 機械手臂組 (Robot Arm Team)
3. AOI - 刀具組 (AOI - Tool Team)
4. 3D 組 (3D Team)
5. 嵌入式電腦視覺組 (Embedded Computer Vision Team)

Relationship between 5 Teams



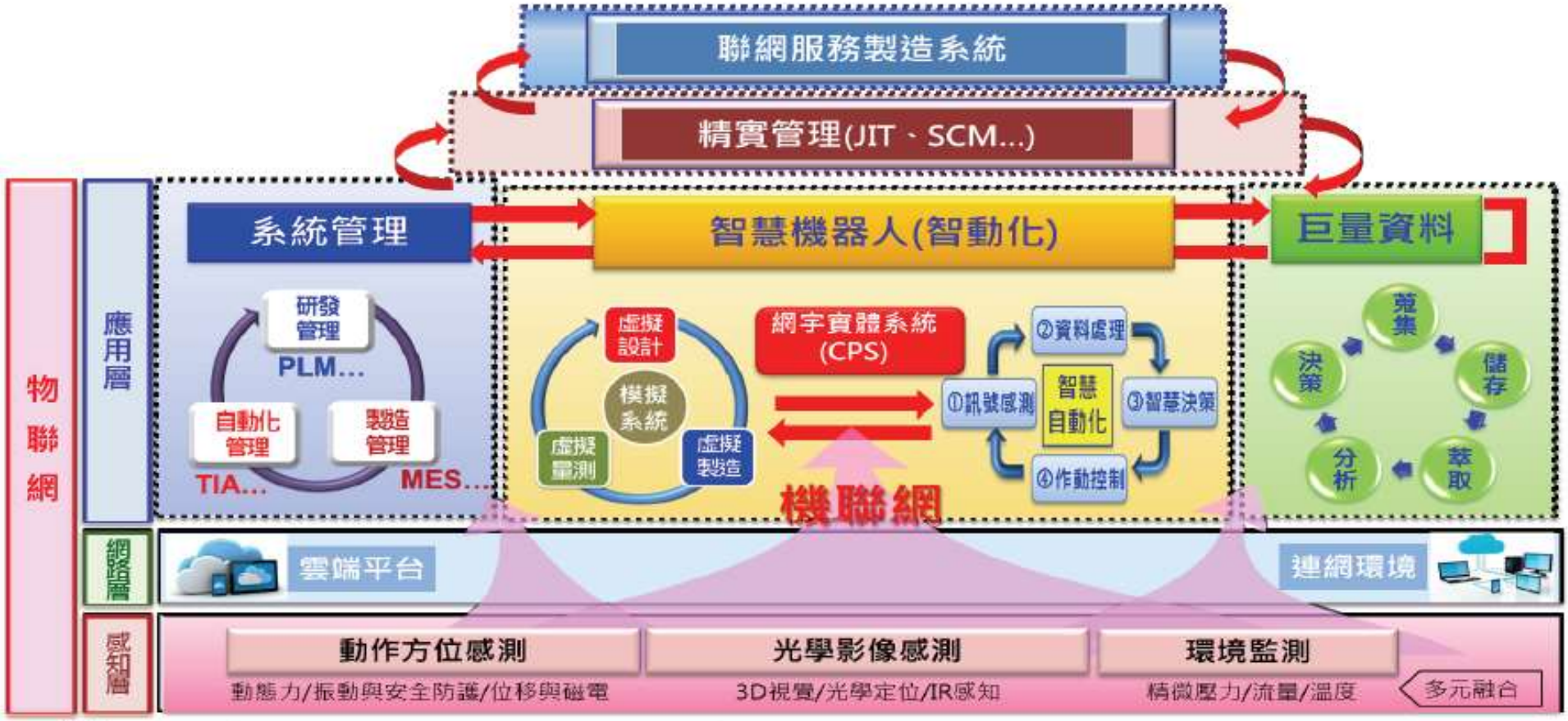
ERP（Enterprise Resource Planning）:企業資源計劃或稱企業資源規劃簡稱，應具備以下功能：基本資料與管理維護、庫存管理、採購進貨管理、配銷管理、財務管理、人資/事務管理、生產管理與決策支援管理等系統功能..

Industry 4.0: IoT (Internet of Things) and Automation

CPS: Cyber Physical System

台灣競爭力優勢分析(生產力4.0系統架構)

– 透過物聯網將生產資訊數位化，並延伸至機器端形成機聯網，再藉由系統管理、巨量資料(製造+服務)技術、以及精實管理，達成聯網服務製造系統之創新營運模式。



備註：1.JIT(Just In Time) ；SCM(Supply Chain Management)
2.PLM (Product Lifecycle Management) ；MES (Manufacturing Execution System) ；TIA (Totally Integrated Automation)

RL Research Fields

1. 人工智慧、機器學習—深度學習
   (Artificial Intelligence、Machine Learning — Deep Learning)

2. 智慧型機器手臂控制及自動化
   (Intelligent Robot Arm Control and Automation)

3. 影像處理、3D電腦視覺及圖形辨識
   (Image Processing、3D Computer Vision and Pattern Recognition)

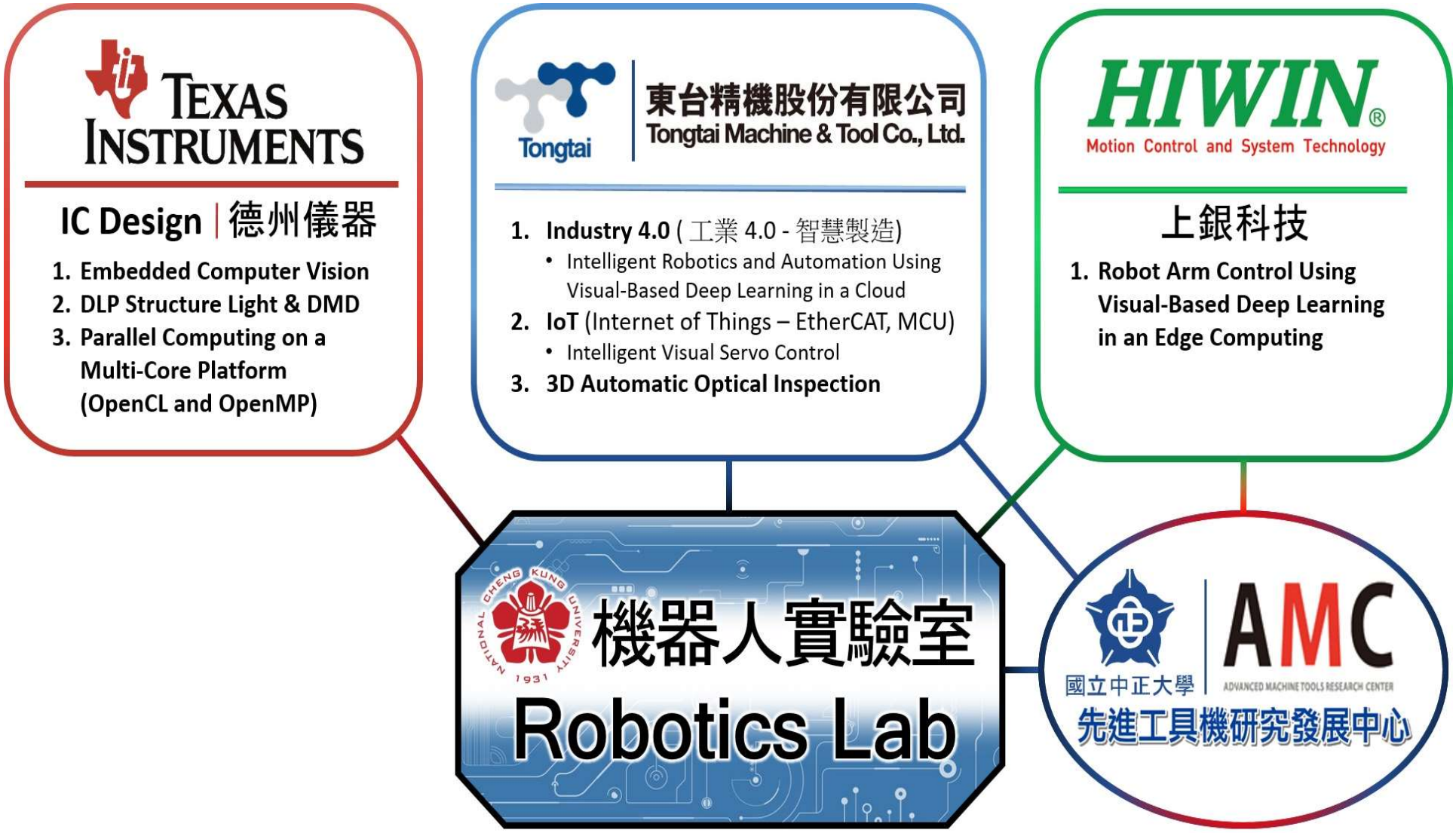
4. 3D自動光學檢測
   (3D Automatic Optical Inspection)

5. 人機互動及擴增實境
   (Human-Computer Interaction and Augmented Reality)

6. 嵌入式系統
   (Embedded System)

7. 雲端智慧型監控服務
   (Intelligent Video Surveillance as a Service (VSaaS, Cloud Computing))

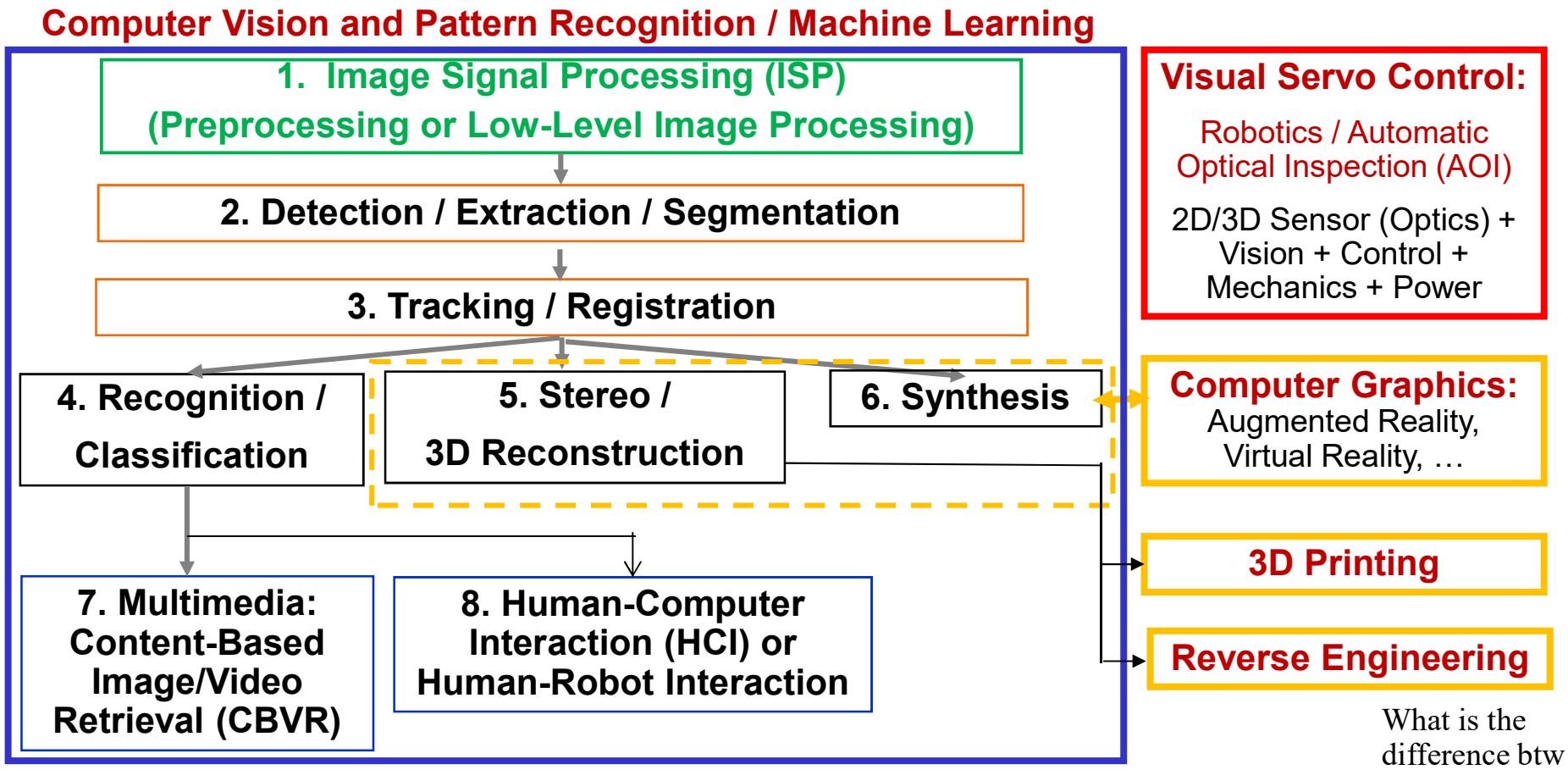
Partners


TEXAS INSTRUMENTS
IC Design | 德州儀器
1. Embedded Computer Vision
2. DLP Structure Light & DMD
3. Parallel Computing on a Multi-Core Platform (OpenCL and OpenMP)
Tongtai
東台精機股份有限公司
Tongtai Machine & Tool Co., Ltd.
1. Industry 4.0 (工業 4.0 - 智慧製造)
• Intelligent Robotics and Automation Using Visual-Based Deep Learning in a Cloud
2. IoT (Internet of Things – EtherCAT, MCU)
• Intelligent Visual Servo Control
3. 3D Automatic Optical Inspection
HIWIN
Motion Control and System Technology
上銀科技
1. Robot Arm Control Using Visual-Based Deep Learning in an Edge Computing
NATIONAL CHENG KUNG UNIVERSITY 1931
機器人實驗室
Robotics Lab
國立中正大學
AMC
ADVANCED MACHINE TOOLS RESEARCH CENTER
先進工具機研究發展中心

From Computer Vision to Visual Servo Control Techniques

CSIE, EE or ME students…

Welcome to Join RL!!



What is the difference btw CV and CG?

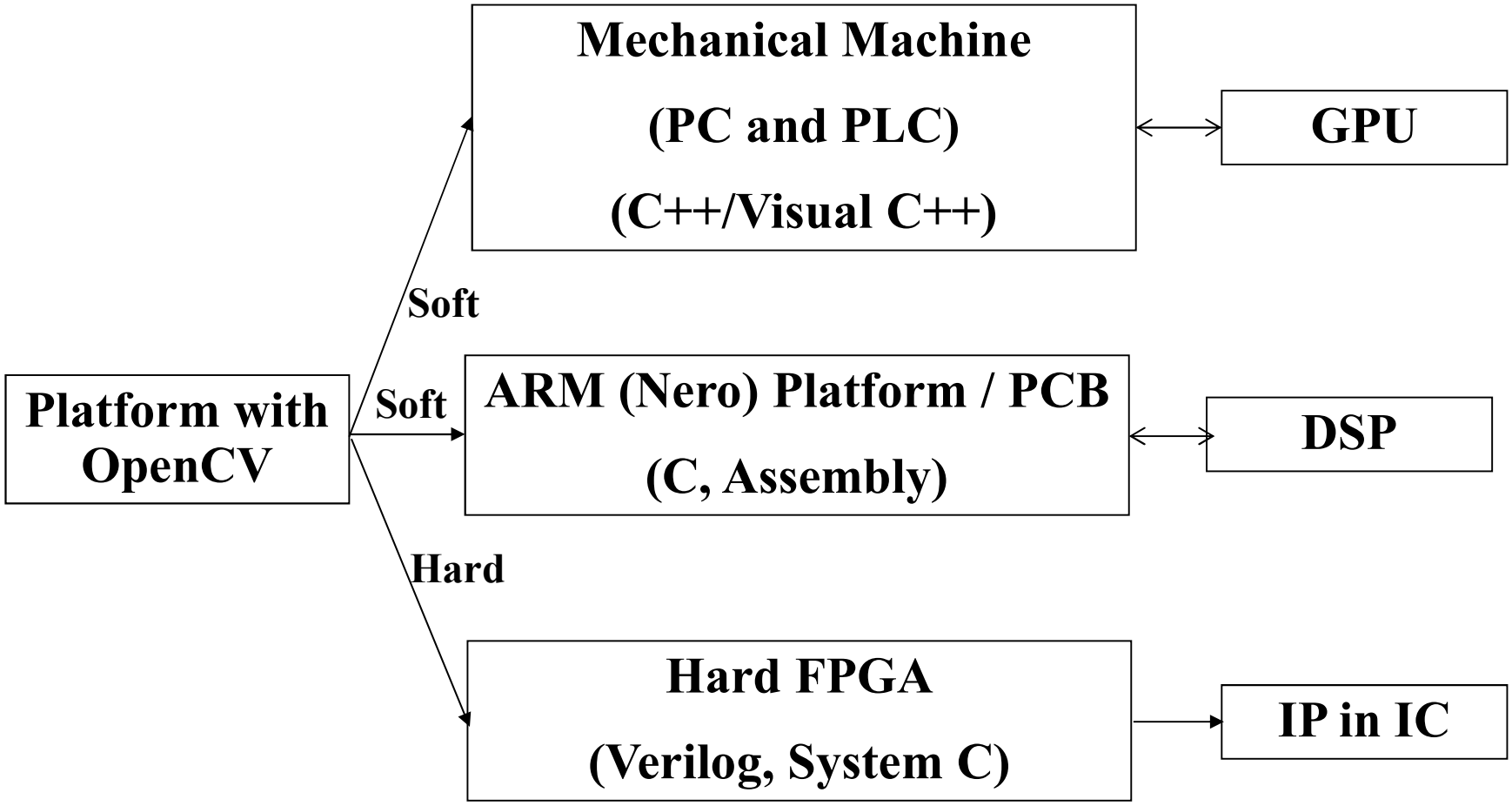
OpenCV

OpenGL

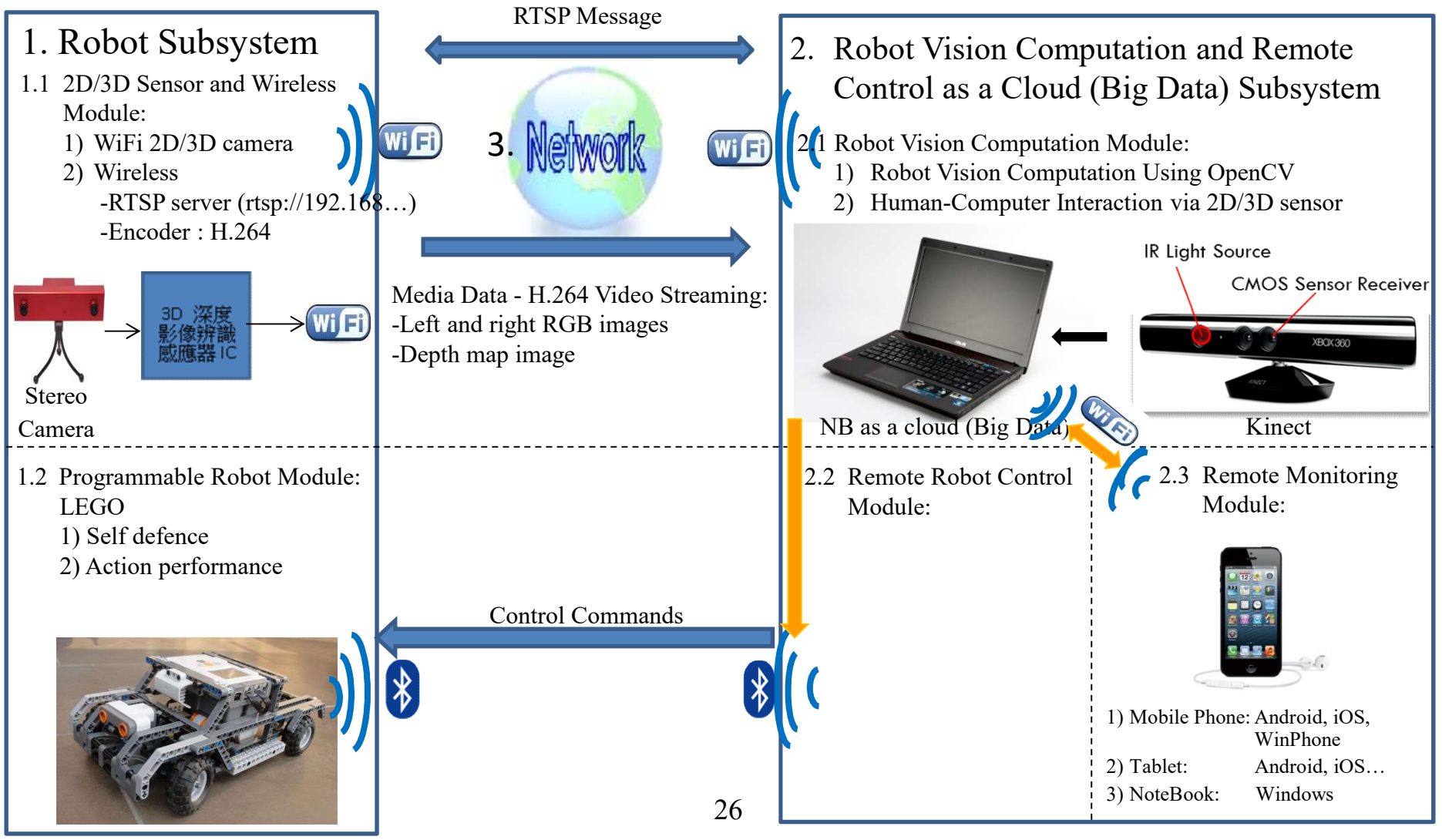
ISP (Image Signal Processing): Function Block 1
Embedded IVA (Intelligent Video Analytics): Function Blocks 2~7

ARM: Function Blocks 1~3
DSP: Function Blocks 3~7

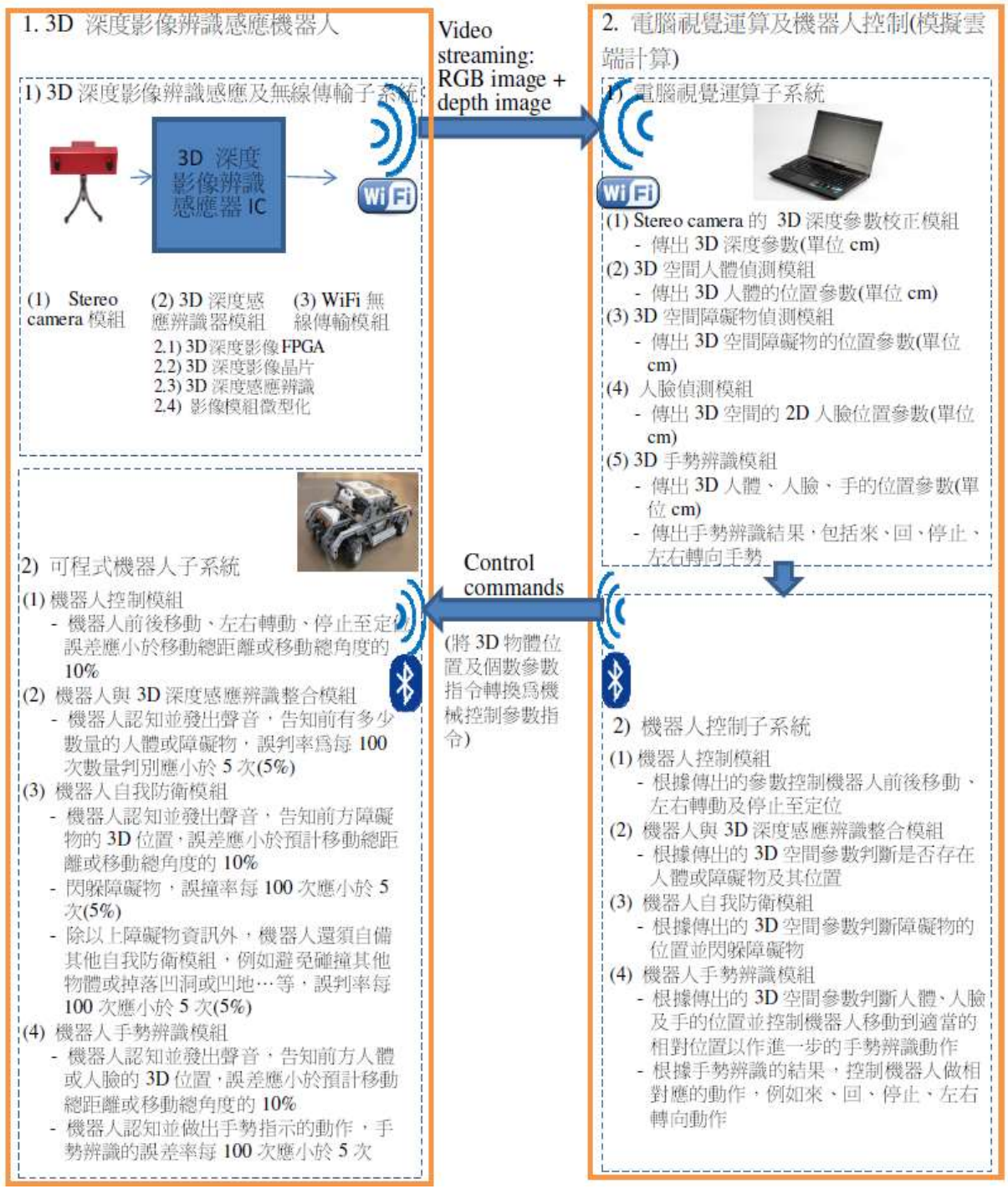
Embedded Computer Vision



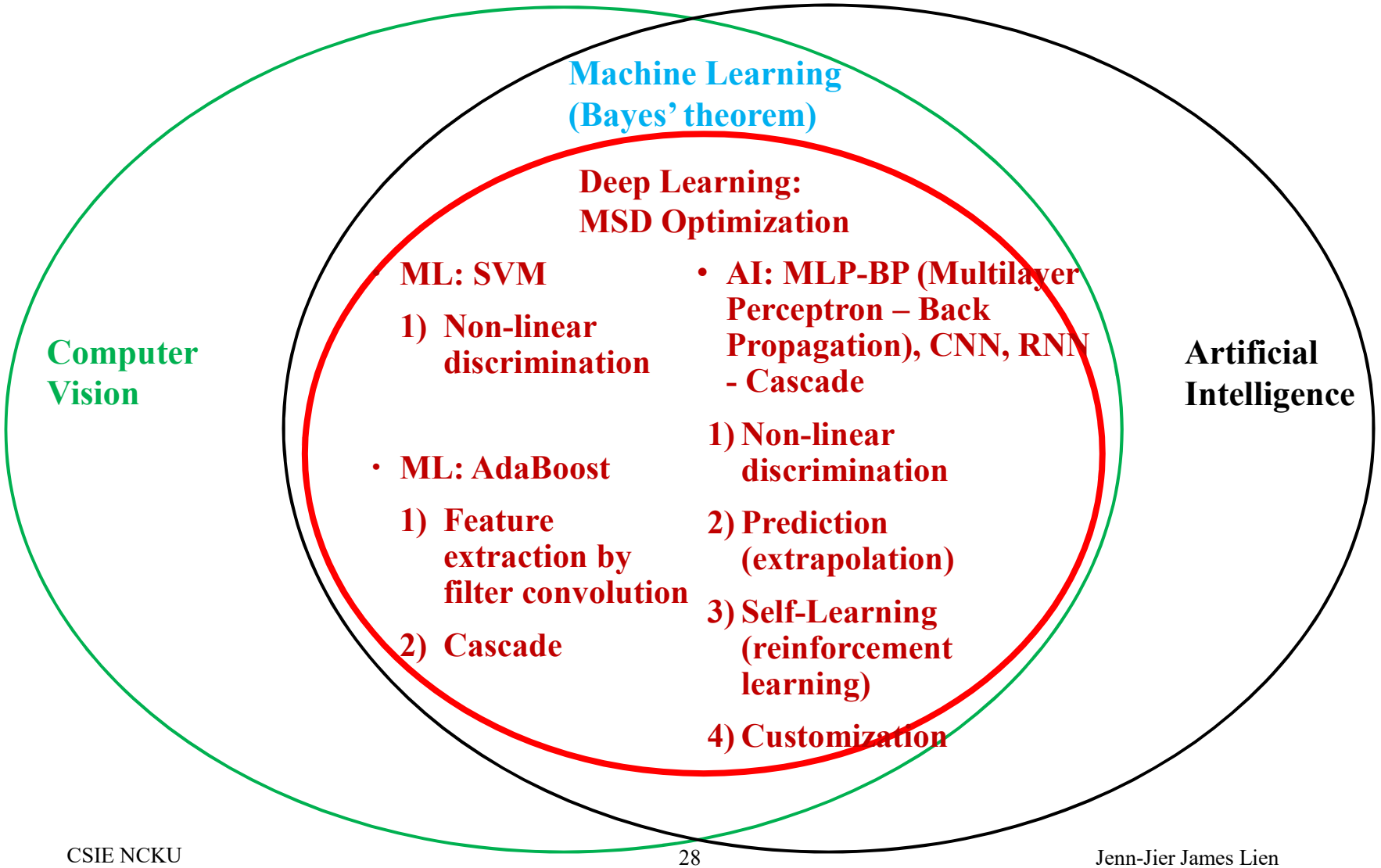
Embedded Robot Vision and Human-Computer Interaction



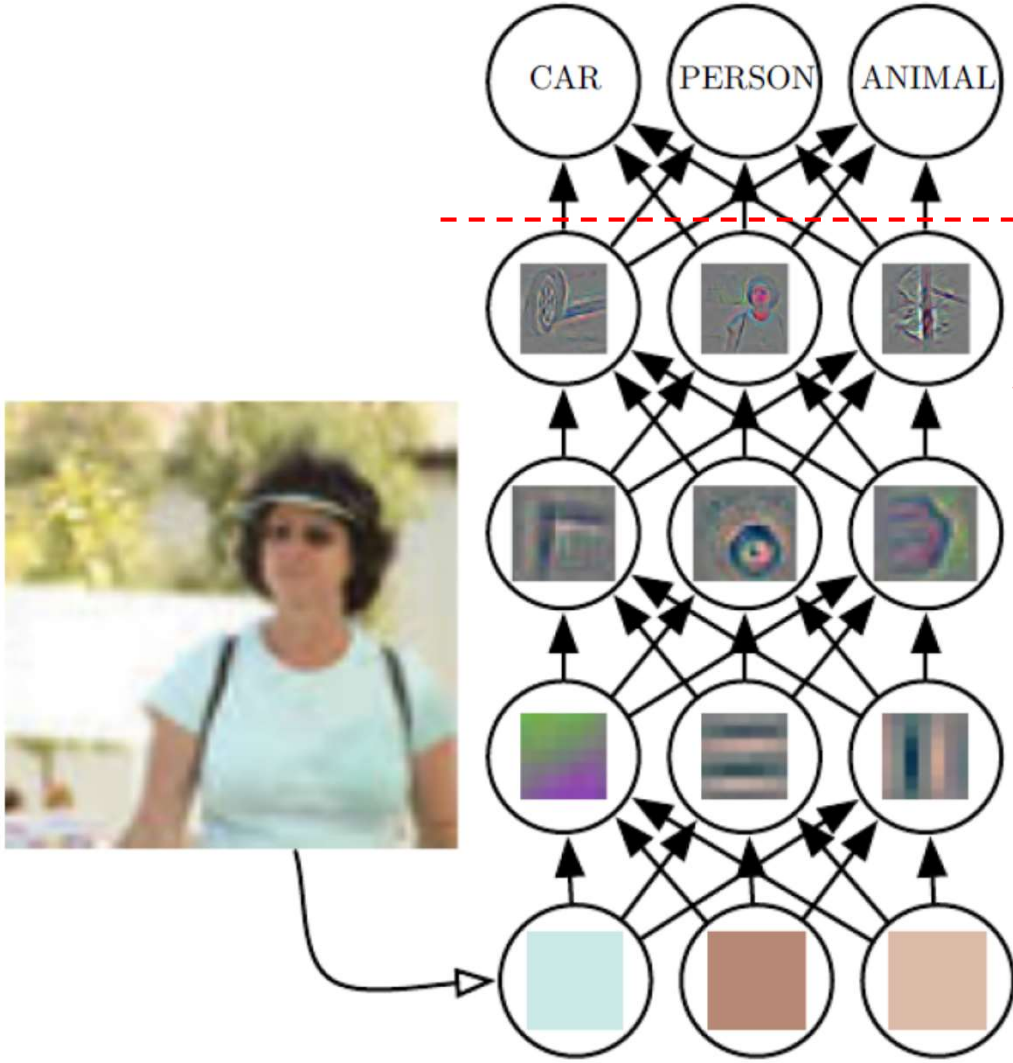
- **One example**


1. 3D 深度影像辨識感應機器人
1) 3D 深度影像辨識感應及無線傳輸子系統
3D 深度影像辨識感應器 IC
WiFi
(1) Stereo camera 模組
(2) 3D 深度感應辨識器模組
(3) WiFi 無線傳輸模組
2.1) 3D深度影像FPGA
2.2) 3D 深度影像晶片
2.3) 3D 深度感應辨識
2.4) 影像模組微型化
Video streaming: RGB image + depth image
2. 電腦視覺運算及機器人控制(模擬雲端計算)
1) 電腦視覺運算子系統
WiFi
(1) Stereo camera 的 3D 深度參數校正模組
- 傳出 3D 深度參數(單位 cm)
(2) 3D 空間人體偵測模組
- 傳出 3D 人體的位置參數(單位 cm)
(3) 3D 空間障礙物偵測模組
- 傳出 3D 空間障礙物的位置參數(單位 cm)
(4) 人臉偵測模組
- 傳出 3D 空間的 2D 人臉位置參數(單位 cm)
(5) 3D 手勢辨識模組
- 傳出 3D 人體、人臉、手的位置參數(單位 cm)
- 傳出手勢辨識結果，包括來、回、停止、左右轉向手勢
2) 可程式機器人子系統
(1) 機器人控制模組
- 機器人前後移動、左右轉動、停止至定位誤差應小於移動總距離或移動總角度的 10%
(2) 機器人與 3D 深度感應辨識整合模組
- 機器人認知並發出聲音，告知前有多少數量的人體或障礙物，誤判率爲每 100 次數量判別應小於 5 次(5%)
(3) 機器人自我防衛模組
- 機器人認知並發出聲音，告知前方障礙物的 3D 位置，誤差應小於預計移動總距離或移動總角度的 10%
- 閃躲障礙物，誤撞率每 100 次應小於 5 次(5%)
- 除以上障礙物資訊外，機器人還須自備其他自我防衛模組，例如避免碰撞其他物體或掉落凹洞或凹地…等，誤判率每 100 次應小於 5 次(5%)
(4) 機器人手勢辨識模組
- 機器人認知並發出聲音，告知前方人體或人臉的 3D 位置，誤差應小於預計移動總距離或移動總角度的 10%
- 機器人認知並做出手勢指示的動作，手勢辨識的誤差率每 100 次應小於 5 次
Control commands
(將 3D 物體位置及個數參數指令轉換爲機械控制參數指令)
2) 機器人控制子系統
(1) 機器人控制模組
- 根據傳出的參數控制機器人前後移動、左右轉動及停止至定位
(2) 機器人與 3D 深度感應辨識整合模組
- 根據傳出的 3D 空間參數判斷是否存在人體或障礙物及其位置
(3) 機器人自我防衛模組
- 根據傳出的 3D 空間參數判斷障礙物的位置並閃躲障礙物
(4) 機器人手勢辨識模組
- 根據傳出的 3D 空間參數判斷人體、人臉及手的位置並控制機器人移動到適當的相對位置以作進一步的手勢辨識動作
- 根據手勢辨識的結果，控制機器人做相對應的動作，例如來、回、停止、左右轉向動作

Deep Learning:
Computer Vision Vs. Artificial Intelligence



Deep Learning: CNN



Classification

CNN: Feature Extraction

3) Global (geometric) feature

2) Medium feature

1) Local feature

Feature:

1) Texture
   > Sensitive to light
2) Shape
   > Sensitive to noise
3) Color
   > RGB to HSV
   > RGB to YUV for compression
4) RGB+Depth

Jenn-Jier James Lien

PROC. OF THE IEEE, NOVEMBER 1998

Deep Learning: CNN

Gradient-Based Learning Applied to Document Recognition

Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner

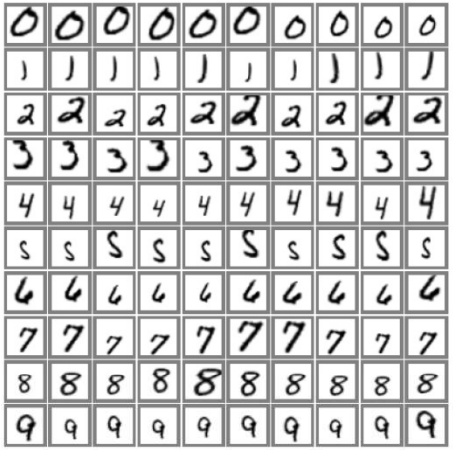


Fig. 7. Examples of distortions of ten training patterns.

Training **Test**

Fig. 4. Size-normalized examples from the MNIST database.

CNN: Extraction **Classification**

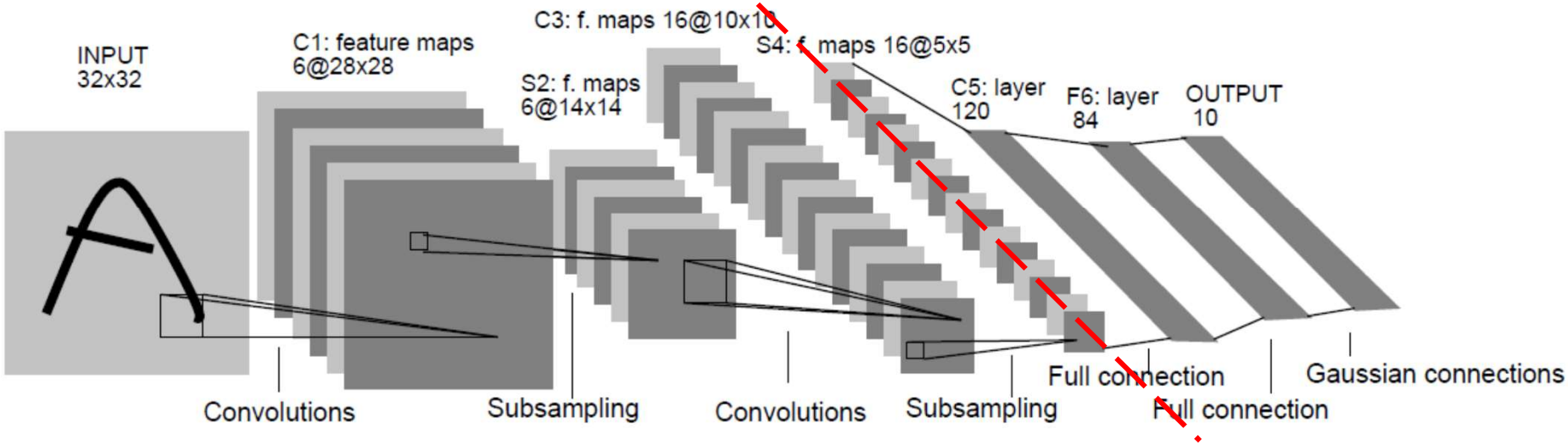


Fig. 2. Architecture of LeNet-5, a Convolutional Neural Network, here for digits recognition. Each plane is a feature map, i.e. a set of units whose weights are constrained to be identical.

MLP-BP

圖 1-2 人工神經元模型

$$y_j = f\left(\sum_i W_{ij} x_i - \theta_j\right)$$

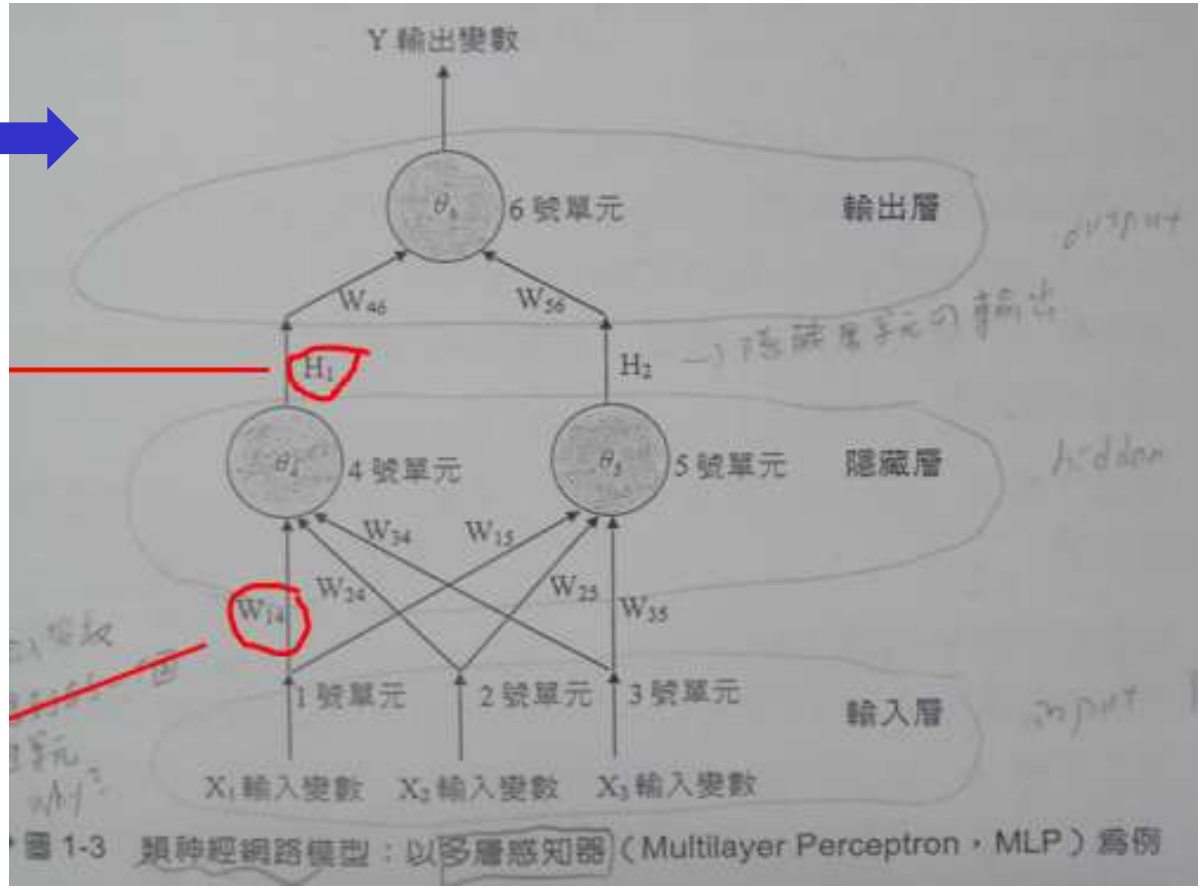
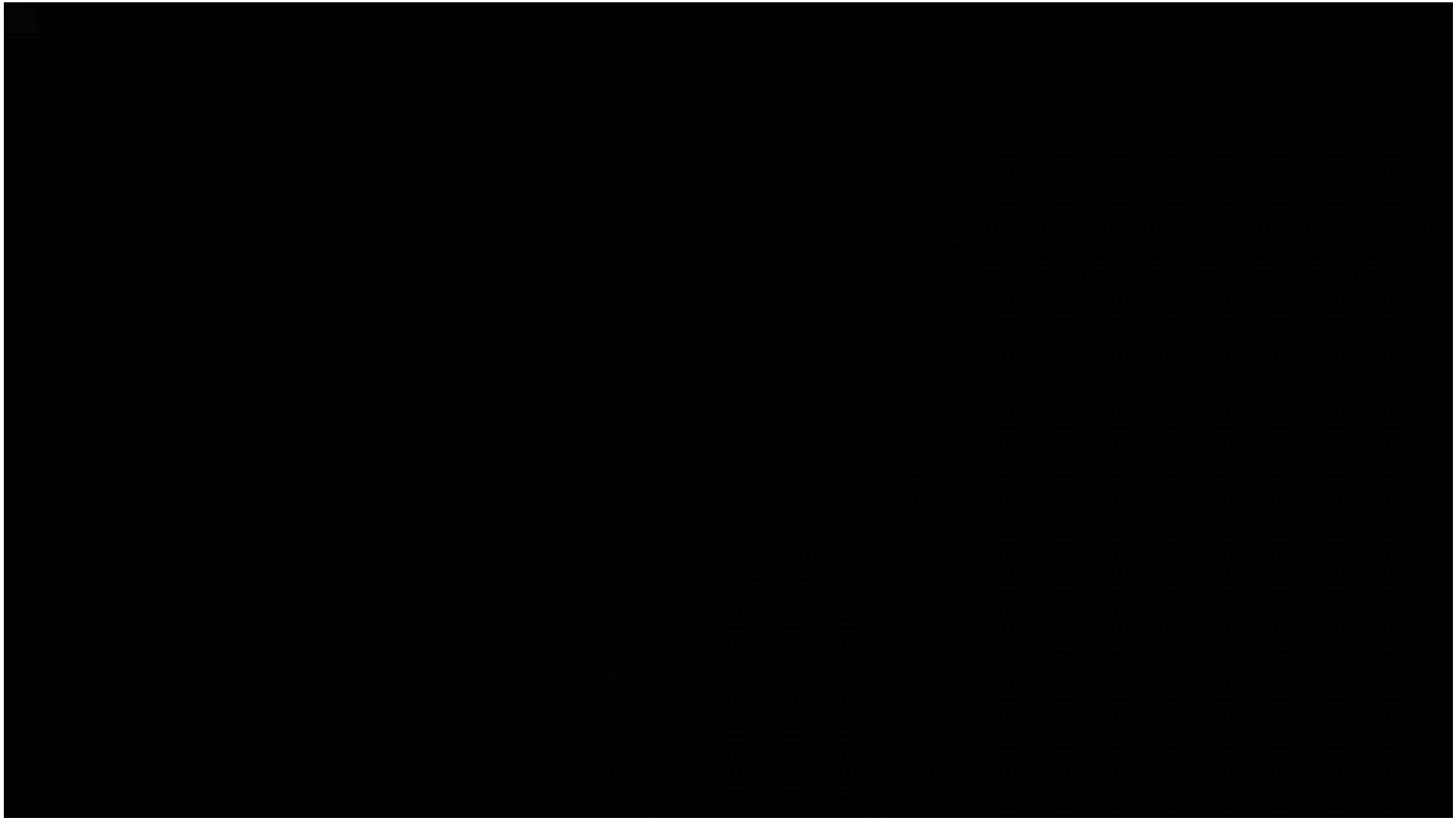


圖 1-3 類神經網路模型：以多層感知器（Multilayer Perceptron，MLP）爲例

Demo: Fanuc - Robot + Vision

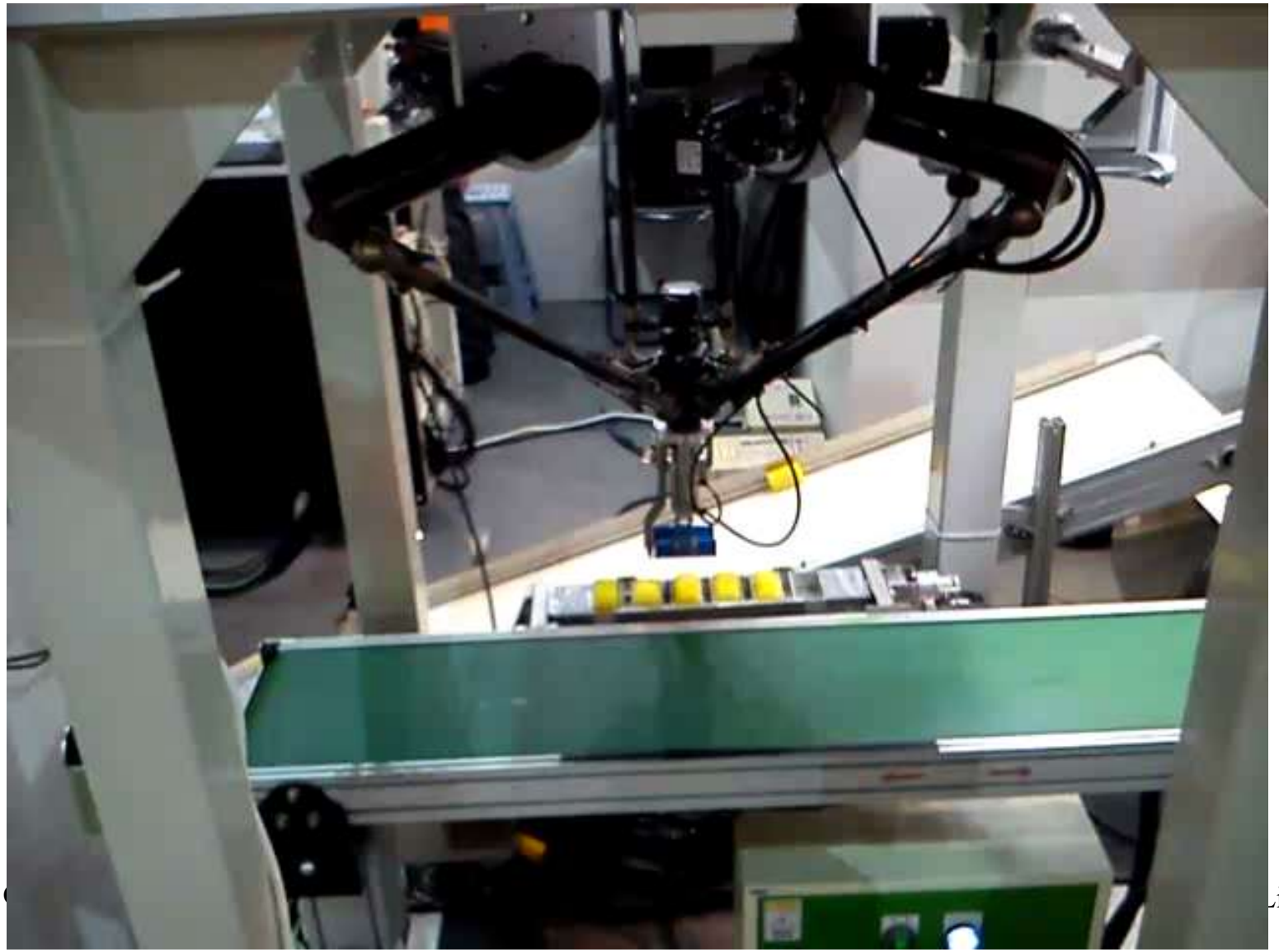


CV: 01:55~02:16

Demo: Robot Arm – Pick and Place (1/2)

C ien

Demo: Robot Arm – Pick and Place (2/2)



Demo: Robot Arm – Assembly (1/2)

Demo: Robot Arm – Assembly (2/2)

References

1. P. Ekman, and W.V. Friesen, *The Facial Action Coding System*, Consulting Psychologists Press Inc., San Francisco, CA, 1978.
2. J.J. Lien, "*Automatic Recognition of Facial Expressions Using Hidden Markov Models and Estimation of Expression Intensity*," Ph.D. Dissertation, Technical Report, Carnegie Mellon University, Robotics Institute, CMU-RI-TR-98-31, May, 1998.
   - Available at http://www.cs.cmu.edu/~jjlien and http://robotics.csie.ncku.edu.tw

❑ Relative Websites

1) www.ri.cmu.edu
2) www.ri.cmu.edu/labs/lab_51.html
3) www.cs.cmu.edu/~tk
4) www.engr.pitt.edu/electrical/about/faculty/li_chingchung.html