

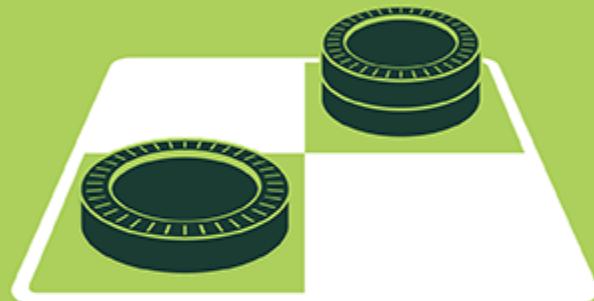
機器學習於材料資訊的應用 Machine Learning on Material Informatics

陳南佑(NAN-YOW CHEN)

楊安正(AN-CHENG YANG)

ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



1950's

1960's

1970's

1980's

1990's

2000's

2010's

MACHINE LEARNING

Machine learning begins to flourish.



DEEP LEARNING

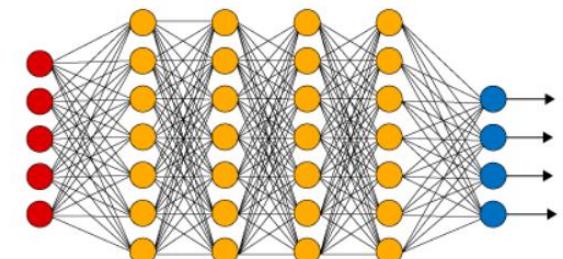
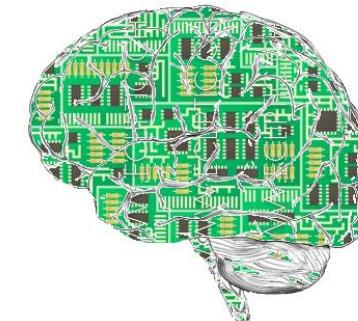
Deep learning breakthroughs drive AI boom.



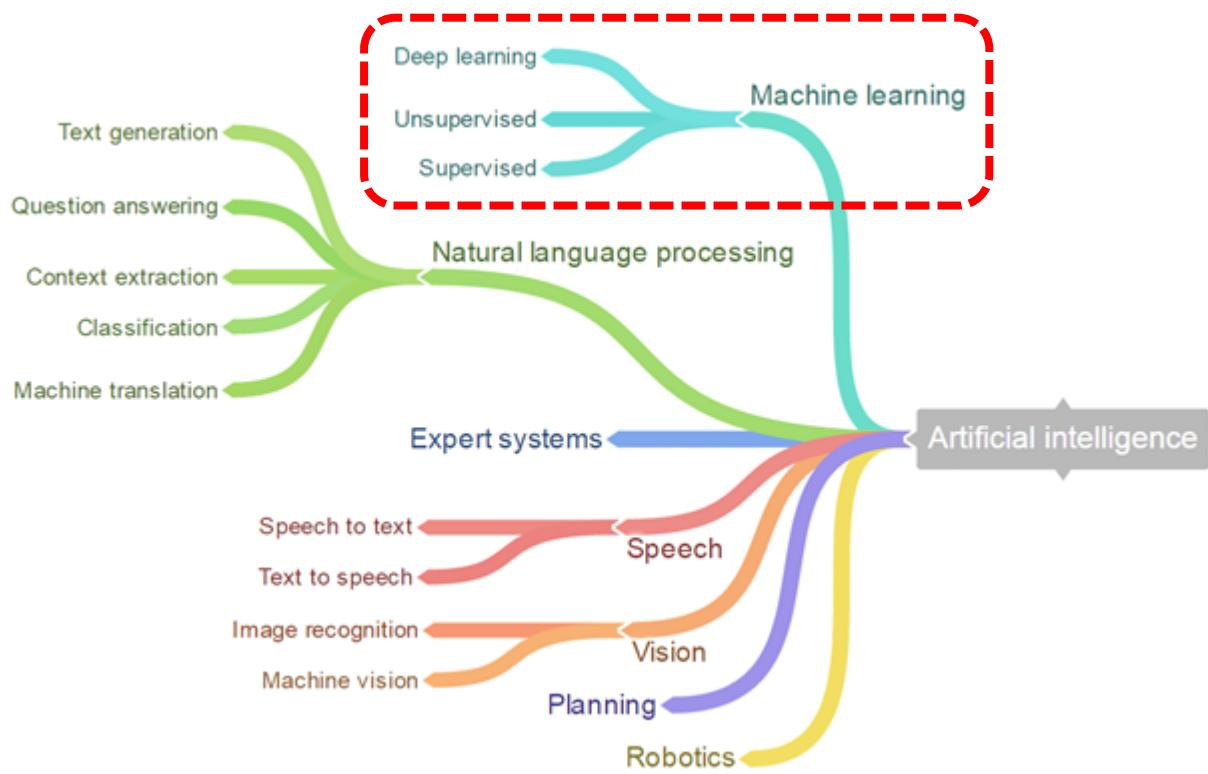
Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

What is AI, ML, DL?

- **AI** (Artificial Intelligence): *the ability of the computer to simulate/model human thinking processes to imitate human ability or behavior.*
(consciousness, cognitive, mind, ...)
- **ML** (Machine Learning): *is the scientific study of **algorithms** and **statistical models** that computer systems use to effectively perform a specific task without using explicit instructions, relying on patterns and inference instead.*
(learning from the data, **data-driven model**)
- **DL** (Deep Learning): *use a cascade of **multiple layers of nonlinear processing units** for feature extraction and transformation. Each successive layer uses the output from the previous layer as input.*
(learning from the data, **multi-layer neural network**)



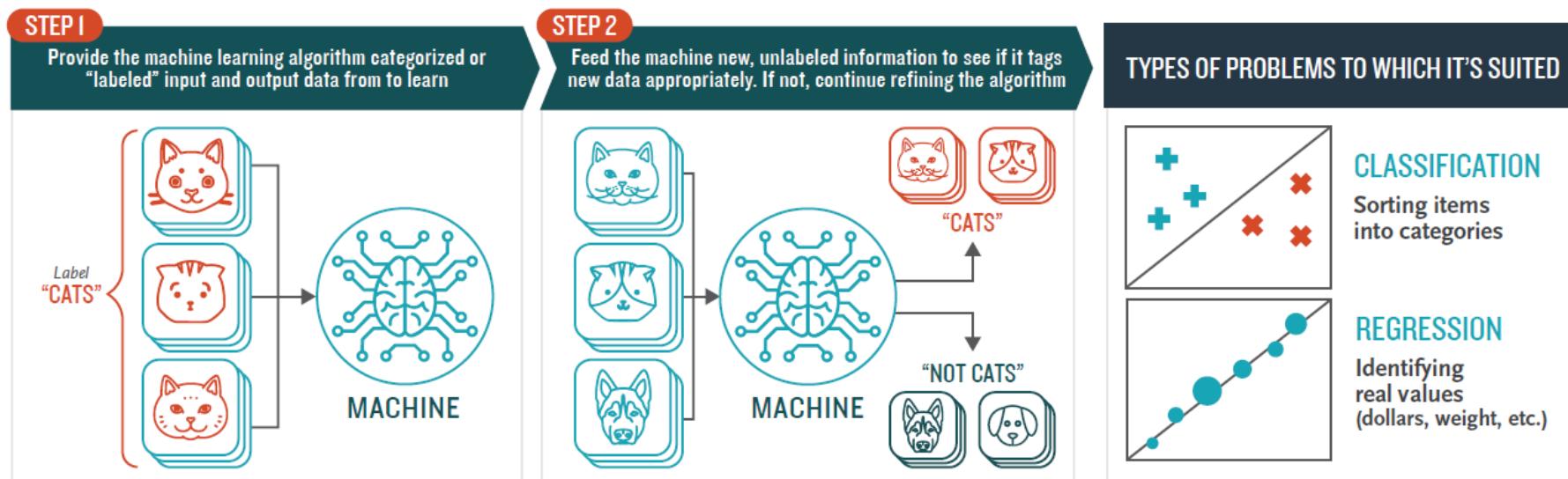
Applications of AI



Types of ML

□ Supervised Learning (監督式學習):

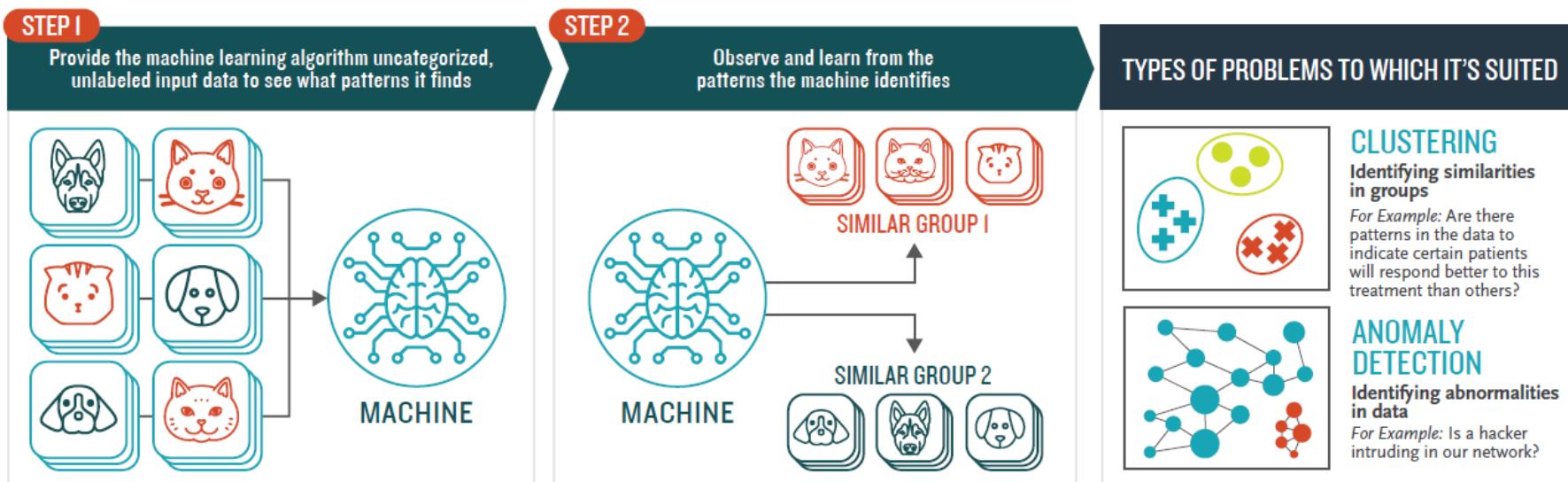
- 可以由訓練資料中學到或建立一個模式(函數 / learning model)，並依此模式推測新的實例。訓練資料是由輸入物件(通常是向量)和預期輸出所組成。函數的輸出可以是一個連續的值(稱為迴歸分析)，或是預測一個分類標籤(稱作分類)。
- 人工神經網路(Artificial Neural Network, ANN)、決策樹(Decision Tree)、支援向量機(Support Vector Machine, SVM)、...



Types of ML

□ Unsupervised Learning (非監督式學習):

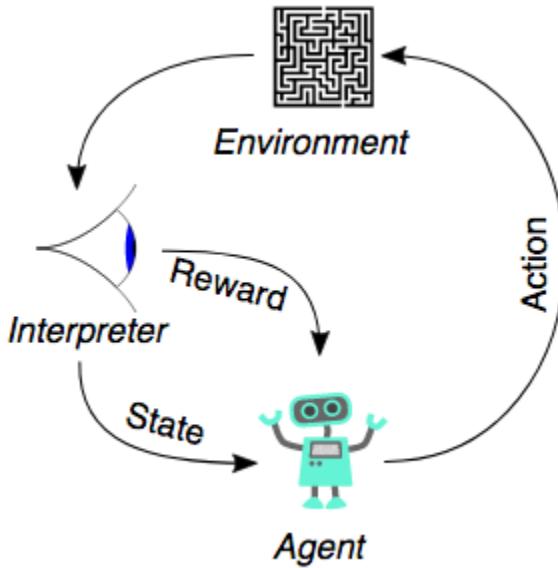
- 沒有給定事先標記過的訓練範例，自動對輸入的資料進行分類或分群。
- 分群(Cluster Analysis)、關聯規則(Association Rule)、維度縮減(Dimensionality Reduce)、...



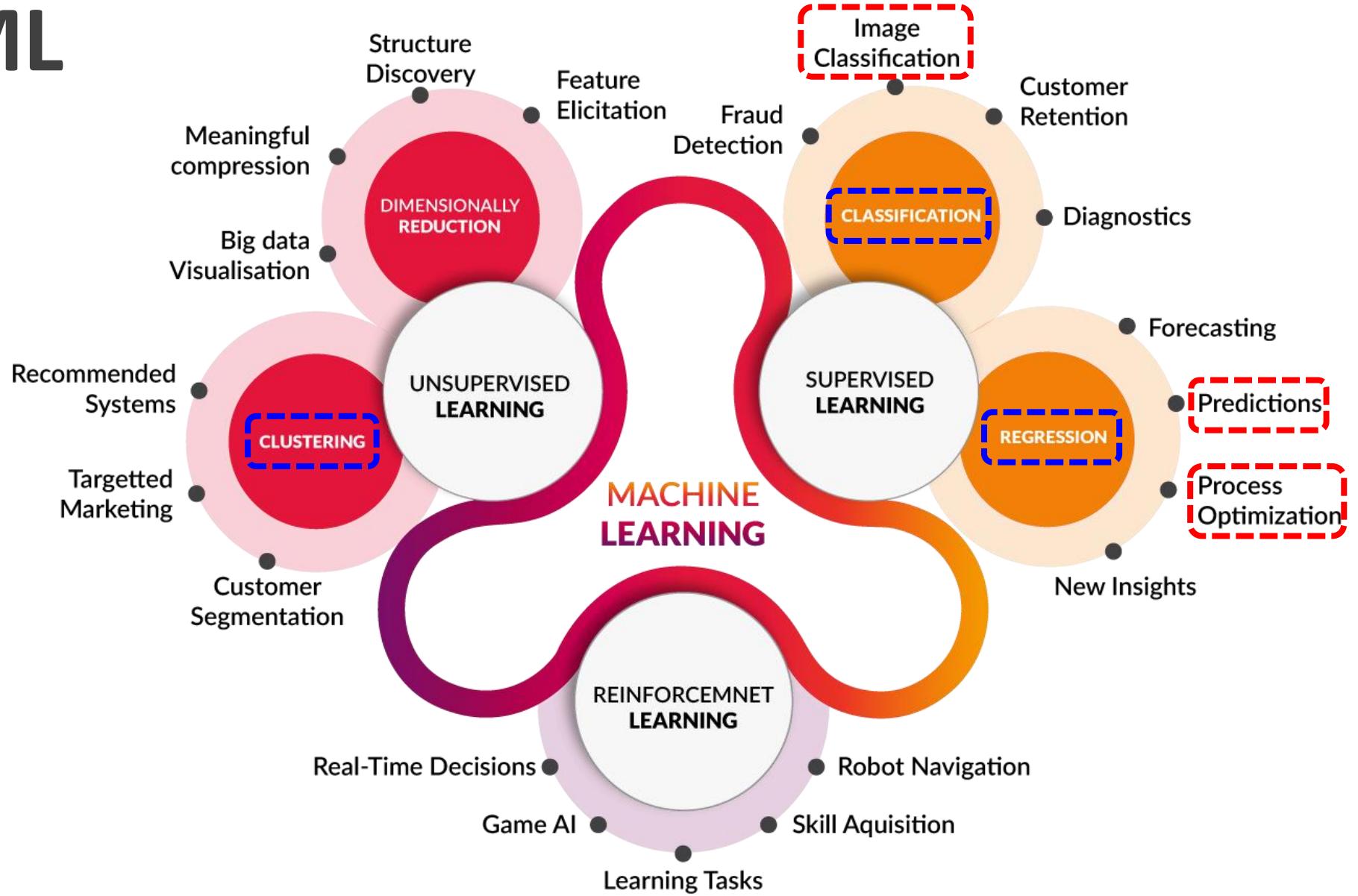
Types of ML

□ Reinforcement Learning (強化學習):

- 強調如何基於環境而行動，以取得最大化的預期利益。其靈感來源於心理學中的行為主義理論，即有機體如何在環境給予的獎勵或懲罰的刺激下，逐步形成對刺激的預期，產生能獲得最大利益的習慣性行為。

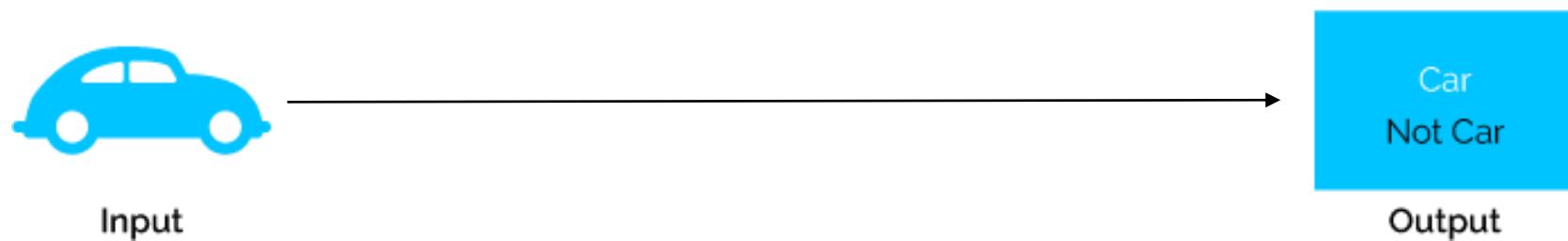
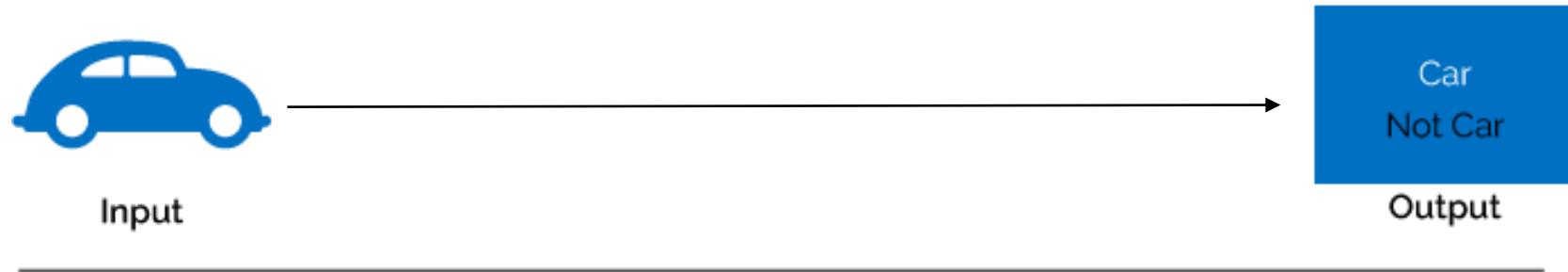


Types of ML

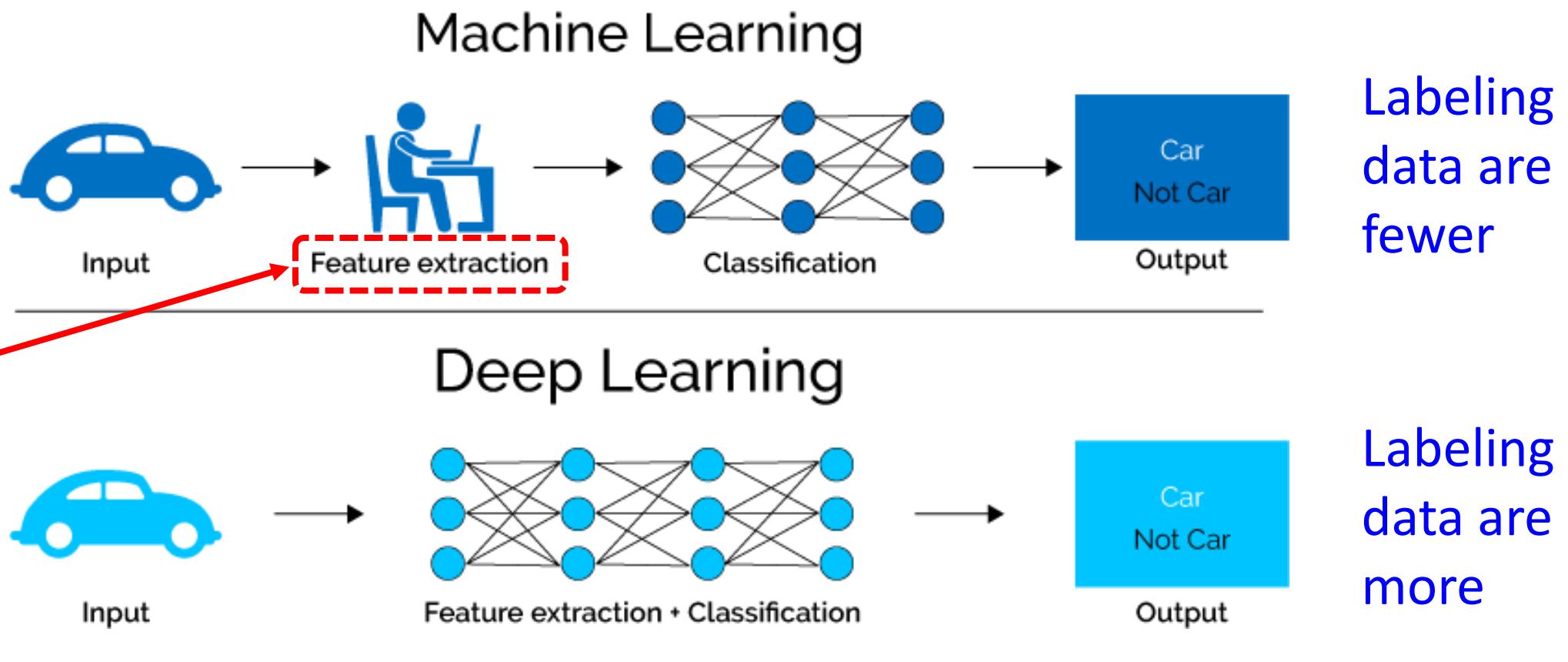


What the same between ML & DL?

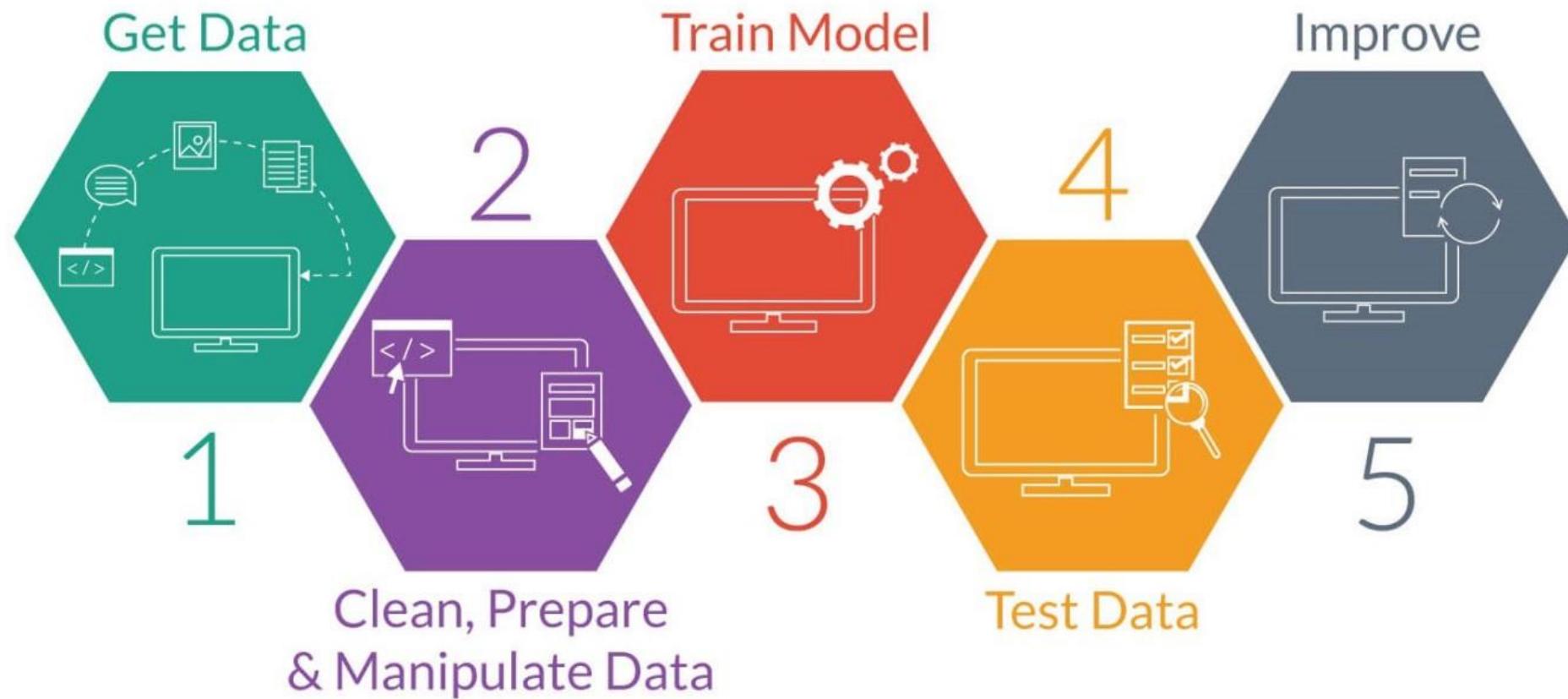
Need labeling data for training



What difference between ML & DL?

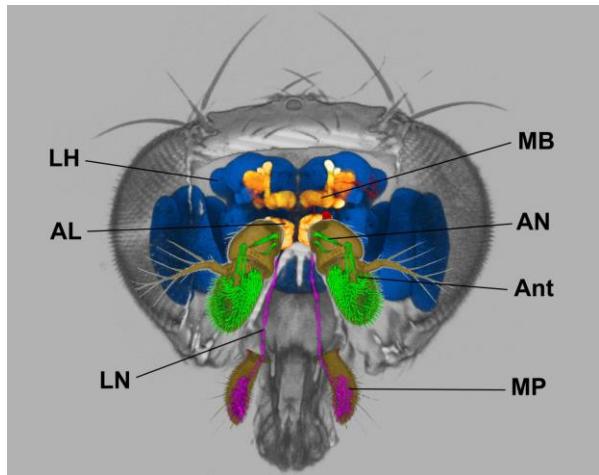


Work flow for ML

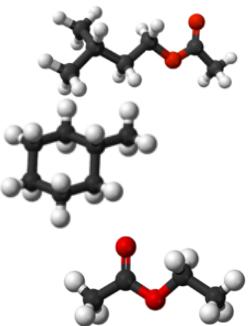


What is neuron?

Example 1: 成年果蠅的嗅覺系統

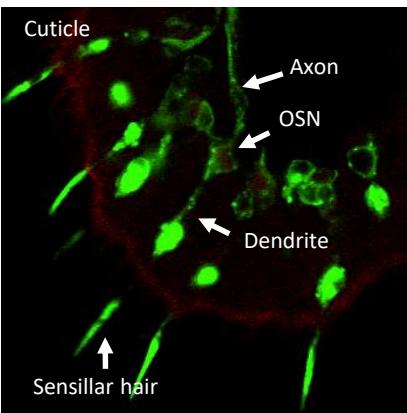
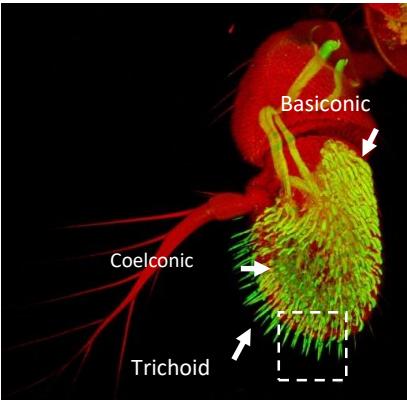


Odorant molecules



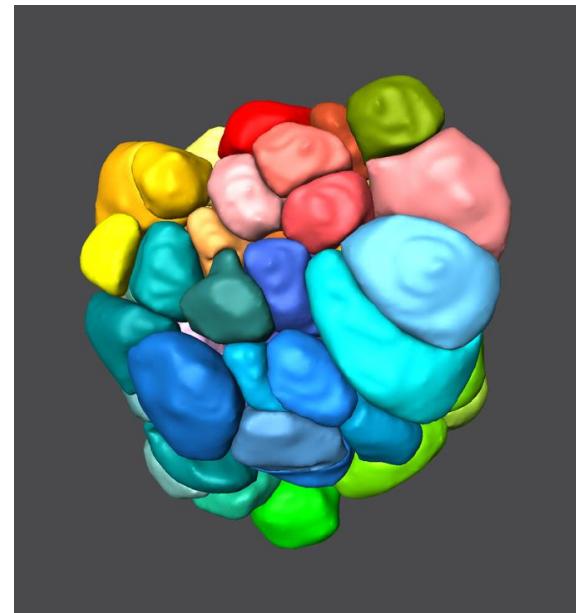
Input

Olfactory Sensory Neurons



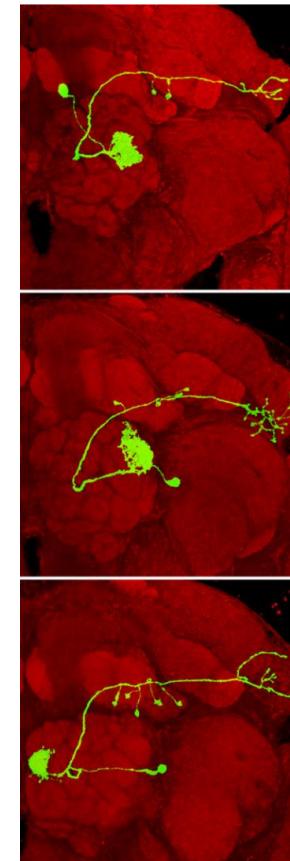
Antennal Lobe

Local processing unit for olfactory information



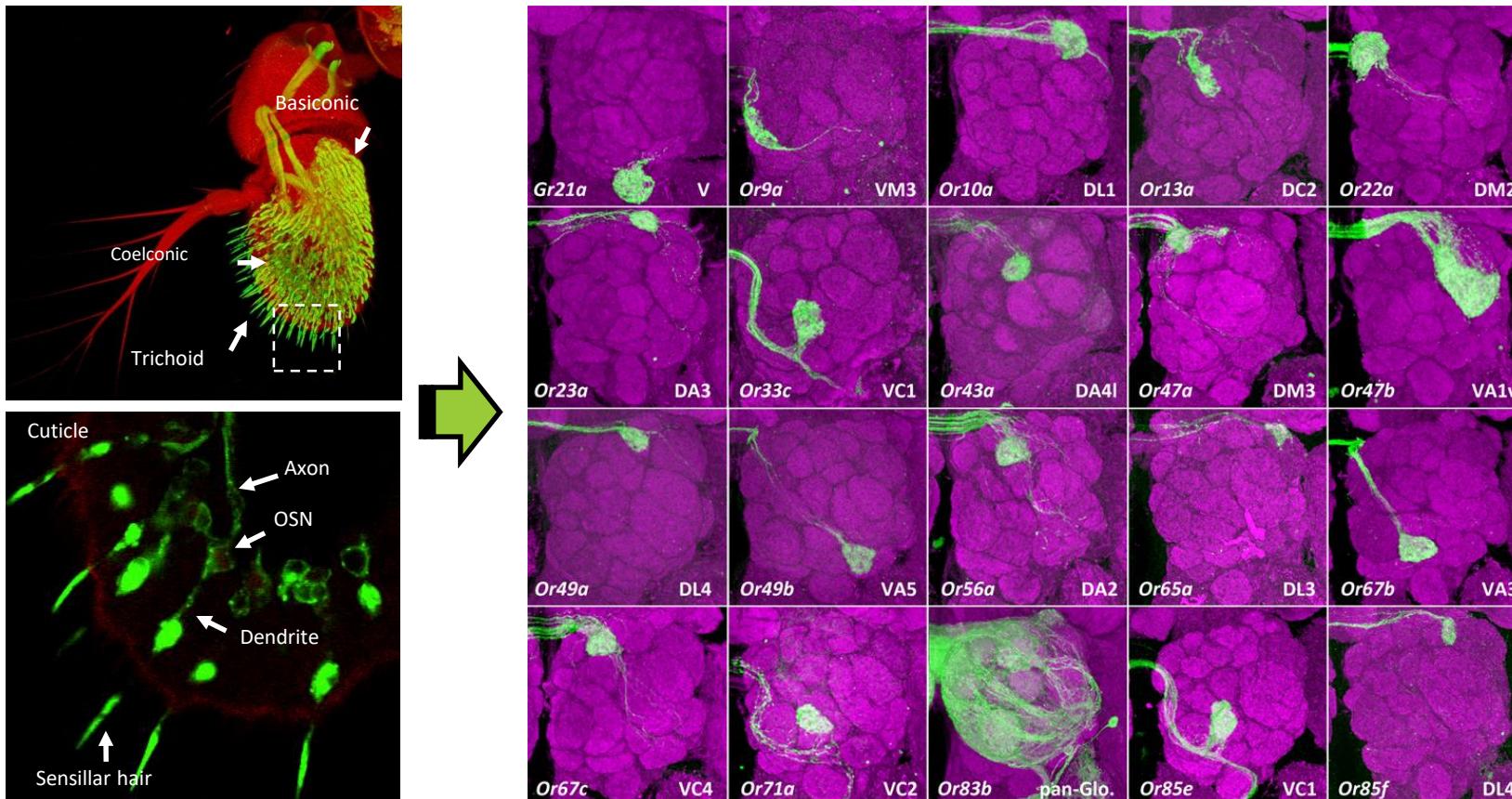
Output

Projection Neurons



AL Input – Olfactory Sensory Neurons

~1300 OSNs expressing 62 odorant receptors convey olfactory information from antennae and maxillary pulps to ~50 glomeruli



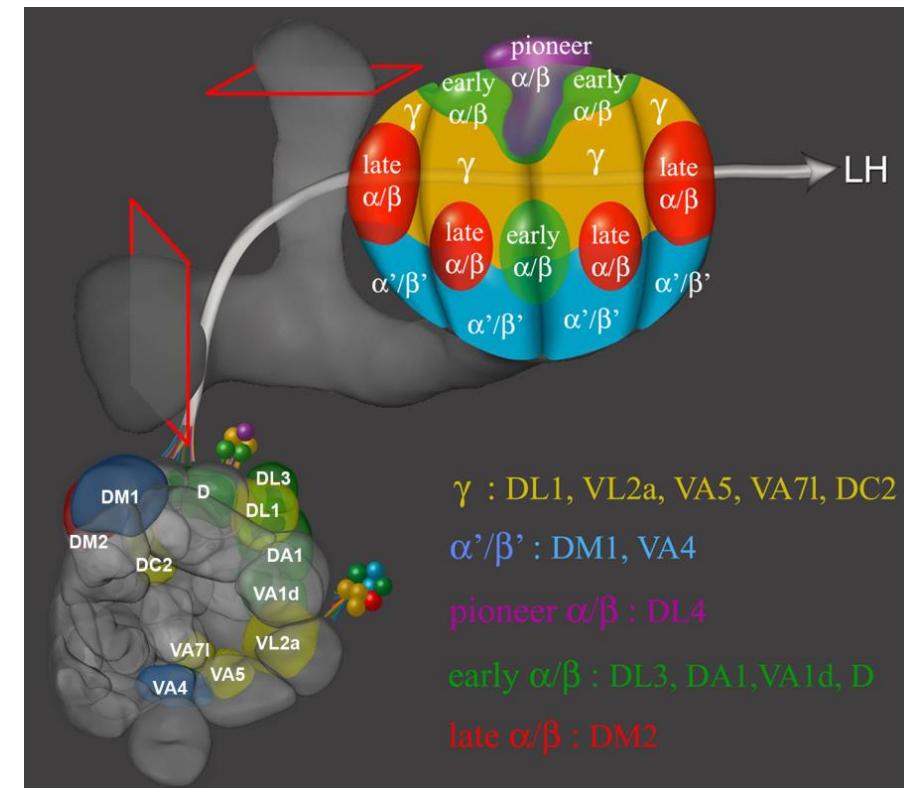
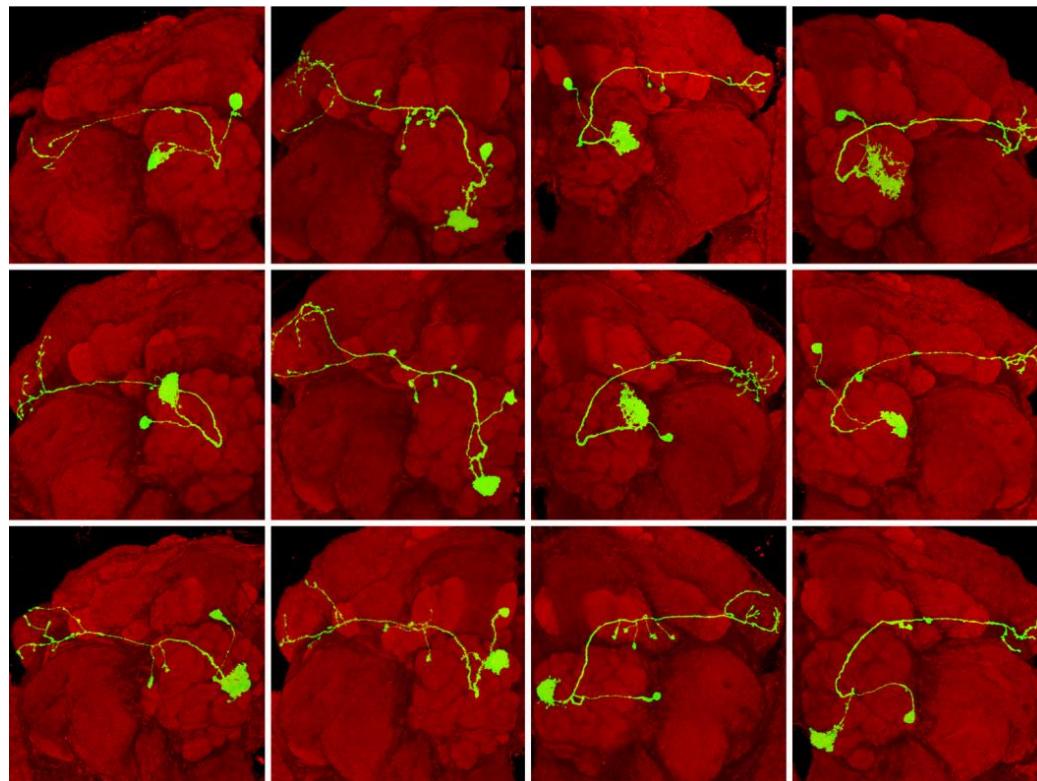
“one neuron – one
receptor”

“one glomerulus – one receptor”

AL Output – Projection Neurons

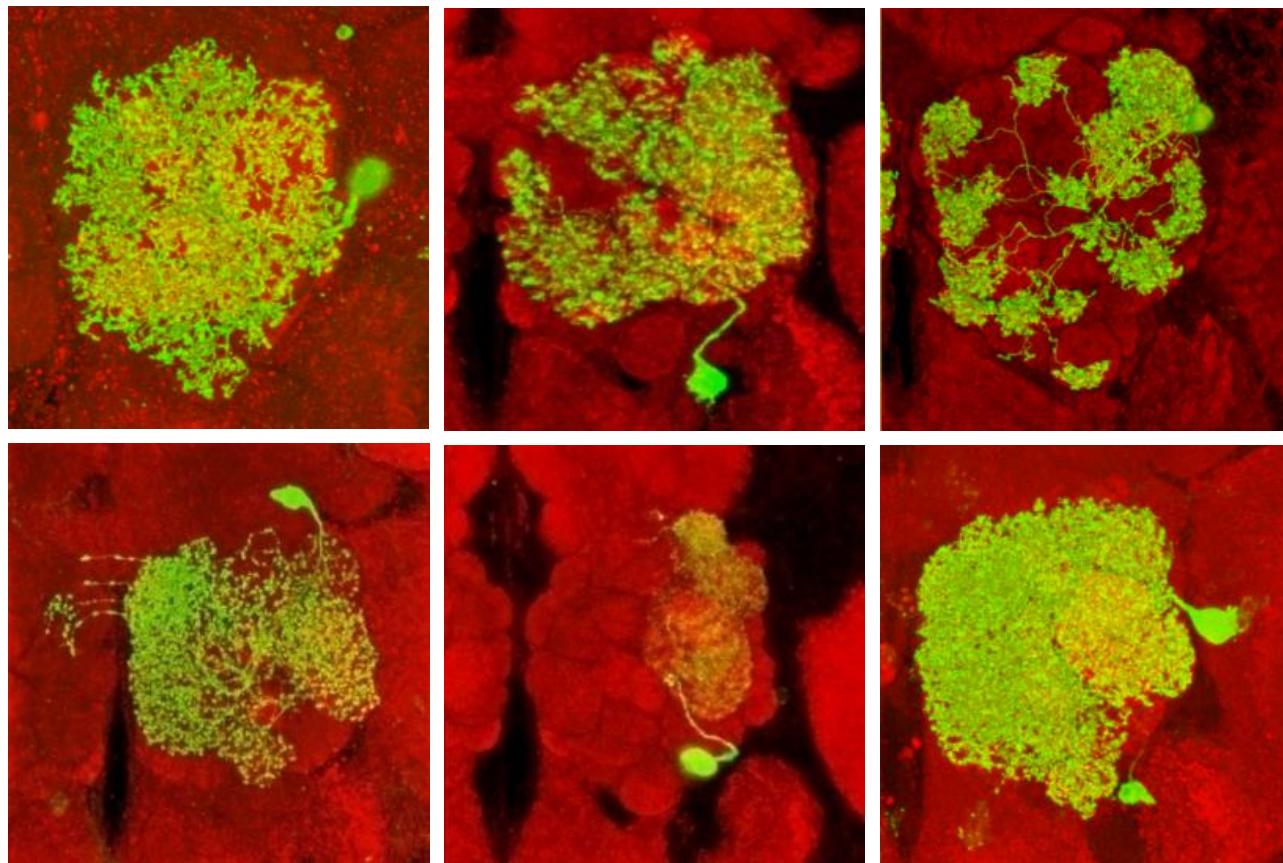
~150 typical PNs convey olfactory information from antennal lobe to higher brain centers

Stereotyped PN terminals

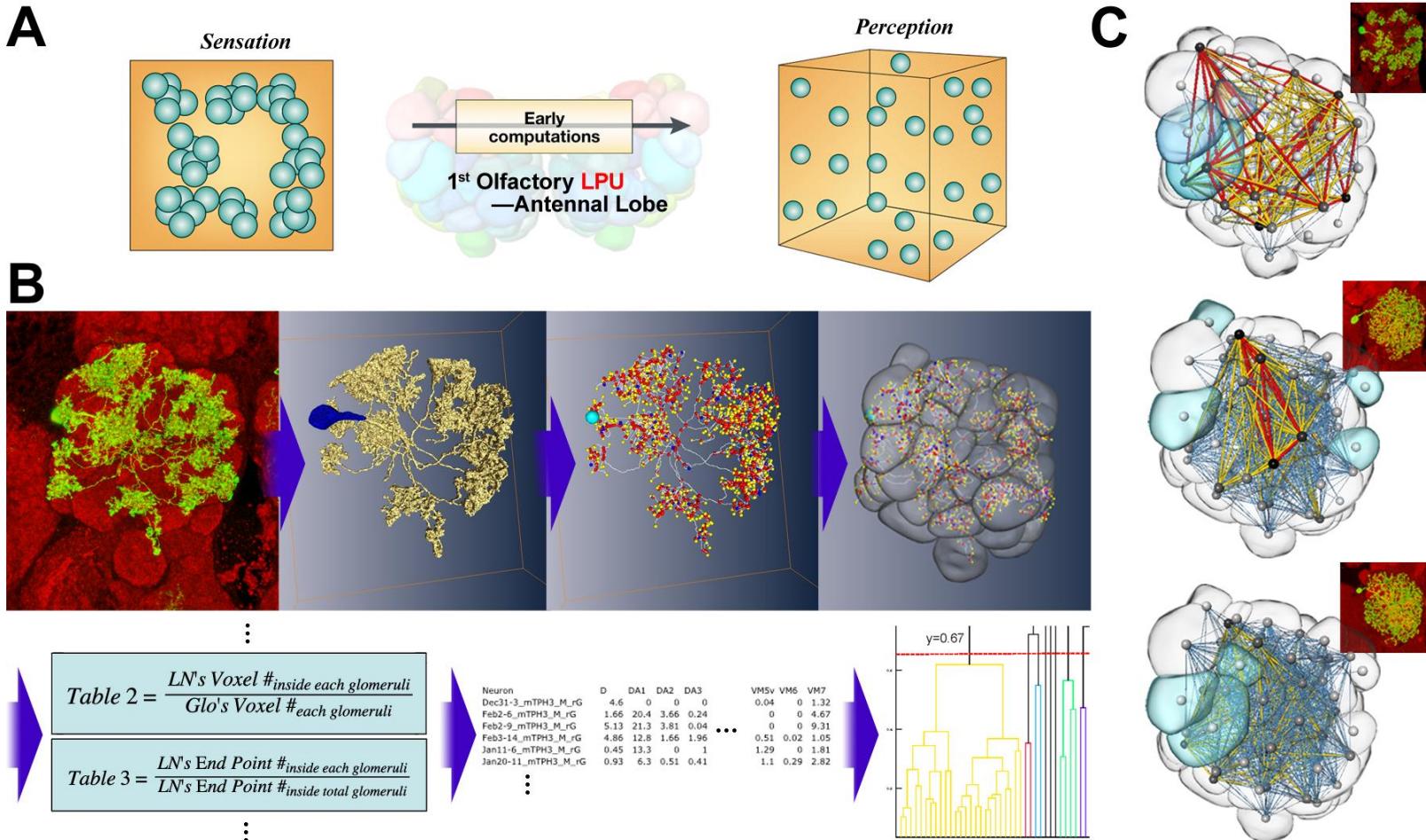


Antennal Lobe – Local Neurons

Local processing unit for olfactory information

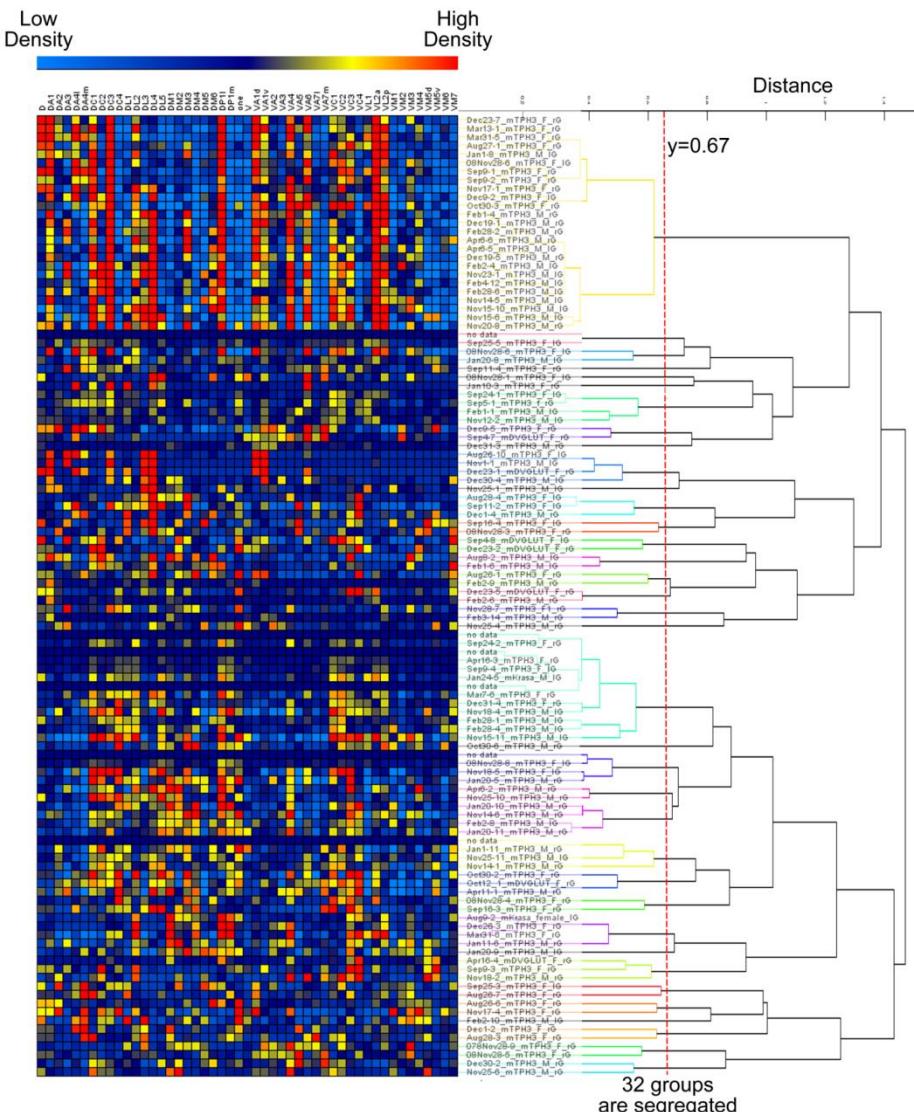


建構果蠅嗅覺神經元連結網路的程序

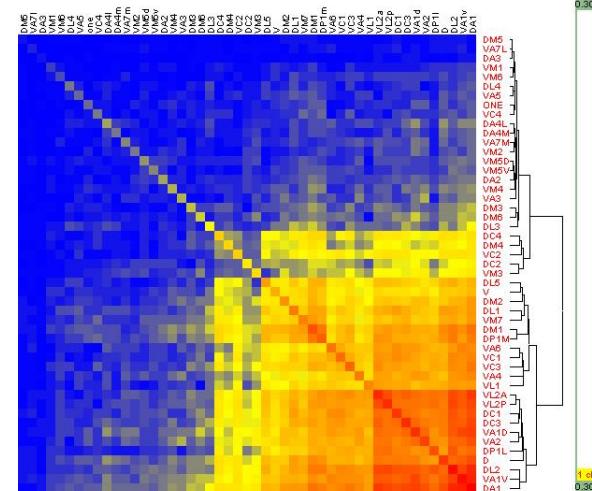


果蠅嗅覺神經元連結網路

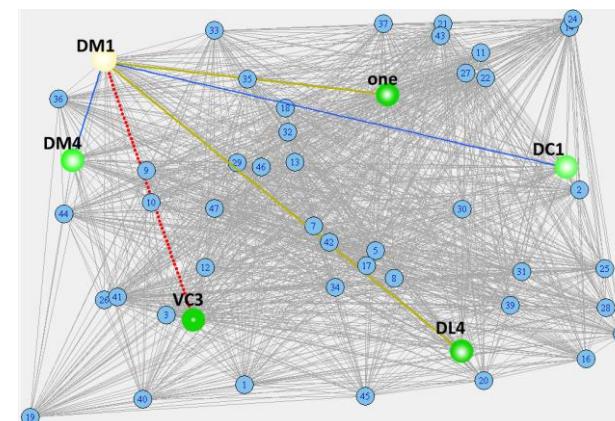
Innervation table (神經元分佈表)



Association matrix (關聯矩陣)



Network of local neurons
(局域神經元連結網路)



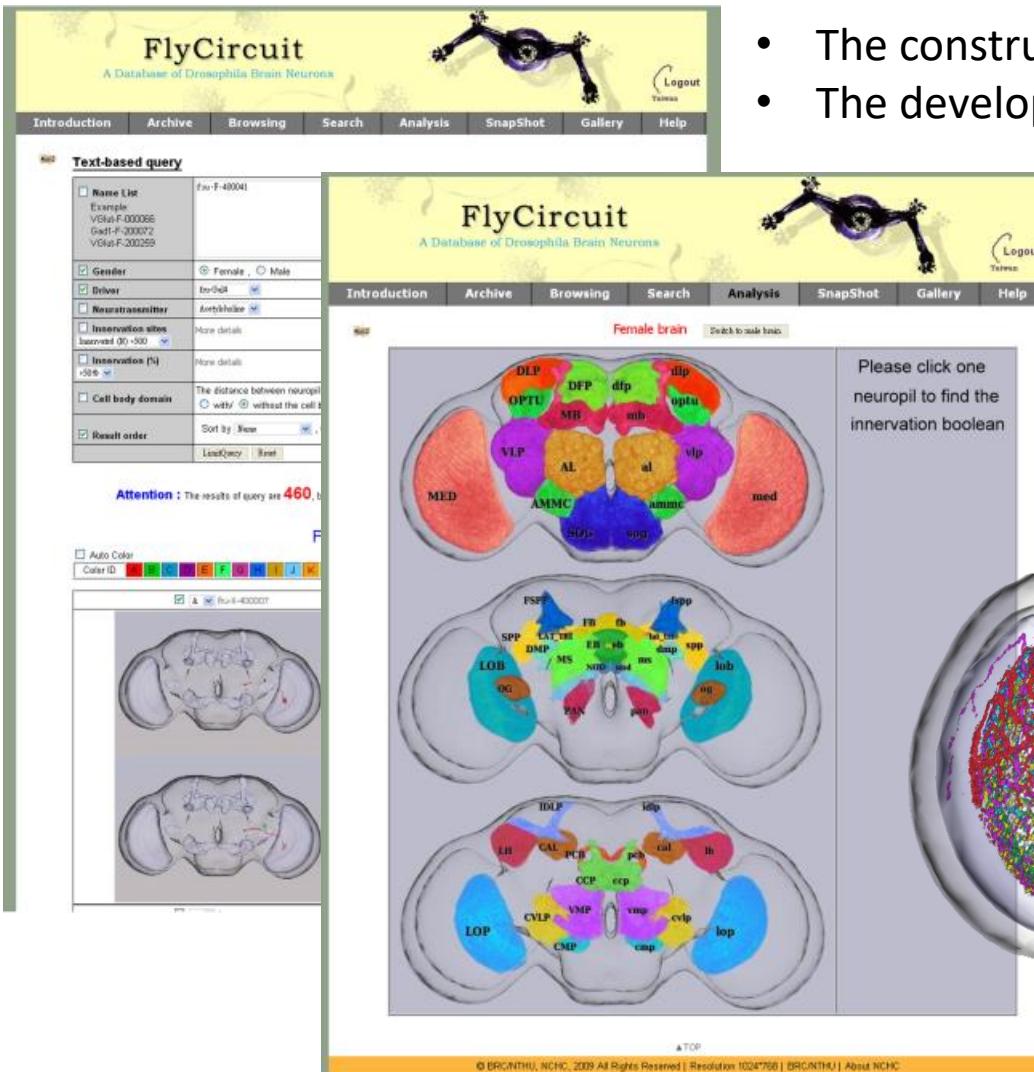
Association matrix A for each ALLN image

If $A_{ij} < \text{threshold}$, $A_{ij} = 0$ ($A_{ij} = Table2_i \cdot Table2_j$)

Example 2 : *FlyCircuit* v1.0

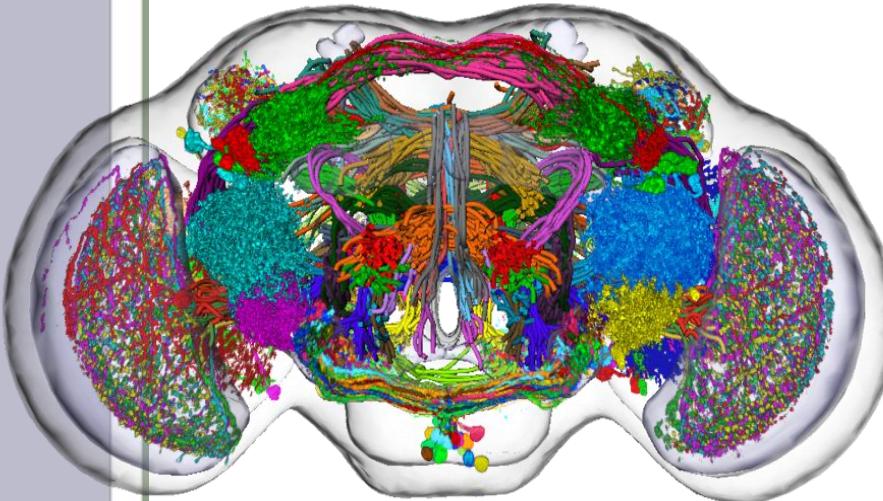
An Image Database of *Drosophila* Brain Neurons

3-D map of fly brain is to neuroscience what genome is to genetics



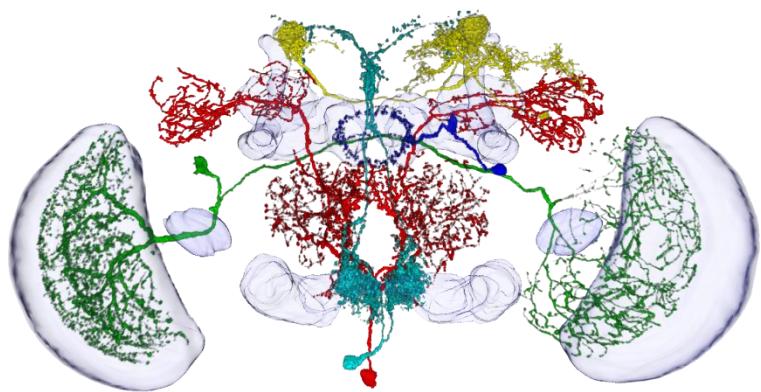
- The construction of 3D neuron images database
 - The development of searching and browsing tool

Current status: 30,000 neurons
(23.5% of 130,000 total brain)

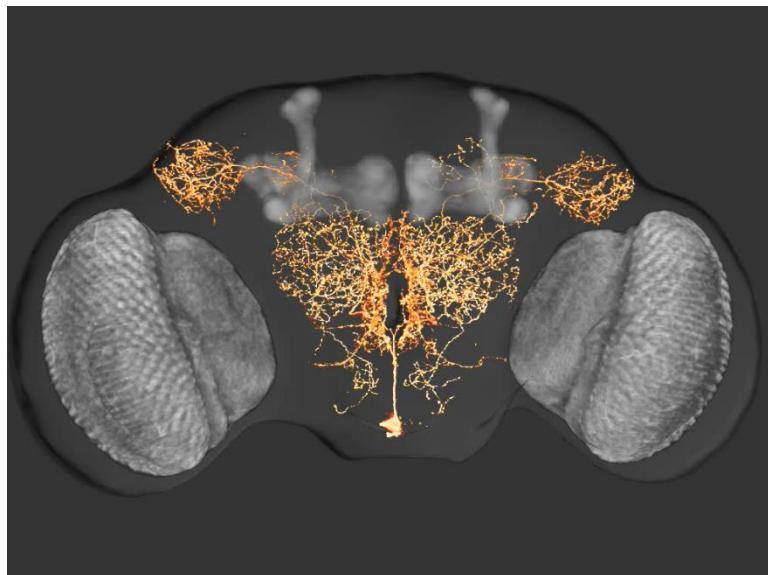
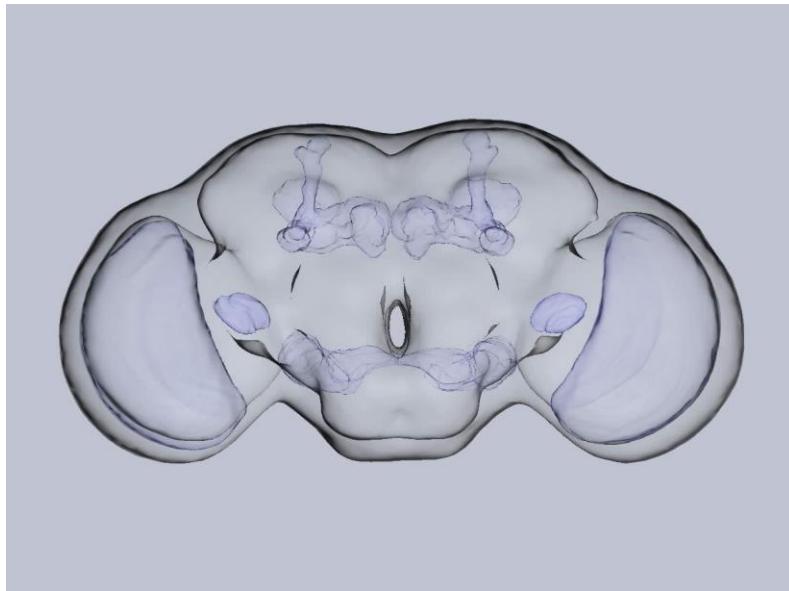


Current Biology 21:1–11 (2011)

3D Compilation & Visualization



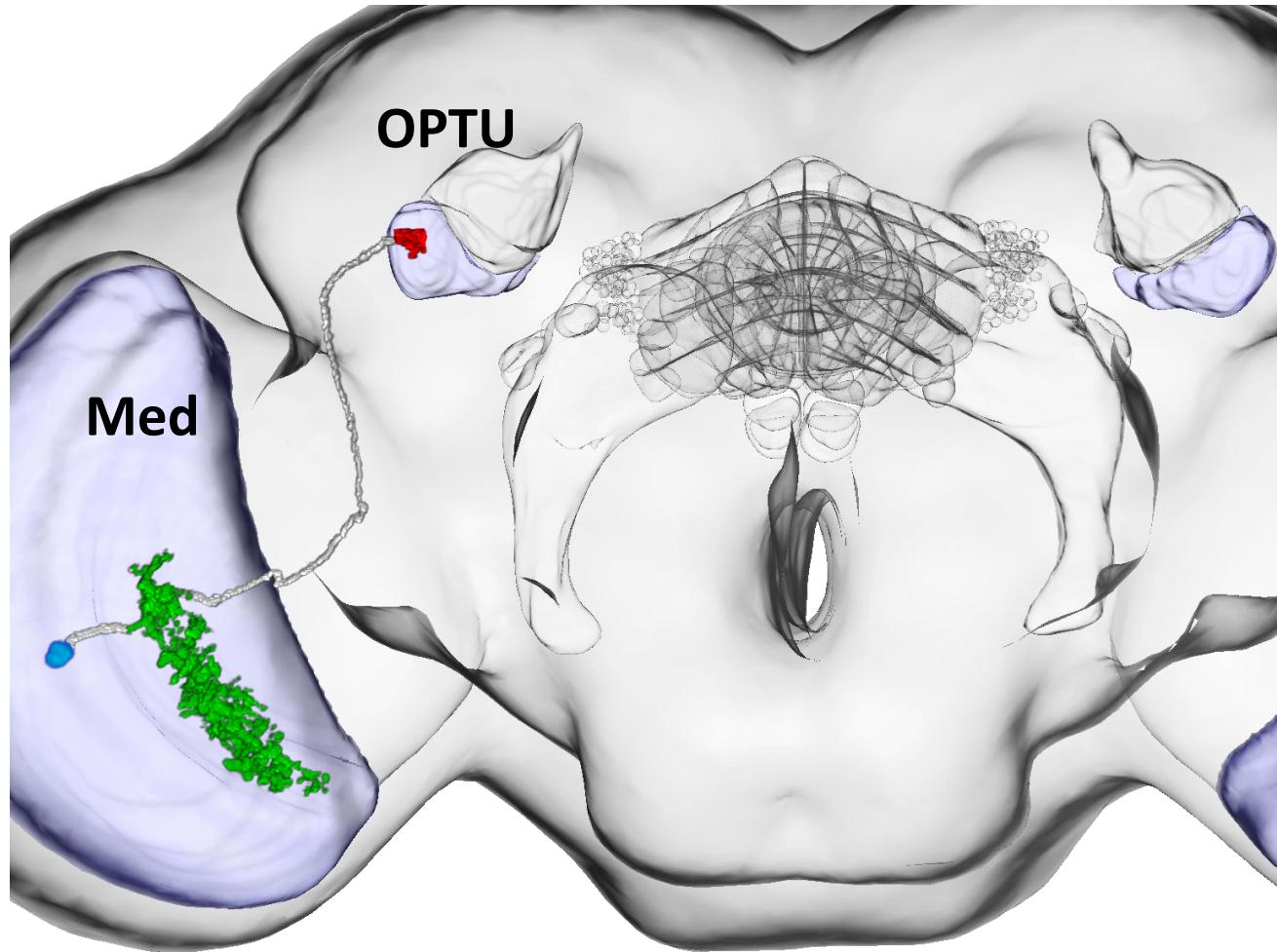
VGlut TpH TH GAD Tdc2



FlyCircuit v2.0

Brain-wide wiring networks in the *Drosophila* Brain

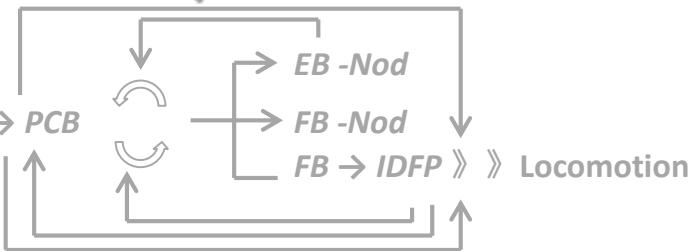
Mapping the neural network



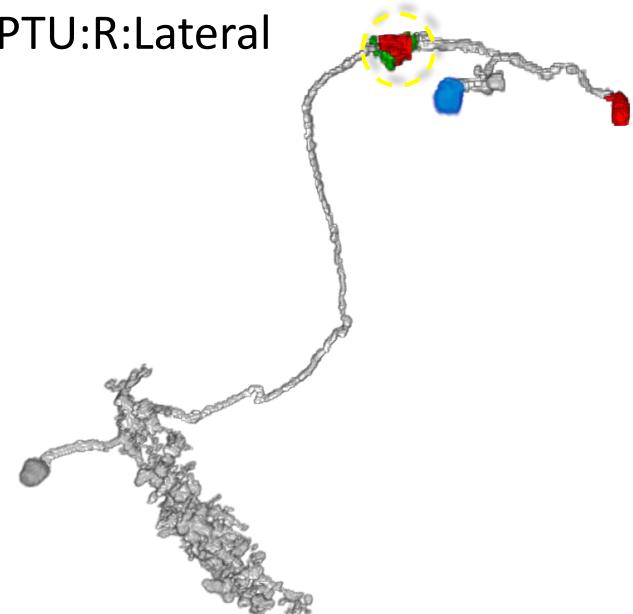
Visual input → Med:R → OPTU:R:Laterl

An UV Circuit from Eye to Pre-motor Center

Med → *OPTU* → *LTR* → *EB* → *PCB*



OPTU:R:Lateral



Med:R → *OPTU:R:Laterl*



OPTU:R:Laterl → *LTR:R*

Welcome to FlyCircuit Database

FlyCircuit
A Database of Drosophila Brain Neurons

Introduction Archive Browsing Search Analysis SnapShot Gallery Help

Login Taiwan

FlyCircuit is a public database for online archiving, cell type inventory, browsing, searching, analysis and 3D visualization of individual neurons in the Drosophila brain. For more details, please read the associated manuscript – “Three-dimensional reconstruction of brain-wide sensory neurons in Drosophila” (Chen et al., 2010; doi:10.1016/j.cub.2010.11.050). Registration is only necessary for users who want to upload/download data or activate 3D viewer. All images are copyright protected.

Acknowledging the NCHC
We appreciate everyone using data from the FlyCircuit. Nevertheless, we suggest that the following statement be used: “Images from FlyCircuit were obtained from the The National Center for High-performance Computing, Hsinchu, Taiwan”.

Java [Download] is needed for interactive 3D visualization. Your computer may already have one.

For the novice, we recommend to start from “Gallery” for application examples and “Help/Tutor” for step-by-step demonstrations.

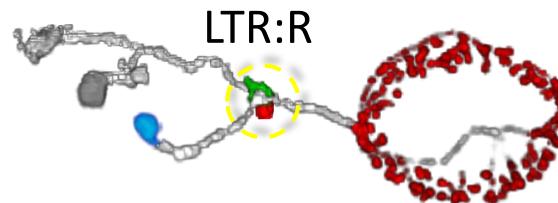
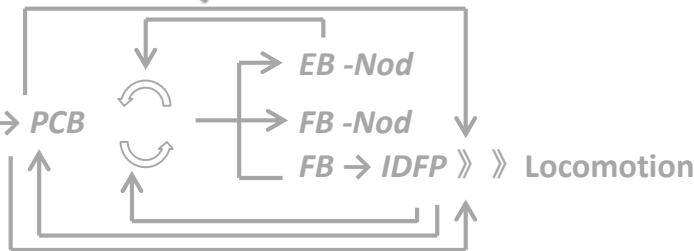
▲ TOP

© NCHC & BRC/NTHU, 2009 All Rights Reserved

NCHC BRC/NTHU

An UV Circuit from Eye to Pre-motor Center

Med → OPTU → LTR → EB → PCB



OPTU:R:LaterI → LTR:R
 ↓
LTR:R → EB:R:O1~8/L:O1~8

FlyCircuit
 A Database of Drosophila Brain Neurons

Login
 Taiwan

Introduction Archive Browsing Search Analysis SnapShot Gallery Help

Welcome to FlyCircuit Database

FlyCircuit is a public database for online archiving, cell type inventory, browsing, searching, analysis and 3D visualization of individual neurons in the *Drosophila* brain. For more details, please read the [associated manuscript](#) – “Three-dimensional reconstruction of brain-wide sensory neurons in *Drosophila* using optical sectioning and image processing” (Bhandawat et al., 2010, *Dev Biol* 340(2): 2010–2019). All images are copyright protected.

Acknowledging the NCNC
 We appreciate everyone using data from the FlyCircuit. Nevertheless, we suggest that the following statement be used: “Images from FlyCircuit were obtained from the The National Center for High-performance Computing, Hsinchu, Taiwan”.

Java ([Download](#)) is needed for interactive 3D visualization. Your computer may already have one.

For the novice, we recommend to start from “Gallery” for application examples and “Help/Tutor” for step-by-step demonstrations.

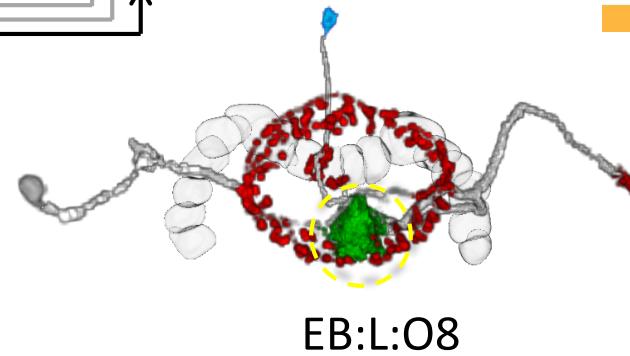
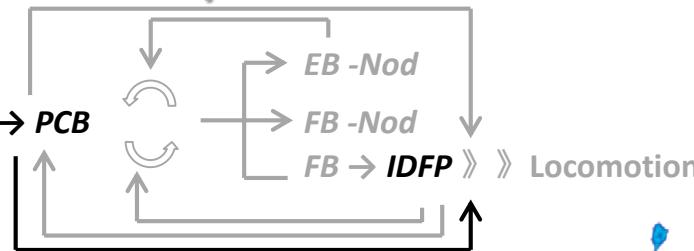
▲ TOP

© NCNC & BRC/NTHU, 2009 All Rights Reserved

NCNC BRC/NTHU

An UV Circuit from Eye to Pre-motor Center

Med → OPTU → LTR → EB → PCB



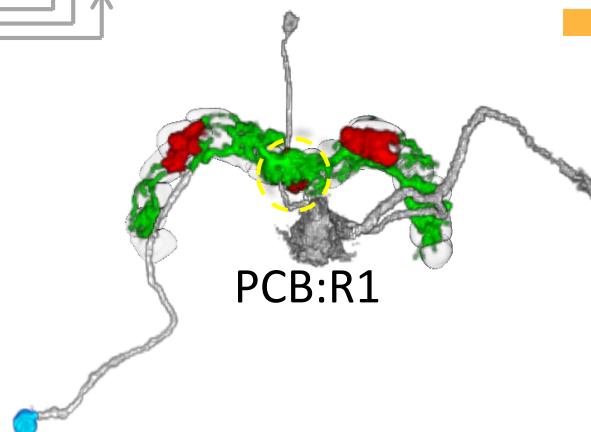
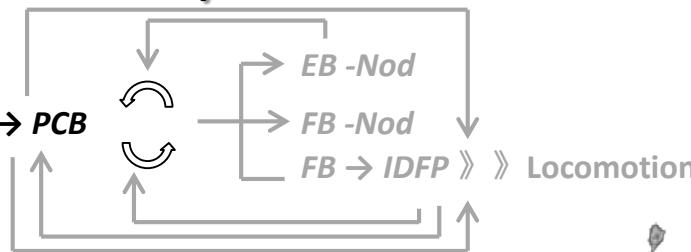
$\text{LTR:R} \rightarrow \text{EB:R:O1~8/L:O1~8}$

↓

Class II $\text{PCB:R1} \leftarrow / \text{EB:L:C8,P8,A8,O8} \rightarrow \text{IDFP:L:DSB}$

An UV Circuit from Eye to Pre-motor Center

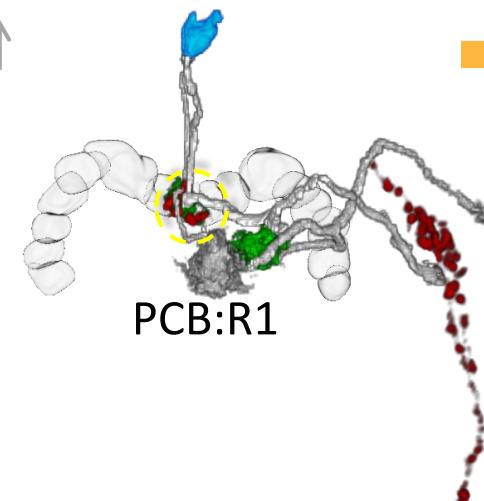
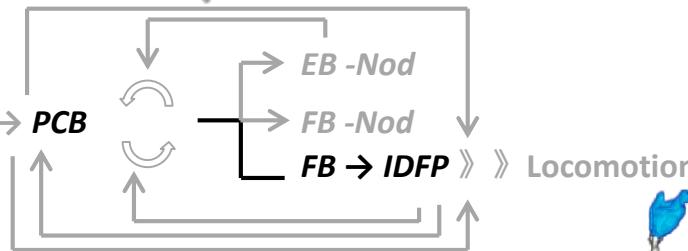
Med → OPTU → LTR → EB → PCB



Class II **PCB:R1** \leftarrow /EB:L:C8,P8,A8,O8 \rightarrow IDFP:L:DSB
 ↓
 Class I **PCB:R1~3,5~8/L1,2,4~7** \rightarrow **PCB:R4/L3**

An UV Circuit from Eye to Pre-motor Center

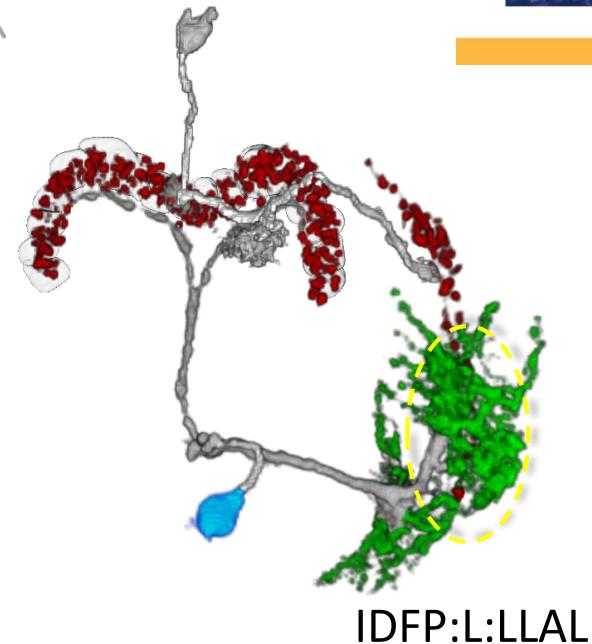
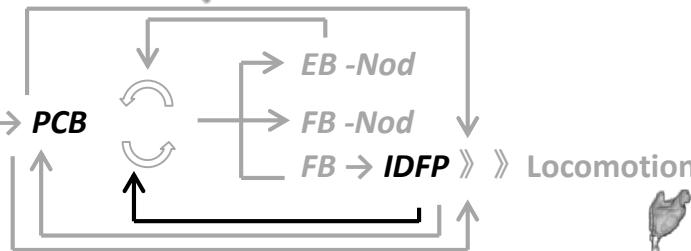
Med → OPTU → LTR → EB → PCB



Class II **PCB:R1←/EB:L:C8,P8,A8,O8→IDFP:L:DSB**
 ↓
 Class VI **PCB:R1-/FB:A5→IDFP:L:LLAL**

An UV Circuit from Eye to Pre-motor Center

Med → OPTU → LTR → EB → PCB

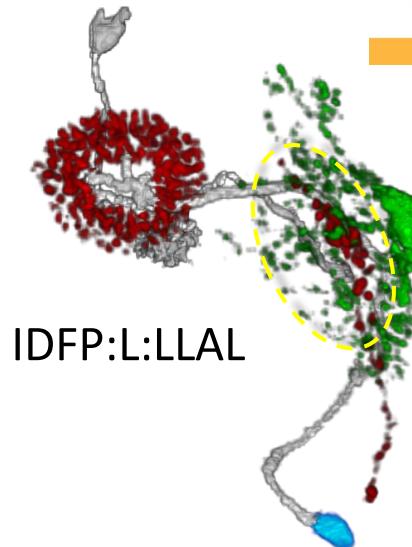
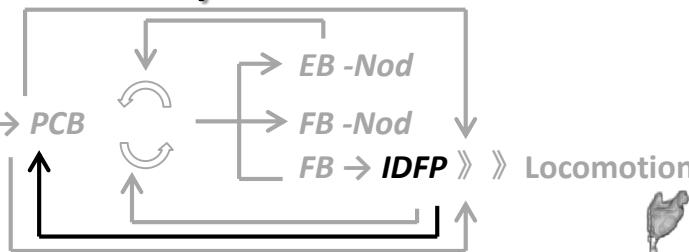


Class VI $\text{PCB:R1-/FB:A5} \rightarrow \text{IDFP:L:LLAL}$

↓
Class III $\text{CVLP:L-VMP:L-IDFP:L:LLAL} \rightarrow \text{PCB:R1~8/L1~8}$

An UV Circuit from Eye to Pre-motor Center

Med → OPTU → LTR → EB → PCB



IDFP:L:LLAL

Class VI $\text{PCB:R1-/FB:A5} \rightarrow \text{IDFP:L:LLAL}$

\downarrow
 $\text{IDFP:L:LLAL, MLAL, DSB, VSB} \rightarrow \text{EB:R:P1}^{\sim 8}/\text{L:P1}^{\sim 8}$

FlyCircuit
A Database of Drosophila Brain Neurons

Login
Taiwan

Introduction Archive Browsing Search Analysis SnapShot Gallery Help

Welcome to FlyCircuit Database

FlyCircuit is a public database for online archiving, cell type inventory, browsing, searching, analysis and 3D visualization of individual neurons in the *Drosophila* brain. For more details, please read the associated manuscript – “Three-dimensional reconstruction of brain-wide sensory neurons in *Drosophila*”, *Nature Protocols*, Volume 2010, Issue 10 (2010) cub 2010:11 (66). Registration is only necessary for users who want to upload/download data or activate 3D viewer. All images are copyright protected.

Acknowledging the NCHC
We appreciate everyone using data from the FlyCircuit. Nevertheless, we suggest that the following statement be used: “Images from FlyCircuit were obtained from the The National Center for High-performance Computing, Hsinchu, Taiwan”.

Java (Download) is needed for interactive 3D visualization. Your computer may already have one.

For the novice, we recommend to start from “Gallery” for application examples and “Help/Tutor” for step-by-step demonstrations.

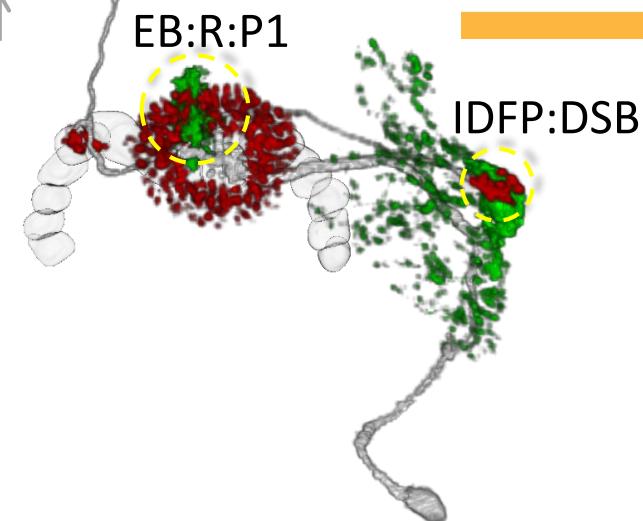
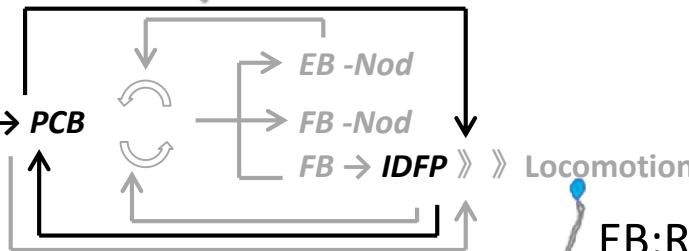
TOP

© NCHC & BRC/NTHU, 2009 All Rights Reserved

NCHC

An UV Circuit from Eye to Pre-motor Center

Med → OPTU → LTR → EB → PCB

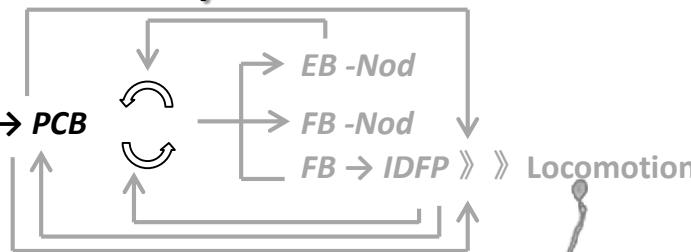


IDFP:L:LLAL,MLAL,DSB,VSB → EB:R:P1~8/L:P1~8

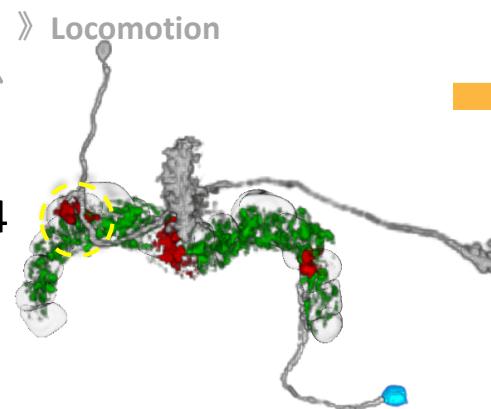
Class II PCB:R4 ← EB:R:C1,P1,A1,O1 → / IDFP:L:DSB

An UV Circuit from Eye to Pre-motor Center

Med → OPTU → LTR → EB → PCB



PCB:R4



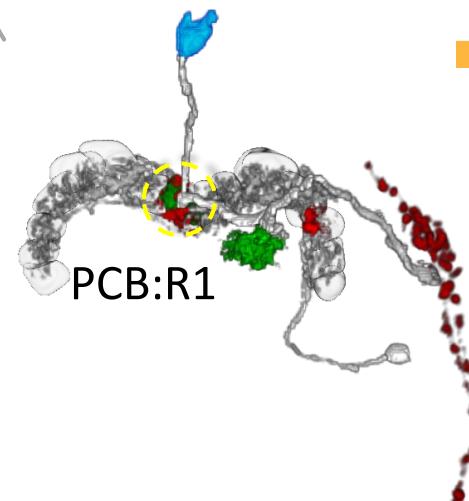
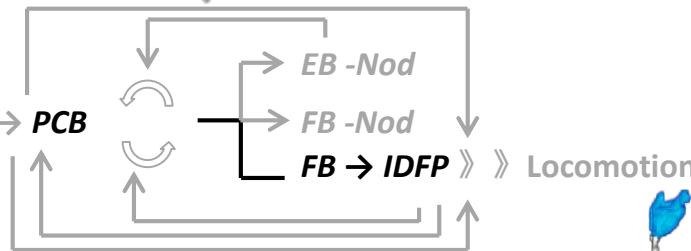
Class II **PCB:R4 ← EB:R:C1,P1,A1,O1 → /IDFP:L:DSB**

↓

Class III **PCB:R2~7/L1~5,7,8 → PCB:R1/L6**

An UV Circuit from Eye to Pre-motor Center

Med → *OPTU* → *LTR* → *EB* → *PCB*

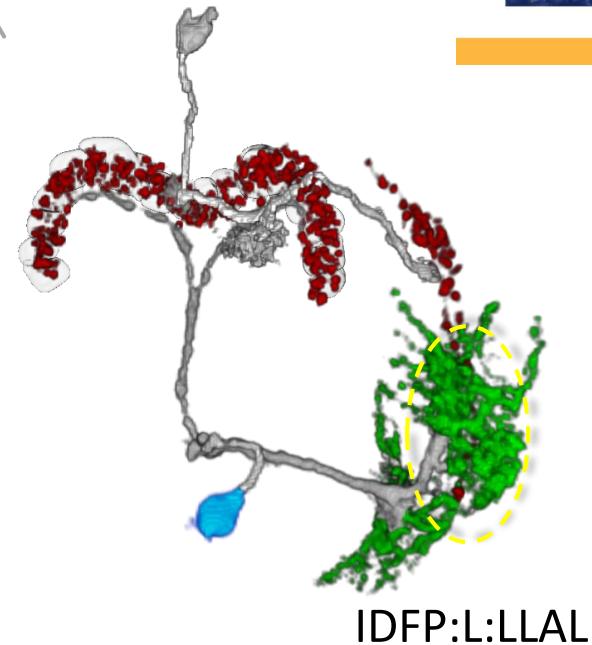
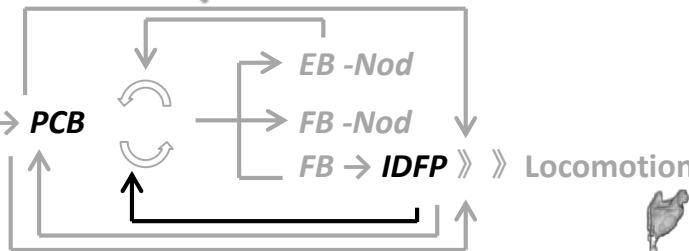


The screenshot shows the FlyCircuit database homepage. At the top, there is a logo of a fly head with neurons and a 'Login' button. Below the logo, a navigation bar includes 'Introduction', 'Archive', 'Browsing', 'Search', 'Analysis', 'SnapShot', 'Gallery', and 'Help'. A welcome message 'Welcome to FlyCircuit Database' is displayed. The main content area contains a brief description of the database, copyright information, and links to Java download and help files. At the bottom, there is a map of the world and logos for NCHC and BRC/Taiwan.

Class II **PCB:R2~7/L1~5,7,8 → PCB:R1/L6**
 ↓
 Class III **PCB:R1-/FB:A5 → IDFP:L:LLAL**

An UV Circuit from Eye to Pre-motor Center

Med → *OPTU* → *LTR* → *EB* → *PCB*



Class VI $\text{PCB:R1-}/\text{FB:A5} \rightarrow \text{IDFP:L:LLAL}$



Class III $\text{CVLP:L-VMP:L-IDFP:L:LLAL} \rightarrow \text{PCB:R1~8/L1~8}$

FlyCircuit
A Database of Drosophila Brain Neurons

Login
Taiwan

Introduction Archive Browsing Search Analysis SnapShot Gallery Help

Welcome to FlyCircuit Database

FlyCircuit is a public database for online archiving, cell type inventory, browsing, searching, analysis and 3D visualization of individual neurons in the *Drosophila* brain. For more details, please read the [associated manuscript](#) – “Three-dimensional reconstruction of brain-wide sensory neurons in *Drosophila* using confocal microscopy”, *Nature* 2010, doi:10.1038/nature09166. Registration is only necessary for users who want to upload/download data or activate 3D viewer. All images are copyright protected.

Acknowledging the NCNC
We appreciate everyone using data from the FlyCircuit. Nevertheless, we suggest that the following statement be used: “Images from FlyCircuit were obtained from the The National Center for High-performance Computing, Hsinchu, Taiwan”.

Java ([Download](#)) is needed for interactive 3D visualization. Your computer may already have one.

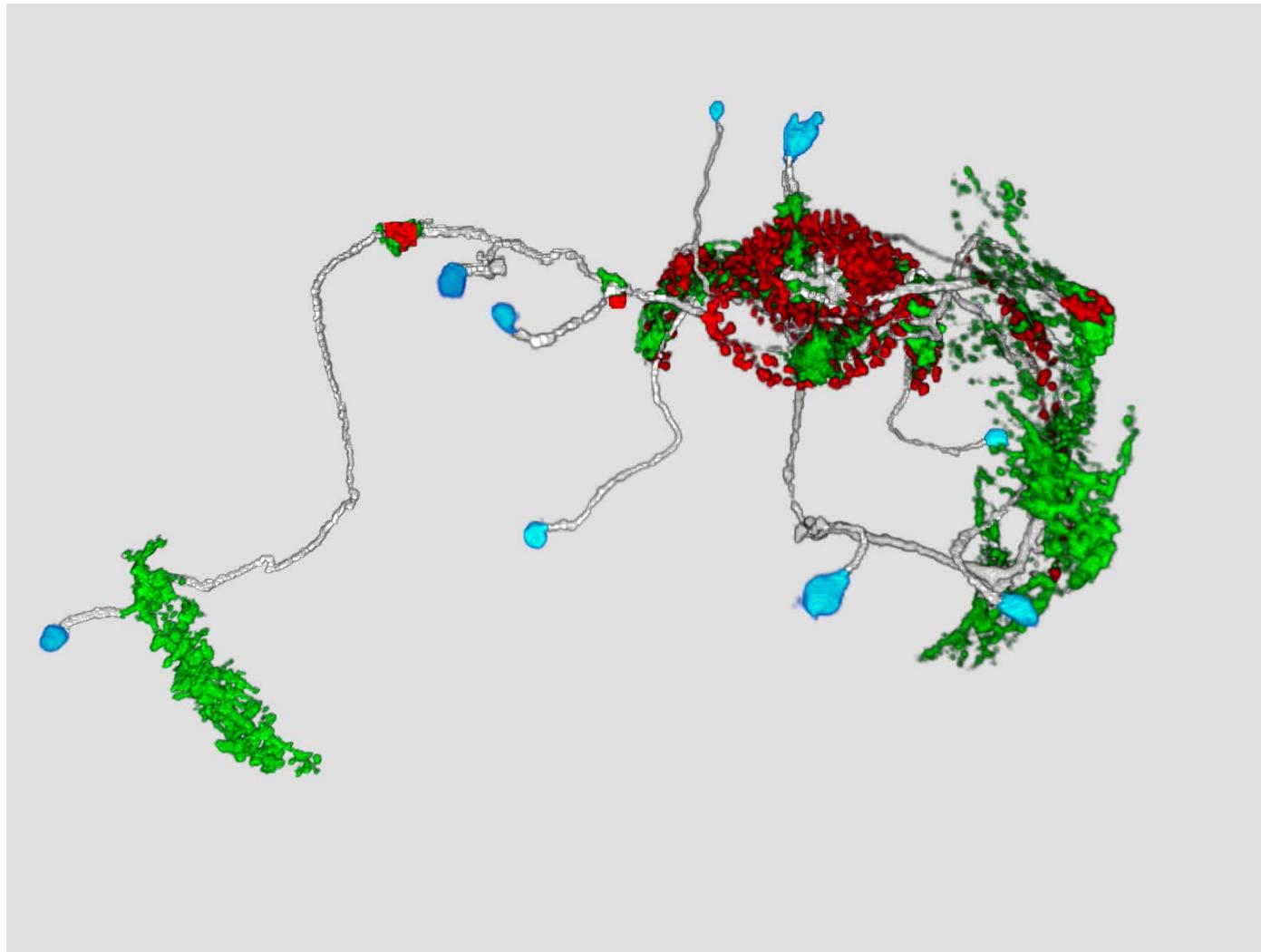
For the novice, we recommend to start from “Gallery” for application examples and “Help/Tutor” for step-by-step demonstrations.

▲ TOP

© NCNC & BRC/NTHU, 2009 All Rights Reserved

NCNC BRC/NTHU

An UV Circuit from Eye to Pre-motor Center



Pathway Movie

FlyCircuit
A Database of Drosophila Brain Neurons

Introduction Archive Browsing Search Analysis SnapShot Gallery Help

Welcome to FlyCircuit Database

FlyCircuit is a public database for online archiving, cell type inventory, browsing, searching, analysis and 3D visualization of individual neurons in the *Drosophila* brain. For more details, please read the associated manuscript – “Three-dimensional reconstruction of brain-wide sensory networks in *Drosophila*”, Berman et al., Current Biology, 2010, doi:10.1016/j.cub.2010.11.050. Registration is only necessary for users who want to upload/download data or activate 3D viewer. All images are copyright protected.

Acknowledging the NCHC
We appreciate everyone using data from the FlyCircuit. Nevertheless, we suggest that the following statement be used: “Images from FlyCircuit were obtained from the The National Center for High-performance Computing, Hsinchu, Taiwan”.

Java [Download] is needed for interactive 3D visualization. Your computer may already have one.

For the novice, we recommend to start from “Gallery” for application examples and “Help/Tutor” for step-by-step demonstrations.

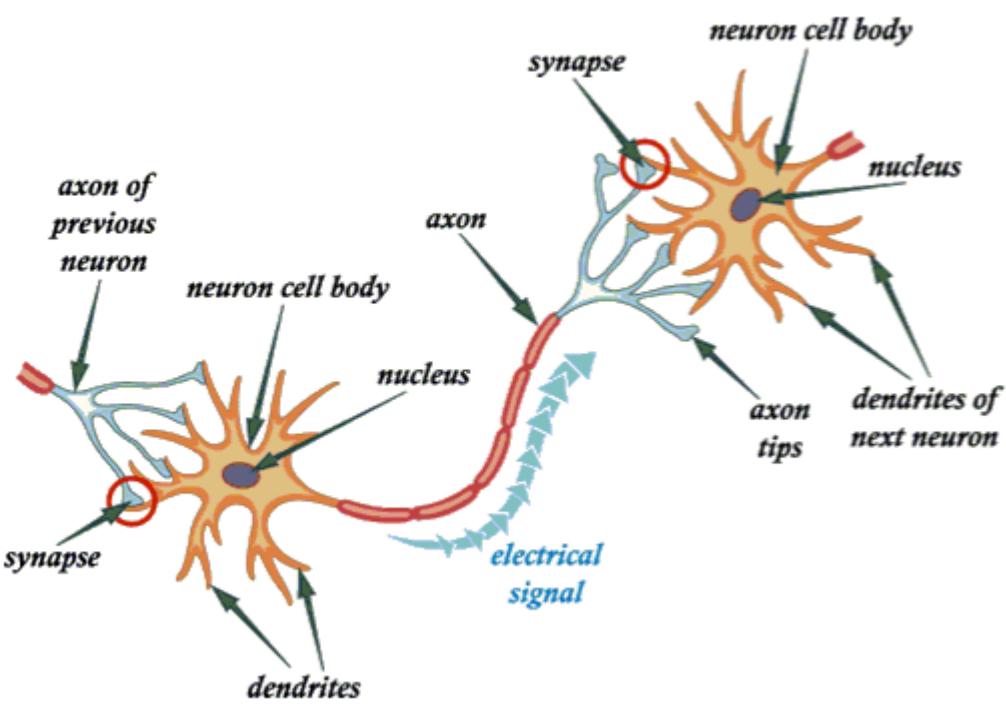
▲ TOP

© NCHC & BRC/NTHU, 2009 All Rights Reserved

NCHC

Cell Report 3,1739 (2013)

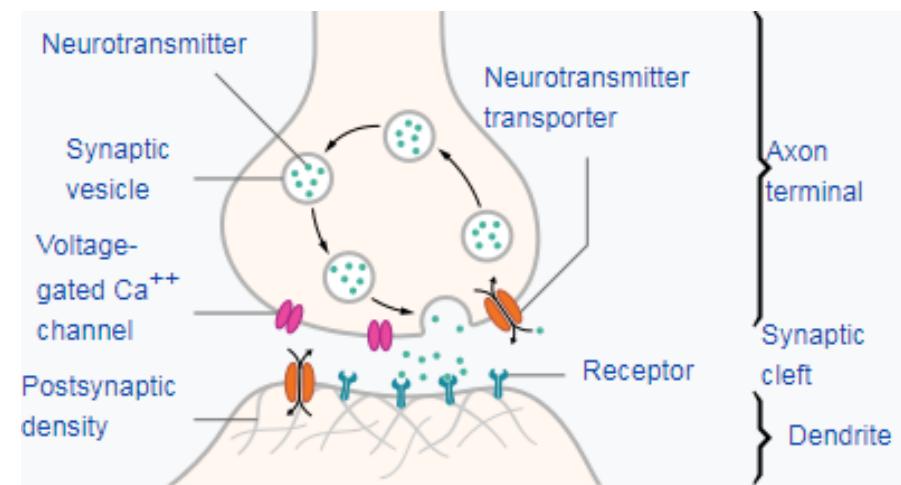
A single neuron



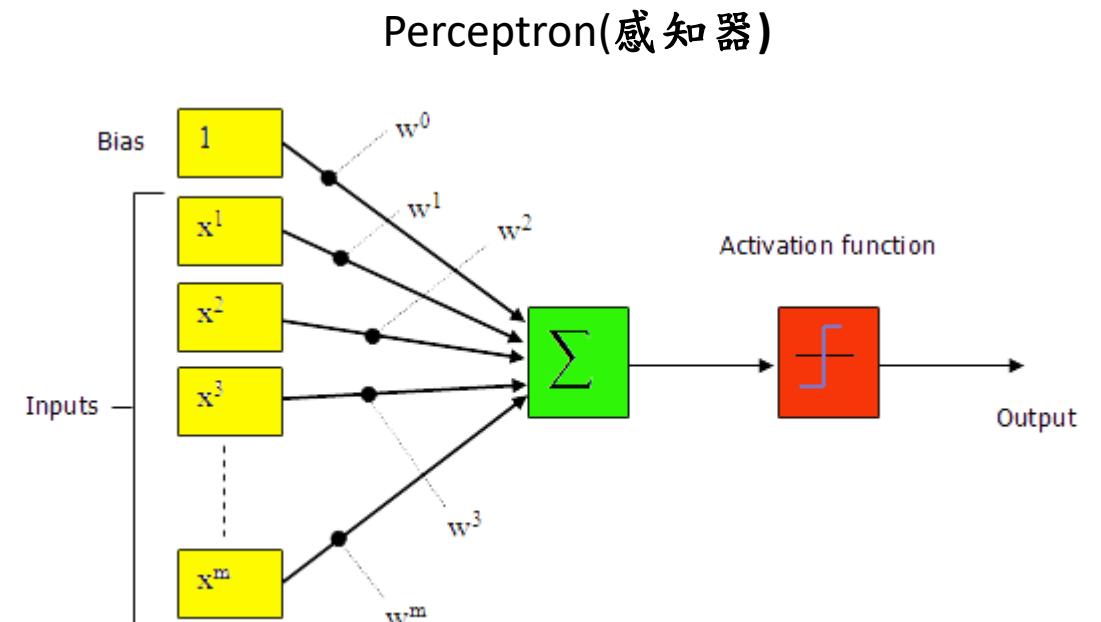
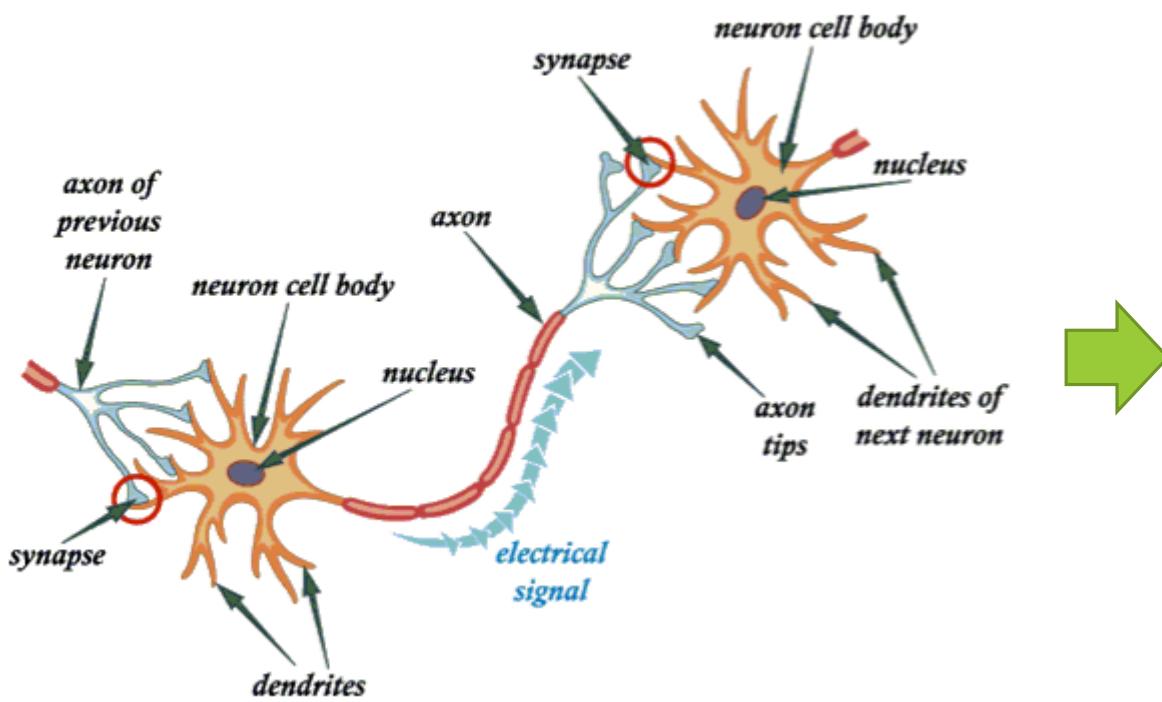
生物神經細胞基本構造：

- 細胞本體(soma)
- 樹突(dendrites)
- 軸突(axon)
- 突觸(synapse)

→ input
→ output

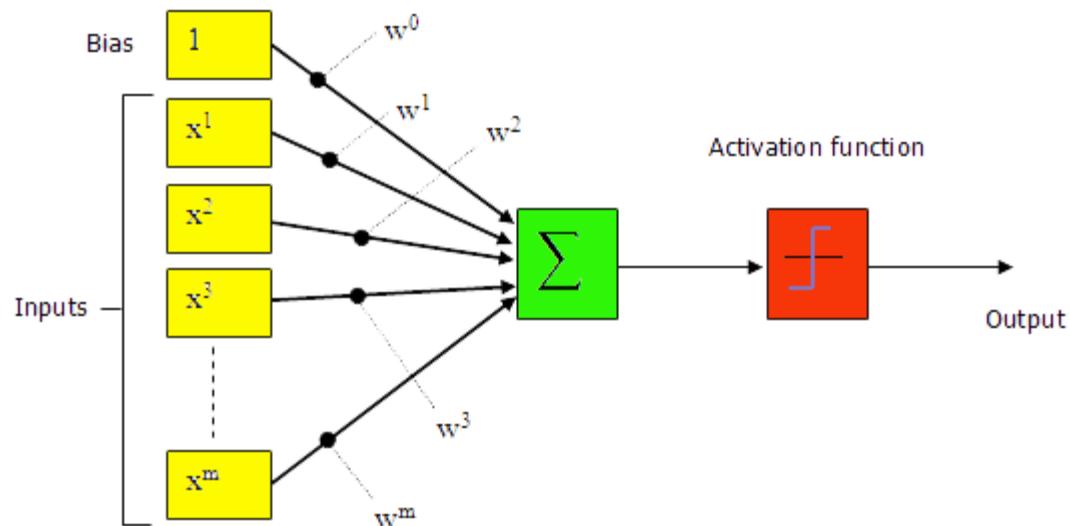


Neuron → Perceptron



A single perceptron

Weights and Bias (權重與偏置)

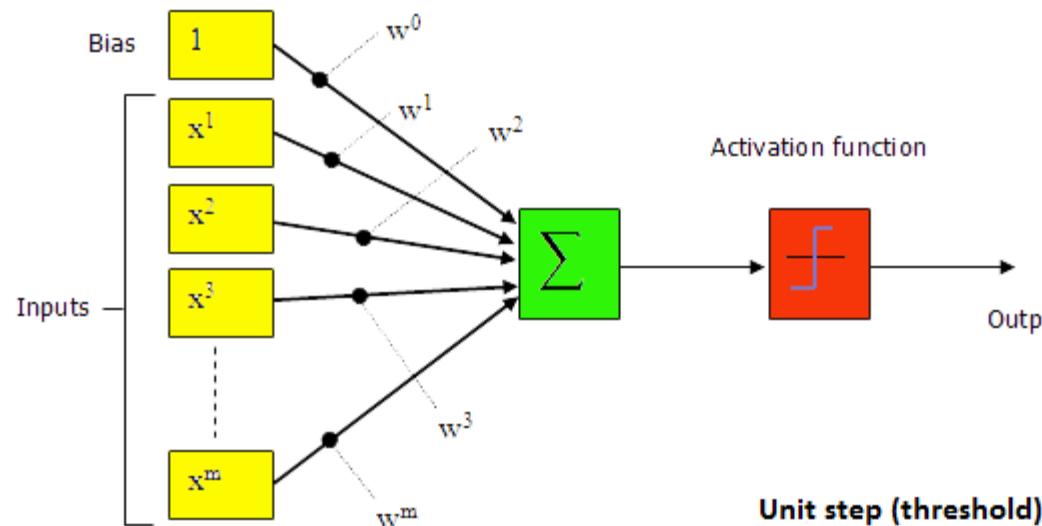


$$Output = f \left(\sum_{i=1}^m w^i x^i + w^0 \right) = f(w^1 x^1 + w^2 x^2 + w^3 x^3 + \dots + w^m x^m + w^0)$$

Learning parameters

A single perceptron

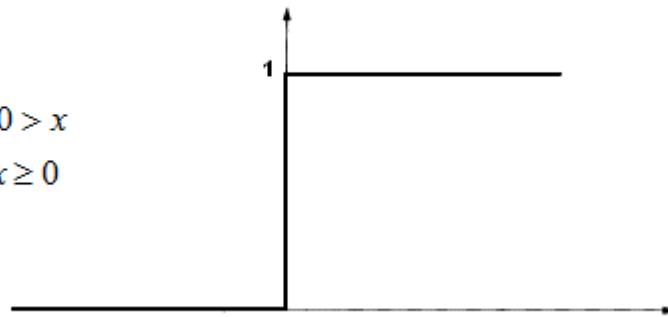
□ Activation Function (激活函數)



$$Output = f\left(\sum_{i=1}^m w^i x^i + w^0\right)$$



$$f(x) = \begin{cases} 0 & \text{if } x > 0 \\ 1 & \text{if } x \geq 0 \end{cases}$$

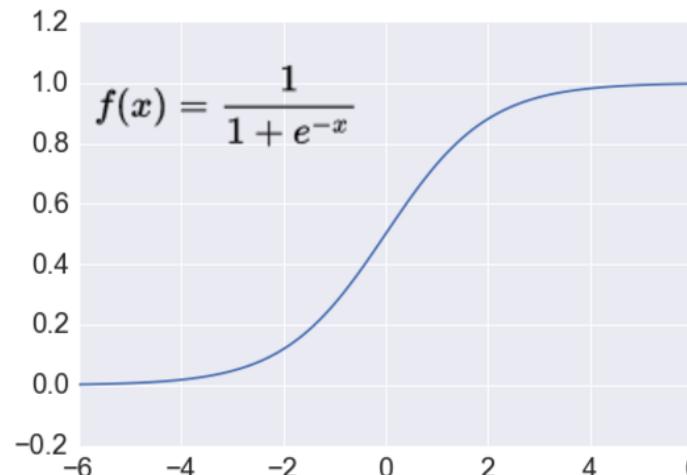


A single perceptron

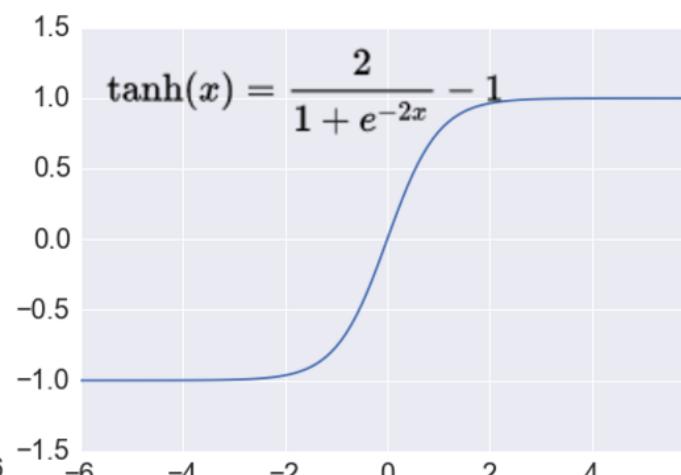
□ Activation Function(激活函數)

$$Output = f\left(\sum_{i=1}^m w^i x^i + w^0\right)$$

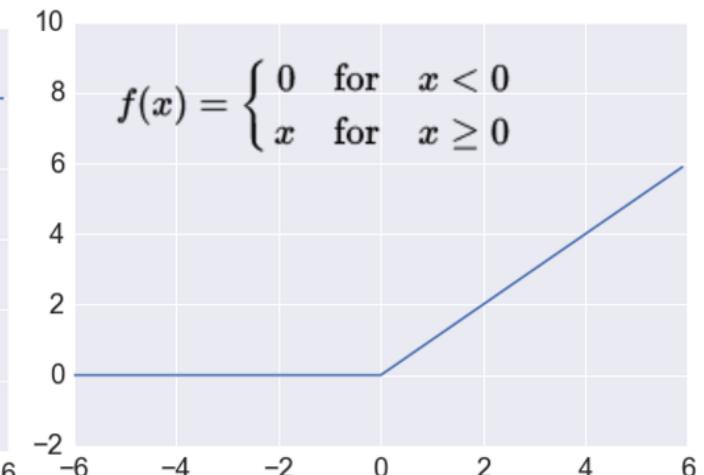
Sigmoid



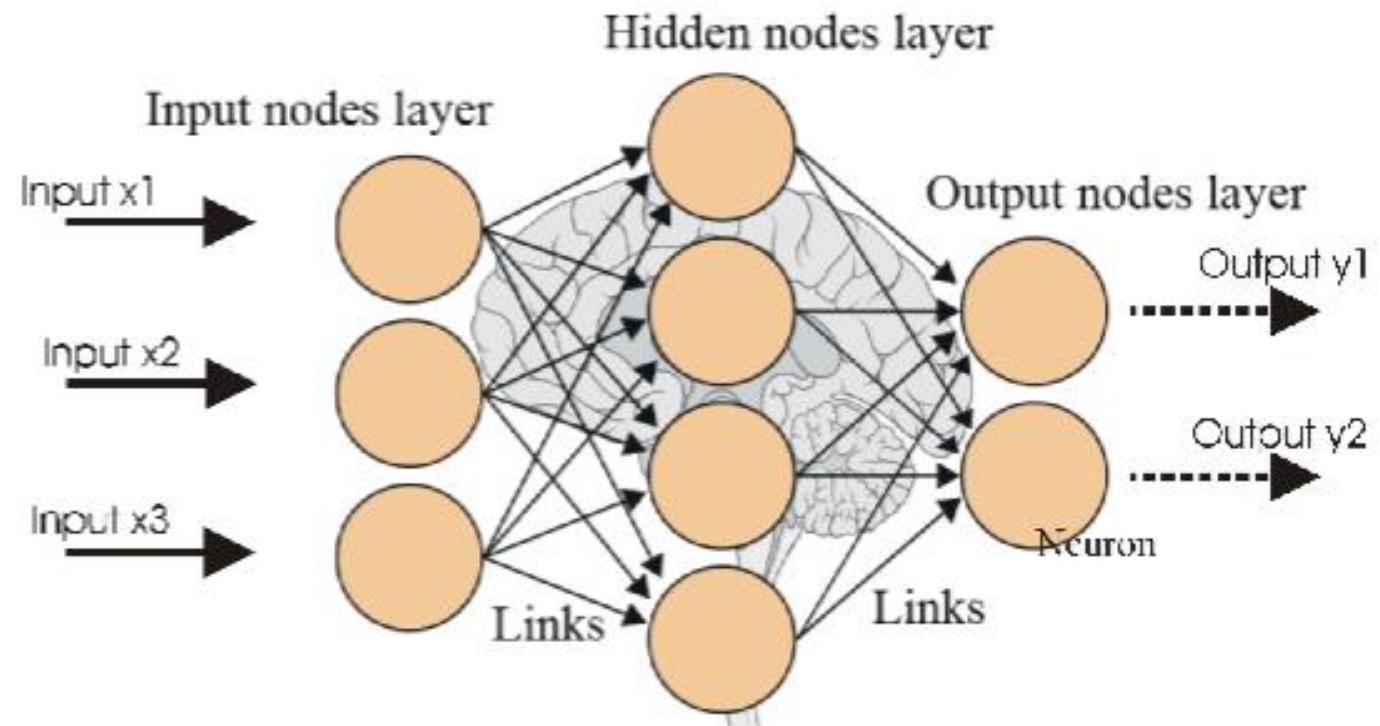
TanH



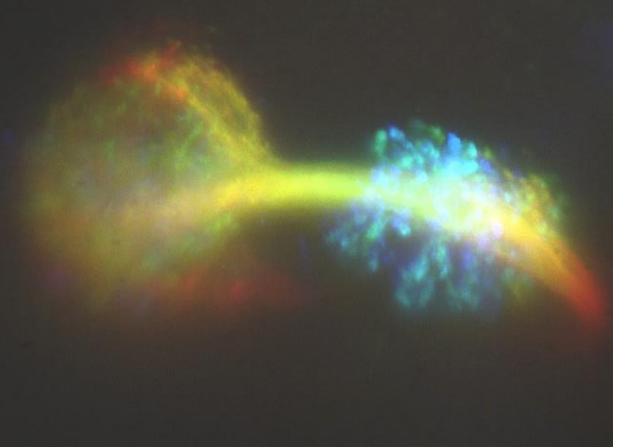
ReLU



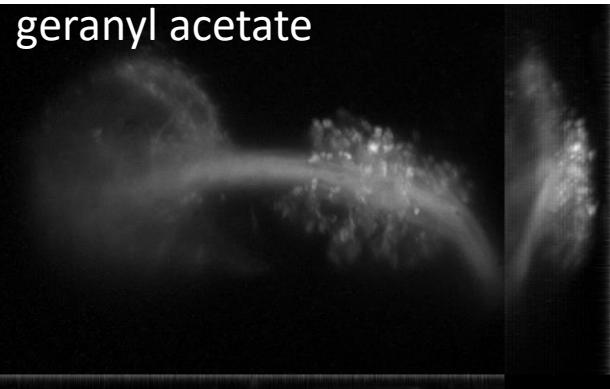
ANN = Multilayer Perceptron



Real neurons'
firing



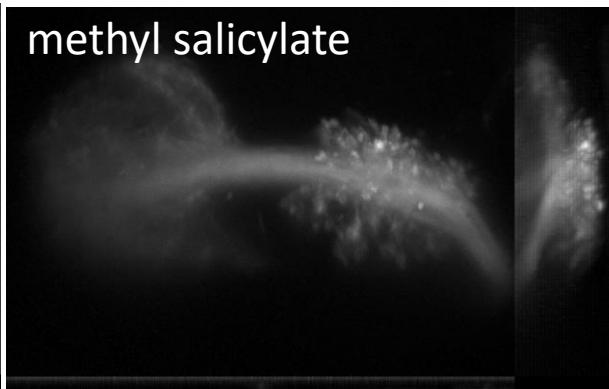
geranyl acetate



4-methylcyclohexanol



methyl salicylate

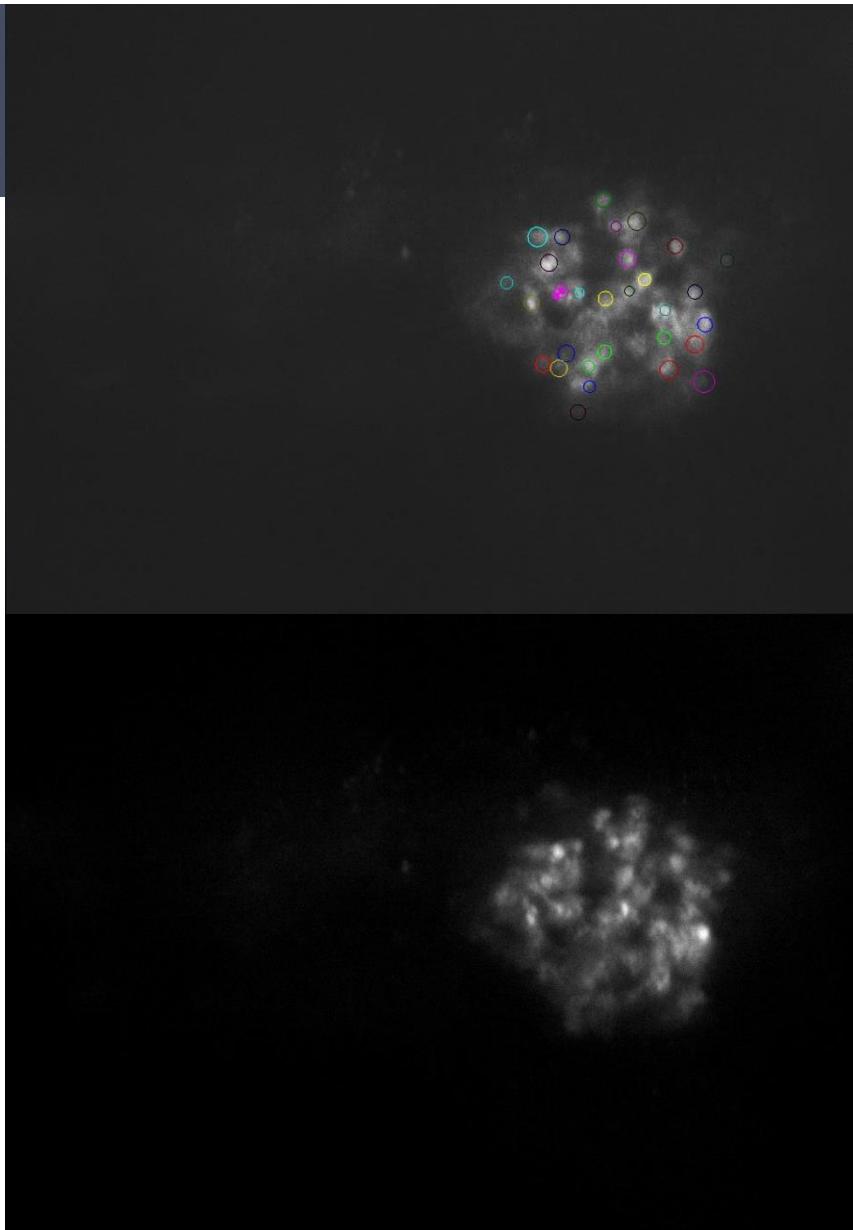
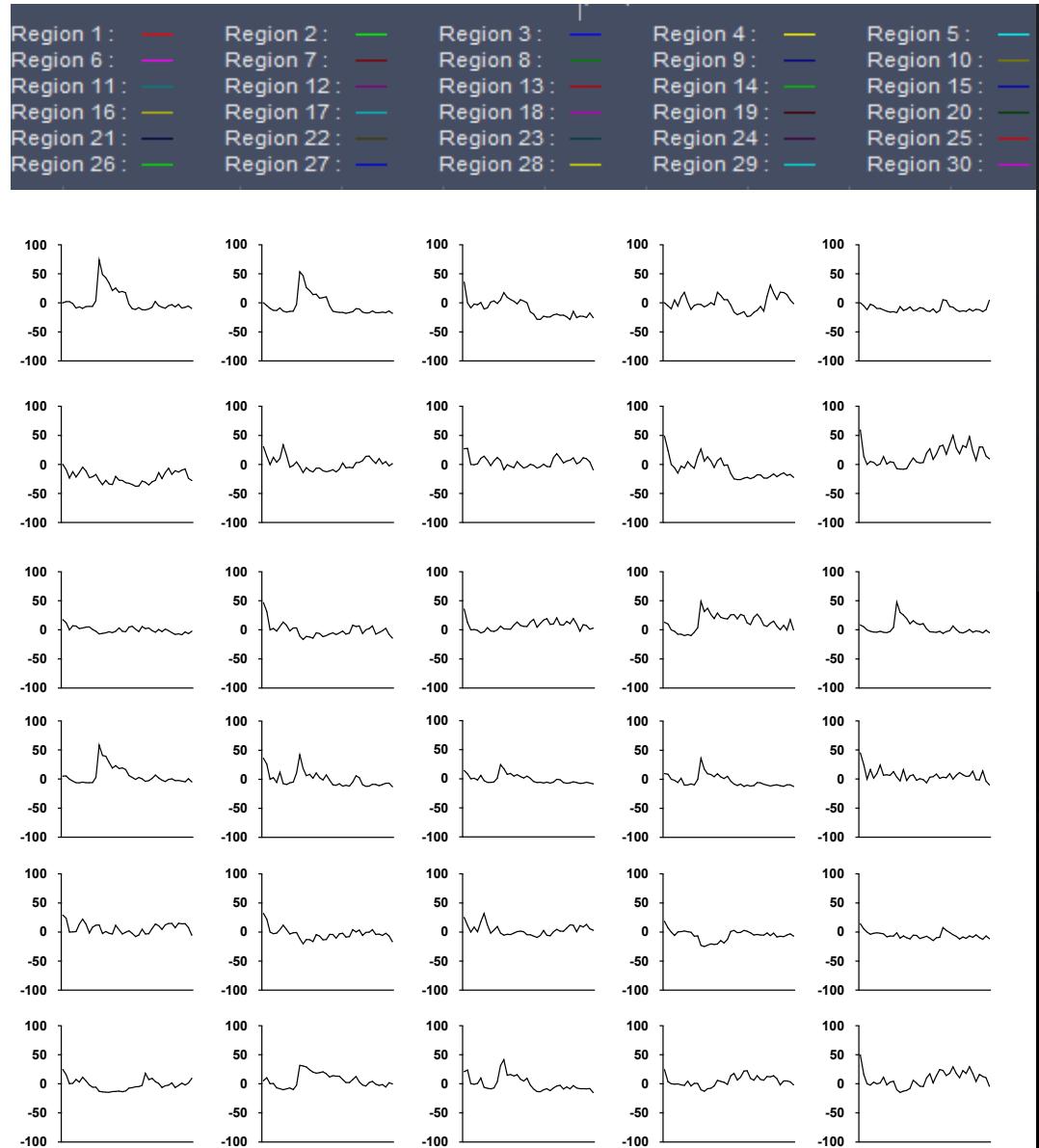


Ethanol

3-octanol

Apple Cider Vinegar

Real activation functions



How ANN learn?

Labeling data
(training set)

$$Input = (x^1, x^2, x^3, \dots, x^m)_j$$

$$Target = (y^1, y^2, y^3, \dots, y^n)_j$$



$$Output = f\left(\left(\sum_{i=1}^m w^i x^i + w^0\right)_j\right) = (\hat{y}^1, \hat{y}^2, \hat{y}^3, \dots, \hat{y}^n)_j$$



$$Loss function = L\left(\sum_j (Target_j - Output_j)\right)$$

**The main goal is finding appropriate weights
and biases to minimize the loss function!!!**

Next class

- ❑ Regression
- ❑ Loss function
- ❑ Logistic regression
- ❑ Forward propagation
- ❑ Backward propagation
- ❑ Gradient descent
- ❑ Optimizers
- ❑

