WASEDA UNIVERSITY

Use Only for JFWM 2025

Image Processing for Mechatronics

Hideyuki Sawada, Ph.D.
Professor

Dept. of Applied Physics
Waseda University
sawada@waseda.jp

Topics in this Lecture

My research related with robotics and image processing techniques

Image processing techniques in industrial applications

Basic programming for image processing

Mechatronics

Mechatronics is a multidisciplinary field of science that includes a combination of mechanical engineering, electronics, computer engineering, telecommunications engineering, systems engineering and control engineering.

https://en.wikipedia.org/wiki/Mechatronics

Approach aiming at the synergistic integration of mechanics, electronics, control theory, and computer science within product design and manufacturing, in order to improve and/or optimize its functionality

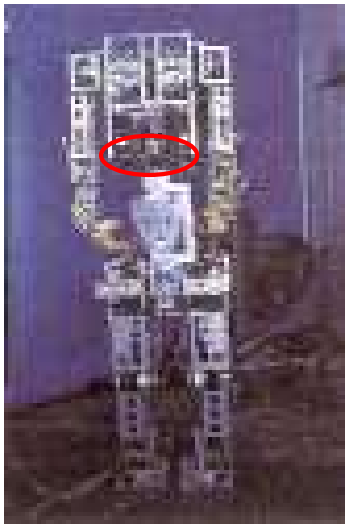
Definition by French standard NF E 01-010

Application fields ….

Machine vision
Expert systems
Industrial goods
Mechatronics systems
Medical mechatronics
Healthcare techs
Human-machine interface
Automation and robotics
Sensing and control systems
Consumer products
Structural dynamic systems
Medical imaging systems
Assistive techs
Human-system interactive systems
Computer-aided design
Computer aided and integrated manufacturing systems
Engineering and manufacturing systems
Automotive engineering, automotive equipment
Autopilot systems
Transportation and vehicular systems
…….

History of Robot Research in Waseda University

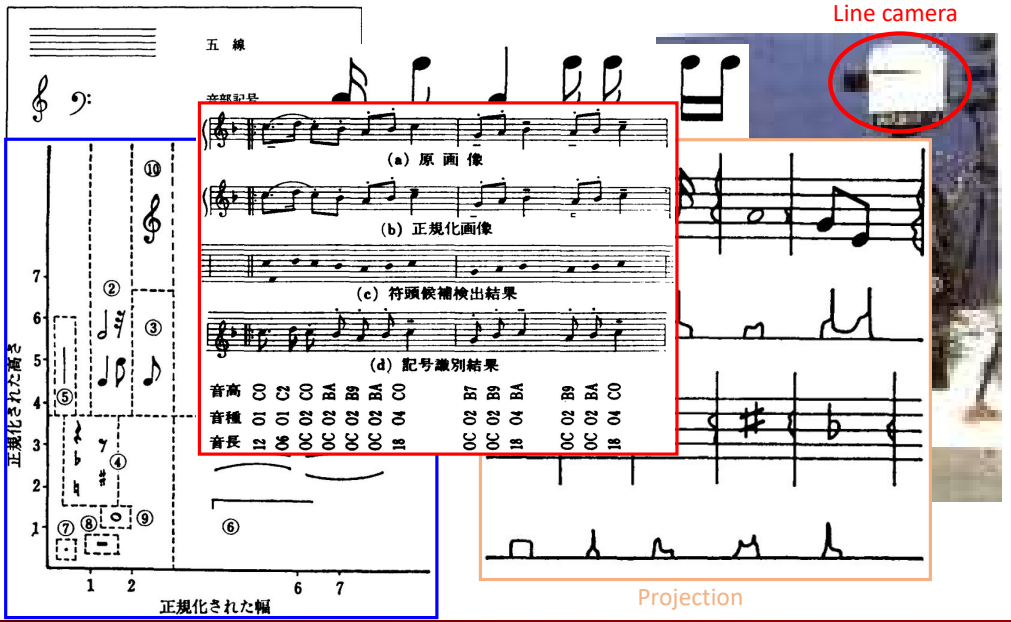
WABOT-1 (1973)
World's first biped walking robot



WABOT-2 (1984)
Keyboard playing robot

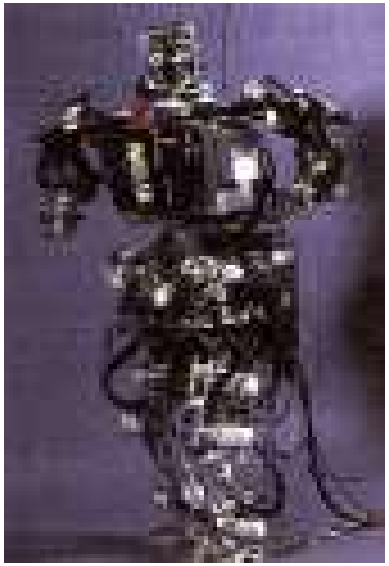
Collaborative Performance
with Real Orchestra
in **Tsukuba Expo 1985**

WABOT-2 (1984)
Automatic Recognition of Music Scores

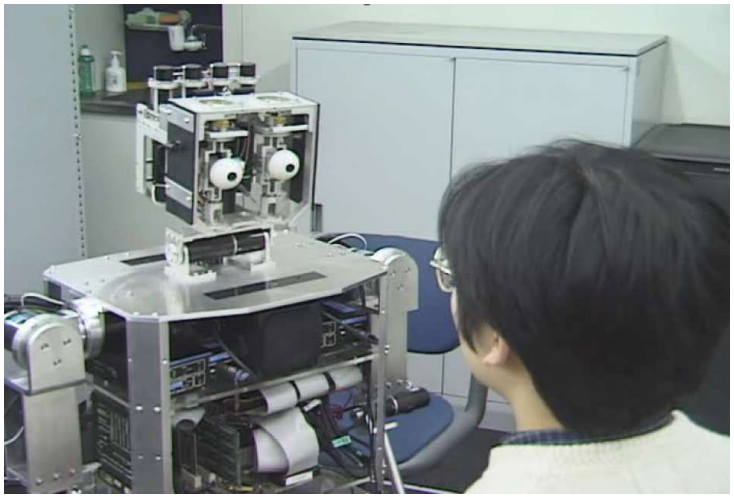


Hadaly-2 (1997)
Communication robot

WABIAN (1997)
Biped walking with Dynamic Balance



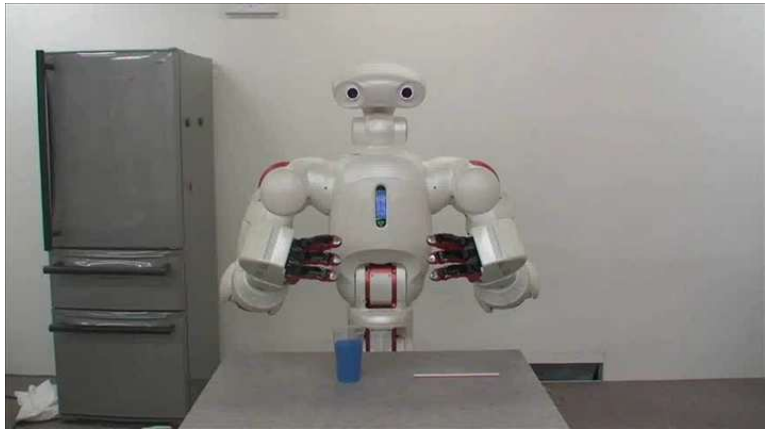
Autonomous humanoid robot
Robita (2000)



WENDY (2000)

TWENDY-ONE

WASEDA UNIVERSITY
Physics, Informatics and Robotics

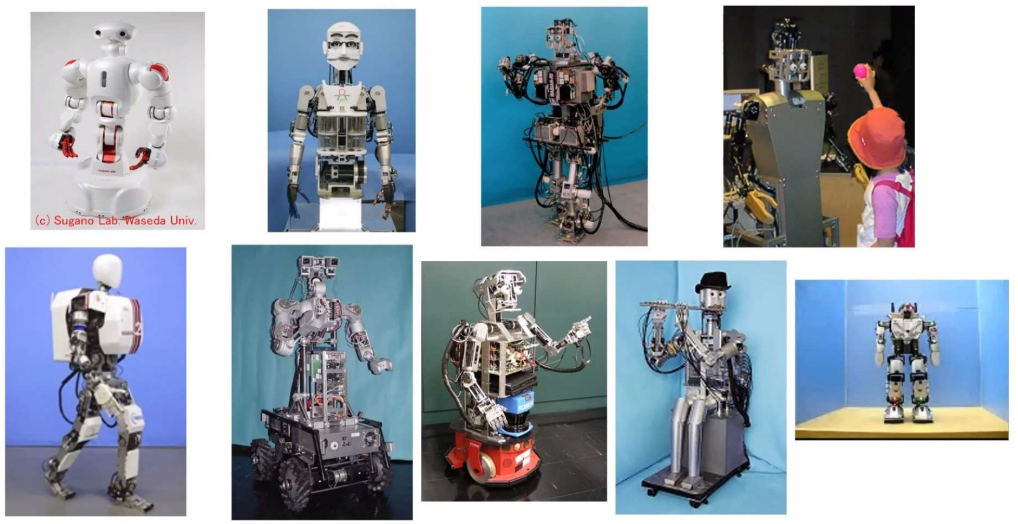


Dexterous handling

Humanoids in Waseda University

WASEDA UNIVERSITY
Physics, Informatics and Robotics

https://www.waseda.jp/inst/fro/



Robots in Society

WASEDA UNIVERSITY
Physics, Informatics and Robotics

Secondary sector of industries: Industrial robots
Machine and Materials

Primary sector of industries: Agricultural robots
Machine and Natural environment

Tertiary sector of industries: Service robots
Machine and Human

Directions of Robotics

WASEDA UNIVERSITY
Physics, Informatics and Robotics

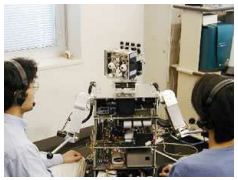
From Factory to Home
Human friendly robots to work with human
in human-living environment

From Superman to Human
Human simulator with sensor complex
to evaluate living environment in society

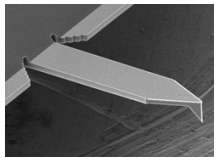

Prof. A. Takanishi

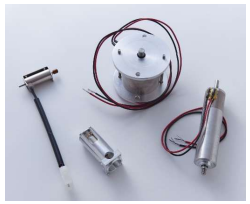
From GUI to PostGUI
Humanoid as multimodal communication terminal
to be connected with network

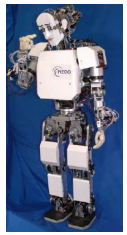
From Virtual to Real
Humanoid as media with body
to bridge virtual space and real world


Prof. T. Kobayashi

"Force" in Different Scale


MEMS cantilever





quark


nucleus


atom/molecule


living creatures


earth


universe

10^{-18} Quantum Physics
10^{-15} Thermo-dynamics
10^{-10} Electro-magnetism
10^{0} Newton mechanics
Mechanism
10^{11} Civil engineering
Earth science
10^{26} m Astro-physics

Are we satisfied with Metaverse?

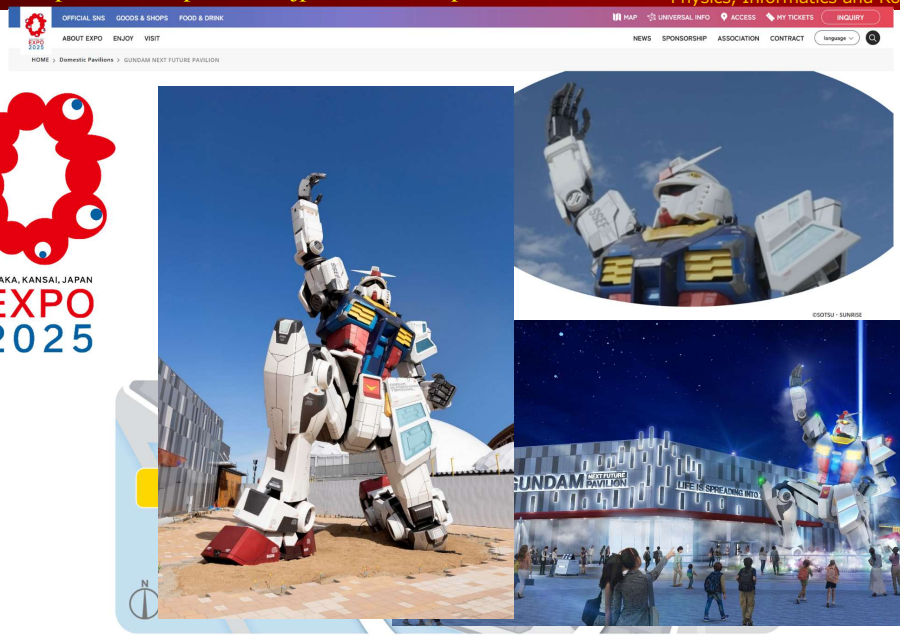
Real size GUNDAM in Yokohama

18 meters height
25 tons weight
34 DoF

GUNDAM had been exhibited
From April 2021 until March 2024

GUMDAM was sent to OSAKA EXPO !

https://www.expo2025.or.jp/en/domestic-pv/bandai-namco/



From Tool to Partner

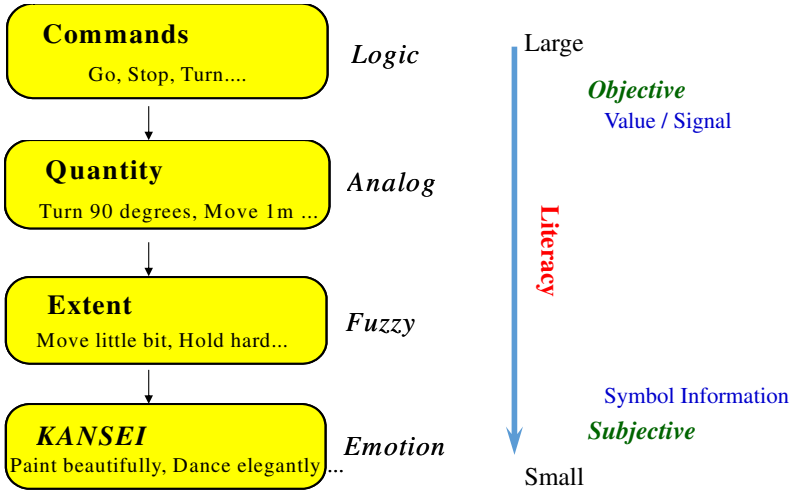
Establishing a new relationship between human and machine

Automation
Plant, Observation, Control, Measurement
"Physics"

Robot in Factory
Production, Planning, Learning, Adaptation
"Physics + Logic"

Robot in Human Environment
Autonomous, Collaborative with Human
"Physics + Logic + KANSEI"
Emotion, Affective, Heartfelt, Feeling, Sympathy, Fun, Intuition ...
Another human ability of understanding

Communication between human and machine



Role of Mechatronics / Robots with Intelligence

Manufacturing – Goods, Materials, ...

Construction – Infrastructure, Environment, ...

Agriculture – Foods, Housing, Gardening, ...

Human interaction – Service

- **Knowledge**
- **Amusement**
- **Arts**
- **Life support**
- **Medical appliance**
- .....

Robotics Technology in Industries

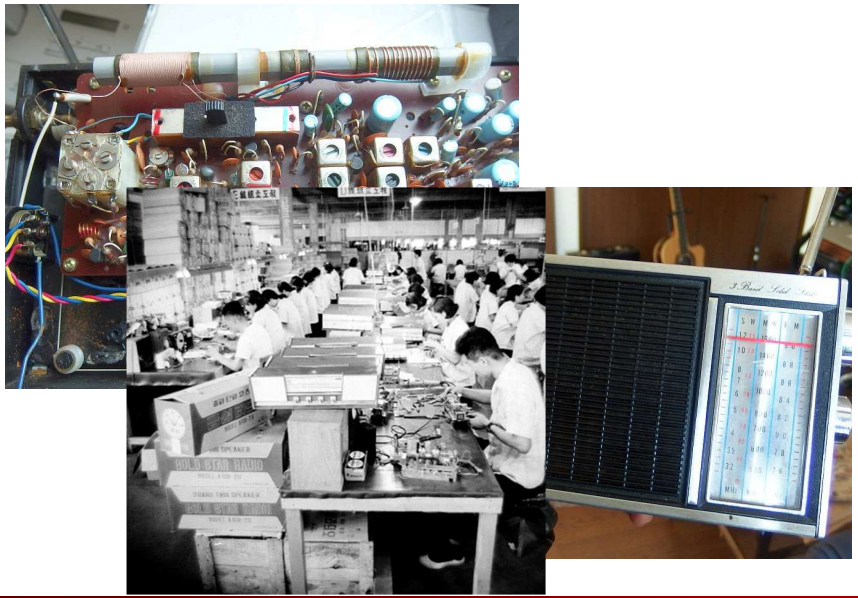
Robot Industry

- **Hardware (Material, Parts, Body, Computer)**
- **Software (OS, Middleware, Application)**
- **I/O, Peripherals (Display, Printer, Communication)**
- **Sensors**
- **Actuators**
- **Design**
- **Assembly**

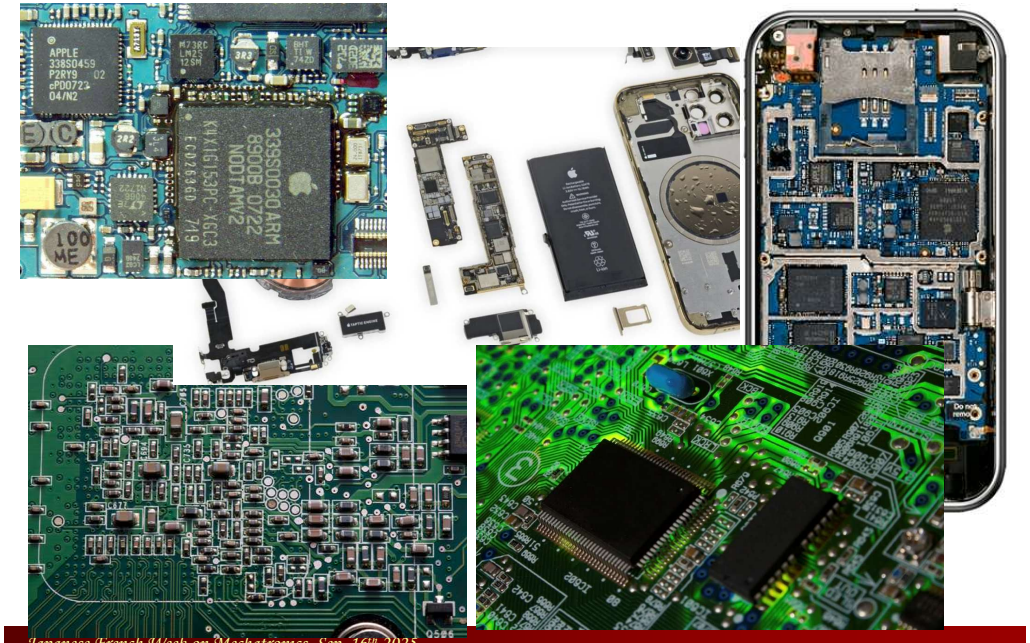
Computer Industry

- **Semiconductors**
- **Hardware (Material, Circuit boards, Electric parts)**
- **Software (OS, Middleware, Applications, AIs)**
- **I/O, Peripherals (Display, Printer, Communication)**

Electronic devices 60 years ago



Recent Electronic Devices



Quality control of manufacturing electronic devices

The only possible approach for such small stuff is to effectively use

Computer vision with human intelligence

Artificial Intelligence: AI

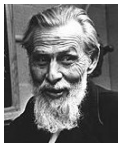
Wikipedia

Artificial intelligence (AI) is intelligence demonstrated by machines, in contrast to the natural intelligence displayed by humans. Colloquially, the term "artificial intelligence" is often used to describe machines (or computers) that mimic "cognitive" functions that humans associate with the human mind, such as "learning" and "problem solving".

Born: 1943 Neuron model
1956 Dartmouth workshop
1958 Perceptron

2nd boom: 1980s Backpropagation algorithm

3rd boom: 2006 - Deep learning



McCulloch & Pitts

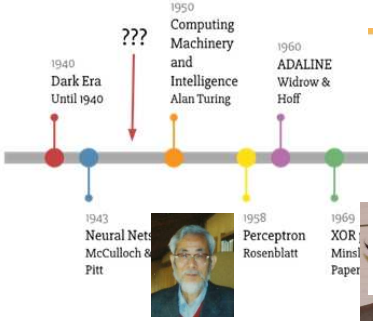

D. Rumelhart


F. Rosenblatt
Wikipedia

Fast computation, Memory, Big Data, and Network
Human-like understanding
Data holders are the Winners

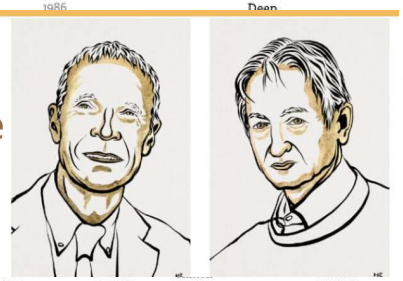
History of Artificial Intelligence

J. v. Neumann, D. O. Hebb, M. Minsky, S. Papert, T. Kohonen, J. Hopfield



Late Prof. Ohteru

Prof. Amari



Williams

Deep Learning Timeline
Made by Favio Vazquez

Information of One Picture

1920x1080 pixels 8bits for each **RGB** (256 levels)

Information given by one picture

$$= 8 \times 1{,}920 \times 1{,}080 \times 3 = 49{,}766{,}400 \text{ bit}$$

$$= 49.77 \text{ Mbit}$$

1 second moving picture with 30 fps

$$49.77 \times 30 = 1{,}492.99 \text{ Mbit} \fallingdotseq 1.5\text{Gbit}$$

Possible variation of 1 second moving pictures

$\rightarrow 2^{1{,}492{,}992{,}000}$ variations !

c.f. Number of atoms in the universe is approx. 10^{80}

Needs for AIs and Machine learning !

Human Brain – comparison with computer

Computer	Human brain
• Billions bytes RAM, and trillions bytes in SSD/HD	• 100 billion neurons and 30 trillion synapses
• Element size: 10^{-9} m	• Neuron size: 10^{-6} m
• Energy: Several tens Watts (CPU)	• Energy use: approx. 25W
• Processing speed: 10^{9} Hz	• Processing speed: 100 Hz
• Centralized processor, Serial com	• Distributed and Parallel proc.
• Generally not Fault Tolerant	• Redundant, Fault Tolerant
• 0 / 1 Logics	• Autonomous Learning
• Intelligence by programming	• Intelligent and Conscious

Human-like Intelligence

Recognition / Understanding of objects, environment as humans do

Realization of ***Human Intelligence*** by using computers

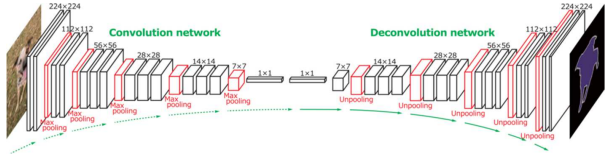
- Analytical ability
- Problem-solving skills
- Quantitative / Qualitative recognition ability
- Subjective / Objective understanding
- Verbal ability
- Ability to learn and adapt to the experiences of everyday life

Realization of Human-like Intelligence

Artificial Neural Network (NN)

Deep Learning (DL)

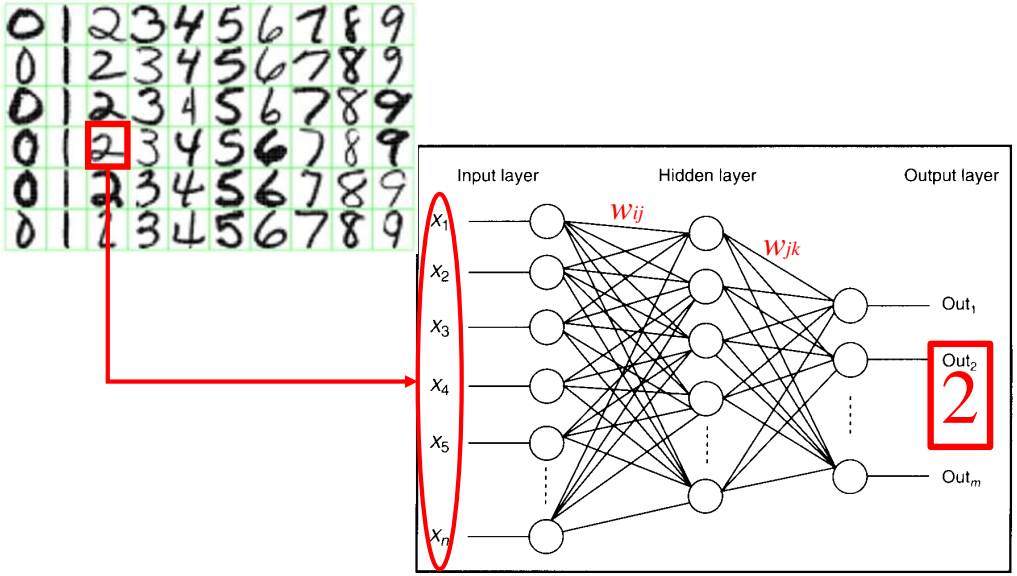
= neural network having
multiple layers of nodes between input and output.



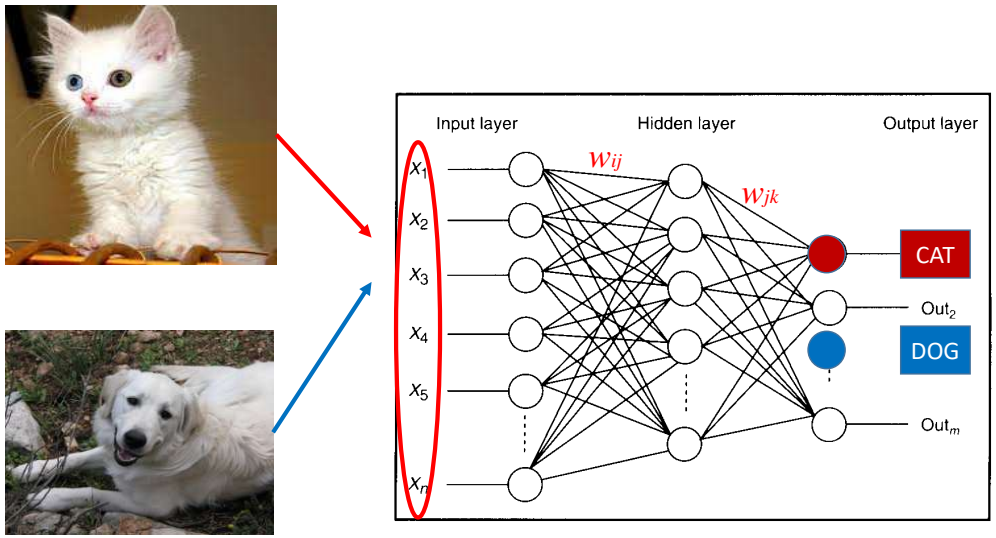
DL has been providing breakthrough results in image classification, speech recognition, pattern recognition …

ChatGPT: Generative AI + LLM

Multiple layer NN



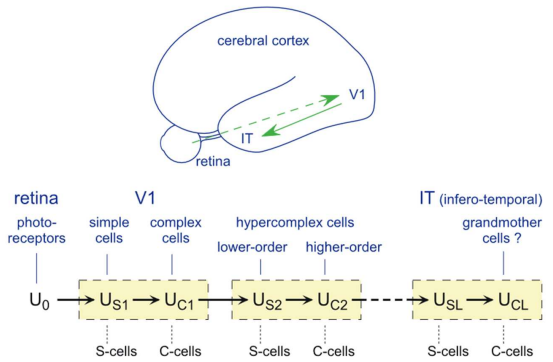
Multiple layer NN



Neocognitron and Deep Learning

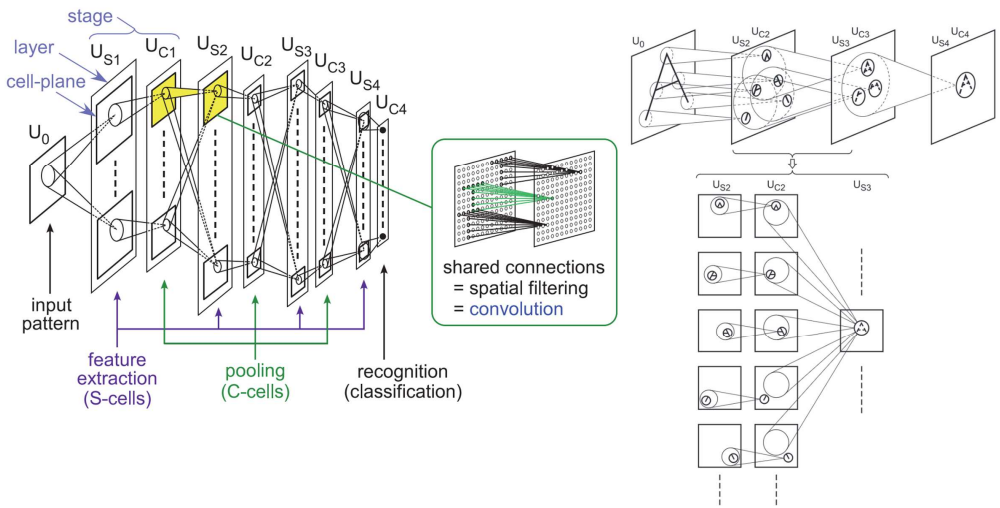
Neocognitron was introduced by Kunihiko Fukushima in 1979.

Multi-layered network with S-cells and C-cells referring to biological visual perception.



S-cells: Feature extraction
C-cells: pooling of the response of S-cells in retinotopic neighborhood

Neocognitron and Deep Learning



Large Language Model (LLM) and Machine Learning

A Model to Estimate
Generation Prob. $P(y_1, y_2, \dots, y_n)$ of Token Sequence $y_1, y_2, \dots, y_n$

Products of Conditional Probabilities:

$$P(y_1, \dots y_n) = P(y_1|BoS)P(y_2|BoS, y_1) \cdots P(y_n|BoS, y_1, \dots, y_{n-1})P(EoS|BoS, \dots, y_n)$$

↑ First word ↑ 2nd word ↑ *n*-th word ↑ End of Sequence

This model predicts the next word from a sequence of text !

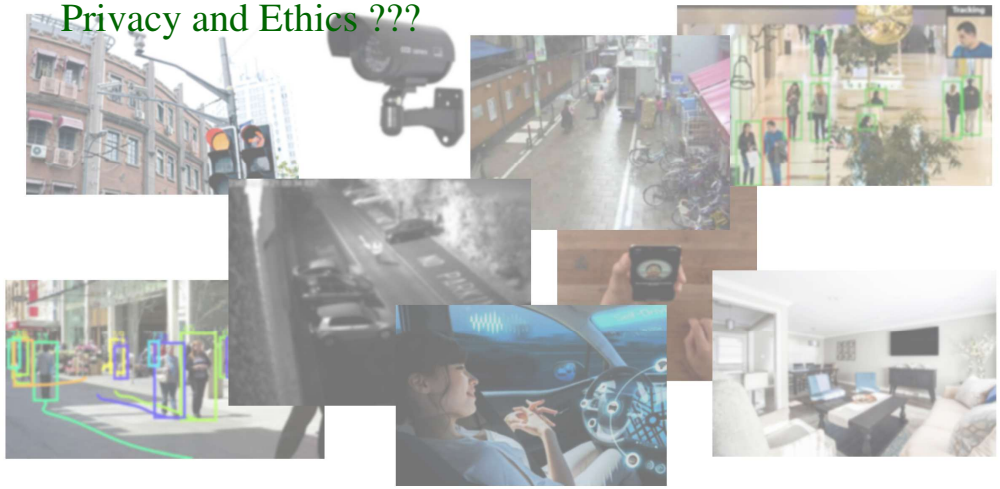
$$\underset{y \in V}{\operatorname{argmax}}\, P(The\ captal, of, France, is, y) = \underset{y \in V}{\operatorname{argmax}}\, P(y|The\ captal, of, France, is)$$

The Prob. of "y = Paris" will be expected to be the highest

ChatGPT has been trained with 4 trillion words with 175 billion parameters.
GPT-4 has been trained with > 10 trillion words !

Applications of AI

Cameras in Street, Buildings and your Mobile Devices
Voice and Privacy Sounds in Smart Devices
Privacy and Ethics ???



Deep Learning for Image Analysis

Surveillance for social safety and security network
Quality Control in Production
Medicine, Health and Healthcare

System and Computational Intelligence Studies

Study of Human ability and behavior
by way of computers

Analysis by Synthesis

Trial and Error

Our Approach for Human Intelligence

Image processing
Acoustic signal processing
Tactile and touch sensation
System and information integration
Robotics
Application to Industrial process

Conclusions

- Realization of Human Visual Ability
- Enhancement of Human Sensing Ability
- Understanding and Control of “Soft” Bodies
- Embodied AI

Alternatives of Human and the Abilities
Supporting Techs for Human
Multimodal interface employing our body

Presented by
Hideyuki Sawada Laboratory
Waseda University