인공지능 기법을 적용한 반도체공정 데이터 분석

성균관대학교

채 희 엽

2020. 11. 6.

제18회 진공기술 실무 수련회

SUNG KYUN KWAN UNIVERSITY

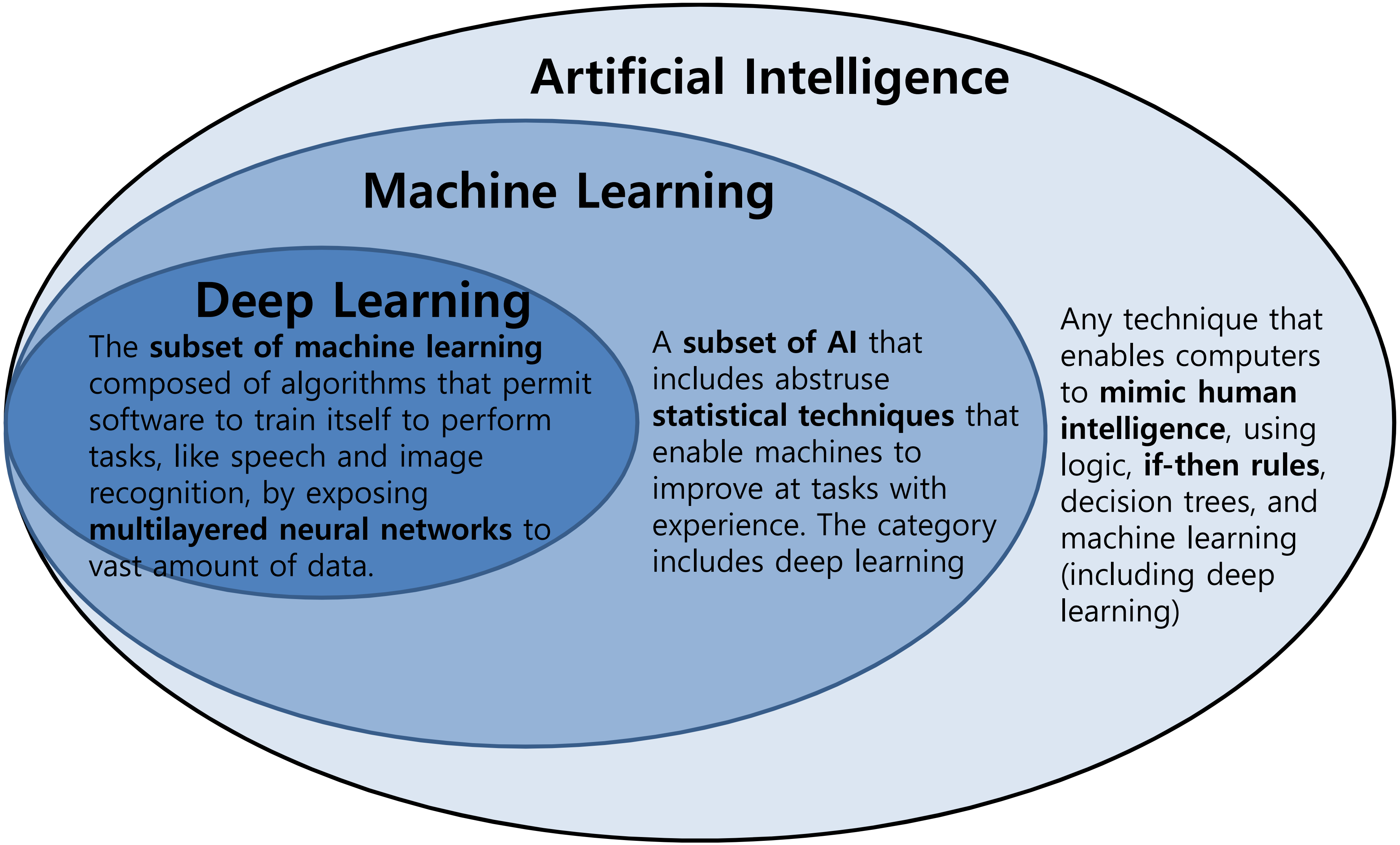
목차

- 인공지능(AI)과 기계학습(Machine Learning)의 분류
- 반도체 공정
- 플라즈마 공정
- 반도체 공정 데이터 분석 예

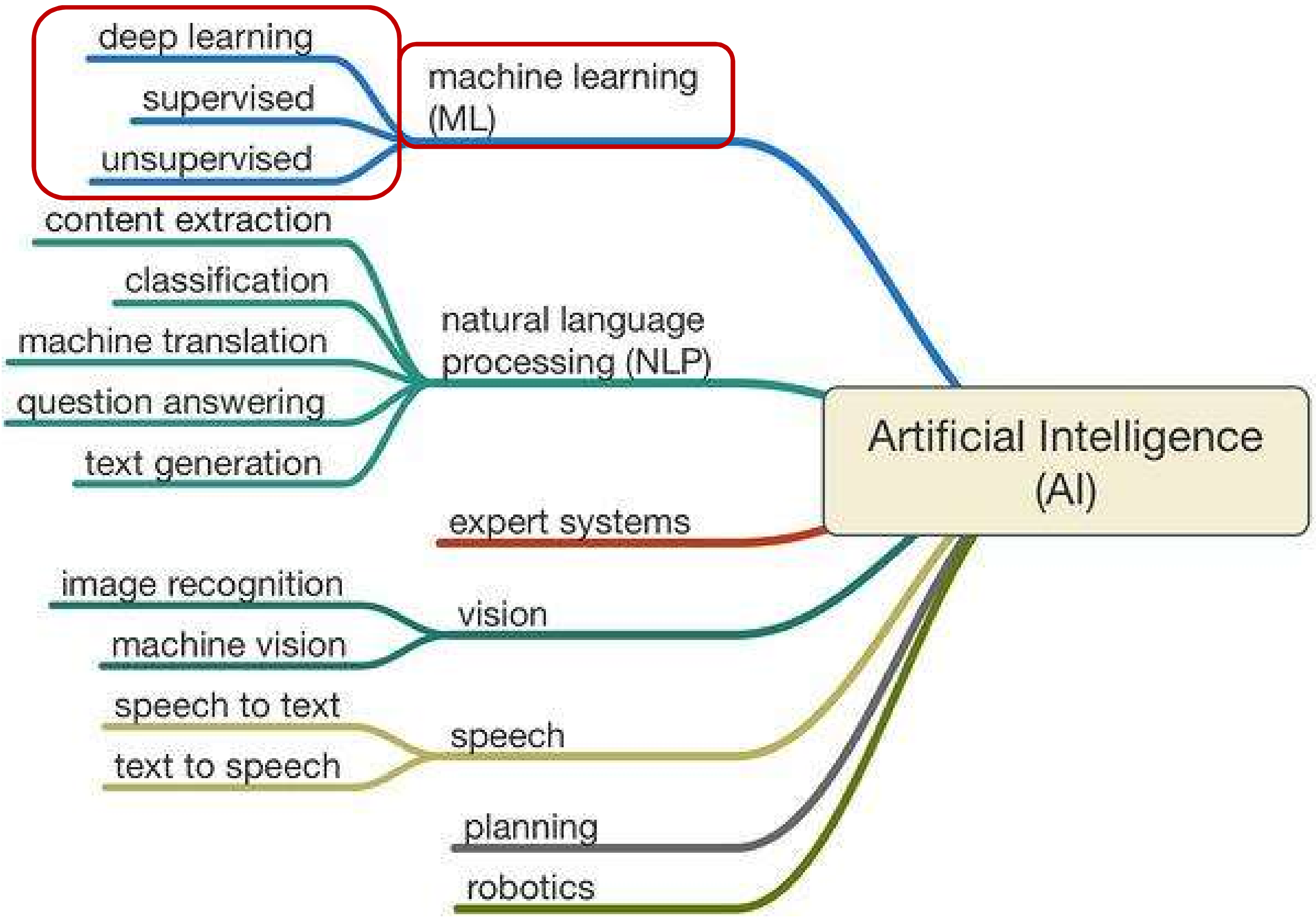
AI/Machine Learning/Deep Learning

SUNG KYUN KWAN UNIVERSITY

Artificial Intelligence vs Machine Learning vs Deep Learning

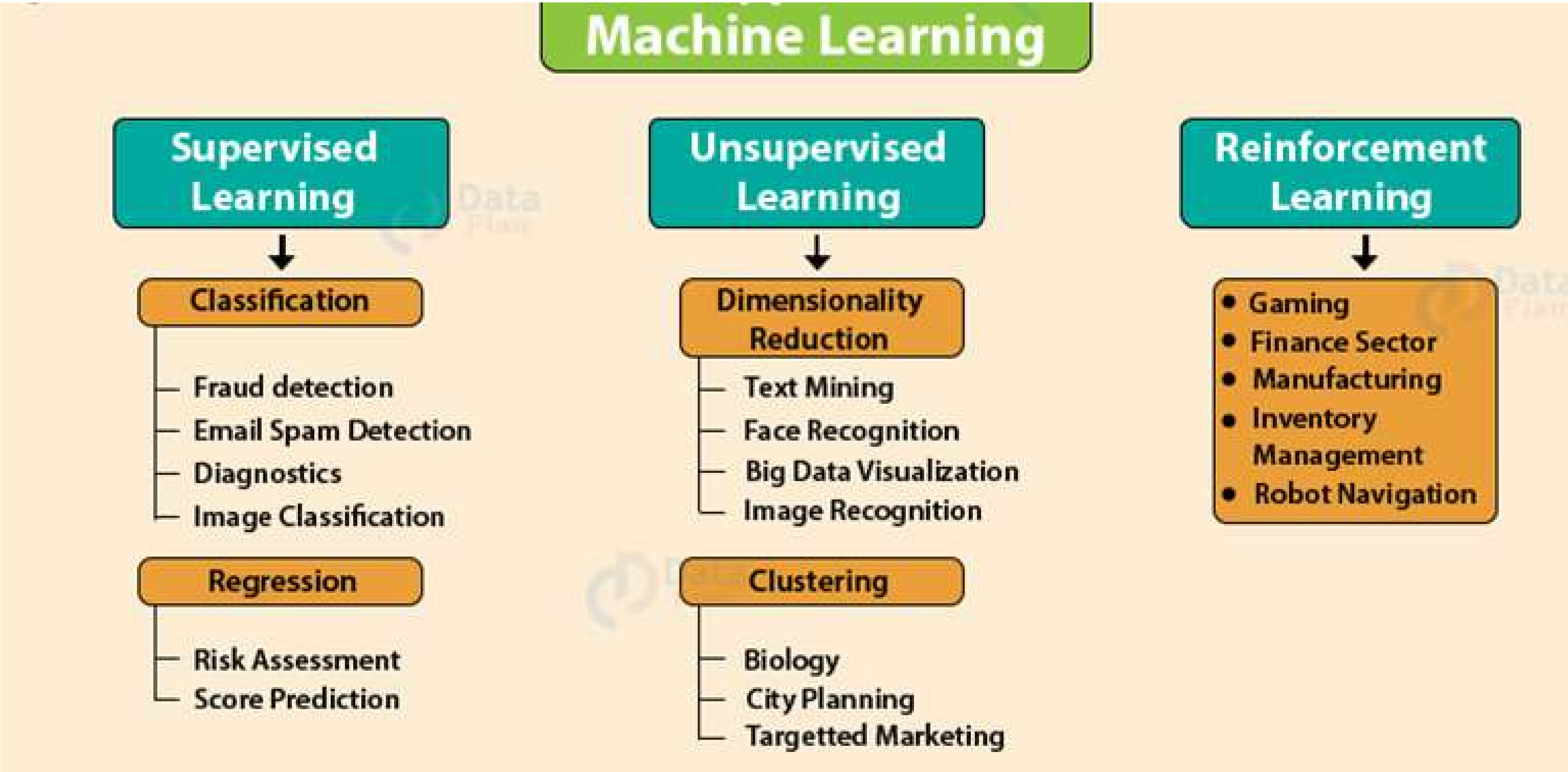


인공지능(AI)의 분류



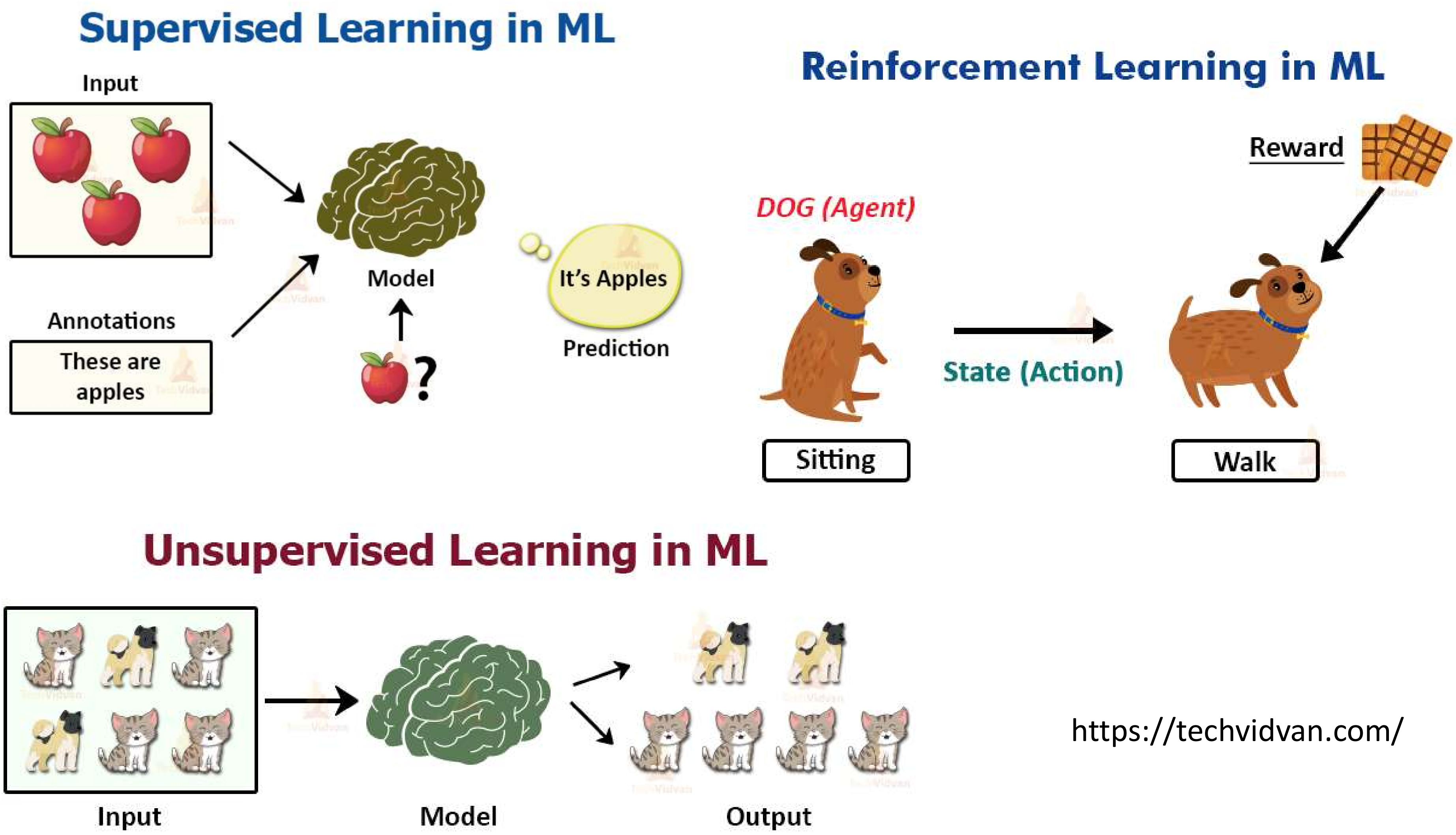
Source: *Chethan Kumar, Iqreate Infotech*

기계학습(Machine Learning)의 분류



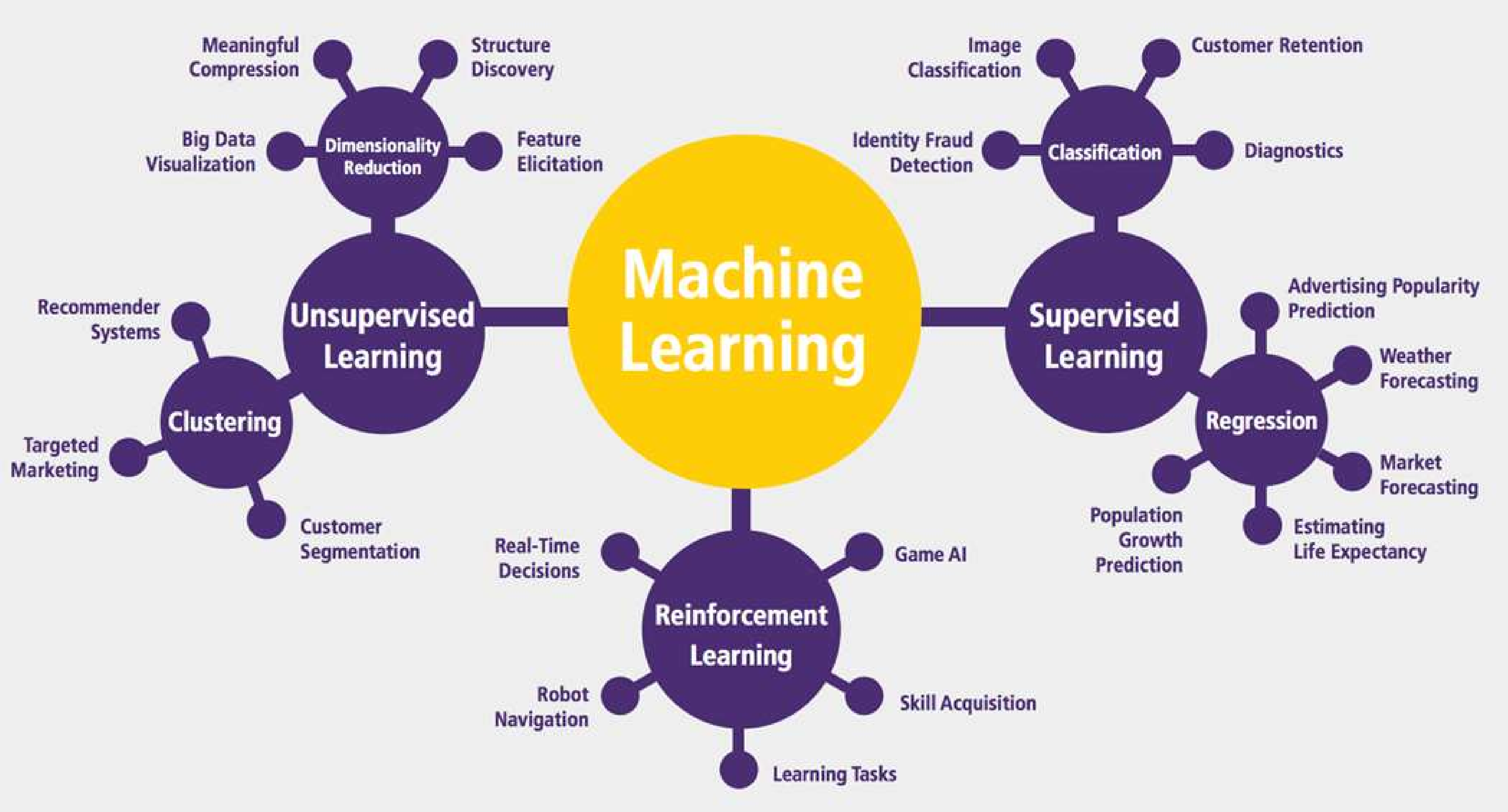
https://data-flair.training/

Machine Learning



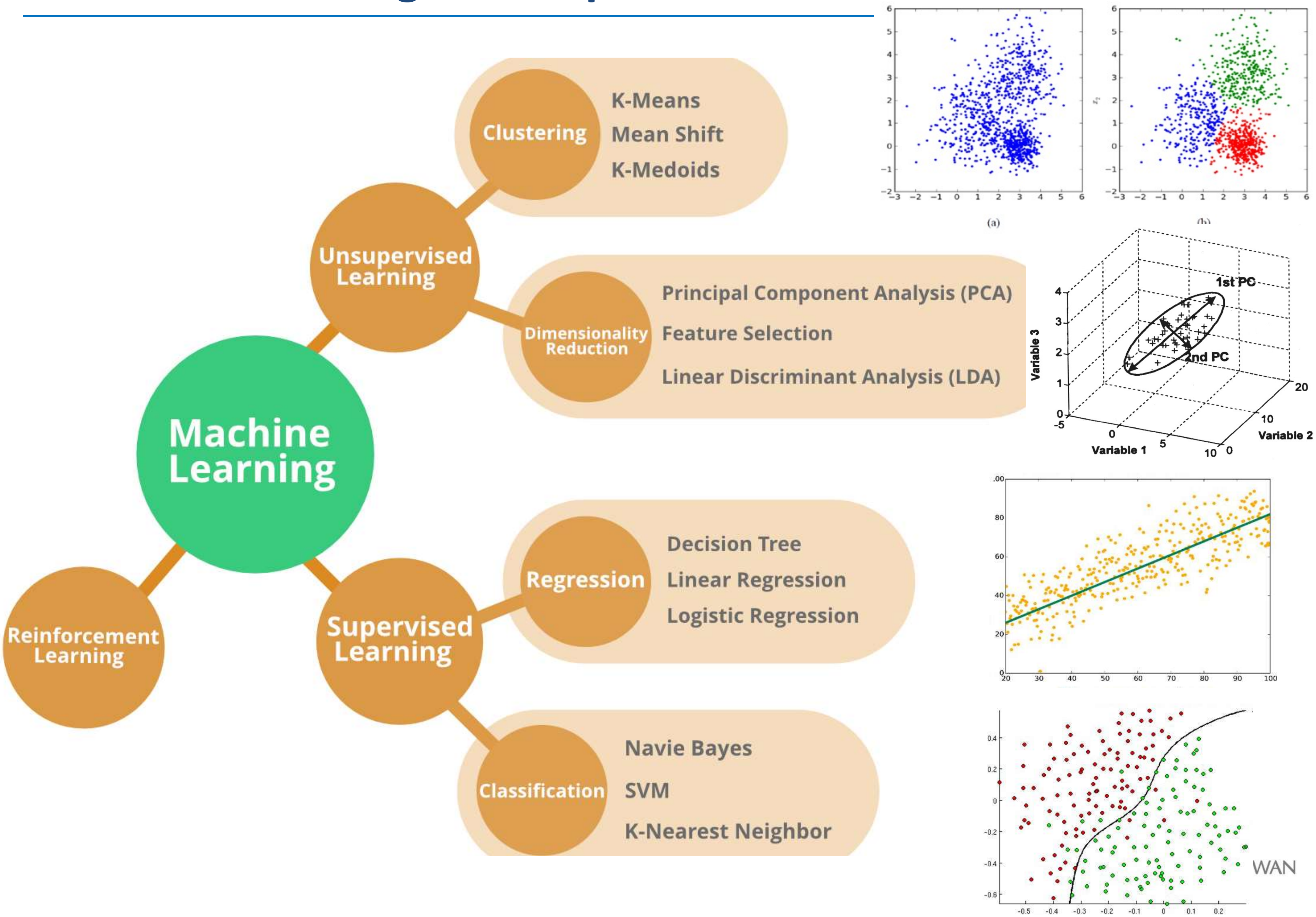
https://techvidvan.com/

Machine Learning and Applications

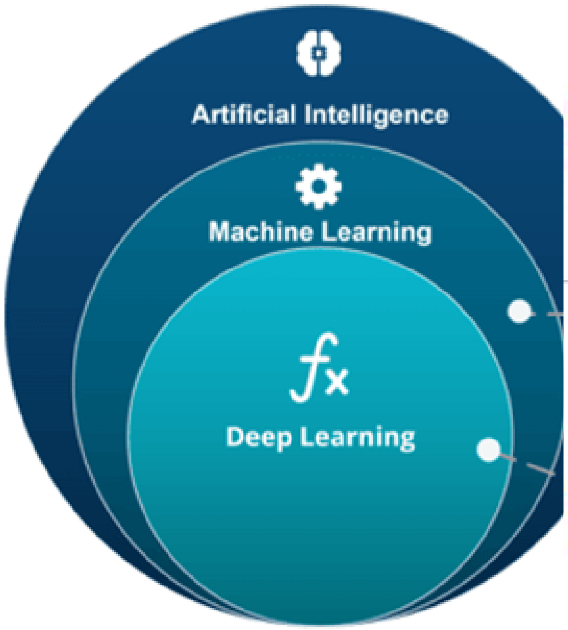


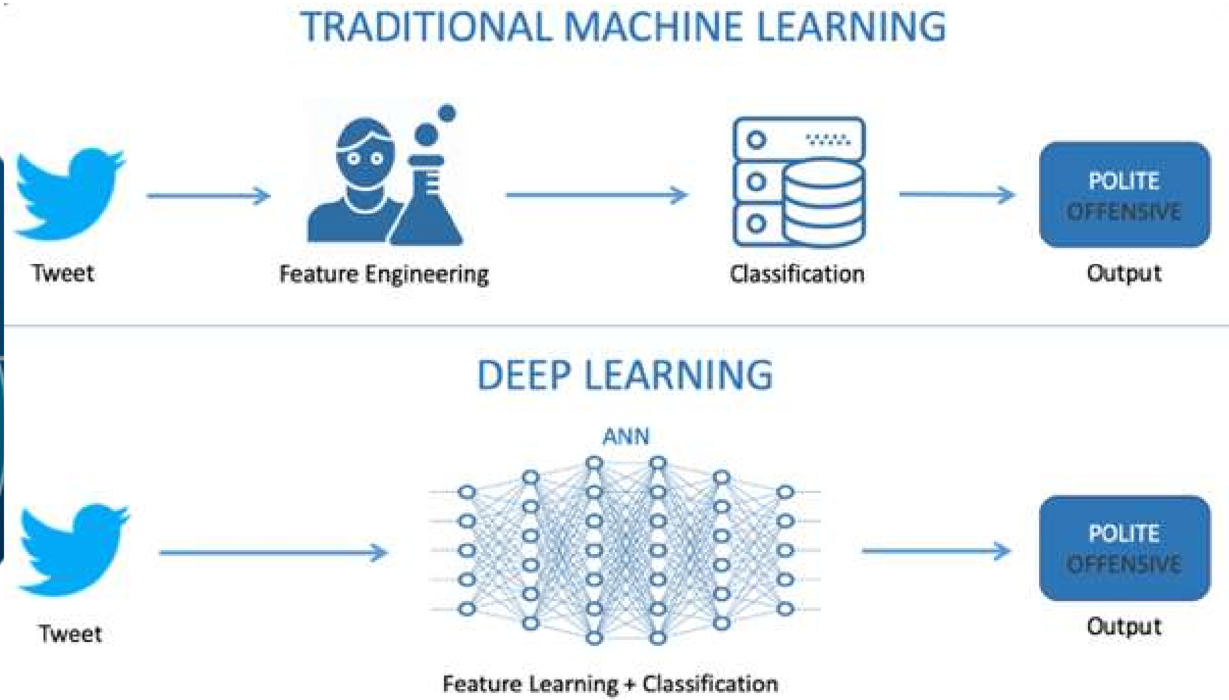
https://www.guru99.com/machine-learning-tutorial.html

Machine Learning Techniques



Deep Learning

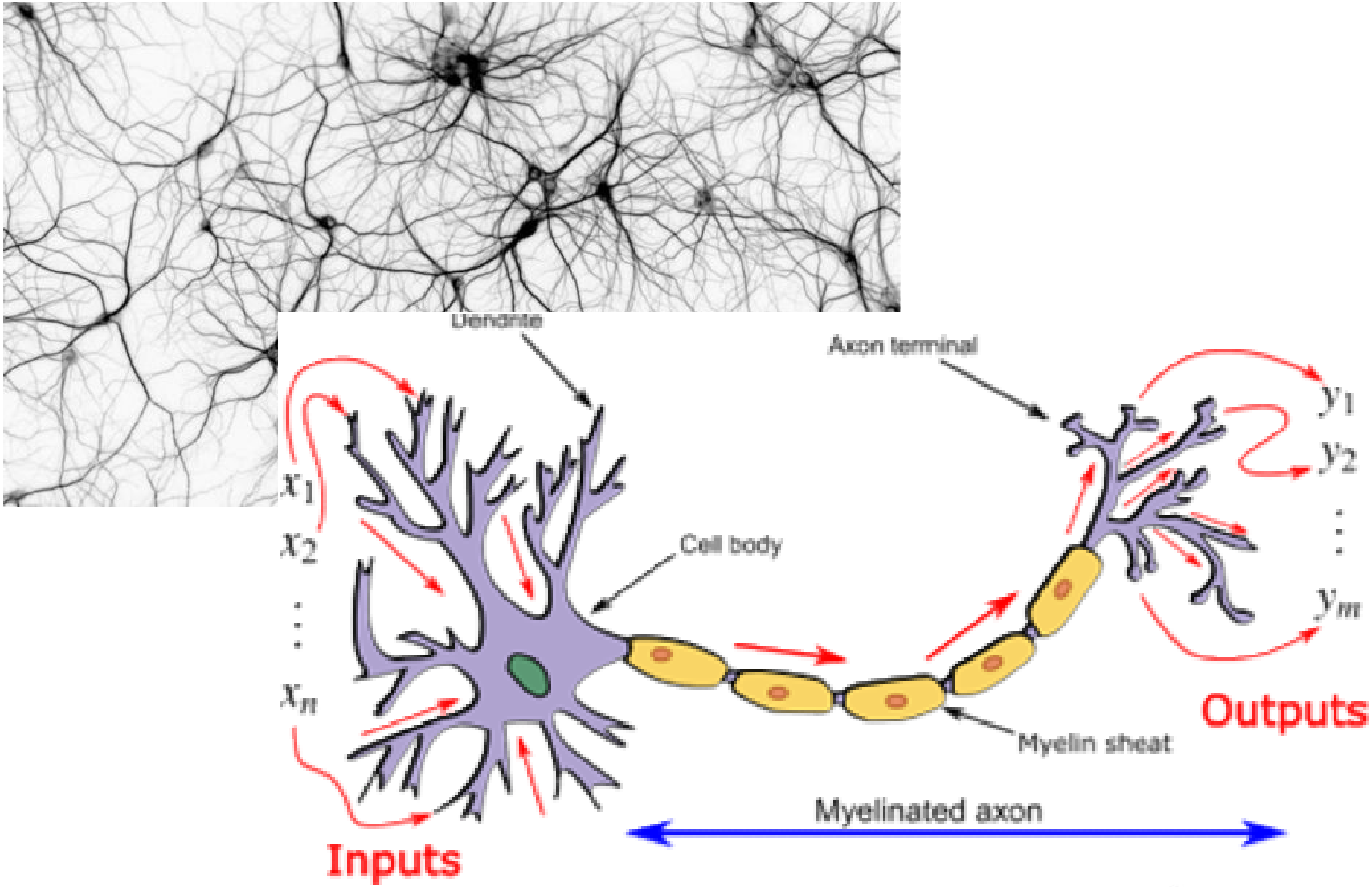




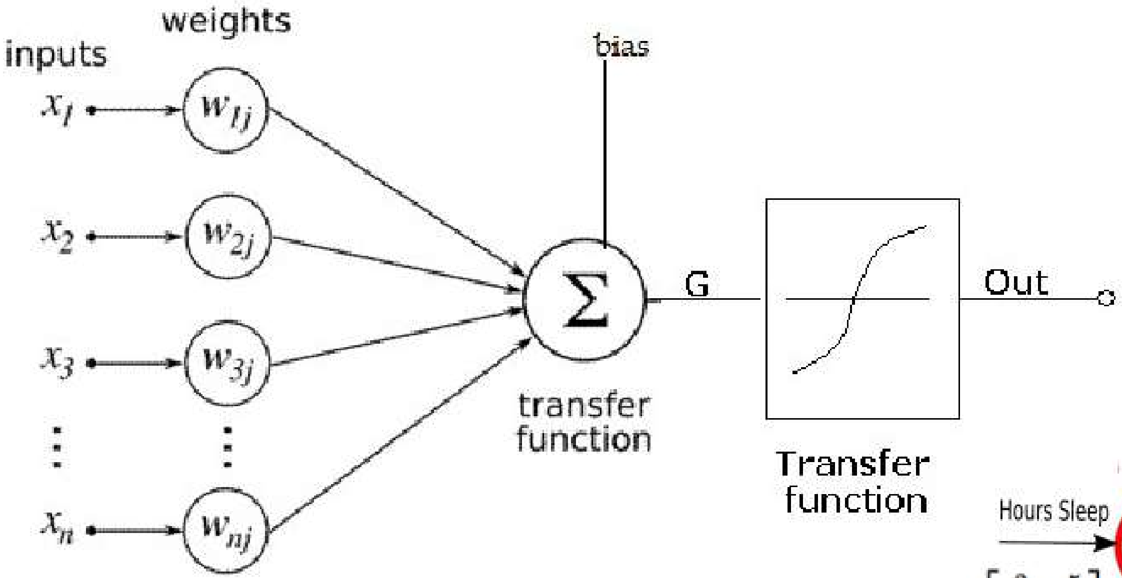
https://datawider.com/

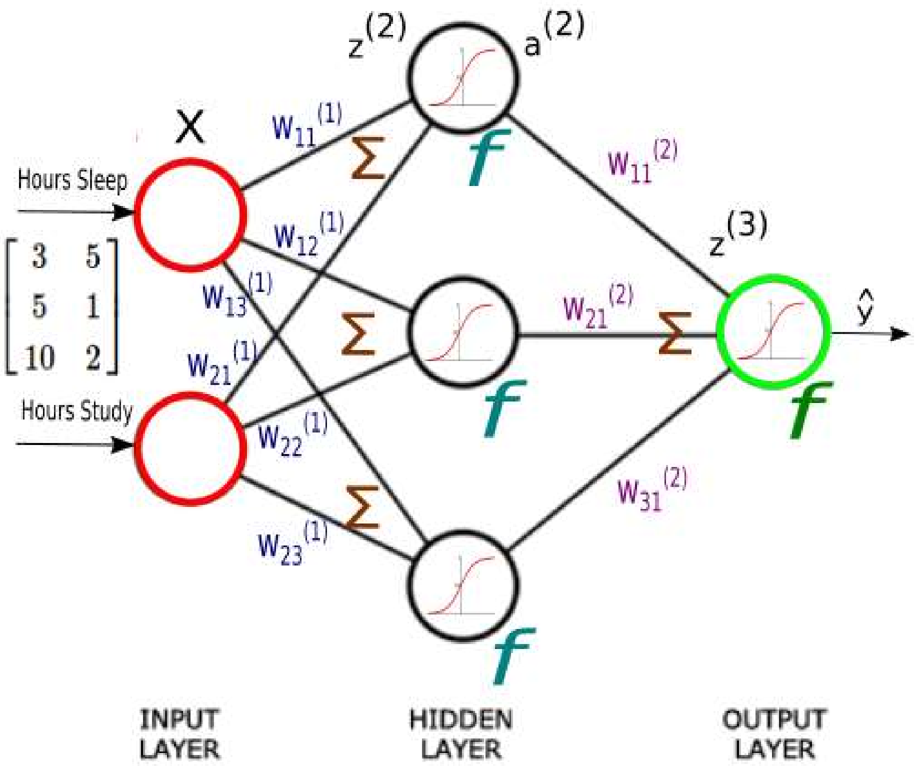
https://thenewstack.io/demystifying-deep-learning-and-artificial-intelligence/

Neural Network (NN)



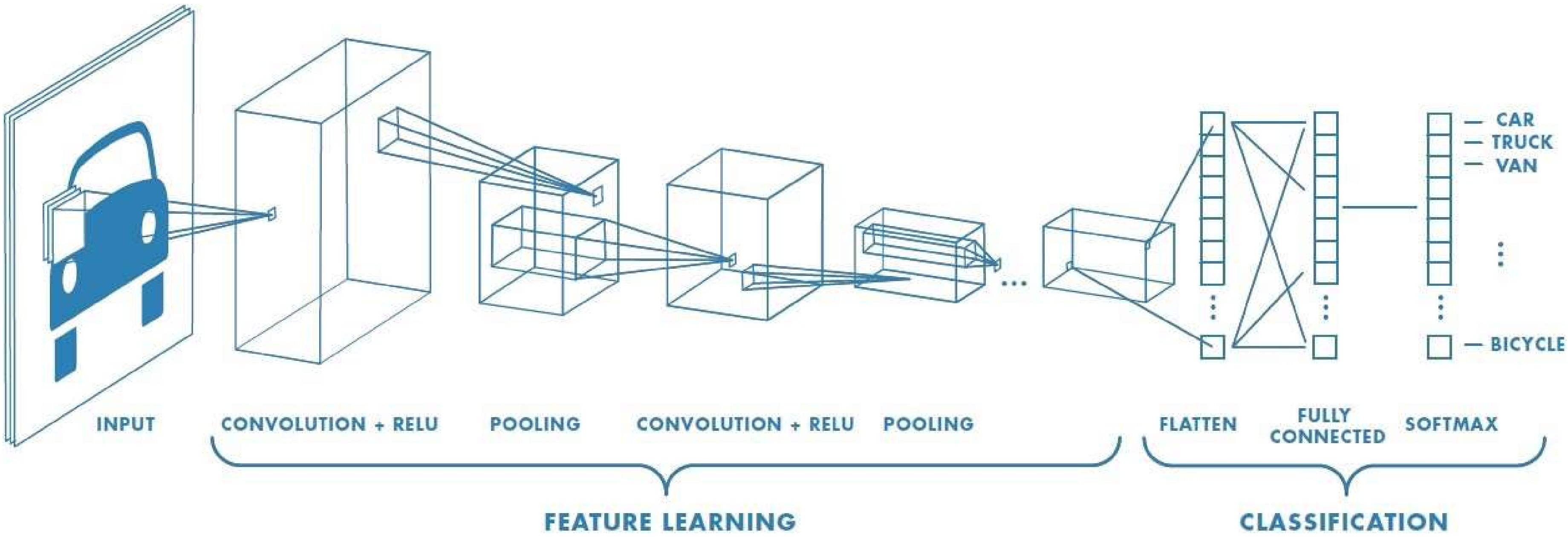
Artificial Neural Network (ANN)





Deep Learning: Convolutional Neural Network (CNN)

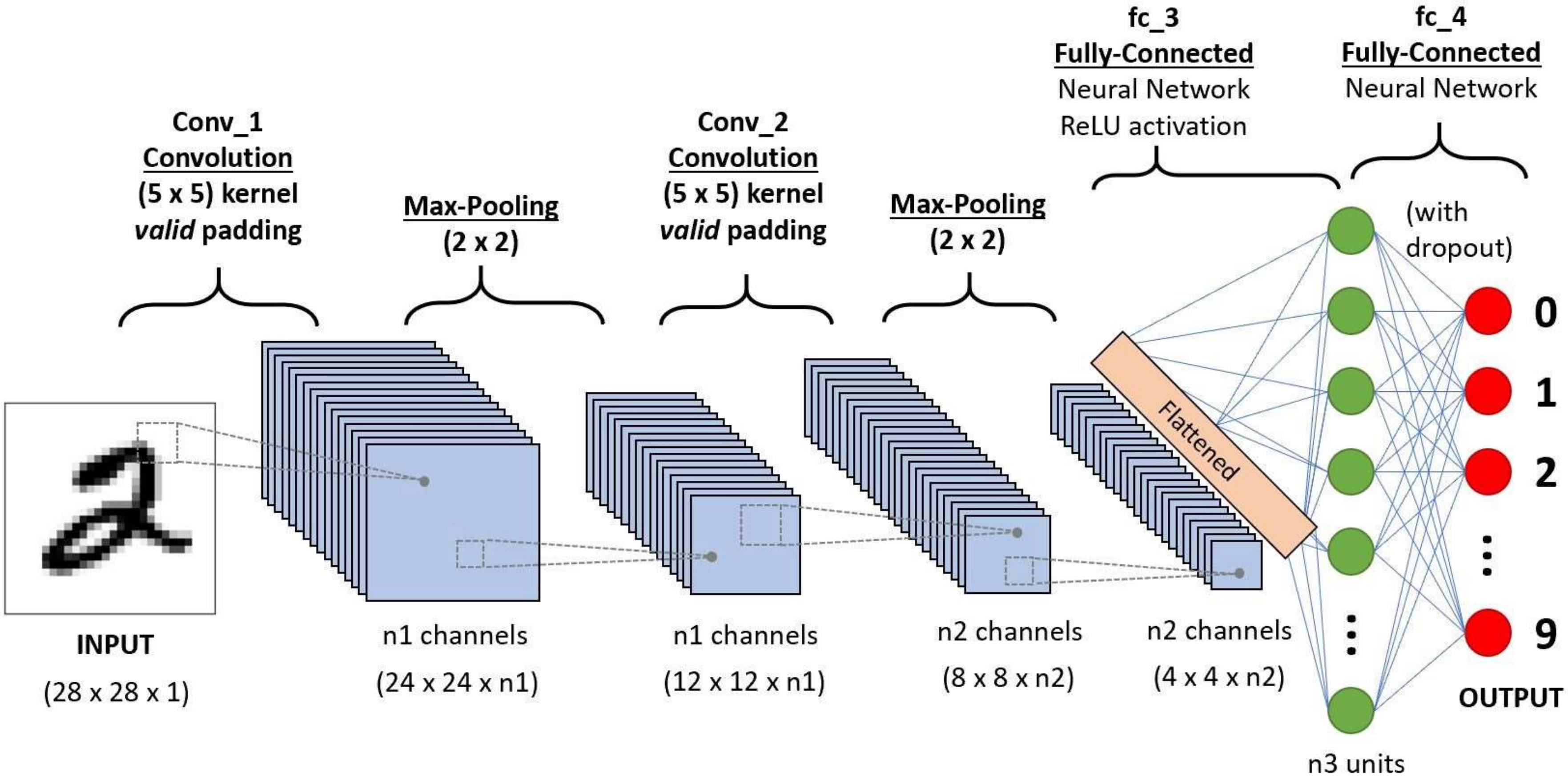
- CNN is powerful for image classification.



SUNG KYUN KWAN UNIVERSITY

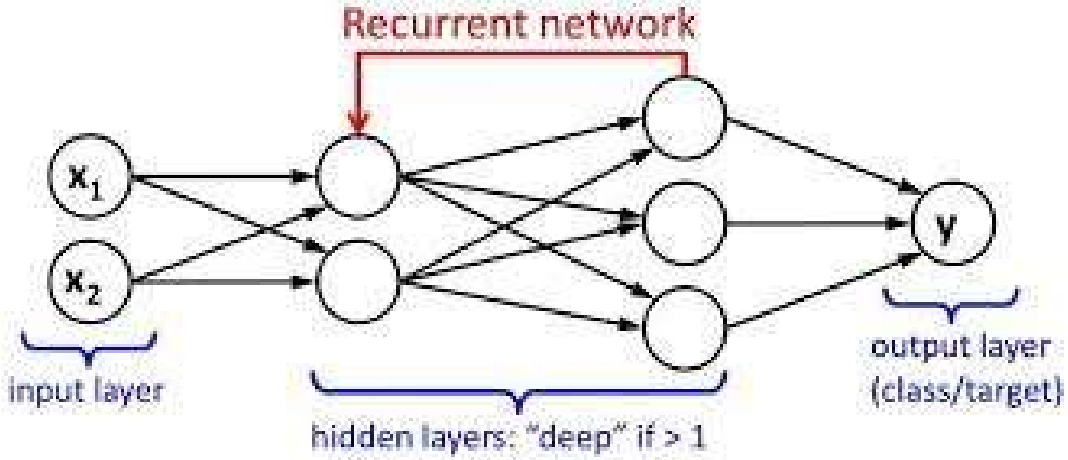
Deep Learning: CNN

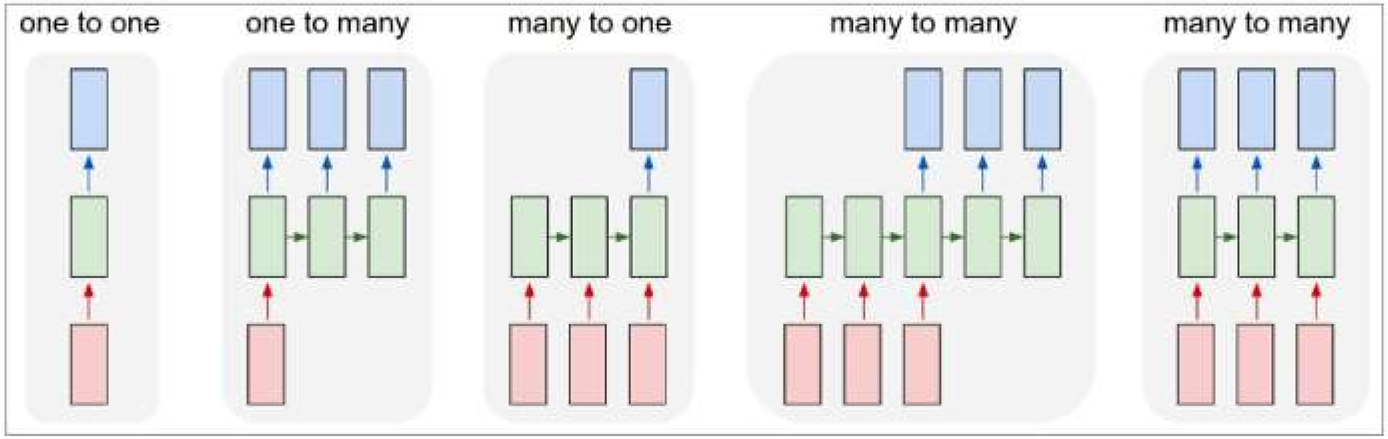
- Example for handwritten digits



Deep Learning: Recurrent Neural Network (RNN)

- Powerful for tasks that are dependent on a sequence of successive states





SUNG KYUN KWAN UNIVERSITY

플라즈마 분석을 위한 빅데이터 분석 기법

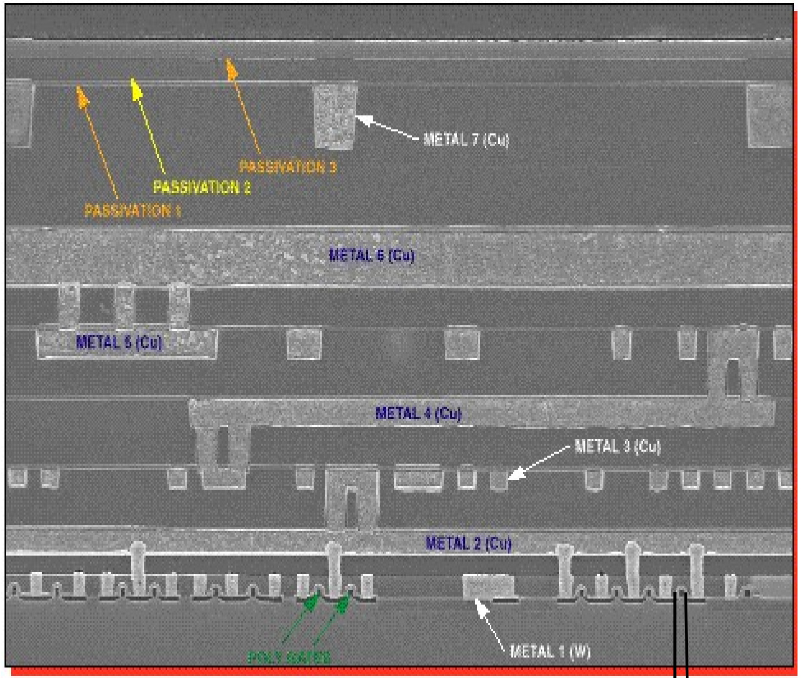
- Multivariate Analysis Techniques
  - Principal Component Analysis (PCA)
  - Real-time Density-based Clustering Analysis
  - K-Means Clustering Analysis
  - Gaussian Mixture Model

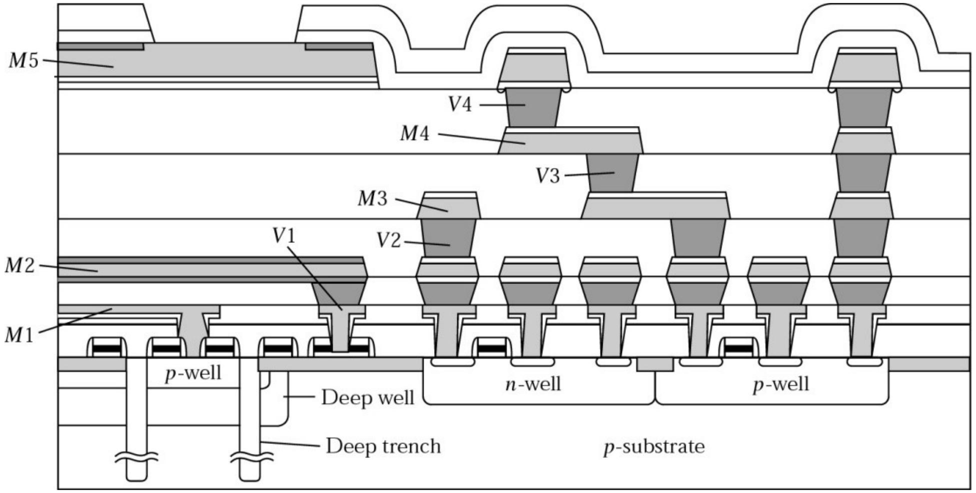
- Artificial Neural Network/Deep Learning

반도체 공정

Micro/Nano-scale Integrated Circuit (IC)

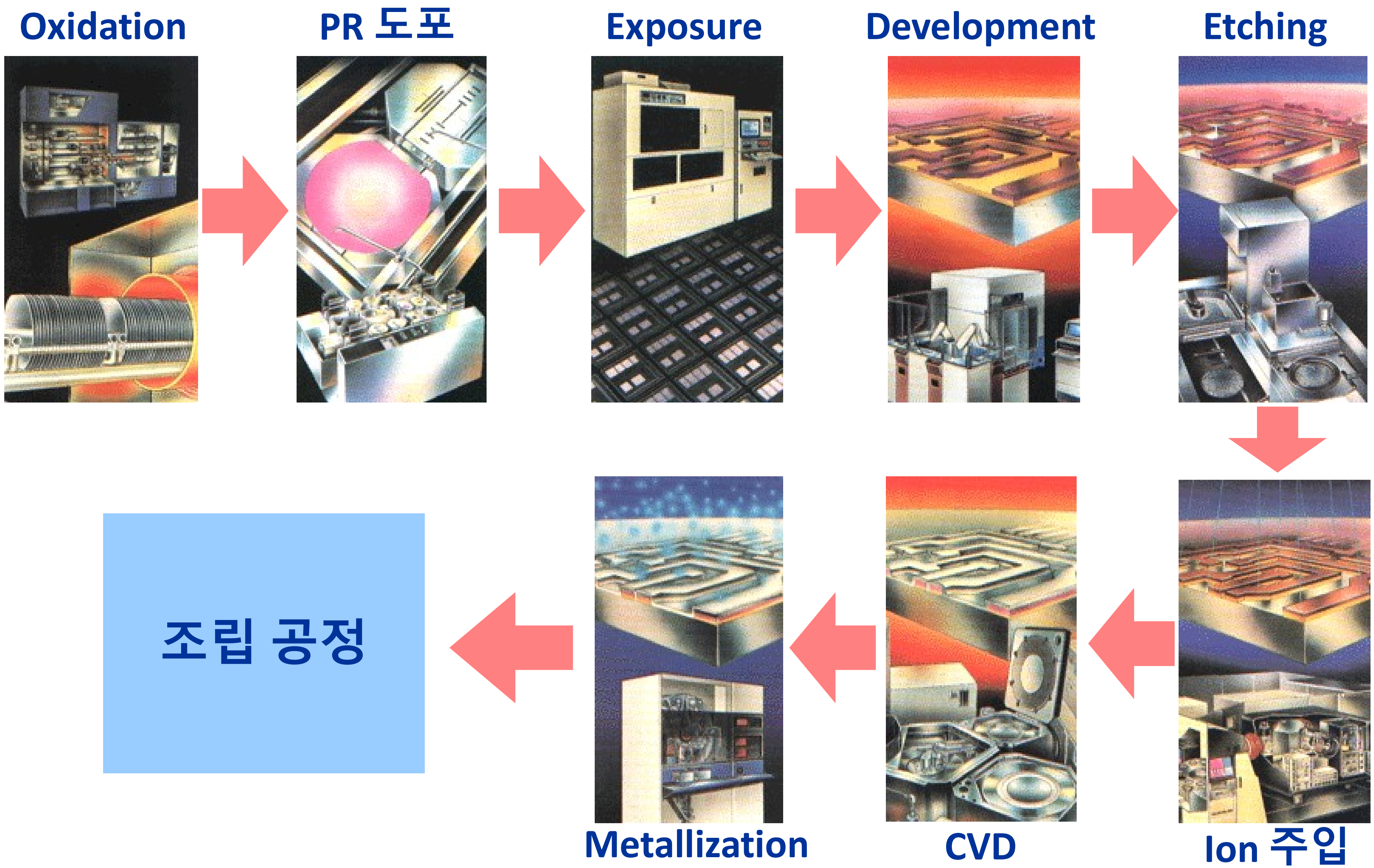
- Plasma processing steps are 30~40% of IC fabrication processing.



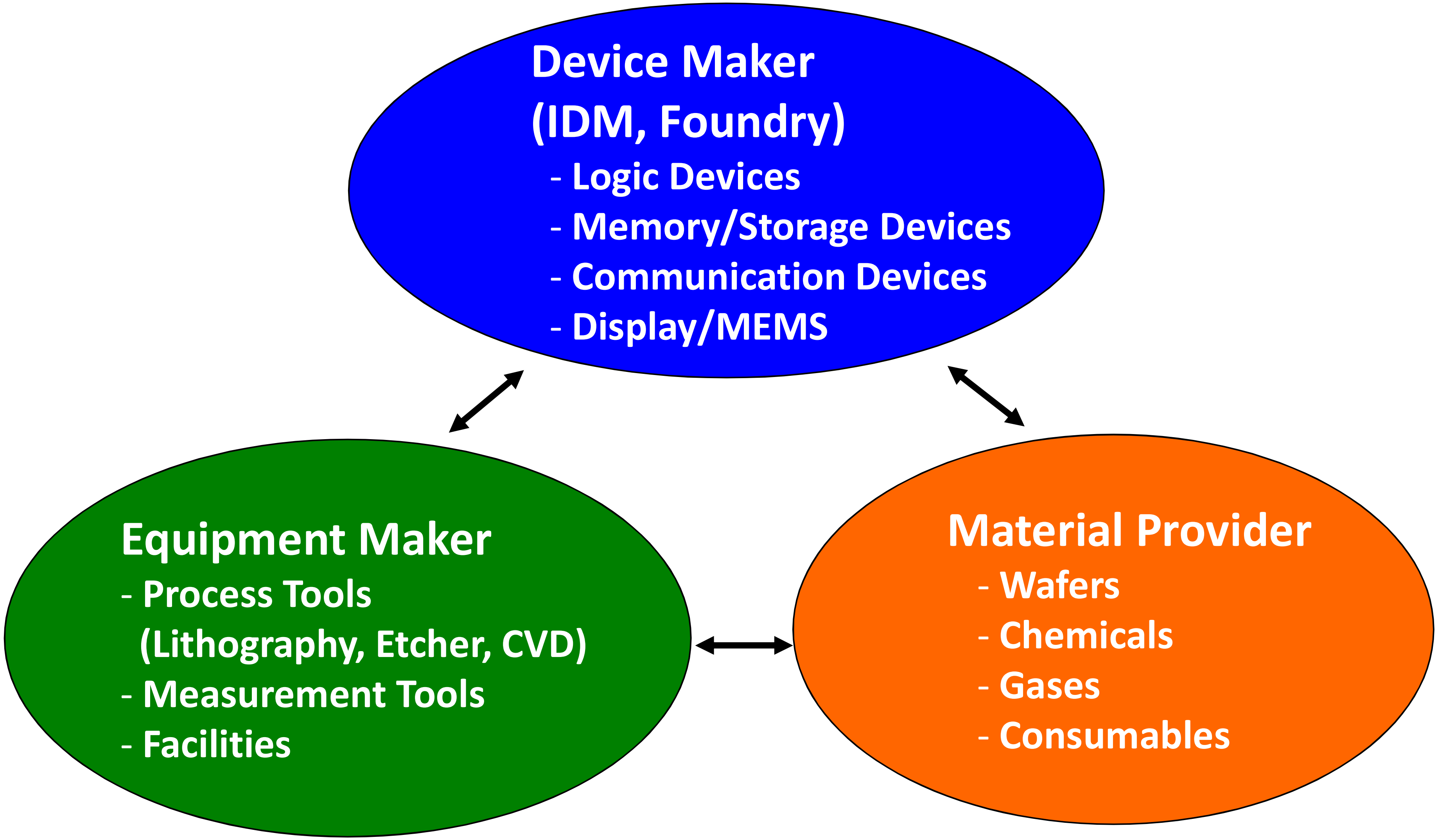


Minimum CD
< 10nm @ 2020

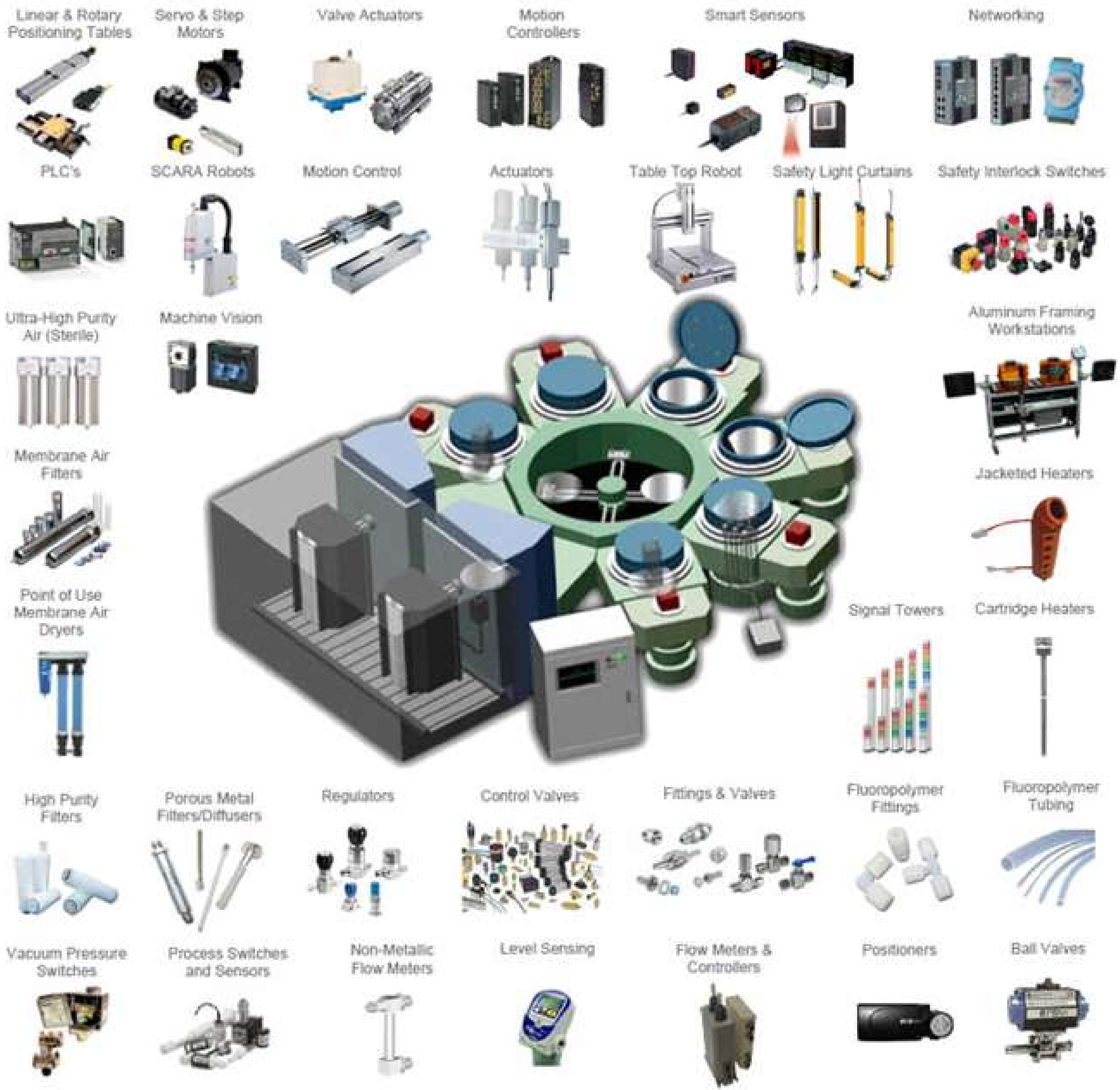
Major Processing Steps



Semiconductor Industry



Processing Tools and Parts



https://www.valin.com/

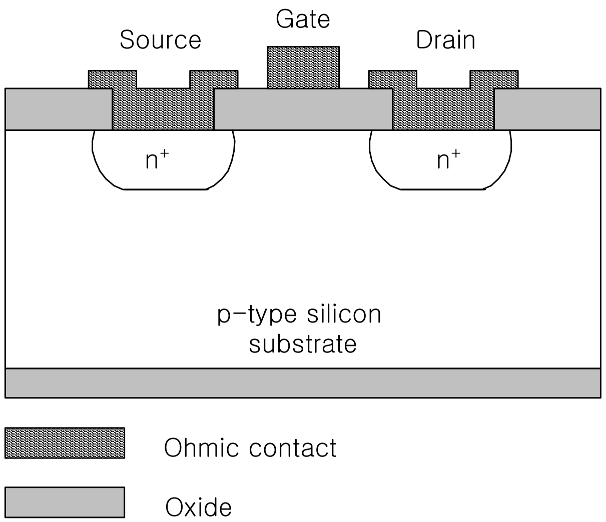
SUNG KYUN KWAN UNIVERSITY

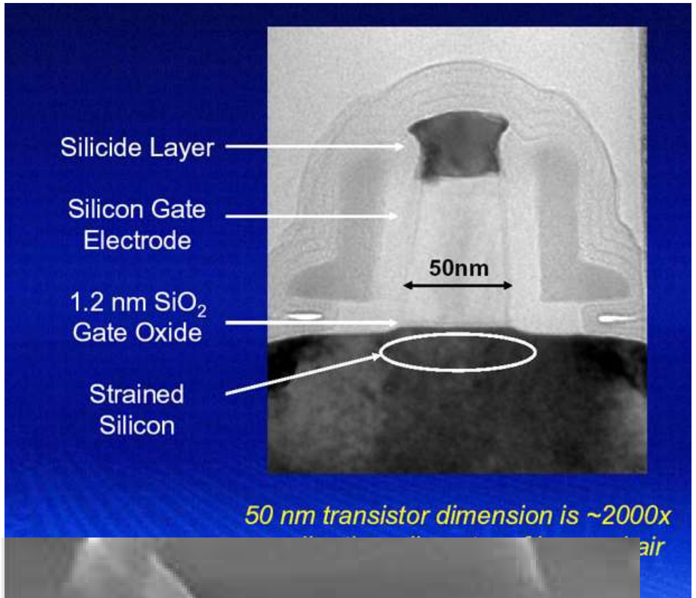
반도체 공정 기술 동향

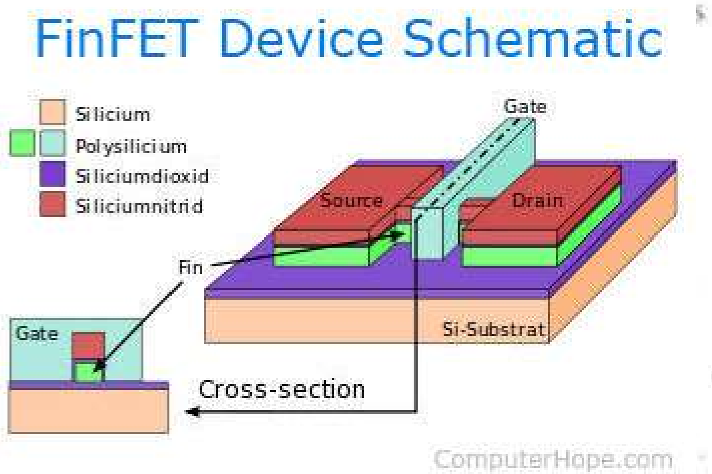
- 소자의 초미세화: ~10nm 크기 패터닝 인쇄 및 식각 기술
  - Multiple patterning technology
  - EUV technology
- 소자의 3차원화: 초박막 적층기술, 고종횡비 식각 기술
  - FinFET transitors and 3D NAND structure
  - Atomic layer processing
  - High aspect ratio processing
- 화학반응의 복잡화
  - More and more elements are adopted
  - Diversified precursors
- 공정의 저온화: 플라즈마 기술
  - Plasma processing
- 기계학습/AI데이터 분석 기법의 적용

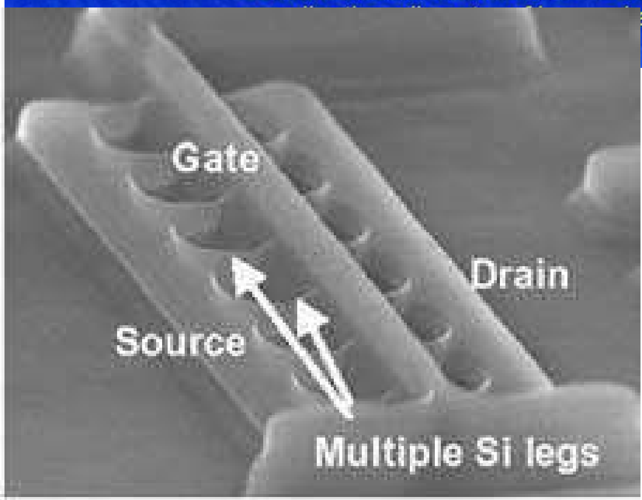
Transistor

- MOSFET (metal-oxide-semiconductor field effect transistor)



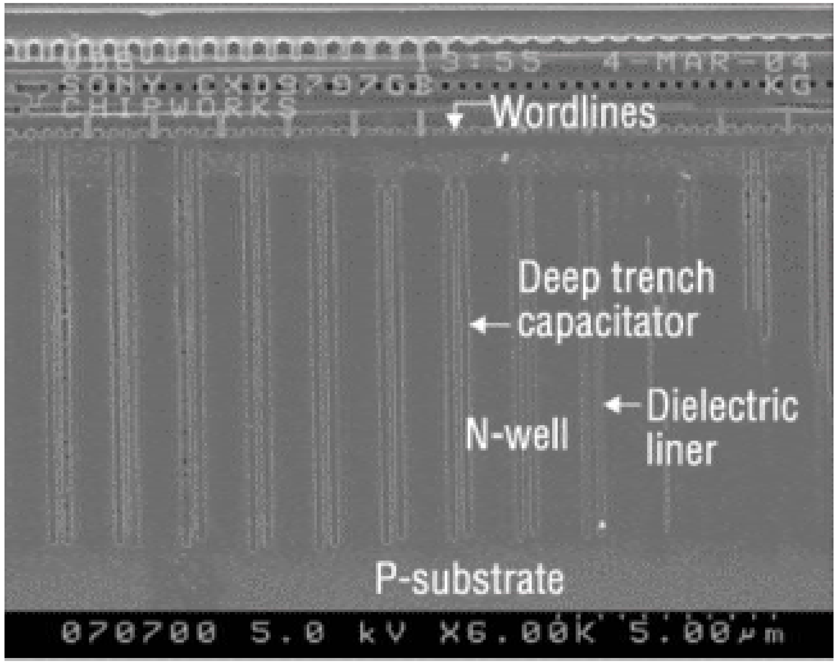


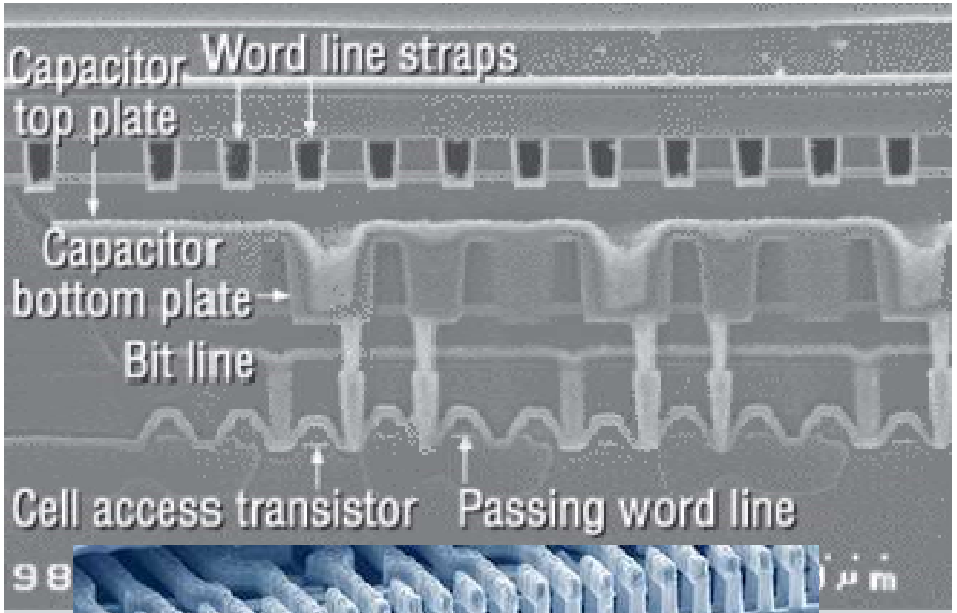


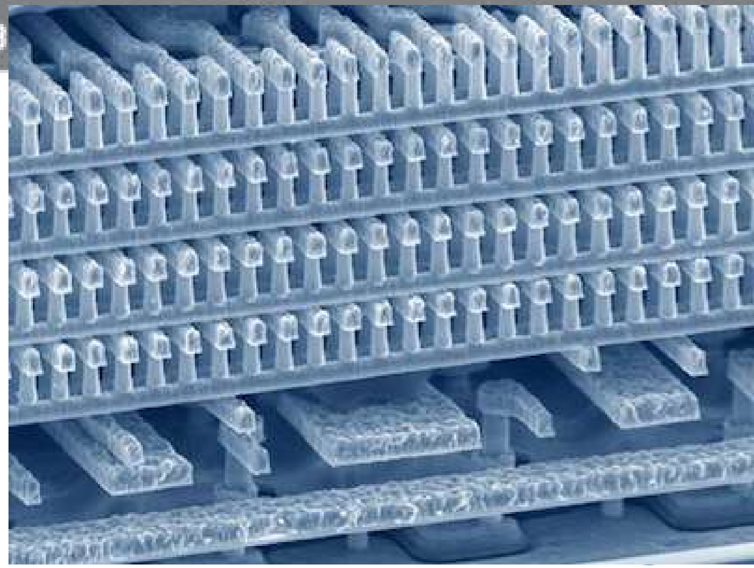


Cross-section of DRAM and Flash Memory Devices

- **Minimum critical dimension is < 20 nm in 2018.**

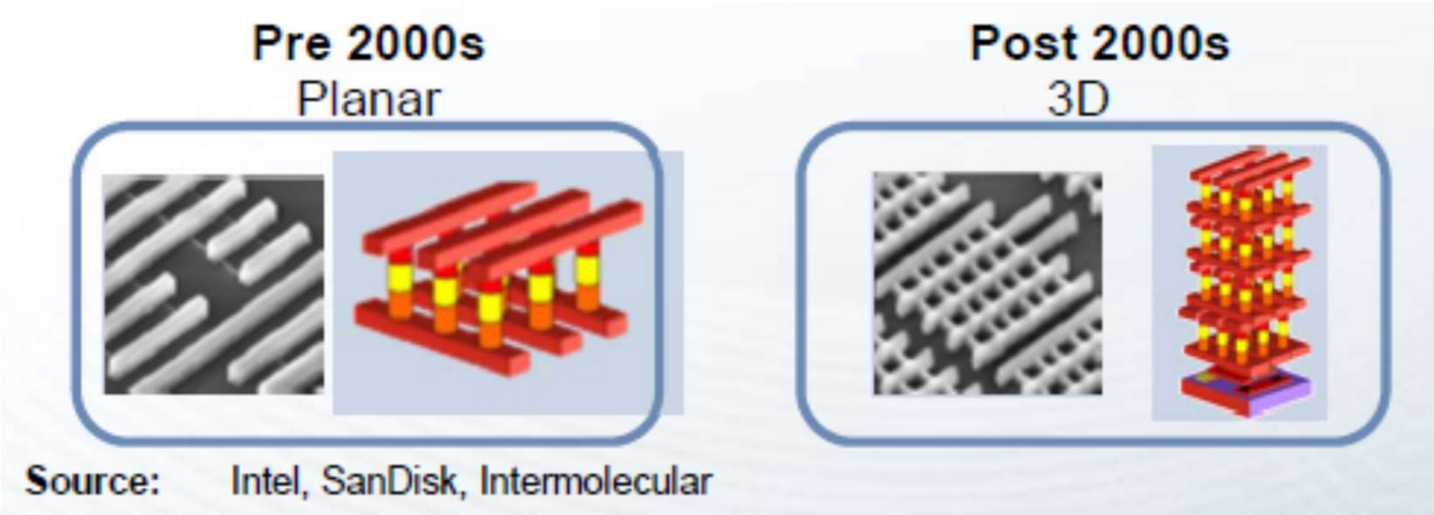


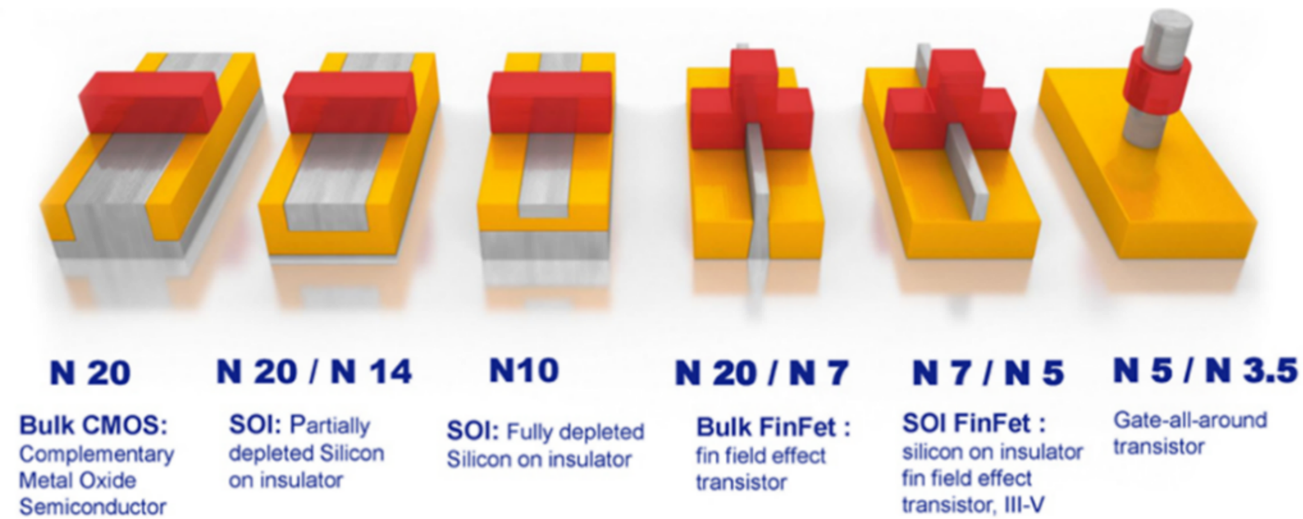


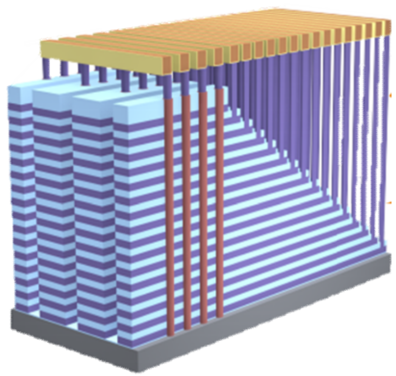


Minimum CD
< 20nm @ 2018

소자의 입체화: 3D Structure

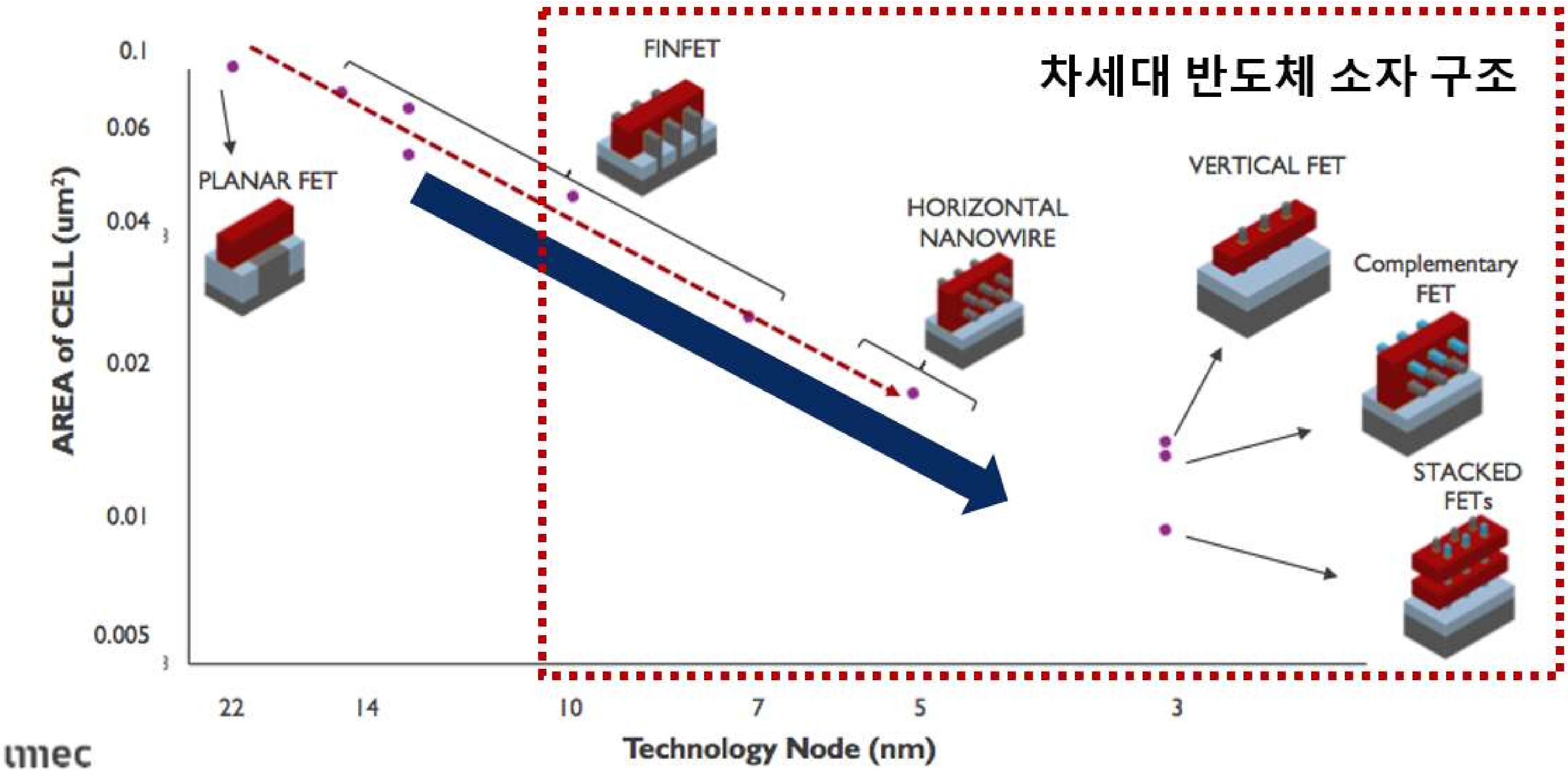






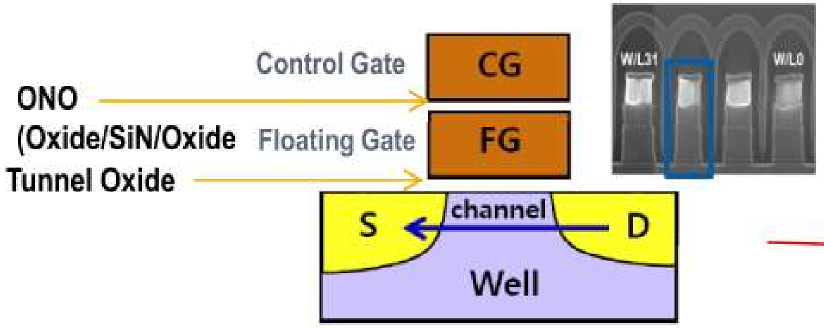
3D NAND

소자의 입체화: 3D Structure

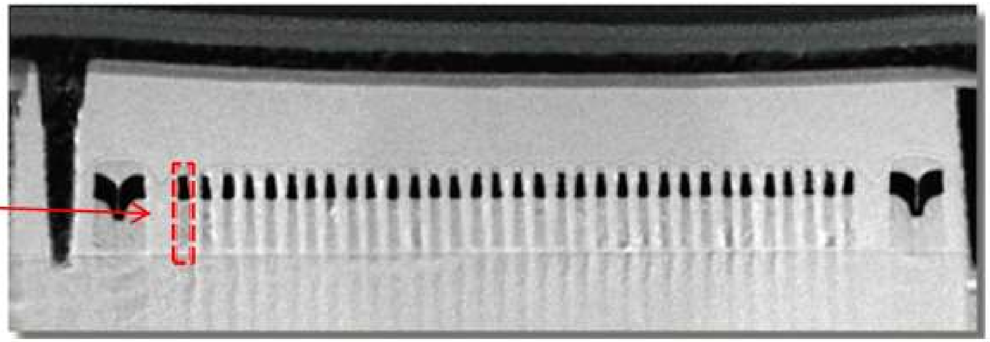


3D cells require high aspect ratio plasma etching

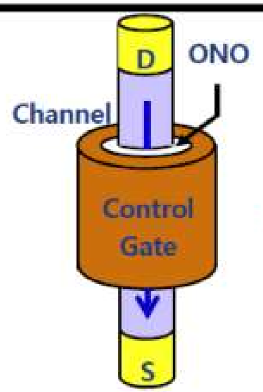
2D cell



- Single crystal Si channel
- Floating gate (or TANOS)
- 1-side gate



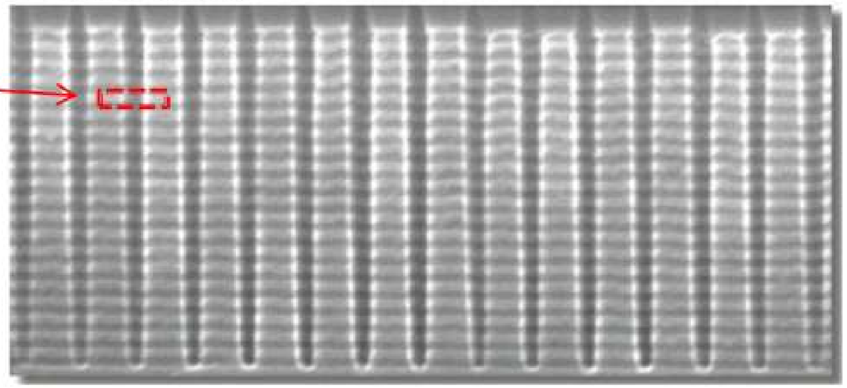
3D cell (Vertical String)



- Poly-Si channel
- SONOS (Si / Oxide / SiN / Oxide / SiN)
- All-around gate
- Channel-last process
- 1 step litho (hole)

http://gigglehd.com/zbxe/613729(

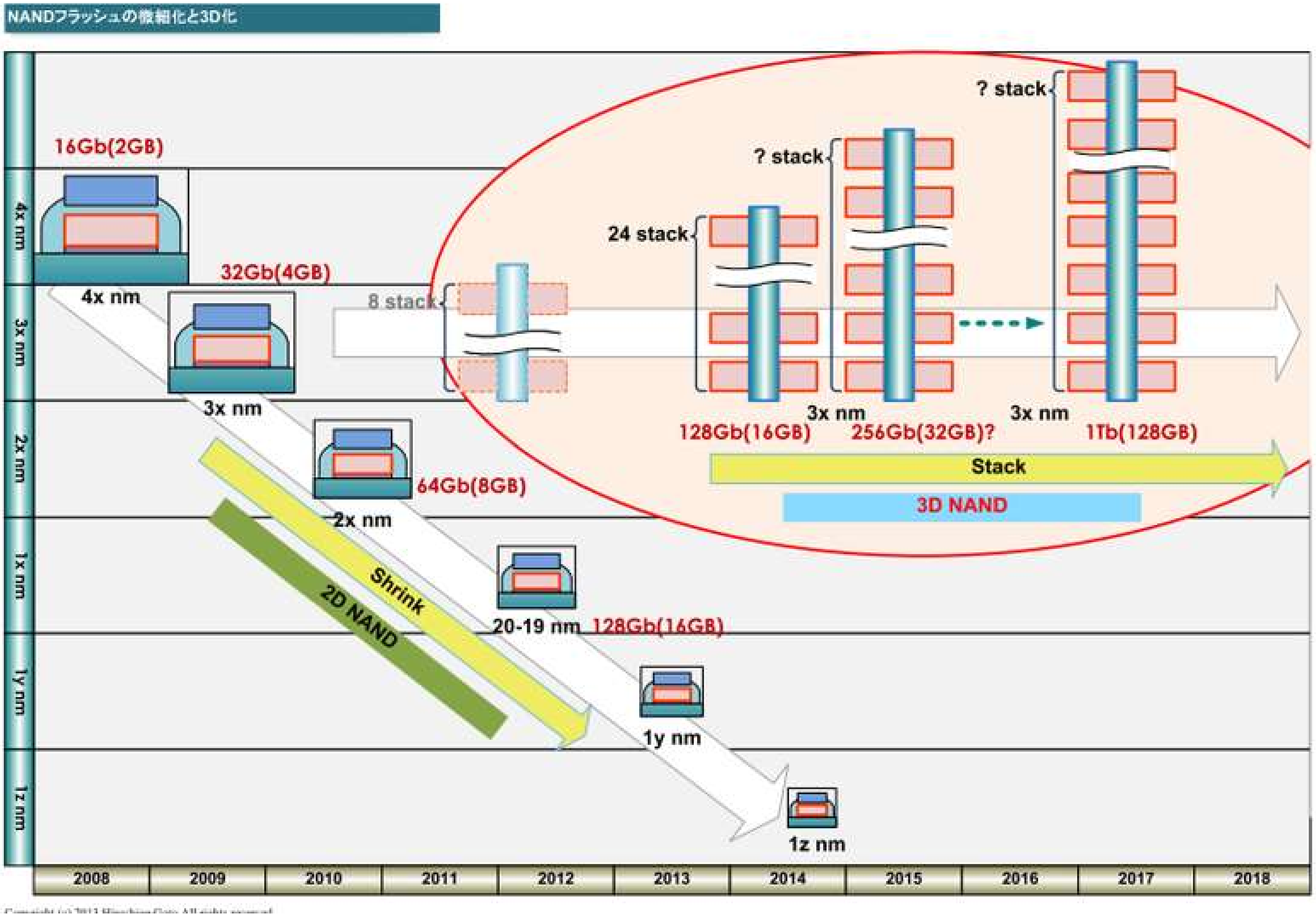
Micron
Wednesday, August 10, 2011



Charges stored in FG
Charges in/out through the tunnel oxide.

SUNG KYUN KWAN UNIVERSITY

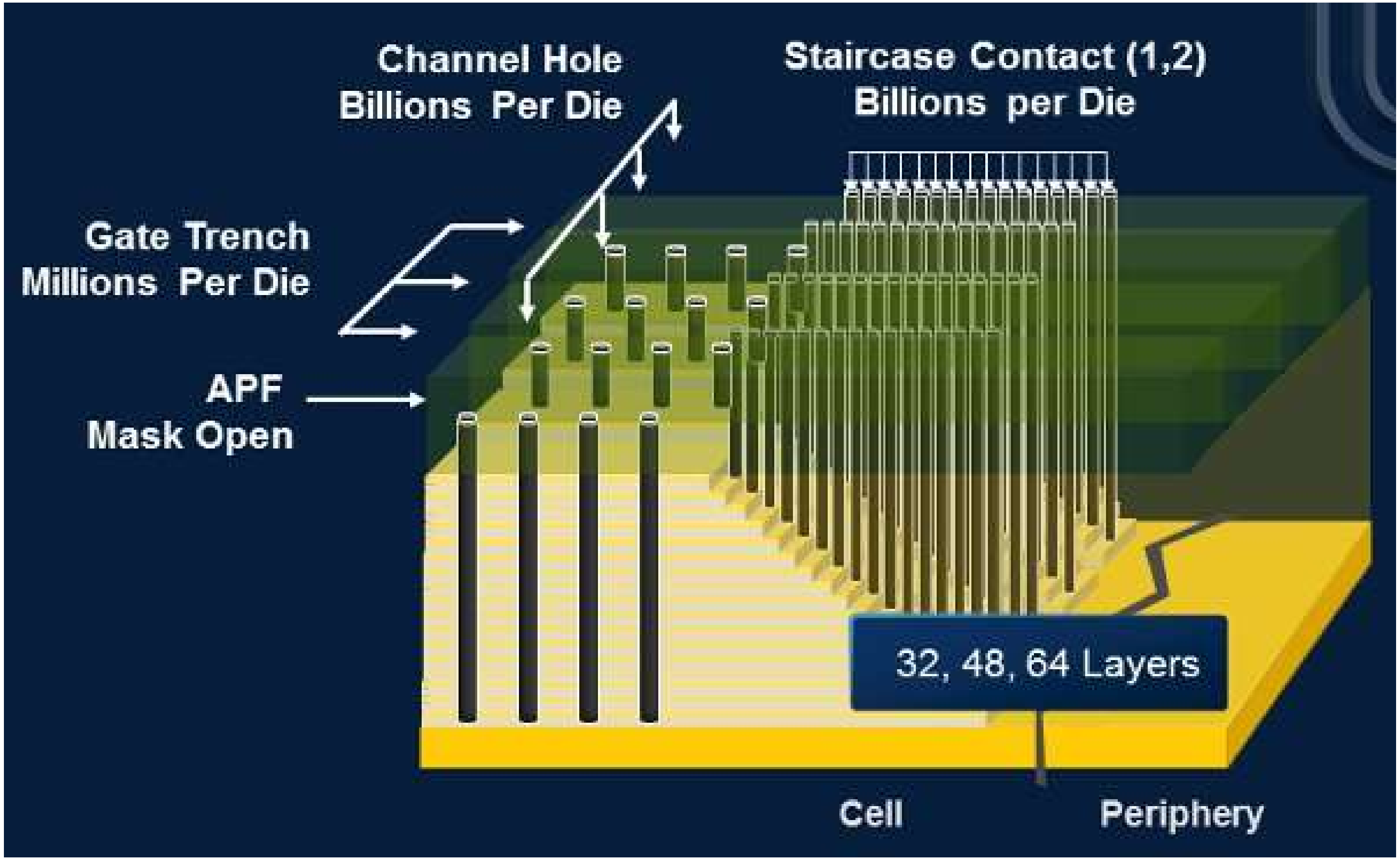
Scaling of 2D Planar and 3D Vertical NAND



http://pc.watch.impress.co.jp/docs/column/kaigai/612339.html

SUNG KYUN KWAN UNIVERSITY

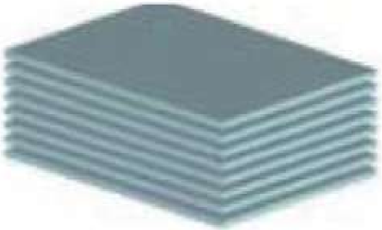
Billions of Holes per Die



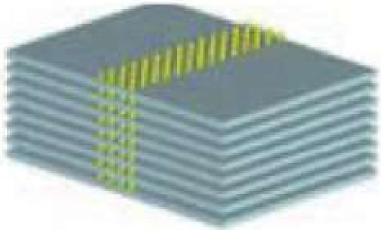
Source: Applied Materials

3D NAND: TCAT Process

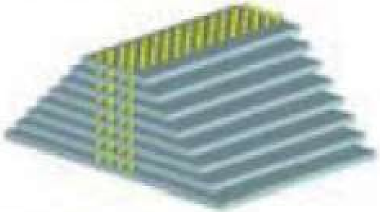
- Channel hole at Ox/SiN multi-stacks followed by WL cut and cell formation with W CG
- Pad formation and BEOL



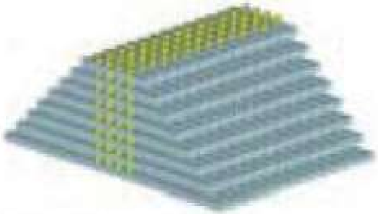
(a) Oxide/Nitride Multi- Layer Deposition



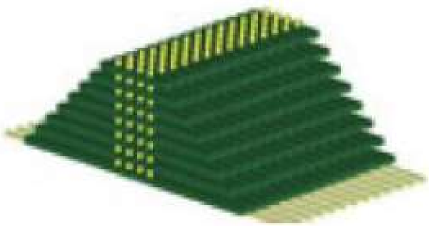
(b) Channel Hole



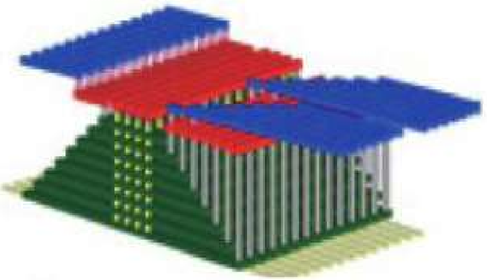
(c) Gate Pad



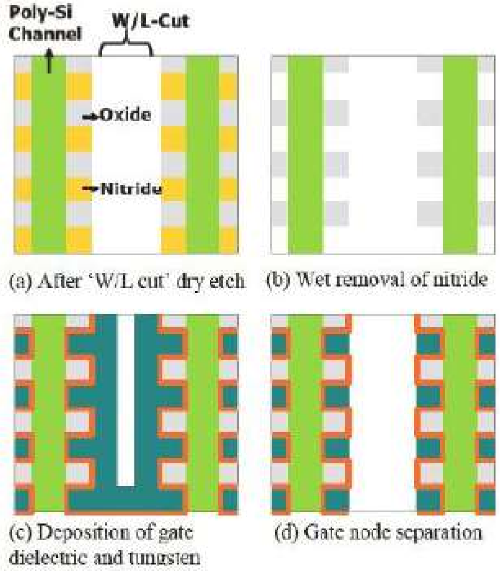
(d) W/L Cut Etch



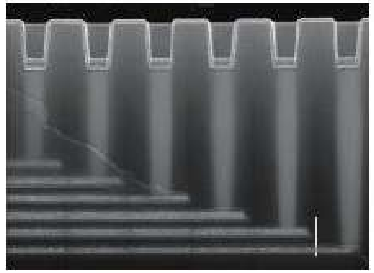
(e) Gate replacement



(f) BEOL



[Gate Replacement Process]

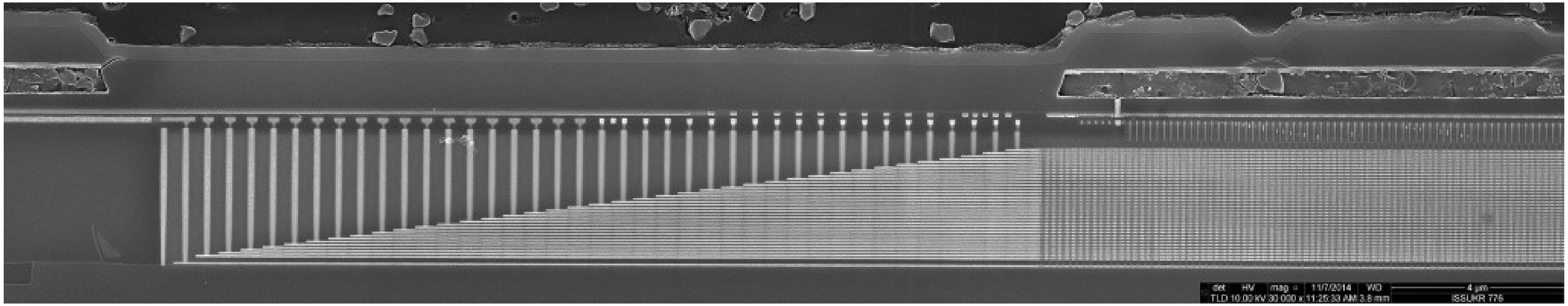


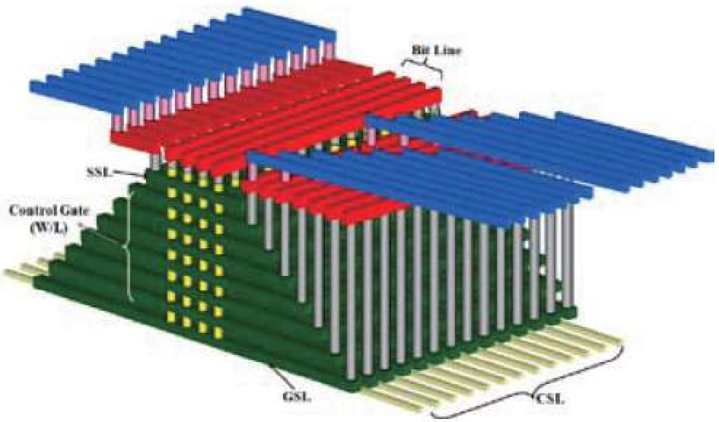
[BEOL]

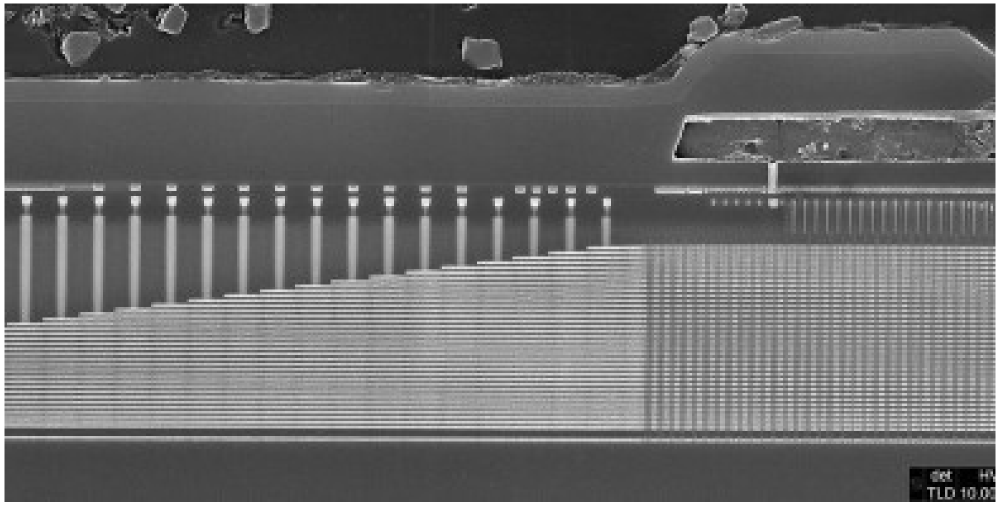
TCAT; Tetra bit Cell Array Transistors
Jaehoon Jang, VLSI 2009

SUNG KYUN KWAN UNIVERSITY

3D NAND Flash Memory

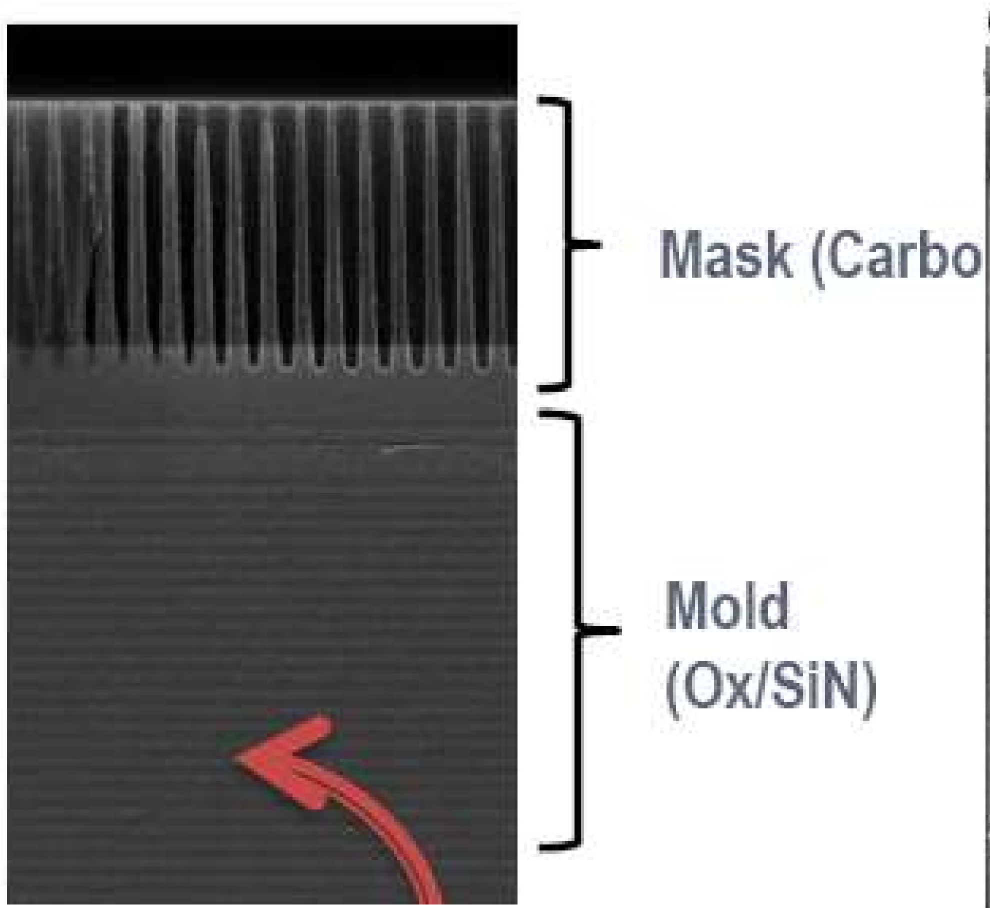


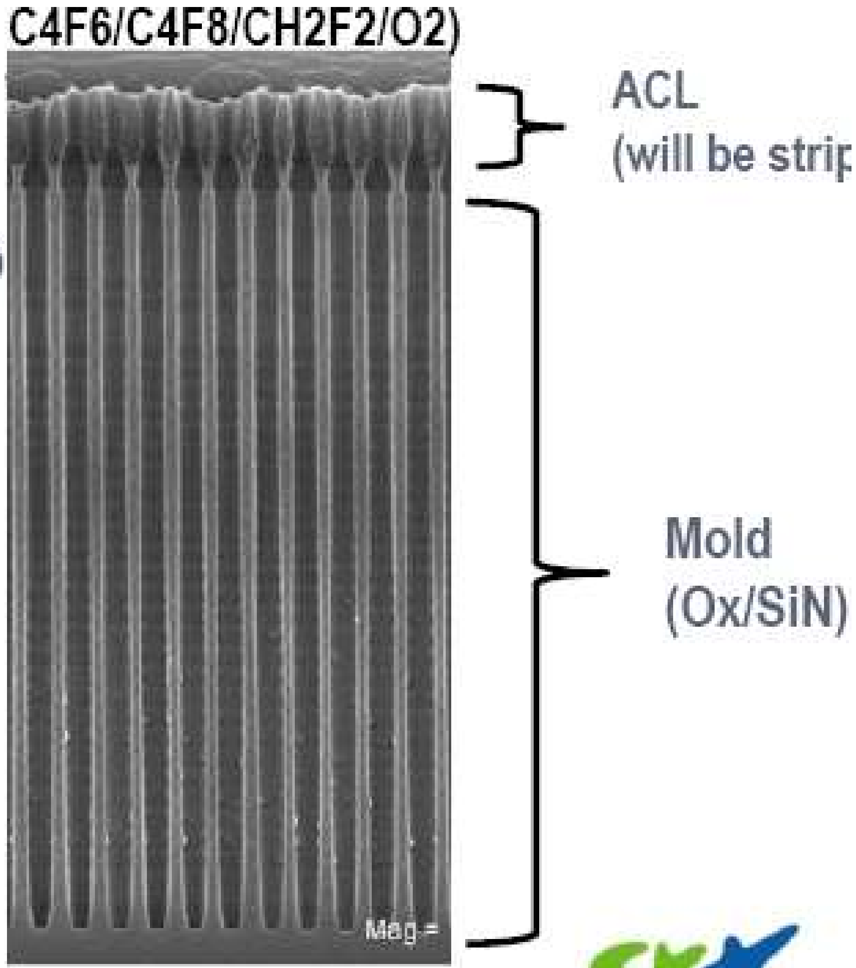




SUNG KYUN KWAN
UNIVERSITY

Process Challenges in 3D NAND





Jungdal Choi, et al., (Samsung), VLSI 2011

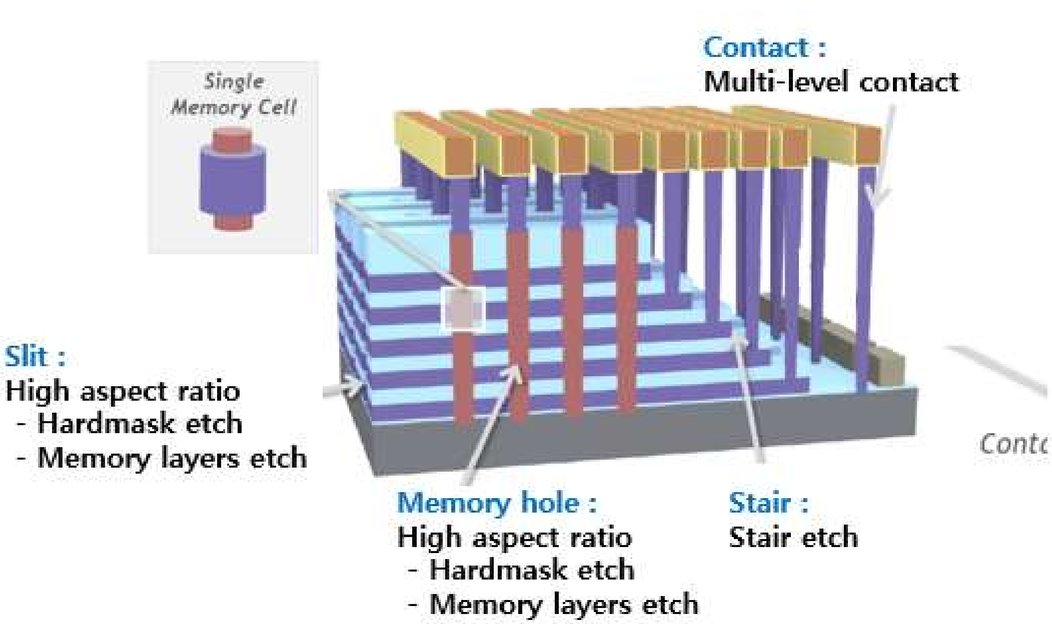
3D V-NAND Challenge

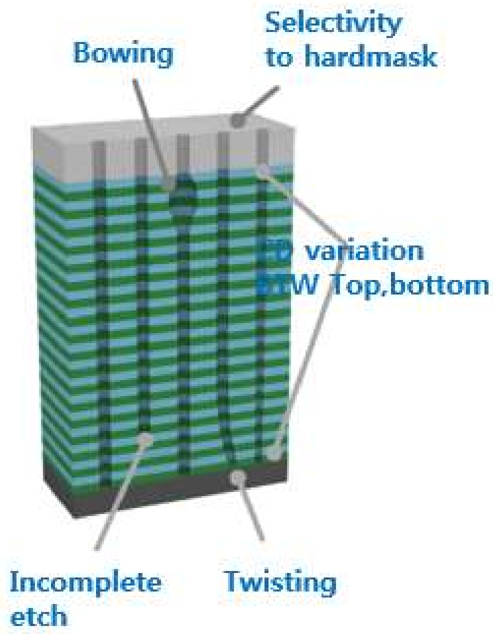
Etch

- ▶ 적층 수 증가에 따른 **고종횡비 식각**필요
  - → 선택비 상향, 반응부산물 효율적인 제거
- ▶ Charging Effect에 따른 Profile 변화
  - → Pulse 적용
- ▶ 공정 진단 기술

Deposition

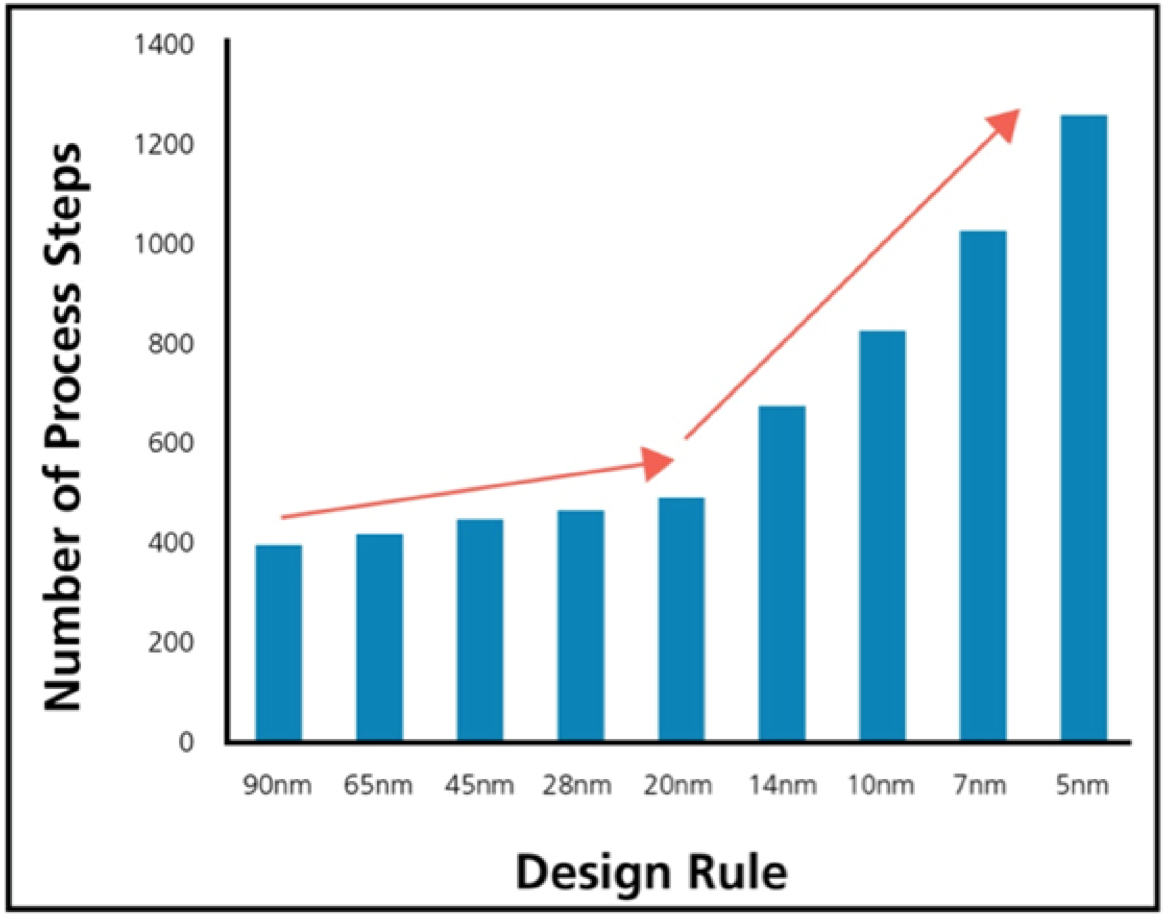
- ▶ 적층높이 증가에 따른 증착공정 수 증가
  - → 균일한 막질의 증착조건, 높은 생산성
- ▶ ALD 공정
  - → 높은 Step Coverage, Void Free조건 확보
- ▶ Defect 제어





※ SEMICON China, LAM

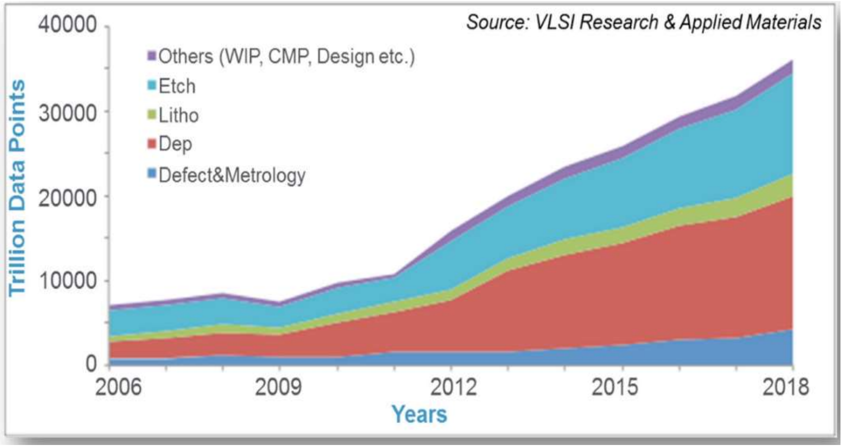
공정의 복잡화

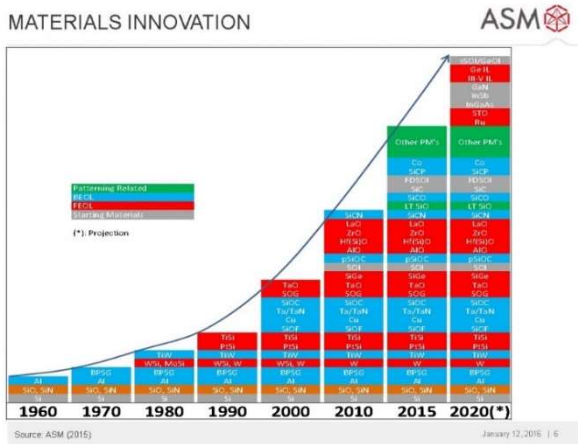


[IC Knowledge Strategic Cost Model, KLA-Tencor internal data]

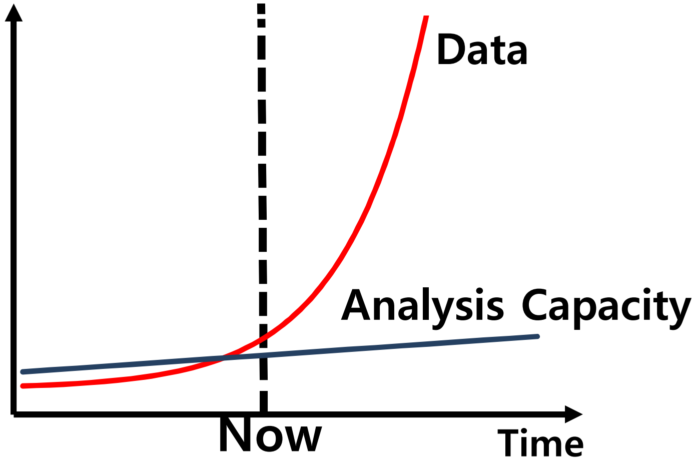
Necessity of Machine Learning

- 플라즈마를 이용한 반도체 및 디스플레이 패널 생산 공정이 세분화되고 다양한 소재들이 공정에 도입됨에 따라, 식각 및 증착 공정에서 발생하는 공정 데이터의 크기가 증가하고 있음.





- 데이터의 저장 및 계산 능력의 발전보다 데이터의 크기 증가가 넘어서고 있다. 이를 해결하기 위해 인공지능 및 머신러닝, 딥러닝의 도입이 요구되고 있다.

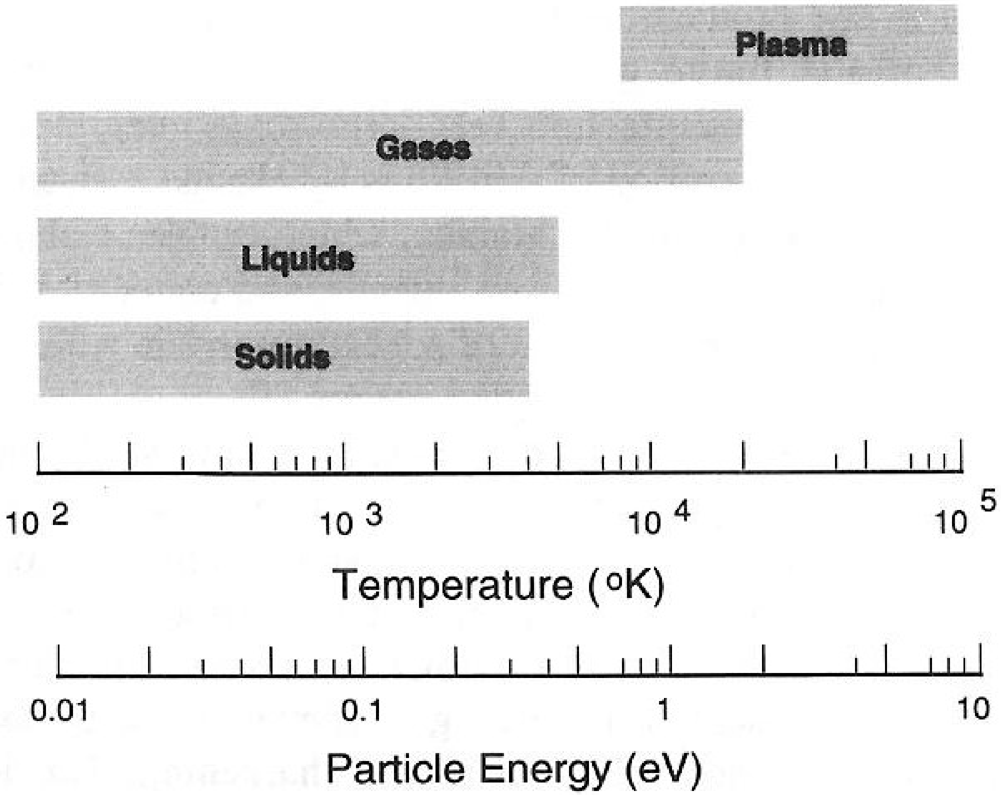


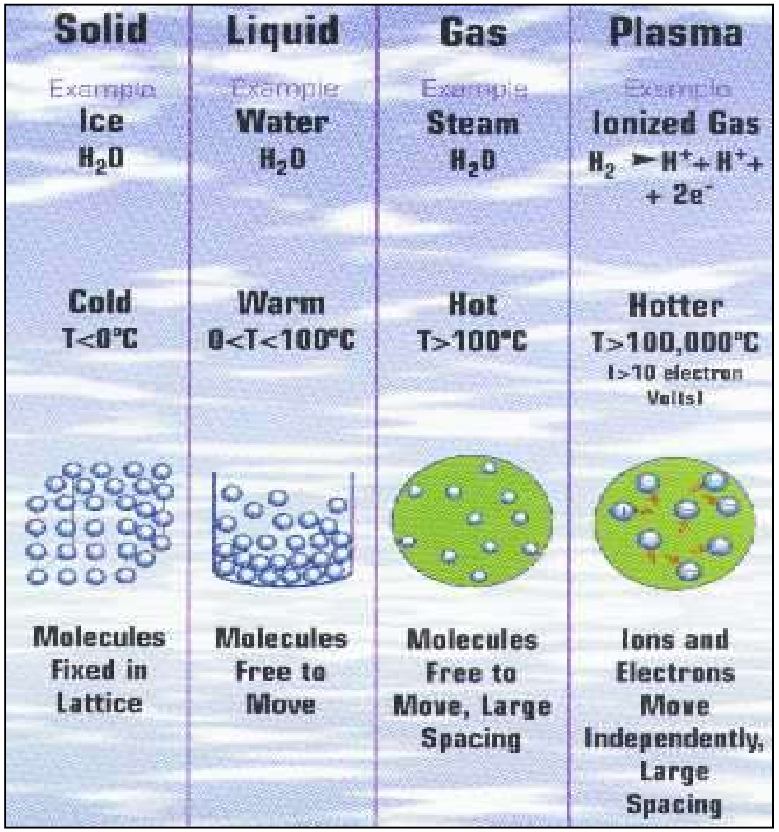
In the past two decades
Storage X 3,000
Computing Power X 10,000
Biological Data X 2,000,000
Data-Driven Bioengineering

플라즈마 공정

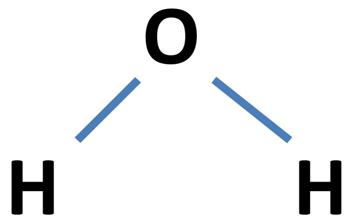
Plasma

- **Plasma (soup of ions, electrons & neutrals)**
  - **4th state of matter**
  - **Ionized Gas**





플라즈마내의 분자, 원자, 라디칼, 이온, 전자



분자 (Molecules)

원자 (Atom)

$$H \longrightarrow H^+ + e^-$$

$$O \longrightarrow O^+ + e^-$$

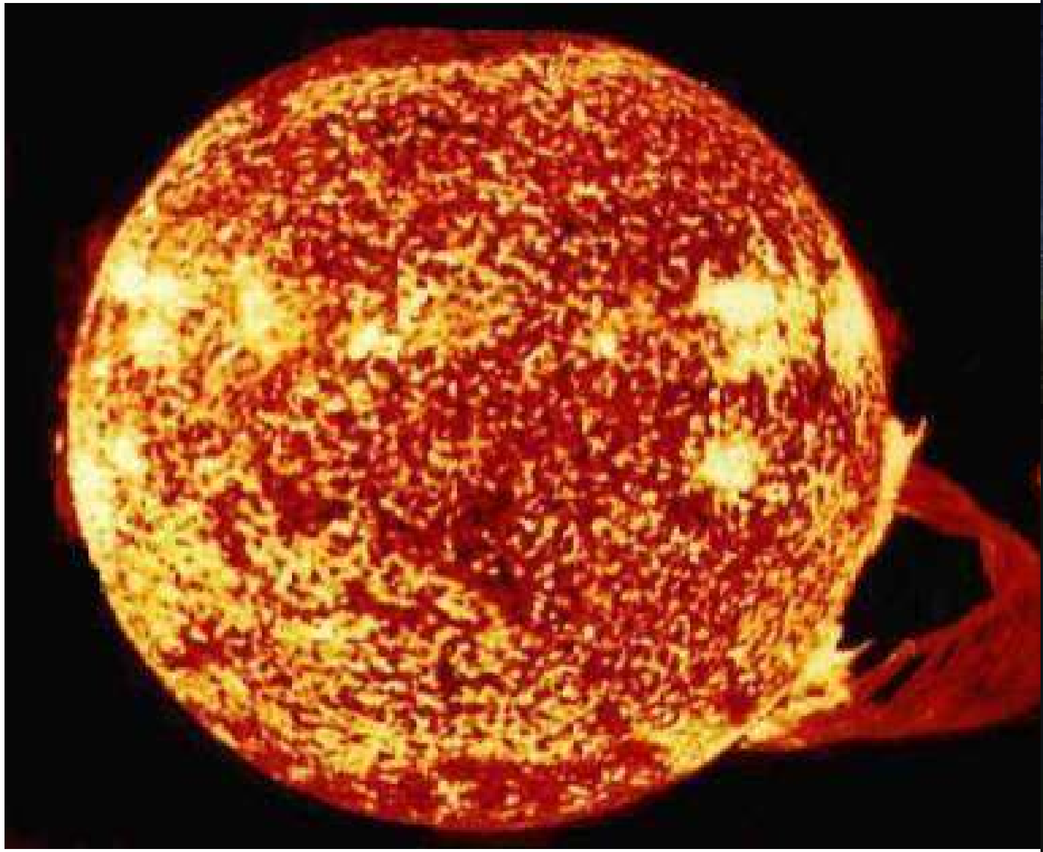
라디칼 (Radicals)

$$O{-}H \longrightarrow OH^+ + e^-$$

이온 (Ions)

전자 (Electrons)

자연의 플라즈마



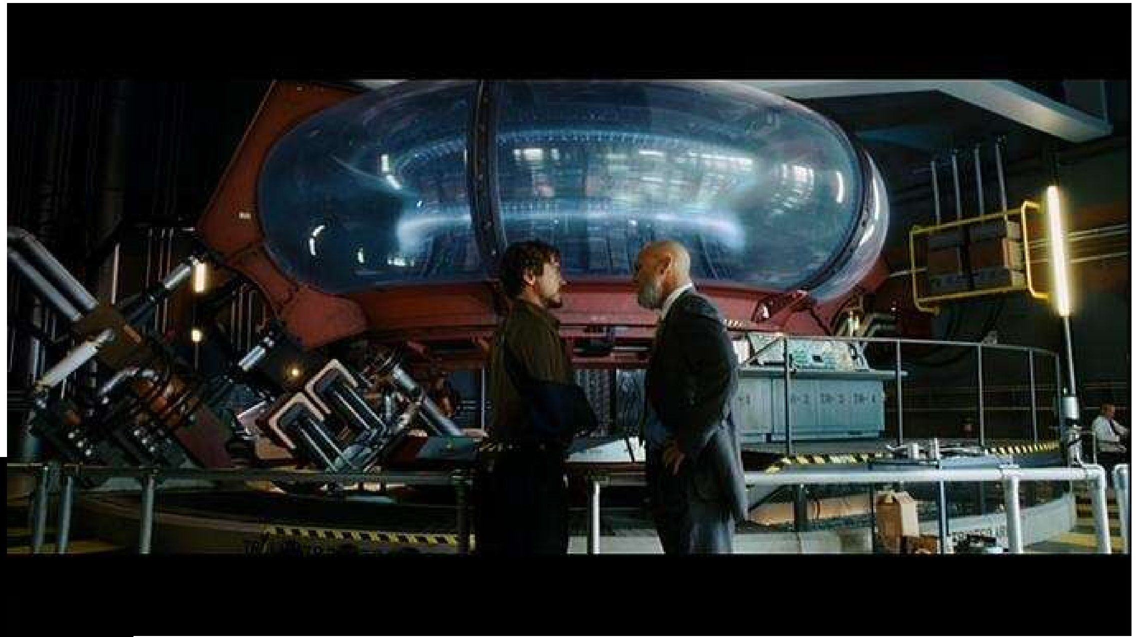
SUNG KYUN KWAN UNIVERSITY

영화속의 플라즈마

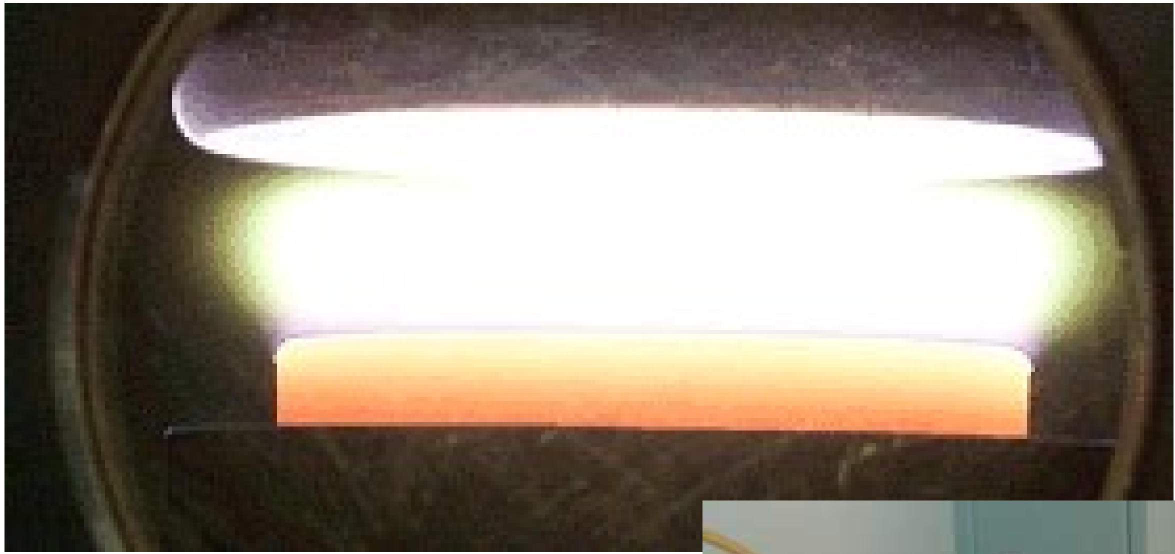
Back to the future (1985)

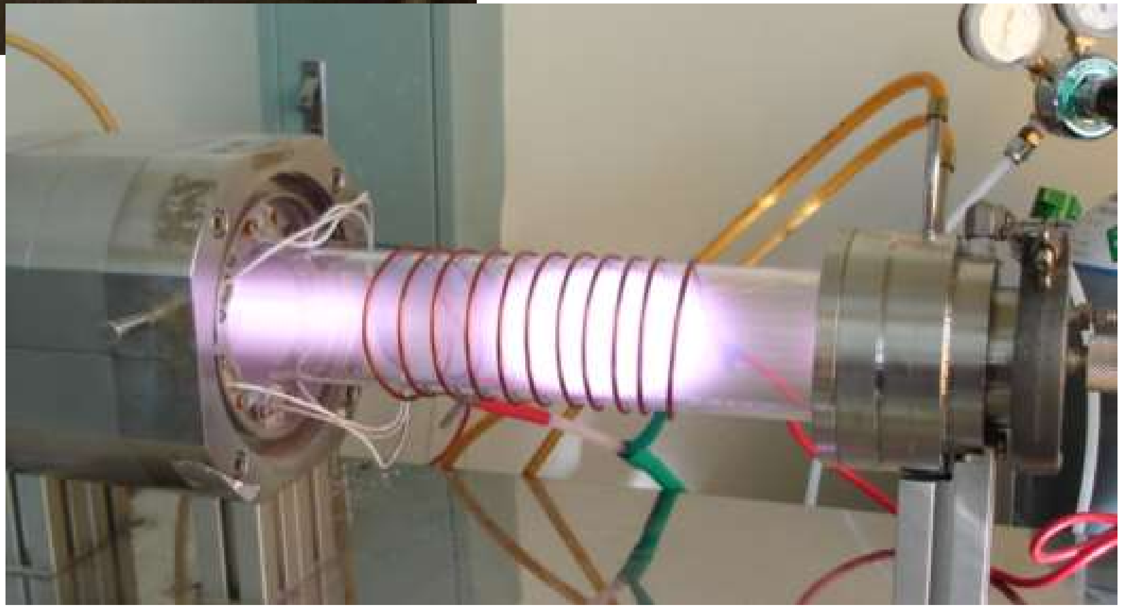
영화속의 플라즈마

Ironman (2008)



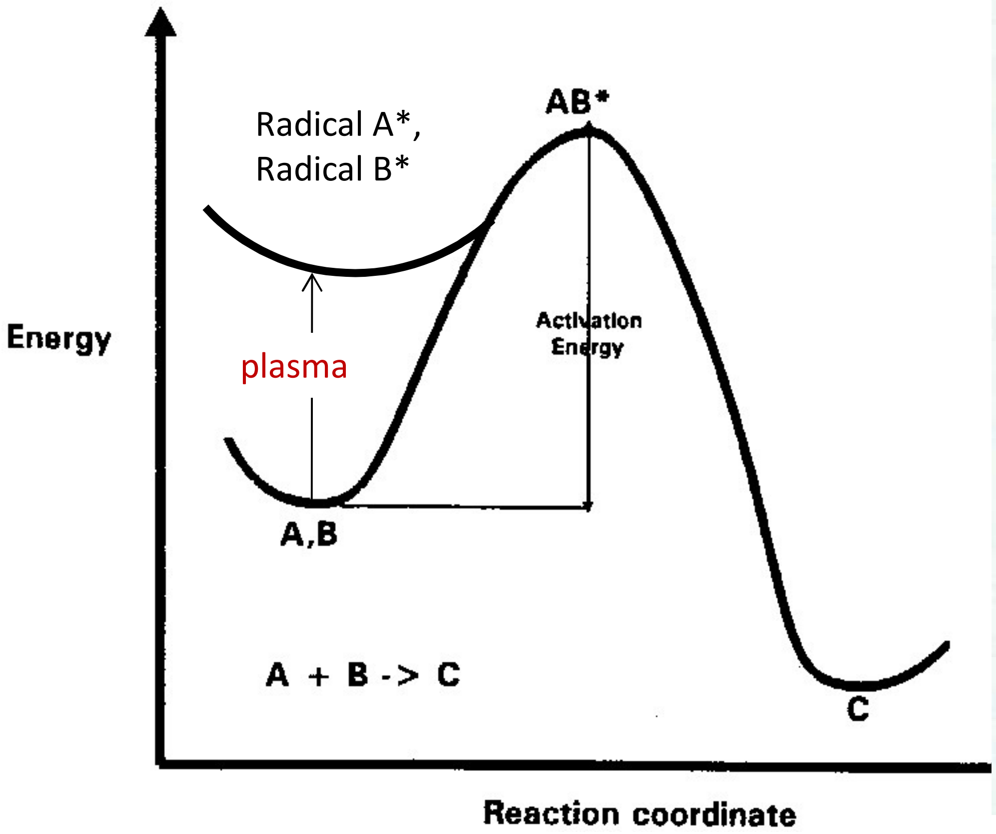
반도체공정에 사용되는 플라즈마



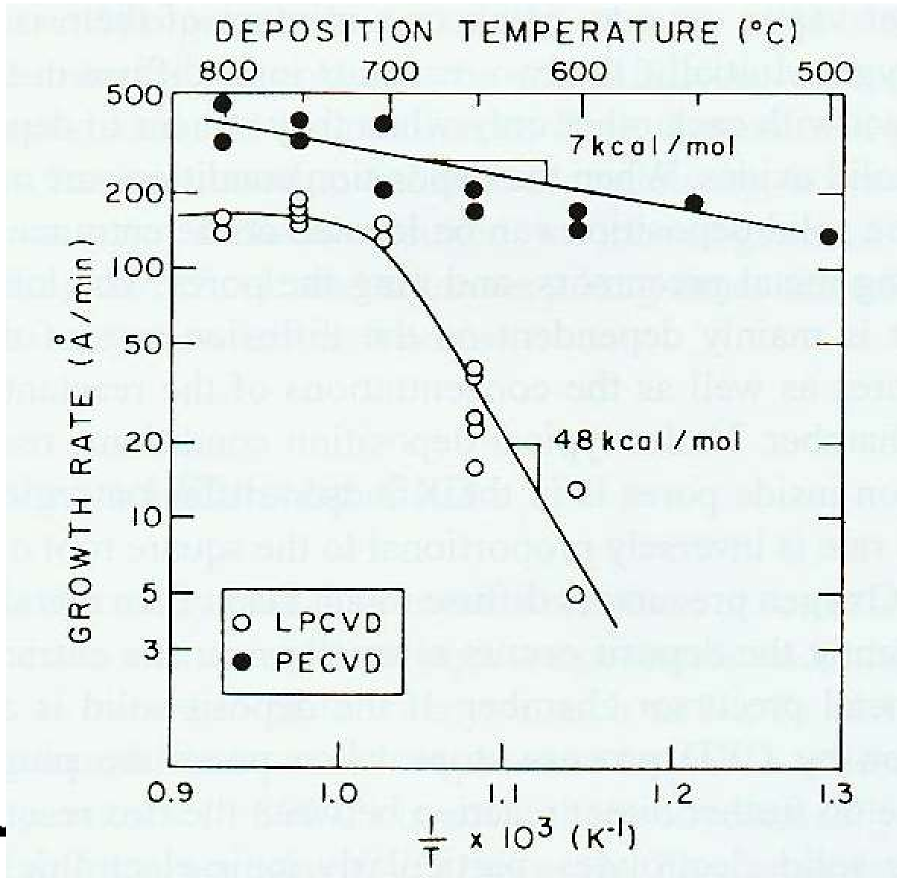


- 진공내에서 전기장을 형성하여 플라즈마 형성

Useful Plasma: Energy Barrier Reduction



Arrhenius Plot of CVD and PECVD



- **Radicals in plasma reduces energy barrier.**
- **Plasma makes low temperature process possible.**

SUNG KYUN KWAN UNIVERSITY

Process Development for Low-T Processes

◆ Plasma sources
- Density control
  . CCP
  . ICP
  . Microwave
  . Atmospheric
  …

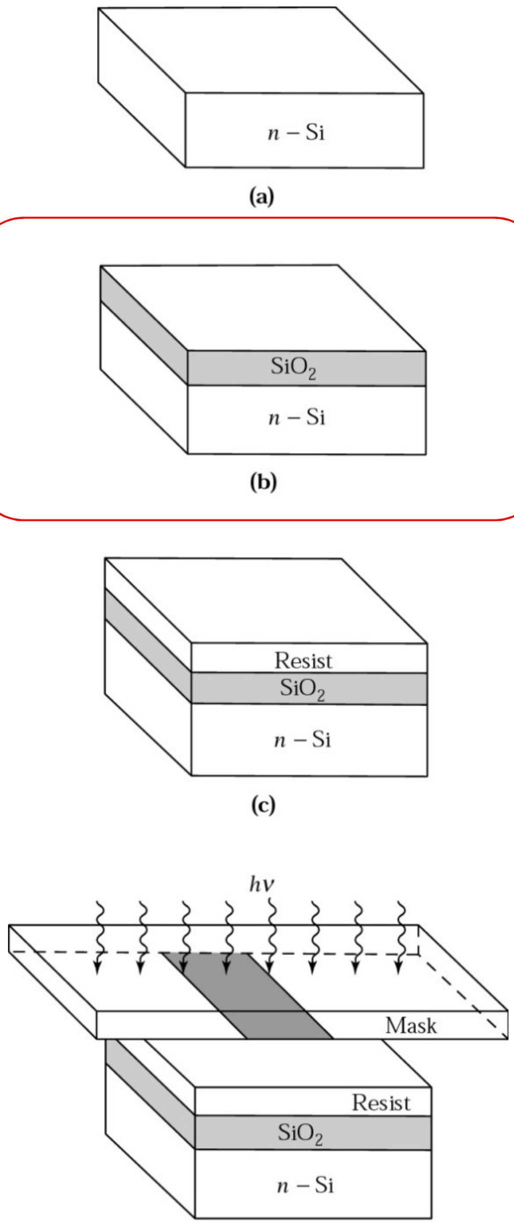
◆ Precursors
- Chemical bonding (ex.)
  . Si-H
  . Si-Cl
  . Si-C
  . Si-N
- Molecular weight
- Chemical structure

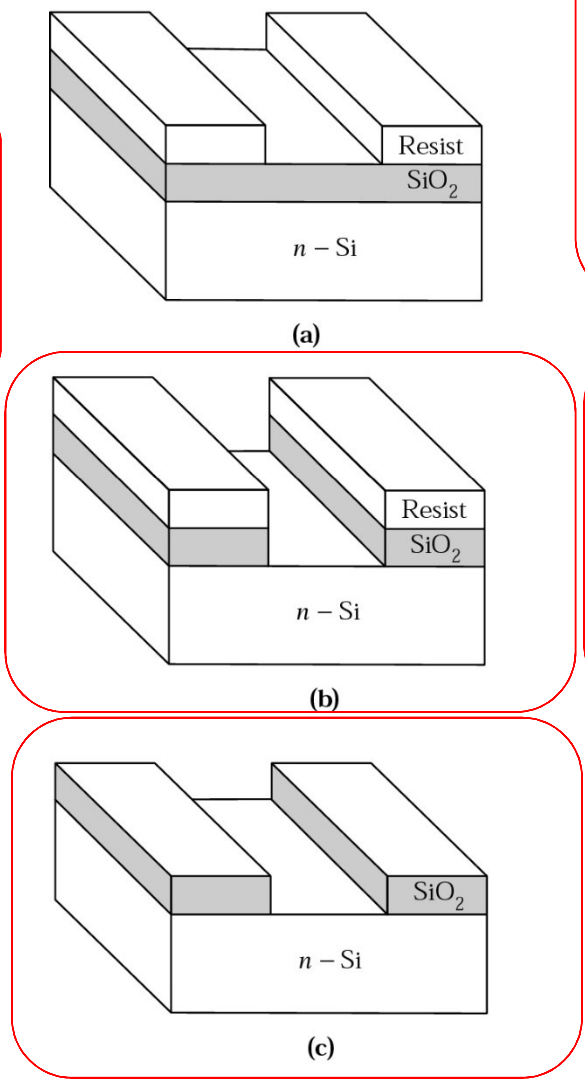
◆ Processes
- Temperature
- Pressure
- Composition
- Additives
- Flow control
…

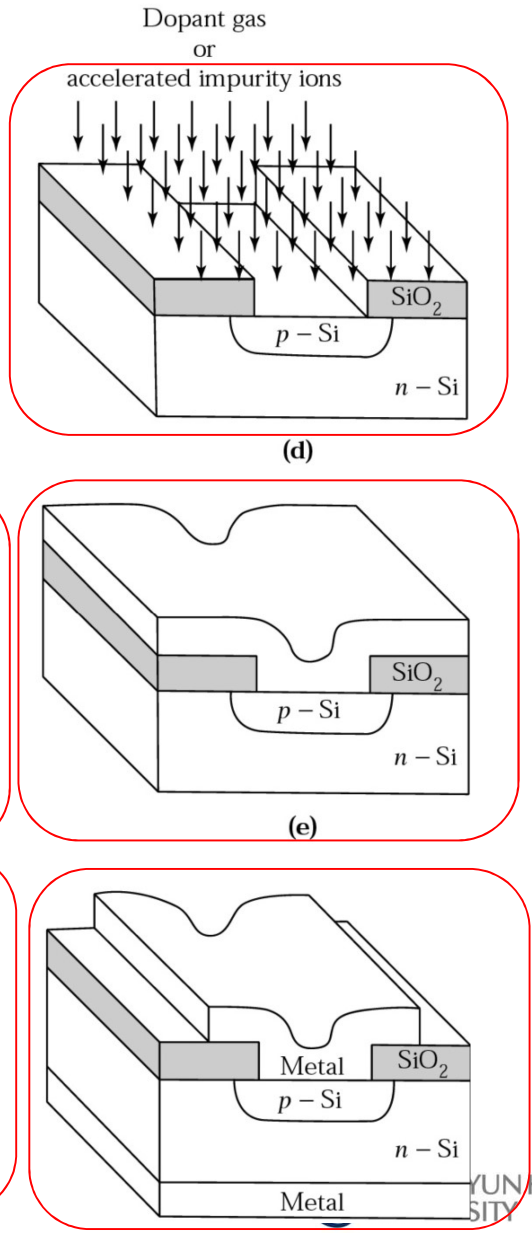
Processes in Semiconductor Device Fabrication

	Thermal Processes	Wet Chemical Processes	Plasma Processes	Physical/ Mechanical Processes
Thin Films	Thermal oxidation Epitaxial Evaporation	Electroplating	Sputtering PECVD HDP-CVD	Sputtering
Lithography	Baking	UV exposure Developing	Light source	Spin Coating Optics
Film Removal	-	Wet etching	Plasma etching	CMP
Cleaning	-	Wet cleaning	Plasma ashing	Ultrasonic
Doping	Diffusion	-	Plasma implantation	Ion implantation

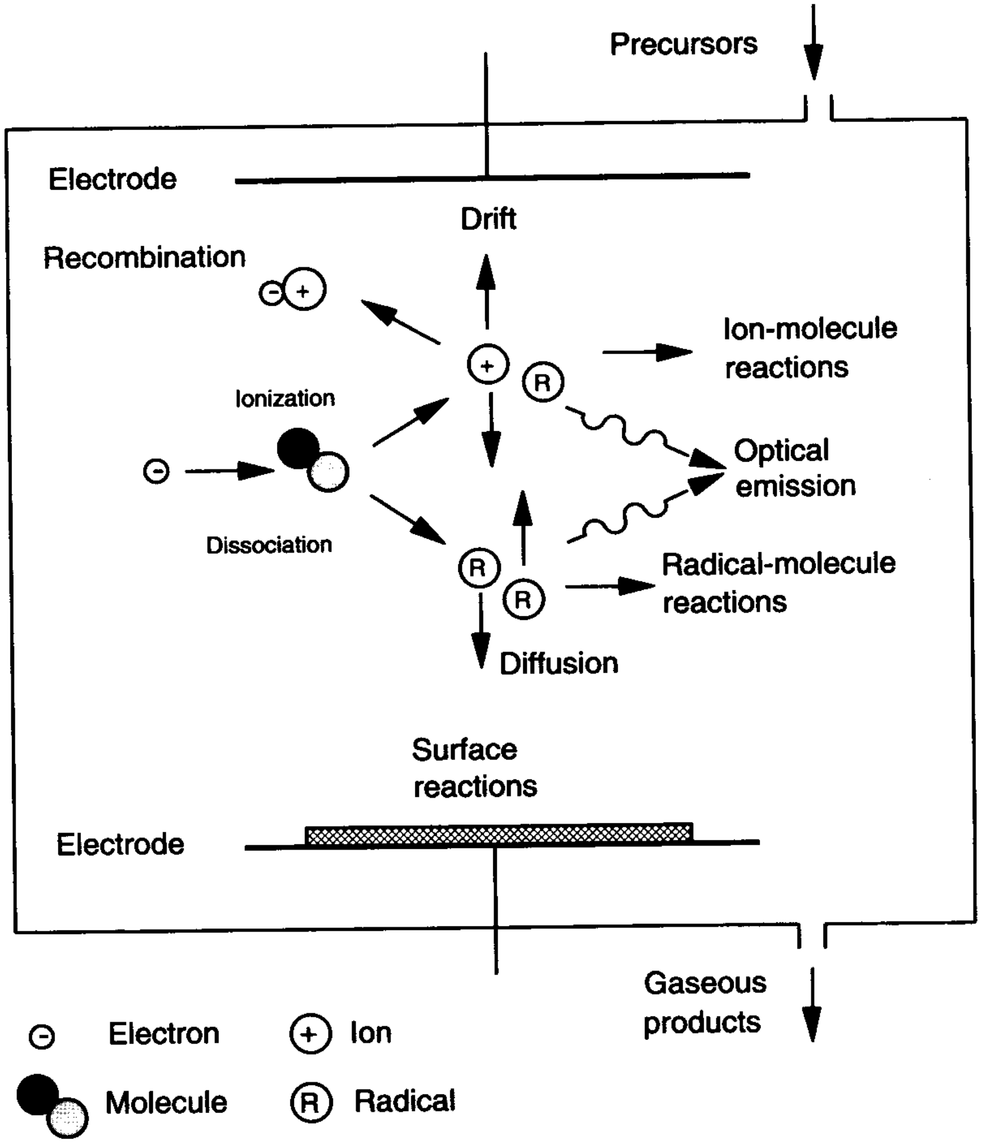
Process Overview and Plasma Processes







Reactions in Plasma



Homogeneous Reactions

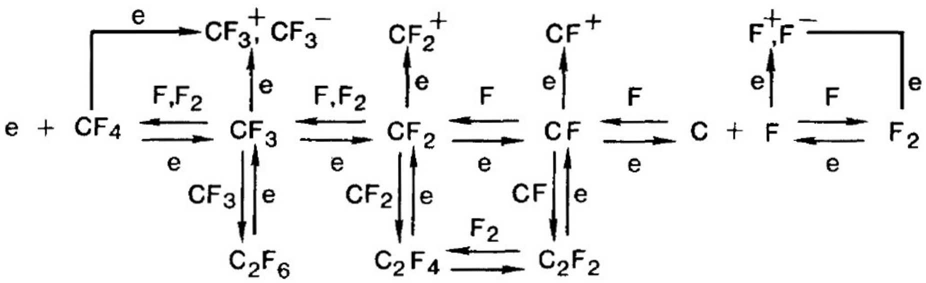
- Recombination of Ions:
- Charge Transfer:
- Transfer of heavy reactants:
- Radical-Molecule Reaction:
  - Electron Transfer
  - Penning Ionization
  - Attachment of Atoms
  - Recombination of radicals
  - Chemiluminescence

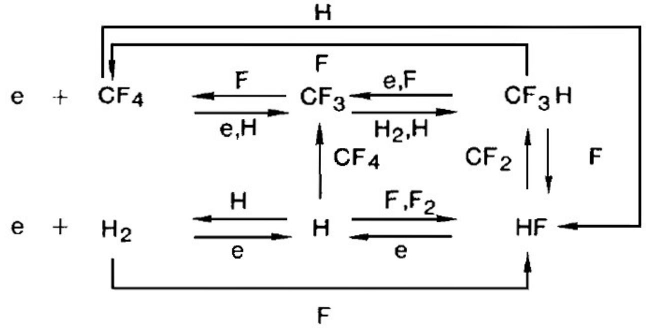
Hetetrogenous Reactions

- Adsorption
- Metastable deexcitation
- Polymerization

플라즈마공정의 복잡성 (플라즈마 내 화학반응)

- CF_4 플라즈마에서 발생하는 Ions, Radicals, Electrons(예시)





- O_2 플라즈마내의 화학반응

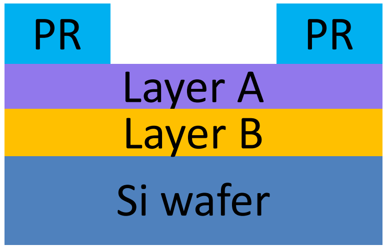
Reaction	k	σ_{max} (cm^2)
Ionization		
1. $e + O_2 \to O_2^+ + 2e$		2.72×10^{-16}
2. $e + O \to O^+ + 2e$		1.54×10^{-18}
Dissociative ionization		
3. $e + O_2 \to O^+ + O$		1.0×10^{-16}
Dissociative attachment		
4. $e + O_2 \to O^- + O$		1.41×10^{-18}
5. $e + O_2 \to O^- + O + e$		4.85×10^{-19}
Dissociation		
6. $e + O_2 \to 2O + e$		2.25×10^{-18}
Metastable formation		
7. $e + O_2 \to O_2(^1\Delta_g) + e$		3.0×10^{-20}
Charge transfer		
8. $O^+ + O_2 \to O_2^+ + O$	2×10^{-11} cm^3/sec	
9. $O_2^+ + O \to O + O_2$		8×10^{-16}
10. $O_2^+ + O_2 \to O_3^+ + O$		1×10^{-16}
11. $O_2^+ + 2O_2 \to O_4^+ + O_2$	2.8×10^{-30} cm^6/sec 2.5×10^{-14} cm^3/sec at E/p = 20 V/cm torr	
12. $O^- + O_2 \to O_2^- + O$	3.4×10^{-12} cm^3/sec at E/p = 45 V/cm torr	
13. $O^- + O_3 \to O_3^- + O$	5.3×10^{-10} cm^3/sec	
14. $O^- + 2O_2 \to O_3^- + O_2$	$1.0 \pm 0.2 \times 10^{-30}$ cm^6/sec	
15. $O_2^- + O \to O^- + O_2$	5×10^{-10} cm^3/sec	
16. $O_2^- + O_2 \to O_3^- + O$		$< 10^{-18}$
17. $O_2^- + O_3 \to O_3^- + O_2$	4.0×10^{-10} cm^3/sec	
18. $O_2^- + 2O_2 \to O_4^- + O_2$	3×10^{-31} cm^6/sec	
19. $O_3^- + O_2 \to O_2^- + O_3$		4×10^{-17}
20. $O_4^- + O \to O_3^- + O_2$	4×10^{-10} cm^3/sec	
21. $O_4^- + O_2 \to O_2^- + 2O_2$	6×10^{-15} cm^3/sec	
Detachment		
22. $O^- + O \to O_2 + e$	3.0×10^{-10} cm^3/sec	
23. $O^- + O_2 \to O + O_2 + e$		7×10^{-16}
24. $O^- + O_2(^1\Delta_g) \to O_3 + e$	$\sim 3 \times 10^{-10}$ cm^3/sec	
25. $O_2^- + O \to O_3 + e$	5.0×10^{-10} cm^3/sec	
26. $O_2^- + O_2 \to 2O_2 + e$		7×10^{-16}
27. $O_2^- + O_2(^1\Delta_g) \to 2O_2 + e$	$\sim 2 \times 10^{-10}$ cm^3/sec	
Electron-ion recombination		
28. $e + \left\{\begin{matrix} O \\ O_2^+ \\ O_3^+ \\ O_4^+ \end{matrix}\right\} \to \left\{\begin{matrix} O \\ 2O \\ O + O_2 \\ 2O_2 \end{matrix}\right\}$	$\lesssim 10^{-7}$ cm^3/sec	
Ion-ion recombination		
29. $\left\{\begin{matrix} O^- \\ O_2^- \\ O_3^- \\ O_4^- \end{matrix}\right\} + \left\{\begin{matrix} O^+ \\ O_2^+ \\ O_3^+ \\ O_4^+ \end{matrix}\right\} \to \left\{\begin{matrix} O \\ O_2 \end{matrix}\right\}$	$\sim 10^{-7}$ cm^3/sec	
Atom recombination		
30. $2O + O_2 \to 2O_2$	2.3×10^{-33} cm^6/sec	
31. $3O \to O + O_2$	1.5×10^{-34} cm^6/sec	
32. $O + 2O_2 \to O_2 + O_2$	1.9×10^{-35} exp $(2100/RT)$ cm^6/sec	
33. $O + O_3 \to 2O_2$	2.0×10^{-11} exp $(-4790/RT)$ cm^3/sec	
34. $O \xrightarrow[\text{wall}]{} O_2$	$\gamma = 1.6 \times 10^{-4}$ to 1.4×10^{-2} (T = 20 − 600 °C)	

Plasma Monitoring: Endpoint Detection

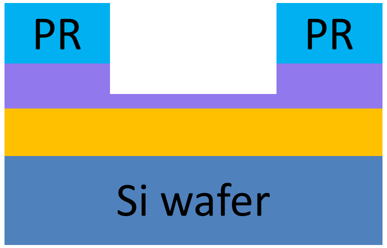
- To control etching rate is important for IC manufacturing.
- It is necessary to avoid incomplete-etching & over-etching.
- Decreasing feature size, it becomes more and more challenging to detect endpoint.

$$SiN_x + CF_4 + O_2 \rightarrow \underbrace{SiF_x + F + CF_x}_{\text{표면}} + CN + FCN + C_2N_2 + N$$

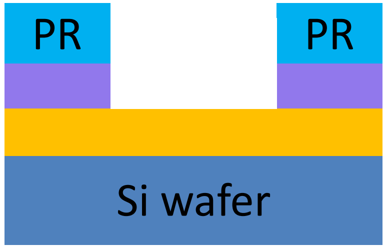
플라즈마 내부



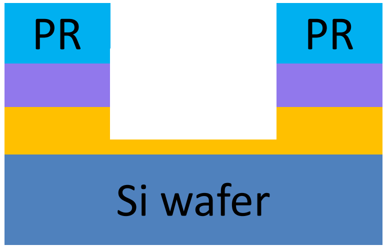
Before etching



After etching Incomplete-etching (X)



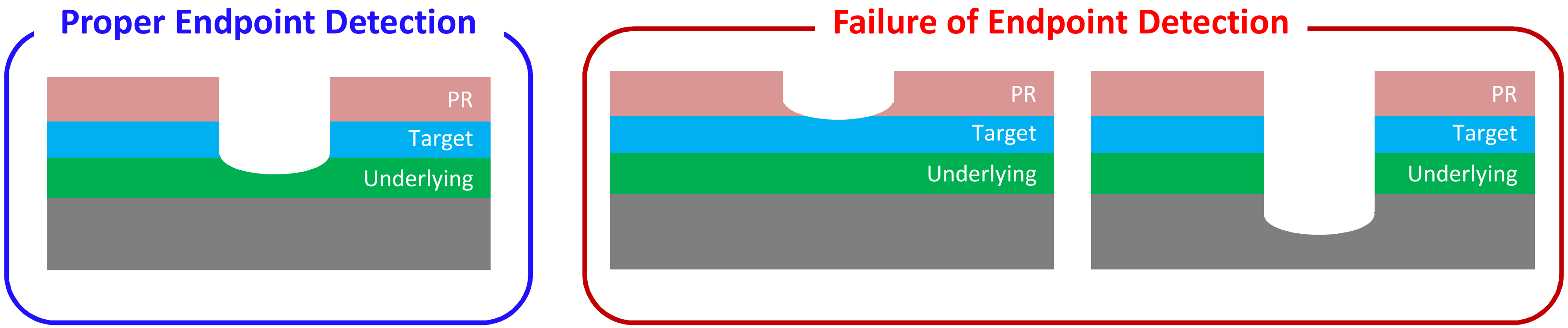
After etching Complete-etching (O)



After etching Over etching (X)

- **It is critical to end the plasma etching process at target depth.**
- **Sensitive plasma monitoring required.**

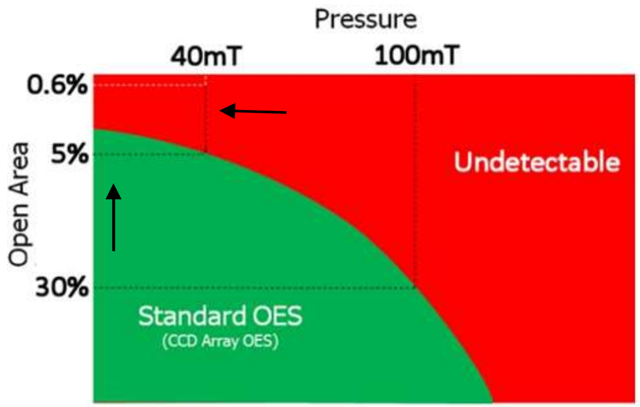
Endpoint Detection



Target layer etching

Incomplete-etching & Over-etching

- Need to stop the etching process at a proper moment, known as **endpoint detection(EPD)**
- Failure of endpoint detection → Device failure & Yield reduction



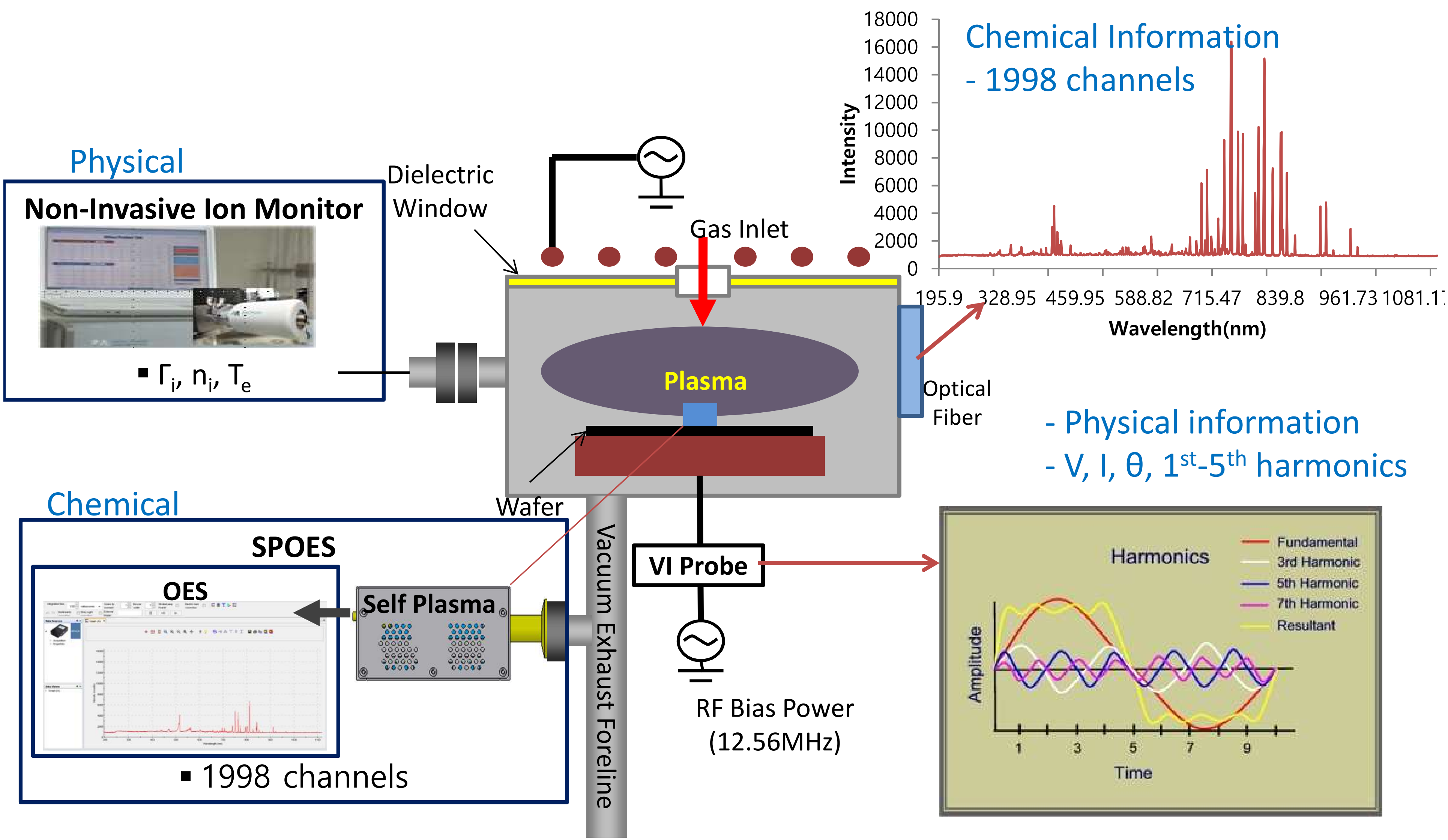
https://www.orbotech.com/spts/about/resources/tech-insights/mems-tech-insights/plasma-etch-end-point-control

- **EPD issue**

Cannot detect endpoint in **small open area** and **low pressure**

→ Need to enhance sensitivity of signal using multivariate analysis

Non-Invasive Plasma Monitoring Tools



반도체 공정데이터 분석의 예

Plasma Monitoring: Principal Component Analysis (PCA)

Measured variable 'x, y, z'

measuring	x	y	z
1	x_1	y_1	z_1
2	x_2	y_2	z_2
. .		Data	
n	x_n	y_n	z_n

t_i : new independent variables

$$\begin{aligned} t_1 &= p_{11}x + p_{12}y + p_{13}z \\ t_2 &= p_{21}x + p_{22}y + p_{23}z \\ t_3 &= p_{31}x + p_{32}y + p_{33}z \end{aligned}$$

(Matrix) $t_i = Xp_i$

Contribution ratio

measuring	t_1	t_2	t_3
1			
2		Principal Component Score	
. .			
n			

Increasing information intensity of t_i

Maximizing s_{t_i}(variance of t_i)

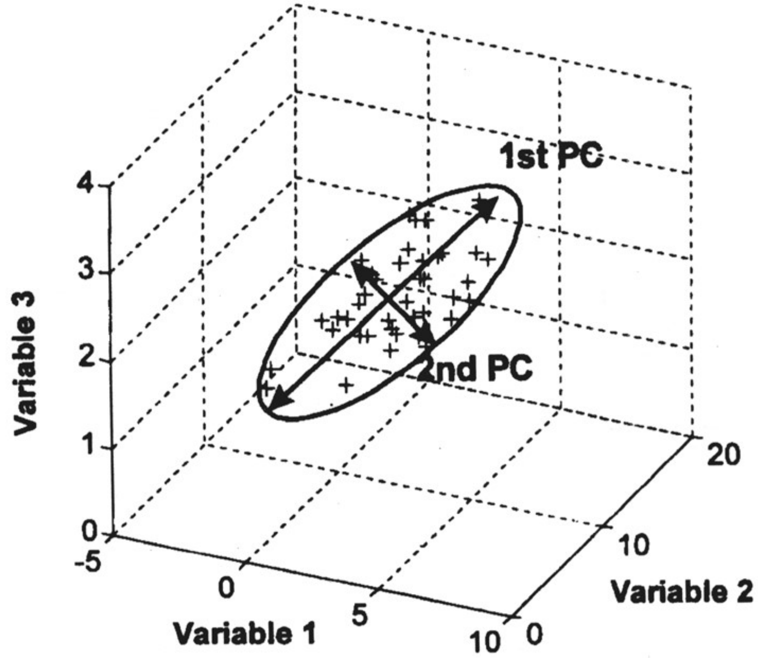
Constraint : $p_{i1}^2 + p_{i2}^2 + p_{i3}^2 = 1$

$$Sp_i = \lambda p_i \qquad S = \left(\frac{1}{n-1}\right) X^T X$$

Finding λ & p from S(variance-covariance matrix)

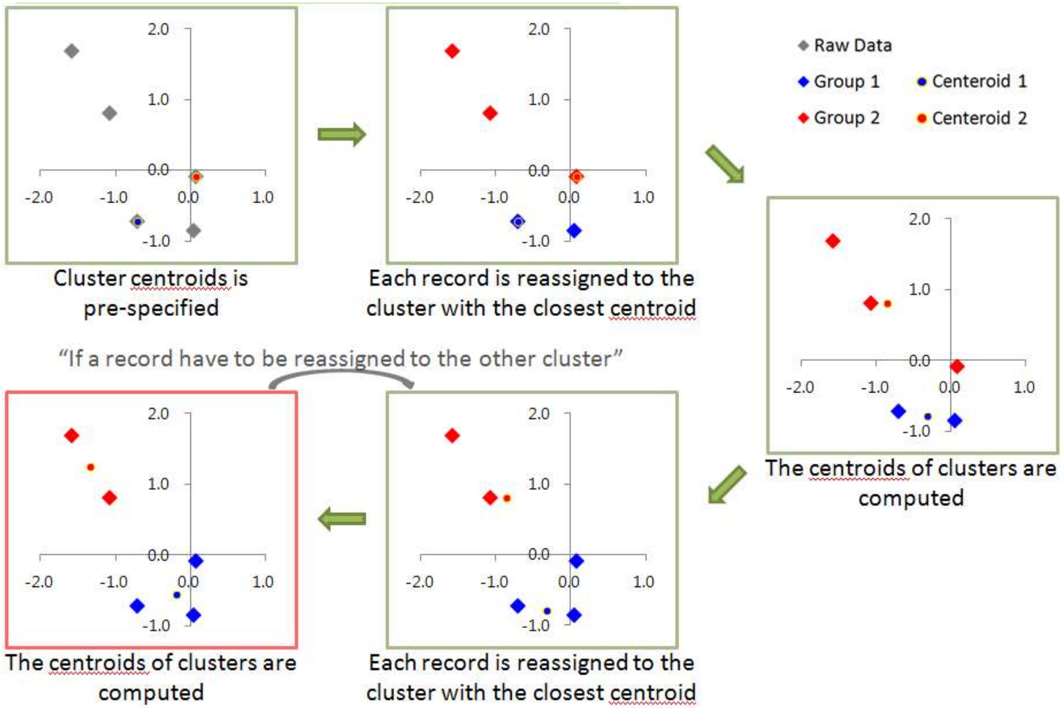
$$|S - \lambda I| = 0$$

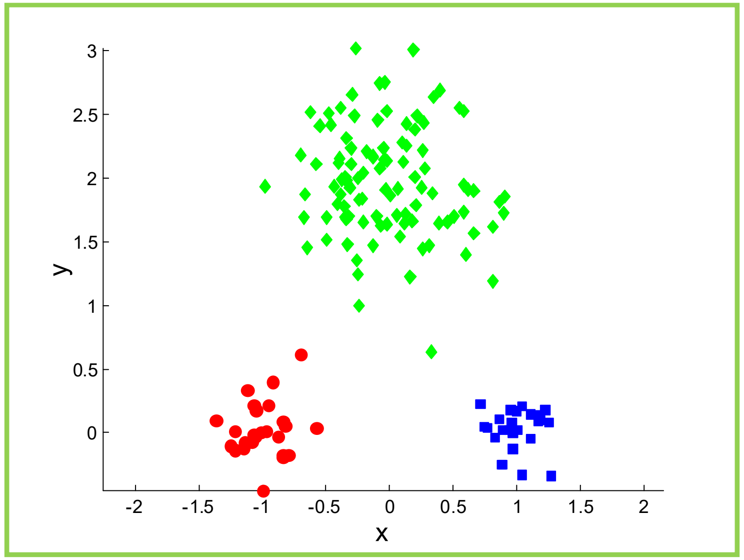
- λ (eigenvalue) = Relative information intensity
- p_i (eigenvector) = loading vector
  = coefficient of principal components
- t_i = score vector



K-Means Clustering Analysis

- K-means Cluster Analysis (KMC)





$$R = \frac{\sum_{i=1}^{n}\sum_{j=1}^{2}(x_i - \bar{x}_j)(x_i - \bar{x}_j)^T a_{ij}}{\sum_{i=1}^{n}(x_i - \bar{x}_0)(x_i - \bar{x}_0)^T}$$

$\bar{x}_0$: Centeroid of all records

$\bar{x}_1$: Centeroid of cluster 1

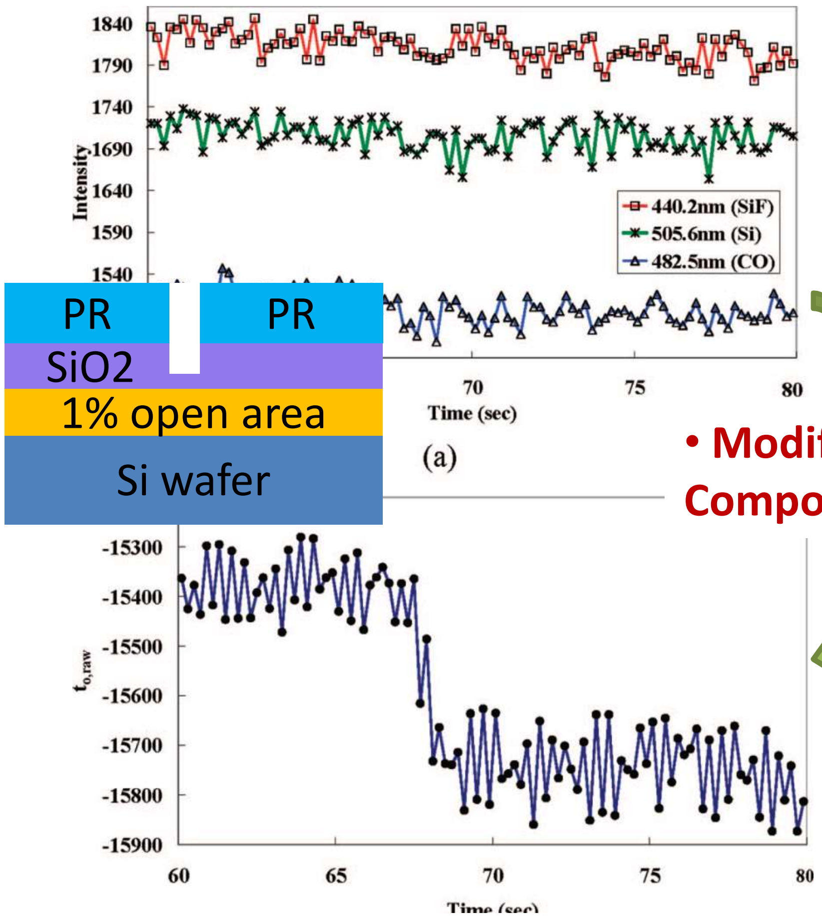
$\bar{x}_2$: Centeroid of cluster 2

$$a_{ij} = \begin{cases} 1 & \text{if cluster } j \text{ includes record } i \\ 0 & \text{else} \end{cases}$$

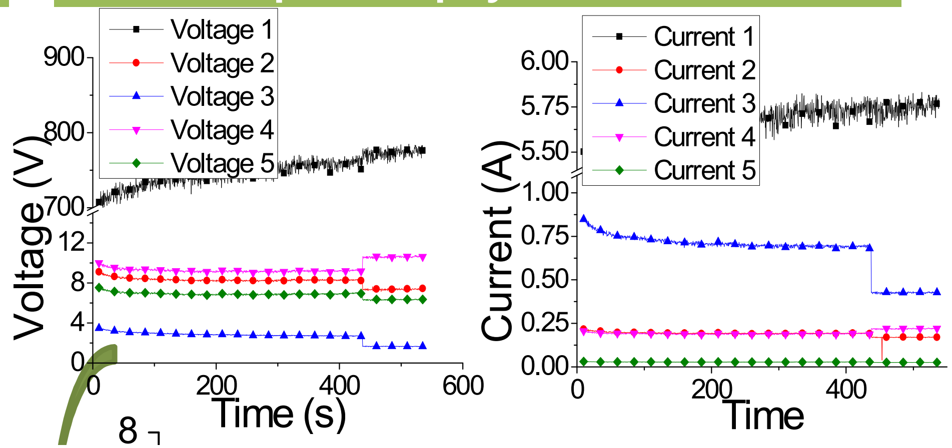
- **Cluster analysis is suitable for fault identification in complex systems**

Plasma Monitoring: VI Probe & PCA

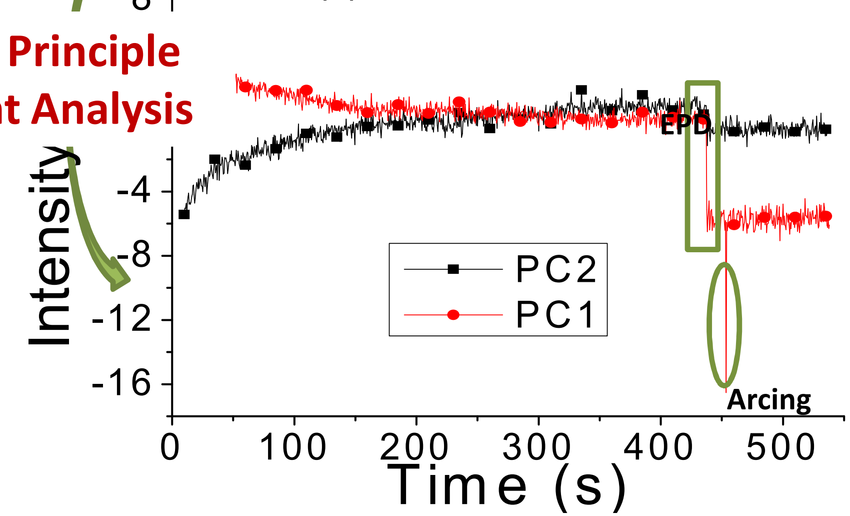
EPD (OES) : chemical information



EPD (VI probe): physical information



• **Modified Principle Component Analysis**



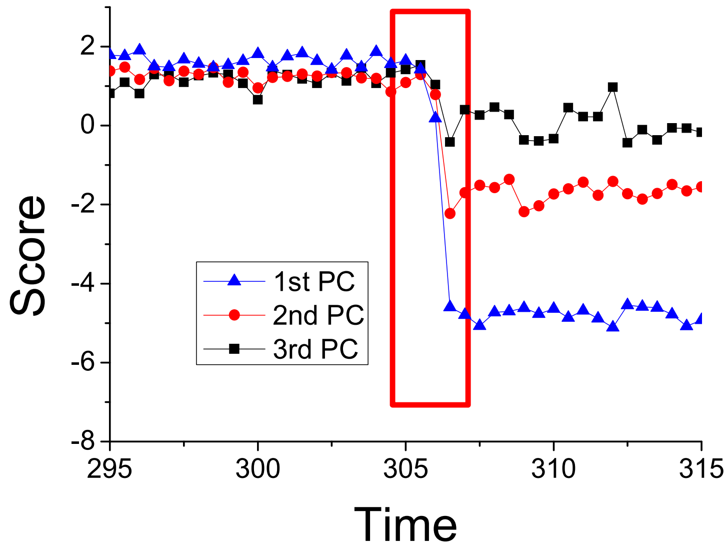
Ind. Eng. Chem. Res, 47, 11, (2008)   Plasma process polym. 10, 850 (2013)

- **Endpoint detection sensitivity improved by PCA algorithm**

SUNG KYUN KWAN UNIVERSITY

Modified PCA with Impedance Monitoring

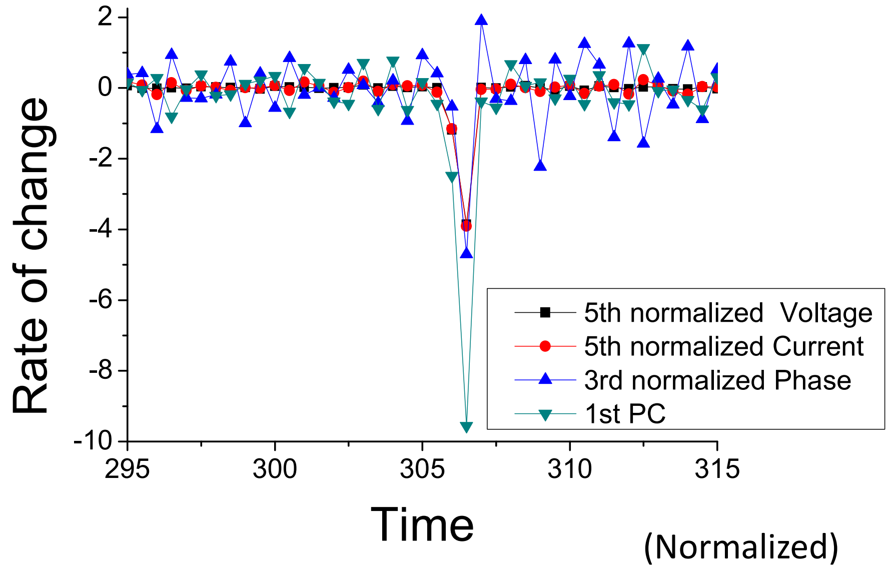
- Plasma Impedance monitoring with PCA



	Contribution Ratio
1^{st} PC	76.40%
2^{nd} PC	16.49%
3^{rd} PC	4.25%

- The 1^{st} principal component (PC1) is chosen : The most sensitive PC

- Comparison : Rate of Change



(Normalized)

	Voltage	Current	Phase	PC
1^{st}	0.82	-2.13	0.68	**-9.56**
2^{nd}	-2.97	-2.92	-2.56	
3^{rd}	-2.84	-2.99	**-4.70**	
4^{th}	2.86	2.82	-4.42	
5^{th}	**-3.85**	**-3.91**	-3.04	

- (about 2 times)

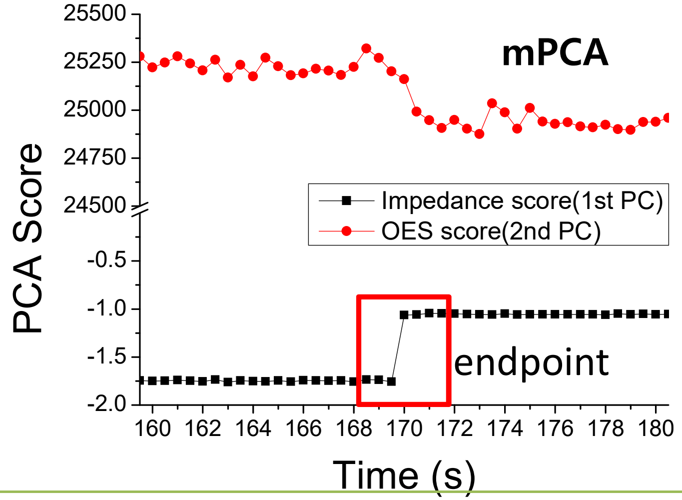
Plasma Process. Polym. 10, 850 (2013)

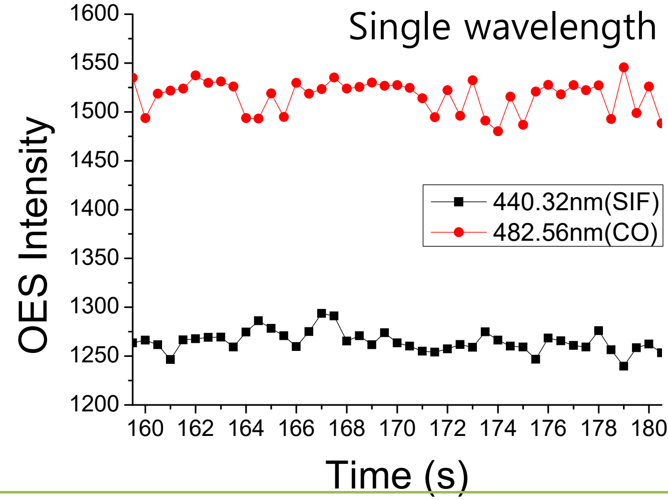
SUNG KYUN KWAN UNIVERSITY

Modified PCA with Impedance Monitoring

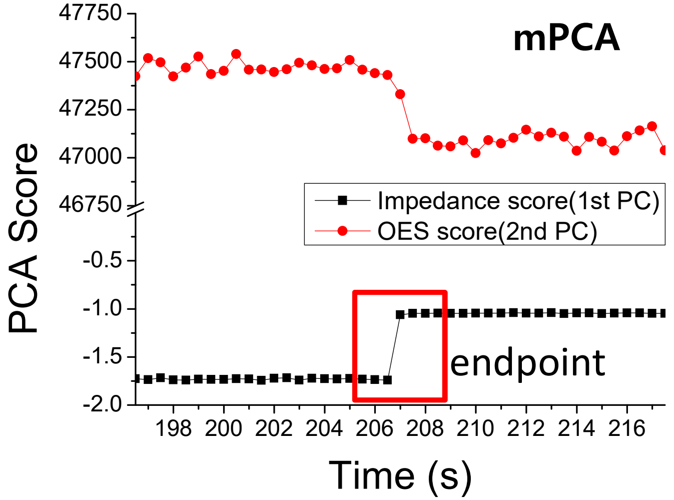
- Target 1.0% SiO_2 area

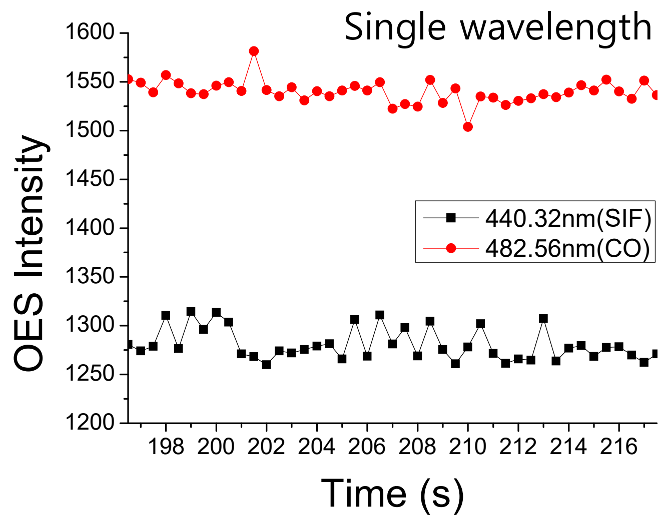
(The most sensitive PC in OES: 2nd PC)





- Target 0.5% SiO_2 area



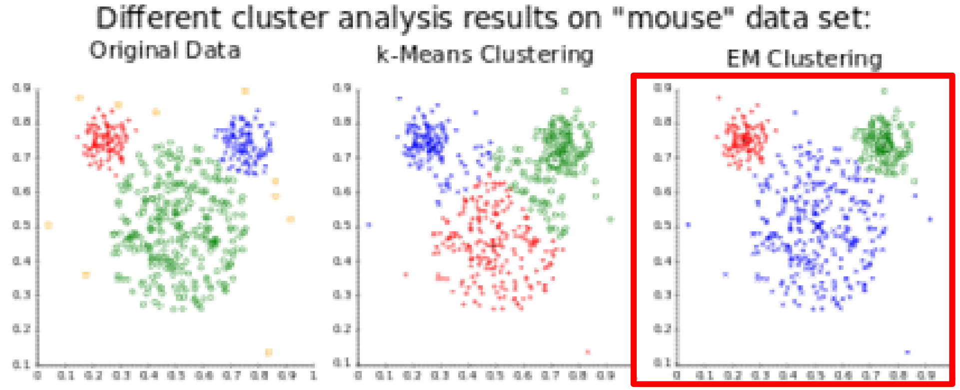


Plasma Process. Polym. 10, 850 (2013)

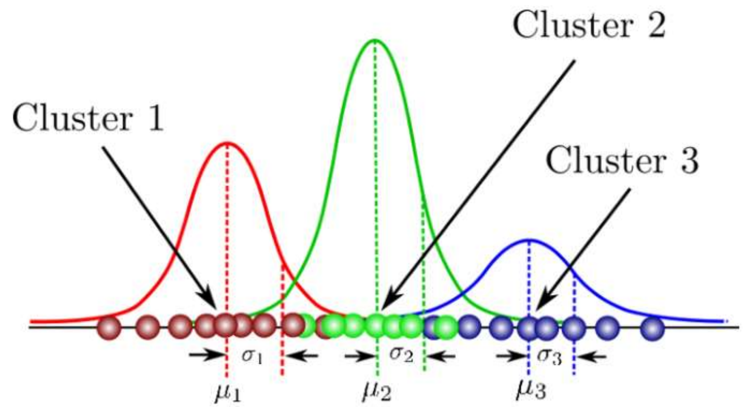
SUNG KYUN KWAN UNIVERSITY

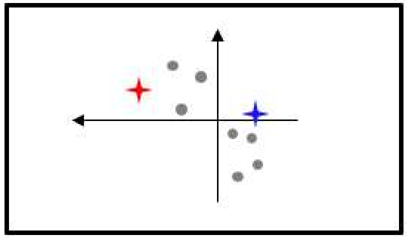
Plasma Monitoring: Cluster Analysis

▪ Gaussian Mixture Model (GMM)

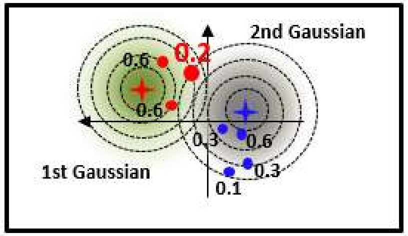


Raw data | K-means cluster analysis | Gaussian mixture model

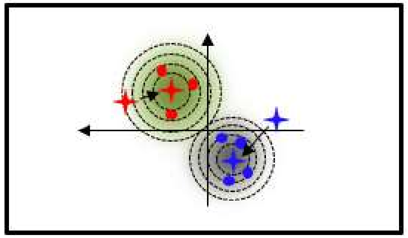




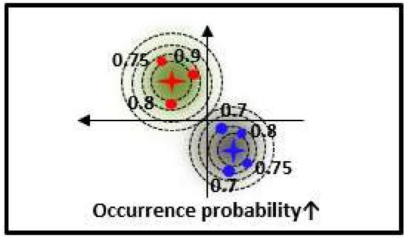
Centroid of gaussian is specified in advance



Expectation step (E-step) : Estimate that records belong to which gaussian

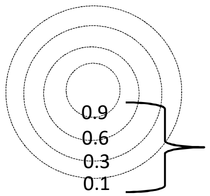


Maximization step (M-step) : Find appropriate centroid and variance of gaussian



The centroids of clusters are computed

➢ **Gaussian distribution parameters**



▶ **Gaussian distribution**

Occurrence probability of record in cluster

- Raw data
- Group 1
- Group 2
- Centroid of gaussian 1
- Centroid of gaussian 2

Sensitivity Enhancement by Select Wavelengths

- **Common spectral lines in SiO_2 etching applications**

  6144 channels → 95 select wavelengths

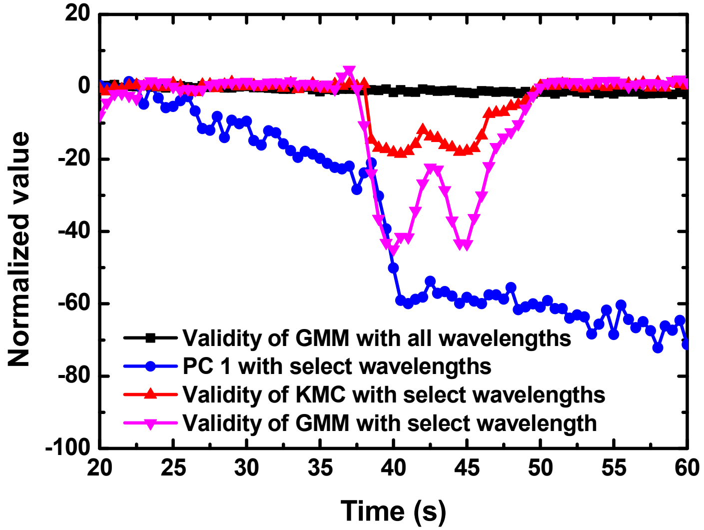
Species	Wavelength (nm)
Ar	434.8, 476.5, 488.0, 696.5, 706.7, 738.4, 750.4, 751.5, 763.5, 772.4, 794.8
C	283.7, 426.7, 732.6
CF	240.0, 247.4, 255.8
CF_2	248.8, 251.9, 259.5, 262.9, 271.1, 275.0, 280.0, 292.1, 321.4
CO	238.9, 269.8, 283.3, 292.5, 302.8, 313.4, 313.8, 325.3, 330.6, 349.3, 451.1, 482.5, 483.5, 519.8, 561.0, 608.0, 662.0
F	623.9, 634.8, 641.4, 677.4, 683.4, 685.4, 685.6, 6870, 690.2, 691.0, 696.6, 703.7, 712.8, 720.2, 733.2, 739.9, 742.6, 755.2, 757.3, 760.7, 775.5, 780.0
O	391.2, 397.3, 407.6, 419.0, 464.9, 615.6, 615.7, 615.8, 645.6, 725.4, 777.2, 844.7
Si	288.2, 504.1, 505.5, 634.7, 637.1
SiF	334.6, 336.3, 436.8, 440.1, 777.0
SiF_2	390.2, 395.5
SiO	229.9, 234.4, 241.4, 248.7, 266.9, 269.4

http://www.verityinst.com/

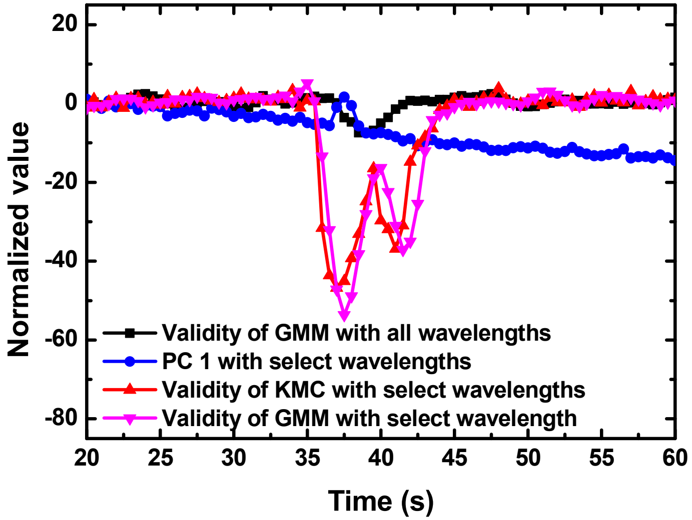
➢ Endpoint detection by using multivariate techniques for 95# select wavelengths

EPD with Select Wavelengths: SiO_2 Etching

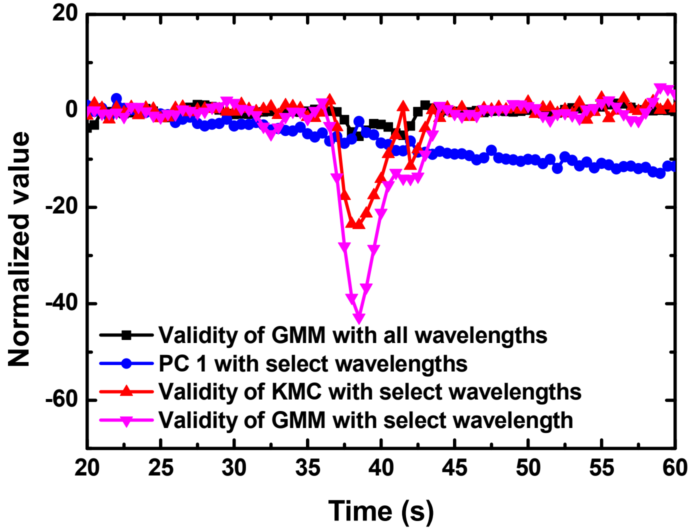
8% area SiO_2



4% area SiO_2

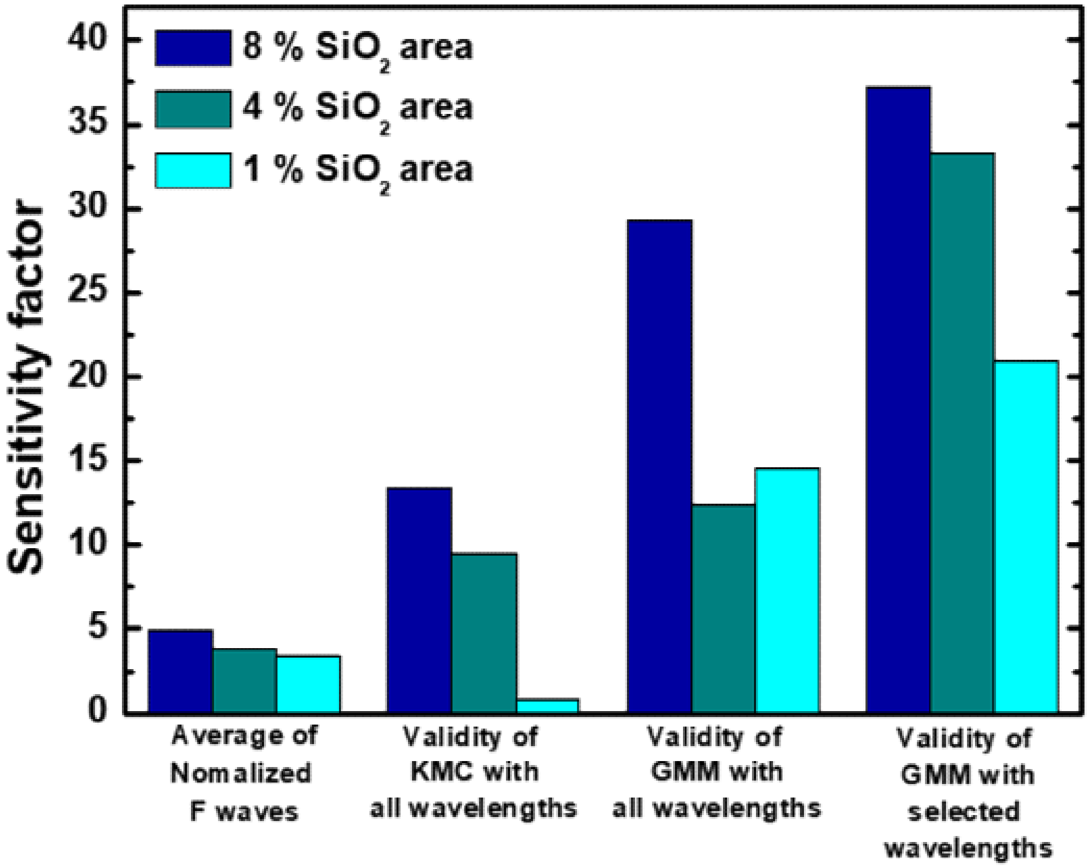


1% area SiO_2



EPD	8%		2%		1%	
GMM with all wavelengths	19.76	O	9.10	O	17.44	O
PCA (1^{st} PC) with select wavelengths	20.36	O	6.41	O	2.93	X
KMC with select wavelengths	29.48	O	32.76	O	16.04	O
GMM with select wavelengths	37.25	O	33.27	O	20.91	O

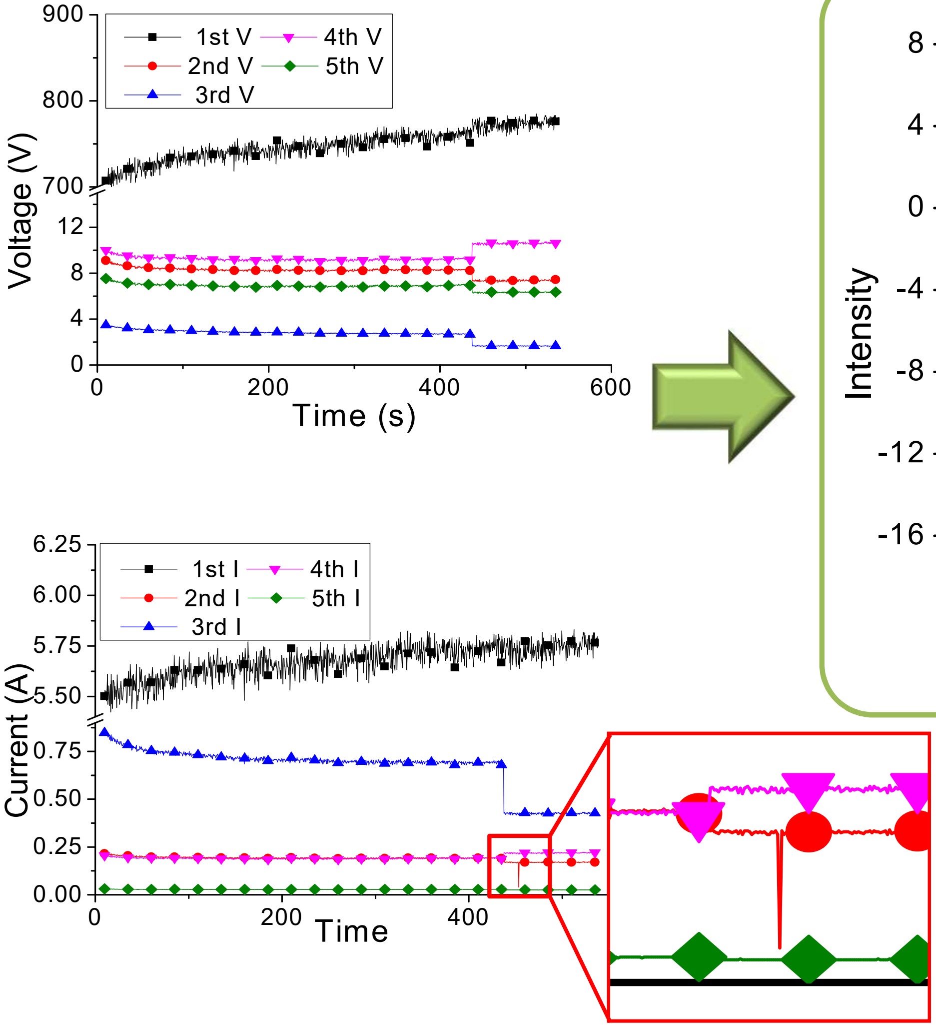
Sensitivity

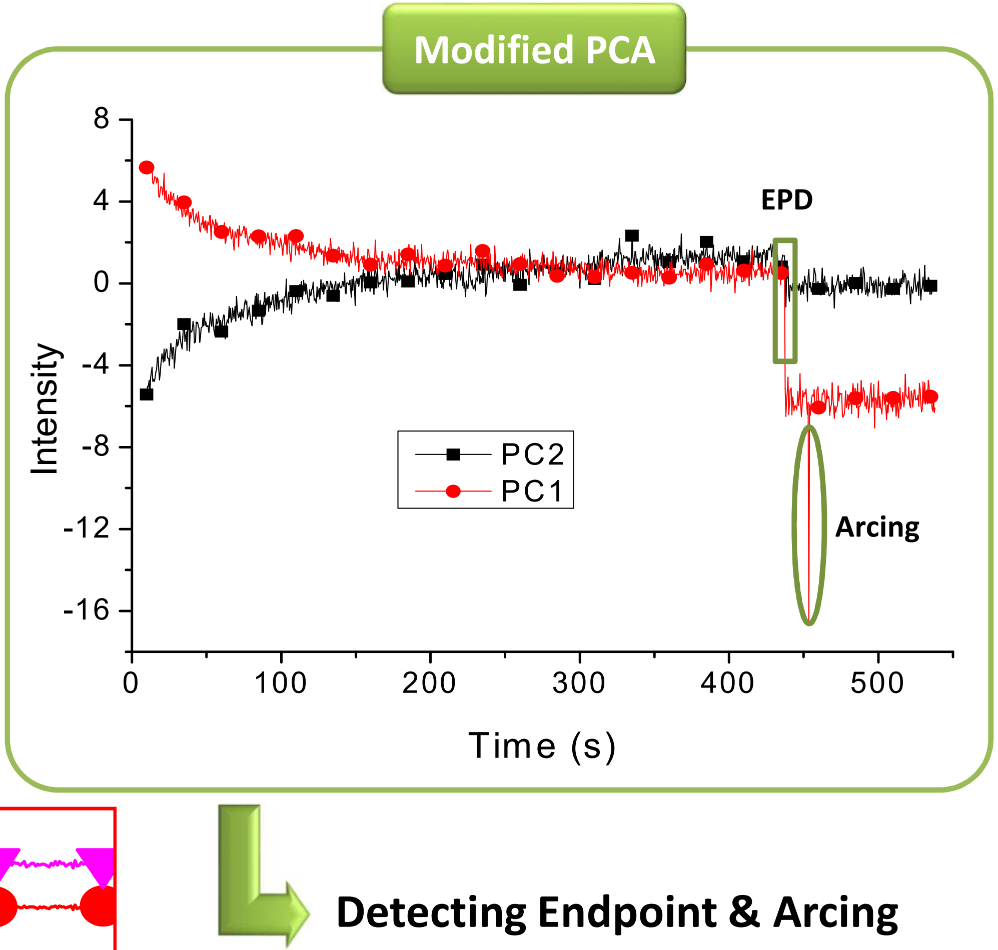


$$Sensitivity\ factor = \frac{|m_{\Delta t_2} - m_{\Delta t_1}|}{s_{\Delta t_1}}.$$

Sensitivity Analysis and Enhancement: Bias Power

- Examples: Endpoint & Arcing Detection

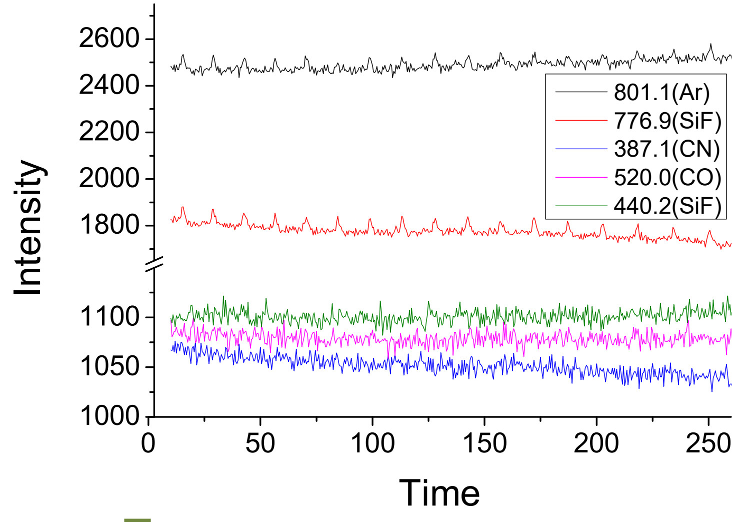




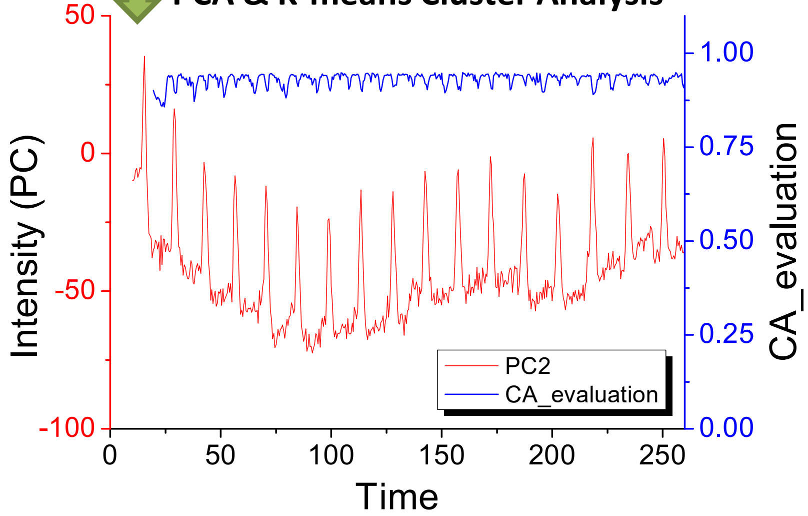
Detecting Endpoint & Arcing

Fault Detection: MFC malfunction

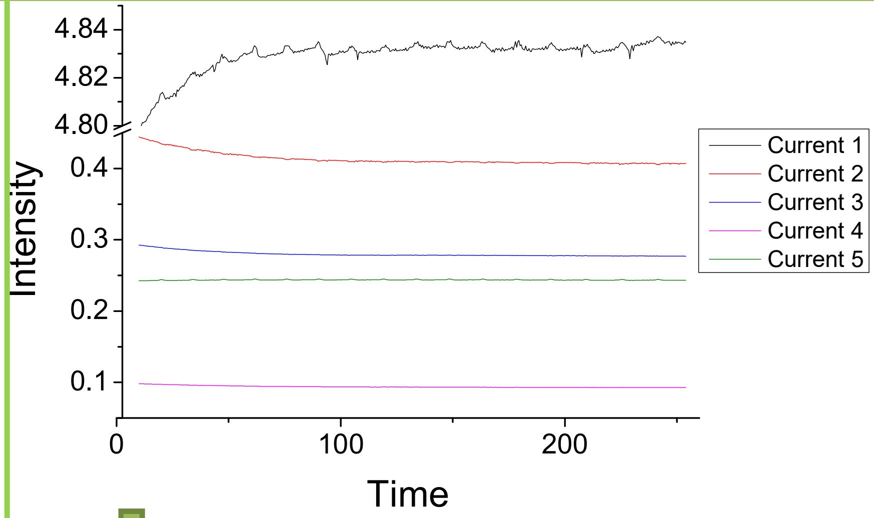
Optical Emission Spectroscopy



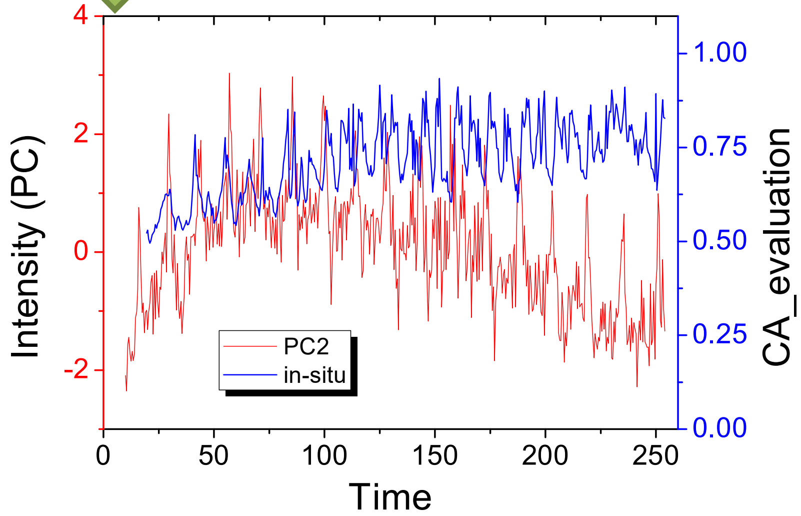
PCA & K-means Cluster Analysis



V-I probe



PCA & K-means Cluster Analysis



SUNG KYUN KWAN UNIVERSITY

Machine Learning for Semiconductor Processing

Area	Process	Measurement (Variable)	Data (size)	Topic	Ref
Fault Detection	**CVD (simulation)**	10 variables (Temperature, Pressure, Flow rate, etc.)	10,000 wafer (6 types classification) 105seceds(0.2 time interval)	1D convolution layer structure for higher speed	*IEEE TRANSACTIONS ON SEMICONDUCTOR MANUFACTURING, VOL. 30, NO. 2, MAY 2017*
Fault Detection	**Monitoring**	OES data(150nm to 1000nm)	2048 wavelengths	DWT(3 layers, soft thresholding)	*Computers and Chemical Engineering 94 (2016) 362–369*
Fault Detection	**Wafer Classification**	112 process variables	76 wafers (60 for training, 16 for test)	LASSO Multi-level LASSO	*IEEE International Conference on Automation Science and Engineering Trieste (2011)*
Fault Detection	**Wafer Classification**	17 variables (flow rate, power impedance, pressure, tuner, etc.)	100 times of 107 wafers (97 for training, 10 for test)	kNN, PC-kNN, RP-kNN	*IEEE TRANSACTIONS ON SEMICONDUCTOR MANUFACTURING, VOL. 28, NO. 1, FEBRUARY 2015*
Fault Detection	**Plasma Etching**	12 variables(Pressure, Tuner, Power, valve, etc.)	120 wafers (normal 100 wafers and 20 types of fault wafer)	PCA(preprocessing) GMM(NLLP, MD)	*Yu et al. IEEE TRANSACTIONS ON SEMICONDUCTOR MANUFACTURING, VOL. 24, NO. 3, (2011)*
Fault Detection	**Plasma Etching**	31 sensors readings for each wafer	782 wafers with 8 recipes (490 for training, 292 for test)	SVM, K-Means Clustering Self-Organizing MAP	*Rostami et al, 15th IEEE International Conference on Machine Learning and Applications (2016)*
Optimization	**Lithography (EPC)**	Variable : 34 parameters(24 local densities and 10 optical kernel signals)	1,600 segments(1,000 for training, 600 for test)	34 input nodes, 3 hidden layers(10 hidden nodes), 1 output node	*Advanced Etch Technology for Nano patterning V, 97820O (2016)*
Optimization	**Lithography (OPC)**	Spatial frequency, wave length, pattern error	3 types of pattern(contacts, multiple gates, complex one)	Stochastic gradient descent(SGD)	*Jia et al. Machine learning for inverse lithography J. Opt. 12. (2010)*
Virtual Metrology	**Plasma Etching**	OES data(1747 wavelength)	1747 cases of process	CNN with image input	*M. Terzi at al, IEEE 3rd International Forum on Research and Technologies for Society and Industry (2017)*

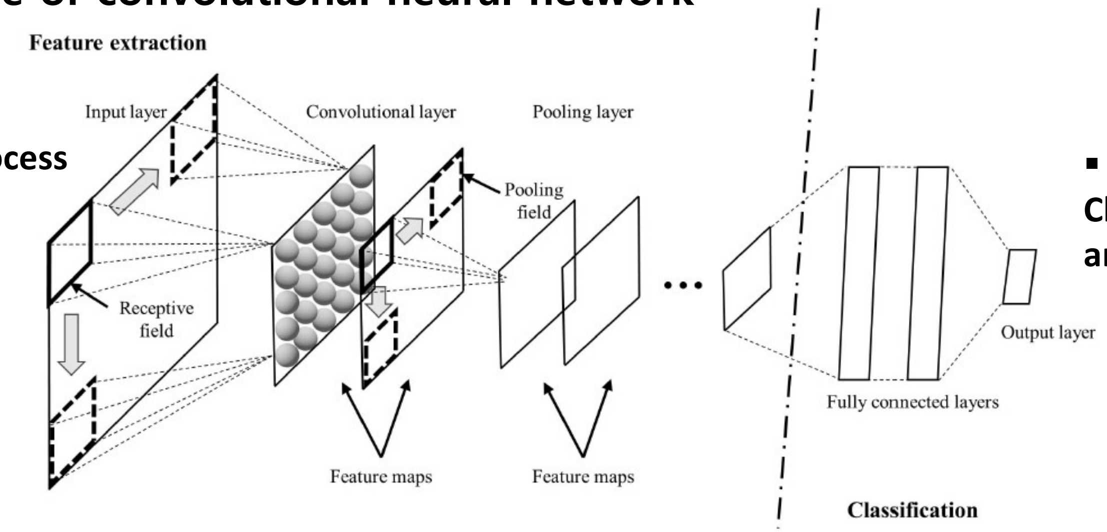
CNN for Fault Classification and Diagnosis

Lee et al, IEEE TRANSACTIONS ON SEMICONDUCTOR MANUFACTURING, VOL. 30, NO. 2 (2017)

◆ **Fault Detection and Classification Convolution Neural Network(FDC-CNN)**

➢ **Structure of convolutional neural network**

- **Input data:** **wafer with 10 process variables**



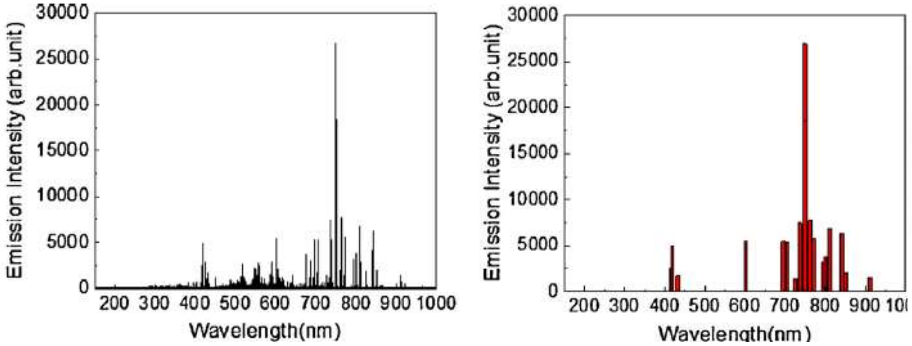
- **Output:** **Class of wafer(normal and 5 types of fault)**

Process	Chemical Vapor Deposition(CVD)
Time(scale)	105 seconds(0.2 second)
Variable(number)	Temperature, pressure, gas flow rate for each wafer, etc.(10)
Preprocessing	Scaling to a range of 0 to 1
Output	Class(total 6 with normal and 5 types of faults)
Training data	5,000 normal wafers and 5,000 fault wafers data(1,000 for each fault type)
Test data	1,000 normal wafers and 1,000 fault wafers data(200 for each fault type)

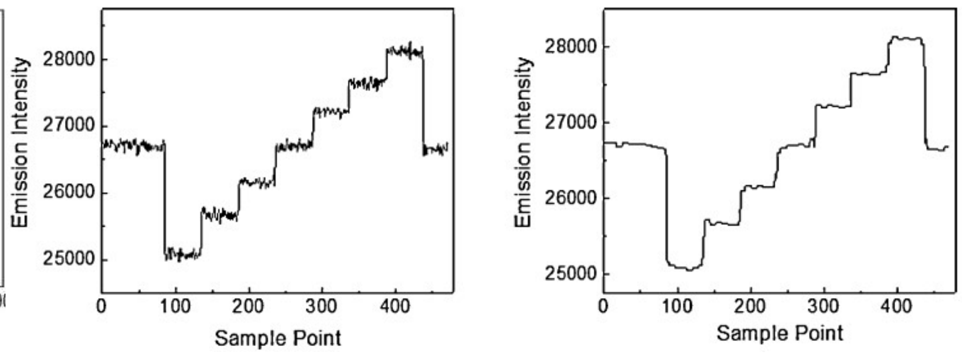
Sensitivity Enhancement with PCA Modeling in Plasma Etch

Ha et al, Computers and Chemical Engineering 94 (2016) 362–369

◆ Improvement of PCA modeling through DWT and automatic variable



➢ **OES spectrum data before and after applying peak wavelength selection algorithm.**



➢ **Time resolved signal of Ar emission at 750 nm wavelength before and after DWT.**

Process	Plasma Etching(Fault Detection)
Data type	Optical Emission Spectroscopy(OES) data
Wavelengths(resolution)	150nm to 1000nm(0.4nm)
Preprocessing	Automatic variable selection algorithm, Discrete Wavelet Transform
Output	Signal(Sensitivity enhancement, noise reduction)
Plasma condition	20 mT of pressure, 300 W of60 MHz RF power, 400 sc cm of Ar flow rate, and 16 sccm of SF_6 flow rate

Fault Detection using Random Projection KNN

Zhou et al. IEEE TRANSACTIONS ON SEMICONDUCTOR MANUFACTURING, VOL. 28, NO. 1, (2015)

◆ Fault Detection using Random Projections and k-Nearest Neighbor

➢ Input data

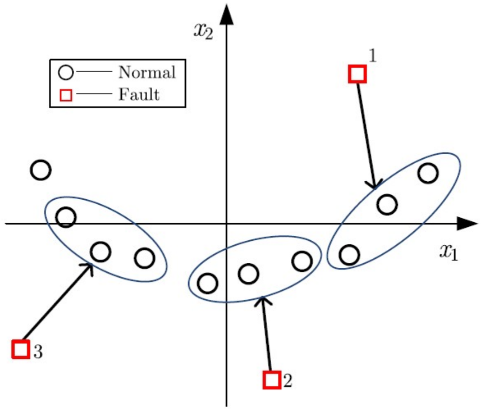
TABLE I
INDUCED FAULTS

No.	Fault	No.	Fault
1	TCP+50	11	CL_2+5
2	RF-12	12	BCL_3-5
3	RF+10	13	Pressure+2
4	Pressure+3	14	TCP-20
5	TCP+10	15	TCP-15
6	BCL_3+5	16	CL_2-10
7	Pressure-2	17	RF-12
8	CL_2-5	18	BCL_2+10
9	He Chuck	19	Pressure+1
10	TCP+30	20	TCP+20

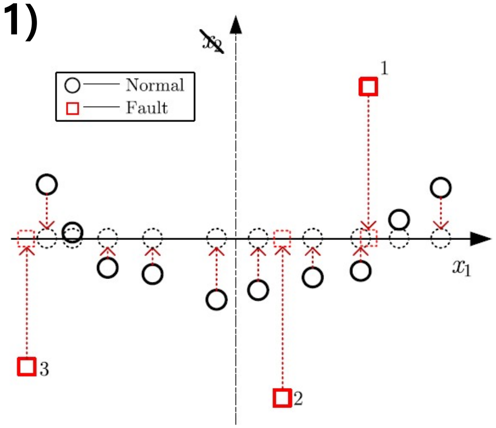
TABLE II
PROCESS VARIABLES USED FOR MONITORING

No.	Variables	No.	Variables
1	BCL_3 flow	10	RF power
2	CL_2 flow	11	RF impedance
3	RF bottom power	12	TCP tuner
4	Endpoint A detector	13	TCP phase error
5	Helium pressure	14	TCP impedance
6	Chamber pressure	15	TCP top power
7	RF tuner	16	TCP load
8	RF load	17	Vat valve
9	Phase error		

➢ Fault detection using FD-kNN (k=3)



➢ Fault detection using PC-kNN (PC1)



Process	Plasma Etching(TiN/A1-0.5% Cu/TiN/oxide stack with an inductively coupled BCl_3/Cl_2 plasma.)
Variables (number)	flow rate, power impedance, pressure, tuner, etc.(17)
Preprocessing	Scaling to zero mean and unit variance for variables
Output	Classification(20 types of faults)
Training data	97 wafers selected randomly for 100 times
Test data	10 wafers for validation respectively

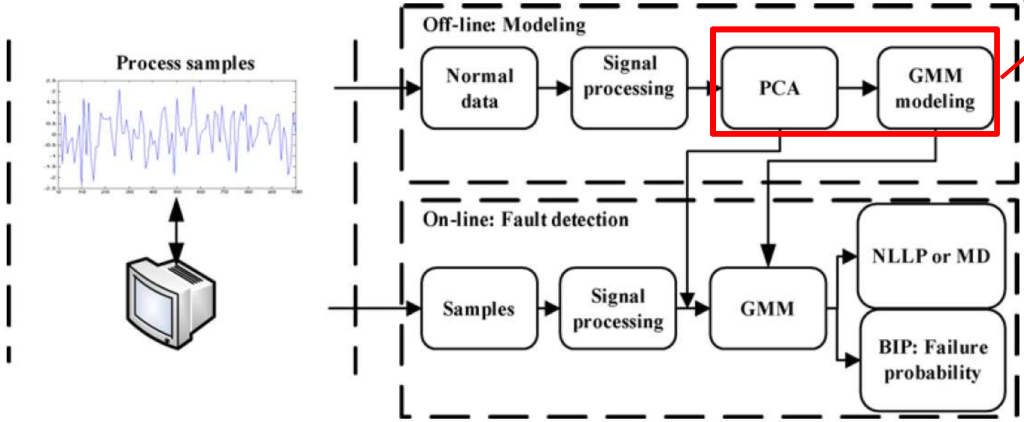
➢ Comparison of computation speed

	FD-kNN	PC-kNN SVD	PC-kNN kNN	RPkNN
Average time of building model (s)	0.3195	0.9813 (0.9626)	0.9813 (0.0187)	0.0840
Average time of processing a test sample (s)	1.2×10^{-3}	0.18×10^{-3}		0.60×10^{-3}

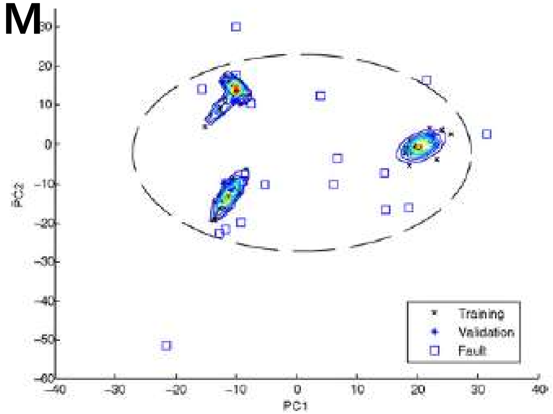
Fault Detection Using PCA-based GMM

Yu et al. IEEE TRANSACTIONS ON SEMICONDUCTOR MANUFACTURING, VOL. 24, NO. 3, (2011)

◆ Fault Detection Using Principal Components-Based Gaussian Mixture Model



PCGMM



➢ **Bivariate scores of PCGMM with PC1 and PC2**

Process	Fault detection after plasm etching
Variable(number)	End point A detector, Helium pressure, RF tuner, RF load, RF phase error, RF power, RF impedance, Transformer-coupled plasma tuner, Transformer-coupled plasma phase error, Transformer-coupled plasma reflected power, Transformer-coupled plasma load, VAT valve(12)
Preprocessing	Principal component Analysis(PCA) for dimension reduction
Output	NLLP AND MD for Fault classification with threshold
Training data	100 wafers with 12 process variables
Test data	100 wafers with 12 process variables

Equipment Condition Diagnosis & Fault Fingerprint Extraction

Rostami et al, 15th IEEE International Conference on Machine Learning and Applications (2016)

◆ Equipment condition diagnosis(ECD) and fault fingerprint extraction

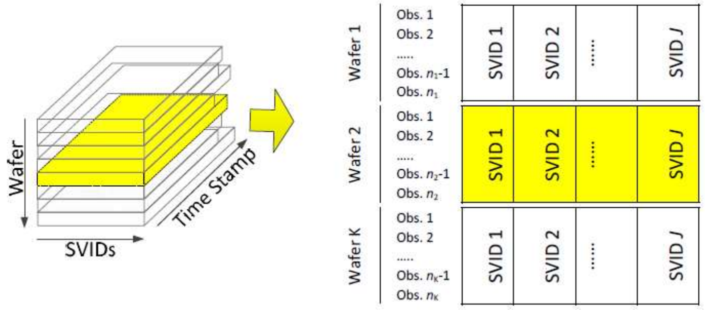
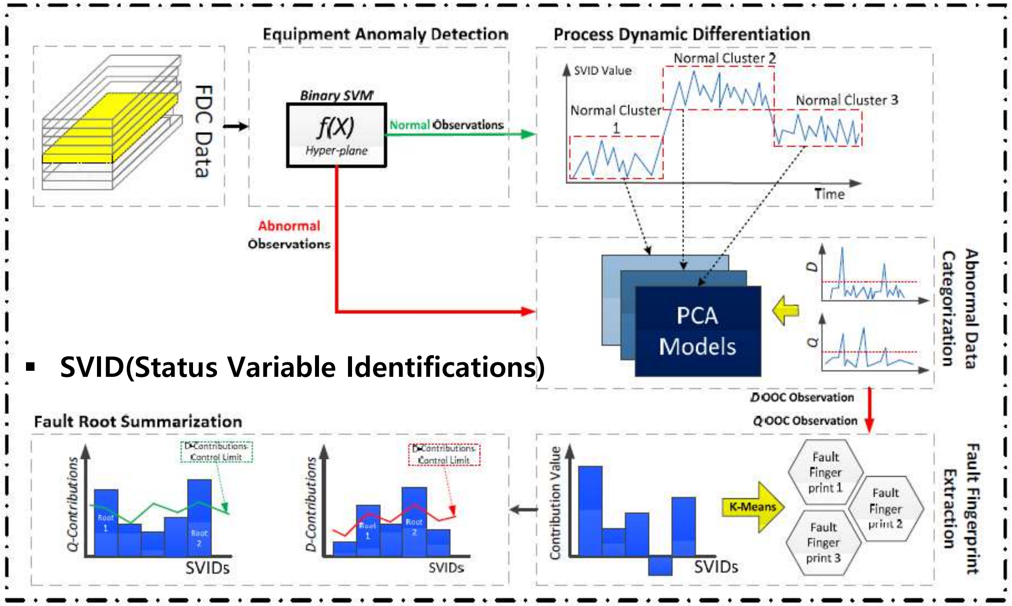


Fig. 1. The irregular FDC data cube is illustrated and unfolded.

$$W(k) = \begin{bmatrix} V^k_{1,1} & V^k_{1,2} & \cdots & V^k_{1,J} \\ V^k_{2,1} & V^k_{2,2} & \cdots & V^k_{2,J} \\ \vdots & \vdots & \ddots & \vdots \\ V^k_{n_k,1} & V^k_{n_k,2} & \cdots & V^k_{n_k,J} \end{bmatrix}_{n_k \times J},$$



- SVID(Status Variable Identifications)

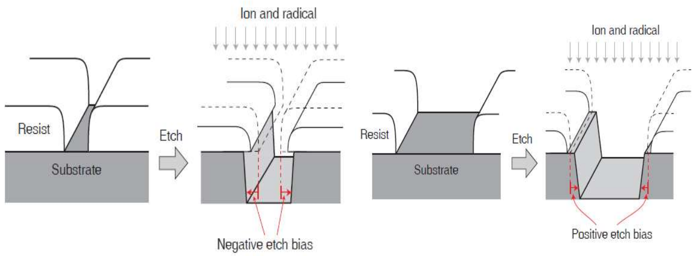
Process	Fault detection of equipment
Variable(number)	31 sensors reading for each wafer in plasma etching process
Preprocessing	SVM as a state-of-the-art classification method Principal component Analysis(PCA) for dimension reduction
Output	Fault finger prints
Training data	490 wafers with 8 recipes(197,562 observations)
Test data	292 wafers with 8 recipes(66,835 observations)

EPC through ML-Driven Etch Bias Model

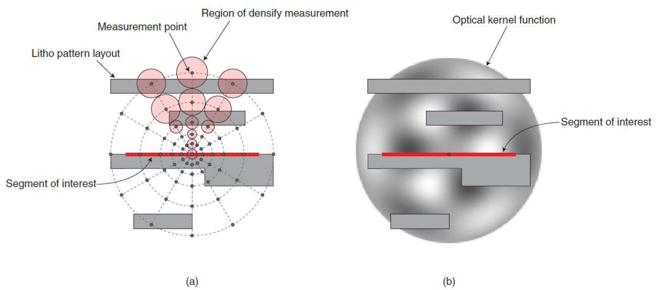
Shim et al. Advanced Etch Technology for Nano patterning V, 97820O (2016)

◆ **Etch Proximity Correction(EPC) through Machine Learning-Driven Etch Bias Model**

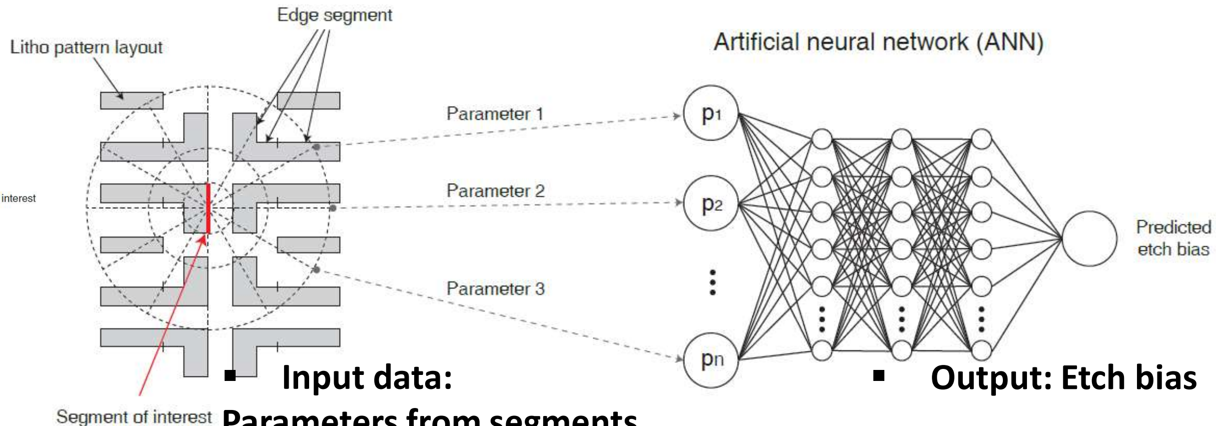
Process	Lithography
Variable (number)	24 local densities and 10 optical kernel signals(34)
Preprocessing	Representative segments selected by K-mean method
Output	Predicted etch bias
Training data	1,000 segments with local/optical variables
Test data	600 segments with local/optical variables



➢ **Etch bias by etch proximity effect**



➢ **Parameterization of a pattern**



- **Input data: Parameters from segments**
- **Output: Etch bias**

➢ **Structure of EPC with ANN**

SUNG KYUN KWAN UNIVERSITY

Machine Learning for Inverse Lithography

Jia et al. Machine learning for inverse lithography J. Opt. 12. (2010)

◆ **ML for inverse lithography: Stochastic gradient descent for robust photomask synthesis**

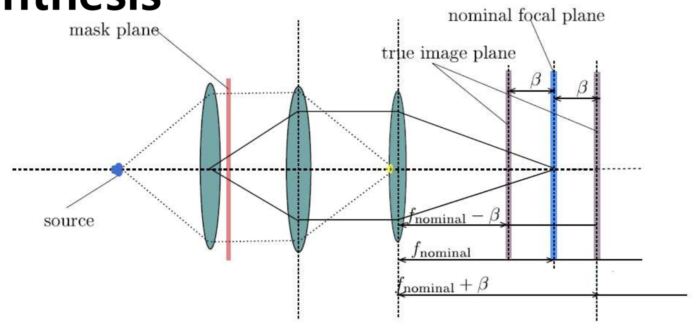


Figure 1. The defocus model in an optical projection lithography system.

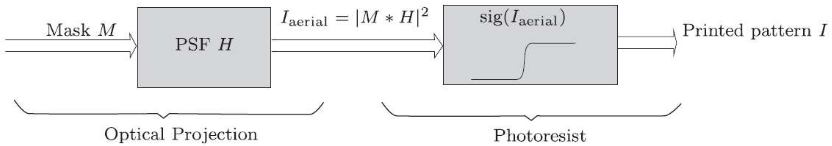
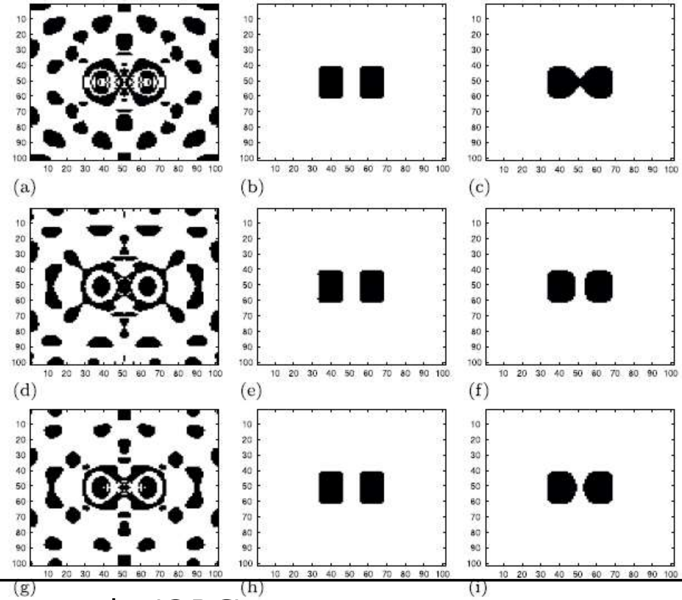
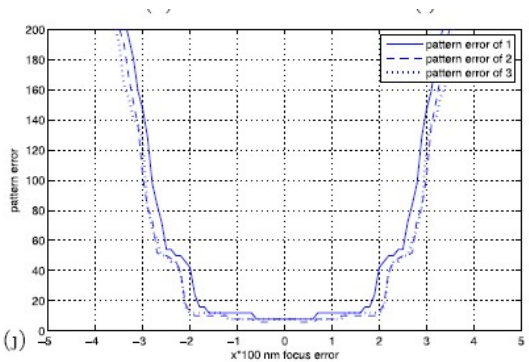


Figure 2. Forward model of the optical lithography

- **GD**
- **SGD**
- **BGD**





➢ **Result of pattern**

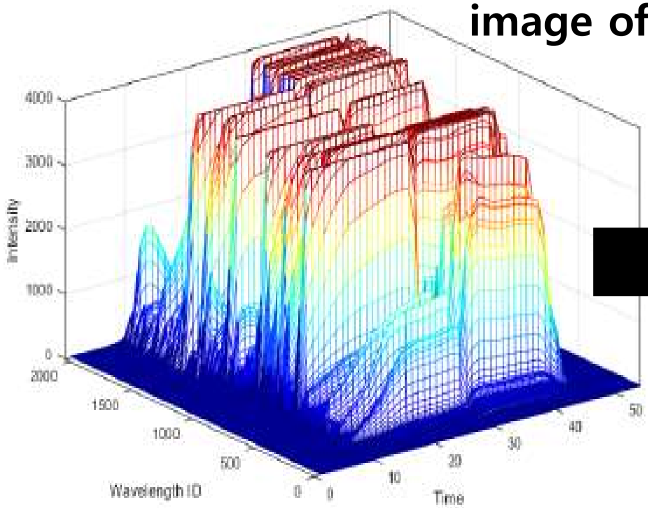
Process	Lithography(OPC)
Variable (number)	Spatial frequency, wave length, pattern error
Preprocessing	Representative segments selected by K-mean method
Output	mask pattern
Training data	3 types of pattern(contacts, multiple gates, complex one)
Comparative method	Standard gradient descent(GD) batch gradient descent(BGD)

SUNG KYUN KWAN UNIVERSITY

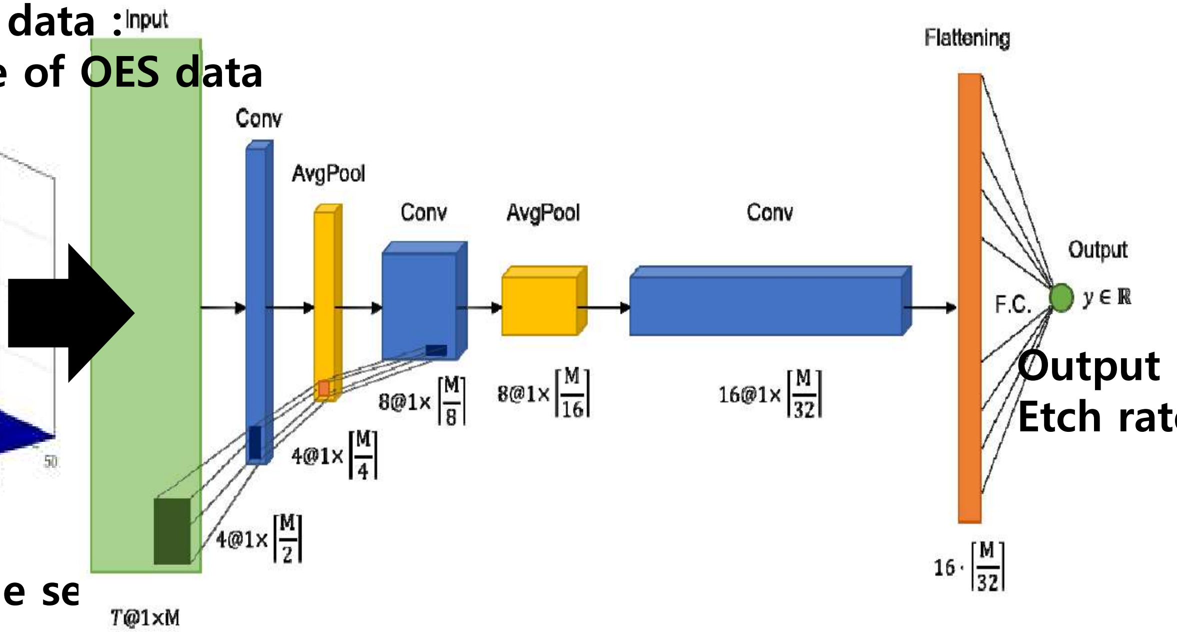
Virtual Metrology with Deep Learning: Etch Rate

M. Terzi at al, 2017 IEEE 3rd International Forum on Research and Technologies for Society and Industry (2017)

◆ Modeling with OES Data in plasma etching process



Input data : image of OES data



Output data : Etch rate

▶ **OES spectrum with the time se**

Process	Plasma Etching
Data type	Slice image of OES data
Data size	1747 process for 2048 time series
Preprocessing	Principal Component Analysis(PCA) for dimension reduction to 100 input variables
Output	Etch rate

➤ **Comparison of Ridge Regression and CNN**

TABLE I. CROSS-VALIDATED RESULTS.

Comparison of models	Accuracy	
	Ridge	*CNN*
R^2-score (± 3 std. dev.)	0.8 ± 0.05	0.91 ± 0.04
Number of parameters	100	877

Tab. 1. Comparison of 10-fold cross-validation results between Ridge Regression (with regularization parameter $\alpha = 0.1$) and CNN.

SUNG KYUN KWAN UNIVERSITY

요약

- 다양한 Machine Learning과 Deep Learning 기법이 개발되고 있음.
- 반도체 공정 데이터의 폭발적인 증가가 진행중임.
- 플라즈마 공정을 포함한 복잡한 공정에 Machine Learning 기법이 적용되고 있음.
- 반도체 공정 데이터 분석 사례가 늘어남.
  - Sensitivity Enhancement
  - Fault Detection
  - Optimization
  - Virtual Metrology