

# 프로젝트 #1 중간 발표

2022. 5. 25

충북대학교 산업인공지능학과

[20-8조] 고정재, 유용주

# 수행방법 및 기여도

## 수행방법

- 같은 회사에 재직중이나 팀이 달라 일과 시간 이후에 프로젝트 수행.
- 회사 업무가 전공분야가 아니라, 코딩 방법, CNN에 대해 공부를 하면서 진행 함.
- 코딩을 할 수 있는 능력이 안되어 최대한 논문을 참고하여 유사하게 수행 하려고 함.

## 업무분장 및 기여도

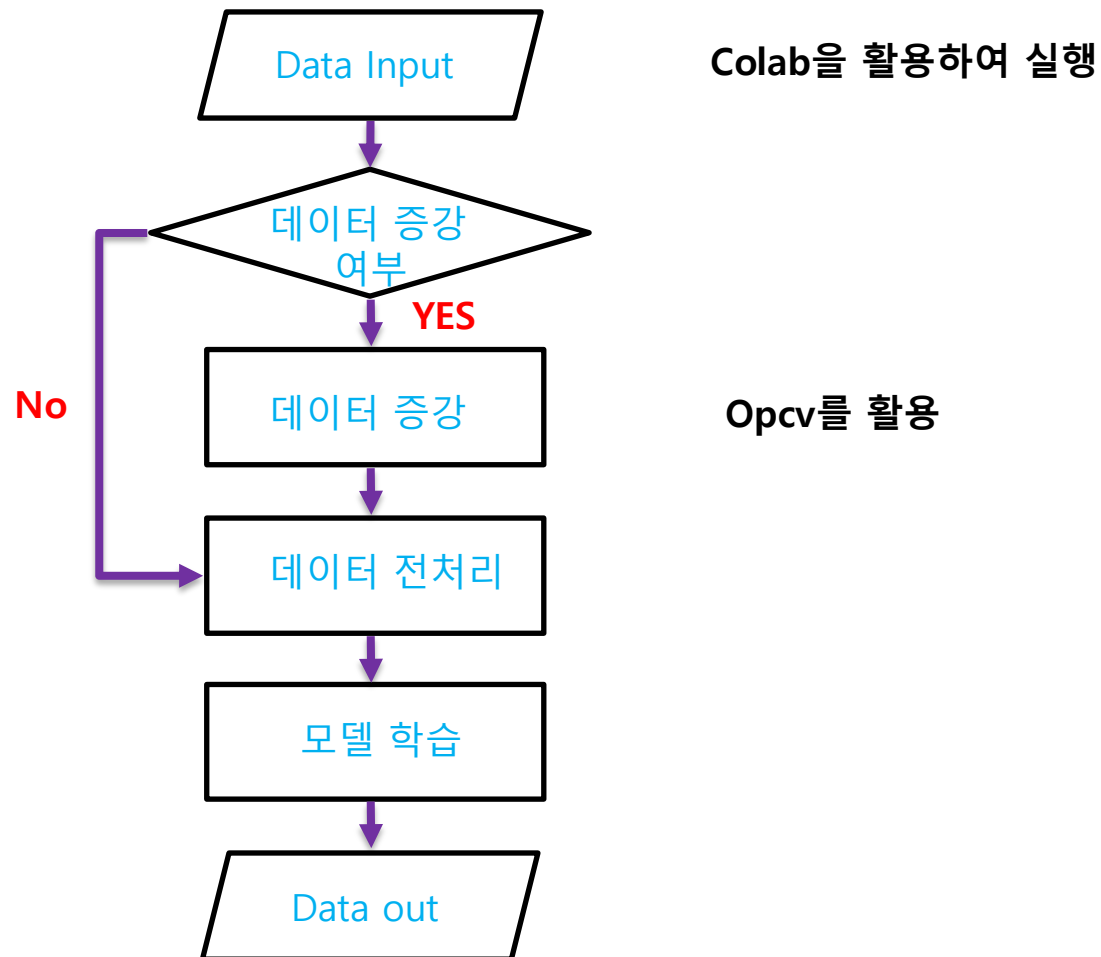
이름	비중	수행내용	비고
고정재	50%	<ul style="list-style-type: none"><li>• 데이터 증량</li><li>• 데이터 전처리</li><li>• 성능 평가</li><li>• 발표</li></ul>	
유용주	50%	<ul style="list-style-type: none"><li>• Kaggle dataset 활용법, colab 사용법, CNN에 대한 공부 자료 수집</li><li>• 발표 자료 작성</li><li>• 발표</li></ul>	

# 데이터셋

## Dataset

- Kaggle에 있는 wm811k-wafer-map을 활용
- Total 811,457개의 WM Data 구성

## 진행 흐름도



# 데이터셋(Augmentation)

## - Data 전처리를 하여 Class 불균형 해결

With Label  
172,950 (21.3%)

None	147,431(18.17%)
Center	4,294 (0.53%)
Donut	555 (0.07%)
Edge-Loc	5189 (0.64%)
Edge-Ring	9680 (1.19%)
Local	3593 (0.44%)
Random	866 (0.11%)
Scratch	1193 (0.15%)
Near-full	149 (0.02%)



논문 증강 기법 적용

10°의 무작위 회전 : 20%  
좌우 대칭 및 너비 이동 : 20%  
높이 이동 : 15%  
전단 범위 : 10%  
채널 이동 및 확대/축소 : 10%  
100% 비율로 맞춤 예정

None	10,000
Center	10,000
Donut	10,000
Edge-Loc	10,000
Edge-Ring	10,000
Local	10,000
Random	10,000
Scratch	10,000
Near-full	10,000

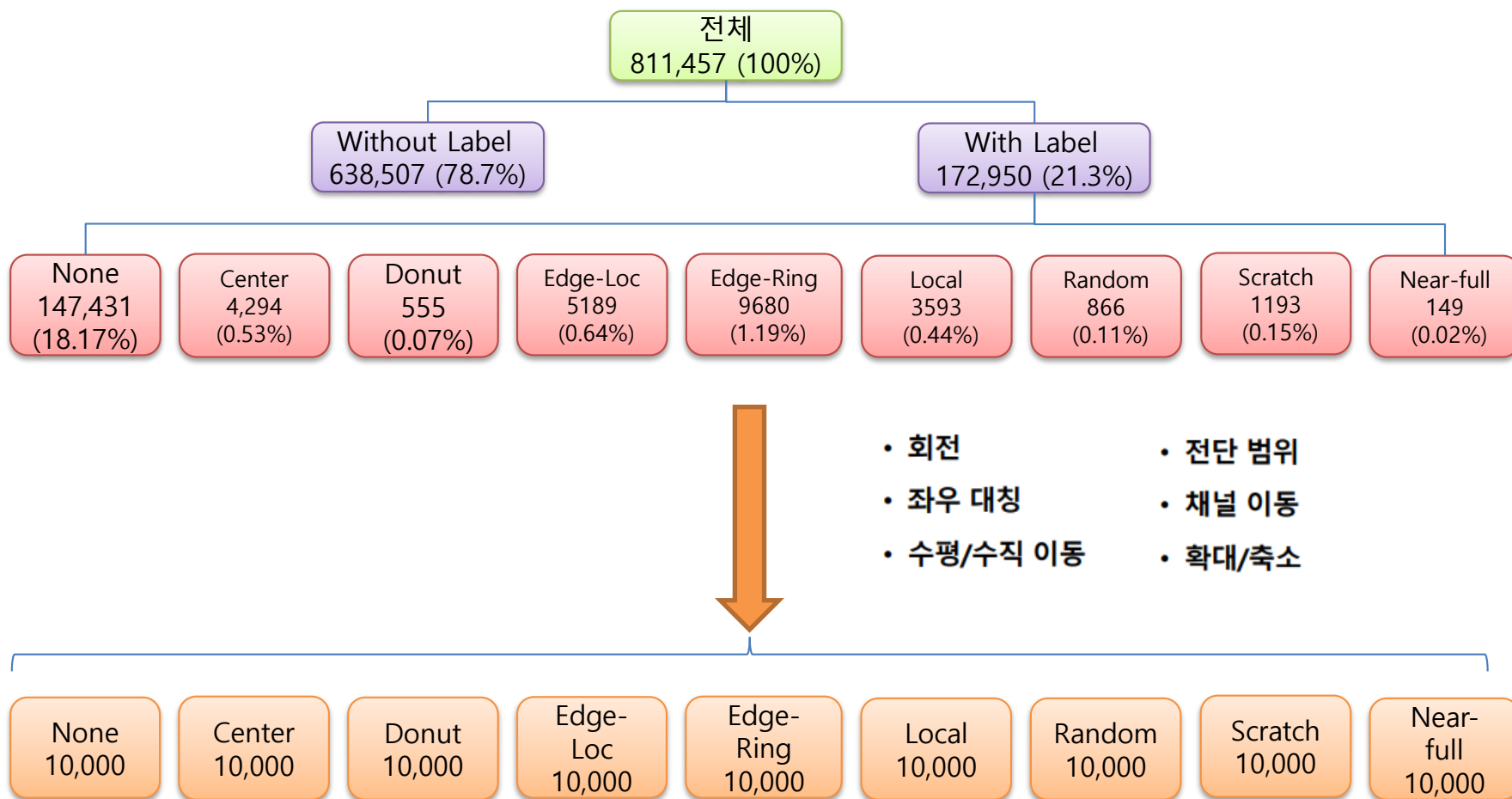
기존 Dataset

변경 Dataset

# 데이터셋

## 데이터 구성

- Kaggle wm811k-wafer-map



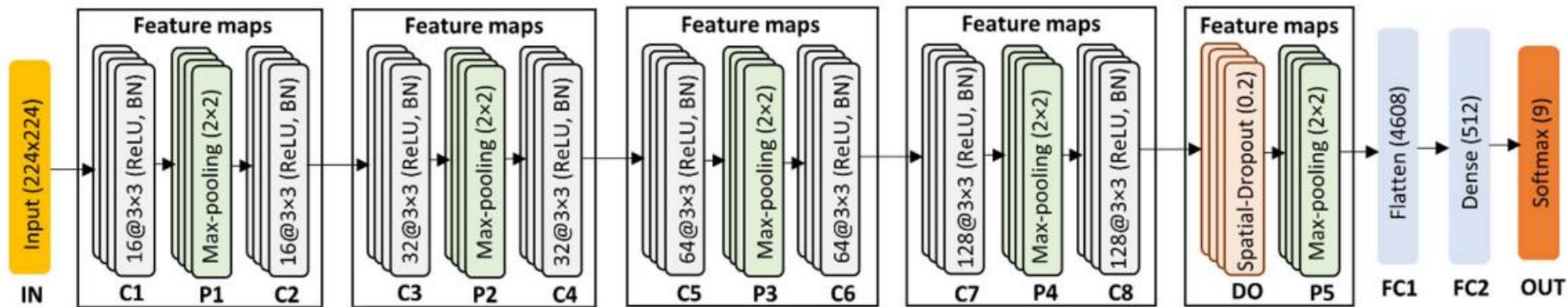
Train : Validation : Test = 65 : 20 : 15

Train : Validation : Test = 65 : 20 : 15

# CNN 구조

## CNN 구조

- 그림/표/모델정보(실행화면)로 표현 (논문과 동일? 다르면 어떤 부분이 다른지 등)



\*Note: IN denotes input layer; C convolution layer; P pooling layer; DO dropout layer; FC fully connected layer; OUT output layer; and BN batch normalization

C2	Convolution2	16	111×111	3×3	Yes	ReLU
C3	Convolution3	32	111×111	3×3	Yes	ReLU
P2	Max Pooling2	32	55×55	2×2	No	-
C4	Convolution4	32	55×55	3×3	Yes	ReLU
C5	Convolution5	64	55×55	3×3	Yes	ReLU
P3	Max Pooling3	64	27×27	2×2	No	-
C6	Convolution6	64	27×27	3×3	Yes	ReLU
C7	Convolution7	128	27×27	3×3	Yes	ReLU
P4	Max Pooling4	128	13×13	2×2	No	-
C8	Convolution8	128	13×13	3×3	Yes	ReLU
P5	Max Pooling5	128	6×6	2×2	No	-
FC1	Fully-Connected1	1	4608	-	-	ReLU
FC2	Fully Connected2	1	512	-	-	ReLU
OUT	Output	1	9	-	-	Softmax

## 과적합을 방지하기 위한 규제화(regulation)

- Batch Normalization(정규화)
- Spatial Dropout = 0.2

# CNN 구조

## 주요 코드 및 실행 결과

- 현재까지는 Kaggle에 있는 Dataset을 다운로드 한 후 colab에 데이터 로딩

```
[1] 1 from google.colab import drive  
2 drive.mount('/content/drive')
```

Mounted at /content/drive

```
1 !cp '/content/drive/MyDrive/tmp/archive.zip' ./
```

```
[3] 1 !unzip archive.zip
```

Archive: archive.zip  
inflating: LSWMD.pkl

```
[4] 1 !ls -alh
```

```
total 2.1G  
drwxr-xr-x 1 root root 4.0K May 24 23:28 .  
drwxr-xr-x 1 root root 4.0K May 24 23:26 ..  
-rw----- 1 root root 150M May 24 23:28 archive.zip  
drwxr-xr-x 4 root root 4.0K May 17 13:38 .config  
drwx----- 6 root root 4.0K May 24 23:28 drive  
-rw-r--r-- 1 root root 2.0G Sep 27 2019 LSWMD.pkl  
drwxr-xr-x 1 root root 4.0K May 17 13:39 sample_data
```

```
[5] 1 import numpy as np  
2 import pandas as pd  
3 import matplotlib.pyplot as plt
```

```
[6] 1 df=pd.read_pickle("LSWMD.pkl")
```

```
[7] 1 df.head()
```

# CNN 구조

## 주요 코드 및 실행 결과

- 현재까지는 Kaggle에 있는 Dataset을 다운로드 한 후 colab에 데이터 로딩

	waferMap	dieSize	lotName	waferIndex	trianTestLabel	failureType
0	[[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,...	1683.0	lot1	1.0	[[Training]]	[[none]]
1	[[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,...	1683.0	lot1	2.0	[[Training]]	[[none]]
2	[[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,...	1683.0	lot1	3.0	[[Training]]	[[none]]
3	[[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,...	1683.0	lot1	4.0	[[Training]]	[[none]]
4	[[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,...	1683.0	lot1	5.0	[[Training]]	[[none]]



```
1 df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 811457 entries, 0 to 811456
Data columns (total 6 columns):
#   Column          Non-Null Count  Dtype  
---  -
0   waferMap         811457 non-null object  
1   dieSize          811457 non-null float64
2   lotName          811457 non-null object  
3   waferIndex       811457 non-null float64
4   trianTestLabel   811457 non-null object  
5   failureType      811457 non-null object  
dtypes: float64(2), object(4)
memory usage: 37.1+ MB
```



# 학습 방법

---

## 딥러닝 학습 조건

- 노트북 PC 사양, 학습시간

CPU : Intel(R) Core(TM) i7-6500U CPU @ 2.50GHz    2.59 GHz

RAM : 8GB

GPU : Intel(R)HD Graphics 520

- 하이퍼파라미터 : epoch수, 학습률, batch size, optimizer, loss 함수 등
- 학습추이 그래프 등
- 학습 중 알게 된 내용 등 ... 학습률/optimizer에 따른 학습추이 비교 등...

**감사합니다**