Foundations of Deep Learning

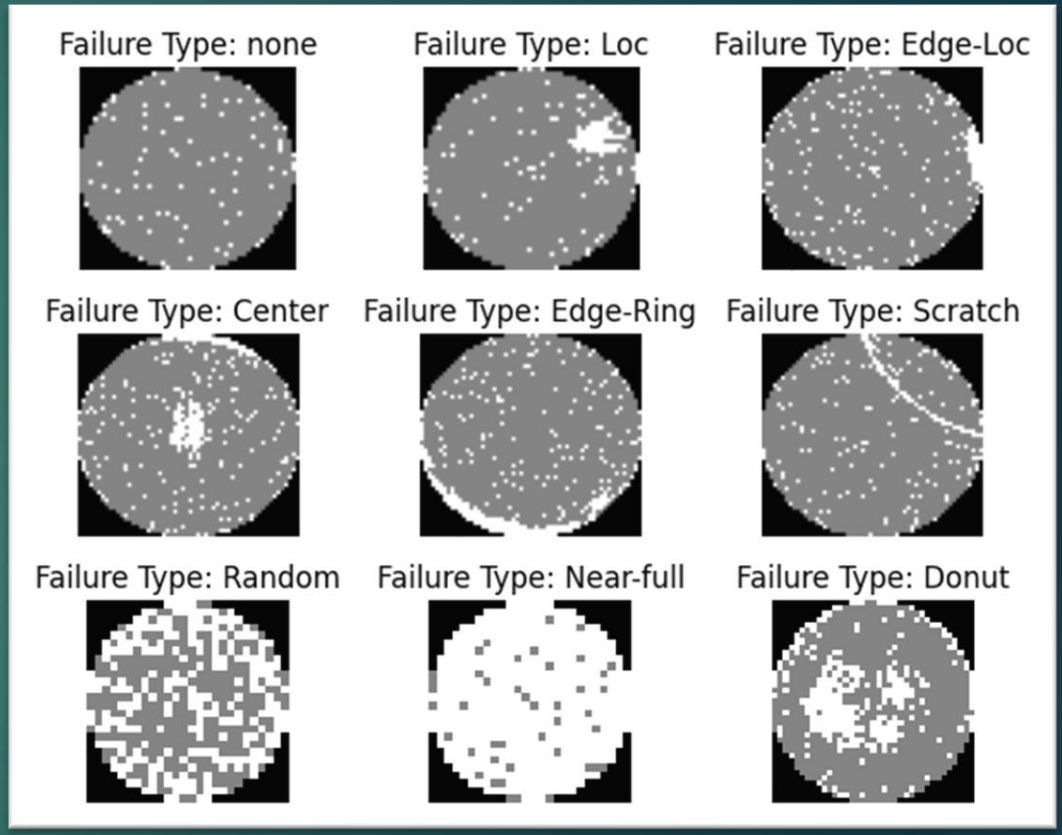
WAFERMAP CLASSIFICATION USING CNN

PAOLO CAGGIANO 902877
ALBERTO ODIERNA 907341

JUNE 2023

The Problem

- 811.457 images:
  - 172.950 with the label
  - 638.507 without the label(to delete)
- 3 pixel types:
  - 0 background
  - 1 normal chip
  - 2 anomalies
- Images with different size
- 9 types of failure
- Unbalanced classes



Data Preprocessing

- Deletion of the images without the label
- Resize of the images in a 32x32 form

	failureType	waferMap	waferMapDim
232339	Edge-Ring	[[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,...	(53, 66)
268835	Edge-Ring	[[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,...	(46, 46)
706495	none	[[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2, 1, 2,...	(25, 27)
754475	none	[[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 2, 1,...	(27, 25)
279190	Donut	[[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,...	(41, 42)

```
(25, 27)     18781
(26, 26)     14366
(30, 34)     12400
(29, 26)     11751
(27, 25)     10682
                 ...
(50, 68)         1
(29, 45)         1
(38, 41)         1
(75, 90)         1
(68, 75)         1
```

Frequencies of the different images size

- Split in train, validation and test
- Two approaches for rebalancing the classes:
  - Data Augmentation
  - SMOTE Method

Data Augmentation

- Rotation
- Shift
- Zoom

Before

none	103202
Edge-Ring	6776
Edge-Loc	3632
Center	3006
Loc	2515
Scratch	835
Random	606
Donut	389
Near-full	104

After

none	103202
Edge-Ring	13552
Edge-Loc	7264
Center	6012
Loc	5030
Scratch	1670
Random	1212
Donut	778
Near-full	208

Smote

- Euclidean distance to determine the nearest observations
- We use the five nearest neighbors in order to generate new samples

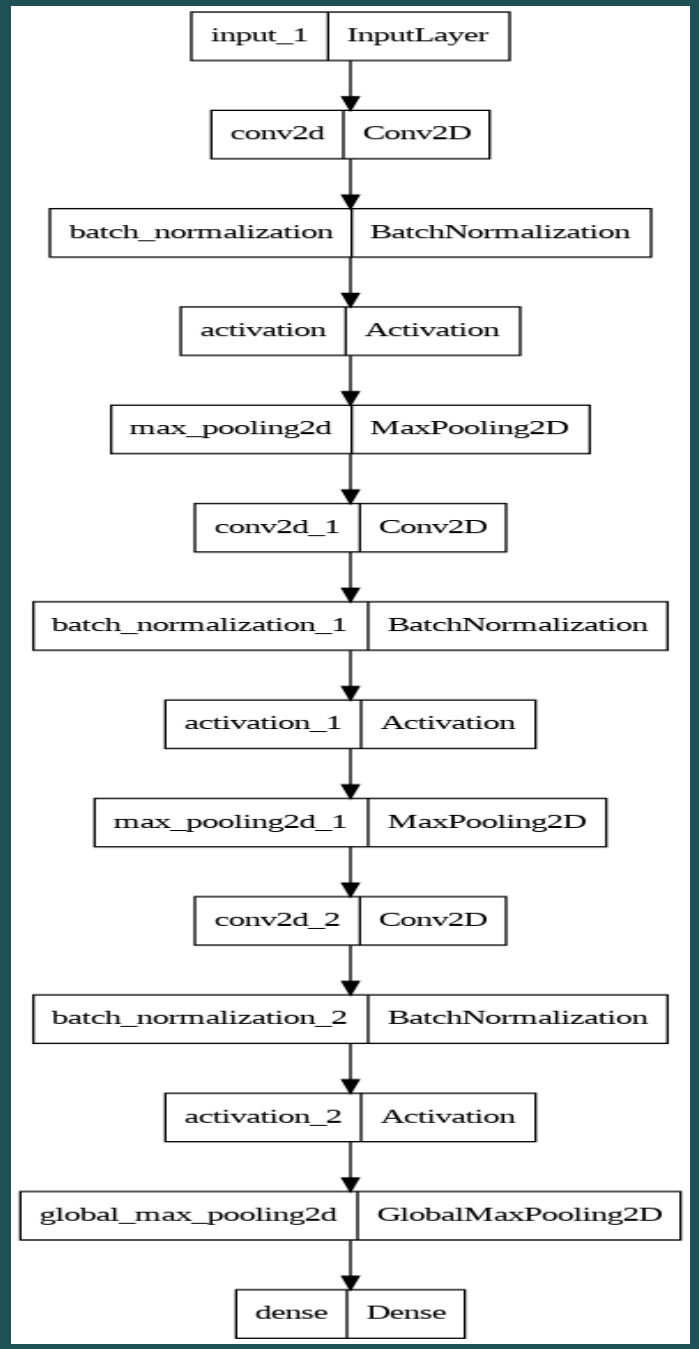
none	103202
Edge-Ring	6776
Edge-Loc	3632
Center	3006
Loc	2515
Scratch	835
Random	606
Donut	389
Near-full	104

none	103202
Edge-Loc	7264
Edge-Ring	6776
Center	6012
Loc	5030
Scratch	1670
Random	1212
Donut	778
Near-full	208

ONLY ON TRAINING DATA

The architecture

- Three convolutional layers
- Three pooling layers
- Activation Function(relu)
- One Dense Layer
- Batch size=32
- Optimizer=Adam
- Callback(patience=5)- Early Stopping
- 20 epochs

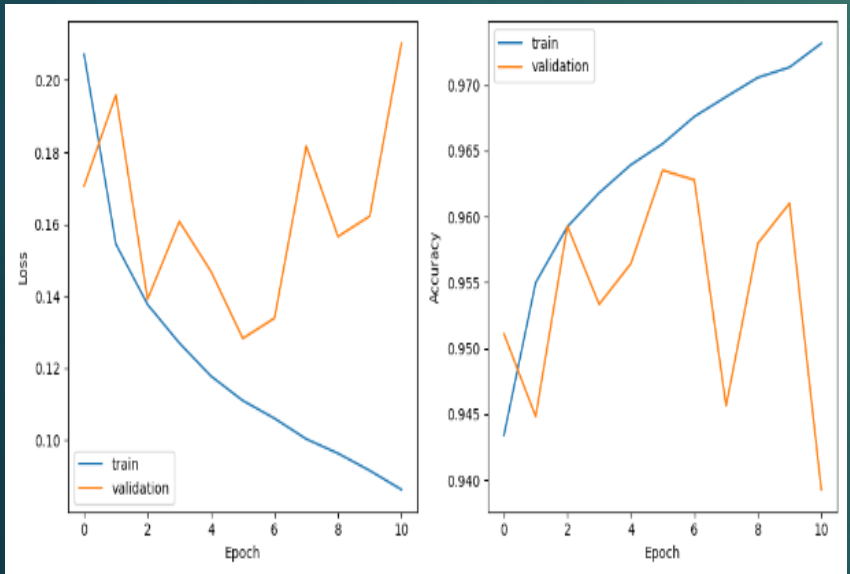


FIRST MODEL

Performances on the test set

Precision	Recall	F-score
0.72	0.68	0.73

LEARNING RATE = 0.01



Precision, F-score, Recall per class

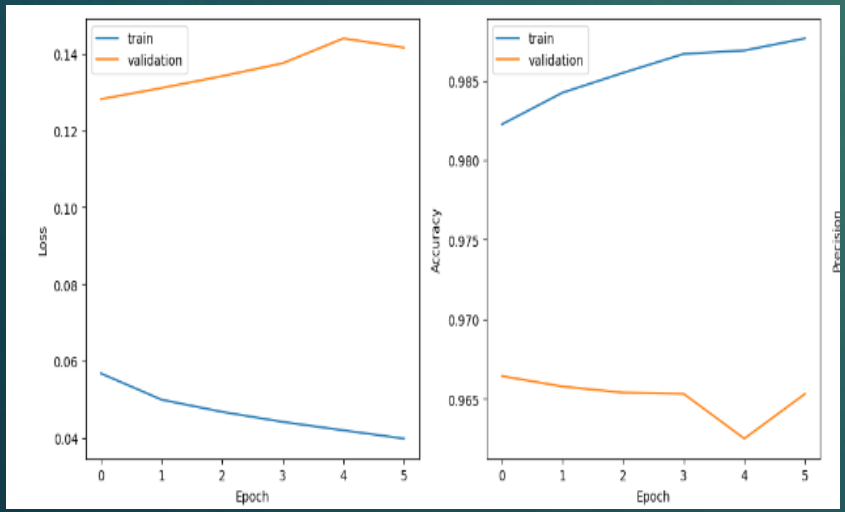
CLASS	PRECISION	F_SCORE	RECALL
Near-full	0.81	0.89	0.89
Center	0.44	0.60	0.93
Edge-Loc	0.58	0.69	0.86
None	0.95	0.94	0.94
Loc	0.64	0.59	0.56
Donut	1	0.35	0.21
Random	0.87	0.82	0.78
Edge-Ring	0.22	0.30	0.47
Scratch	0.98	0.97	0.96

SECOND MODEL

Performances on the test set

Precision	Recall	F-score
0.84	0.82	0.80

LEARNING RATE = 0.001



Precision, F-score, Recall per class

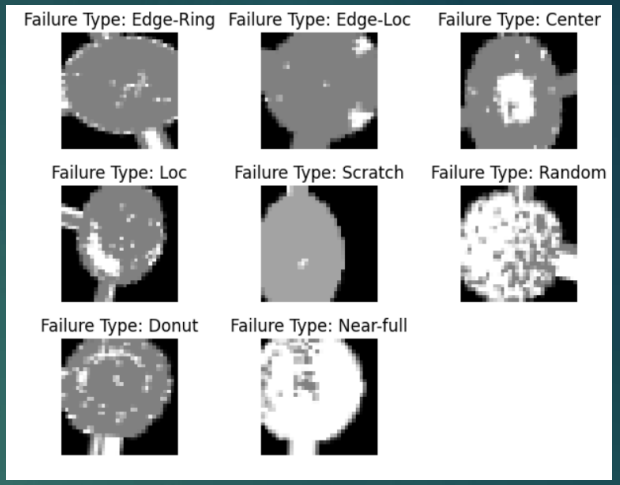
CLASS	PRECISION	F_SCORE	RECALL
Near-full	0.90	0.90	0.89
Center	0.81	0.81	0.80
Edge-Loc	0.78	0.76	0.73
None	0.97	0.96	0.95
Loc	0.69	0.69	0.69
Donut	0.95	0.95	0.95
Random	0.90	0.87	0.84
Edge-Ring	0.53	0.44	0.37
Scratch	0.98	0.98	0.99

THIRD MODEL – DATA AUGMENTATION

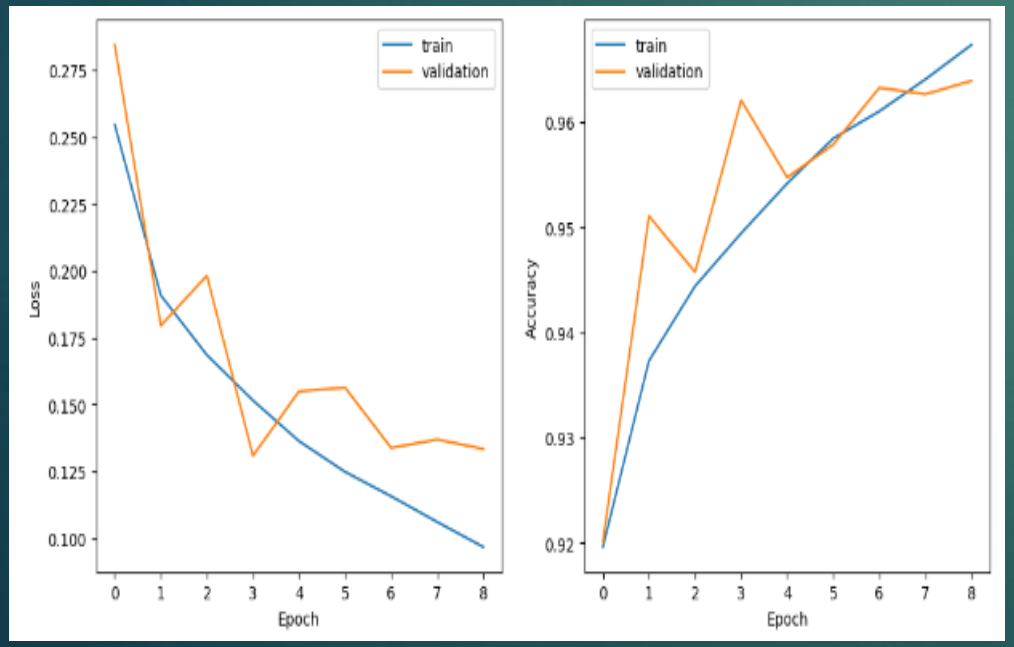
Performances on the test set

Precision	Recall	F-score
0.86	0.81	0.78

Examples of images generated with Data Augmentation



LEARNING RATE=0.001



Precision, F-score, Recall per class

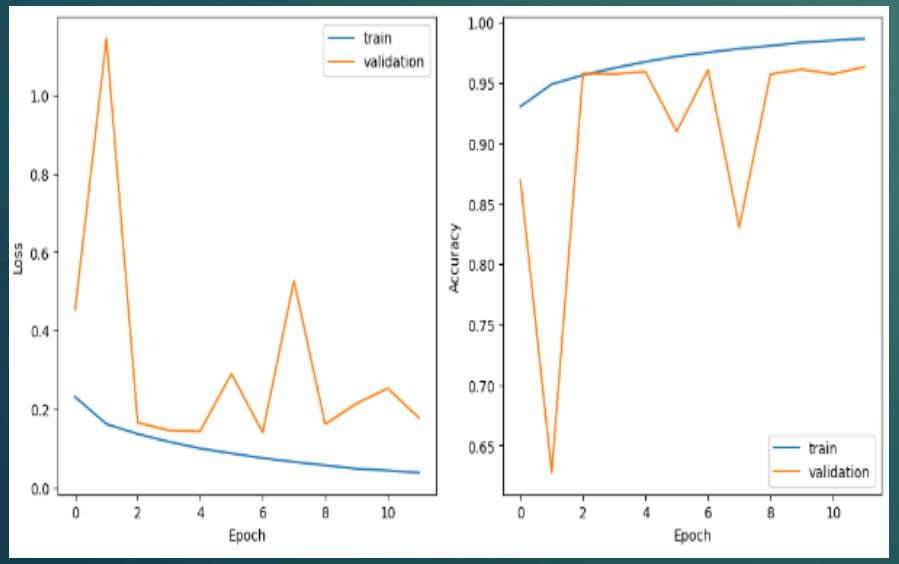
CLASS	PRECISION	F_SCORE	RECALL
Near-full	0.91	0.90	0.88
Center	0.93	0.80	0.71
Edge-Loc	0.81	0.77	0.74
None	0.97	0.96	0.94
Loc	0.78	0.64	0.54
Donut	1	0.97	0.95
Random	0.91	0.88	0.85
Edge-Ring	0.43	0.43	0.44
Scratch	0.97	0.98	0.89

FOURTH MODEL – SMOTE

Performances on the test set

Precision	Recall	F-score
0.86	0.79	0.75

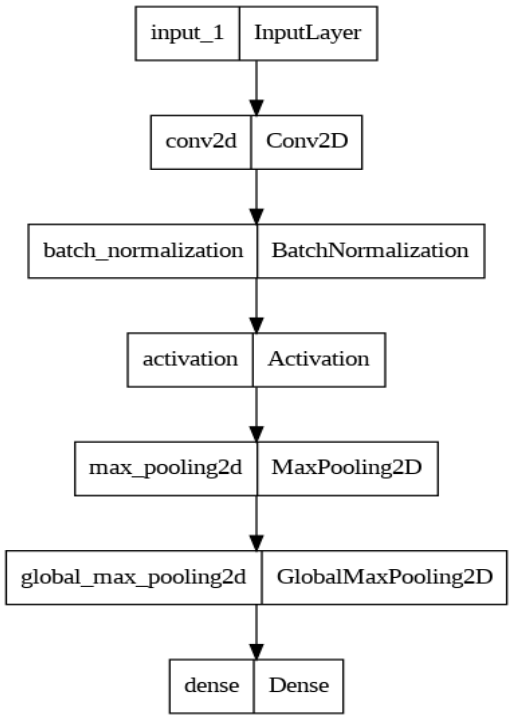
LEARNING RATE=0.001

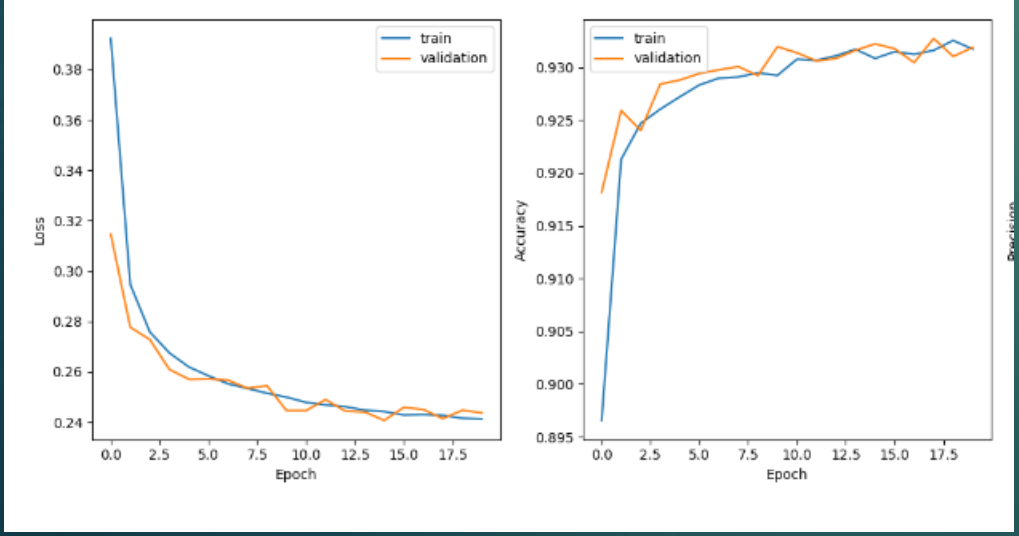


Precision, F-score, Recall per class

CLASS	PRECISION	F_SCORE	RECALL
Near-full	0.90	0.89	0.89
Center	0.94	0.71	0.57
Edge-Loc	0.73	0.75	0.78
None	0.99	0.95	0.91
Loc	0.70	0.66	0.62
Donut	1	0.95	0.91
Random	0.87	0.88	0.90
Edge-Ring	0.62	0.31	0.21
Scratch	0.97	0.98	0.99

REDUCING THE COMPLEXITY





Performances on the test set

Precision	Recall	F-score
0.60	0.51	0.50

Precision, F-score, Recall per class

CLASS	PRECISION	F_SCORE	RECALL
Near-full	0.63	0.67	0.72
Center	0.41	0.38	0.36
Edge-Loc	0.61	0.50	0.43
None	0.89	0.88	0.88
Loc	0.63	0.24	0.15
Donut	0.80	0.28	0.17
Random	0.53	0.63	0.78
Edge-Ring	0	0	0
Scratch	0.95	0.97	0.99

RECAP

RECALL PER CLASS OF THE FIVE MODELS

CLASS	First model	Second model	Third model	Fourth model	Fifth model
Near-full	0.89	0.89	0.88	0.89	0.72
Center	0.93	0.80	0.71	0.57	0.36
Edge-Loc	0.86	0.73	0.74	0.78	0.43
None	0.94	0.95	0.94	0.91	0.88
Loc	0.56	0.69	0.54	0.62	0.15
Donut	0.21	0.95	0.95	0.91	0.17
Random	0.78	0.84	0.85	0.90	0.78
Edge-Ring	0.47	0.37	0.44	0.21	0.0
Scratch	0.96	0.99	0.89	0.99	0.99

THANKS FOR THE ATTENTION!