Enhancement of existing ML pipeline via DL approaches for the detection of patterns in wafermaps

AI

ARROWHEAD TOOLS

Methodology

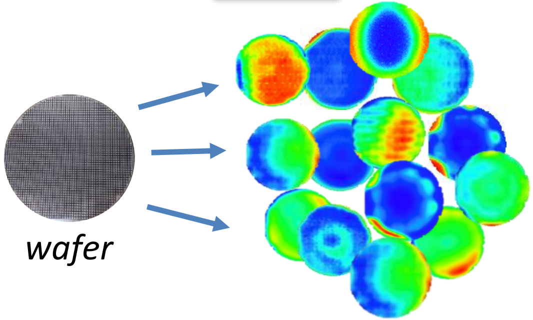
1. Machine Learning pipeline
2. Scope of the Master Thesis
3. Motivation for using Deep Learning

Wafer Health Factor – Machine Learning pipeline

Concept:

The wafer health factor informs you about the presence of pre-defined (critical) process patterns, based on analog wafer test data (before the yield is affected).

Input → WHF calculation → Output

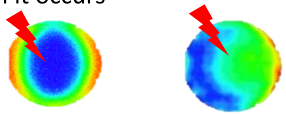


wafer

wafermaps

Pattern criticality

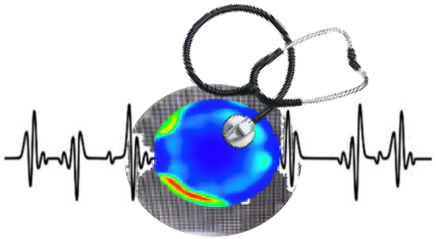
the **expert** defines which pattern he wants to be informed about when it occurs



Pattern classification

an **ML algorithm** is trained on data (electrical measurements) to detect and group the different patterns (visible as wafermaps)

pattern A | pattern B



[0%, 100%]

ARROWHEAD TOOLS

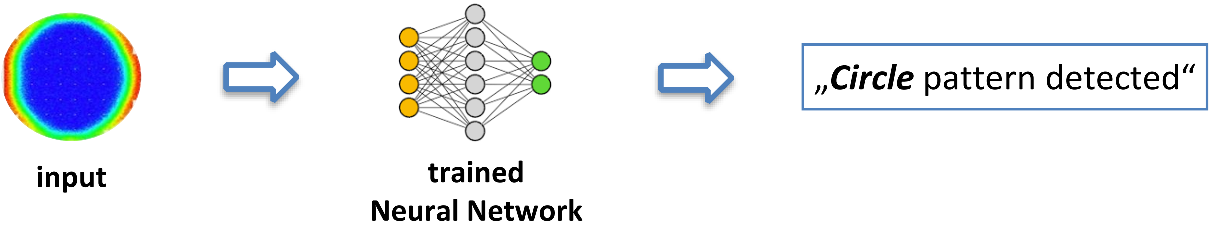
Scope of the Master Thesis

Already existing: **Machine-Learning** pipeline already *verified…*

- … due to few **labelled** data available in real-world applications
- … and **high variety** of pattern types
- … with **good performance** for few training data

Aim of master thesis: Implementation of a Deep Learning pipeline under certain restrictions, e.g.

- few training data (~**20 wafermaps** per pattern type)
- as low computation time as possible
- better classification performance compared to Machine Learning pipeline



Classification Task

1. Data Aquisition
2. Description of Data Set
3. Classification Task
4. Desired Outcome

Data Aquisition

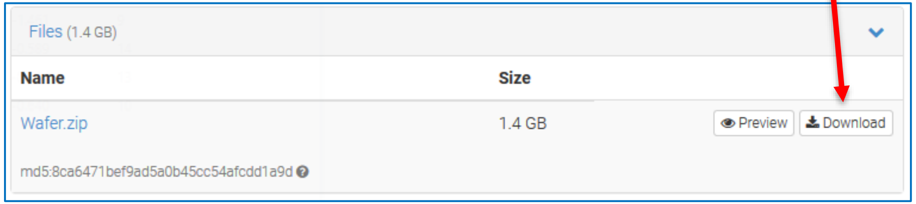
- The data set is publicly available: https://zenodo.org/record/2542504#.X3rgiO2xXqY
- Information on WHF methodology: https://ieeexplore.ieee.org/abstract/document/8914309

Paper Title: *A Health Factor for Process Patterns Enhancing Semiconductor Manufacturing by Pattern Recognition in Analog Wafermaps*

Detailed information of the following points can be found on next slides

- Data set generation description
- Brief description of occurring pattern types
- All necessary files
  - Data Set: *Wafer.zip*
  - Data set description as text file
  - 4 files with training data sets (training_*x*.csv)
  - 1 file with test data set (test.csv)

Download data set



Description of Data Set (1)

Available in file: *Data Set Decription.txt*

- Bullet points 1-5: General description
- Bullet points 6-8: Data set structure description
    - 14 attributes (columns)
    - 1000 wafermaps (~17500 rows per wafermap)
    - Each wafer has 10 wafermaps

NOTE_1: Labels (ground truth) are available in **column headers** !!

NOTE_2: Pattern Types *xxx_pure* not relevant for analysis (as they contain the same wafermaps without noise)!

				Test values with noise					Original test values				
Lot	**Wafer**	**X**	**Y**	**Ring**	**Spot**	**Trend**	**Twospots**	**Crescent**	**Ring_pure**	**Spot_pure**	**Trend_pure**	**Twospots_pure**	**Crescent_pure**
Lot_1	1	1	67	0,033	6,48e-05	0,312	0,039	15,58	0,032	6,54e-10	0,285	0,039	-4,843e-17
Lot_1	1	1	68	0,032	-4,66e-06	0,288	0,040	15,64	0,032	8,76e-10	0,281	0,040	-5,07e-17
…	…	…	…	…	…	…	…	…	…	…	…	…	…
Lot_1	1	115	104	0,032	-4,09e-05	0,131	0,044	16,57	0,032	1,50e-10	0,130	0,044	0,420
Lot_1	2	1	67	0,016	0,000	0,8332	0,016	15,12	0,016	0,000	0,826	0,015	-4,33e-17
Lot_1	2	1	68	0,032	-4,09e-05	0,131	0,044	15,22	0,032	1,50e-10	0,130	0,044	0,420
…	…	…	…	…	…	…	…	…	…	…	…	…	…
Lot_1	2	115	104	0,016	3,67e-05	1,340	0,019	16,11	0,016	3,29e-09	1,399	0,019	0,393
…	…	…	…	…	…	…	…	…	…	…	…	…	…

Data of all wafermaps for wafer 1

Data of all wafermaps for wafer 2

Description of Data Set (2)

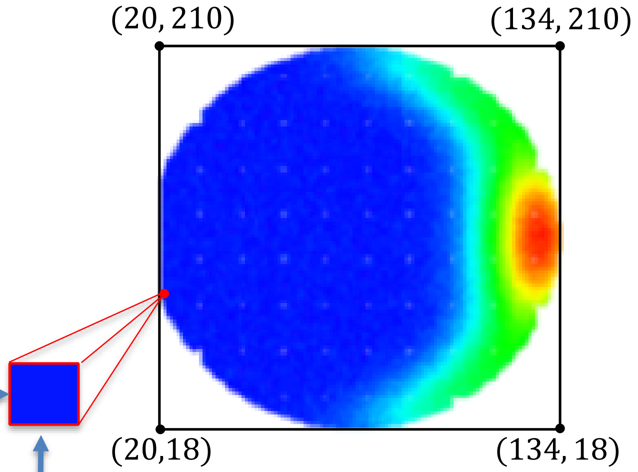
- How to get data of a specific wafermap, e.g. *Crescent* pattern of wafer 1

1. get all rows where *Wafer*==1

Lot	Wafer	X	Y	Ring	Spot	Trend	Twospots	Crescent
Lot_1	1	20	84	0,033	6,48e-05	0,312	0,039	15,58
Lot_1	1	20	85	0,032	-4,66e-06	0,288	0,040	15,64
…	…	…	…	…	…	…	…	…
Lot_1	1	134	121	0,032	-4,09e-05	0,131	0,044	16,57

2. get test values for each device by checking X/Y coordinates

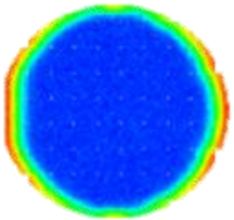
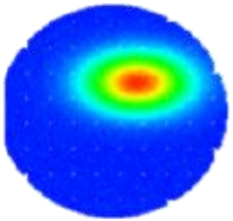
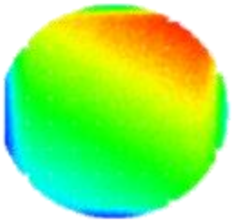
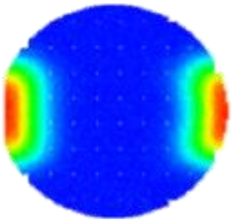
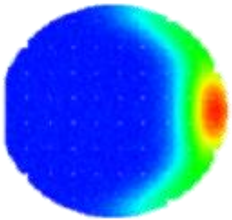
Lot	Wafer	X	Y	Ring	Spot	Trend	Twospots	Crescent
Lot_1	1	20	84	0,033	6,48e-05	0,312	0,039	15,58
Lot_1	1	20	85	0,032	-4,66e-06	0,288	0,040	15,64
…	…	…	…	…	…	…	…	…
Lot_1	1	134	121	0,032	-4,09e-05	0,131	0,044	16,57



Interpretation: *Wafer 1, value of* 5^{th} *test of device with position* $(20,18)$ *– the value 15,58 gets printed as a blue-ish color.*

Classification Task

- **Pattern Types**

	Ring	**Spot**	**Trend**	**Twospots**	**Crescent**
Example Wafermap					
Characteristics	Ring around wafermap • Size can differ but not position	Single spot on wafermap • Size, position and orientation can differ	Trend over whole wafermap	Two spots on opposite sides of wafermap • Size can differ	Single spot on one side of wafermap with crescent • Size can differ

Desired Outcome

- **Tasks:**

1. Implementation of a **Neural Network model** which can reasonably distinguish between the pattern types *Ring, Spot, Trend, Twospots* and *Crescent* by using a *maximum of **20 wafermaps** per pattern type* from *training_x.csv* in the training phase

2. Incorporation of **Data Augmentation Techniques** in order to increase the number of **relevant** data samples. These techniques include, but are not limited to:
   - Fliping, Rotating, Translating the images
   - more sophisticated methods like SMOTE

3. Prediction of all pattern types on **all provided wafermaps** in file *test.csv* and compare these predicted labels to the provided labeling (respective column headers)
   - Evaluation measure(s): F1 score[1] (**mandatory**) and others (optional)

[1] https://en.wikipedia.org/wiki/F1_score: $F1 = 2 * \frac{Precision * Recall}{Precision + Recall}$

Enhancement of existing ML pipeline via DL approaches for the detection of patterns in wafermaps

ARROWHEAD TOOLS