

# Inzichtelijk luik

## 1. Data Preprocessing

Standardisatie van de dataset:

- size
- color space
- afbeelding formaat
- resolutie

## 2. Kenmerk extractie

Bepaal relevante kenmerken/eigenschappen in de afbeeldingen die kunnen helpen bij het clusteren:

- kleuren
- texturen
- vormen
- ...

(optioneel) Dimensionality reduction:

- PCA
- LDA
- ICA

Optionele stap, Data augmentatie

Deze stap is enkel nodig als er niet genoeg data in de dataset beschikbaar is

## 3. Clustering algoritme

- K-means
- DBSCAN
- Hierarchische clustering

## 4. Aantal clusters bepalen

Om het optimale aantal clusters te bepalen voor het gekozen algoritme kan je elleboogmethode of de silhouetmethode gebruiken.

## 5. Clusteren afbeeldingen

Nu kan je het gekozen algoritme toepassen met het optimale aantal clusters bepaald in vorige stap.

## 6. Categorizeer de clusters

Begrijp welke soorten defecten worden gedetecteerd. Benoem en identificeer welke afbeeldingen bij elke cluster behoren, dit helpt met het categoriseren van de defecten.

## 7. Evaluatie en fine tuning:

Beoordeel de kwaliteit van de clusters en consistentie van de clusters. Pas zo nodig de kenmerk

selectie. extractie of het aantal clusters aan en voer opnieuw uit.

#### 8. Defectclassificatie:

Eenmaal je tevreden bent met de clusters, kun je een classificatiemodel trainen om nieuwe afbeeldingen te classificeren in een defect of niet en welk soort.