MASTER THESIS

ANALYSIS OF DATA SYNCHRONIZATION METHODS AND FRAMEWORK IN WIRE ARC ADDITIVE MANUFACTURING (WAAM)

Computer Sciences, 45483

Student: Sergejs Kopils (st.c.83519)

Supervisor: Professor, Dr. sc. ing. Mihails Savrasovs

Consultant: Invited lecturer, Mg. sc. ing. Arsenii Kisarev

CONTENT

MOTIVATION & RELEVANCE

- WAAM generates complex, asynchronous data (Scan, Process, Initial) without consistent timestamps.
- Existing QA manual processes are , time-consuming, and error-prone.
- Industry 4.0 demands synchronized and explainable data.

RESEARCH AIM AND QUESTIONS

- **Aim**: To develop a modular framework for synchronizing Scan, Process, and Initial data streams in WAAM systems and to detect anomalies in the deposited geometry using rule-based logic and interactive visualization.
- **Research Questions**:
- 1. What type of synchronization method is most reliable without timestamps?
- 2. How effective is geometry-based segmentation vs clustering?
- 3. Can rule-based logic detect anomalies from geometry data?

OBJECT AND SUBJECT OF THE RESEARCH

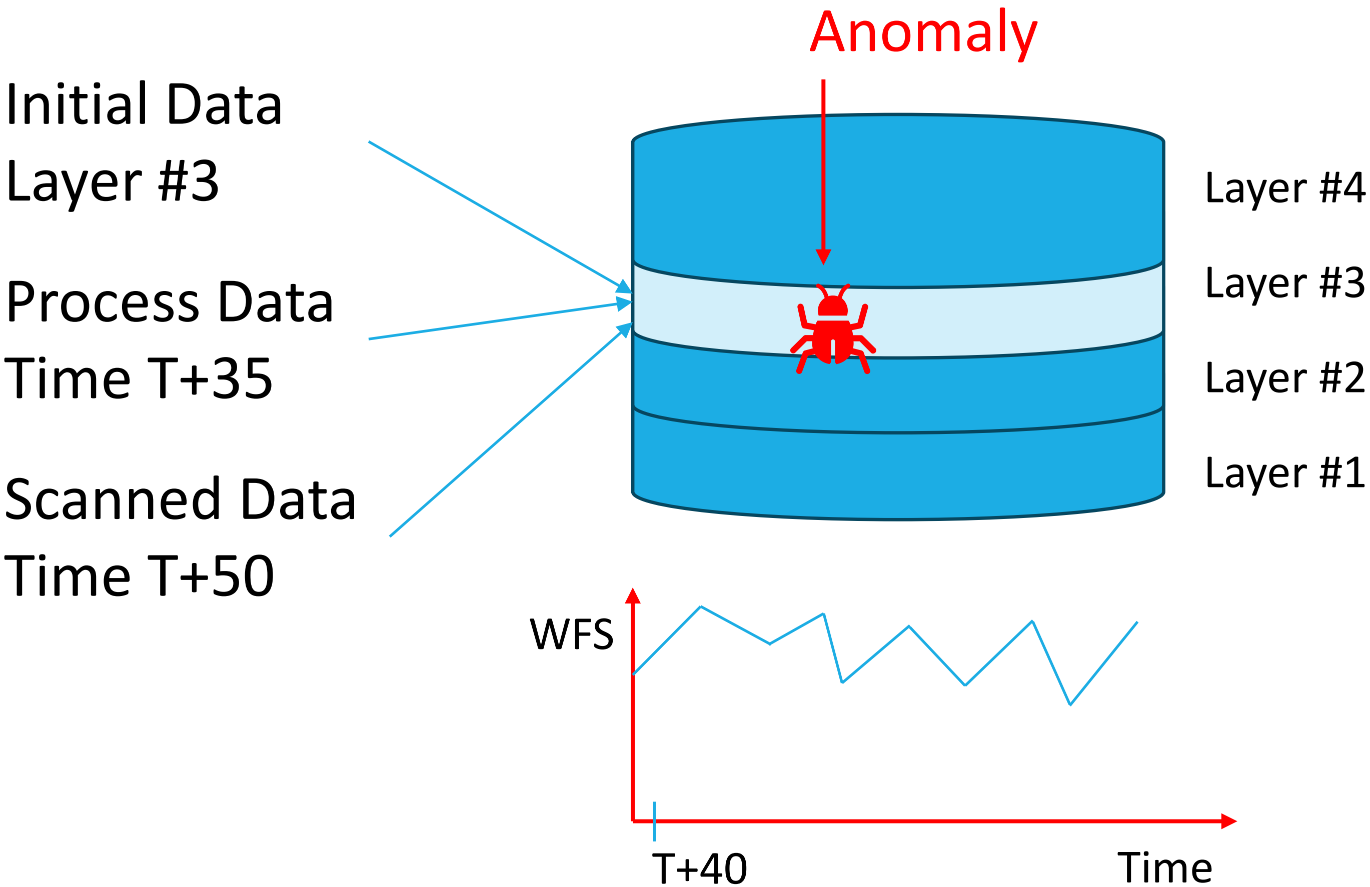
- Object: Multi-stream data collected during WAAM-based additive manufacturing processes.
- Subject: Synchronization methods, segmentation techniques, and anomaly detection algorithms applied to scan-process-initial data integration.

METHODOLOGY OVERVIEW

- Multi-stream input: **Scan** data, **Process** data, **Initial** data
- Modular structure: segmentation, synchronization, anomaly detection, visualization
- Implemented in Python 3.11 using open-source libraries

EXPECTED RESULT

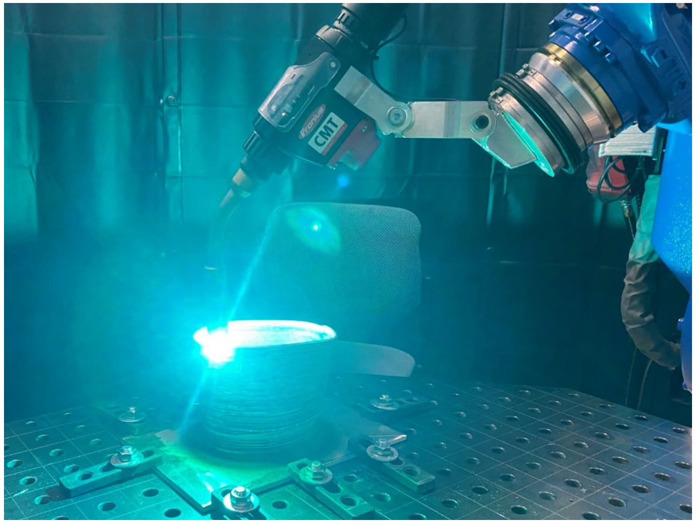
A functioning framework that allows accurate synchronization of multi-stream data and reliable detection of anomalies in deposition geometry

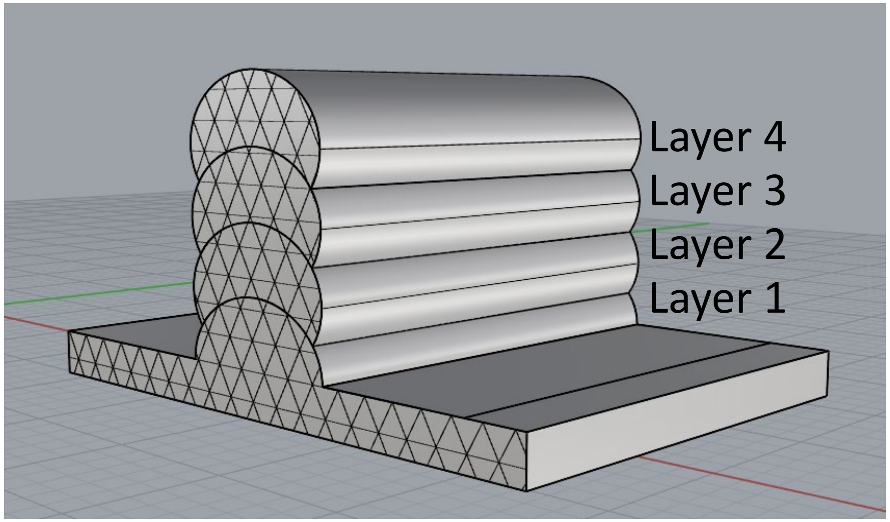


WIRE ARC ADDITIVE MANUFACTURING (WAAM) TSI ADDITIVE LAB

WAAM PROCESS

3D printing Layer by Layer. From different grades of steel.



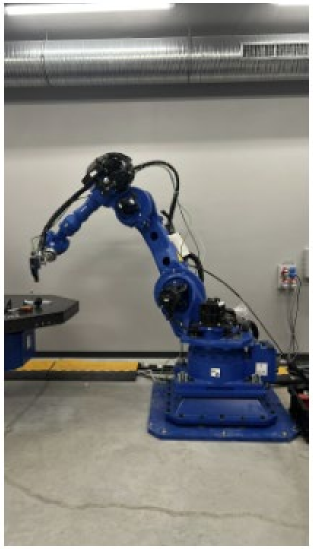


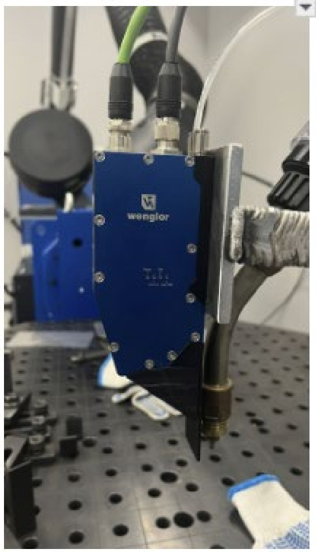
Functional concept “Layer"

DATA SOURCES

The experimental data was collected in TSI Additive Lab equipped:

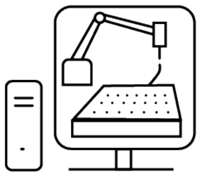
- Industrial robot (Yaskawa)
- Welding power source (Fronius)
- Laser scanner (Wenglor)





DATA TYPES

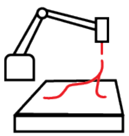
Slicer



INITIAL

- Layer
- Path Planning (X, Y, Z, Rx, Ry, Rz)
- TS (Travel Speed, mm/min)
- WFS (Wire Feed Speed)
- Job Nr.

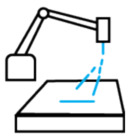
Process On



PROCESS

- Robot Coordinates (Path)
- TS
- WFS
- DateTime
- Weld Current
- Weld Voltage

Scanning



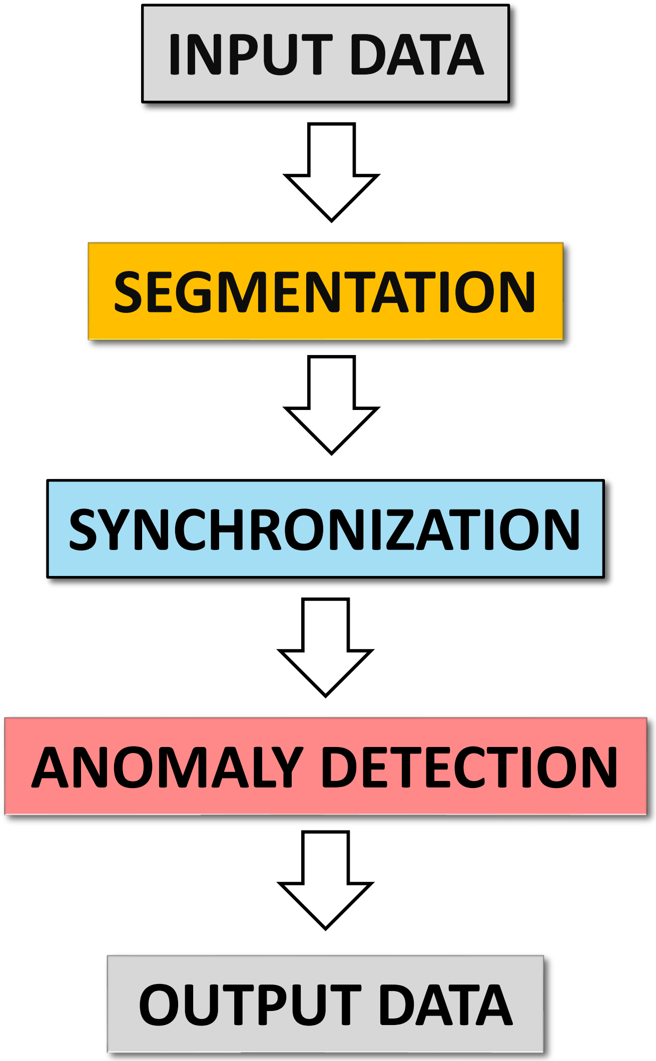
SCAN

- DateTime
- Laser Coordinates (X, Z)
- Intensity

FRAMEWORK ARCHITECTURE

General architecture of the WAAM segmentation, synchronization and anomaly detection framework.

From raw Input Data to Output Data



REFERENCE METHOD WORKFLOW

Key Points:

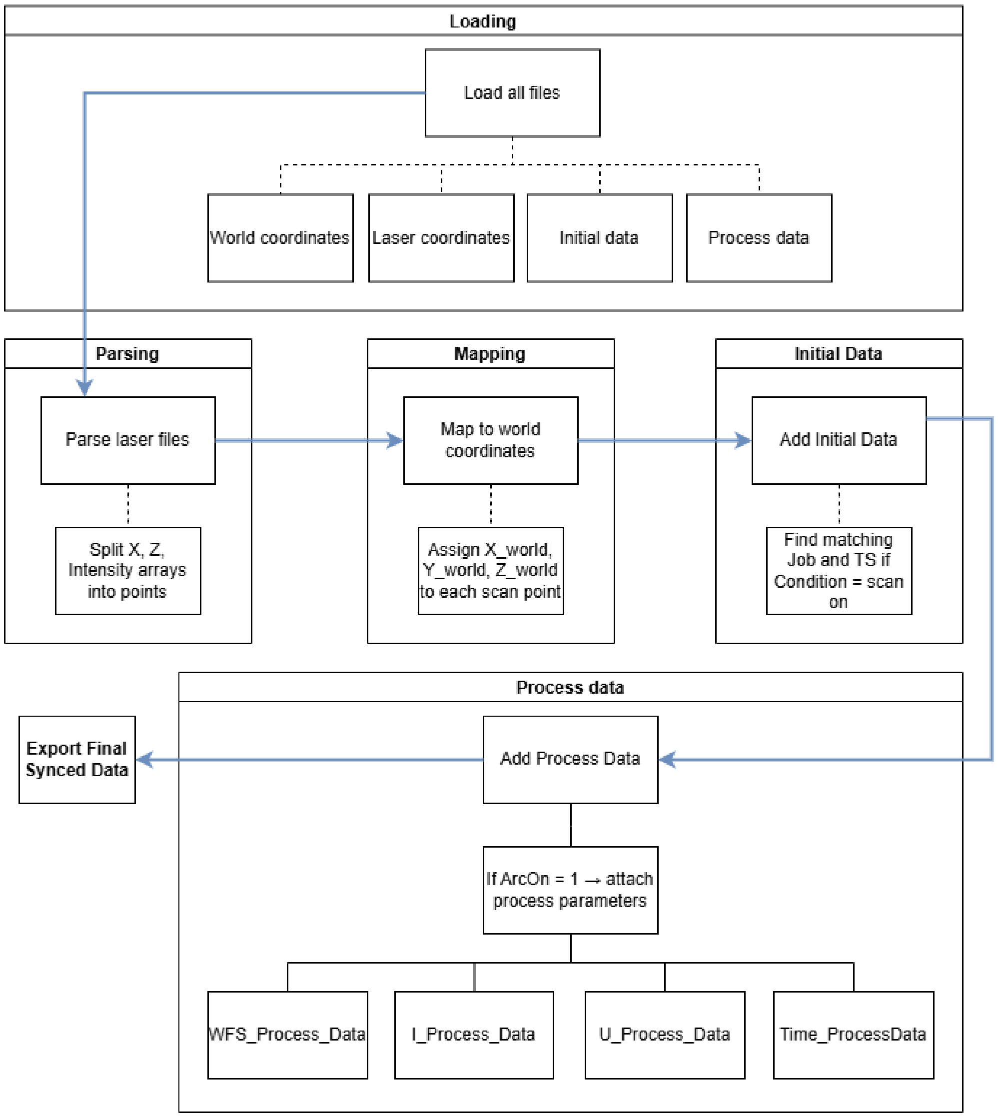
Laser scan data is parsed into individual points.

Mapped to world coordinates using transformation matrices.

Initial data is matched based on scan start conditions.

Process parameters are added.

Final output is a synchronized dataset.



REFERENCE LABELS FOR EVALUATION

Scan File	Scan Layer ID	Points Start Index	Points End Index	Num Points
LaserCoordinate1.csv	0	0	1244	1245
LaserCoordinate2.csv	1	1245	2679	1435
LaserCoordinate3.csv	2	2680	4077	1398
LaserCoordinate4.csv	3	4078	5461	1384
LaserCoordinate5.csv	4	5462	6872	1411
LaserCoordinate6.csv	5	6873	8319	1447
LaserCoordinate7.csv	6	8320	9700	1381

This method produces a labeled point cloud with absolute confidence in layer structure and is used as the ground truth reference for evaluating the performance of other clustering and synchronization strategies.

SEGMENTATION METHODS

Five methods were evaluated and compared against a reference **Method 1: Ground Truth** :

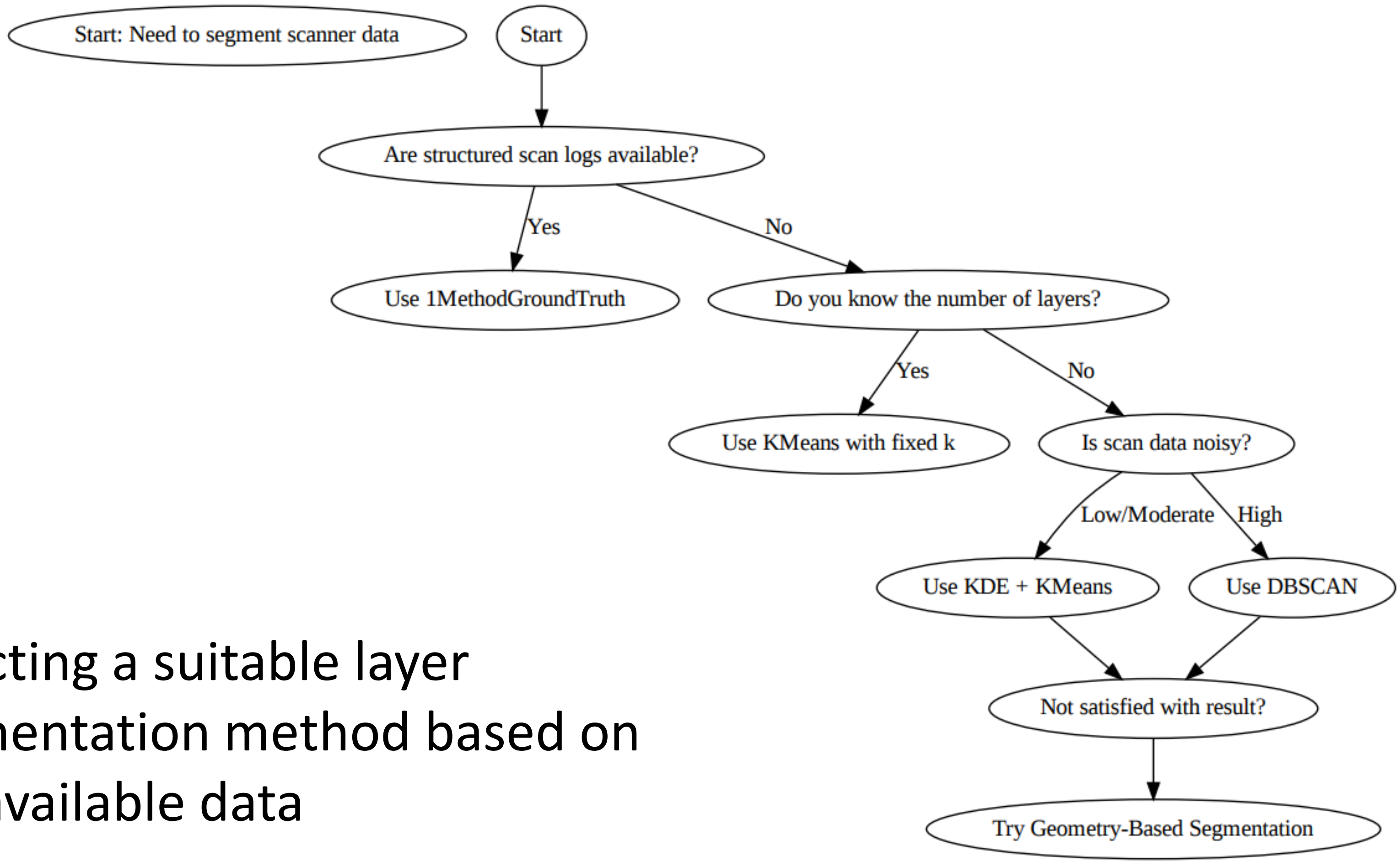
- **Method 2:** Cluster K-means Known – clustering by Z using a predefined number of clusters corresponding to expected layers;
- **Method 3:** Cluster K-means Auto – automatic layer detection using kernel density estimation and peak-based K-Means clustering;
- **Method 4:** Cluster DBSCAN Auto – density-based clustering of Z-values using DBSCAN, requiring no prior knowledge of cluster count;
- **Method 5:** Geometry Based – segmentation based on spatial jumps in Euclidean distance between consecutive scan points;
- **Method 6:** Hierarchical Auto – hierarchical clustering on Z-values with automatic dendrogram cutoff.

COMPARISON OF LAYER SYNCHRONIZATION METHODS

Method	Type	Automatic	Uses Process Logs	Needs n_layers	Robust to Noise	Sensitive to Parameters
2MethodClusterKMeansKnown	Clustering	No	No	Yes	Low	Yes
3MethodClusterKMeansAuto	Clustering	Yes	No	No	Medium	Yes
4MethodClusterDBSCANAuto	Clustering	Yes	No	No	High	Yes
5MethodGeometryBased	Geometric Segmentation	Yes	No	No	Medium	Yes
6MethodHierarchicalAuto	Clustering	Yes	No	No	Medium	Yes

- Each method is evaluated based on:
  - automation level,
  - reliance on process logs,
  - requirement of known number of layers,
  - robustness to noise,
  - sensitivity to parameter tuning.

DECISION TREE



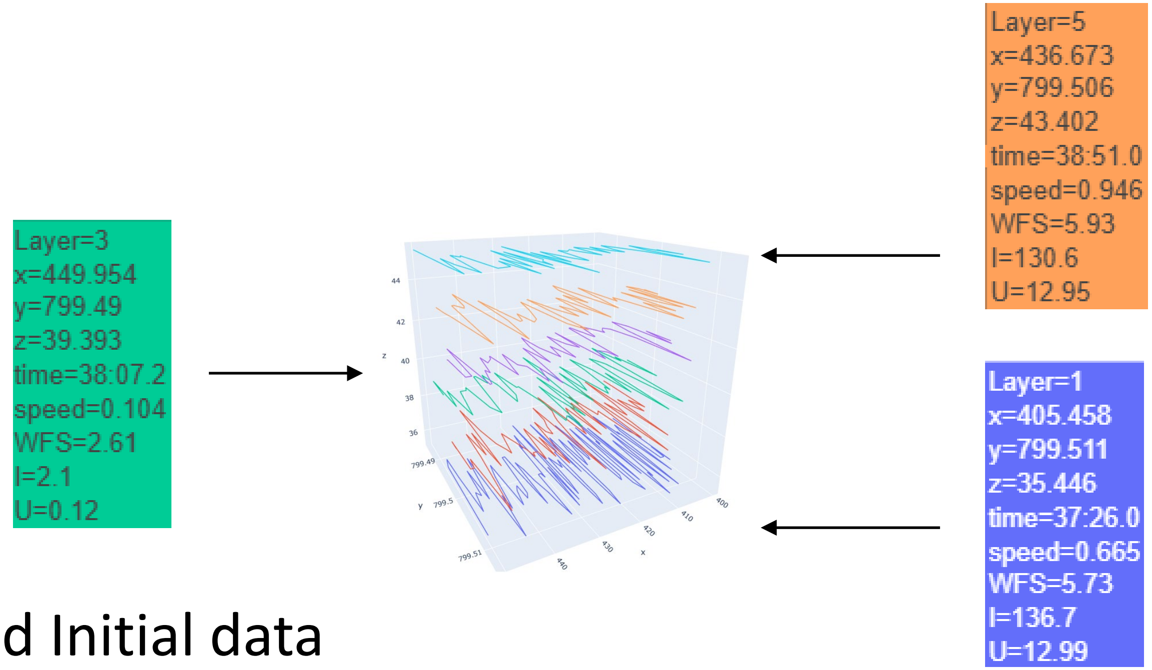
Selecting a suitable layer segmentation method based on the available data

COMPARISON OF SEGMENTATION METHODS

- Geometry-based method 5 achieved perfect alignment (100%); clustering methods showed varied and lower performance.

Method	Accuracy (%)	Precision	Recall	F1-score	IoU
2MethodClusterKMeansKnown	57.1	53.9	49.3	51	36.7
3MethodClusterKMeansAuto	14.4	4.7	16.7	7.4	4.7
4MethodClusterDBSCANAuto	14.4	2.4	16.7	4.2	2.4
5MethodGeometryBased	100	100	100	100	100
6MethodHierarchicalAuto	55.3	53.5	48	49.1	34.8

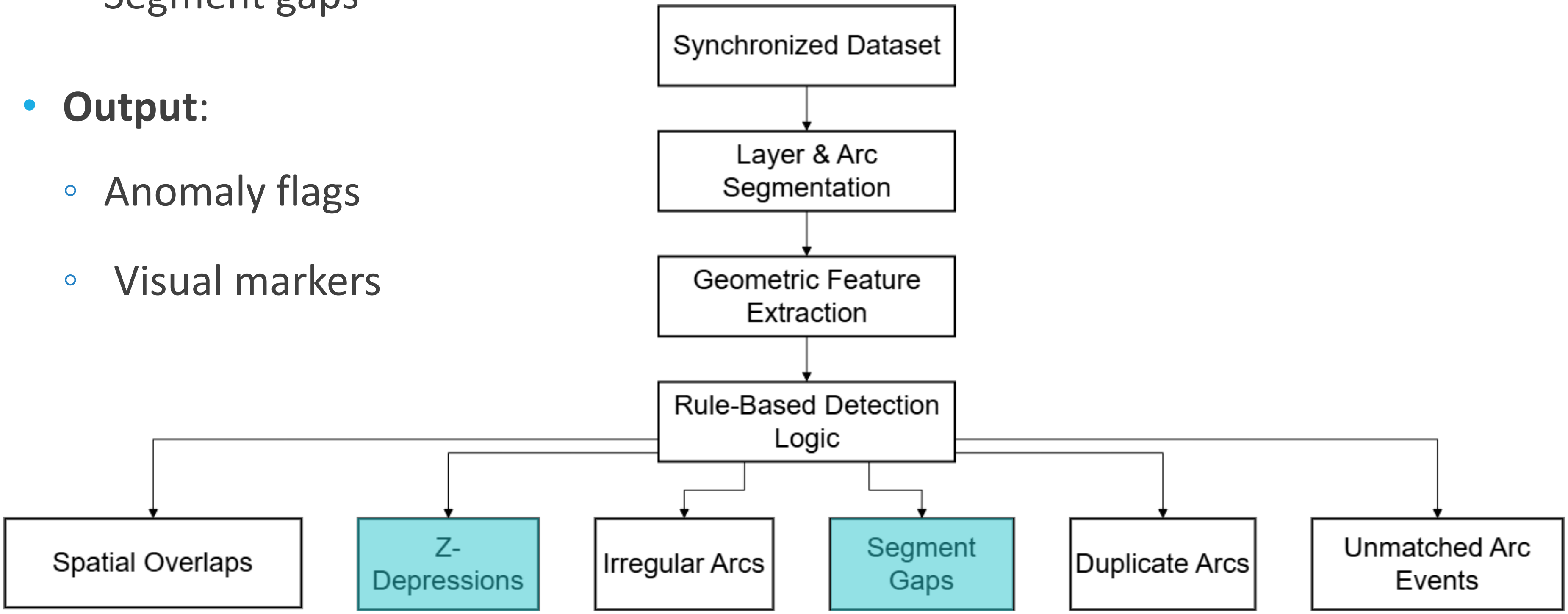
SYNCHRONIZATION DATA



- Process and Initial data synchronized with Scan points.
- Process data Layers segmented using DBSCAN along the Z-axis.

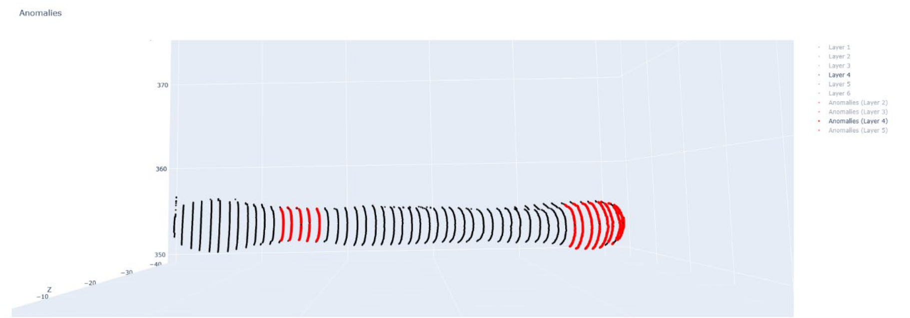
THE ANOMALY MODULE

- **Rule-based engine**:
  - Z-depressions
  - Segment gaps
- **Output**:
  - Anomaly flags
  - Visual markers

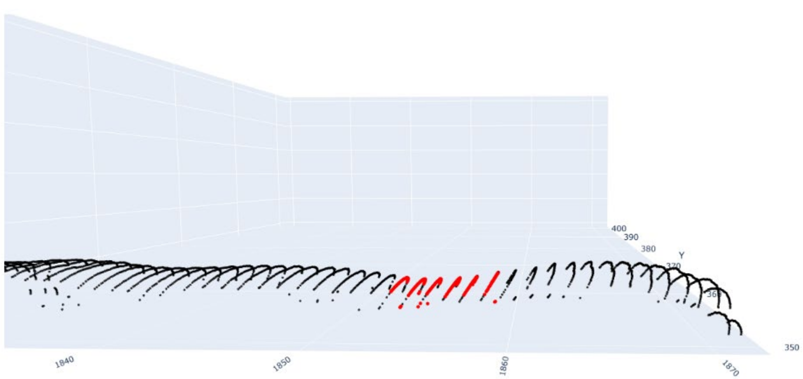


VISUAL INSPECTION

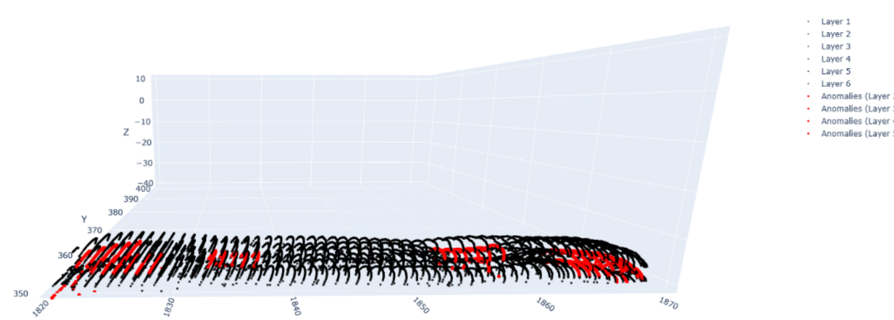
- Features:
  - 3D layer-by-layer inspection
  - Toggle visibility, anomaly overlays



top-down view of Layer 4



Side view of layer 5



Side view of all layers

ANOMALY DETECTION SUMMARY

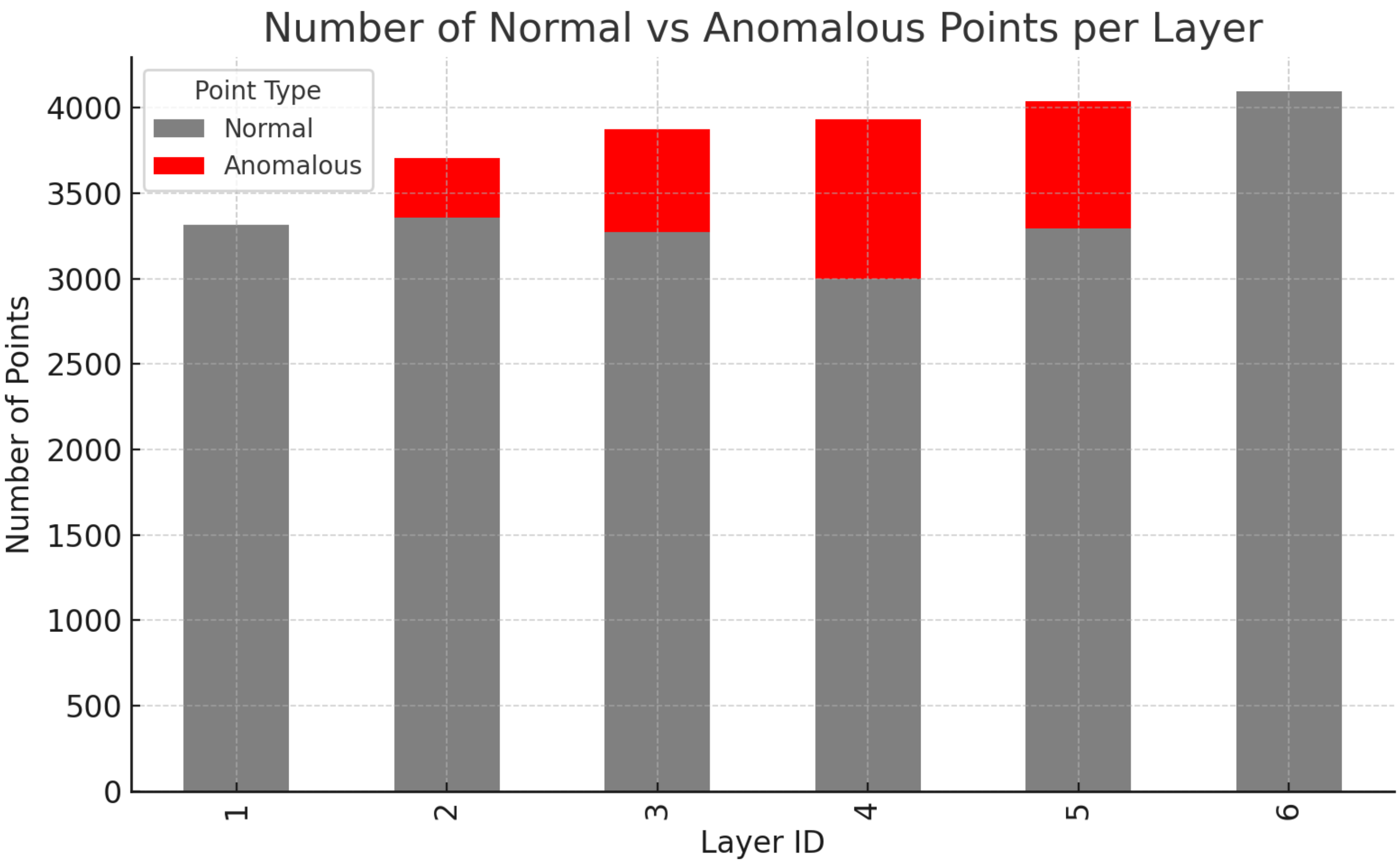
- 2,625 anomalies flagged
- Detected issues: Z-depressions, gaps
- Most errors found in Layers 3–5

Number of anomalous points per layer

Layer	Normal Points	Anomalous Points
1	3,316	0
2	3,360	344
3	3,274	599
4	2,999	937
5	3,296	745
6	4,095	0

DISTRIBUTION OF ANOMALIES

Visual confirmation that most anomalies are concentrated in mid-layers.



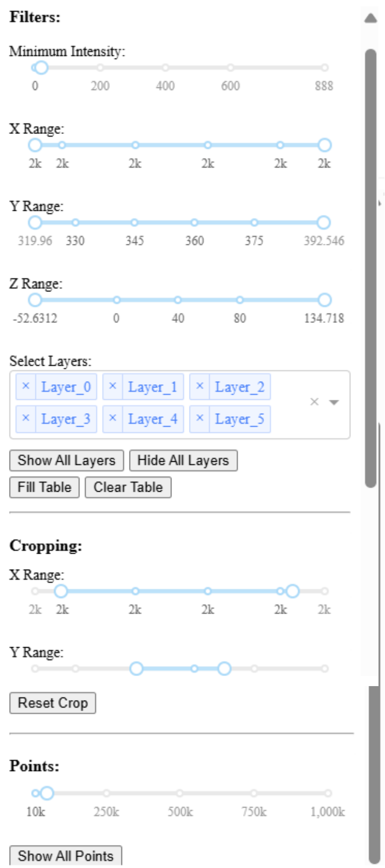
OUTPUT DATA SCRUCTURE

- The final output is a unified pandas DataFrame where each **Scan** point is enriched with both **Initial** and **Process** and **Anomaly data**.

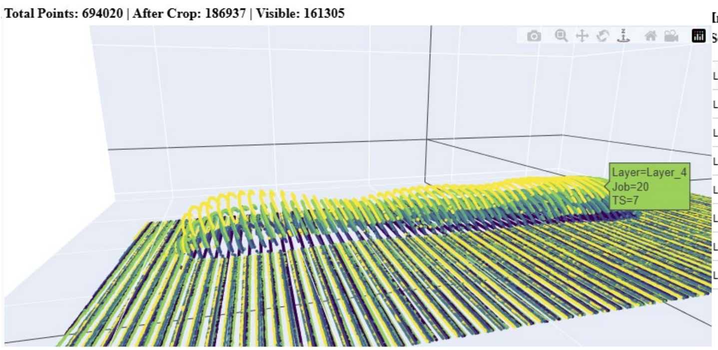
Column	Description
X, Y, Z	3D coordinates of the scan point
layer_id_*	Layer label from selected segmentation method
ArcOn	Arc status at that point (1 or 0)
JobID	Job number from control metadata
TS	Travel speed (mm/s)
anomaly_flag	Detected anomaly at point level (if any)

FINAL DASHBOARD INTERFACE

Unified interface for interactive inspection of scan data, process parameters, and detected anomalies across deposition layers.



Laser Scan Data

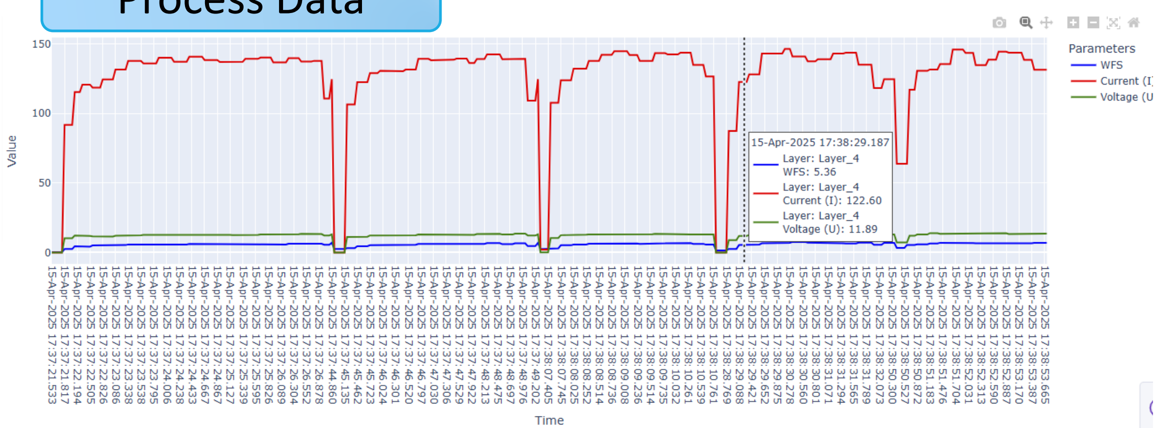


Initial Data

Initial Data
Selected Layers:

Layer	Initial_Job	Initial_TS
Layer_0	20	7
Layer_1	20	7
Layer_2	20	7
Layer_3	20	7
Layer_4	20	7
Layer_5	20	7
Layer_6	20	7

Process Data



FUTURE WORK

- Real-time integration
- Feedback-driven process correction
- Application in robotic welding, CNC, quality control

CONCLUSION

- Modular, timestamp-independent framework created
- Accurate and scalable segmentation
- Practical anomaly detection
- Adaptable to various industrial contexts

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to:

- **Dr.sc.ing., Professor Mihails Savrasovs** – for his invaluable supervision, continuous support, and academic guidance throughout this research.
- **Dipl. Engineer Arseniy Kisarev, Head of the TSI Additive Lab** – for his expert consultations and technical insights related to the WAAM process.
- **The AdditiveLab team at TSI** – for providing access to equipment, data, and collaborative support during the project.
- **Transport and Telecommunication Institute (TSI)** – for creating the academic and technical environment that made this work possible

Thank you!

1. How the data of Table 6 (Precision, Recall, F1-score , and especially IoU) it was obtained? How to calculate IoU?

- The metrics in Table 6 (Precision, Recall, F1-score, and IoU) were obtained by comparing the predicted layer assignments from each method (e.g., KMeans, DBSCAN, Geometry-based) with the reference labels from the Ground Truth method (Method 1), on a per-point basis.

1. How to calculate IoU?

- IoU (Intersection over Union) is a metric that measures how well two sets of points overlap.
  - One set is the predicted points for a layer (from a clustering method like KMeans or Geometry-based).
  - The other set is the ground truth points for that layer, obtained from structured scan files.

$$IoU = \frac{|\text{Intersection}|}{|\text{Union}|}$$

- Intersection: Number of points that appear in both the predicted and ground truth layer.
- Union: Total number of unique points that appear in either the prediction or the ground truth.

2. In what form the anomalies (Arc absence; Arc duration anomalies; Uncovered regions; Geometric jumps; Duplicated arcs) from p.45 are included into the Metrics for evaluation?

- The anomaly types listed on page 45 represent a conceptual overview of possible deviations in the WAAM process.
- Only a subset of anomalies was implemented and evaluated in the final framework using a rule-based detection module, as described in Section 3.3 and 4.3.

2. Metrics for evaluation?

- Out of the 5 theoretical anomaly types, only 2 were implemented in research:
    - Z-depressions: based on vertical deviations below a defined Z threshold.
    - Segment gaps: identified via large spatial discontinuities in X between arcs.
- These were evaluated per arc and recorded as a binary anomaly flag (True / False) in the final dataset (DetectAnomalies.csv).
- Anomalies are not used for numeric evaluation because there is no reference dataset; they are shown to help understand and visualize process issues.

3. How to estimate the "predefined precision threshold δ\deltaδ:" (p.51, expression (1))?

- In the presented method (Method 5 – Geometry-Based Scan Segmentation by Distance Jumps), the threshold δ defines the minimum Euclidean distance between two consecutive 3D points that triggers the start of a new segment or arc.
- The value of δ is not constant across all datasets—it depends on scan characteristics. In the current implementation, δ was empirically set to 1.0 mm. This estimation was based on the average inter-point distance within continuous scans (i.e., regions without deposition gaps or robot retractions).

4. The best Geometry Based Method shows 100% accuracy, precision, recall, and IoU (Conclusions 5.1, p.68). Are there any limitations or conditions for these absolute results? Or is it always present?

- Conditions for these absolute results:
  - The scanner captured one layer per scan file or in the correct sequential order.
  - No points were duplicated in the scan files.
  - The process had no overlapping or partial layers.
- Limitations:
  - If scanner data nonsequential, the method may over-segment or under-segment.
  - If layer transitions are smooth (e.g., no jump in geometry), this method may fail to detect boundaries.

4. Or is it always present?

- Conclusion:
  - While this method achieved absolute results accuracy on the test dataset, it is not universally guaranteed. Its performance depends on scanner behavior, data integrity, and point density.