Journal of Engineering and Technology

ISSN 2180-3811 eISSN 2289-814X https://jet.utem.edu.my/jet/index

DOI: https://doi.org/10.54554/jet.2025.16.2.020

VISUAL DATA ACQUISITION USING YOLO AND OCR FOR SPUTTRING PROCESS MONITORING

N. Y. Loong[1,2], S. L. Kok*[1,2], S. A. Radzi[1,3] and S. P. Tong[4]

[1] Faculty of Electrical and Electronic Engineering Technology, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia.

[2] Advanced Sensors & Embedded Controls System (ASECs), Centre for Telecommunication Research and Innovation (CeTRI), Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia.

[3] Machine Leaning & Signal Processing (MLSP), Centre for Telecommunication Research and Innovation (CeTRI), Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia.

[4] Texas Instruments Electronics Malaysia, Free Trade Zone, Batu Berendam, 75350 Melaka, Malaysia

corresponding: sweeleong@utem.edu.my

Article history:
Received Date: 1 October 2025
Revised Date: 6 December 2025
Accepted Date: 24 December 2025

Abstract— This paper presents a real-time data visualization and fault detection model for sputtering process monitoring, focusing on parameters such as deposition rate, film thickness, material types (Ti, Ag, Ni), and process status. The objective is to model multi-level anomaly outliers for detecting potential OCR errors and sputtering process deviations, develop a real-time embedded

Keywords:
Machine Learning, Sputtering Process, Fault Detection, YOLO, OCR

vision system for automated data acquisition from the sputtering equipment's display, and implement a complete monitoring application with real-time visualization and post-process analysis for industrial deployment. Data acquisition is carried out using a high-resolution camera, where YOLO achieves 99.5% mAP@0.5 in supervised detection of visual indicators, and PaddleOCR attains 99.57% accuracy in extracting numerical parameters. Preprocessing incorporates a median filter to suppress noise, while DBSCAN identifies sudden OCR fluctuations and linear regression models parameter trends. The postprocessed data are stored in structured CSV files. By integrating supervised and unsupervised learning with data science techniques, the proposed system enables reliable monitoring, early anomaly detection, and predictive maintenance in industrial sputtering operations.

I. Introduction

In Industry 4.0, real-time sensor data acquisition is essential for predictive maintenance, anomaly detection, and process optimization in manufacturing systems [1-2]. Sputtering processes in semiconductor and thin-film coating require precise monitoring of parameters like power, deposition rate, and film thickness. The sputtering process involves different materials, such as Titanium (TI), Silver (AG), and Nickel (NI), each with unique deposition characteristics. Many legacy sputtering machines lack remote data access, making hardware retrofitting impractical.

This paper proposes vision-based monitoring system that captures process parameters directly from machine display panels using YOLO-based object detection [3-4] and Optical Character Recognition (OCR). The system leverages the computational power of NVIDIA Jetson Orin Nano for real-time image processing, enabling efficient data extraction. Initially, YOLO localizes and classifies visual indicators on the display, while OCR extracts numerical parameters such as power, deposition rate, and film thickness. To enhance the data quality, the system applies median filtering for noise reduction. It then uses DBSCAN (Density-Based Spatial Clustering of Applications with Noise) for anomaly detection, identifying sudden OCR reading jumps and distinguishing genuine faults from normal fluctuations. Finally, linear regression models the deposition behavior to track trends and identify machine-level faults. The acquired, time-stamped data is structured in CSV format, enabling real-time fault detection and informed decision-making.

The main contributions of this work can be summarized as follows: the development of a non-invasive, embedded vision-based system for real-time data acquisition from sputtering equipment, the implementation of a multi-level anomaly detection framework combining OCR fault detection and machine fault modeling, and a comprehensive performance evaluation of detection models, OCR engines, and anomaly detection methods to benchmark accuracy, efficiency, and robustness.

II. Literature Review

Optical Character Recognition (OCR) has evolved from template-based methods to learning-based systems capable of interpreting text in unconstrained scenes and documents [5-6]. Neural networks and fuzzy classifiers outperform traditional templates, handling diverse fonts, distortions, and environmental variations [6]. Deep learning pipelines, including CNN-RNN-CTC and region-based detectors, achieve real-time recognition on datasets such as ICDAR and SVT [7].

Applications like Automatic License Plate Recognition (ALPR) illustrate OCR pipelines relevant to industrial panels, addressing challenges such as lighting variation, occlusion, and motion blur [5][8]. Real-time frameworks using External Regions and two-stage classification show techniques adaptable to machine display monitoring [7].

OCR has been reported to be applied to numeric instrument displays such as seven-segment LEDs, LCDs, and character panels. Early work targeted measuring instruments [11], while recent studies extend to medical devices [12]. These methods combine geometric analysis, preprocessing, and machine-learning classifiers for high accuracy.

In industrial settings, OCR enables data captured from legacy equipment lacking digital interfaces. Camera-based monitoring with Tesseract or similar engines, coupled with edge devices, converts visual readings into structured telemetry for MES, dashboards, and OEE/OEU monitoring [9-10, 13]. Hybrid computer-vision approaches are also being used to read analog gauges by detecting needles and scale marks [14].

OCR's advantages include non-invasiveness, low cost, and rapid deployment. Limitations include sensitivity to lighting, reflections, viewing angles, and displaying refresh artifacts. Successful deployment requires careful imaging, preprocessing, and validation.

Although effective, gaps remain for OCR deployment, which includes robustness under factory conditions, limited industrial datasets, and real-time telemetry extraction [15]. Despite these, OCR is a practical, cost-effective solution for acquiring measurement data from machine displays where direct digital acquisition is unavailable.

III. Methodology

A. Flow Chart

In the process begins with Data Acquisition, where a high-resolution camera captures real-time images of the sputtering machine's human-machine interface (HMI). These images contain key operational parameters, such as power, deposition rate, and film thickness. To extract these parameters, the system utilizes YOLO (You Only Look Once), an object detection model that localizes and classifies visual indicators on the HMI screen. This non-invasive approach enables accurate data extraction without requiring additional sensors or hardware modifications, providing a seamless way to monitor critical process metrics.

Once the data is acquired, the next phase is Data Cleaning, which ensures that the extracted data is accurate and free from noise. In this phase, OCR (Optical Character Recognition) is employed to recognize and clean the text and numerical values displayed on the HMI. For example, OCR extracts parameters like power, deposition rate, and film thickness from the image. This step corrects any errors or distortions caused by variations in font, lighting, or other environmental factors, ensuring that the data is ready for further analysis.

After the data is cleaned, the system proceeds with Data Modeling, which focuses on identifying any anomalies or faults in the sputtering process. The system uses the DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm to detect sudden fluctuations or outliers in the OCR data, signaling potential process irregularities. Additionally, linear regression models are employed to track trends in key parameters, such as deposition rate and film thickness, over time. By integrating these methods, the system can perform real-time anomaly detection and predictive maintenance, ensuring efficient and reliable operation of the sputtering process.

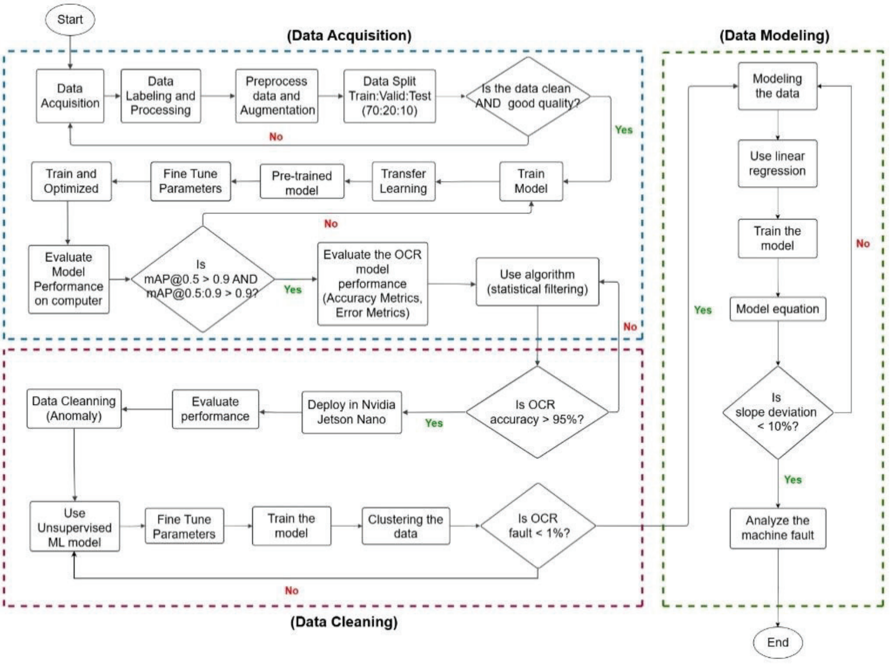

Figure 1: Flow Chart

B. Data Collection

Standard practice in machine learning research adopts dataset splits such as 70:20:10 or 80:10:10 to achieve a balance between sufficient training data and reliable evaluation sets for model tuning and assessment [16]. In this study, a 70:20:10 split was used, with 2,877 images for training, 800 for validation, and 417 for testing, shown in Figure 2.

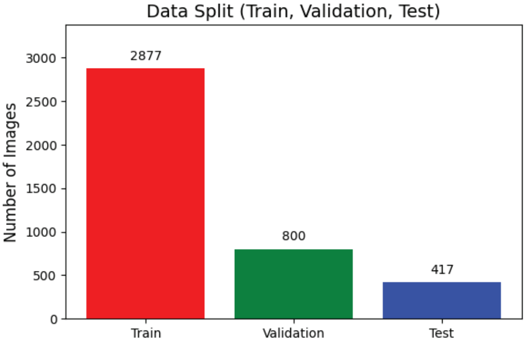

Figure 2: Data Split (Train, Validation, Test)

C. Training Process

Model training is conducted using transfer learning on pre-trained models, followed by fine-tuning and optimization. The training setup and parameters are summarized in Table 1. The training parameters were selected based on established practices in YOLO based object detection and transfer learning literature. Pre-trained YOLO weights were adopted to accelerate convergence and improve generalization on a moderate sized custom dataset. The image resolution, batch size, learning rate, and optimizer were chosen to balance detection accuracy, training stability, and computational efficiency, following commonly reported configurations in recent YOLO studies [17].

Table 1: YOLO Training Configuration Parameters

Symbol	Quantity
Task specification	detect
Model	yolo11s.pt
Dataset location	{dataset.location}/data.yaml
Number of epochs	100
Image size (imgsz)	640, 480
Batch size	16
Optimizer	AdamW
Learning rate	0.002
Early stopping	10
Weight decay	0.005

D. System Deployment

The method utilizes a high-resolution camera for data acquisition, where the captured images are localized and classified using YOLO. OCR is then applied to recognize key process parameters such as power, deposition rate, and film thickness. Anomaly detection is performed using DBSCAN to identify sudden fluctuations in the OCR data, while linear regression models track process trends over time. This integrated approach enables real-time monitoring and fault detection without hardware modifications as shown in Figure 3.

Figure 3: System Deployment

IV. Experiment Result

A. YOLO Model Results

Table 2 shows that YOLOv8s achieved 99.5% mean average precision (mAP@0.5) in detecting visual indicators, demonstrating its robustness and accuracy in the sputtering process monitoring system. While all models exhibited perfect precision and recall (1.0), YOLOv8s offers the best overall performance with its high mAP@0.5 and a strong mAP@0.5:0.95 of 0.907. Its balance of 11.13 million parameters and 168 layers ensures computational efficiency without sacrificing accuracy, making it the most suitable choice for real-time deployment. The model size of 21.4MB is manageable for high-speed processing, and its architectural efficiency supports stable and reliable operation in industrial environments. Therefore, YOLOv8s stands out as the optimal model for this application.

Table 2: YOLO Model Performance Comparison

Model	Precision	Recall	mAP@0.5	mAP@0.5:0.95	Parameters (M)	Layers	Model size (MB)
YOLOv5s	1	1	0.995	0.915	9.11	193	17.6
YOLOv8s	1	1	0.995	0.907	11.13	168	21.4
YOLOv11s	1	1	0.995	0.907	9.41	238	18.2
YOLOv12s	1	1	0.995	0.915	9.08	376	17.7

B. Anomaly Detection on OCR

Table 3 shows the 1000 video frames were extracted from the source video using Python. A script was implemented to capture every frame from the video, ensuring that a consistent sample size was obtained for analysis. Each frame was then processed through OCR (Optical Character Recognition) tools to extract the relevant process parameters. The extracted data was used to evaluate the accuracy and performance of different OCR tools, including PaddleOCR, EasyOCR, and PyTesseract. As shown in Table 3, PaddleOCR demonstrated the highest accuracy at 99.9%, significantly outperforming the other OCR tools in correctly reading values from the frames.

Table 3: OCR Performance Comparison for 1000 Video Frames

OCR Tools	Correctly read values by OCR	Incorrectly read values by OCR	Accuracy (%)	Error (%)
PaddleOCR	999	1	99.9	0.1
EasyOCR	483	517	48.3	51.7
PyTesseract	553	447	55.3	44.7

C. OCR Anomaly Detection

Figure 4 illustrates the use of DBSCAN for anomaly detection on the value parameter across sequential frames. The plot shows the overall trend of the data, with a sudden sharp drop occurring around frame 1500.

DBSCAN successfully detects this irregular deviation and flags the outlier, marking it with a red point on the graph. After the anomaly occurs, the values increase again and return to their expected pattern. This demonstrates DBSCAN's effectiveness in identifying and flagging abnormal points that deviate from the normal process behavior.

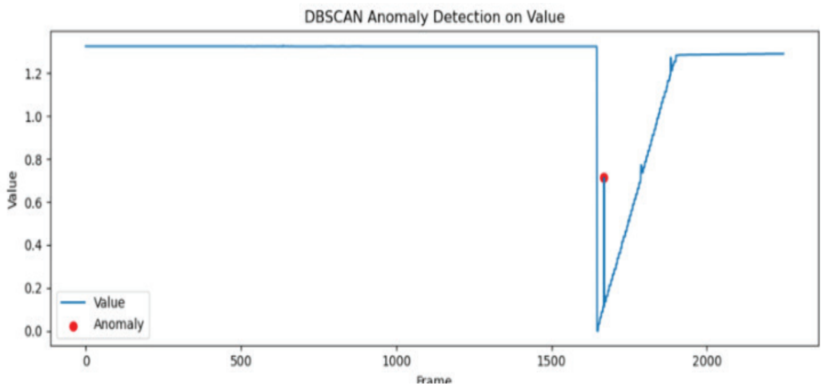

Figure 4: DBSCAN Anomaly Detection on Dataset

D. Machine Fault Detection

To evaluate machine-related faults, linear regression models were developed for three sputtering sources: Titanium (TI), Silver (AG), and Nickel (NI).

Figure 5 compares the fitted regression models. The graphs illustrate the gradient of each regression line reflects the deposition rate, where AG exhibited the highest growth rate of 0.0351 kÅ/s, followed by TI at 0.0245kÅ/s, and NI with the lowest rate of 0.0060kÅ/s. These differences demonstrate the source-dependent behavior of the sputtering process, where each material contributes distinct deposition dynamics. By establishing these regression models as reference baselines, subsequent runs can be compared to detect deviations in slope or growth rate, allowing for reliable identification of potential machine faults.

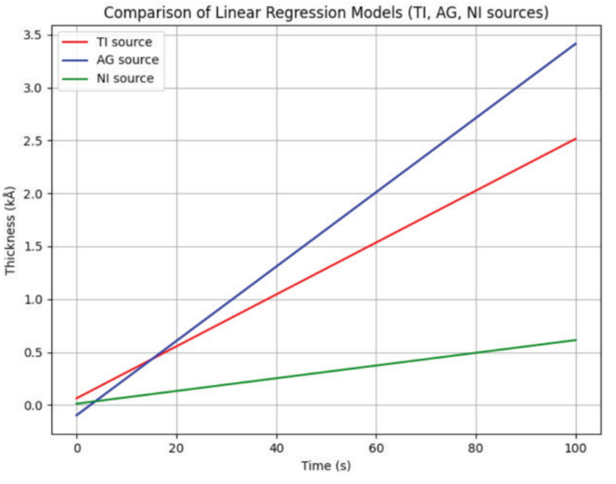

Figure 5: Comparison of Linear Regression Models (TI, AG, NI sources)

The fitted regressions are expressed as Equation (1) to (3).

For TI source:

$$k\text{Å} = 0.0634 + 0.0245t \quad (1)$$

For AG source:

$$k\text{Å} = 0.0972 + 0.0351t \quad (2)$$

For NI source:

$$k\text{Å} = 0.0126 + 0.006t \quad (3)$$

where:

kÅ = thickness

t = time in seconds

The general deposition model is defined as Equation (4).

$$Thickness = \beta_0 + Rt \quad (4)$$

where:

β_0 = y-intercept of the graph

R = deposition rate

t = time in seconds

Empirical analysis indicates that β_0 is negligible compared to linear growth term, allowing the model to be simplified as, Thickness ≈ *Rt*, which assumes a constant deposition rate effectively originating from zero.

V. Conclusion

Experimental results presented a real-time sputtering process monitoring system integrating YOLOv8s for display detection and PaddleOCR for text recognition, achieving over 99.9% accuracy. Anomaly detection using DBSCAN and linear regression proved effective for fault identification. The system provides a robust, automated solution for industrial process supervision.

VI. Acknowledgements

The authors would like to acknowledge the support of this work by Universiti Teknikal Malaysia Melaka (UTeM) on short term research grant INDUSTRI/ TEXASINSTRUMENTS/ FTKEK/ 2025/ IO0101.

VII.References

[1] J. Yan, Y. Meng, L. Lu and L. Li, "Industrial Big Data in an Industry 4.0 Environment: Challenges, Schemes, and Applications for Predictive Maintenance," in *IEEE Access*, vol. 5, pp. 23484-23491, 2017.

[2] F. M. Abdelillah, H. Nora, O. Samir and S. M. Benslimane, "Predictive Maintenance Approaches in Industry 4.0: A Systematic Literature Review," *2023 IEEE International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)*, Paris, France, pp. 1-6, 2023.

[3] M. Ponika, K. Jahnavi, P. S. V. S. Sridhar and K. Veena, "Developing a YOLO based Object Detection Application using OpenCV," *2023 7th International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India, pp. 662-668, 2023.

[4] N. P, M. S. K, A. M. Gurusigaamani, B. S. Vasanth, V. S. S. Saranya and P. G. K. Reddy, "Traffic Detection and Enhancing Traffic Safety: YOLO V8 Framework and OCR for Violation Detection Using Deep Learning Techniques," *2024 International Conference on Science Technology Engineering and Management (ICSTEM)*, Coimbatore, India, pp. 1-6, 2024.

[5] S. Mori, C. Y. Suen and K. Yamamoto, "Historical review of OCR research and development," in *Proceedings of the IEEE*, vol. 80, no. 7, pp. 1029-1058, July 1992.

[6] C. -N. E. Anagnostopoulos, I. E. Anagnostopoulos, I. D. Psoroulas, V. Loumos and E. Kayafas, "License Plate Recognition From Still Images and Video Sequences: A Survey," in *IEEE Transactions on Intelligent Transportation Systems,* vol. 9, no. 3, pp. 377-391, Sept. 2008.

[7] L. Neumann and J. Matas, "Real-time scene text localization and recognition," *2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence*, RI, USA, pp. 3538-3545, 2012.

[8] S. Du, M. Ibrahim, M. Shehata and W. Badawy, "Automatic License Plate Recognition (ALPR): A State-of-the-Art Review," in *IEEE*

Transactions on Circuits and Systems for Video Technology, vol. 23, no. 2, pp. 311-325, Feb. 2013.

[9] H. Kim, W.-K. Jung, I.-G. Choi, and S.-H. Ahn, "A low-cost vision-based monitoring of computer numerical control (CNC) machine tools for small and medium-sized enterprises (SMEs)," *Sensors*, vol. 19, no. 20, p. 4506, Oct. 2019.

[10] R. Kaur, J. Acharya, and S. Gaur, "Non-invasive data extraction from machine display units using video analytics," *World Academy of Science, Engineering and Technology, International Journal of Computer and Information Engineering*, vol. 13, no. 7, pp. 419-424, 2019.

[11] Rakhi P. Ghugardare, Sandip P. Narote, P. Mukherja, Parthamesh M. Kulkarni, "Optical Character Recognition System for Seven Segment Display Images of Measuring Instruments, " IEEE TENCON 2009, Singapore, 2009.

[12] Noppakun Boonsim, Saranya Kanjaruek, "An Integrated Technique for Detecting Seven-Segment Digits on Medical Devices, " in *27th International Computer Science and Engineering Conference (ICSEC)*, pp. 1-4, 2023.

[13] S.N.H. Abd Rahim and A.H. Embong, "Overall Equipment Utilization (OEU) Monitoring and Remote Quality Check in Legacy machine with Raspberry Pi," in *Journal of Integrated and Advanced Engineering,* Vol. 1(2) p.135-144, 2021.

[14] C.H. Wang, K.K Huang, R.I. Chang and C.K. Huang, "Scale-Mark-Based Gauge Reading for Gauge Sensors in Real Environments with Light and Perspective Distortions, " in *Sensors*, 22, pp. 7490, 2022.

[15] Q. Tang, Y.S. Lee and H. Jung, "The Industrial Application of Artificial Intelligence-Based Optical Character Recognition in Modern Manufacturing Innovations," in *Sustainability*, 16(5), pp. 2161, 2024.

[16] A. Gholamy, V. Kreinovich and O. Kosheleva, "Why 70/30 or 80/20 Relation Between Training and Testing Sets: A pedagogical Explanation, " in *Department Technical Reports (CS)*, pp. 1209, 2018.

[17] A.M Moraes, L. F. Pugliese, R.Fd. Santos, G.B. Vitor and R.Ad.S. Braga, "Effectiveness of YOLO Architectures in Tree Detection: Impact of Hyperparameter Tuning and SGD, " Adam, and AdamW Optimizers. *Standards*. 2025; 5(1):9.