Jurnal Teknologi Informasi dan Pendidikan
Volume 18, No. 2, September 2025
https://doi.org/10.24036/jtip.v18i2.1065

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

Optimization Analysis of Mig Welding Technique Learning Through Artificial Intelligence-Based Weld Defect Analysis: A Comparative Study of Monica AI, Claude, And Perplexity

Muhammad Rama Khairu Syifa[1], Marsono[1*], Imam Muhtarom[1], Bahrul Rizky Kurniawan[2]

[1]Department of Engineering Education, Universitas Negeri Malang, Malang, Indonesia
[2]Institute of Science Education, National Taiwan Normal University, Taipei, Taiwan
*Corresponding Author: marsono.ft@um.ac.id

Article Information

Article history:

No. 1065
Rec. October 03, 2025
Rev. November 04, 2025
Acc. November 04, 2025
Pub. November 07, 2025
Page. 1110 – 1126

Keywords:

- *Artificial Intelligence*
- *Monica AI*
- *Claude*
- *Perplexity*

ABSTRACT

This study aims to examine efforts to optimize MIG welding technique learning through the use of artificial intelligence (AI). This study compares three platforms: Monica, Claude, and Perplexity. The results were then evaluated using expert judgment. The background to this research stems from the difficulty students face in recognizing various types of defects in welding results, which results in decreased joint quality and a poor understanding of the welding process as a whole. This study employed a comparative experimental method involving expert validators to validate the analysis results. Three images of MIG welding defects were analyzed separately by the three AI platforms, and the identification result were validated by welding experts. In this study, three welded workpiece samples were used, each containing three analysis points: undercut, spatters, and overlap. The statistical approach utilized was the Crosstabs method, which generated comparative data for each AI platform. The resulting data were then examined using Chi-Square and Kappa tests to assess agreement levels and determine which AI system demonstrated superior performance in defect analysis. Based on the findings, Perplexity proved to have the highest accuracy in identifying defects, followed by Monica, while Claude performed the lowest. These findings indicate that AI integration can accelerate the identification process, simplify analysis, and make a positive contribution to welding technique learning. However, the study also revealed limitations, particularly related to the use of complex technical language and local contexts that AI has not yet fully accommodated. Therefore, it can be concluded that the application of AI, specifically Perplexity, has the potential to be an effective innovation to support the welding learning process in educational settings.

How to Cite:

Syifa, M.R.K., & et al. (2025). Optimization Analysis of Mig Welding Technique Learning Through Artificial Intelligence-Based Weld Defect Analysis: A Comparative Study of Monica AI, Claude, And Perplexity. Jurnal Teknologi Informasi Dan Pendidikan, 18(2), 1110-1126. https://doi.org/10.24036/jtip.v18i2.1065

1. INTRODUCTION

Innovation in engineering learning is a necessity in the digital era, especially in vocational education that emphasizes the mastery of practical skills. In the field of welding engineering, the Metal Inert Gas (MIG) welding process is one of the core competencies that must be mastered by students and vocational welding engineering students as well as job training institutions as cited in research[1]. One of the main challenges in learning MIG welding techniques is the difficulty of students in recognizing and analyzing welding defects accurately. The inability to recognize welding defects such as undercuts, spatter, and overlap in the Metal Inert Gas (MIG) welding process can significantly reduce the quality of the joint as cited in research[2], reduce the efficiency of the production process, and increase the potential risk to work safety. Spatter in MIG welding is generally caused by improper current and voltage settings, unstable shielding gas flow, and inappropriate welding techniques as stated in research[3]. This molten metal splash has an impact on reducing the quality of the weld surface, increasing time and costs because it requires an additional cleaning process. Overlap has the potential to disrupt load distribution and reduce the aesthetics of the weld. Meanwhile, undercuts can create thinning at the joint edges, which can become vulnerable to cracking.

Avoiding these three defects is crucial for ensuring the strength, durability, and reliability of welds, while minimizing material waste and repair time[4]. Therefore, an innovative learning approach is needed to equip students with visual analysis skills, as cited in research[5]. The use of artificial intelligence (AI) has significant potential to contribute to improving visual analysis-based learning processes, enabling students to become more skilled at identifying and preventing MIG weld defects effectively and measurably. This is in line with research by Maret et al.[6], which showed that variations in MIG welding speed on SC-80 pipe significantly affected the microstructure and hardness of the weld, thus technology-based understanding can support students' practical skills.

AI technologies such as Monica, Claude, and Perplexity are beginning to be explored as educational tools that enable students to obtain fast and accurate feedback in identifying weld defects. These three AI platforms were specifically selected because they represent different reasoning frameworks and interaction models, offering diverse analytical perspectives on visual data. In addition, they provide open accessibility, multimodal support for image-based analysis, and proven adaptability in processing technical terminology related to welding processes, making them suitable for comparative evaluation in this study. Based on this background, this study has three main focuses that serve as the problem formulation. First, this study highlights the performance of Monica, Claude, and Perplexity in analyzing visual defects in MIG welds. Second, the study evaluated the accuracy of each AI in identifying weld defect types based on validation from welding expert judgment. Third, the impact of AI use on improving student understanding in MIG welding techniques was also examined.

This study also aimed to measure the accuracy of the analysis results provided by the three AIs by comparing them with expert judgment assessments. This research is theoretically expected to enrich the development of digital technology-oriented science, with an emphasis on the integration of artificial intelligence in the welding learning process. This aligns with research by Sutijo et al[7], which emphasizes the importance of weld defect analysis and the use of technological approaches to support welding quality improvement in the maritime industry. For teachers, this research provides practical benefits as a reference in creating creative teaching strategies that align with the needs of the Industry 4.0 era.

For students, this research provides benefits in improving visual and analytical skills regarding weld defects in a more interactive and efficient manner, as cited in research[8]. For AI developers in education, the results of this study can support the design of AI systems that are more adaptive to the needs of welding engineering learning in Indonesia. This is in line with the study of Yulianti et al.[9] who emphasized that AI is able to personalize the learning experience, provide an adaptive assessment system, and create an inclusive virtual classroom to improve the quality of education in Indonesia. This study is limited to the use of three types of AI platforms, namely Monica, Claude, and Perplexity. The focus of the study is only on the types of visual defects in MIG welding results obtained from student practice. The study does not include non-visual defects, other types of welding outside MIG.

Research in the book Advances of Welding by Kunar et al.[10] states that Metal Inert Gas (MIG) welding is an electric arc welding technique that uses a shielding gas to prevent oxidation of molten metal during the joining process. MIG is known for its efficiency in producing neat and strong joints, and is easy to apply to various types of metal. The MIG welding procedure involves automatically feeding an electrode wire into the electric arc formed between the electrode and the base metal, with a flow of shielding gas such as argon or a CO_2 mixture. Despite its practicality, MIG welding still carries the risk of various visual defects such as undercuts, overlaps, and spatter. These defects can reduce the integrity of the weld joint and compromise work safety. Therefore, a thorough understanding of the types of defects and MIG work procedures is a crucial element in learning welding techniques in vocational education environments.

Artificial intelligence (AI) is a branch of computer science that focuses on developing systems capable of mimicking human cognitive abilities, in line with research findings[11], such as recognizing patterns, analyzing data, and making decisions. In the educational context, AI is developing into an adaptive and interactive learning tool. Based on its

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

classification, AI is divided into two main forms: generative AI and analytical AI. Generative AI, such as Claude and Perplexity, functions in generating narratives, answering questions, and constructing explanations based on big data. Meanwhile, analytical AI, such as Monica AI, operates based on visual processing and image pattern recognition. In engineering learning, both types of AI have great potential to support visual and verbal conceptual understanding and provide contextual feedback on learners' practice.

Monica AI is a computer vision-based platform designed to analyze technical drawings and identify specific visual features, such as welding defect patterns. This is in line with research by Sun et al.[12], which demonstrated that a vision-based seam tracking and multi-modal flaw detection system for GMAW welding using artificial intelligence was able to recognize welding defects in real time with a high degree of accuracy. The mechanism focuses on training the AI model using industrial datasets and image processing techniques. Claude, developed by Anthropic, excels at understanding context and constructing technical narratives, and is capable of providing detailed explanations based on user commands. Perplexity AI, on the other hand, relies on real-time information retrieval and explicit references from trusted sources to answer technical questions. In the context of MIG welding, Monica AI is used to visually recognize defect types, Claude to explain defect types and repair procedures, and Perplexity supports the learning process by providing relevant and up-to-date additional references.

Constructivism theory underpins the approach to vocational engineering learning, where students actively construct knowledge through hands-on experience and reflection on practical activities. Digitalization in vocational education increasingly enables the implementation of this theory through the use of technologies such as AI. The use of AI in engineering learning provides opportunities for creating interactive simulations, automated feedback, and personalized materials based on student needs. This aligns with the findings of Suryadi et al[13], who emphasized that artificial intelligence, augmented reality, big data, and digital technology are priority skills to support the transformation of education and future jobs. Pedagogically designed digital media can increase student engagement and accelerate the process of mastering technical skills. In the context of welding, AI can enrich the learning environment by providing real-time visualizations, automatic defect recognition, and narrative explanations that help students grasp concepts comprehensively.

Previous research has shown that the integration of technologies such as AI and augmented reality in engineering learning can improve students' conceptual understanding and practical skills. This is in line with the research of Barokah, Wahyono et al.[14] who developed a PBL and PjBL based learning model with the help of Black Box ARVR media, and was proven to significantly increase student interest and learning outcomes. The use of an image recognition system in welding training increased the accuracy of welding defect detection by up to 93.3%, this opinion is in accordance with the research of Maulana et al[15]. However, there has not been much research that directly compares the effectiveness of several different AI platforms in the context of learning MIG welding techniques. Therefore, this research has added value in filling the literature gap related to the use of AI for supports comprehensive welding defect analysis.

This research framework is based on the assumption that the use of AI can help improve students' ability to identify visual defects in MIG welding, in line with the findings of Faizal et al [16], who emphasized the importance of technological support in welding learning in vocational schools, which in turn will improve their overall technical understanding. Monica, Claude, and Perplexity were used as input in the form of image analysis tools and technical knowledge. The process involved visual defect image analysis and practice-based learning through interaction with AI. Ultimately, the intended outcome was the development of a more effective and contextual AI-based learning strategy.

2. RESEARCH METHOD

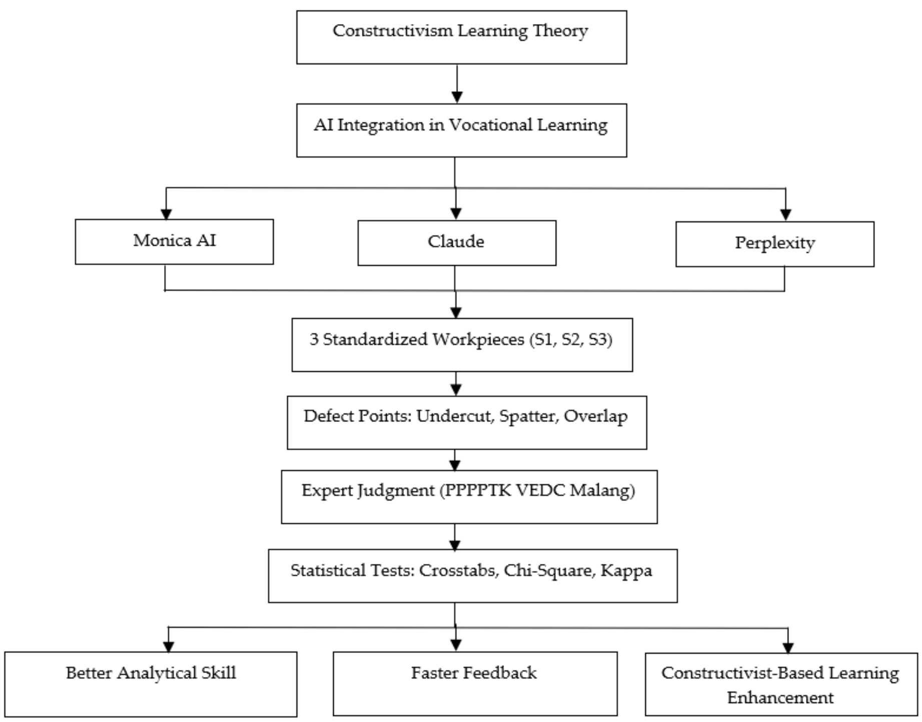


Figure 1. Diagram linking the empirical results to the theoretical framework

This study used a comparative experimental design to compare the accuracy and effectiveness of three artificial intelligence platforms: Monica AI, Claude, and Perplexity, in analyzing weld defect images during the Metal Inert Gas (MIG) welding process. This study focused on evaluating the ability of each AI to recognize and interpret common visual defects found in welding practices in engineering education environments, as evidenced by research[17]. All weld defect images were captured using a single device under consistent angle and contrast conditions, resulting in standardized photographs for workpieces 1, 2,

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

and 3. Each analysis point undercut, spatters, and overlap was pre-assessed by an expert possessing a valid Welding Inspector certification until 2026 from PPPPTK VEDC Malang, East Java, Indonesia. The collected data consisted of numerical values and concise qualitative descriptions, where the numerical outputs from each AI analysis were entered into SPSS Crosstabs with Chi-Square and Kappa tests applied to examine the statistical agreement and proximity of AI-generated results to expert judgment. This study was conducted in a vocational school and relevant engineering education workshops, involving expert judgment. This aligns with research by Ramadhani et al[18], which demonstrated that the use of AI at SMK Negeri 2 Tanjung Balai was able to enhance student creativity through technology-based practical activities.

Validation was conducted using expert methods, as evidenced by research[19], to ensure the accuracy of defect identification produced by each AI compared to professional welding practice standards. With this approach, the study is expected to provide an objective overview of AI's contribution to supporting more effective, technology-based welding technique learning. This is in line with the findings of Yahya et al.[20] who emphasized that the implementation of AI in vocational education is able to personalize learning, increase efficiency, and create learning experiences that are more interactive and relevant to industry needs.

2.1. Research Subjects

Table 1. Research subjects

Research Subjects		Standard	AI 1	AI 2	AI 3
S1	S1.1	S1.1 A	S1.1 B	S1.1 C	S1.1 D
	S1.2	S1.2 A	S1.2 B	S1.2 C	S1.2 D
	S1.3	S1.3 A	S1.3 B	S1.3 C	S1.3 D
S2	S2.1	S2.1 A	S2.1 B	S2.1 C	S2.1 D
	S2.2	S2.2 A	S2.2 B	S2.2 C	S2.2 D
	S2.3	S2.3 A	S2.3 B	S2.3 C	S2.3 D
S3	S3.1	S3.1 A	S3.1 B	S3.1 C	S3.1 D
	S3.2	S3.2 A	S3.2 B	S3.2 C	S3.2 D
	S3.3	S3.3 A	S3.3 B	S3.3 C	S3.3 D

Information :
S1 : Welding Workpiece (MIG) 1
S2 : Welding Workpiece (MIG) 2
S3 : Welding Workpiece (MIG) 3
S1.1 : Welding Defects (Undercut) in Welding Workpieces (MIG) 1
S1.2 : Welding Defects (Spatters) on Welding Workpieces (MIG) 1
S1.3 : Welding Defects (Overlap) on Welding Workpieces (MIG) 1

S2.1 : Welding Defects (Undercut) on Welding Workpieces (MIG) 2
S2.2 : Welding Defects (Spatters) on Welding Workpieces (MIG) 2
S2.3 : Welding Defects (Overlap) on Welding Workpieces (MIG) 2
S3.1 : Welding Defects (Undercut) on Welding Workpieces (MIG) 3
S3.2 : Welding Defects (Spatters) on welding workpieces (MIG) 3
S3.3 : Welding Defects (Overlap) in MIG Welding Workpiece 3
A : Standard welding defect analysis using (expert judgment)
B : Welding defect analysis using Monica AI
C : Welding defect analysis using Claude
D : Welding defect analysis using Perplexity

2.2. Data Collection

Table 2. Weld Defect Dataset and Validation Criteria

No	Number of Data Samples	Type of Weld Defect	Workpiece Code	Validation Criteria (Expert Judgment)
1	3 × 1	Undercut	S1.1, S2.1, S3.1	Evaluated based on ISO 5817:2023 Grade D and BNSP rubric 1227-P2-12/13 by a certified Welding Inspector (PPPPTK VEDC Malang, valid until 2026). Each image analyzed under identical lighting, angle, and contrast; assessment includes feasibility levels: Feasible, Fairly Feasible, Less Feasible, Not Feasible.
2	3 × 1	Spatter	S1.2, S2.2, S3.2	
3	3 × 1	Overlap	S1.3, S2.3, S3.3	

Data collection in this study was carried out using a random sampling method on photos of MIG welding results that had been standardized previously. Each photo was randomly selected from three workpieces (S1, S2, and S3) that represented certain types of welding defects such as undercut, spatter, and overlap. This approach is in line with the research of Irawan et al.[21] who identified various types of welding defects including undercut, spatter, porosity, and overlap in ST41 carbon steel using a non-destructive testing method. The workpieces are shown in the following photos:

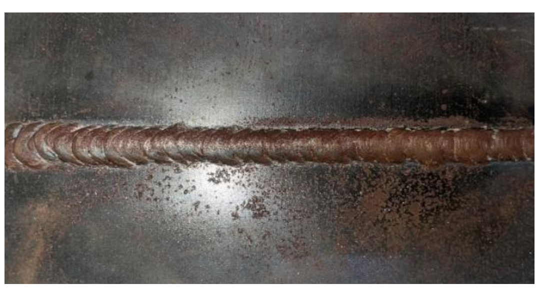


Figure 2. Welding Workpiece (MIG)1 (S1)

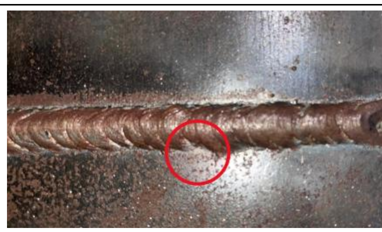
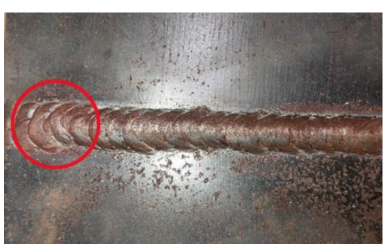
P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

Volume 18, No. 2, September 2025
https://doi.org/10.24036/jtip.v18i2.1065

Figure 3. Welding Workpiece (MIG)2 (S2)

Figure 4. Welding Workpiece (MIG)3 (S3)

Table 3. Standard Photos of Research Subjects (S1,S2,S3)

Subject	Standard Photo	Subject	Standard Photo
S1.1	 **Figure 5.** Welding Defects (Undercut) in Welding Workpieces (MIG) 1	S1.2	 **Figure 6.** Welding Defects (Spatters) on Welding Workpieces (MIG) 1
S1.3	 **Figure 7.** Welding Defects (Overlap) on Welding Workpieces (MIG) 1	S2.1	 **Figure 8.** Welding Defects (Undercut) on Welding Workpieces (MIG) 2

S2.2

Figure 9. Welding Defects (Spatters) on Welding Workpieces (MIG) 2

S2.3

Figure 10. Welding Defects (Overlap) on Welding Workpieces (MIG) 2

S3.1

Figure 11. Welding Defects (Undercut) on Welding Workpieces (MIG) 3

S3.2

Figure 12. Welding Defects (Spatters) on welding workpieces (MIG) 3

S3.3

Figure 13. Welding Defects (Overlap) in MIG Welding Workpiece 3

The photos were then analyzed using three artificial intelligence platforms: B (Monica AI), C (Claude), and D (Perplexity). The analysis results from these three AI platforms were then compared with the reference results from A (expert judgment), which serves as the standard for welding defect analysis. Thus, this data collection process ensured that the evaluation was conducted objectively, representatively, and aligned with the experimental procedures outlined in the research design table.

2.3. Measurement Scale

The measurement scale in this study uses a standardized welding defect assessment rubric reference and legalized by BNSP with copyright Kemendikbud 1227-P2-12/13. This rubric provides quantitative and qualitative criteria for welding results, in line with the research of Khamdan et al. [22] who emphasized the importance of industry standard assessment instruments such as AWS in evaluating the welding quality of vocational high school students in the National LKS, especially in identifying spatter, undercut, and overlap defects. With this measurement scale, the analysis results of the three AIs tested can be evaluated objectively and compared based on valid national standards.

Table 4. Measurement Scale

Assessment components	Indicator	Range Score
Undercut	No undercut	9,0 – 10
	Undercut 0.2mm2 x 10%	8,0 – 8,9
	Undercut 0.5mm 2x 10%	7,0 – 7,9
	Undercut <0.5mm2 x 10%	0
Spatter	No Spatter	9,0 – 10
	Spatter has 2 points	8,0 – 8,9
	Spatter has 4 points	7,0 – 7,9
	Spatter contains > 5 points	0
Overlap	100% even and smooth	9,0 – 10
	90% even and smooth	8,0 – 8,9
	85% even and smooth	7,0 – 7,9
	<80% even and smooth	0

To measure the quality of Artificial Intelligence (AI) in terms of accuracy, the calculation results were analyzed and compared with the score interpretation criteria obtained using the provisions listed in the measurement scale table above.

This measurement scale includes an expert judgment indicator referring to ISO 5817:2023 Grade D to support the scores for each research subject, with the stipulation that the Grade D standard has the highest tolerance for weld defect analysis. This aligns with research by Hadi et al[23], which confirms that international standards such as ASME and ISO 5817 are used to determine weld defect tolerance limits in non-destructive testing.

Table 5. Expert Judgment Indicator

Research Subject	Skor	Information
S1.1A	7,5	Undercut 0.5mm 2x 10%
S1.2A	8,7	Spatter has 2 points
S1.3A	8,0	90% even and smooth
S2.1A	7,3	Undercut 0.5mm 2x 10%
S2.2A	7,5	Spatter has 4 points
S2.3A	7,0	85% even and smooth
S3.1A	8,0	Undercut 0.2mm2 x 10%
S3.2A	8,5	Spatter has 2 points
S3.3A	6,0	<80% even and smooth

The results obtained from the expert judgment data in the table above indicate that workpiece 1 (S1), categorized as Fairly Feasible according to the score interpretation criteria, while S2 and S3 have an average score of Less Feasible.

The indicator data obtained from the expert judgment are categorized into four categories:

Score Interpretation Criteria:

4 = Feasible: $\sum$ score ≥ 9

Jurnal Teknologi Informasi dan Pendidikan
Volume 18, No. 2, September 2025
https://doi.org/10.24036/jtip.v18i2.1065

3 = Fairly Feasible: 8.0 < ∑ score ≤ 8.9
2 = Less Feasible: 7.0 < ∑ score ≤ 7.9
1 = Not Feasible: ∑ score ≤ Mi 7.0 Sbi.

Table 6. Example of display results from each AI

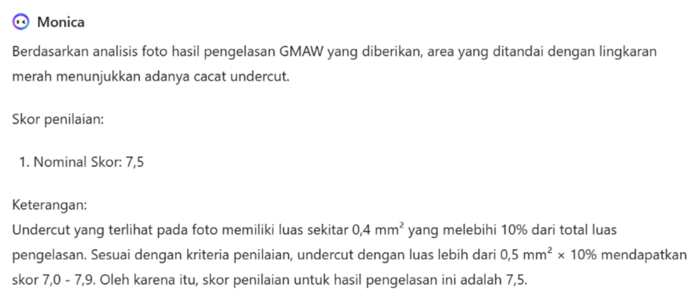
<table>
<tr><th>Sample</th><th>Example of coding display results/from each AI</th></tr>
<tr><td>AI 1</td><td>


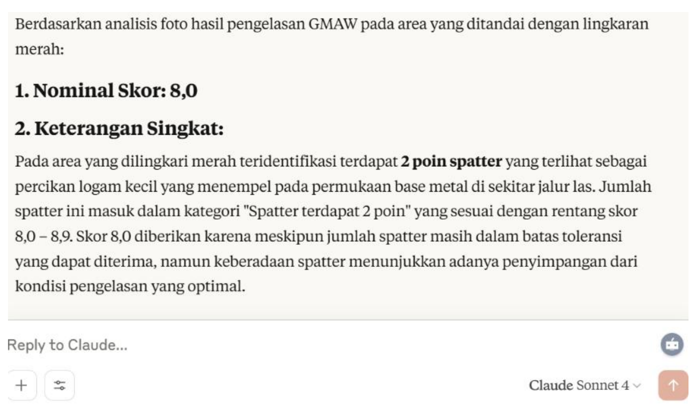
Figure 14. Results of the subject S1.1B display for analysis on Monica AI
</td></tr>
<tr><td>AI 2</td><td>


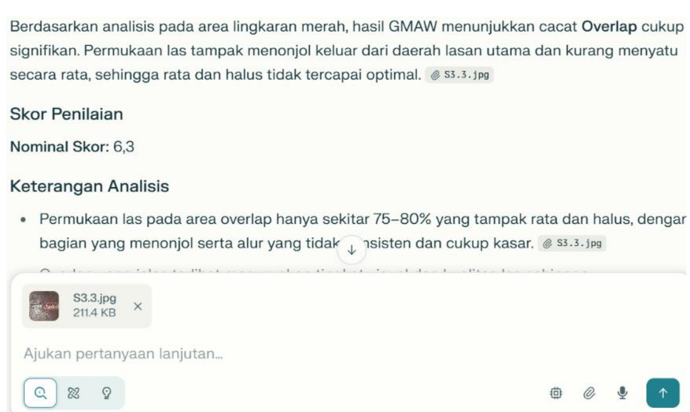
Figure 15. Results of the S2.2C subject display for analysis on Claude
</td></tr>
<tr><td>AI 3</td><td>


Figure 16. Results of the S3.3 D subject display for analysis on Perplexity
</td></tr>
</table>

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

Example of the results of the analysis display on each AI based on table 6 using subject S1.1 B for example on Monica AI which shows an output score of 7.5 and a brief description then continued with subject S2.2 C on Claude and shows the same output, namely a score of 8.0 and a brief description followed by subject S3.3 D on Perplexity with an output score of 6.3 and a brief description. The data from the three AI results can be seen in table 6 below.

Table 7. Three AI Results Data (Monica AI, Claude, and Perlexity)

Research Subject	**Analysis Results**					
	AI 1 (B)		AI 2 (C)		AI 3 (D)	
		Information		Information		Information
S1.1	7,5	ut 0.5mm 2x 10%	7,5	ut 0.5mm 2x 10%	7,5	ut 0.5mm 2x 10%
S1.2	7,5	has 4 points	7,5	has 4 points	8,5	has 2 points
S1.3	8,5	en and smooth	7,5	en and smooth	7,5	en and smooth
S2.1	7,0	ut 0.5mm2 x 10%	7,5	ut 0.5mm 2x 10%	8,5	ut 0.2mm2 x 10%
S2.2	8,5	has 2 points	8,0	has 2 points	8,5	has 2 points
S2.3	7,5	en and smooth	7,5	en and smooth	7,0	en and smooth
S3.1	8,5	ut 0.2mm2 x 10%	7,3	ut 0.5mm2 x 10%	7,5	ut 0.5mm 2x 10%
S3.2	8,5	has 2 points	9,5	tter	9,5	tter
S3.3	7,0	en and smooth	8,5	en and smooth	6,3	ven and smooth

Based on Table 7, the scores and interpretation criteria generated from each AI show the scores and brief descriptions according to the rubric in Table 4.

2.4. Data Analysis

Data analysis in this study used a crosstable to compare the data of three AI platforms with expert judgment opinions. This was followed by a Chi-Square test which functions to see whether there is a significant relationship between expert judgment and AI. This is in line with the research of Susanti et al.[24] who used the Chi-Square test to analyze the relationship between climate factors and Dengue Fever cases in Surabaya. This was followed by a Kappa analysis which functions to assess the level of agreement (inter-rater agreement) between the results of the AI analysis and expert judgment. The importance of this agreement test is in line with the research of Sari et al.[25] This emphasizes that instrument validity needs to be strengthened through appropriate statistical analysis to ensure the reliability of learning outcomes. It will be divided into four categories:

- If p-value > 0.05: no significant relationship
- If p-value < 0.05: significant relationship
- If r-value > 0.05: no significant agreement
- If p-value < 0.05: significant agreement

Therefore, this analysis aims to determine the best AI platform that best aligns with expert judgment standards. The interpretation of the criteria from the SPSS statistical analysis will determine which AI (AI1, AI2, or AI3) is valid and suitable for use in MIG welding defect analysis.

Jurnal Teknologi Informasi dan Pendidikan
Volume 18, No. 2, September 2025
https://doi.org/10.24036/jtip.v18i2.1065

3. RESULTS AND DISCUSSION

3.1. Implementation Results

Table 8. Average Score Result

Research Subject		Averange Score Result			
		Expert Judgement	AI 1	AI 2	AI 3
S1	S1.1	7,5	7,5	7,5	7,5
	S1.2	8,7	7,5	7,5	8,5
	S1.3	8,0	8,5	7,5	7,5
S2	S2.1	7,3	7,0	7,5	8,5
	S2.2	7,5	8,5	8,0	8,5
	S2.3	7,0	7,5	7,5	7,0
S3	S3.1	8,0	8,5	7,3	7,5
	S3.2	8,5	8,5	9,5	9,5
	S3.3	6,0	7,0	8,5	6,3

From the data in table 8, it has been grouped based on the categories of three AI platforms and expert judgment which shows the score results for each AI which refers to the expert judgment opinion.

Table 9. Results of crosstabs test between Expert and Monica AI

		AI 1			Total
		1.00	2.00	3.00	
Expert	Not Suitable	1	1	0	2
	Less Suitable	1	1	1	3
	Quite Suitable	0	1	3	4
Total		2	3	4	9

Table 10. Results of crosstabs test between Expert and Claude

		AI 1			Total
		2.00	3.00	4.00	
Expert	Not Suitable	1	1	0	2
	Less Suitable	2	1	0	3
	Quite Suitable	3	0	1	4
Total		6	2	1	9

Table 11. Crosstab test results between Expert and Perplexity

		AI 1				Total
		1.00	2.00	3.00	4.00	
Expert	Not Suitable	2	0	0	0	2
	Less Suitable	0	1	2	0	3
	Quite Suitable	0	2	1	1	4
Total		2	3	3	1	9

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

Table 12. Relationship between Expert and the three Ais

AI	p	r
AI 1	0.432	0.201
AI 2	0.537	0.422
AI 3	0.088	0.353

Information:
p = Chi-Square
r = Kappa

3.2. Discussion

Based on the results of the crosstab analysis between the expert and AI 1 in table 9, it shows that in the Fairly Feasible category, AI1 is quite consistent because the majority of cases (3 out of 4) were given a score of 3.00 according to the expert. However, in the Not Feasible category, AI1 is not completely consistent because there are cases that are rated 2.00 (higher than the expert), and in the Less Feasible category the distribution of AI results is spread across all categories (1.00, 2.00, and 3.00) which shows inconsistency. This means that AI1 has a moderate level of agreement with the expert, better than AI2, but is not yet stable in all categories. Continued with the results of the Chi-Square analysis on AI1 based on table 11 shows a significance of 0.432 (> 0.05), so there is no significant relationship between AI1 and the expert's assessment. Meanwhile, the Kappa test produces a value of 0.308 with a significance of 0.201 (> 0.05), which means there is low and insignificant agreement between AI1 and the expert. Thus, AI1 cannot be considered valid and still falls far short of expert judgment standards. With further development, AI1 has no potential to be used to help optimize students' ability to independently identify welding defects because the results obtained still fall short of expert judgment standards according to SPSS analysis. This condition aligns with research by Siahaan et al[26], who found that the Rasch-Cohen Kappa value on the Self-Perceived Employability instrument approached zero, indicating a low level of agreement between experts, thus requiring significant revision before the instrument can be relied upon.

The crosstab analysis between the expert and AI2 in Table 10 shows that the majority of AI2 scores are concentrated around 2.00 (6 out of 9 cases), resulting in an unbalanced distribution compared to the more varied expert assessments. In the "Sufficiently Adequate" category according to the expert, AI2 actually gave many low scores (2.00), resulting in a discrepancy. This finding indicates that AI2 tends to underestimate its assessment, providing lower results than the expert, resulting in a low level of agreement between AI2 and the expert. Meanwhile, the Chi-Square test results for AI2 showed a significance of 0.537 (>0.05), indicating no significant relationship between AI2 and the expert. The Kappa test showed a significance of 0.422 (>0.05), indicating not only insignificance but also no meaningful agreement between AI2 and the expert's assessment. This is consistent with research that explains AI2 as the platform with the lowest level of validity and is less suitable for use in MIG weld defect analysis. This condition is similar to the findings of Wazira et al[27] who reported that instruments with very low Kappa values (average 0.026) showed minimal reliability and inter-rater agreement, with some items even yielding negative scores, thus declaring the instrument unsuitable for use without revision and improvement.

The results of the crosstab analysis between the expert and AI 3 show that in the Unsuitable category, AI 3 is highly consistent, with all cases scored 1.00, exactly matching the expert's. In the Less Feasible category, most cases were scored higher (3.00) resulting in a mismatch, and in the Fairly Feasible category, AI results were more varied (2.00, 3.00, 4.00) with a tendency to give higher scores than expert judgment. This finding means that AI3 has the best fit in the extreme category (Not Feasible) compared to AI1 and AI2, but still shows a tendency to overestimate in the middle category. Meanwhile, the Chi-Square test results on AI3 showed a significance of 0.088 (> 0.05), so statistically there is no significant relationship, even though the value is close to the 0.05 limit. The Kappa test produced a significance of 0.353 (> 0.05), which means there is a weak but not significant fit with expert judgment. Thus, AI3 has relatively better results compared to AI1 and AI2, although it is not yet fully valid to replace expert judgment in MIG welding defect analysis. The consistency of AI3 in the extreme category can be utilized in MIG learning to train students to recognize cases of welding defects that are clearly inappropriate, thereby strengthening their basic skills in evaluating the quality of welding results. Utilizing this consistency is in line with research by Fadholi et al[28] which shows that direct experience-based learning, such as guided inquiry assisted by PhET media, can significantly improve students' analytical and problem-solving skills.

4. CONCLUSION

Based on the discussion, it can be concluded that of the three AI platforms analyzed, AI3 demonstrated relatively better performance than AI1 and AI2. AI3 had the highest consistency, especially in the extreme unsuitable category, supported by better Chi-Square and Kappa values than the other two AI platforms, although they remained at an insignificant level. Meanwhile, AI1 had a moderate but unstable level of suitability, and AI2 showed the lowest results, with a tendency to underestimate expert judgment.

Therefore, although these three platforms cannot yet replace expert judgment, AI3 can be considered a starting point for developing an AI-based evaluation system to support the optimization of MIG welding technique learning, particularly in training students to recognize clearly unsuitable weld defects. These findings also confirm that the integration of AI into the learning process serves not only as an analytical tool but also as a tool for students to conduct independent checks during welding lessons. It is recommended that further development of the AI model, including improvements to the training database and

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

analysis system, is needed to produce more valid, consistent, and expert-standard assessments, making them truly useful in welding technique learning.

REFERENCES

[1] C. Sutowo and I. Budiawan, "Analysis of the Effect of Tig and Mig Welding on Welded Joints with SS316 and SS304 Materials"Universitas Muhammadiyah Jakarta, pp. 46-57 Jan 2020.

[2] S. B. Handana and R. E. Pramitasari, "Tensile Strength Analysis of Smaw and Mig Welding of St 41 Steel Using V and X Weld Variations" Jurnal Ilmiah Surya Teknika Vol 8, No 1, pp 6-10, April 2024.

[3] G. I. Huka, L. S. Loppies, J. Matheus, A. D. Wilujeng, and A. Hamid, "The Effect of Gas Metal Arc Welding (GMA) Current Variations on the Mechanical Properties of Seamless Pipes.," vol. 1, no. 3, 2023.

[4] P. Diah Larasati et al., "Improving the Quality and Quantity of Welding Production Using MIG Welding Machines at the Rahman Jaya Welding Workshop in Gedawang Village, Banyumanik District, Semarang City" JPkMN, vol. 5, no. 2, pp. 2057–2063, May 2024, doi: 10.55338/jpkmn.v5i2.3140.

[5] T. Surawan et al., "Training and Application of MIG Welding Machine Technology in Repairing Metal Plates to Become Skilled Workers in the Automotive Industry at SMK Nasional Depok, West Java" Dedikasi, vol. 3, no. 1, pp. 67–76, Dec. 2022, doi: 10.31479/dedikasi.v3i1.204.

[6] R. N. Maret, S. Syaripuddin, and F. B. Susetyo, "The Effect of MIG Welding Speed on SC-80 Pipes on Microstructure and Hardness with 1g Welding Position," JKTM, vol. 4, no. 2, pp. 76–80, Sept. 2023, doi: 10.52447/jktm.v4i2.1774.

[7] B. Sutijo, Z. Hilmy, F. T. Adi, and T. Susilo, "Analysis of Welding Defects (Cracks) in Butt Weld Joints The Influence of Crane Pedestal Welding on the WTIV Project (Wind Turbine Installation Vessel)," Jurnal Sains dan Teknologi Maritim (JSTM) Vol. 25 No. 2 Maret 2025 e-ISSN: 2623-2030; p-ISSN: 1412-6826,pp. 160-177 DOI: 10.33556/jstm.

[8] A. Agustina and Y. Suharya, "The Application of Artificial Intelligence (AI) Technology in the Field of Education Towards Indonesia's Golden Generation 2045," Prosiding Temu Ilmiah Nasional Guru XVI vol. 16,pp. 129-138, 2024.

[9] G. Yulianti, N. Permana, and F. A. K. Wijayanti, "Transforming Indonesian Education: Implementing the Potential of Artificial Intelligence (AI),"journal of information systems and management(JISMA), Vol. 2 No. 6, pp. 102-106, Dec 2023.

[10] S. Kunar and G. Mandal, Advanced Welding Technologies. John Wiley & Sons, ISBN : 1394331908, 9781394331901,pp. 608, April 2025.

[11] S. Q. Moinuddin, S. H. Saheb, A. K. Dewangan, M. M. Cheepu, and S. Balamurugan, Automation in the Welding Industry: Incorporating Artificial Intelligence, Machine Learning and Other Technologies. John Wiley & Sons, Feb 2024.

[12] M. Mobaraki, "Vision-Based Seam Tracking and Multi-Modal Defect Detection in GMAW Fillet Welding Using Artificial Intelligence"The University of British Columbia pp. 1-122, 2025.

[13] S. Suryadi and F. A. P. Nasution, "The Industrial Revolution, Future Job Trends, and Indonesia's Position," J-naker, vol. 18, no. 2, pp. 124–141, Aug. 2023, doi: 10.47198/jnaker.v18i2.237.

[14] A. I. Barokah, H. Wahyono, and S. Listyarini, “Development of a Learning Model for Special Characteristics of Animals Using PBL and PjBL Assisted by Black Box ARVR Media,” basicedu, vol. 6, no. 1, pp. 978–986, Jan. 2022, doi: 10.31004/basicedu.v6i1.1949.
[15] A. Maulana, D. E. Yuliana, and D. A. W. Kusumastutie, “Weld Defect Classifier Using GLCM Extraction and ANN,” JTECS, vol. 2, no. 2, p. 133, July 2022, doi: 10.32503/jtecs.v2i2.2701.
[16] E. Faizal, S. H. Susilo, A. Walid, and E. Yudiyanto, “Welding Training for Students as a Medium for Entrepreneurship Learning at Muhammadiyah 1 Vocational School, Malang” Bulletin of Community Engagement, Vol. 4, No. 2, pp.668-674, August 2024.
[17] T. Schaupp, M. Rhode, H. Yahyaoui, and T. Kannengiesser, “Hydrogen-assisted cracking in GMA welding of high-strength structural steels using the modified spray arc process,” Weld World, vol. 64, no. 12, pp. 1997–2009, Dec. 2020, doi: 10.1007/s40194-020-00978-0.
[18] A. Ramadhani and R. T. A. Agus, “Utilization of Artificial Intelligence (AI) to Improve the Creativity of Students at State Vocational School 2 Tanjung Balai” Gudang Jurnal Pengabdian Masyarakat Vol.3 No. 1, Januari 2025, pp 343-347 Doi : https://doi.org/10.59435/gjpm.v3i1.1419.
[19] N. Sari, A. Thamrin, and A. Nurhidayati, “Contribution of Occupational K3 Knowledge and Attitudes of Vocational High School Students towards Occupational K3 Behavior Awareness,” IJCEE, vol. 8, no. 1, p. 67, Dec. 2022, doi: 10.20961/ijcee.v8i1.68044.
[20] Muhammad Yahya, Wahyudi, and Akmal Hidayat, “Implementation of Artificial Intelligence (AI) in Vocational Education in the Era of the Industrial Revolution 4.0,” semnasdies, vol. 1, pp. 190–199, July 2023, doi: 10.59562/semnasdies.v1i1.794.
[21] A. Irawan, A. Finali, L. Umar, and R. E. P. Utomo, “Analysis of Defects Resulting from Smaw Welding on Steel Carbon St 41 with Cooling Variations,” . Vol., vol. 6, no. 1, 2024.
[22] N. Khamdan, R. D. Djatmiko, and S. A. Ramadhani, “Pressure Vessel Welding Quality Analysis in the National Vocational High School Competency Competition with AWS Standards,” J. Din. Vok. Tek. Mesin, vol. 5, no. 1, pp. 34–41, Mar. 2020, doi: 10.21831/dinamika.v5i1.30993.
[23] M. N. Hadi, H. Abizar, and A. Maulana, “Analysis of Non-Destructive Testing Results on Boiler Superheater Connections,” Vocational Education National Seminar (VENS) Vol.3 No.1, pp 12-16, Dec 2024.
[24] N. P. Ghaisani, S. Sulistiawati, and M. L. I. Lusida, “Correlation Between Climate Factors with Dengue Hemorrhagic Fever Cases in Surabaya 2007 – 2017,” IJTID, vol. 9, no. 1, p. 39, Apr. 2021, doi: 10.20473/ijtid.v9i1.16075.
[25] Titin Astina, D. Wardiah, Muhammad Ali, J. M. Santos, K. M. Villanueva, and L. G. Bautista, “Development of Project-Based Student Worksheets (LKPD) on Students' Ability to Design and Create a Product in Expository Learning,” pembahsi, vol. 15, no. 1, pp. 60–73, Jan. 2025, doi: 10.31851/pembahsi.v15i1.16300.
[26] F. Siahaan and A. Ridwan, “Development of Self-Perceived Employability Instrument for Vocational Students' Work Readiness,” vol. 5, no. 1, pp. 179-192, 2024 doi : 10.47841/semnasadpi.v5i1.156.
[27] A. Wazira and K. Azmi, “Testing Instruments of Teacher Discipline on Student Academic Achievement,” Jurnal Riset Edukasi Vokasi, Inovasi, dan Sosial (J-REVISI) vol. 1, no. 2,pp. 44-53, Dec 2024.
[28] A. H. Fadholi, K. Kosim, M. Zuhdi, and S. Ayub, “The Effect of the Guided Inquiry Learning Model Assisted by PhET Media on Problem-Solving Ability,” Kappa, vol. 9, no. 1, pp. 121–128, Apr. 2025, doi: 10.29408/kpj.v9i1.29651.

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id