ISSN 2442-6571

Automatic coloring of ulos motif using large language model and non-dominated sorting differential evolution algorithm

Arlinta Christy Barus [a,1,*], Humasak Simanjuntak [a,2], Samuel Indra Gunawan Situmeang [a,3], Robert Aritonang [a,4], Nicholas Hutabarat [a,5], Grase Panjaitan [a,6]

[a] Institut Teknologi Del, Toba, Indonesia
[1] arlinta@del.ac.id; [2] humasak@del.ac.id; [3] samsitumeang@gmail.com; [4] robertaritonang2002@gmail.com; [5] nicholashutabarat0@gmail.com; [5] gracethessa1@gmail.com
* corresponding author

ARTICLE INFO

Article history
Received December 11, 2025
Revised January 28, 2026
Accepted January 31, 2026
Available online May 31, 2026

Keywords
Ulos
Automatic coloring
Large language model
Multi-objective optimization
NSDE

ABSTRACT

Ulos is a traditional Batak textile with high cultural significance, yet artisans lack tools to modernize designs while preserving heritage. Previous research primarily uses single-objective methods that fail to accommodate user preferences. This study introduces the first multi-objective framework for traditional textile coloring, differentiating itself by using a Large Language Model (LLM) as a natural language interface for non-technical weavers. The LLM translates user design preferences into dynamic objective functions to guide a Non-dominated Sorting Differential Evolution (NSDE) algorithm. Performance is assessed using the epsilon indicator, aesthetic metrics (contrast and colorfulness), and user preference scores. The system achieved high convergence with an epsilon-indicator value, and user preference scores reaching 0.66. Additionally, a paired t-test (p=0.07) demonstrated the algorithm's robustness across parameter variations. This approach enables artisans to create culturally authentic, aesthetically optimized designs through intuitive interaction.

1. Introduction

Ulos, a traditional woven textile from the Batak ethnic group in North Sumatra, Indonesia, holds significant philosophical, spiritual, and economic value [1]. However, its sustainability is threatened by declining interest among younger generations due to the time-intensive production process and limited economic returns [2], [3]. This challenge reflects the broader struggle of safeguarding Intangible Cultural Heritage (ICH) through digital preservation [4]–[6]. To meet modern market demands and improve artisan livelihoods, contemporary Ulos production requires innovation that maintains cultural authenticity while streamlining design processes [7].

Technological innovation, particularly artificial intelligence (AI), has already begun transforming the fashion and textile industries [8]–[12]. Digital design tools offer pathways to sustain traditional crafts by reducing production time [13]. The DiTenun platform pioneered this for Ulos through automated motif generation and coloring [14], [15]. However, DiTenun's current color optimization, which relies on Genetic Algorithms, faces two critical limitations [16]. First, its single-objective focus often results in imbalanced color distributions that sacrifice cultural authenticity for computational efficiency [17]–[19]. Second, the lack of an intuitive interface prevents artisans from expressing personal preferences, thus limiting personalization. While computational tools have supported other craft traditions [20]–[26], existing systems often prioritize single-objective optimization [27]–[29] or require technical expertise that is incompatible with traditional artisan workflows. Consequently, there is a clear research gap in multi-criteria optimization specifically tailored for culturally constrained textiles.

To address these limitations, we propose the use of the Non-dominated Sorting Differential Evolution (NSDE) algorithm [30], [31]. NSDE is a well-established multi-objective optimization technique previously applied to complex problems in civil engineering [32], materials science [33], electrical engineering [34], and logistics [35]. In this study, we adapt NSDE to simultaneously optimize multiple aesthetic criteria, including color contrast, colorfulness, and pattern diversity, while strictly adhering to cultural constraints. However, NSDE's requirement for explicit objective functions creates a usability barrier for artisans.

We bridge this gap by integrating Large Language Models (LLMs) as natural language interfaces. Building on recent advances in LLM-guided optimization [36]–[40], our system translates an artisan's culturally grounded preferences into formal algorithmic parameters. Aesthetic quality is quantified using established metrics such as RMS Contrast, Michelson Contrast, and Colorfulness [41]. Cultural authenticity is maintained through thread color dictionaries derived from traditional Ulos classifications like Harungguan and Sadum. Finally, the system's effectiveness is validated through subjective interviews with 20 expert weavers.

The primary objective of this research is to develop a user-centric design framework that empowers artisans through technology. Specifically, this work aims to:

- develop the first multi-objective optimization framework for traditional textile coloring that balances aesthetic and cultural criteria.
- create a novel LLM-mediated interface that allows non-technical artisans to control design parameters through natural language.
- provide empirical validation demonstrating that the generated designs maintain high levels of usability and cultural authenticity.

2. Method

The proposed system follows a three-stage workflow: data preparation through color quantization, LLM-based preference translation, and multi-objective optimization using the NSDE algorithm.

2.1. Data Preparation and Color Quantization

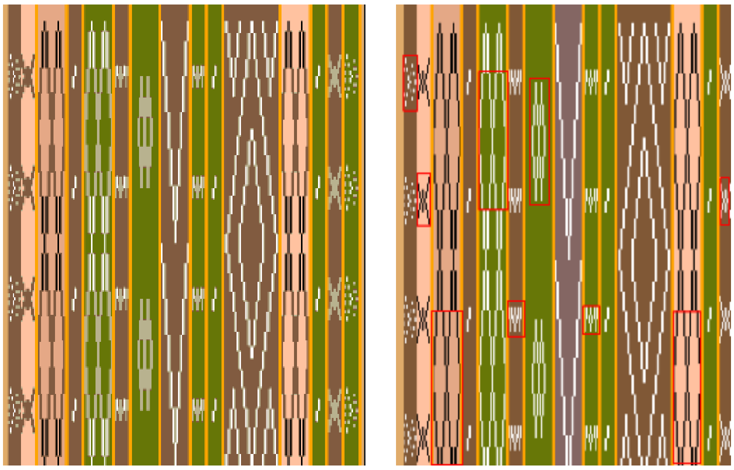
The research utilizes the "Ulos Batak DiTenun" dataset (https://ieee-dataport.org/documents/dataset-ulos-batak), comprising 94 high-resolution images of traditional Ulos motifs, specifically the Harungguan, Sadum, and Puca types. Original Ulos images exhibit high heterogeneity in pixel intensity due to lighting and texture, often containing thousands of unique shade variations (e.g., hundreds of red tones). As traditional weavers work with a limited set of dyed threads, this continuous color space is impractical for direct optimization.

To align the digital images with the physical constraints of traditional weaving, the continuous color space of the 94 motif images is discretized using K-Means-based Color Quantization [42], as follows:

- Feature Extraction: Pixel intensities are converted to the HSV color space to facilitate culturally relevant color manipulation.
- Clustering: K-Means clustering is applied to each motif image to identify the dominant color groups.
- Optimal K Selection: The number of clusters (K) is determined by the highest silhouette coefficient, resulting in $K = 5$ for Puca, $K = 9$ for Sadum, and $K = 10$ for Harungguan.

- Quantization: Each pixel is mapped to its nearest cluster centroid, reducing the search space from thousands of shades to a discrete thread-based palette.

This process reduces the search space D from a per-pixel definition to a discrete palette definition, where the genome for the optimization algorithm represents the centroids of these K clusters. Fig. 1 shows a comparison of Ulos images before and after the color quantization process.



Before After

Fig. 1. Before and after color quantization

In (a) before color quantization, it can be seen that the Ulos image has very heterogeneous colors, with various color variants such as green, brown, and cream which have different levels of intensity. This heterogeneity enriches the visual details, but is difficult to produce manually by traditional weavers due to limited materials to produce many color variants. In image (b) after color quantization, the results of the K-Means based color quantization process are shown, where each type has a K value or a predetermined number of color clusters.

2.2. Large Language Model Integration

An LLM serves as a natural language interface to bridge the gap between qualitative artisan preferences and quantitative optimization. The LLM's primary role is to generate a Python-based objective function, $f_{pref}(x)$, which evaluates how closely a candidate color palette matches user intent. The model configuration and standardized prompt are defined as follows:

- Experiments utilized GPT-4 (https://platform.openai.com/docs/models/gpt-4), DeepSeek V3 (https://huggingface.co/deepseek-ai/DeepSeek-V3), and Gemini 2.0 Flash with a temperature setting of 1.00 (https://docs.cloud.google.com/vertex-ai/generative-ai/docs/models/gemini/2-0-flash) to encourage creative color suggestions. To ensure reproducibility, all models were accessed via API using the configurations detailed in Table 1, and the standardized prompt was applied to each (Fig. 2).

Table 1. LLM Parameter Configuration

Model	Temperatur	Top-P	Max Input Token Limit
GPT-4	1.00	1.00	8.000
DeepSeek V3	1.00	0.70	128.000
Gemini 2.0 Flash	1.00	0.95	1.048.576

- The models are provided with a structured prompt containing the Ulos type, pattern characteristics, and the list of user-selected HSV color codes.

- The generated function must accept an image matrix as input and return a fitness value in the range [0,1], where 1 represents a perfect match with the user's desired palette.

```
1.  Create a Python function named calculate_user_color_preferences.

2.  This function is the objective function of the NSDE (Non-dominated Sorting
    Differential Evolution) algorithm that optimizes color combinations in Ulos motifs
    with the following characteristics:
3.  - Line: {line}
4.  - Pattern: {pattern}
5.  - Accent color: {accent_color}
6.  - Contrast color: {contrast}

7.  The list of user-preferred Hue, Saturation, and Value combinations is
    {list_colors_hsv}. This function accepts as input an image matrix in the HSV color
    space (with Hue 0-179, Saturation 0-255, Value 0-255, and dtype np.uint8).

8.  From this image, extract all unique combinations of Hue, Saturation, and Value.

9.  If the list of Hue, Saturation, and Value combinations matches the user-preferred
    Hue, Saturation, and Value combination, then the fitness value is 1 (one). The
    fitness value will decrease according to the extent to which the color matches the
    given preference, based on the specified criteria.

10. Ensure the Hue value is in the range 0-360, explicitly convert to int32, and ensure
    hue_img is an int before modulo.
```

Fig. 2. Standardized prompt

The generated function should accept three main parameters: "ulos type name" to specify the type of Ulos to be optimized, an "api key" to access the API, and "ulos selected color codes", which contains a list of user-selected color codes. First, this function retrieves Ulos characteristics from the database, including information about stripes, patterns, accent colors, and color contrast (see Table 2).

Table 2. Ulos Characteristics

Element	Harungguan	Sadum	Puca
Lines	Dominated by diagonal lines that form the main structure of the fabric.	Horizontal lines are used to divide the pattern and structure.	Motifs are arranged horizontally, dividing the pattern and motif.
Patterns	Traditional patterns include curved line ornaments and plant-like motifs.	Symmetrical and geometric patterns, including crosses or small, evenly distributed motifs.	Geometric motifs are the main element, including traditional house/Gorga motifs and repeating lines.
Dominant Colors	There are many color combinations, with the base color tending to be darker than the ulos motif. The color combinations are complex.	Each ulos has one to two main colors, while the border color tends to be darker.	The dominant color forms the overall base of the fabric, generally red, cream, and blue.
Accent Colors	Accent colors (gold, white, and black) emphasize lines or patterns in the form of thin lines or small motifs around the main pattern.	Accent colors used are gold, white, or black. Accents form lines or detailed elements on the fabric.	Accent colors such as white, black, gold, or gray are used to emphasize motifs or lines. Accents highlight the pattern.
Color Contrast	Light colors as the lines of the motif contrast with the base color of the ulos, which tends to be dark.	Each ulos has a contrast between light colors (yellow, white) and dark colors (black, dark red) to emphasize the pattern.	High contrast is achieved through a combination of light colors (yellow, white) and dark colors (black, dark red).

Next, the user-selected color codes are converted into HSV format using a predefined thread colors database (see Fig. 3).

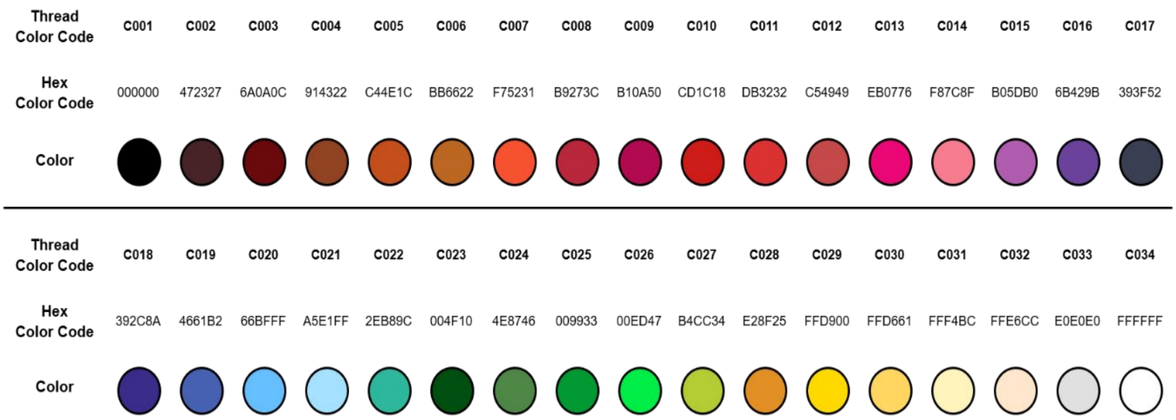


Fig. 3. Representative Thread Color Samples for Authenticated Ulos Weaving, integrated into the LLM Prompting Logic

2.3. Large Language Models in Evolutionary Optimization

Recent research has seen LLMs transition from generative text tools to active components in optimization pipelines, a paradigm often termed "LLM-as-optimizer." Studies by Yao et al. [40] and Liu et al. [39] demonstrate the ability of models like GPT-4 and Gemini to evolve heuristics and generate dynamic objective functions. Our work extends these benchmarks by applying them to the domain of Intangible Cultural Heritage (ICH). Unlike general optimization tasks, this application requires a specialized translation layer that maps subjective cultural descriptions to strict, predefined color dictionaries, bridging the gap between artisan intent and algorithmic execution.

2.4. Mathematical Definitions of Objective Functions

The NSDE algorithm optimizes a vector of objective functions $F(x) = [f_{RMS}(x), f_{Mich}(x), f_{color}(x), f_{unique}(x), f_{pref}(x)]$, where x represents a candidate solution (color palette.

2.4.1. Contrast Metrics

RMS Contrast (f_{RMS}) calculates the standard deviation of normalized pixel intensity at pixel (i, j) (I_{ij}) relative to the mean intensity ($\bar{I}$) [43], [44], as in (1).

$$f_{RMS} = \sqrt{\frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (I_{ij} - \bar{I})^2} \tag{1}$$

Michelson Contrast (f_{Mich}) evaluates the luminance range between maximum (I_{max}) and minimum (I_{min}) intensities [43], [44], as in (2).

$$f_{Mich} = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \tag{2}$$

2.4.2. Aesthetic Metrics

Colorfulness (f_{color}) quantifies visual vividness using the mean (μ) and standard deviation (σ) of the opponent color spaces rg and yb [41], as in (3):

$$f_{color} = \sqrt{\sigma_{rg}^2 + \sigma_{yb}^2} + 0.3\sqrt{\mu_{rg}^2 + \mu_{yb}^2} \tag{3}$$

Optimal Unique Color (f_{Unique}) maximizes diversity by counting perceptually distinct colors up to the limit K.

2.4.3. Preference Metrics

(f_{pref}): This dynamic function, generated by the LLM, maximizes similarity between the candidate palette (C_{sol}) and the user's preferred colors (C_{user}), as in (4).

$$f_{pref}(X) = \frac{1}{1+\sum_{u\in C_{user}} \min_{v\in C_{sol}} \|u-v\|_2} \tag{4}$$

2.5. Non-Dominated Sorting Differential Evolution Algorithm Configuration

The optimization core utilizes the NSDE algorithm with the following specifications:

- Genome Representation. Each individual (x) is a vector of size $3 \times K$, representing H, S, V channels for each of the K cluster.
- Algorithm Parameters. The system uses a population size (NP) of 100, a crossover rate (CR) of 0.7, and a mutation factor (F) of 0.5 to ensure robust convergence across 200 generations (G_{max}).

2.6. System Architecture and User Interface

The proposed framework follows a modular, user-centric architecture designed to bridge qualitative artisan intent with quantitative optimization, as illustrated in Fig. 4.

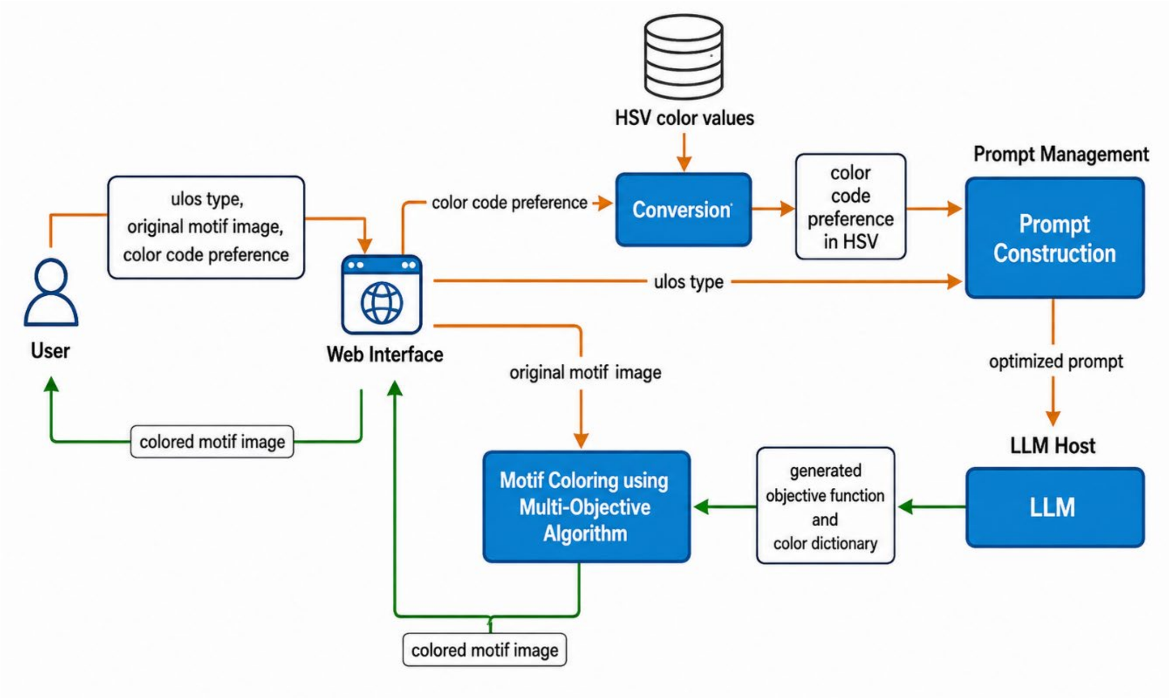


Fig. 4. System Architecture of the LLM-NSDE Framework for User-Centric Ulos Design

The workflow is structured into four primary phases:

- Multi-Modal Input Acquisition. The process begins with the End User providing three distinct data streams via the User Interface: the cultural category (Ulos type), the visual source (original motif image), and aesthetic preferences (preferred color codes).
- Prompt Engineering and LLM Interaction. The user's color preferences are transformed into the HSV (Hue, Saturation, Value) color space and passed to the Prompt Management module. This module applies a logic-tier template that combines cultural constraints with user intent to construct an optimized prompt. This prompt is transmitted via API to the Large Language Model (LLM), which functions as a code generator.

- Code and Dictionary Generation. The LLM returns two critical artifacts: a customized objective function, $f_{pref}(x)$, and a color dictionary. The $f_{pref}(x)$ is a Python-executable snippet that mathematically represents the user's preference, allowing it to be integrated directly into the optimization pipeline
- Multi-Objective Optimization (NSDE). The Non-dominated Sorting Differential Evolution (NSDE) algorithm serves as the core engine. It ingests the original motif and utilizes the LLM-generated $f_{pref}(x)$ alongside standard aesthetic metrics (contrast and colorfulness) to guide the evolutionary search. By evaluating candidate color palettes against this dynamic fitness function, the NSDE identifies the Pareto-optimal solutions that best balance cultural authenticity with user preference

3. Results and Discussion

The scope of this research is defined by the use of the Ulos dataset from DiTenun. Consequently, its findings may not be generalizable to Ulos types outside this specific collection. The study is limited to the digital coloration phase of existing Ulos motifs and does not encompass the physical creation process itself, such as thread selection, weaving techniques, or the generation of new motifs. The developed coloring system is evaluated based on its aesthetic quality and alignment with traditional Ulos aspects, an evaluation which involves Ulos weavers in Toba Regency to ensure its relevance and objectivity against established standards. While their expertise provides invaluable feedback, it is acknowledged that this cohort may not represent the full spectrum of preferences across all Ulos communities.

3.1. Experimental Environment and Parameter Configuration

To ensure reproducibility, all experiments were conducted in a standardized computational environment. The system was implemented using Python 3.10.12 with core libraries including NumPy 1.23.5 for numerical operations and OpenCV 4.8.0 for image processing and HSV color space transformations. LLM interactions were facilitated through their respective official APIs (GPT-4, DeepSeek V3, and Gemini 2.0 Flash) with the specific configurations previously detailed in Table 1.

The optimization process utilized the NSDE algorithm parameters defined in Section 2.4. For the objective evaluations, we compared the Default Parameter Set against the Reference Set from [45] to assess algorithmic robustness. Table 3 summarizes these configurations to facilitate a direct comparison of the search intensity across generations (G_{max}) and population sizes (NP).

Table 3. NSDE Parameter Configuration

Parameter	Symbol	Default Value	Reference Value
Max Generations	G_{max}	200	300
Population Size	NP	100	100
Crossover Rate	CR	0.7	0.4
Mutation Factor	F	0.5	0.85

3.2. Evaluation Metrics and Baseline

The performance of the LLM-NSDE framework is evaluated against a baseline NSDE (without LLM preference integration). Quantitative assessment is performed using core metrics defined in Section 2.4.

- Convergence: Measured via the ϵ-indicator.
- Aesthetic Quality: Evaluated through Colorfulness (f_{color}), RMS Contrast (f_{RMS}), and Michelson Contrast (f_{Mich}).
- Diversity and Intent: Assessed via Optimal Unique Color (f_{unique}) and the User Preference Score (f_{pref}) generated by the LLM interface.

3.3. LLM Configuration and Functional Reliability

To ensure the robustness of the system, we measured the Functional Code Rate (FCR), defined as the percentage of LLM responses that produced executable Python code without syntax errors. Gemini 2.0 Flash, DeepSeek V3, and GPT-4 achieved an FCR of 100% in 10 trials. The Prompt Management module utilizes a three-tier logic.

- Context Injection: Retrieval of Ulos-specific color dictionaries.
- Preference Mapping: Processing of user-defined natural language inputs.
- Syntactic Constraints: Structuring the prompt to mandate a specific Python $f(x)$ interface

3.4. Quantitative Analysis of Convergence and Aesthetic Quality

The convergence of the LLM-NSDE algorithm was evaluated using the ϵ indicator across different Ulos motif categories. As shown in Table 4, the performance metrics for the Harungguan Ulos indicate that the integration of DeepSeek V3 yielded the most favorable convergence, achieving the lowest ε-indicator value of 2.18×10^{-3}. In terms of aesthetic quality, DeepSeek V3 also achieved the highest Michelson Contrast (0.86) among the LLM-integrated models, surpassing the baseline NSDE. While the Gemini 2.0 Flash model recorded the highest User Preference score (0.28), it experienced a significant drop in RMS contrast compared to the baseline. Conversely, DeepSeek V3 provided a more balanced optimization, maintaining high colorfulness (168.65) and contrast levels comparable to the baseline while successfully incorporating user preferences better than GPT-4o.

Table 4. Experimental results of *Harungguan* Ulos

LLM	ε-indicator	Colorfulness	RMS Contrast	Michelson Contrast	Optimal Unique Color	User Preferences
NSDE	2.27×10^{-3}	168.46	30.43	0.85	1	-
NSDE + Gemini 2.0 Flash	4.10×10^{-3}	168.71	23.86	0.76	1	0.28
NSDE + DeepSeek V3	$\mathbf{2.18 \times 10^{-3}}$	168.65	29.62	0.86	1	0.08
NSDE + GPT-4o	3.12×10^{-3}	165.25	24.72	0.66	1	0.01

Table 5 details the experimental results for the Sadum Ulos, highlighting a significant disparity in the models' ability to adhere to user constraints. DeepSeek V3 demonstrated superior performance in capturing user intent, achieving a User Preference score of 0.60, which substantially outperforms Gemini 2.0 Flash (0.20×10^{-2}) and GPT-4o (0). Additionally, both DeepSeek V3 and GPT-4o maximized the Michelson Contrast to 0.99, ensuring high visibility of the motif patterns. Although the baseline NSDE produced the highest RMS contrast (32.92), the DeepSeek V3 integration offered the best trade-off, securing high preference alignment and maximum colorfulness (169.93) without critically compromising the structural clarity of the motif.

Table 5. Experimental Results of *Sadum* Ulos

LLM	ε-indicator	Colorfulness	RMS Contrast	Michelson Contrast	Optimal Unique Color	User Preferences
NSDE	5.61×10^{-2}	169.44	32.92	0.84	1	-
NSDE + Gemini 2.0 Flash	1.33×10^{-2}	169.99	22.54	0.90	1	0.20×10^{-2}
NSDE + DeepSeek V3	1.36×10^{-1}	169.93	23.37	0.99	1	0.60
NSDE + GPT-4o	5.61×10^{-2}	169.77	22.69	0.99	1	0

Table 6 shows the experimental results for the Puca Ulos. In this scenario, Gemini 2.0 Flash exhibited the strongest capability in satisfying user constraints, achieving the highest preference score of 0.66 and the best ε-indicator value of 1.75×10^{-3}. However, regarding visual properties, DeepSeek V3 proved to be the most robust in preserving the motif's definition, matching the baseline's RMS contrast (32.21) and achieving the highest Michelson contrast (0.64). While GPT-4o generated the most colorful results (171.16), it failed to register any alignment with user preferences. This suggests that for the Puca motif, Gemini offers the best adherence to specific color requests, whereas DeepSeek ensures the best preservation of contrast and pattern details.

Table 6. Experimental results of *Puca* Ulos

LLM	ε-indicator	Colorfulness	RMS Contrast	Michelson Contrast	Optimal Unique Color	User Preferences
NSDE	5.22×10^{-3}	170.19	32.22	0.59	1	-
NSDE + Gemini 2.0 Flash	1.75×10^{-3}	169.69	21.79	0.63	1	0.66
NSDE + DeepSeek V3	1.94×10^{-2}	170.90	32.21	0.64	1	0.48
NSDE + GPT-4o	1.91×10^{-3}	171.16	29.93	0.63	1	0

Based on the results, the use of a Large Language Model (LLM) combined with the Non-Dominated Sorting Differential Evolution (NSDE) algorithm can generate color combinations for Ulos motif images that better match user preferences compared to using the NSDE algorithm alone. In addition to producing preferred colors, the use of LLMs also introduces new colors that closely resemble users' desired hues. These new colors are selected from a set of traditional Ulos thread colors, enriching the variety of colors in the Ulos motif images while still aligning with user preferences.

Among the three LLMs used in this experiment, DeepSeek V3 outperformed Gemini 2.0 Flash and GPT-4o. This model achieved higher fitness scores compared to the others and also excelled at generating custom objective function code used to calculate fitness values based on user preferences. While other models often produced incorrect or non-functional code, DeepSeek consistently generated logical, well-structured, and executable code. In addition, a paired t-test was conducted on the ε-indicator values to compare the performance of the default NSDE parameter set against the reference parameter set. The analysis yielded a $p - value$ of 0.0700. Since this value exceeds the significance level ($\alpha = 0.05$), we fail to reject the null hypothesis.

3.5. Qualitative Results

This study serves as a qualitative validation, and future work will scale this to a larger demographic to increase statistical power. Based on the results of qualitative evaluation conducted through interviews with 20 weavers from Silaen Subdistrict, Toba Regency, statistical analysis of seven questions revealed variations in perception toward the three evaluated types of Ulos: Harungguan, Puca, and Sadum (Table 7). The questions posed to the weavers to gauge their perceptions based on the evaluation rubric include:

- Q1: Are the color combinations on the Ulos harmonious and appropriate?
- Q2: Are the color combinations consistent with the distinctive beauty of the Ulos?
- Q3: Are the colors evenly distributed and complementary to each other within the motifs?
- Q4: Are these colors still appropriate for weaving today?
- Q5: Do the selected colors reflect the meaning and values of Batak culture?
- Q6: Have dyeing results ever differed from expectations?
- Q7: If so, can they continue, or must the results be as planned?

The evaluation method used is a face-to-face interview, in which each weaver will try dyeing each type of Ulos and complete the evaluation questionnaire three times. To maintain consistency and accuracy in their responses, particularly for the seventh question, weavers are instructed to choose Strongly Agree or Agree if the different dyeing results can still be continued, and Disagree or Strongly Disagree if the dyeing results must be as planned. This direction is given so that weavers do not answer carelessly and still consider the meaning of each choice.

Table 7. Subjective evaluation results of the colored Ulos images

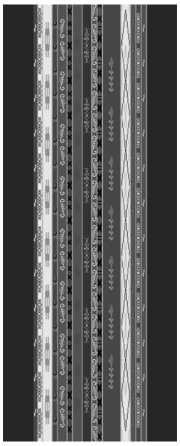
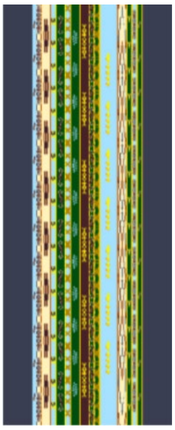
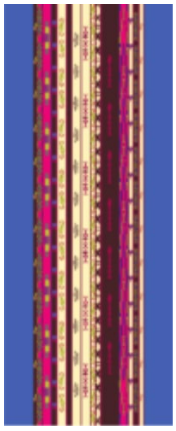
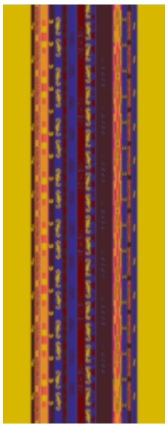
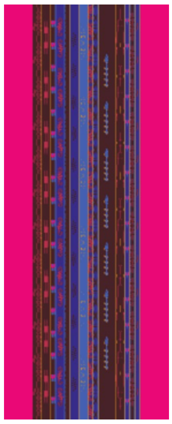
Measures	Ulos Type	Q1	Q2	Q3	Q4	Q5	Q6	Q7
Mean	*Harungguan*	3.50	3.70	3.20	3.16	3.25	3.35	2.50
	Sadum	3.95	3.95	3.85	3.60	3.85	3.40	3.45
	Puca	3.50	3.35	3.52	3.35	3.50	2.85	3.00
Standard Deviation	*Harungguan*	0.51	0.47	0.71	0.69	0.55	0.59	0.83
	Sadum	0.23	0.23	0.37	0.60	0.36	0.75	0.69
	Puca	0.51	0.81	0.61	0.74	0.36	0.59	0.80
Median	*Harungguan*	4.00	4.00	3.00	3.00	3.00	3.00	2.00
	Sadum	4.00	4.00	4.00	4.00	4.00	4.00	4.00
	Puca	4.00	3.50	4.00	3.50	3.50	3.00	3.00
Mode	*Harungguan*	4.00	4.00	3.00	3.00	3.00	3.00	2.00
	Sadum	4.00	4.00	4.00	4.00	4.00	4.00	4.00
	Puca	4.00	4.00	4.00	4.00	4.00	3.00	3.00
Agreement Percentage	*Harungguan*	100	100	85	85	95	95	82.35
	Sadum	100	100	100	95	100	85	88.20
	Puca	95	90	95	85	100	75	41.20

Harungguan Ulos was positively received for its color harmony and aesthetic appeal (Q1 and Q2: 100% agreement). However, lower agreement on Q7 (2.5 mean; 82.35%) and high variability (SD: 0.83) indicate concerns regarding consistency and alignment with user expectations, especially when colors differ from the original plan. Sadum Ulos achieved the highest overall satisfaction, with strong consensus (100% agreement on Q1–Q3, Q5) and low variability (Q1–Q2 SD: 0.23). It effectively balanced aesthetic quality and cultural values. Still, Q6 and Q7 showed moderate disagreement, suggesting the need to improve how deviations from user intent are handled. Puca Ulos was rated positively for visual harmony and cultural representation (high scores on Q1–Q5), but showed the lowest agreement on Q7 (41.2%) and higher standard deviations. This indicates that unexpected outcomes reduced user satisfaction despite good overall aesthetic results.

Table 8 presents a comparative visualization of the automatic coloring results for the Harungguan Ulos motif, illustrating the efficacy of integrating Large Language Models (LLMs) to enforce user constraints. The "Input Colors" column establishes a specific user-defined palette consisting of maroon, gold, dark blue, dark brown, and pink hues. The result labeled "NSDE" serves as a baseline, utilizing the optimization algorithm without LLM guidance; this output notably fails to adhere to the user's requested palette, instead generating a combination of unrelated greens and cyan.

In contrast, the results guided by LLMs, specifically DeepSeek V3, Gemini 2.0 Flash, and GPT-4o, successfully translate the user's preferences into the final output, producing motifs that strictly align with the provided color set. Among these, the study notes that DeepSeek V3 (displayed in the top right) outperformed other models in generating objective functions that match user preferences, visually resulting in a high-fidelity mapping of the requested pink and maroon tones compared to the discordant baseline.

Table 8. Coloring results of *Harungguan* Ulos

Original Motif	Input Colors	LLM	Output Motif	LLM	Output Motif
		NSDE		NSDE + DeepSeek V3	
		NSDE + Gemini 2.0 Flash		NSDE + GPT-4o	

3.6. Discussion

The experimental results validate the proposed framework's ability to modernize Ulos design while maintaining heritage. The discussion below connects the quantitative and qualitative findings to the primary research objectives. The first objective focused on developing a multi-objective framework that balances visual appeal with cultural constraints. The ϵ-indicator values using DeepSeek V3 (mean 1.94×10^{-2} for Puca, 1.36×10^{-1} for Sadum and 1.36×10^{-1} for Harungguan) confirm that the NSDE algorithm successfully converges toward a diverse Pareto front without sacrificing computational efficiency. While traditional single-objective approaches often prioritize a single aesthetic, leading to "over-optimized" but culturally unrecognizable designs, the high scores in **RMS Contrast** and **Colorfulness** demonstrate that our multi-objective approach produces vibrant results that remain within the boundary of traditional color dictionaries. The Pareto-based optimization ensures that improvements in visual contrast do not come at the expense of cultural authenticity.

The second objective sought to lower the technical barrier for artisans through an LLM-mediated interface. The experimental data reveals that the effectiveness of this interface varies significantly depending on the Ulos type and the specific model utilized. The highest performance was recorded in the Puca category using Gemini 2.0 Flash, which achieved a User Preference Score (f_{pref}) of 0.66, followed by the Sadum category using DeepSeek V3 with a score of 0.60. These results provide empirical evidence that LLMs can successfully translate qualitative artisan intent into formal objective functions, effectively transforming a complex mathematical process into an intuitive design session.

In contrast, the Harungguan category yielded a lower peak score of 0.28 with Gemini 2.0 Flash. This discrepancy suggests that motifs with higher cultural or structural complexity remain a significant challenge for current natural language translation layers. The variance in these scores highlights that while the LLM serves as a vital bridge for non-technical users, model selection and prompt engineering must be carefully tailored to the specific characteristics of the textile tradition to ensure high-fidelity design outcomes. By utilizing the LLM as a "human-in-the-loop" mechanism, the system narrows the search space, fulfilling the goal of making advanced AI tools accessible to traditional cultural practitioners, even as it identifies areas for further model refinement.

The final objective was to validate the system's usability and its ability to provide personalized designs. The paired t-test result (p=0.0700) is particularly revealing; while it indicates that the algorithm is robust and consistent, the variance in scores suggests that artisan preferences are distinct and subjective. This finding justifies the shift from a "one-size-fits-all" automated coloring template to a personalized optimization framework. The visual comparisons in Table 8 further confirm that while the algorithm introduces modern color variations, the core cultural identifiers, the "soul" of the Ulos, remain intact, demonstrating that digital innovation can successfully coexist with traditional safeguarding.

4. Conclusion

This research successfully established a multi-objective framework for the automatic coloring of Ulos motifs by integrating the Non-dominated Sorting Differential Evolution (NSDE) algorithm with Large Language Model (LLM) preference translation. By utilizing the LLM as a natural language interface, the system effectively bridged the gap between complex mathematical optimization and the qualitative, culturally grounded intent of traditional artisans. The experimental findings confirmed that this approach achieves high convergence on the Pareto front while significantly improving user satisfaction compared to standard automated methods. The primary contribution of this work lies in providing a scalable methodology for the digital preservation of Intangible Cultural Heritage (ICH), demonstrating that advanced AI can empower rather than replace traditional craftsmanship by lowering technical barriers to design innovation. While the framework showed high performance for Puca and Sadum motifs, the lower preference scores for Harungguan motifs highlight a need for more nuanced prompt engineering to handle complex cultural hierarchies. Future research will explore the use of multi-modal LLMs to analyze motif geometry alongside textual preferences to improve design fidelity. Additionally, we plan to implement a real-time feedback loop within the optimization process, enabling weavers to adjust objective weights dynamically, and scale the qualitative validation to a larger demographic to increase statistical power. Expanding the color dictionaries to include a broader range of Indonesian textiles will also be a priority to test the framework's cross-cultural scalability.

Acknowledgment

The authors gratefully acknowledge the research funding provided by the Ministry of Higher Education, Science, and Technology (KemdiktiSaintek) of the Republic of Indonesia through the BIMA Grant for Applied Research - Prototype Output scheme. The authors also thank Febriani Gultom for her support and administrative assistance throughout this research. Furthermore, the authors are deeply grateful to the traditional Ulos weavers in Silaen Village, Toba Regency, who generously shared their time and expertise by serving as evaluators. Finally, the authors acknowledge the support of Institut Teknologi Del for providing the necessary resources and environment to conduct this research.

Declarations

Author contribution. Author contributions are presented using the CRediT (Contributor Roles Taxonomy) system.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
First Author	✓	✓		✓	✓	✓	✓	✓		✓		✓	✓	✓
Second Author	✓	✓	✓		✓	✓				✓	✓	✓		
Third Author		✓		✓	✓	✓				✓		✓	✓	✓
Fourth Author		✓	✓		✓	✓		✓	✓		✓			
Fifth Author			✓		✓			✓	✓		✓			
Sixth Author			✓		✓			✓	✓		✓			

C	: Conceptualization	I	: Investigation	Vi	Visualization
M	: Methodology	R	: Resources	Su	Supervision
So	: Software	D	: Data Curation	P	Project Administration
Va	: Validation	O	: Writing Original Draft	Fu	Funding Acquisition
Fo	: Formal analysis	E	: Writing – Review & Editing		

Funding statement. The authors gratefully acknowledge the funding received from the Ministry of Higher Education, Science, and Technology (KemdiktiSaintek) of the Republic of Indonesia under contract number 46/SPK/LL1/AL.04.03/PL/2025, 027/ITDel/LPPM/Kemdiktisaintek/Pen/VI/2025.
Conflict of interest. The authors declare no conflict of interest.
Additional information. No additional information is available for this paper.

Data and Software Availability Statements

The data and software that support the findings of this study are available from the corresponding author, Samuel Situmeang, upon reasonable request.

References

[1] R. J. Siagian, "The Symbolic Meaning of Traditional Woven Fabric Ulos as A Spiritual Expression in Batak Toba Rituals," *Int. J. Relig.*, vol. 5, no. 5, pp. 200–209, Apr. 2024, doi: 10.61707/stw03g83.

[2] Zulkifli and M. Ridwan, "Revitalization of The Traditional Values Lost Due to The Commodification of Art/Crafts: A Case Study of Bataknese Traditional Ulos," *Asian Ethn.*, vol. 20, no. 4, pp. 541–554, Oct. 2019, doi: 10.1080/14631369.2019.1608812.

[3] R. S. B. Simbolon, M. Sihotang, and K. Tobing, "Threads of Tradition: Silalahi Ulos Weaving and Cultural Significance in North Sumatra," *J. Ilmu Pendidik. dan Hum.*, vol. 11, no. 2, pp. 138–152, 2022, doi: 10.35335/jiph.v11i2.20.

[4] M. Z. Idris, N. B. Mustaffa, and S. O. S. Yusoff, "Preservation of Intangible Cultural Heritage Using Advance Digital Technology: Issues and Challenges," *Harmon. J. Arts Res. Educ.*, vol. 16, no. 1, p. 1, Dec. 2016, doi: 10.15294/harmonia.v16i1.6353.

[5] W. M. Wan Isa, N. A. Mat Zin, F. Rosdi, and H. Mohd Sarim, "Digital Preservation of Intangible Cultural Heritage," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 12, no. 3, p. 1373, Dec. 2018, doi: 10.11591/ijeecs.v12.i3.pp1373-1379.

[6] S. Mazzetto, "Integrating Emerging Technologies with Digital Twins for Heritage Building Conservation: An Interdisciplinary Approach with Expert Insights and Bibliometric Analysis," *Heritage*, vol. 7, no. 11, pp. 6432–6479, Nov. 2024, doi: 10.3390/heritage7110300.

[7] C. Nugroho, I. K. Nurhayati, K. Nasionalita, and R. M. U. Malau, "Weaving and Cultural Identity of Batak Toba Women," *J. Asian Afr. Stud.*, vol. 56, no. 6, pp. 1165–1177, Sep. 2021, doi: 10.1177/0021909620958032.

[8] M. P. Sikka, A. Sarkar, and S. Garg, "Artificial Intelligence (AI) in Textile Industry Operational Modernization," *Res. J. Text. Appar.*, vol. 28, no. 1, pp. 67–83, Jan. 2024, doi: 10.1108/RJTA-04-2021-0046.

[9] H. Anwer, M. Ali, and H. Jamshaid, "Applications of Artificial Intelligence in Textiles and Fashion," in *Creative Textile Industry*, 2024, pp. 175–194, doi: 10.1007/978-981-97-3802-1_8.

[10] X. Wu and L. Li, "An Application of Generative AI for Knitted Textile Design in Fashion," *Des. J.*, vol. 27, no. 2, pp. 270–290, Mar. 2024, doi: 10.1080/14606925.2024.2303236.

[11] Y. Zhang, C. Liu, and S. Xia, "From Hype to Value: Harnessing Generative AI in Fashion Retailing From A Technology-Organization-Environment Perspective," *J. Electron. Bus. Digit. Econ.*, vol. 4, no. 2, pp. 203–219, Dec. 2025, doi: 10.1108/JEBDE-11-2024-0042.

[12] G. Rizzi and P. Bertola, "Exploring the generative AI potential in the fashion design process: an experimental experience on the collaboration between fashion design practitioners and generative AI tools," *Eur. J. Cult. Manag. Policy*, vol. 15, pp. 1–22, May 2025, doi: 10.3389/ejcmp.2025.13875.

[13] LIAOANSHUN and S. Banerjee, "Investigating Cultural Sustainability via The Fusion and Innovation of Traditional Crafts and Modern Design," *Cuest. Fisioter.*, vol. 53, no. 3, pp. 4739–4747, 2024, doi: 10.48047/v5bakm73.

[14] A. C. Barus, M. Simanjuntak, and V. Situmorang, "DiTenun, Smart Application Producing Ulos Motif," *SHS Web Conf.*, vol. 86, p. 01025, Nov. 2020, doi: 10.1051/shsconf/20208601025.

[15] H. Simanjuntak, E. Panjaitan, S. Siregar, U. Manalu, S. Situmeang, and A. Barus, "Generating New Ulos Motif with Generative AI Method in Digital Tenun Nusantara (DiTenun) Platform," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 7, pp. 1125–1134, 2024, doi: 10.14569/IJACSA.2024.01507109.

[16] A. Barus, S. Situmeang, M. Napitupulu, S. Saragih, and R. Theresia, "Automatic Ulos Coloring Using Genetic Algorithm Approach," *Eng. Appl. Artif. Intell.*, no. September, pp. 1–36, 2024, doi: 10.2139/ssrn.5059671.

[17] Y. Bo, J. Yu, and K. Zhang, "Computational Aesthetics and Applications," *Vis. Comput. Ind. Biomed. Art*, vol. 1, no. 1, p. 6, Dec. 2018, doi: 10.1186/s42492-018-0006-1.

[18] S. Wang, J. Liu, J. Jiang, Y. Jiang, and J. Lan, "Attribute analysis and modeling of color harmony based on multi-color feature extraction in real-life scenes," *Front. Psychol.*, vol. 13, pp. 1–15, Sep. 2022, doi: 10.3389/fpsyg.2022.945951.

[19] Y. Du, X. Liu, M. Cai, and K. Park, "A Product's Kansei Appearance Design Method Based on Conditional-Controlled AI Image Generation," *Sustainability*, vol. 16, no. 20, p. 8837, Oct. 2024, doi: 10.3390/su16208837.

[20] G. Tian, Q. Yuan, T. Hu, and Y. Shi, "Auto-Generation System Based on Fractal Geometry for Batik Pattern Design," *Appl. Sci.*, vol. 9, no. 11, p. 2383, Jun. 2019, doi: 10.3390/app9112383.

[21] V. M. Ngo, T.-V. T. Duong, T.-B.-T. Nguyen, P. T. Nguyen, and O. Conlan, "An Efficient Classification Algorithm for Traditional Textile Patterns from Different Cultures Based on Structures," *J. Comput. Cult. Herit.*, vol. 14, no. 4, pp. 1–22, Dec. 2021, doi: 10.1145/3465381.

[22] R. F. Alya, M. Wibowo, and P. Paradise, "Classification of Batik Motif Using Transfer Learning on Convolutional Neural Network (CNN)," *J. Tek. Inform.*, vol. 4, no. 1, pp. 161–170, Feb. 2023, doi: 10.52436/1.jutif.2023.4.1.564.

[23] A. Das, A. Deka, K. Medhi, and M. J. Saikia, "Deep Learning to Authenticate Traditional Handloom Textile," *Information*, vol. 15, no. 8, p. 465, Aug. 2024, doi: 10.3390/info15080465.

[24] N. Yusof, N. A. Abd. Majid, A. Ismail, and N. H. Hussain, "Preserving Songket Heritage Through Intelligent Image Retrieval: A PCA and QGD-Rotational-Based Model," *Computers*, vol. 14, no. 10, p. 416, Oct. 2025, doi: 10.3390/computers14100416.

[25] Y. Chen and Y. Lou, "Deep Learning-Driven Craft Design: Integrating AI Into Traditional Handicraft Creation," *IEEE Access*, vol. 13, pp. 176141–176156, 2025, doi: 10.1109/ACCESS.2025.3614185.

[26] V. Melnyk, "Parametric Pattern Design For Manually Knitted Textile Panels," *Digit. Creat.*, vol. 36, no. 3, pp. 167–177, Jul. 2025, doi: 10.1080/14626268.2025.2514120.

[27] X. Cong and W. Zhang, "Design of geometric flower pattern for clothing based on deep learning and interactive genetic algorithm," *J. Intell. Syst.*, vol. 33, no. 1, pp. 1–13, Jun. 2024, doi: 10.1515/jisys-2023-0269.

[28] C. Cui and S. Nazlina, "Integrating Genetic Algorithms with Biotechnological Insights for the Application of Chinese Traditional Geometric Patterns in Modern Textile Design," *J. Commer. Biotechnol.*, vol. 29, no. 3, pp. 364–376, 2024. [Online]. Available at: https://commercialbiotechnology.com/article-detail/?id=1488.

[29] H. Du, X. Huang, Y. Sheng, and Z. Liu, "Fuzzy Design Structure Matrix-Guided Genetic Algorithm for Optimizing Apparel Production Line Layouts," *Int. J. Cloth. Sci. Technol.*, pp. 1–21, Nov. 2025, doi: 10.1108/IJCST-11-2024-0220.

[30] R. Angira and B. V. Babu, "Non-dominated Sorting Differential Evolution (NSDE): An Extension of Differential Evolution for Multi-objective Optimization," *Proc. 2nd Indian Int. Conf. Artif. Intell. IICAI 2005*, pp. 1428–1443, 2005. [Online]. Available at: https://www.semanticscholar.org/paper/Non-

dominated-Sorting-Differential-Evolution-An-of-Angira-Babu/b995941cf069cfe76b05a9f96608dd5f555f8ae8.

[31] J. Yazdi, Y. H. Choi, and J. H. Kim, "Non-Dominated Sorting Harmony Search Differential Evolution (NS-HS-DE): A Hybrid Algorithm for Multi-Objective Design of Water Distribution Networks," *Water*, vol. 9, no. 8, p. 587, Aug. 2017, doi: 10.3390/w9080587.

[32] A. Cemiloglu, L. Zhu, B. Chen, L. Lu, and Y. A. Nanehkaran, "Enhancing Urban Surface Runoff Conveying System Dimensions through Optimization Using the Non-Dominated Sorting Differential Evolution (NSDE) Metaheuristic Algorithm," *Water*, vol. 15, no. 16, p. 2927, 2023, doi: 10.3390/w15162927.

[33] A. Zade, S. Neogi, and R. R. P. Kuppusamy, "Development of Non-Dominated Sorting Differential Evolution Algorithm for Cure Cycle Optimization of Industrial Composite Parts," *J. Reinf. Plast. Compos.*, vol. 44, no. 1–2, pp. 96–111, 2025, doi: 10.1177/07316844231212527.

[34] L. Wu *et al.*, "Multi-Objective Distribution Network Planning Method with Distributed Generation Based on Non Dominated Sorting Differential Evolution Algorithm," *J. Phys. Conf. Ser.*, vol. 2247, no. 1, p. 12019, Apr. 2022, doi: 10.1088/1742-6596/2247/1/012019.

[35] L. Duan, Y. Liu, H. Li, K.-H. Park, and K. Cao, "Nondominated Sorting Differential Evolution Algorithm to Solve the Biobjective Multi-AGV Routing Problem in Hazardous Chemicals Warehouse," *Math. Probl. Eng.*, vol. 2022, no. 1, p. 3785039, 2022, doi: 10.1155/2022/3785039.

[36] F. Liu *et al.*, "Large Language Model for Multi-objective Evolutionary Optimization." in *Evolutionary Multi-Criterion Optimization: 13th International Conference*, 2024, pp. 1–13 2024. [Online]. Available at: https://dl.acm.org/doi/abs/10.1007/978-981-96-3538-2_13.

[37] W. M. Freire *et al.*, "Large Language Model-Based Suggestion of Objective Functions for Search-Based Product Line Architecture Design," in *Anais do XVIII Simpósio Brasileiro de Componentes, Arquiteturas e Reutilização de Software (SBCARS 2024)*, Sep. 2024, pp. 21–30, doi: 10.5753/sbcars.2024.3833.

[38] N. van Stein and T. Bäck, "LLaMEA: A Large Language Model Evolutionary Algorithm for Automatically Generating Metaheuristics," *IEEE Trans. Evol. Comput.*, vol. 29, no. 2, pp. 331–345, Apr. 2025, doi: 10.1109/TEVC.2024.3497793.

[39] Y. Huang, S. Wu, W. Zhang, J. Wu, L. Feng, and K. C. Tan, "Autonomous Multi-Objective Optimization Using Large Language Model," *IEEE Trans. Evol. Comput.*, pp. 1–1, 2025, doi: 10.1109/TEVC.2025.3561001.

[40] S. Yao, F. Liu, X. Lin, Z. Lu, Z. Wang, and Q. Zhang, "Multi-Objective Evolution of Heuristic Using Large Language Model," *Proc. AAAI Conf. Artif. Intell.*, vol. 39, no. 25, pp. 27144–27152, Apr. 2025, doi: 10.1609/aaai.v39i25.34922.

[41] D. Hasler and S. E. Suesstrunk, "Measuring Colorfulness in Natural Images," Jun. 2003, p. 87, doi: 10.1117/12.477378.

[42] M. Beyeler, *Machine Learning for OpenCV: Intelligent Image Processing with Python*. London, England: Packt Publishing, 2017. [Online]. Available at: https://books.google.co.id/books/about/Machine_Learning_for_OpenCV.

[43] B. Moulden, F. Kingdom, and L. F. Gatley, "The Standard Deviation of Luminance as a Metric for Contrast in Random-Dot Images," *Perception*, vol. 19, no. 1, pp. 79–101, Feb. 1990, doi: 10.1068/p190079.

[44] H. Kukkonen, J. Rovamo, K. Tiippana, and R. Näsänen, "Michelson Contrast, RMS Contrast and Energy of Various Spatial Stimuli at Threshold," *Vision Res.*, vol. 33, no. 10, pp. 1431–1436, Jul. 1993, doi: 10.1016/0042-6989(93)90049-3.

[45] C. Peng, H. Sun, and J. Guo, "Multi-Objective Optimal PMU Placement Using A Non-Dominated Sorting Differential Evolution Algorithm," *Int. J. Electr. Power Energy Syst.*, vol. 32, no. 8, pp. 886–892, Oct. 2010, doi: 10.1016/j.ijepes.2010.01.024.