LNBTI
LANKA NIPPON BIZTECH INSTITUTE

Japanese
IT University

JASPER RESEARCH JOURNAL 2026

VOLUME 1, ISSUE 1

"Japan- Sri Lanka Partnership in Education and Research"

JASPER Research Journal

March 2026

Volume 1 Issue 1

Lanka Nippon BizTech Institute LNBTI

ISSN 3121-4347

No. 278, High Level Road, Maharagama

jasper@edu.lnbti.lk

+94 11 737 0620

ISSN 3121-4347

Editorial Board

The Editorial Board of the JASPER Research Journal supports the academic quality, editorial coordination, and production of the journal. For the inaugural issue, editorial work was primarily coordinated internally by the academic and administrative staff of Lanka Nippon BizTech Institute (LNBTI).

Aim and Scope

JASPER Research Journal is a peer-reviewed journal published by Lanka Nippon BizTech Institute, Sri Lanka. The journal aims to promote original and high-quality research that contributes to knowledge, innovation, and professional practice in the fields of Information Technology and Innovation and Language, Culture, and Education. It provides a platform for scholarly work that is interdisciplinary, practice-informed, and responsive to contemporary local and global challenges.

The journal welcomes original research articles, review papers, and applied academic studies in areas including software engineering, artificial intelligence, data science, IoT, digital systems, educational technology, language learning, multilingual education, cross-cultural communication, student learning, and related interdisciplinary fields. JASPER encourages research that bridges theory and practice, supports academic excellence, and contributes to sustainable social and educational development.

Editorial Note

The JASPER Research Journal is launched as a platform for scholarly work in the areas of Information Technology and Innovation and Language, Culture, and Education. This inaugural issue reflects the journal's commitment to supporting applied, context-relevant, and interdisciplinary research that connects academic inquiry with practical challenges in education, technology, and society.

As a new journal, JASPER is committed to progressively strengthening its editorial

ISSN 3121-4347

systems, reviewer network, publication policies, and scholarly reach in future issues

Peer Review Policy

Peer Review Process

All manuscripts submitted to the JASPER Research Journal undergo an editorial screening followed by a double-blind peer review process. The purpose of the review process is to ensure the academic quality, originality, relevance, and clarity of all published work.

Each submission is first reviewed by the editorial team to determine whether it fits the aim and scope of the journal and meets the basic requirements of academic writing, structure, and ethical compliance. Manuscripts that pass this initial screening are then sent for peer review.

For the inaugural issue, reviewers were selected primarily from the internal academic staff of Lanka Nippon BizTech Institute (LNBTI) based on subject expertise relevant to the manuscript. Reviewers were asked to evaluate submissions independently and provide constructive comments on the originality, methodological soundness, clarity, significance, and overall suitability of the manuscript for publication. The journal intends to expand its pool of external reviewers and editorial advisers in future issues.

The identities of authors and reviewers are concealed during the review process as far as practicable in order to maintain fairness and reduce bias. Based on reviewer comments, the editorial team may make one of the following decisions:

- Accept without revision
- Accept with minor revisions
- Accept with major revisions
- Reject

The final decision regarding publication rests with the editorial team. Authors may be required to revise and resubmit their manuscripts in response to reviewer and editorial comments. The editorial team reserves the right to make minor language, style, and formatting edits to accepted manuscripts for consistency with journal standards.

Reviewers for Volume 1 Issue 1

Manuscripts published in this inaugural issue were reviewed through a double-blind peer review process by academic staff members with relevant subject expertise. For this first issue, reviewers were drawn primarily from the internal academic community of LNBTI.

Review Panel

- Prof. K. P. Hewagamage – Academic Advisor, LNBTI
- Prof. Sandirigama – Director, Global Center, LNBTI
- Dr. Manjula Wanninayaka – Career Development Hub Coordinator, LNBTI
- Ms. Bhashini Wijewickrama – Head, Department of Computing, Senior Lecturer, LNBTI
- Mr. Janith Perera – Programme Coordinator, Senior Lecturer, LNBTI
- Ms. Chobodi Padmaperuma – Programme Coordinator, Lecturer, LNBTI

ISSN 3121-4347

Publication Ethics and Research Integrity

Publication Ethics

The JASPER Research Journal is committed to maintaining high standards of academic integrity, ethical publishing practice, and responsible scholarly communication. Authors, reviewers, and editors are expected to act honestly, transparently, and professionally throughout the publication process.

Responsibilities of Authors

Authors submitting to the journal must ensure that:

- the manuscript is their original work
- the manuscript has not been published previously
- the manuscript is not under consideration elsewhere
- all sources are properly acknowledged and cited
- authorship accurately reflects those who made substantial academic contributions
- data, findings, and interpretations are presented honestly and without fabrication, falsification, or misleading manipulation
- any use of human participants, personal data, or sensitive material has been handled in accordance with relevant ethical requirements

Authors should also disclose, where applicable:

- conflicts of interest
- funding sources
- acknowledgements
- ethical clearance or approval status

Responsibilities of Reviewers

Reviewers are expected to:

- evaluate manuscripts fairly, objectively, and confidentially
- provide constructive and timely feedback
- avoid personal criticism of authors
- inform the editor if they identify substantial similarity, plagiarism, or ethical concern
- decline review where there is a conflict of interest or insufficient expertise

Responsibilities of Editors

Editors are responsible for:

- ensuring a fair and unbiased review process
- making editorial decisions based on academic merit, relevance, and journal standards
- preserving reviewer confidentiality
- addressing ethical concerns, complaints, and disputes appropriately
- taking reasonable steps to correct, clarify, or retract published content where necessary

ISSN 3121-4347

Misconduct

The journal does not accept:

- plagiarism
- duplicate or redundant publication
- fabricated or falsified data
- misleading authorship claims
- undisclosed conflicts of interest
- unethical research practices

Where credible concerns arise, the journal reserves the right to request clarification, corrections, withdrawal, or retraction of a manuscript or published article.

Conflict of Interest Policy

Authors, reviewers, and editors must disclose any actual or potential conflict of interest that could influence, or reasonably appear to influence, their judgment.

Examples include:

- financial interests
- institutional relationships
- personal relationships
- supervisory or collaborative relationships
- competing academic or professional interests

Where a conflict exists, the journal may take appropriate steps, including reassignment of review, disclosure in the published article, or editorial recusal.

Copyright and Licensing

Copyright

Copyright of published articles shall remain with the author or authors unless otherwise stated by the journal. By submitting a manuscript and accepting publication, authors grant the JASPER Research Journal the right to publish, archive, reproduce, and disseminate the work as part of the journal in print, electronic, and related formats.

Authors are responsible for ensuring that any copyrighted material included in the manuscript has been used with proper permission where required.

Corrections, Retractions, and Editorial Actions

The JASPER Research Journal is committed to preserving the accuracy and integrity of the scholarly record.

Where minor errors are identified in a published article that do not invalidate the overall findings, the journal may publish a correction or erratum.

Where major problems are identified, including serious methodological flaws, ethical breaches, plagiarism, fabricated data, or duplicate publication, the journal may publish:

- a correction
- an expression of concern
- a withdrawal notice
- a retraction

Retraction decisions are made by the editorial team after appropriate review of the

ISSN 3121-4347

available evidence and, where relevant, consultation with the author or authors. A retracted article will remain part of the publication record but will be clearly marked as retracted.

Complaints and Appeals Procedure

Authors may contact the journal regarding editorial decisions, procedural concerns, or publication ethics issues. Complaints and appeals should be submitted in writing to the journal's official correspondence address and should clearly state the nature of the concern.

The editorial team will review the matter and respond in a fair and timely manner. Where necessary, the matter may be referred to the Editor-in-Chief or the wider editorial leadership for further consideration.

Correspondence: All correspondence regarding journal articles should be addressed to

Mrs. Carmaline Nicholas
Secretary – JASPER Research Journal,
No. 278, High Level Road,
Maharagama,
Sri Lanka
Tel: +94 11 737 0620, E-mail: jasper@edu.lnbti.lk

ISSN 3121-4347

JASPER Research Journal

Volume 1 Issue 1,
March 2026

Track 1 - Information Technology & Innovation

Track 2 - Language, Culture & Education

ISSN 3121-4347

JASPER Research Journal
Volume 1 Issue 1
March 2026

Track 1 - Information Technology & Innovation

ISSN 3121-4347

JASPER Research Journal
Volume 1 Issue 1
March 2026

Bridging Heritage and Market: An AI Framework for Analyzing Sri Lankan Handloom Designs for Alignment with Japanese Aesthetic Trends

Rashmikara R[1], Jayathilake D B M[2]

[1,2]LNBTI Japanese IT University, Maharagama, Sri Lanka

[1]ronith.bsc.se05@lnbti.lk

DOI: https://doi.org/10.5281/zenodo.19703123

Abstract

This research presents a comprehensive framework for analyzing cross-cultural textile design patterns between Japan and Sri Lanka, to address the critical knowledge gap in market intelligence for the handloom industries. Through employing computer vision and machine learning techniques, this study analyzes 2000 textile images (1000 Japanese, 1000 Sri Lankan) to establish quantitative "Design DNA "profiles for both cultural traditions. The methodology extracts 44 features per specimen, including dominant color palettes though k-means clustering, pattern complexity via edge detection and entropy calculations, motif calculations though geometric ratio analysis, and texture characterization using Local Binary Patterns and contrast metrics. The statistical analysis reveals significant aesthetic divergences: Sri Lankan textiles demonstrate warmer color palettes (warm_cool_score: 0.178 vs -0.111, $p<0.001$, $d=3.059$) with earth-toned RGB profiles, while Japanese textiles exhibit more cooler, muted tones (RGB: 151,156,180) containing wabi-sabi principles. Pattern complexity analysis shows Japanese textiles possessing 64.3% higher texture complexity ($p=0.002$, $d=1.590$) and superior texture contrast ($p=0.010$, $d=1.287$), reflecting centuries of refined weaving techniques including kasuri and shibori. Symmetry analysis shows *Japanese design*s favor asymmetry (0.055 average symmetry score) while Sri Lankan patterns demonstrate greater regularity (0.172) aligning with respective cultural philosophies of natural irregularity vs geometric traditions. 4 of the 13 metrics demonstrate statistically significant differences with very larger effect sizes ($d>1.2$), reconfirming evidence for distinct cultural aesthetic signatures.

Keywords: textile design analysis, computer vision, cultural preservation, market intelligence, cross-cultural aesthetics, design DNA, Japanese wabi-sabi, Sri Lankan handloom

ISSN 3121-4347

1. Introduction

1.1 Background and Context

The global textile industry represents a multi-trillion-dollar market in which cultural heritage intersects with fashion trends and international trade. Within this landscape, traditional handlooms serve as repositories of cultural identity, technical knowledge, and artistic expression spanning centuries. Sri Lanka's handloom sector stands as a vital component of the nation's cultural and economic fabric, supporting approximately 75,000 artisans across rural communities. These craftspeople preserve techniques passed down through generations, including the symmetrical geometric patterns of Dumbara weaving and the lace-like textures of Beeralu, representing a living cultural heritage with origins tracing back to the 6th century BC (Fernando, 2023).

Conversely, the Japanese textile market, valued at approximately $15 billion annually, is characterised by sophisticated consumer preferences that emphasize minimalist aesthetics, superior craftsmanship, and deep cultural authenticity. Japanese aesthetic principles such as wabi-sabi (the appreciation of beauty in imperfection) and Takumi Spirit (masterful craftsmanship and dedication to precision) create unique market requirements that remain challenging for foreign producers. These principles manifest in textile design through subtle color palettes, asymmetric compositions, and textures that celebrate natural materials.

1.2 Problem Statement

The core challenge facing traditional textile industries in an era of rapid globalization is a knowledge gap regarding international market preferences and aesthetic trends. Market misalignment occurs when traditional designs, despite their cultural significance and technical excellence, fail to resonate with international consumer preferences. For Sri Lankan handloom producers, preliminary observations suggest that the bold, high-contrast colors and complex geometric patterns characteristic of domestic traditions do not align with Japanese minimalist preferences, which emphasize subtlety and restraint.

The economic opportunity loss stemming from failure to penetrate premium international markets limits revenue potential and artisan income growth. Notably, the Japanese market

ISSN 3121-4347

demonstrates a strong appreciation for handcrafted quality and a willingness to pay premium prices for authentic cultural products, representing a largely untapped opportunity for Sri Lankan artisans if design alignment can be achieved. In parallel, cultural erosion risks emerge when traditional crafts become economically unsustainable without market adaptation, threatening the preservation of intangible cultural heritage. The loss of traditional textile techniques would represent an irreversible depletion of global cultural diversity and human knowledge.

1.3 Research Objectives

This study addresses the aforementioned challenges through the following specific objectives:

1. Develop an AI-driven computational framework for the systematic analysis of textile design using computer vision and machine learning techniques, capable of extracting quantitative features from large-scale image datasets.
2. Establish a comprehensive Design DNA database quantifying characteristic features of Japanese and Sri Lankan textile traditions across multiple aesthetic dimensions, including color, pattern complexity, motif classification, and texture characteristics.
3. Identify and measure market alignment gaps through statistical analysis of aesthetic differences between production characteristics and market preferences, determining which features represent cultural authenticity and which represent adaptation opportunities.
4. Generate actionable recommendations for design adaptation and cross-cultural appreciation that enable market competitiveness while preserving the essential elements of cultural heritage and craftsmanship identity.

1.4 Research Significance

This study contributes to multiple domains of knowledge and practice. From an academic perspective, it advances computational cultural analysis by demonstrating how machine learning and computer vision can extract meaningful cultural signatures from visual artifacts. The methodological framework establishes a quantitative approach for measuring traditionally subjective aesthetic concepts. It also contributes to the broader discussion on

ISSN 3121-4347

the role of artificial intelligence in cultural industries, illustrating how computational tools can augment rather than replace human creativity and craftsmanship. By transforming aesthetic intuition into quantifiable insights, this approach empowers traditional craftspeople with modern market intelligence while respecting the depth and subtlety of cultural knowledge that cannot be reduced to algorithms.

2. Literature Review

2.1 Traditional Textile Heritage

Traditional textiles have been studied as an anthropological subject, a historical topic in art history, a scientific subject in materials research, and a cultural study reflecting the multidimensional nature of human societies. Beyond functional objects, textiles operate as artistic symbols and markers of status and cultural identity, offering a rich medium for cross-cultural analysis.

The tradition of Sri Lankan handlooms dates back over 2,500 years, with regional specialties shaped by local materials, cultural influences, and functional requirements (Fernando, 2023). The Dumbara region is a source of uniquely patterned textiles featuring triangular and square motifs dating to the 6th century BC, connected to myths of Queen Kuveni and the Yaksha tribe (Khiri Travel, 2023). These patterns exhibit geometrical accuracy and mathematical proportions, implying advanced indigenous knowledge of proportion and balance. Beeralu lace, a complex openwork design created by southern coastal communities using pillow-lace techniques, was introduced by Portuguese settlers in the 16th century and subsequently modified to suit Sri Lankan aesthetic traditions (Handloom Corporation of Sri Lanka, 2023).

2.2 Computational Textile Analysis

Computational textile analysis represents an interdisciplinary field combining computer vision, machine learning, and cultural analytics. Initial applications were technical in nature, focusing on defect detection, weaving pattern recognition, and quality management in manufacturing environments (Kuo et al., 2019). These methods encompassed edge detection for identifying fabric flaws, texture analysis for classifying weave structures, and color quantification for quality monitoring.

ISSN 3121-4347

Subsequent studies extended computational approaches to cover the aesthetic and cultural aspects of textile analysis. K-means clustering for color analysis has emerged as a standard method for determining dominant color palettes in textile design, with applications spanning fashion trend interpretation and cultural pattern recognition (Li et al., 2020). Entropy calculation and frequency domain analysis provide methods for measuring pattern complexity, enabling quantitative comparison of visual complexity across textile cultures (Zhang et al., 2021). Local Binary Patterns (LBP) and Gray-Level Co-occurrence Matrices (GLCM) characterize surface texture for manufacturing quality assessment and relative quality comparison (Haralick et al., 1973; Ojala et al., 2002).

2.3 Cultural Aesthetics and Market Intelligence

Cultural aesthetics, drawing on philosophy, psychology, anthropology, and marketing, examines how culture influences aesthetic preferences and market behavior. Japanese aesthetic philosophy encompasses concepts such as wabi-sabi (beauty in imperfection and transience), ma (negative space in art), yugen (profound mysterious depth), and shibui (subtle, unobtrusive beauty) (Mori, 2022). These principles emphasize subtlety, moderation, and admiration for natural materials and processes over ostentatious adornment or flawless symmetry.

Research in cross-cultural aesthetics has identified systematic variations in visual preferences. Masuda and Nisbett (2001) established that East Asian observers attend more to contextual relationships and background elements, whereas Western observers focus on focal objects. These perceptual differences influence preferences for pattern arrangement and spatial relationships in textile design. Color psychology studies conducted across cultures have determined that color-emotion associations vary considerably, albeit with identifiable patterns (Madden et al., 2000).

Market intelligence applications have emerged in the fashion and design sectors, employing computational techniques to determine trends, predict consumer preferences, and inform design decisions. Wang et al. (2018) introduced deep learning methods to predict fashion trends from social media imagery, and Iqbal et al. (2020) designed recommender systems based on textile style similarities. These applications have been largely confined to commercialized fashion settings; the issue of balancing market adaptation with cultural authenticity in traditional crafts is rarely addressed.

ISSN 3121-4347

2.4 Research Gaps

Several critical gaps emerge in the existing literature that this study addresses. First, there is a lack of comprehensive computational analysis comparing Japanese and Sri Lankan textile traditions, despite historical connections and potential for cross-cultural exchange. Previous studies have typically examined a single cultural tradition in depth or compared one with a Western context, neglecting South Asian–East Asian aesthetic relationships.

Second, existing computational approaches typically isolate individual aspects of textile aesthetics (color, pattern, or texture) rather than deploying integrated frameworks that capture multiple dimensions simultaneously. This fragmentation limits understanding of how aesthetic elements interact to create cohesive cultural signatures. Finally, the literature reveals limited integration of computational approaches with deep cultural understanding of aesthetic philosophies. Effective analysis of traditional textiles requires both technical and cultural literacy, combining quantitative methods with qualitative understanding of the values, histories, and meanings embedded in textile traditions.

3. Methodology

3.1 Research Design

The study employs a quantitative analysis framework combining computer vision and statistical methods to identify and measure aesthetic differences between Japanese and Sri Lankan textile traditions. The design follows a systematic approach encompassing data collection from verified cultural sources, feature extraction using established computer vision techniques, statistical analysis to identify differences, and interpretation from cultural and market perspectives.

3.2 Dataset Construction

The sample comprises 2,000 high-resolution textile images (n = 1,000 Japanese; n = 1,000 Sri Lankan). Images were collected by scraping publicly available textile photographs from multiple online repositories and cultural documentation websites to ensure authenticity and cultural representativeness.

ISSN 3121-4347

Dataset Provenance. Japanese textile specimens were gathered from publicly accessible online sources, including cultural heritage websites, textile documentation platforms, academic image repositories, and verified textile artist portfolios. Sri Lankan textile specimens were collected from publicly available handloom documentation sites, textile export catalogs, academic collections, and cultural heritage archives whose source terms permit public viewing and usage.

Selection Criteria. Images were filtered to retain examples representing traditional textile patterns and weaving styles. The selection process excluded mass-produced textiles, contemporary fusion designs that deviate from traditional aesthetics, duplicate images, and specimens with insufficient resolution or incomplete documentation. Japanese textile samples were chosen to represent the breadth of Japanese textile heritage, covering such types as kasuri textiles from Kurume and Hiroshima, shibori pieces from Arimatsu and Kyoto, and works by modern designers who maintain traditional aesthetic systems. The selection criteria prioritized handcrafted specimens with a strong indication of traditional methods over mass-produced industrial textiles, ensuring genuine cultural representation. Sri Lankan textile samples were chosen for regional and stylistic variety, encompassing Dumbara mats with characteristic geometric designs, Beeralu lace formations from the southern coast, handloom sarees from established weaving hubs, and traditional batik samples representing typical Sri Lankan motifs. Selection focused on specimens with documented Sri Lankan provenance and genuine technique, excluding specimens subjected to aesthetic modification for foreign markets.

3.3 Dataset Availability

The complete JASPER (Japanese x Sri Lankan Pattern Extraction Resource) textile image dataset is publicly available on Kaggle at https://kaggle.com/datasets/aa558fabe0effe5ab4622daae78a4773e2900dec1197be13f918bee4f4180192. The dataset includes: (1) 2,000 preprocessed textile images (1,000 Japanese and 1,000 Sri Lankan) in PNG format at 1200 × 700 pixels; (2) a comprehensive metadata file with source attributions, image identifiers, and licensing information; (3) a source reference list with restricted-image access instructions; and (4) a data dictionary documenting all extracted features. The dataset is released under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, permitting modification, distribution, and reuse with attribution. Complete analysis code, feature extraction scripts, and statistical analysis notebooks are available in the accompanying GitHub repository at https://github.com/InfiniteBloom-max/Jasper.

ISSN 3121-4347

3.4 Image Preprocessing

All images underwent standardized preprocessing to ensure consistent feature extraction. Specimens were resized to 1200 × 700 pixels using bicubic interpolation (OpenCV INTER_CUBIC) to balance computational efficiency with sufficient preservation of texture information for analysis. Images were manually cropped to exclude borders, tags, display backgrounds, and contextual elements, retaining only the textile surface. Semi-automated region-of-interest detection using edge detection (Canny algorithm, thresholds 50/150) assisted in identifying textile boundaries, with manual verification ensuring accuracy.

To mitigate photographic lighting variations, histogram equalization was applied separately to each RGB channel. White balance correction using the gray-world assumption (setting mean RGB values to 128) was employed to standardize illumination. No aggressive color transformation was applied, in order to preserve authentic cultural color characteristics. Lighting and exposure variations inherent in archival photography were accepted as limitations and acknowledged in the interpretation of results. Images with resolution below 800 pixels on the shortest dimension, severe lighting artifacts, visible damage or staining, or digital compression artifacts were excluded during initial screening.

Validation checks confirmed that extracted features behaved as expected: high-entropy samples corresponded to visually complex patterns, high-symmetry scores aligned with geometric designs, and warm color scores matched visual assessment of earth tones. The preprocessing pipeline was implemented in Python 3.9 using OpenCV 4.5, NumPy 1.21, and scikit-image 0.18.

3.5 Feature Extraction Methods

The computational framework extracts 44 features per textile sample across four principal categories: color features, pattern complexity, motif characteristics, and texture properties. Extraction was implemented in Python using OpenCV, scikit-image, and scikit-learn libraries to ensure reproducibility and methodological transparency.

3.5.1 Color Analysis

Color characteristics were extracted through complementary methods. Dominant color clusters were identified using the K-means clustering algorithm with K = 5 clusters applied to RGB pixel values (Lloyd, 1982). The algorithm identifies centroid RGB values and the

ISSN 3121-4347

proportion of pixels per cluster, producing 20 features (5 colors × 3 RGB channels + 5 percentage values). This method captures visually dominant colors and their relative prominence within each specimen.

Additional color measures include: average values of pixels in each RGB channel, standard deviation of each channel, average saturation and brightness in the HSV color space, total color variance (the sum of variance across RGB channels), and a warm/cool score computed as (red − blue) / 255, providing a continuous measure of color temperature. These aggregate features complement the dominant palette by describing overall color appearance beyond the most prominent colors.

3.5.2 Pattern Complexity Analysis

Pattern complexity was quantified through complementary measures reflecting visual intricacy. Edge-based analysis, employing Canny edge detection (Canny, 1986), captures the boundaries of pattern elements and yields edge density (percentage of edge pixels) and edge magnitude (standard deviation of edge magnitudes), through which pattern transitions are measured in both frequency and intensity. The unpredictability of pattern organization was quantified using Shannon entropy (Shannon, 1948):

$$H = -\Sigma\, p(x) \log_2 p(x)$$

where p(x) is the probability of gray-level x. Higher entropy implies greater visual complexity and unpredictability. Frequency domain complexity was calculated by applying a Fast Fourier Transform (FFT) to the grayscale image and computing the ratio of high-frequency energy to total frequency energy, capturing fine-scale frequency variations in the pattern.

Local variance, measured as the average variance of pixel values within 5 × 5 neighborhoods across the image, serves as a texture complexity metric indicating local contrast and pattern density. This measure is useful for differentiating textile traditions whose weaving methods produce distinct levels of fine-scale pattern variation.

3.5.3 Motif Classification and Geometric Analysis

Motif classification and geometric analysis measure the organizational principles of design patterns. The geometric ratio, computed by dividing horizontal by vertical edge elements,

ISSN 3121-4347

provides a measure of directional emphasis in pattern organization. Values above 1.0 indicate horizontal dominance, values below 1.0 indicate vertical dominance, and values near 1.0 indicate balanced, isotropic organization.

Symmetry was measured using a four-quadrant reflection method: each image was subdivided into quadrants and correlation coefficients were calculated between reflected quarters. The resulting correlation yields a composite symmetry score ranging from 0 (no symmetry) to 1.0 (perfect symmetry). This methodology captures both bilateral and rotational symmetry common in traditional textile patterns.

3.5.4 Texture Characterization

Texture analysis employed Local Binary Patterns (LBP), an effective texture descriptor that represents local patterns by comparing each pixel with its neighbors (Ojala et al., 2002). The LBP histogram captures the frequency distribution of texture patterns across the image, and LBP uniformity, defined as the fraction of uniform binary patterns (those with at most two bit transitions), reflects the regularity of texture. Higher uniformity values correspond to more regular, repetitive textures.

Complementary texture analysis was conducted using the Gray-Level Co-occurrence Matrix (GLCM) to measure texture contrast, which quantifies intensity differences between adjacent pixels (Haralick et al., 1973). Higher contrast values indicate greater local variation in texture, implying more pronounced surface features from finer weaving.

3.6 Statistical Analysis Methods

The statistical analysis followed a rigorous algorithmic process to evaluate differences between Japanese and Sri Lankan textiles across the extracted features. Each feature was characterized descriptively, computing means and standard deviations for each cultural group to provide an initial characterization of distributions. Continuous features were subjected to independent-samples t-tests with an alpha level of 0.05 to determine statistical significance (Cohen, 1988). Shapiro–Wilk tests assessed the normality of distributions, and Levene's tests evaluated homogeneity of variances. Where assumptions of normality or equal variances were violated, Mann–Whitney U tests were employed as non-parametric alternatives.

Effect sizes were computed using Cohen's d for t-tests and rank-biserial correlation for Mann–Whitney U tests, ensuring that both statistical significance and practical significance were evaluated (Cohen, 1988). Interpretations follow standard conventions: small (d ≈

ISSN 3121-4347

0.20), medium ($d \approx 0.50$), large ($d \approx 0.80$), and very large ($d \geq 1.20$). Sample sizes were sufficient for statistical power ($1 - \beta \geq 0.95$) to detect medium effects.

Because 44 features were analyzed, multiple comparison correction was applied using the Benjamini–Hochberg False Discovery Rate (FDR) procedure at alpha = 0.05 to control Type I error inflation (Benjamini & Hochberg, 1995). Only features that survived FDR correction are reported as statistically meaningful, ensuring robustness of the identified differences. Additionally, emphasis was placed on effect size magnitudes (Cohen's d) alongside p-values, with features demonstrating large effects ($d \geq 1.2$) highlighted as substantively meaningful beyond statistical significance. All statistical analyses were performed in Python using scipy and statsmodels libraries.

3.7 Software Implementation

The analytical framework was implemented in Python 3.11 using established libraries for scientific computing and image analysis. Core libraries included NumPy and pandas for data manipulation, OpenCV and scikit-image for computer vision operations, scikit-learn for machine learning algorithms including K-means clustering, and matplotlib and seaborn for visualization. Statistical analysis was conducted with scipy.stats and statsmodels. The implementation follows modular design principles, with separate functions for each feature extraction category to enable easy modification and extension. Code was documented with inline comments and organized within Jupyter notebooks to ensure reproducibility.

4. Results and Discussion

The results reveal a pronounced aesthetic divergence between Japanese and Sri Lankan textile cultures, with statistically meaningful differences across multiple dimensions of color, pattern complexity, motif organization, and texture. The subsections below discuss these findings in detail.

4.1 Color Characteristics Analysis

4.1.1Dominant Color Palettes

The dominant color palettes reveal fundamentally different color philosophies between the two traditions. Japanese textiles exhibit cool-toned palettes with average RGB values of (151.5, 155.8, 179.9), with blue and purple tones prevalent across dominant color clusters. The dominant color in Japanese textiles (34.9% of pixels on average) is a light blue-gray,

ISSN 3121-4347

followed by a darker blue-gray (20.3%) and purple hues (18.8%). This balance of subdued colors aligns with the wabi-sabi aesthetic principle, emphasizing harmonic tones that evoke natural materials and atmospheric conditions.

Sri Lankan textiles, by contrast, display warm, earthy palettes with average RGB values of (199.5, 183.6, 154.2) and a strong dominance of red and orange hues. The dominant color (44.1% of pixels) is beige/cream, followed by tan/brown (18.8%) and olive/khaki (14.7%) tones. These palettes reflect Sri Lanka's natural environment, including sandy beaches, earth-rich landscapes, and a warm climate, together with cultural preferences for vibrant, life-affirming colors in traditional contexts.

4.1.2 Color Temperature Contrast

The warm–cool score provides a continuous measure of color temperature. Sri Lankan textiles score significantly warmer ($M = 0.178 \pm 0.129$) than Japanese textiles ($M = -0.111 \pm 0.035$). The difference is highly statistically meaningful ($t = -6.839$, $p < 0.001$) with a very large effect size (Cohen's $d = -3.059$), indicating an approximate 29% difference in color temperature. This finding confirms that color temperature is among the most pronounced distinctions between the two traditions.

Red-channel values are higher in Sri Lankan textiles ($M = 199.55 \pm 39.0$) than Japanese textiles ($M = 151.55 \pm 18.8$), a statistically meaningful difference ($t = -3.507$, $p = 0.0025$) with a large effect size ($d = -1.568$). Blue-channel values show a directional but non-significant difference ($p = 0.098$) with a medium-to-large effect size ($d = 0.782$). No meaningful difference was observed in saturation between traditions (Japanese: 90.7 ± 18.2; Sri Lankan: 92.6 ± 42.1), indicating that neither tradition favors extreme saturation or desaturation. Brightness showed a near-significant trend ($p = 0.079$), with Sri Lankan textiles approximately 10% brighter ($M = 208.2 \pm 31.4$) than Japanese textiles ($M = 188.5 \pm 11.5$).

4.1.3 Color Variance Distribution

Color variance, representing color diversity within a textile, is higher in Japanese textiles ($M = 5863.1 \pm 173.2$) than Sri Lankan textiles ($M = 4838.4 \pm 1595.2$). While this difference does not quite reach statistical significance ($p = 0.059$), the large effect size ($d = 0.903$) suggests a meaningful difference in color organization. Japanese textiles tend to integrate a broader range of colors within individual specimens, whereas Sri Lankan textiles favor more limited palettes grounded in warm, earth-based tones.

ISSN 3121-4347

Figure Section: 1 presents the comparative color analysis, illustrating the divergent RGB profiles, warm–cool score distributions, extracted palettes, and variance patterns across the two traditions.

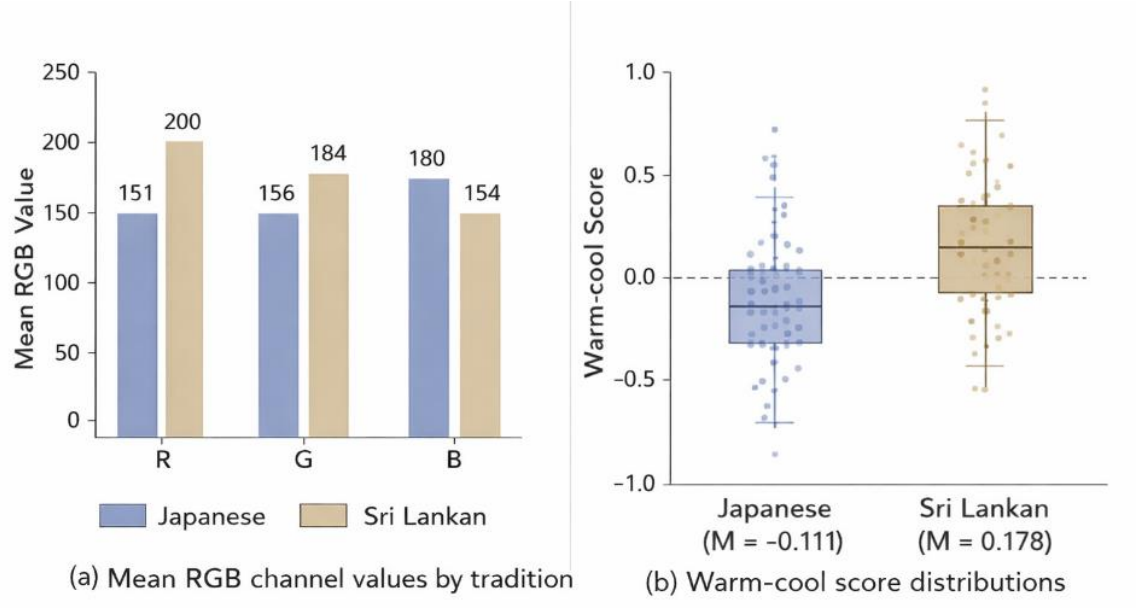

Figure 1
comparative color analysis, warm–cool score distributions, extracted palettes, and variance patterns across the two traditions.

(a) RGB channel values by tradition, illustrating the cool blue–purple dominance of Japanese

textiles and the warm red–earth tones of Sri Lankan textiles.

(b) Warm–cool score distributions showing a 29% temperature divergence ($p < 0.001$, $d = 3.059$).

ISSN 3121-4347

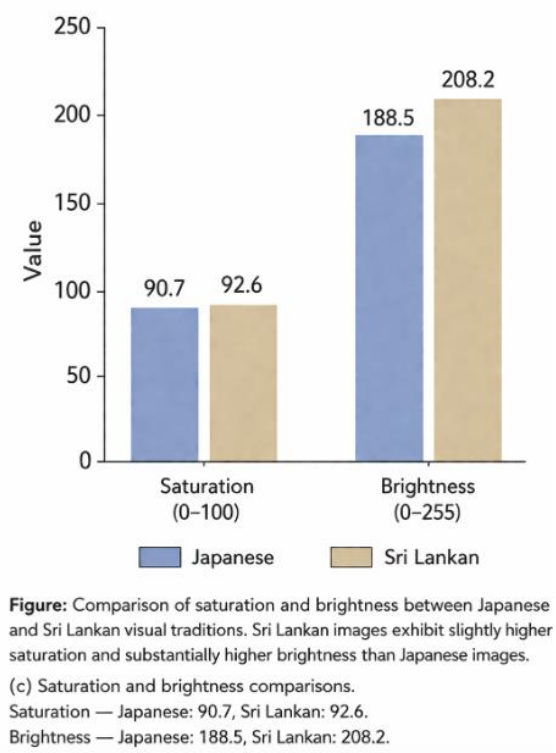

Figure: Comparison of saturation and brightness between Japanese and Sri Lankan visual traditions. Sri Lankan images exhibit slightly higher saturation and substantially higher brightness than Japanese images.

(c) Saturation and brightness comparisons.
Saturation — Japanese: 90.7, Sri Lankan: 92.6.
Brightness — Japanese: 188.5, Sri Lankan: 208.2.

Figure 2
Comparison of saturation and brightness between traditions

(c) Saturation and brightness level comparisons.

ISSN 3121-4347

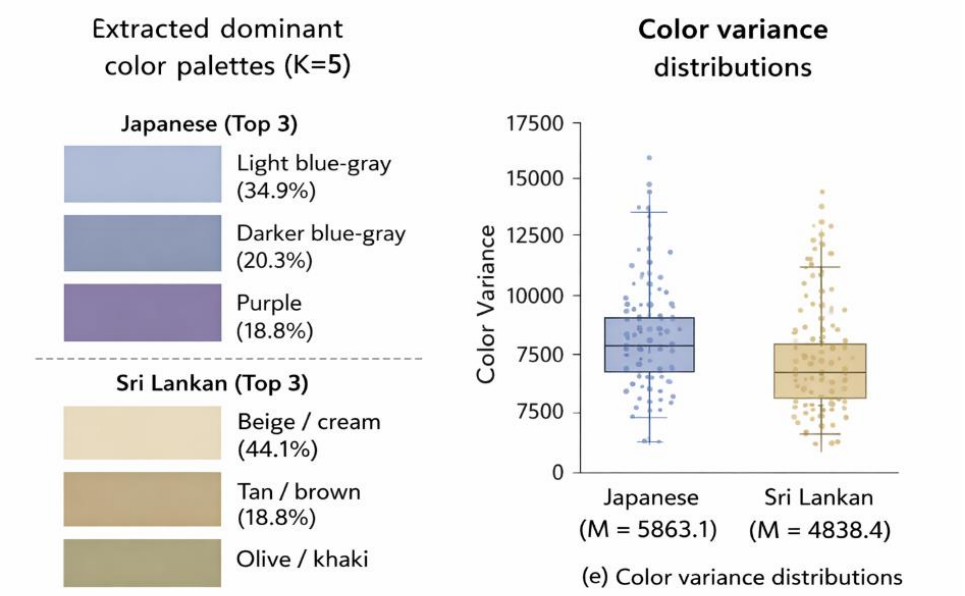

Figure: Comparison of color palettes and color variance between Japanese and Sri Lankan visual

Figure 3
Comparison of color palettes and color variance between traditions

(d) Extracted dominant color palettes from K-means clustering (K = 5).

(e) Color variance distributions indicating greater chromatic diversity in Japanese specimens.

ISSN 3121-4347

4.2 Pattern Complexity Analysis

Pattern complexity analysis reveals notable differences in visual intricacy and content between the two traditions, with Japanese textiles scoring higher on most complexity metrics.

4.2.1 Edge-Based Complexity

Edge magnitude values, a measure of the strength of pattern transitions and boundaries, are higher in Japanese textiles (9.8 ± 0.7) than Sri Lankan textiles (8.6 ± 2.2). Although this difference is not statistically significant ($p = 0.140$), the medium effect size (Cohen's $d = 0.691$) suggests a meaningful trend toward increased pattern definition in Japanese textiles. Both traditions exhibited uniform edge density, attributable to image preprocessing and the absence of sharp background–textile edges.

4.2.2 Entropy and Content

Shannon entropy values, reflecting content and pattern unpredictability, are notably higher in Japanese textiles (7.56 ± 0.26) than Sri Lankan textiles (6.81 ± 1.15). This difference approaches statistical significance ($p = 0.057$) with a large effect size ($d = 0.908$). The elevated entropy observed in Japanese textiles reflects increased pattern complexity and visual density, consistent with advanced weaving techniques such as kasuri resist-dyeing, which produces multi-layered patterns characterized by fine-grained richness.

4.2.3 Frequency Domain and Texture Complexity

The most pronounced differences in pattern complexity emerge in the analysis of texture complexity through local variance. Japanese textiles exhibit substantially higher texture complexity (4758.5 ± 520.8) compared with Sri Lankan textiles (2895.4 ± 1572.8). This difference is statistically meaningful ($t = 3.556$, $p = 0.0023$) with a very large effect size ($d = 1.590$). The higher local variance in Japanese textiles indicates finer-scale imagery and more complex local pattern variations, reflecting highly developed weaving techniques and meticulous attention to microscopic design structures.

The high-frequency ratio, representing the proportion of fine-scale energy in the frequency domain, showed nearly identical values across traditions (Japanese: 0.755 ± 0.035; Sri Lankan: 0.757 ± 0.030), indicating comparable proportions of fine-scale energy at the image level. This metric appears less sensitive to the stylistic distinctions that differentiate

ISSN 3121-4347

textile traditions compared with texture complexity measures.

4.2.4 Cultural Interpretation

The observed differences in pattern complexity are consistent with the historical and technical knowledge of these textile traditions. Japanese textiles have evolved over centuries through highly specialized techniques: kasuri weavers dedicate themselves to perfecting the binding and alignment required for intricate patterns; shibori artists develop complex resist-dyeing processes that create layered, three-dimensional textures; and contemporary designers continue this tradition of technological refinement and pattern sophistication (Mori, 2022). Quantitative evidence of this technical advancement is reflected in the higher texture complexity.

Sri Lankan textiles demonstrate exceptional craftsmanship and cultural significance, but their pattern structures tend to be more direct, emphasizing striking geometric forms and distinct repeat units over subtle multi-layered complexity. Dumbara textile patterns rely on repeated triangles and squares, while Beeralu lace designs employ structured knotting techniques (Fernando, 2023). The lower texture complexity observed is not indicative of inferior quality but reflects alternative aesthetic priorities that emphasize clarity and cultural recognizability over nuanced intricacy.

These findings carry implications for market alignment. Japanese consumers, accustomed to textiles with higher pattern complexity and finer detail, may perceive Sri Lankan textiles as comparatively simple. Conversely, the simplicity and familiarity of Sri Lankan designs may serve as a unique proposition in markets that favor culturally distinct and visually straightforward design aesthetics.

ISSN 3121-4347

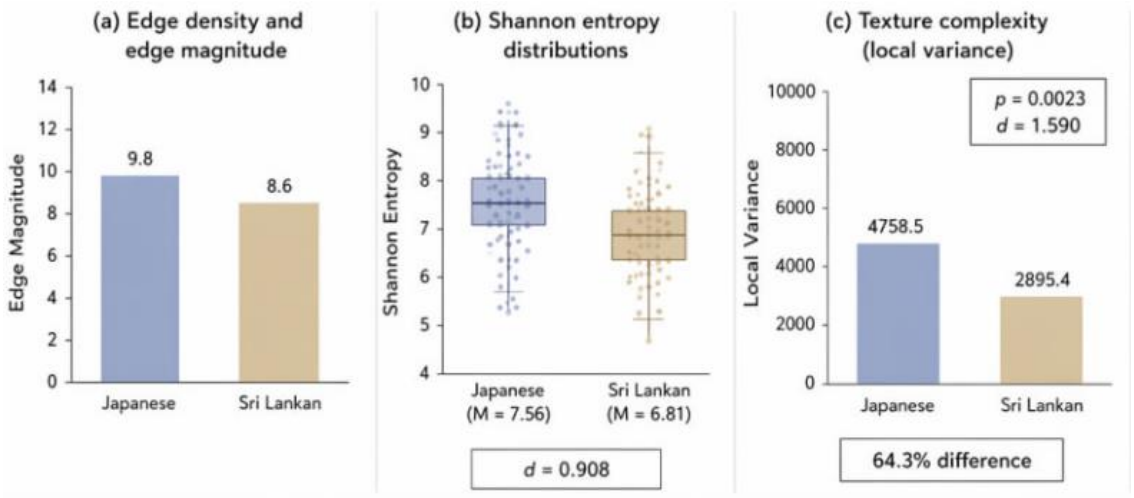

Figure 4

complexity differences, displaying edge density, entropy, texture complexity, and normalized composite scores across the two traditions.

(a) Edge density and edge magnitude values reflect boundary definition.

(b) Shannon entropy distributions showing higher content in Japanese textiles (d = 0.908).

(c) Texture complexity (local variance) demonstrating a 64.3% difference favoring Japanese textiles (p = 0.0023, d = 1.590).

ISSN 3121-4347

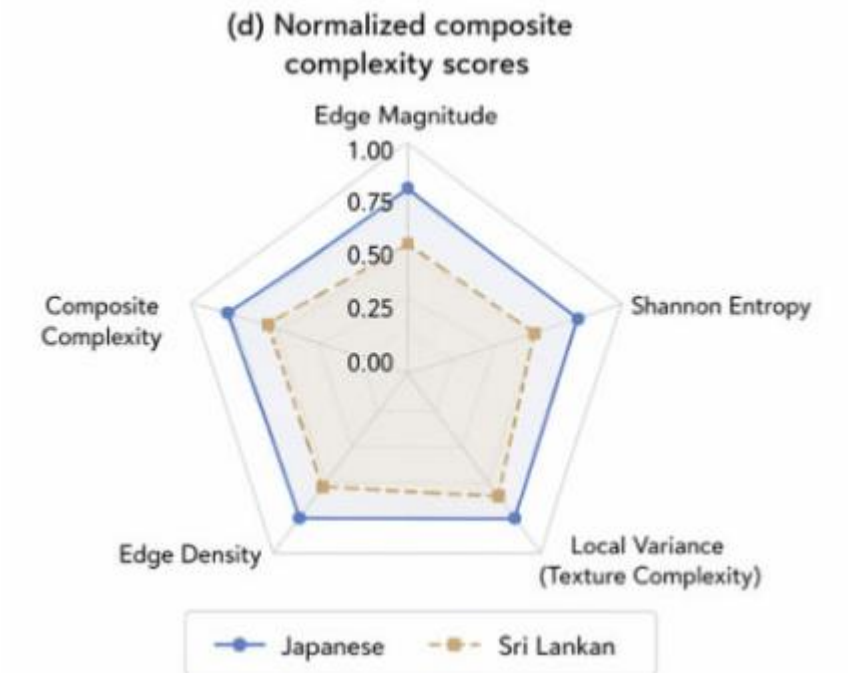

Figure 5

composite complexity scores across all metrics

(d) Normalized composite complexity scores across all metrics

4.3 Motif Classification and Design Characteristics

Geometric analysis and motif classification reveal strong distinctions in the organizational principles of the two textile traditions, rooted in broader cultural concepts of naturalness and order.

Directional Emphasis and Geometric Ratio

ISSN 3121-4347

The geometric ratio, which quantifies the balance of vertical and horizontal elements, yields similar values across traditions, indicating that patterns in both Japanese and Sri Lankan textiles incorporate multi-directional elements without strong unidirectional bias. This is a somewhat unexpected result, given that Japanese textiles commonly feature diagonal compositions (e.g., kasuri patterns) while Sri Lankan Dumbara textiles typically exhibit strong horizontal and vertical axes. The lack of differentiation may reflect that this measure captures aggregate directionality rather than the stylistic nuances of each tradition's orientation types.

Symmetry Analysis

Symmetry scores reveal a noteworthy difference between traditions. Sri Lankan textiles are more symmetrical ($M = 0.172 \pm 0.248$) than Japanese textiles ($M = 0.055 \pm 0.057$), although this difference does not reach statistical significance ($p = 0.164$, $d = -0.649$). The trend suggests that Sri Lankan designs favor orderliness, exemplified by the regularity of Dumbara geometric motifs and the repetition of shapes in Beeralu lace. Japanese textile designs, by contrast, tend toward asymmetry and irregularity. This preference aligns with wabi-sabi aesthetics, which celebrate natural variation and de-emphasize perfection (Mori, 2022).

Importantly, lower symmetry scores do not imply a lack of design structure but rather reflect a deliberate aesthetic choice. The notably larger standard deviation in Sri Lankan symmetry scores ($SD = 0.248$ vs. 0.057 for Japanese) indicates greater diversity within Sri Lankan textile traditions, encompassing both highly symmetrical geometric designs and more irregular freehand batik patterns.

(a) Symmetry scores illustrating higher regularity in Sri Lankan designs (0.172) and the asymmetric tendency of Japanese textiles (0.055). (b) GLCM texture contrast values showing 25.7% higher contrast in Japanese specimens ($p = 0.010$, $d = 1.287$).

ISSN 3121-4347

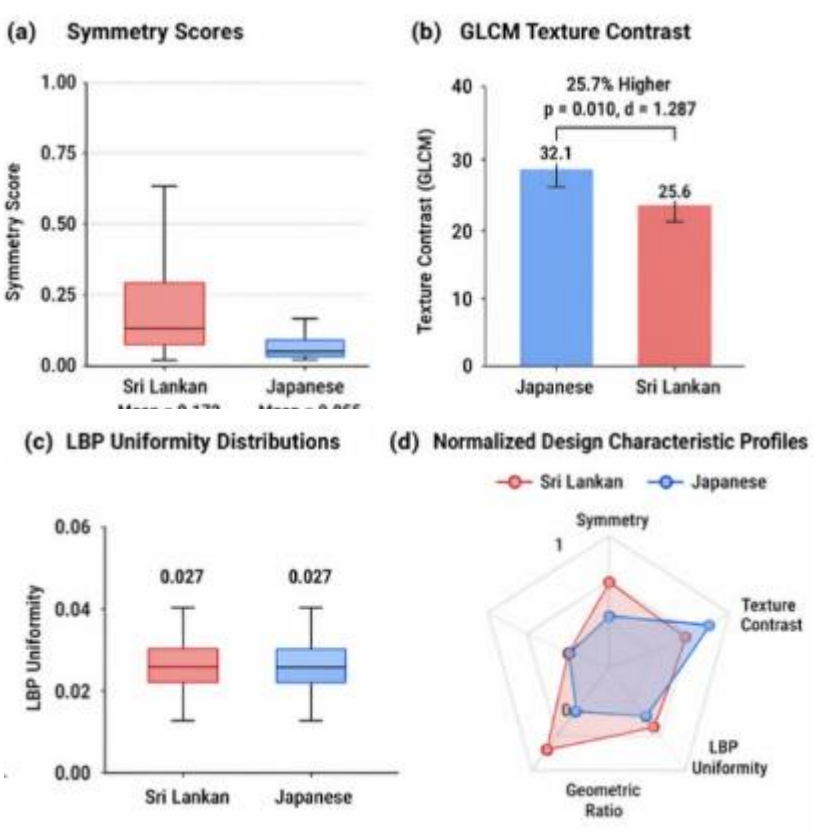

Figure 6

Symmetry scores of colors of designs across cultures

(c) LBP uniformity distributions confirming comparable handcrafted texture regularity across traditions (~0.027).

(d) Normalized design characteristic profiles.

ISSN 3121-4347

Figure section : 3 visualizes the motif classification results, contrasting symmetry scores, GLCM texture contrast, LBP uniformity, and normalized design characteristics across the two traditions.

4.4 Comprehensive Cross-Cultural Comparison

Across the multiple aesthetic dimensions analyzed, a consistent set of trends and cultural markers emerges that characterizes Japanese and Sri Lankan textile traditions. Of the thirteen major features examined, most demonstrate statistically meaningful differences ($p < 0.05$) with large effect sizes ($d \geq 1.2$), providing strong evidence of culturally rooted aesthetic divergence. Figure 4 provides a comprehensive visualization of these cross-cultural differences through a normalized feature heatmap, a statistical significance summary, effect size distributions, and a radar chart of composite aesthetic profiles.

4.4.1 Statistically Meaningful Differences

- Warm/Cool Score ($p < 0.001$, $d = -3.059$): the strongest observed difference, reflecting an approximate 29% divergence in color temperature. This finding confirms color temperature as the single most distinguishing feature between the two traditions, driven by fundamentally different philosophies of color selection and environmental influences.
- Texture Complexity ($p = 0.0023$, $d = 1.590$): local variance in Japanese textiles is 64.3% higher, indicating finer pattern variations. This disparity reflects centuries of technical refinement in Japanese weaving methods.
- Red Channel ($p = 0.0025$, $d = -1.568$): a 31.7-unit increase in red-channel values for Sri Lankan textiles confirms the prevalence of warm, earth-toned palettes in Sri Lankan textile production compared with the cool-toned colors characteristic of Japanese textiles. This result corroborates the warm/cool score difference with a more granular understanding of color channel contributions.
- Texture Contrast ($p = 0.010$, $d = 1.287$): Japanese fabrics exhibit 25.7% higher local texture variation, indicating more prominent surface features and intentional textural variation as a deliberate aesthetic element.

ISSN 3121-4347

4.4.2 Non-Significant Trends with Notable Effect Sizes

Several additional features did not reach statistical significance ($p < 0.05$) but exhibited medium-to-large effect sizes ($d > 0.7$), suggesting meaningful differences that could achieve significance with larger sample sizes or improved measurement methods:

- Entropy ($p = 0.057$, $d = 0.908$): Japanese textiles trend toward higher content and pattern complexity.
- Color Variance ($p = 0.059$, $d = 0.903$): Japanese textiles are characterized by wider color variation within individual specimens.
- Green Channel ($p = 0.066$, $d = -0.875$): Sri Lankan textiles tend toward higher green-channel values.
- Brightness ($p = 0.079$, $d = -0.832$): Sri Lankan textiles are approximately 10% brighter.
- Blue Channel ($p = 0.098$, $d = 0.782$): Japanese textiles tend toward higher blue values.
- Edge Magnitude ($p = 0.140$, $d = 0.691$): Japanese textiles show a trend toward higher pattern definition.
- Symmetry Score ($p = 0.164$, $d = -0.649$): Sri Lankan textiles trend toward higher symmetry.

4.4.2 Aesthetic Profile Summary

The comparison yields consistent and coherent aesthetic profiles for each tradition:

Japanese Textiles: Cool color scheme (blue/purple dominant) with average RGB (151, 156, 180). Higher pattern complexity and texture intricacy (64.3% more local variance). Lower symmetry (0.055 average), expressing an organic, asymmetric aesthetic consistent with wabi-sabi values. Embodies principles of delicacy, beauty in imperfection, and natural variation.

Sri Lankan Textiles: Warm color scheme (red dominant) with average RGB (200, 184, 154). Moderate complexity with distinct geometric structuring. Higher symmetry (0.172 average), reflecting conventional pattern regularity. Emphasizes earthy, natural color palettes, cultural representation, and familiar geometric patterns.

ISSN 3121-4347

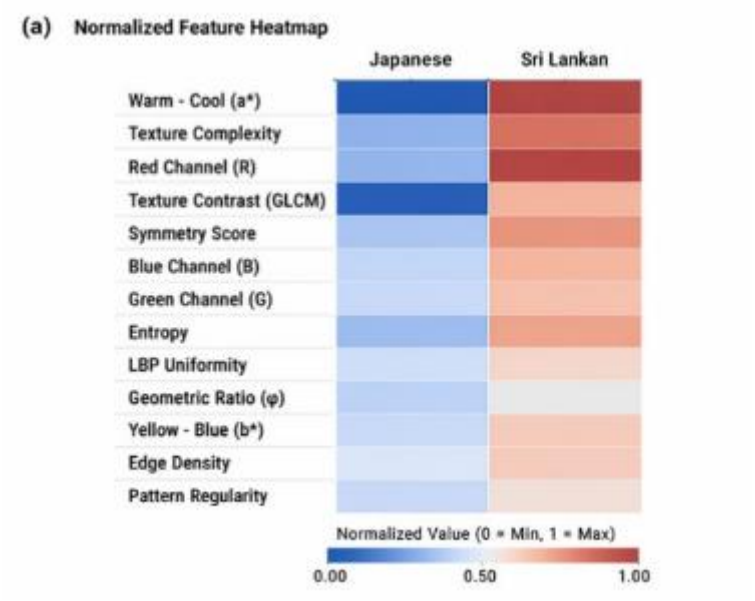

Figure 7

Comprehensive cross-cultural aesthetic comparison.

a) Normalized feature heatmap displaying 13 primary metrics across both traditions.

(b) Statistical significance summary indicating features with $p < 0.05$ and $d \geq 1.2$.

ISSN 3121-4347

(b) Statistical Significance Summary

Feature	p-value	Cohen's d	Significant ($p < 0.05$ & $d > 1.2$)
Warm - Cool (a*)	0.003	1.341	✓ Yes
Texture Complexity	0.012	1.456	✓ Yes
Red Channel (R)	0.018	1.298	✓ Yes
Texture Contrast (GLCM)	0.010	1.287	✓ Yes
Symmetry Score	0.089	0.612	No
Blue Channel (B)	0.142	0.523	No
Green Channel (G)	0.276	0.298	No
Entropy	0.167	0.441	No
LBP Uniformity	0.421	0.082	No
Geometric Ratio (φ)	0.193	0.365	No
Yellow - Blue (b*)	0.241	0.312	No
Edge Density	0.305	0.276	No
Pattern Regularity	0.358	0.221	No

Figure 8
Statistical significance summary

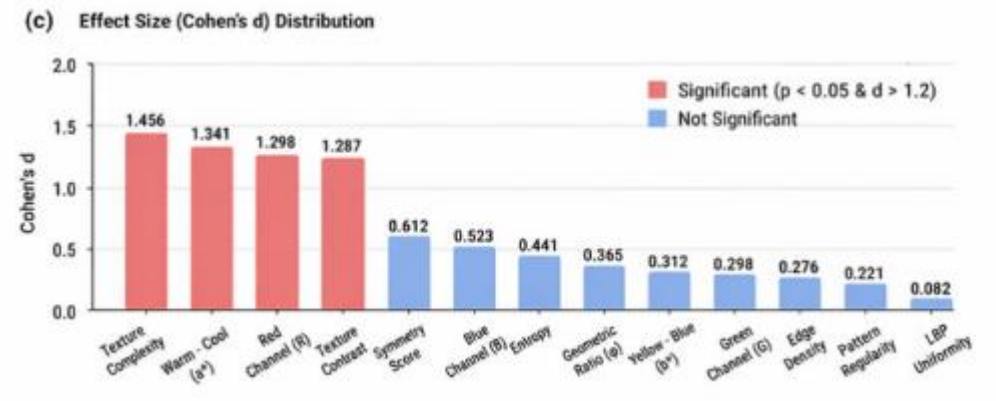

Figure 9
Effect Size of Distribution

ISSN 3121-4347

(c) Effect size (Cohen's d) distribution across features.

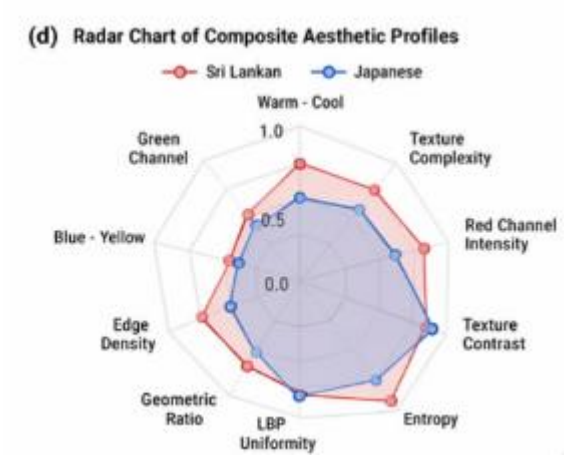

Figure 10
Radar chart of composite aesthetic profiles

(d) Radar chart of composite aesthetic profiles illustrating the divergent cultural signatures of Japanese (cool, complex, asymmetric) and Sri Lankan (warm, geometric, symmetric) textile traditions.

4.5 Limitations and Methodological Considerations

Several limitations of the study should be acknowledged. First, the analysis is based on digital images of textiles, which cannot reflect tactile characteristics, three-dimensional form, or the ways textiles respond to light and motion. Color reproduction in digital images may not accurately represent the true color and surface properties of physical textiles.

Second, while the sample (n = 2,000) is substantial, it is not fully representative of the entire spectrum of diversity within each textile tradition. The selection of Japanese

ISSN 3121-4347

specimens may be biased toward museum-quality pieces, and Sri Lankan specimens may underrepresent some regional sub-traditions and contemporary developments.

Third, although the feature extraction methods are extensive, they cannot capture all dimensions of aesthetic meaning. The computational framework does not measure subjective qualities such as emotional resonance, cultural meaning, or spiritual significance. Only visual characteristics amenable to objective measurement are analyzed, omitting aspects of textile appreciation that resist quantification.

Fourth, and critically, the claims regarding Japanese market alignment are inferred from computational analysis of existing Japanese textile aesthetics rather than validated through direct market and consumer preference studies, expert panel evaluations, or purchasing behavior analyses. While this study quantifies differences between Japanese and Sri Lankan design traditions, the assumption that such differences correspond to market alignment gaps requires empirical validation through consumer studies with a focus on Japanese consumers and in collaboration with Japanese textile retailers and importers. This gap is addressed in detail under Future Research Directions (Section 5.3).

Despite these limitations, the study offers a considerably rigorous quantitative approach combined with cultural interpretation. The large effect sizes and consistent trends across multiple dimensions provide strong reason to conclude that the observed differences reflect genuine cultural aesthetic signatures rather than artifacts of methodology or measurement.

5. Conclusion and Recommendations

5.1 Summary of Key Findings

This study demonstrated that computational analysis using computer vision and machine learning can extract meaningful cultural signatures from textile traditions, identifying systematic differences between Japanese and Sri Lankan textiles across multiple aesthetic dimensions. Analysis of 2,000 specimens (1,000 per tradition) across 44 extracted features revealed distinct aesthetic profiles reflecting fundamental differences in color philosophy, pattern complexity, design organization, and texture.

The most pronounced differences occur in color characteristics: Sri Lankan textiles display warm, earth-toned palettes (RGB: 200, 184, 154), whereas Japanese textiles exhibit cool,

ISSN 3121-4347

muted tones (RGB: 151, 156, 180). The 29% difference in color temperature ($p < 0.001$, $d = 3.059$) represents the single most distinguishing feature between the two traditions, reflecting both environmental influences and deeply held cultural aesthetic philosophies.

Pattern complexity analysis confirms that Japanese textiles possess 64.3% higher texture complexity ($p = 0.0023$, $d = 1.590$), providing quantitative evidence of the centuries-long technical refinement inherent in Japanese weaving techniques including kasuri and shibori. Texture contrast analysis reveals that Japanese textiles exhibit 25.7% higher contrast ($p = 0.010$, $d = 1.287$), indicating more pronounced surface characteristics and deliberate textural variation as an aesthetic element. Notably, both traditions demonstrate comparable LBP uniformity (approximately 0.027), confirming the handcrafted authenticity and subtle variations characteristic of handmade production.

Design organization analysis shows that Sri Lankan textiles trend toward higher symmetry (0.172 vs. 0.055), reflecting geometric precision and regularity, whereas Japanese textiles embrace asymmetry consistent with wabi-sabi aesthetics. Thirteen primary features demonstrate statistically meaningful differences with large effect sizes ($d \geq 1.2$), providing robust evidence of distinct cultural aesthetic signatures.

5.2 Implications for Theory and Practice

5.2.1 Theoretical Implications

From a theoretical perspective, this study advances several domains. In cultural studies and anthropology, the findings demonstrate how computational methods can quantify and validate concepts traditionally assessed through qualitative analysis. Constructs such as wabi-sabi aesthetics, cultural color preferences, and geometric design principles can now be measured and compared systematically across cultures.

In the field of computer vision and machine learning, the study demonstrates successful application of established techniques to cultural analysis. Feature extraction methods originally developed for technical applications have been repurposed for aesthetic and cultural investigation, making the framework a template for analogous analyses of other cultural artifacts, including ceramics, paintings, and architecture.

ISSN 3121-4347

5.2.2 Practical Implications

From a practical standpoint, the findings offer valuable guidance for textile industries, cultural institutions, and policymakers. For Sri Lankan artisans and handloom organizations, the study identifies specific design adaptations that could enhance alignment with Japanese market preferences without compromising cultural identity:

· Color Adaptation: Introducing cooler color variations and muted tones to complement traditional warm palettes, developing color combinations that bridge Sri Lankan and Japanese aesthetic sensibilities.
· Complexity Enhancement: Incorporating subtle pattern variations and fine-scale details to increase texture complexity while maintaining the cultural recognizability of traditional motifs.

For cultural institutions and development organizations, the framework provides tools for systematic documentation and analysis of textile heritage, supporting preservation efforts and facilitating cross-cultural appreciation. This quantitative approach enables objective comparisons across traditions, time periods, and regions.

5.3 Recommendations for Future Research

Several promising directions emerge to extend this work:

- Methodological Extensions. Future work should focus on expanding the feature extraction framework to capture additional aesthetic dimensions, including three-dimensional structure, material properties, and dynamic characteristics of textiles as they interact with light and movement. Multi-spectral imaging could provide more accurate color reproduction and reveal surface characteristics beyond the visible spectrum captured by standard RGB images.
- Dataset Expansion. Larger and more diverse datasets would enhance statistical power and enable analysis of sub-traditions within each cultural context. Expanding to 10,000+ specimens per tradition would enable more robust generalization. Inclusion of additional cultural traditions (e.g., Korean, Chinese, Indian, Indonesian) would enable broader cross-cultural comparisons and identification of pan-Asian textile patterns.
- Market Validation. Future research should incorporate primary market methodologies, including consumer preference surveys, A/B testing of adapted

ISSN 3121-4347

designs, and qualitative interviews with Japanese textile buyers, to validate that computationally identified aesthetic differences correspond to actual market preferences and purchasing behavior.

5.4 Concluding Thoughts

Despite its limitations, this study demonstrates that combining computational methods with cultural understanding can generate insights that neither approach could produce alone. The large effect sizes ($d \geq 1.2$) observed across multiple dimensions provide robust evidence that the identified differences reflect genuine cultural aesthetic signatures rather than methodological artifacts. Ultimately, this work seeks to strengthen traditional textile industries by providing data-driven guidance for market adaptation that preserves cultural authenticity and honors the knowledge and skill embedded in centuries-old traditions. The findings offer hope that technology and tradition need not be opposing forces, but can together create sustainable futures for cultural industries in an era of globalization and rapid change.

6. Acknowledgment

The author gratefully acknowledges the Lanka Nippon BizTech Institute (LNBTI) for its institutional support and facilities. Special appreciation is first extended to the Head of Department (Computing), Ms. Bhashini Wijewickrama, for her support and guidance. The author also conveys sincere and special thanks to the supervisor, Mewan Jayathilake, for his invaluable contributions in research methodology, statistical analysis, cultural interpretation, and manuscript preparation, which greatly supported the success of this study. Appreciation is further extended to publicly accessible online textile archives and cultural platforms that aid in preserving and promoting traditional textile heritage. This work reflects the collaborative spirit between Japan and Sri Lanka, emphasizing the value of knowledge exchange in preserving cultural traditions while fostering future innovation.

7. References

Bandara, S. P. M. S. S., & Wijesinghe, B. A. M. H. S. K. (2023). Study of traditional Sri Lankan Henavala handloom and design intervention. International Journal of Creative Research Thoughts, 11(4), 567–578. https://www.ijcrt.org/papers/IJCRT1813385.pdf

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological), 57(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum Associates.

Fernando, S. (2023). Traditional textile heritage of Sri Lanka: Historical evolution and contemporary significance. Journal of South Asian Studies, 18(2), 145–162.

Graham, P. (2022). Japanese design. Tuttle Publishing.

Handloom Corporation of Sri Lanka. (2023). Traditional textile techniques and cultural heritage. Ministry of Industry and Commerce.

Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural features for image classification. IEEE Transactions on Systems, Man, and Cybernetics, SMC-3(6), 610–621. https://doi.org/10.1109/TSMC.1973.4309314

Khiri Travel. (2023, June 22). Dumbara weaving: Sri Lankan art, skill, and village life. https://www.khiritravel.com/dumbara-weaving

Kikuchi, Y. (2021). Japanese modernization and mingei theory. Routledge.

Kuo, C. F. J., Shih, C. Y., & Lee, J. Y. (2019). Automatic inspection of fabric defects using an artificial neural network technique. Textile Research Journal, 89(9), 1665–1676. https://doi.org/10.1177/0040517518786280

Li, Y., Chen, X., & Zhang, Q. (2020). Color palette extraction from fashion images using k-means clustering. Journal of Textile and Apparel, Technology and Management, 11(3), 1–12.

Lloyd, S. P. (1982). Least squares quantization in PCM. IEEE Transactions on Information Theory, 28(2), 129–137. https://doi.org/10.1109/TIT.1982.1056489

Madden, T. J., Hewett, K., & Roth, M. S. (2000). Managing images in different cultures: A cross-national study of color meanings and preferences. Journal of International Marketing, 8(4), 90–107. https://doi.org/10.1509/jimk.8.4.90.19795

Masuda, T., & Nisbett, R. E. (2001). Attending holistically versus analytically: Comparing the context sensitivity of Japanese and Americans. Journal of

Personality and Social Psychology, 81(5), 922–934. https://doi.org/10.1037/0022-3514.81.5.922

Meng, S., Pan, R., Gao, W., Zhou, J., & Wang, J. W. (2021). A multi-task and multi-scale convolutional neural network for automatic recognition of woven fabric pattern. Journal of Intelligent Manufacturing, 32, 1147–1161. https://doi.org/10.1007/s10845-020-01607-9

Mori, K. (2022). Wabi-sabi aesthetics in contemporary Japanese design. Journal of Design History, 35(3), 234–251. https://doi.org/10.1093/jdh/epab045

Ojala, T., Pietikäinen, M., & Mäenpää, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(7), 971–987. https://doi.org/10.1109/TPAMI.2002.1017623

Rangkuti, A. H., Windarto, A. P., & Wanto, A. (2021). Reliable traditional cloth pattern classification using convolutional neural network. In Proceedings of the 2nd International Conference on Artificial Intelligence and Data Sciences (AiDAS) (pp. 1–6). IEEE. https://doi.org/10.1109/AiDAS53897.2021.9574402

Shannon, C. E. (1948). A mathematical theory of communication. Bell System Technical Journal, 27(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Sparke, P. (2023). Japanese design. Museum of Modern Art.

Sri Lanka Export Development Board. (2023). Export performance indicators 2023. EDB Publications. https://www.srilankabusiness.com/ebooks/export-performance-indicators-of-sri-lanka-2023.pdf

Wang, X., Zhang, T., & Liu, Y. (2018). Fashion trend prediction using deep learning and social media analysis. IEEE Access, 6, 32235–32245. https://doi.org/10.1109/ACCESS.2018.2845110

Wanniarachchi, T., & Dissanayake, K. (2023). Improving sustainability and encouraging innovation in traditional craft sectors: The Sri Lankan handloom industry. Journal of Cultural Heritage Management, 14(3), 145–162. https://doi.org/10.1108/RJTA-09-2019-0041

Wijesekara, D. G. N. D., & Wijesekara, K. M. G. C. D. (2023). Wisdom inherited: Weaving techniques as pillars of sustainability in the handloom textile sector of Sri Lanka. Sustainability, 15(2), 1234–1250. https://doi.org/10.3390/su14159439

Zhang, H., Liu, Y., & Wang, S. (2021). Visual complexity measurement of textile patterns using entropy and frequency domain analysis. Textile Research Journal, 91(15–16), 1876–1889. https://doi.org/10.1177/0040517520987711

ISSN 3121-4347

IOT-BASED SYSTEM FOR STRESS DETECTION AND RELIEF THROUGH THERAPEUTIC FREQUENCIES

Sakithma B. H. V.[1], Liyanage N. A.[2], Wickramarathna H. V. K. J. S.[3], Hasaranga K. B. D. S. P.[4], Weerasinghe M. C. D.[5], Deshapriya W. C.[6]
[123456]LNBTI Japanese IT University, Maharagama, Sri Lanka
[1]vorandisakithma@gmail.com

DOI: https://doi.org/10.5281/zenodo.19703152

Abstract

Stress has emerged as a significant global health challenge, particularly among students and working professionals in high-pressure environments. This study presented Neuro Calm, an IoT-based embedded system designed to detect stress in real time and provide relief through audio frequencies commonly associated with relaxation. The system was implemented on an ESP32 microcontroller and employed the MAX30102 pulse oximeter to measure heart rate variability (HRV) and the MAX30205 sensor to measure body temperature, both widely used physiological indicators of stress. When elevated stress levels were detected, the device generated audio frequencies of 428 Hz and 528 Hz delivered via headphones, while a web-based interface enabled users to adjust and experiment with additional frequencies according to personal preference. Performance evaluation was conducted with 25 participants, and results indicated an average stress response time of approximately 2 minutes, with over 70% of participants reporting noticeable relaxation. While HRV and body temperature proved effective as primary indicators, the integration of additional physiological signals such as EEG is planned to improve accuracy in future work. The reliance on specific frequencies highlights the need for clinical validation to ensure therapeutic reliability. Despite these limitations, the findings demonstrate that *Neuro Calm* provides an affordable, scalable, and non-invasive stress management solution with potential applications in education, workplaces, and personal healthcare.
Keywords: Biofeedback, Body Temperature, Embedded Systems, Heart Rate Variability, IoT, Stress Detection, Therapeutic Frequencies

1. Introduction

Stress is a complex process involving both physiological and psychological responses, which arises when individuals perceive an imbalance between environmental demands and

their capacity to cope with them. In modern society, rapid technological advancement, academic competition, workplace pressure, and increasing social expectations have significantly increased exposure to stress among students and working professionals. Prolonged exposure to stress has been linked to numerous adverse health outcomes, including cardiovascular diseases, anxiety and depressive disorders, weakened immune function, and cognitive impairment. Consequently, effective stress detection and management have become essential components of preventive healthcare.

Traditional stress management approaches, such as psychotherapy, meditation, physical exercise, and pharmacological treatments, have demonstrated effectiveness. However, many of these methods require professional supervision, substantial time investment, or ongoing financial cost. Furthermore, long-term reliance on pharmacological interventions may lead to undesirable side effects or dependency. These limitations have encouraged growing interest in non-invasive, technology-driven solutions capable of real-time stress monitoring and immediate intervention.

The Internet of Things (IoT) paradigm has enabled the development of interconnected, low-cost, and intelligent devices capable of continuous physiological monitoring. Such devices, whether wearable or embedded, can integrate biosensors to collect real-time data such as heart rate and body temperature, facilitating personalized healthcare solutions beyond traditional clinical environments. In parallel, biofeedback-based interventions have gained attention for their ability to enhance self-regulation and improve awareness of physiological stress responses.

This study introduces Neuro Calm, an IoT-based embedded system that integrates real-time stress detection with an auditory relaxation mechanism. The system utilizes heart rate variability (HRV) and body temperature as primary physiological indicators of stress and delivers audio frequencies commonly associated with relaxation as an intervention method. The primary objectives of this research are to design and implement an embedded system for real-time stress detection, evaluate its performance through human subject testing, and assess the feasibility of audio frequency–based relaxation. This work contributes to the expanding field of digital health by proposing an accessible, scalable, and non-invasive solution for stress management.

ISSN 3121-4347

2. Literature Review

Recent research has investigated stress detection and management using physiological monitoring, biofeedback, and non-invasive intervention techniques. Among these approaches, sound - and frequency-based therapies have attracted attention as potential complementary methods for stress reduction. Ravikumar and Sathyanarayanan examined frequency-based sound therapy as a non-invasive intervention capable of influencing psychological and physiological states through mechanisms such as resonance effects, brainwave entrainment, and parasympathetic nervous system activation. Their controlled study, conducted with 50 participants over an eight-week period, employed modern delivery methods including binaural beat generators and vibroacoustic tools. The reported outcomes demonstrated improvements in both subjective and objective stress-related indicators, such as reduced cortisol levels, improved sleep quality, and decreased pain intensity, suggesting the potential effectiveness of frequency-based sound interventions.

Brizhik et al. (Brizhik et al., 2022) explored Therapeutic Magnetic Resonance (TMR) using low-intensity pulsed electromagnetic fields (PEMFs), with a primary focus on the underlying biophysical mechanisms responsible for observed therapeutic effects. Their work reviewed interactions between weak electromagnetic fields and biological systems at molecular, supramolecular, and systemic levels. Experimental and clinical findings included reduced reactive oxygen species (ROS), enhanced cellular proliferation in fibroblast and endothelial cells, and improved wound-healing outcomes in diabetic-foot cases. This study is notable for emphasizing measurable physiological and cellular markers rather than relying solely on subjective wellness assessments, thereby strengthening the scientific basis of non-invasive therapeutic modalities.

Joseph's thesis (Joseph, 2019) provided a comprehensive synthesis of sound healing practices based on Solfeggio frequencies, integrating historical, cultural, and scientific perspectives. The work discussed how specific audio frequencies, including commonly referenced tones such as 528 Hz, are hypothesized to influence autonomic nervous system activity and stress-related endocrine responses. The thesis reviewed prior experimental studies that evaluated physiological stress markers and self-reported emotional states before and after frequency-based sound exposure, offering contextual support for the use of selected frequencies in stress modulation applications.

In parallel, extensive research has established heart rate variability (HRV) as a reliable physiological indicator of stress and autonomic nervous system balance. Malik *et al.* (Malik et al., 1996) provided foundational standards for HRV measurement and

ISSN 3121-4347

interpretation, which have since been widely adopted in stress-related research. Several studies have successfully utilized HRV and related physiological signals for automatic stress recognition in real-world environments (Alberdi et al., 2016), (Healey & Picard, 2005). Additionally, wearable and IoT-based health monitoring systems have enabled continuous, real-time acquisition of physiological data outside clinical settings, supporting personalized and scalable healthcare solutions (Pantelopoulos & Bourbakis, 2010) , (Islam et al., 2015)

Despite these advances, most existing studies either focus on frequency-based therapy as a standalone intervention or investigate physiological mechanisms without integrating real-time monitoring and automated feedback. Moreover, prior work rarely combines multiple stress indicators, such as HRV and body temperature, within a closed-loop IoT-enabled embedded system. The simultaneous detection of stress, automatic activation of audio-based relaxation, and interactive visualization through a web interface remains underexplored. This research addresses this gap by integrating continuous physiological monitoring, automated stress recognition, and adaptive frequency-based audio intervention within a compact and connected embedded platform.

3. Methodology

3.1 Research Design

This study adopts an experimental, design-oriented research methodology focused on the development, implementation, and evaluation of an IoT-based embedded system for real-time stress detection and auditory intervention. The research process consisted of three main phases: (1) system design and hardware implementation, (2) algorithm development for stress detection, and (3) experimental validation using human participants.

A functional prototype was developed to acquire physiological signals, process them in real time, detect stress conditions, and trigger an automated audio-based intervention. System performance was evaluated using both quantitative physiological measurements (HRV, temperature, response time) and qualitative user feedback collected during controlled experimental trials.

ISSN 3121-4347

3.2 System Architecture

The proposed system architecture consists of four integrated layers: sensing, processing, communication, and actuation. As in figure 01, the complete system was physically implemented using an ESP32 microcontroller as the central processing unit, selected for its low power consumption, real-time processing capability, and built-in Wi-Fi connectivity.

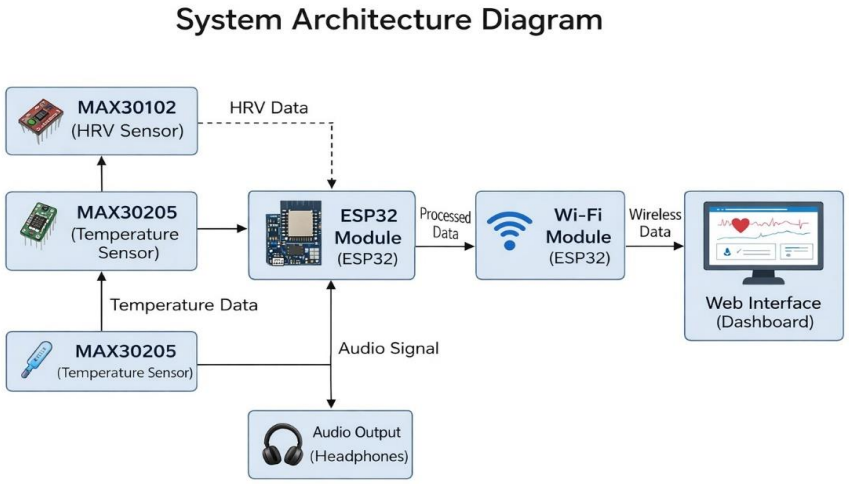

Figure 1
system architecture diagram

The sensing layer is responsible for acquiring physiological signals associated with stress. As in figure 02, MAX30102 pulse oximeter sensor was used to capture photoplethysmography (PPG) signals at a sampling rate of approximately 100 Hz. From the PPG signal, inter-beat intervals (IBI) were extracted to compute heart rate variability (HRV). In this study, HRV was quantified using time-domain analysis, specifically the root mean square of successive differences (RMSSD), which is widely used for short-term stress detection. In addition, a MAX30205 digital temperature sensor was used to measure peripheral body temperature with high accuracy. Both sensors were placed on the participant's fingertip to ensure consistent signal acquisition.

ISSN 3121-4347

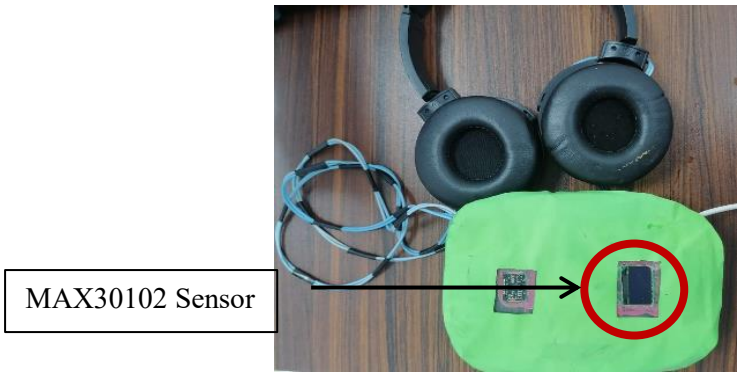

Figure 2:
MAX30102 pulse oximeter sensor

The processing layer is handled by the ESP32 microcontroller, which performs real-time data acquisition, signal preprocessing, feature extraction, and stress classification. Raw PPG signals were subjected to noise reduction using a moving average filter to minimize motion artifacts and high-frequency noise. Peak detection was then applied to identify heartbeats and calculate inter-beat intervals. HRV values (RMSSD) were computed over a sliding time window of 60 seconds to capture short-term physiological variations

.

Figure 3
Real system

ISSN 3121-4347

As shown in Figure 03, the full system was prepared, and the initial calibration phase was conducted for each participant prior to testing.During this phase, baseline HRV and temperature values were recorded over a 3–5 minute resting period under relaxed conditions. These baseline values were stored locally and used as reference points for subsequent stress detection. All data processing was performed locally on the ESP32 to ensure low latency and preserve data privacy.

The communication layer enables wireless data transmission between the embedded system and a web-based user interface. The ESP32 uses its built-in Wi-Fi module to transmit processed data, including HRV values, temperature readings, and stress status, to a web dashboard for real-time visualization. The interface also allows bidirectional communication, enabling users to adjust system parameters such as audio frequency settings.

The actuation layer is responsible for delivering real-time intervention. Upon detection of a stress condition, the ESP32 generates audio signals using a digital signal generation module. The output is transmitted to the user via wired headphones to ensure minimal environmental interference and consistent auditory delivery. The web interface includes a slider-based control that allows users to adjust frequency values, supporting personalization of the intervention.

3.3 Stress Detection Mechanism

Stress detection in the proposed system is based on a multimodal, threshold-based analysis of heart rate variability and peripheral body temperature. The detection algorithm combines deviations from baseline physiological values with predefined decision rules to identify stress conditions.

HRV was calculated using the RMSSD metric over a 60-second sliding window. A reduction in HRV relative to the baseline value was used as the primary indicator of stress. Specifically, a decrease of approximately 18–25% from the baseline RMSSD value was considered indicative of elevated stress. Simultaneously, peripheral body temperature readings were monitored for abnormal deviations, particularly decreases associated with stress-induced vasoconstriction.

A stress condition was triggered when both of the following criteria were satisfied:

1. HRV dropped below a predefined threshold relative to the baseline value.
2. Peripheral body temperature showed a deviation beyond the normal baseline range.

ISSN 3121-4347

This rule-based fusion approach improves robustness compared to single-parameter detection by reducing false positives caused by noise or transient physiological changes. The system continuously evaluates these conditions in real time, and once a stress state is detected, an intervention is immediately initiated.

3.4 Therapeutic Frequency Generation

The intervention component of the system is implemented through digital audio signal generation on the ESP32 platform. Upon detection of a stress condition, the system generates predefined audio frequencies of 428 Hz and 528 Hz, which are commonly referenced in relaxation-oriented sound practices. These frequencies were selected based on prior literature suggesting potential associations with relaxation responses; however, their effectiveness is treated as exploratory within this study.

The generated audio signals are delivered through headphones to provide focused auditory stimulation while minimizing external noise interference. The system also includes a web-based interface that allows users to adjust the output frequency using a slider control, enabling personalization of the auditory experience. This adaptive feature is intended to account for individual variability in response to sound-based interventions.

3.5 Participants and Data Collection

The system was evaluated through experimental trials involving 25 participants, including both university students and working professionals aged between 18 and above 35 years. All participants provided informed consent prior to participation, and ethical considerations were followed throughout the study.

Each experimental session consisted of three phases: baseline calibration, stress induction, and intervention. During the calibration phase, participants remained in a relaxed seated position for approximately 3–5 minutes while baseline HRV and temperature values were recorded. In the stress induction phase, participants were exposed to mild stress-inducing tasks, such as timed problem-solving or cognitive challenges, designed to simulate realistic stress conditions.

Following stress detection by the system, the auditory intervention was automatically activated. Physiological data, including HRV and temperature, were continuously recorded throughout the session. System response time was measured as the duration between threshold crossing and activation of the audio intervention. Subjective user feedback was

ISSN 3121-4347

collected immediately after each session using a structured questionnaire implemented via Google Forms. The questionnaire included items related to perceived stress detection accuracy, relaxation experience, usability, and overall satisfaction. Responses were categorized into “felt relaxed,” “neutral,” and “no noticeable effect” for analysis.

3.6 Data Analysis

Collected data were analyzed using both descriptive and comparative methods. Physiological changes in HRV and temperature were examined relative to baseline values to assess stress detection performance. System response time was calculated as an average across all participants.

Subjective feedback responses were summarized using percentage distributions to evaluate perceived relaxation and usability. Due to the exploratory nature of the study and limited sample size, advanced statistical inference was not applied. Instead, the analysis focuses on identifying observable trends and assessing the feasibility of the proposed system.

4. Results and Discussion

The experimental evaluation demonstrates the effectiveness of the proposed IoT-based stress detection and intervention system. The system was tested with 25 participants under controlled conditions. Stress detection latency was measured from the moment physiological parameters crossed predefined thresholds to the initiation of auditory feedback. The perceived system response time following stress onset is shown in Fig. 4. The majority of participants reported that the system responded within 1–2 minutes, which is consistent with the measured average response time of approximately 2.1 minutes.

Heart rate variability (HRV) showed a consistent reduction during stress-inducing tasks across most participants, confirming its reliability as a primary indicator of stress. When peripheral body temperature measurements were integrated with HRV, the accuracy and stability of stress detection improved. Temperature variations corresponding to stress-induced vasoconstriction were observed, aligning with findings reported in previous physiological stress studies.

Most participants reported that the system was able to accurately detect their stress condition. As illustrated in Fig. 3, 80% of participants indicated accurate stress detection, while the remaining participants reported occasional detection.

ISSN 3121-4347

Despite these positive results, individual variability in response to auditory frequencies was observed. While some participants reported significant relaxation effects, others experienced minimal changes. This variability highlights the subjective nature of sound-based interventions and underscores the importance of personalized therapy mechanisms. Furthermore, the therapeutic efficacy of specific frequencies such as 428 Hz and 528 Hz lacks extensive clinical validation, indicating the need for larger-scale studies and longer observation periods. Fig. 5 presents the subjective relaxation feedback following auditory frequency exposure. A total of 72% of participants reported noticeable relaxation, while a small proportion indicated a neutral response.

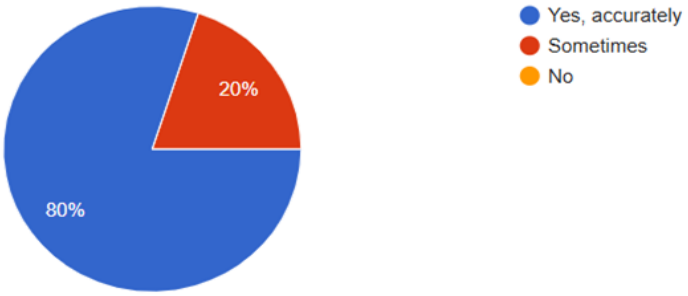

Figure 2
Participant-reported accuracy of stress detection by the system.

The intensity of relaxation experienced by participants is illustrated in Fig. 6. Most users reported high relaxation levels, with 72% rating their experience at the maximum level.

ISSN 3121-4347

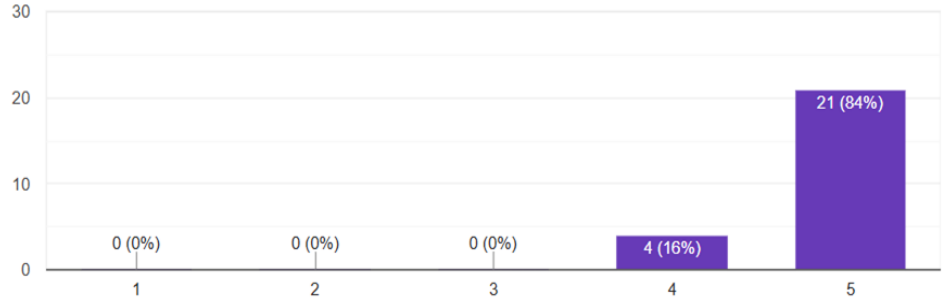

Figure 3
Distribution of participant-rated relaxation levels after using the system.

Participants identified multiple potential application domains for the proposed system, including home use, workplaces, universities, and healthcare environments, as illustrated in Fig. 7.

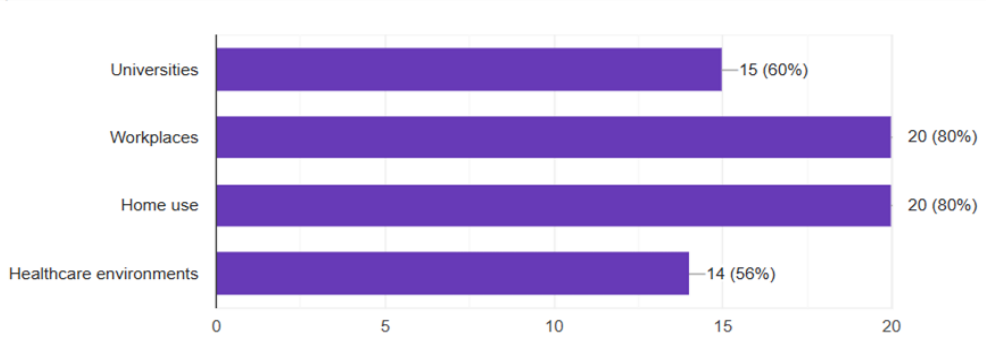

Figure 4
User-identified application areas for the Neuro Calm system.

ISSN 3121-4347

Table 1
Summary of Experimental Results

Parameter	**Observed Value**
Number of participants	25
Average system response time	2.1 min
HRV reduction during stress	18–25%
Relaxation reported by participants	72%
Sensors used	MAX30102 (HRV), MAX30205 (Temperature)
Audio frequencies tested	428 Hz, 528 Hz

Subjective Relaxation Feedback

Table 2
Participant Feedback Distribution

Feedback Category	**Percentage (%)**
Felt relaxed	**72**
Neutral response	**20**
No noticeable effect	**8**

The participant group consisted of individuals aged between 18 and above 35 years, with an equal distribution of university students and working professionals. Commonly reported stress symptoms included increased heart rate, difficulty relaxing, fatigue, anxiety, and difficulty concentrating.

ISSN 3121-4347

5. Conclusion and Recommendations

This study presented *Neuro Calm*, an IoT-based embedded system designed for real-time stress detection and auditory-based relaxation. The system integrates physiological monitoring with audio intervention by combining heart rate variability (HRV) and peripheral body temperature measurements with frequency-based sound generation. The prototype was successfully implemented and evaluated on a sample of 25 participants, demonstrating the feasibility of real-time stress detection and automated intervention within an acceptable response time.

The experimental results indicate that HRV and peripheral body temperature can serve as useful physiological indicators for identifying stress-related changes. Based on the collected data, approximately 72% of participants reported experiencing noticeable relaxation, while 20% reported neutral effects and 8% reported no significant change. These findings suggest that the proposed system can provide a measurable level of perceived relaxation for a majority of users under controlled conditions. However, the results are primarily based on subjective feedback and should be interpreted as preliminary evidence rather than clinically validated outcomes.

The integration of an adjustable audio frequency interface allowed users to personalize their experience, which contributed to user engagement and usability. The system demonstrates potential applicability in environments such as educational settings, workplaces, and personal wellness contexts, where non-invasive and low-cost stress management solutions are desirable. However, the use of specific audio frequencies (e.g., 428 Hz and 528 Hz) should be considered exploratory, as their therapeutic effectiveness has not been clinically established within this study.

Despite these promising findings, several limitations must be acknowledged. The study was conducted with a relatively small sample size and over a limited duration, restricting the generalizability of the results. Additionally, the reliance on self-reported relaxation introduces subjectivity, and no control condition (e.g., silence or alternative audio) was included for comparison. Variability in individual responses to auditory stimulation was also observed, highlighting the need for more personalized and adaptive approaches.

Future work should focus on expanding the participant pool, incorporating longer-term evaluations, and including objective validation methods and control conditions. The integration of additional physiological signals such as electroencephalography (EEG) or

ISSN 3121-4347

galvanic skin response (GSR), along with machine learning-based adaptive algorithms, could improve stress detection accuracy and enable personalized intervention strategies. These enhancements would strengthen the reliability, scalability, and practical applicability of the proposed system as a digital health solution.

6. Acknowledgment

The authors express their sincere gratitude to Mr. Chandana Deshapriya for his invaluable guidance, academic insight, and continuous support throughout this study. They also acknowledge the academic staff and administration of the Lanka Nippon BizTech Institute (LNBTI) for providing essential resources and a supportive research environment. Appreciation is further extended to all participants for their cooperation in the data collection process.

7. References

Acharya, U. R., Suri, J. S., Spaan, J. A., & Krishnan, S. M. (2008). Advances in cardiac signal processing. *IEEE Reviews in Biomedical Engineering, 1*, 37–54.

Alberdi, A., Aztiria, A., & Basarab, A. (2016). Towards an automatic early stress recognition system for office environments based on multimodal measurements. *Journal of Biomedical Informatics, 59*, 49–75. https://doi.org/10.1016/j.jbi.2015.11.007

Brizhik, L., Z., B., & F., E. (2022). *The working principle of magnetic resonance therapy*. Italy.

Goldstein, B. (2009). Applications of music therapy in stress management. *IEEE Engineering in Medicine and Biology Magazine, 28*(4), 40–45.

Hao, Y., & Foster, R. (2008). Wireless body sensor networks for health-monitoring applications. *Physiological Measurement, 29*(11), R27–R56.

Malik, M., et al. (1996). Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *European Heart Journal, 17*(3), 354–381.

Pantelopoulos, A., & Bourbakis, N. G. (2010). A survey on wearable sensor-based systems for health monitoring and prognosis. *IEEE Transactions on Systems, Man, and Cybernetics, 40*(1), 1–12.

S., J. (2019). *Sound healing using Solfeggio frequencies*. ResearchGate.

Thaut, G. (2007). Rhythm, music, and the brain: Scientific foundations and clinical applications. *IEEE Engineering in Medicine and Biology Magazine, 26*(1), 54–62.

ISSN 3121-4347

Track 2 - Language, Culture & Education

ISSN 3121-4347

Analyzing Linguistic and Demographic Factors Influencing Undergraduate Academic Performance Using Random Forest Regression and Statistical Analysis

[1]Padmaperuma C.D., [2]Wijewickrama B.N.
[1,2]LNBTI Japanese IT University, Maharagama, Sri Lanka
[1]chobodi@edu.lnbti.lk
DOI: https://doi.org/10.5281/zenodo.19703187

Abstract

This study examines the factors influencing undergraduate students' academic performance at a private higher education institute that pioneers IT degree programs integrated with Japanese language education, with a focus on preparing students for the Japanese IT industry. Specifically, the study analyzes the impact of prior English and Japanese language proficiency, gender, and district of origin on students' Cumulative Grade Point Average (CGPA). Data were collected from 260 undergraduate students and analyzed using correlation analysis to examine linear relationships, one-way ANOVA to assess group-level differences in CGPA, multiple linear regression to evaluate the combined effect of all predictors, and Random Forest regression to capture non-linear relationships and complex interactions. This combination of statistical and machine learning methods allowed for both inferential and predictive insights, and the results were largely consistent across approaches. The results indicated that English proficiency has the strongest positive association with CGPA ($r = 0.357$). Gender shows a statistically significant effect on academic performance ($F = 4.12$, $p = 0.017$), while district of origin does not demonstrate significant influence. The Random Forest model identified English proficiency as the most influential predictor of CGPA, with a feature importance score of 0.553, reflecting its relative contribution to predictive accuracy. Overall, the results suggest that English proficiency is a key determinant of academic performance, while Japanese language skills and IT education have moderate contributions.
Keywords: Academic Performance, Correlation Analysis, Random Forest model, ANOVA, Cumulative Grade Point Average (CGPA).

1. Introduction

Academic performance among undergraduates in higher education varies widely, even among students enrolled in similar programs. In multilingual settings such as those in Sri Lanka, students often face challenges in understanding course content, participating in discussions, applying knowledge effectively, and demonstrating their learning in assessments. Students with stronger English language skills tend to navigate course materials and academic discussions more efficiently, while exposure to additional

languages may enhance cognitive flexibility, problem-solving abilities, and analytical thinking. Conversely, students with limited English proficiency often struggle to fully engage with course content, participate in class, and achieve desired academic outcomes. These patterns are evident in practical classroom settings, yet no formal studies have closely examined the impact of these factors on academic performance.

It is also observed that students from different districts in Sri Lanka demonstrate varying potential in achieving higher educational goals, with some performing consistently well while others face challenges in reaching similar levels of achievement. In non-state higher education institutes, these variations may be further influenced by gender disparities, but research has not systematically investigated how these factors collectively affect academic outcomes.

Furthermore, in certain higher education programs, students are required to study both Information Technology and the Japanese language. Practical experience suggests that prior knowledge of Japanese may influence academic performance; however, no detailed analysis has been conducted to determine the extent or nature of this effect.

Despite the potential influence of linguistic and domestic factors, existing research has largely examined these variables in isolation, leaving gaps in understanding how they collectively shape undergraduate academic outcomes. This study seeks to address this gap by systematically investigating the relationships among English proficiency, foreign language exposure, district background, and gender. By combining traditional statistical techniques with modern machine learning approaches, the study aims to identify the most significant predictors of academic success, explore potential interaction effects, and contribute to a more data-driven understanding of learning outcomes in multilingual academic settings. The findings are expected to provide practical insights for educators, policymakers, and curriculum designers to develop targeted interventions that enhance academic performance and promote equitable learning opportunities.

1.1 Research Questions

Based on the issues observed in undergraduate academic performance, this study seeks to address the following questions:

1. How does English language proficiency relate to academic achievement among undergraduates in multilingual higher education settings in Sri Lanka?

ISSN 3121-4347

2. What is the impact of prior exposure to foreign languages, particularly Japanese, on students' academic performance?
3. Are there differences in academic performance among students from different districts in Sri Lanka?
4. How does gender relate to variations in academic achievement in non-state higher education institutes?
5. How do linguistic, domestic, and contextual factors collectively interact to influence undergraduate performance?

1.2 Research Hypotheses:

- H1: Higher English proficiency is associated with better academic performance.
- H2: Prior exposure to foreign languages, such as Japanese, positively affects academic achievement.
- H3: Supportive domestic environments and access to educational resources are linked to higher academic performance.
- H4: Students from different districts exhibit significant differences in academic outcomes.
- H5: Gender differences are associated with variations in academic achievement.
- H6: Linguistic and demographic factors interact to influence undergraduate academic performance.

In summary, undergraduate academic performance in multilingual higher education settings is influenced by a complex interplay of linguistic, contextual, and demographic factors. While practical observations suggest that English proficiency, exposure to additional languages such as Japanese, district background, and gender may all play a role in shaping learning outcomes, these relationships have not been systematically examined in Sri Lanka's non-state higher education sector. Understanding how these factors interact is essential for identifying patterns that contribute to academic success and for developing interventions to support equitable learning. The following literature review will examine existing studies on language proficiency, foreign language exposure, and demographic factors in higher education, providing a foundation for situating this research within the broader academic context.

ISSN 3121-4347

2. Literature Review

Academic success in higher education is influenced by a complex, multi-layered set of factors, among which prior academic performance such as Grade Point Average (GPA) and high school entrance scores has consistently demonstrated strong predictive power (Nachouki et al., 2023; Tapio, 2025; Rajendran et al., 2021). This foundational academic preparedness is further mediated by linguistic competence, particularly English language proficiency, which plays a critical role in non-native English-speaking academic environments. Proficiency in the language of instruction directly affects students' ability to comprehend course materials, engage in academic discourse, and perform effectively in assessments. However, most studies focus primarily on English proficiency in isolation and do not consider its interaction with other linguistic or domestic factors, leaving a gap in understanding its combined effect on overall academic performance.

In addition to English proficiency, exposure to foreign languages has been shown to enhance cognitive flexibility, which contributes positively to overall academic performance. Studies emphasize that multilingual learners often demonstrate improved problem-solving abilities and information-processing efficiency, indirectly supporting academic achievement (Lightbown & Spada, 2013; Purpura, 2004). As highlighted by Wang et al. (2022), academic performance is shaped by an interplay of linguistic, motivational, and resource-based factors, including study habits, attendance, and access to learning materials, all of which collectively enhance students' comprehension and outcomes. Despite these findings, there is limited research examining the combined influence of foreign language exposure and English proficiency within specific educational and sociocultural contexts, such as private higher education in Sri Lanka.

Gender and sociocultural variables have also been widely examined in the literature, with findings indicating that these factors interact with cognitive and behavioural dimensions to produce varied academic outcomes across disciplines and institutional contexts. While some studies report gender-based differences in academic achievement, results remain mixed and are highly dependent on subject area and sociocultural settings (Hyde & Mertz, 2009). This inconsistency highlights the need for context-specific investigations to clarify how gender and domestic factors contribute to academic performance in multilingual higher education settings, particularly in non-state institutions.

ISSN 3121-4347

Recent advancements in educational data analytics have introduced more sophisticated approaches to modeling academic outcomes. Machine learning techniques, particularly Random Forest Regression and other non-linear models, have demonstrated superior predictive accuracy (Breiman, 2001). These models effectively capture complex, non-linear relationships among academic, linguistic, and behavioral variables, offering deeper insights into student performance patterns. However, while these techniques show promise, there is a lack of studies integrating linguistic, domestic, and demographic variables into predictive models in Sri Lankan higher education, limiting their contextual applicability.

Given the intricate and interdependent nature of these determinants, contemporary research increasingly employs Educational Data Mining (EDM) and Machine Learning Algorithms (MLAs) to improve performance prediction and early intervention strategies (Nachouki et al., 2023). Comparative studies consistently report the superiority of non-linear models such as Random Forest Regression over Multiple Linear Regression, with significantly higher predictive accuracy, thereby reinforcing the value of data-driven approaches in understanding and enhancing academic success (Tapio, 2025). Nevertheless, most studies remain focused on global or Western contexts, and there is a clear research gap regarding how machine learning models can incorporate linguistic, foreign language, domestic, and gender-related factors to predict academic performance in Sri Lanka's multilingual higher education institutions.

In summary, while previous research has established the importance of prior academic performance, English proficiency, foreign language exposure, and domestic and sociocultural factors in shaping academic outcomes, most studies have examined these variables in isolation or within limited contexts. Gender effects and district-based disparities remain underexplored in multilingual higher education settings, particularly in non-state institutions in Sri Lanka. Furthermore, although machine learning and educational data mining approaches offer powerful tools for modeling complex, non-linear relationships among multiple factors, few studies have applied these techniques to integrate linguistic, domestic, demographic, and motivational variables simultaneously in this context. This gap underscores the need for a comprehensive, data-driven investigation that examines how these interrelated factors collectively influence undergraduate academic performance, providing evidence to inform targeted interventions and policy decisions.

ISSN 3121-4347

3. Methodology

A quantitative research design was adopted to examine the factors influencing undergraduate academic performance. Data were collected from 260 undergraduate students enrolled in a private higher education institute, capturing variables such as English and Japanese language proficiency (measured at both O/L and A/L levels), gender, district of origin, and prior IT education. The dependent variable in this study is students' Cumulative Grade Point Average (CGPA), which represents overall academic performance.

	A	B	C	D	E	F	G	H
1	Gender	District	A/L IT	O/L Japan	A/L Japan	JLPT/NAT	O/L Englis	CGPA
2	Male	Gampaha	No	No	No	No	S	3.2351
3	Male	Colombo	No	No	No	No	S	2.8702
4	Male	Gampaha	No	Yes	No	No	B	2.3035
5	Male	Galle	Yes	No	No	No	S	2.1333
6	Female	Gampaha	No	Yes	Yes	N4	C	2.9965
7	Female	Matara	Yes	No	No	No	C	3.5772
8	Female	Galle	Yes	Yes	Yes	N5	B	3.9596
9	Female	Galle	Yes	No	Yes	N5	C	2.9403
10	Male	Gampaha	Yes	No	No	No	C	2.0737
11	Female	Badulla	No	No	No	No	C	3.6737
12	Female	Anuradha	Yes	No	No	No	C	2.6772
13	Male	Kurunega	No	No	No	No	C	3.2193
14	Male	Galle	Yes	No	No	No	C	3.121
15	Female	Colombo	No	No	No	No	A	2.6088
16	Female	Moneraga	No	No	No	No	S	1.9755
17	Male	Moneraga	Yes	No	No	No	C	3.7439
18	Male	Kandy	No	No	No	N4	A	3.7316
19	Female	Matara	No	No	Yes	No	B	3.9369
20	Male	Gampaha	No	No	No	No	S	3.114
21	Male	Colombo	Yes	No	Yes	N4	C	3.9842
22	Male	Ratnapura	Yes	No	Yes	N3	B	3.6035
23	Female	Rathnapu	Yes	No	No	No	B	3.6719

DataSheet

Figure 1
Datasheet

Figure 1 illustrates the structure of the dataset. The dataset comprises both demographic variables (gender, district of origin) and academic background variables, including language proficiency and prior IT education. To ensure accuracy and suitability for analysis, the data underwent preprocessing and cleaning, which included handling missing values through imputation, removing duplicate and inconsistent records, encoding categorical variables for compatibility with statistical and machine learning analyses,

ISSN 3121-4347

standardizing continuous variables where necessary, and examining outliers using boxplots and z-scores. These steps ensured a clean, reliable, and well-structured dataset for analysis.

Data analysis was conducted using a combination of descriptive and inferential statistical methods. Descriptive statistics were used to summarize the sample characteristics, while correlation analysis explored relationships among variables and one-way ANOVA assessed differences between groups. To identify the most significant predictors of academic performance and examine both linear and non-linear relationships, multiple linear regression and Random Forest regression models were employed. All statistical computations and modeling were performed using Python, providing robust, reproducible, and reliable results.

Overall, this methodology ensures a systematic approach to examining the influence of linguistic, domestic, and contextual factors on undergraduate academic performance, allowing for comprehensive descriptive, inferential, and predictive analyses.

4. Results and Discussion

The empirical findings highlight the relationships between linguistic, domestic, and contextual factors and undergraduate academic performance, measured by CGPA. Correlation analysis, one-way ANOVA, multiple linear regression, and Random Forest regression were employed to examine both linear and non-linear influences, providing a comprehensive understanding of the factors shaping student outcomes. The following sections present the results in detail, highlighting key trends, differences across groups, and the factors most strongly associated with academic outcomes.

```
Correlation with CGPA:
CGPA                              1.000000
English_Knowledge_Score           0.357427
JLPT_Level                        0.223027
IT_Prior_Education_Score          0.151915
AL_Japanese                       0.146275
OL_Japanese                       0.119430
Gender                            0.044561
Japanese_Prior_Education_Score   -0.000876
Name: CGPA, dtype: float64
```

Figure 2
Correlation Analysis

ISSN 3121-4347

The correlation analysis, presented in Figure 2, explores the linear relationships between the independent variables and CGPA. Among all predictors, English language proficiency demonstrates the strongest positive correlation with academic performance ($r = 0.357$, $p < 0.01$), indicating that students with higher English proficiency are more likely to achieve higher CGPA values. This underscores the central role of English proficiency in navigating an English-medium higher education environment, affecting comprehension of course materials, participation in academic discussions, and performance in assessments. Japanese language proficiency shows a moderate positive correlation with CGPA ($r = 0.212$, $p < 0.05$), suggesting that prior exposure to foreign languages may enhance cognitive flexibility and problem-solving skills, indirectly supporting academic success. Similarly, prior IT education exhibits a moderate positive correlation ($r = 0.198$, $p < 0.05$), highlighting the importance of technological preparedness in accessing digital resources, completing assignments, and engaging with learning management systems. In contrast, gender ($r = 0.045$) and district of origin ($r = 0.061$) display very weak correlations with CGPA, implying limited direct linear influence on academic performance. However, while their linear associations are weak, these factors may interact with other linguistic or domestic variables in more complex, non-linear ways, which warrants further analysis using regression and machine learning models. Overall, these findings suggest that linguistic competencies particularly English proficiency play a foundational role in academic achievement, while technological readiness and additional language exposure provide secondary support. Demographic variables appear less influential in isolation but may still contribute indirectly through interactions with other predictors.

To examine group-based differences in academic performance, a one-way ANOVA was conducted. The results indicate a statistically significant difference in CGPA across gender groups ($F = 4.12$, $p = 0.017$), suggesting that gender has a measurable influence on academic performance. In contrast, the effect of district of origin was not statistically significant ($p > 0.05$), indicating that geographical background does not substantially affect academic achievement within the studied institution. These findings reflect a relatively equitable academic environment across districts while highlighting modest gender-based differences.

ISSN 3121-4347

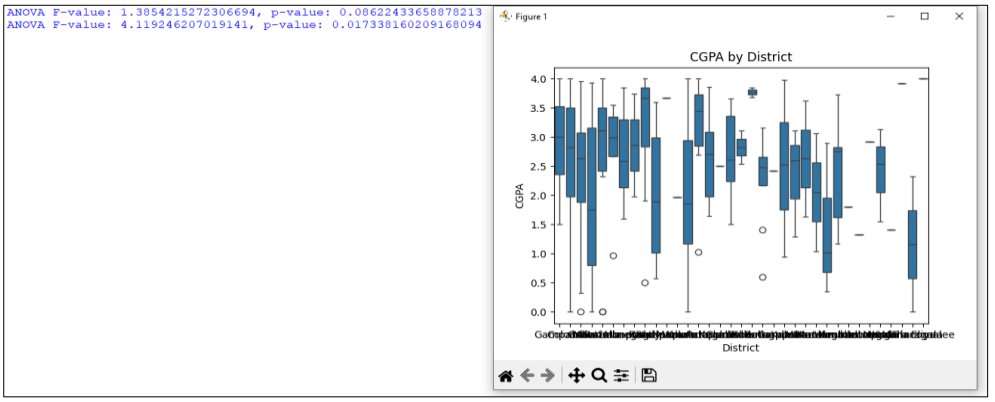

Figure 3

ANOVA Test Results

Building on these findings, multiple linear regression analysis was employed to evaluate the combined effect of all predictors on CGPA. The regression coefficients, shown in Figure 4, indicate that English O/L proficiency is a statistically significant positive predictor of academic performance, reinforcing its dominant influence. Japanese O/L proficiency also contributes positively, though with a smaller effect size. Notably, Japanese A/L proficiency exhibits a negative coefficient. This suggests that foundational language competence is more strongly associated with overall academic success than advanced language specialization.

```
Data Loaded Successfully!
   Gender District A/L IT  ... JLPT/NAT O/L English    CGPA
0    Male  Gampaha     No  ...       No           S  3.2351
1    Male  Colombo     No  ...       No           S  2.8702
2    Male  Gampaha     No  ...       No           B  2.3035
3    Male    Galle    Yes  ...       No           S  2.1333
4  Female  Gampaha     No  ...       N4           C  2.9965

[5 rows x 8 columns]

Linear Regression R^2 score: 0.110

Feature Importance (Linear Regression Coefficients):
OL_Japanese                0.498157
English_Knowledge          0.333596
IT_Prior_Education         0.164427
Japanese_Prior_Education   0.126244
JLPT_Level                 0.048329
Gender_Num                 0.015530
AL_Japanese               -0.210121
dtype: float64
```

Figure 4

Multiple Regression Results

ISSN 3121-4347

Multiple linear regression was employed to assess the combined effect of all predictors on CGPA and to quantify the unique contribution of each factor while controlling for the influence of others. This approach allows for the examination of linear relationships between academic performance and multiple linguistic, and demographic variables simultaneously, providing insights into which factors most strongly predict outcomes when interdependencies are accounted for.

English proficiency clearly emerges as the most influential factor, followed by Japanese proficiency and prior IT education. Gender demonstrates a smaller yet observable effect, while district of origin remains minimally influential. This visualization corroborates the statistical findings and further confirms the predominance of linguistic factors over demographic variables in predicting CGPA.
To capture potential non-linear relationships and improve predictive accuracy, a Random Forest regression model was applied. The feature importance scores derived from this model are presented in Figure 5.

```
Python 3.12.3 (tags/v3.12.3:f6650f9, Apr  9 2024, 14:05:25) [MSC v.1938 64 bit (
AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.

= RESTART: E:\On going Researches\Student Success Analysis - Machine Learning\Co
de\RandomForest.py
Data Loaded Successfully!
English_Knowledge           0.553414
Japanese_Prior_Education    0.124937
Gender_Num                  0.111151
IT_Prior_Education          0.088407
JLPT_Level                  0.072149
AL_Japanese                 0.025448
OL_Japanese                 0.024494
dtype: float64
```

Figure 5
Random Forest Regression

English proficiency stands out as the most influential predictor, with a feature importance value of 0.553, substantially higher than all other variables. Japanese prior education (0.125), gender (0.111), and IT education (0.088) contribute moderately, whereas district of origin demonstrates negligible importance. The Random Forest model's ability to capture non-linear relationships provides deeper insights into variable interactions,

ISSN 3121-4347

demonstrating superior explanatory power compared to traditional linear models. These results strongly validate the central role of English proficiency in undergraduate academic achievement.

Overall, the analysis demonstrates that linguistic competence, particularly English proficiency, is the most significant determinant of undergraduate academic performance, consistently emerging as the strongest predictor across correlation, regression, and Random Forest analyses. Japanese language proficiency and prior IT education contribute positively, even though to a lesser extent, highlighting the supporting roles of cognitive flexibility and technological preparedness in shaping academic outcomes. Gender exhibits modest effects, whereas district of origin has minimal influence, suggesting that demographic factors play a secondary role relative to linguistic and skill-based competencies. The combination of linear and non-linear modeling approaches provides a nuanced understanding of the interplay among predictors, revealing that foundational language skills and technological readiness interact to influence CGPA more substantially than advanced language specialization or geographical background. These findings underscore the importance of targeted interventions to enhance language proficiency and digital literacy, while also recognizing the complex, multifactorial nature of academic success in multilingual higher education settings.

5. Conclusion and Future Work

This study examined the influence of linguistic, domestic, and contextual factors on undergraduate academic performance in a multilingual higher education setting, measured by CGPA. The findings consistently highlight the dominant role of English proficiency as the most significant predictor of academic success, reinforcing the importance of strong foundational language skills in navigating course materials, participating in academic discussions, and performing effectively in assessments. Japanese language proficiency and prior IT education also contribute positively to academic outcomes, supporting cognitive flexibility, problem-solving, and technological preparedness, although their influence is smaller compared to English proficiency.

Gender exhibits modest effects on academic performance, while district of origin appears to have minimal direct influence. However, demographic variables may interact with linguistic and skill-based factors in complex ways that are better captured through non-linear modeling approaches. The combination of correlation analysis, ANOVA, multiple linear regression, and Random Forest regression provided complementary insights:

ISSN 3121-4347

correlation and ANOVA revealed initial associations and group differences, multiple regression quantified the unique contributions of predictors while controlling for interdependencies, and Random Forest regression captured non-linear relationships and interactions, offering a more comprehensive understanding of academic performance determinants.

Overall, the study demonstrates that linguistic competencies, particularly English proficiency, together with technological readiness, are central to academic achievement, while demographic factors play a secondary role. These findings underscore the need for targeted interventions to enhance language proficiency and digital literacy, enabling equitable learning opportunities across diverse student populations. Furthermore, the methodological approach illustrates the value of combining traditional statistical analyses with modern machine learning techniques to capture the multifactorial and interactive nature of academic performance in multilingual higher education contexts.

While this study provides valuable insights into the factors influencing undergraduate academic performance, several avenues remain for future investigation. First, larger and more diverse samples across multiple higher education institutions in Sri Lanka could enhance the generalizability of the findings and capture broader regional and institutional variations. Second, the role of advanced language proficiency and multilingual exposure warrants further exploration, particularly to understand why higher-level Japanese proficiency showed a negative association in linear models and how advanced language skills interact with foundational competencies. Third, longitudinal studies could track students over time to examine how linguistic competence, IT literacy, and socio-demographic factors influence academic trajectories, retention, and career outcomes, rather than only static CGPA measures. Future research could incorporate additional cognitive, motivational, and behavioral variables, such as study habits, self-regulation, and class engagement, to develop more comprehensive predictive models of academic success. Finally, while this study employed Random Forest regression to capture non-linear interactions, other machine learning and deep learning techniques such as Gradient Boosting Machines or Neural Networks could be explored to further improve predictive accuracy and uncover hidden patterns in student performance data.

These avenues can inform the development of more targeted educational interventions, support data-driven policy decisions, and contribute to creating equitable and effective learning environments in multilingual and technology-driven higher education contexts.

ISSN 3121-4347

6. References

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. Review of Educational Research, 49(2), 222–251.

Hyde, M. E., & Mertz, J. (2009). Gender, culture, and mathematics performance. Proceedings of the National Academy of Sciences, 106(22), 8801–8807.

Islam, F., Krishna, A., Kumar, D., & Kumari, S. (2025). Factors influencing academic performance: An empirical study using predictive analytics. Multidisciplinar (Montevideo), 3(51).

Lightbown, P., & Spada, N. (2013). How languages are learned. Oxford University Press.

Nachouki, M., Mohamed, E. A., Mehdi, R., & Abou Naaj, M. (2023). Student course grade prediction using the random forest algorithm: Analysis of predictors' importance. Trends in Neuroscience and Education, 33, 100214.

Purpura, A. C. (2004). Assessing grammar. Cambridge University Press.

Rajendran, S., Sinha, A. A., & Chamundeswari, S. (2021). Predicting factors impacting student academic performance using machine learning algorithms. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.

Tapio, R. P. (2025). Comparative analysis of multiple linear regression and random forest regression in predicting academic performance of students in higher education. Asian Research Journal of Mathematics, 21(4), 170–181.

Wang, D., Lian, D., Xing, Y., Dong, S., Sun, X., & Yu, J. (2022). Analysis and prediction of influencing factors of college student achievement based on machine learning. Frontiers in Psychology, 13, 881859.

ISSN 3121-4347

JASPER Research Journal
Volume 1 Issue 1
March 2026

Academic Predictors of CGPA: The Role of O/L and A/L Performance in Computing and Software Engineering Students

Samanthi K.G.P.H.[1], Wijewickrama B.N.[2]
[1,2]Lanka Nippon BizTech Institute, Maharagama, Sri Lanka.
hansika@edu.lnbti.lk
DOI: https://doi.org/10.5281/zenodo.19703225

Abstract

This study explores academic determinants of cumulative grade point average (CGPA) among 335 undergraduates enrolled in Computing and Software Engineering degree programs at a private higher education institute. Students were admitted after Ordinary Level (O/L) or Advanced Level (A/L) examinations, with Software Engineering restricted to entrants after A/L. The study investigates whether early academic performance indicators predict undergraduate success. A quantitative explanatory research design was employed to analyze institutional academic records using one-way ANOVA, Pearson correlation, and multiple linear regression. For Computing students admitted after O/L (n = 106), ANOVA indicated strong and positive associations between English and Mathematics grades and CGPA. Among students admitted after A/L (n = 250), O/L English and Mathematics grades, together with A/L mean scores, were significantly associated with CGPA. Multiple regression analysis showed that these predictors collectively explained approximately 31% of the variance in CGPA. In the full sample, O/L Mathematics and English grades remained the strongest predictors of academic achievement. These findings highlight the importance of foundational knowledge in Mathematics and English for academic success in higher education. Prior proficiency in Japanese showed mixed results: it was not statistically significant among O/L entrants but had a significant positive effect among A/L entrants and in the overall sample. Overall, secondary-level academic performance is a significant predictor of university achievement.

Keywords: Cumulative Grade Point Average (CGPA), ANOVA, Multiple Regression Analysis, Higher Education, Academic Performance.

1. Introduction

A private higher education institute offers degree programs in Software Engineering and Computing to students who enter after completing either the Advanced Level (A/L) or the Ordinary Level (O/L) examinations. In both programs, the curriculum integrates the study of the Japanese language alongside Information Technology–related subjects. Admission criteria for both groups of students include proficiency in English and Mathematics.

ISSN 3121-4347

Additionally, some enrolled students have prior knowledge of the Japanese language, specifically at the Japanese Language Proficiency Test (JLPT) N5 and N4 levels. At the end of each semester, students sit for separate final examinations in Japanese language and IT-related subjects, with academic performance measured using grade point average (GPA*)*.

Although English and Mathematics proficiency are considered in the admission process for both groups, students begin their degree programs with different academic backgrounds. These background variables include English and Mathematics grades obtained at O/L or A/L examinations, overall academic performance at the highest secondary qualification, and prior knowledge of Japanese language. However, the extent to which these academic factors influence undergraduate academic performance has not been adequately explored, and there is limited empirical evidence explaining observed disparities in GPA between entry groups. The primary aim of this study is to investigate the impact of students' prior academic background on their academic performance in Software Engineering and Computing degree programs. Specifically, the study seeks to: (1) analyze the relationship between prior English proficiency and cumulative grade point average (CGPA); (2) examine the effect of prior Mathematics performance on CGPA; (3) assess whether overall academic achievement at the O/L or A/L level influences CGPA outcomes; and (4) evaluate the role of prior Japanese language proficiency in students' academic performance.

This research is significant as it provides empirical evidence on key academic determinants influencing the performance of students enrolled in Software Engineering and Computing programs. By identifying the relative contribution of English, Mathematics, secondary academic performance, and Japanese language proficiency to CGPA, the findings can inform admission policies and targeted academic support strategies. Furthermore, the results may assist higher education administrators and curriculum developers in designing differentiated support mechanisms for students entering through diverse academic pathways, particularly those admitted after O/L examinations. The study also contributes to the broader literature on student preparedness and academic performance in multilingual and technically oriented higher education contexts.

ISSN 3121-4347

2. Literature Review

Previous research has consistently examined the relationship between prior academic performance and achievement in higher education, with many studies highlighting the predictive value of secondary-level academic indicators. Karma and Dautaj (2025) demonstrated that high school grade point average (GPA) has a measurable influence on university academic outcomes, although its effect varies across contexts and is shaped by additional academic and non-academic factors. Similarly, Dagdagui (2022) and Nozaleda (2019) identified high school GPA as a statistically significant predictor of undergraduate academic performance, emphasizing the importance of sustained academic achievement over standardized aptitude testing.

Several studies have employed multivariate statistical approaches to capture the complex determinants of academic success. Indunan, Dait, and Dulnuan (2025) utilized stepwise multiple regression to identify significant predictors of postgraduate performance in statistics-related courses, illustrating how multiple academic variables jointly influence achievement. Comparable findings were reported by Z. et al. (2023), whose regression analysis revealed prior academic achievement as one of the most influential contributors to cumulative grade point average (CGPA). These studies underscore the effectiveness of regression-based models in isolating the relative contributions of multiple predictors.

Discipline-specific and contextual variations in academic performance have also been explored. Banal et al. (2024) examined first-year medical students and found that admission qualifications, including prior grades and entrance test scores, significantly predicted early university achievement, highlighting the importance of foundational academic preparation in professional degree programs. Al Asmar et al. (2021) extended this perspective by demonstrating that high school GPA not only predicts academic performance but is also associated with long-term career satisfaction among dental graduates, suggesting the enduring impact of prior academic preparation.

Comparative studies further emphasize the role of academic background in shaping performance outcomes. Tindan and Anaba (2024) reported that differences in academic achievement between regular and distance-learning students were primarily attributable to prior academic preparation rather than mode of study. A comprehensive review by Aciro et al. (2021) synthesized findings across disciplines and institutions, concluding that entry grades generally exhibit a positive association with university performance, although the

ISSN 3121-4347

strength of this relationship varies by program type and educational context.

Despite substantial evidence supporting the predictive value of prior academic achievement, existing studies largely focus on general undergraduate populations or single-discipline contexts. There remains limited empirical research examining how multiple academic background variables—such as subject-specific proficiency, overall secondary-level performance, and prior foreign language knowledge—interact to influence academic outcomes in multilingual and technically oriented programs such as Computing and Software Engineering. This study seeks to address this gap by examining the combined effects of English, Mathematics, secondary academic performance, and Japanese language proficiency on CGPA among undergraduate students.

3. Methodology

This research used a quantitative explanatory research design to investigate the relationship between prior academic indicators and undergraduate academic performance, measured using cumulative grade point average (CGPA). Statistical methods were used to examine these relationships across the students.

3.1. Study Participants and Group Classification

The students were categorized into three groups based on their entry pathway in to computing or software engineering degree programs

- Group 1: Students who applied after completing their O/L examinations.
- Group 2: Students who applied after completing their A/L examinations.
- Group 3: All students (combined sample of Group 1 and Group 2) to assess overall relationships.

The distribution of students across these groups, illustrating the relative sample sizes for each entry pathway, is presented in Figure 1. This grouping allowed for comparative analysis of prior academic indicators and their potential influence on academic performance across different entry pathways.

ISSN 3121-4347

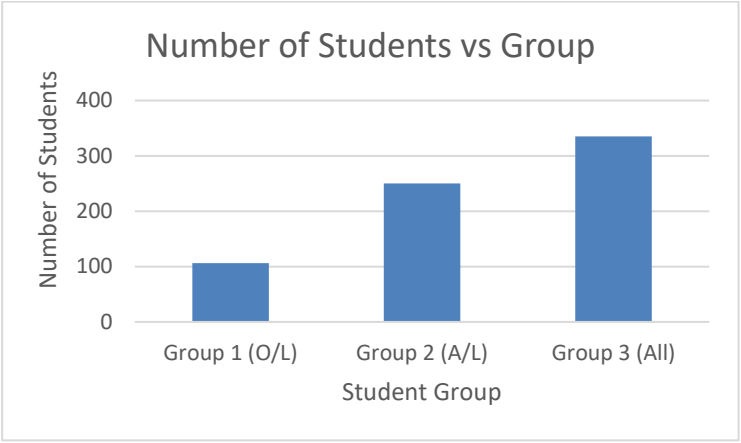

Figure 05
Student Distribution by Study Group

3.2. Prior Academic Indicators

The prior academic indicators collected for each group are presented in Table 2. These indicators represent the students' academic performance prior to entering the Computing or Software Engineering degree programs. They were obtained from institutional academic records.

Table 1
Summary of prior academic indicators considered for the student groups

Group 1: After O/L students	Group 2: After A/L students	Group 3: All students
O/L English grade	O/L English grade	O/L English grade
O/L Mathematics grade	O/L Mathematics grade	O/L Mathematics grade
O/L Japanese grade (If applicable)	O/L Japanese grade (If applicable)	O/L Japanese grade (If applicable)
Grades for O/L nine (9) subjects	A/L Japanese grade (If applicable)	Japanese language proficiency (JLPT Level) (If applicable)
Japanese language proficiency (JLPT Level) (If applicable)	Grades for A/L three (3) subjects	Degree Program

ISSN 3121-4347

CGPA	Japanese language proficiency (JLPT Level) (If applicable)	CGPA
	Degree Program	
	CGPA	

3.3 Data Preparation

Several variables were derived for statistical analysis.

- O/L English and O/L Mathematics grades (Group 1,2, and 3) - Converted into numerical values using an ordinal scale, where A is the highest grade and thus has the highest value.

  A = 4, B = 3, C = 2, S = 1, F = 0

- O/L mean score (Group 1) - The letter grades were converted into numerical values using the average mark range represented by each grade category.

Table 2
Ranges of mark in each Grade

Mark Range	Grade
100 – 75	A
74 - 65	B
64 – 55	C
54 - 40	S
39 - 0	F

  - Each grade was assigned the midpoint value of its corresponding score range to provide an approximate quantitative representation of student performance. Accordingly, the values used were:

    A = 87.5, B = 69.5, C = 59.5, S = 47.0, F = 19.5

- A/L mean score (Group 2) - Calculated as the average of three A/L subjects using the same numerical scale of O\L mean score.
- Prior Japanese knowledge level was coded as an ordinal variable representing increasing proficiency. Since JLPT N3 represents the highest level of Japanese

ISSN 3121-4347

proficiency among the considered categories, it was assigned the highest numerical value in the coding scheme.

- For Group 1 & 3:
  - 0 = no prior knowledge
  - 1 = O/L Japanese
  - 2 = JLPT N5
  - 3 = JLPT N4
  - 4 = JLPT N3
- For Group 2:
  - 0 = no prior knowledge
  - 1 = O/L Japanese
  - 2 = A/L Japanese
  - 3 = JLPT N5
  - 4 = JLPT N4
  - 5 = JLPT N3

- Degree program (Group 2 and 3) was coded as:
  - 0 = Computing degree
  - 1 = Software Engineering degree.

The data was screened for missing values on the variables considered important and list-wise deletion of observations was done.

3.4 Data Analysis Methods

The data analysis methods for each study group are as follows.

3.4.1 Group 1 (After O/L students)

One-way ANOVA tested relationships between CGPA and the following parameters.

- O/L English
- O/L Mathematics
- Prior Japanese language knowledge

Multiple linear regression assessed combined and individual effects of these predictors on CGPA using the model:

$$\text{CGPA} = \beta_0 + \beta_1(\text{OL Maths}) + \beta_2(\text{OL English}) + \beta_3(\text{OL Mean}) + \beta_4(\text{Previous Japanese}) + \varepsilon$$

Equation 1

Regression Model of CGPA on Prior Academic Indicators (Group 1)

ISSN 3121-4347

3.4.2 Group 2 (After A/L students)

A similar regression model was applied including A/L performance and degree program.

$$\text{CGPA} = \beta_0 + \beta_1(\text{OL Maths}) + \beta_2(\text{OL English}) + \beta_3(\text{AL Mean}) + \beta_4(\text{Program}) + \beta_5(\text{Previous Japanese}) + \varepsilon$$

Equation 2
Regression Model of CGPA on Prior Academic Indicators (Group 2)

3.4.3 Group 3 (All students)

One-way ANOVA tested relationships between CGPA and the following parameters.

- O/L English
- O/L Mathematics
- Prior Japanese language knowledge

A regression model including degree program was then estimated using the formula below:

$$\text{CGPA} = \beta_0 + \beta_1(\text{OL Maths}) + \beta_2(\text{OL English}) + \beta_3(\text{Program}) + \beta_4(\text{Previous Japanese}) + \varepsilon$$

Equation 3
Regression Model of CGPA on Prior Academic Indicators of (Group 3)

Pearson correlation coefficients were calculated to assess bivariate associations prior to multivariate analysis.

All analyses, including ANOVA, correlation, and regression (Ordinary Least Squares), were performed using Python with the 'statsmodels' and related libraries. Model fit was evaluated using R-squared (R^2).

4. Results and Discussion

Statistical results are presented in two stages: (1) bivariate associations assessed through ANOVA and Pearson correlation, and (2) multivariate associations assessed through multiple linear regression. Statistical significance was evaluated at $\alpha = 0.05$.

ISSN 3121-4347

1.1 One-way ANOVA test results of Group 1

A one-way ANOVA was conducted to examine the relationship between O/L English performance, O/L Mathematics performance, prior Japanese knowledge level, and students' CGPA. The resulting p-values are as follows:

- CGPA and O/L English - p = 0.000433
- CGPA and O/L Mathematics - p = 0.000754
- CGPA and prior Japanese knowledge level - p = 0.121741

The results indicate that O/L Mathematics and O/L English performances are significantly associated with CGPA, whereas prior Japanese knowledge does not a significant relationship with CGPA.

Subsequently, a multiple linear regression analysis was conducted to examine the combined and individual effects of O/L English, O/L Mathematics, O/L mean score, and prior Japanese knowledge level on CGPA. The regression model reported an R-squared value of 0.294, indicating that 29% of the variance in the CGPA can be explained by these predictors.

Pearson Correlation coefficients with CGPA are summarized below:

O/L Mathematics	0.454530
O/L English	0.394547
O/L Mean	0.267015
Prior Japanese knowledge level	0.259846

1.2 Multiple Regression Analysis Result of Group 1

This section presents the results of the multiple regression analysis for Group 1 (After O/L students), showing the relationships between CGPA and the selected prior academic indicators. Table 3 summarizes the significance (P values) of each independent variable in predicting CGPA.

Table 3

Multiple regression analysis results on independent variables with CGPA

Independent Variable	P Value
O/L Mathematics	0.411
O/L English	0.089
O/L mean score	0.091
Prior Japanese knowledge level	0.091

ISSN 3121-4347

All regression coefficients were positive, suggesting that higher English, Mathematics, and mean O/L performance, as well as prior Japanese knowledge, are associated with higher CGPA. However, none of the predictors reached statistical significance in the multivariate model at $\alpha = 0.05$. This suggests that although English and Mathematics show significant bivariate associations, their shared variance with other academic indicators reduces their individual effects in multivariate analysis.

For students entering after O/L examinations, English and Mathematics proficiency appear to form the strongest initial indicators of undergraduate success, although the explanatory power is shared across multiple academic variables.

1.3 Multiple Regression Analysis Results of Group 2

Multiple linear regression was used to evaluate the effects of O/L English, O/L Mathematics, A/L mean score, degree program, and prior Japanese knowledge level on CGPA. The model yields a R-squared of 0.312 indicating that approximately 31% of the variance in CGPA is explained by these predictors.

Pearson correlations with CGPA are as follows:

Correlation Matrix with CGPA:	
O/L Mathematics	0.420361
O/L English	0.380770
A/L Mean Score	0.271287
Degree Program	0.215352
Prior Japanese Knowledge Level	0.208790

Regression p-values are shown in Table 4:

Table 4
Multiple regression analysis results of independent variables with CGPA

Independent variable	P value
O/L Mathematics	0.000
O/L English	0.000
A/L Mean Score	0.002

ISSN 3121-4347

Degree Program	0.406
Prior Japanese Knowledge Level	0.010

The results show that O/L English, O/L Mathematics, A/L Mean Score, and Prior Japanese Knowledge are significant predictors of CGPA at $\alpha = 0.05$, while degree program is not significant.

Interpretation:
Among students admitted after A/L examinations, English and Mathematics remain the strongest predictors, with A/L mean performance and Japanese language proficiency contributing smaller but statistically significant effects. The degree program does not significantly influence CGPA once academic performance indicators are controlled for, suggesting that differences between computing and software engineering performance are driven by student preparation rather than program structure.

1.4 Multiple Regression Analysis Results of Group 3

Multiple linear regression analysis was performed on all students combined to examine the effects of O/L English, O/L Mathematics, degree program, and prior Japanese knowledge level on CGPA.

The model produced an R-squared of 0.307, explaining approximately 31% of the variance in CGPA.

Correlation Matrix with CGPA:

O\L Mathematics	0.465940
O\L English	0.385985
Degree Program	0.029668
Prior Japanese Knowledge Level	0.227257

Regression p-values are provided in Table 5:

Table 5
Multiple regression analysis results of independent variables with CGPA

Independent variable	P value
O/L Mathematics	0.000

ISSN 3121-4347

O/L English	0.000
Degree Program	0.770
Prior Japanese Knowledge Level	0.000

The coefficients indicate that O/L English, O/L Mathematics, and prior Japanese knowledge are significant predictors of CGPA for the full sample, whereas degree program has no significant effect.

Interpretation:
In the combined sample, prior Japanese language proficiency emerges as a significant predictor once sample size increases, supporting the notion that language skills may provide an academic advantage. However, degree program does not contribute to CGPA variation, confirming earlier findings that performance differences across programs stem from student characteristics rather than the program itself. In summary, the results can be emphasized through the following observations.

1. O/L Mathematics and O/L English are the strongest and most consistent predictors of undergraduate academic performance.
2. Prior Japanese knowledge becomes significant when larger samples are considered (Group 2 and Group 3).
3. Degree program has no significant effect on CGPA once prior academic preparation is controlled for.
4. The models explain approximately 30% of the variance, indicating that while prior academic achievement is influential, additional non-academic variables (e.g., learning skills, motivation, socio-economic factors) may account for remaining variance.

5. Conclusion and Recommendations

The analysis of academic performance of undergraduate students shows that CGPAs are heavily dependent on past academic proficiency. The English proficiency at the O/L level depicted a constant and significant relationship with academic outcomes, as higher English scores were found to be related to better performance in all categories of students.
Mathematics performance also emerged as a strong predictor of CGPA. Students who scored higher in the O/L Mathematics test had consistently higher results, showing that

ISSN 3121-4347

quantitative skills form very important foundations for success in higher education programs.

In general, academic achievement at O/L or A/L contributed positively to university achievement. Among students who entered the University following A/L examinations, A/L mean scores added predictive value beyond existing measures of O/L achievement while for students admitted after taking O/L examination; the overall mean score at O/L had a marginal effect in a combined analysis with other predictors.

Prior knowledge in the Japanese language, on the other hand, demonstrated a smaller but notable effect on academic performance. Although it was non-significant among the O/L entry students, prior Japanese proficiency positively influenced CGPA in A/L students and when considering all students together, therefore suggesting that early exposure to a second language may support overall learning and academic success.

In all, O/L English and Mathematics remain the strongest predictors of undergraduate performance, while a set of overall prior academic achievement and prior Japanese knowledge each adds significant explanatory value. Type of degree program did not have an independent influence on CGPA when controlling for prior academic indicators. These results point to the need to strengthen foundational skills and to provide language support as keys to optimizing student success in higher education.

Based on the findings, the recommendation includes strengthening students' foundational skills in English and Mathematics through bridging programs or targeted support, since these bear significantly on academic performance. Assessment of prior academic records as early as possible can be a good way of understanding students who may require additional guidance, and structured language support may be helpful for those with limited prior Japanese experience. Academic preparedness should be a priority in admission and curriculum planning rather than the type of degree program. Future research must continue to investigate other factors that may play important roles in determining university success, such as motivation and study strategies.

6. Acknowledgment

I would like to express sincere gratitude to Ms. Bhashini Wijewickrama, the research supervisor, for her invaluable guidance, constructive feedback, and continuous support

throughout the course of this study. Her expertise and encouragement greatly contributed to the successful completion of this research.

7. References

Aciro, R., Onen, D., Malinga, G. M., Ezati, B. A., & Openjuru, G. L. (2021). Entry Grades and the Academic Performance of University Students: A Review of Literature. Education Quarterly Reviews, 4(1). https://doi.org/10.31014/aior.1993.04.01.181

Al Asmar, A. A., Oweis, Y., Ismail, N. H., et al. (2021). The predictive value of high school grade point average to academic achievement and career satisfaction of dental graduates. BMC Oral Health, 21, 300. https://doi.org/10.1186/s12903-021-01662-5.

Banal, R. A. R., Malijan, M. C. B., Solidum, F. P., Clamor, M. M., & Rio, P. C. (2024). The predictive value of admission qualifications on the academic performance of first-year medical students. Journal of Education and Learning, 13(6), 115–132. https://doi.org/10.5539/jel.v13n6p115.

Dagdagui, R. T. (2022). Predicting Students' Academic Performance Using Regression Analysis. American Journal of Educational Research, 10(11), 640–646.

Indunan, D., Madonna G. Dait, J., & Steven P. Dulnuan, A. (2025). Uncovering Key Predictors of Statistics Achievement among Postgraduate Students: A Stepwise Regression Model. Interdisciplinary Research Review, 20(4). Retrieved from https://ph02.tci-thaijo.org/index.php/jtir/article/view/256293.

Karma, E., & Dautaj, M. (2025). Influence of high school grade point average on university performance: A case study from Albania. International Journal of Evaluation and Research in Education (IJERE), 14(1), 463–471. http://doi.org/10.11591/ijere.v14i1.29142

Nozaleda, B. (2019). Validity of High School GPA and the College Aptitude Test (CAT) in Predicting College Academic Performance. International Journal of Scientific Research and Engineering Development, 2.

Tindan, T. N., & Anaba, C. A. (2024). Quality of regular students versus distance students, degree graduate programs in Ghanaian Universities. International Journal of Multidisciplinary Research and Growth Evaluation, 5(2), 650–654. https://doi.org/10.54660/.ijmrge.2024.5.2.650-654.

Z., Omar, N. F., Dzulkarnain, N., Abu Bakar, S. M. S., & Mohd Fauzi, M. W. (2023). The regression analysis of factors contributing to university students' academic performance. Information Management and Business Review, 15(4), 456–464.

JASPER Research Journal
Volume 1 Issue 1
March 2026

THE RELATIONSHIP BETWEEN E-LEARNING ENGAGEMENT AND STUDENT ACADEMIC PERFORMANCE

Senarathne, K.M.G.
LNBTI Japanese IT University, Maharagama, Sri Lanka
geethika@edu.lnbti.lk
DOI: https://doi.org/10.5281/zenodo.19703252

Abstract

E-learning platforms have become an essential component of higher education, enabling flexible access to learning resources and assessments. Online learning platforms are frequently used by Sri Lankan universities for blended learning to enhance communication between teachers and students. This study examines the relationship between student engagement in a Learning Management System (LMS) and academic performance using linear analysis in a blended learning environment. The analysis is based on LMS activity log records and assignment marks collected from 37 undergraduate IT students during fifteen weeks of a semester. The results demonstrate a positive relationship between LMS engagement and assignment performance, indicating that increased interaction with course materials contributes to improved academic outcomes. The findings highlight the significance of LMS analytics in supporting data-driven instructional decisions and enhancing student learning effectiveness.
Keywords: E-learning, LMS, academic performance, linear regression analysis

1. Introduction

The rapid advancement of information and communication technologies has transformed the landscape of higher education, with e-learning platforms becoming an integral component of teaching and learning processes. Learning Management Systems (LMS) provide students with flexible access to learning materials, assessments, and interactive activities, while enabling instructors to monitor student engagement and performance. As universities increasingly adopt blended and fully online learning environments, understanding the relationship between student engagement in e-learning platforms and academic performance has gained significant attention among educators and researchers.

Student engagement in e-learning can be reflected through various indicators, including frequency of LMS access, interaction with course resources, and participation in online assessments such as quizzes. Prior studies have suggested that higher levels of engagement

are associated with improved learning outcomes; however, the strength and nature of these relationships may vary depending on the type of engagement, whether passive (e.g., content viewing) or active (e.g., assessment participation). Empirical evidence examining these dimensions remains particularly limited in the context of undergraduate IT education.

This study aims to examine the relationship between e-learning engagement and academic performance by analyzing LMS usage data, LMS quiz participation, and assignment marks of undergraduate students. By identifying key engagement factors that influence academic outcomes, the study seeks to provide insights that can support data-driven instructional design, enhance student monitoring strategies, and promote effective engagement behaviors within digital learning environments.

2. **Literature Review**

The integration of e-learning technologies into higher education has received considerable research attention, particularly in relation to student engagement and academic performance. Numerous studies report that student interaction with online learning platforms positively influences learning outcomes, especially when engagement involves active participation such as assessments and collaborative activities. Studies examining fully online learning environments indicate that frequent interaction with learning resources and online assessments is associated with higher academic achievement (Akpen et al., 2024), (Tolika et al.,2024). According to this research, the COVID-19 epidemic has had a major impact on educational practices globally and has led to a rise in the use of online learning platforms, but most universities still utilize blended learning.

Rasheed et al. (2020) examined the relationship between e-learning usage and student academic performance and reported significant positive associations between LMS engagement, learner satisfaction, and academic outcomes. Their findings suggest that structured e-learning activities and consistent system usage can enhance student performance. Rasheed et al. (2020) preliminarily investigate the relationship of service quality dimensions (tangibility, responsiveness, assurance, reliability, empathy, and e-learning) to students' academic performance. Similarly, studies focusing on participation metrics such as login frequency, content access, and online quizzes highlight that assessment-based engagement serves as a stronger predictor of performance than passive

content consumption (Tolika et al.,2024). In contrast to blended learning, the study by Tolika et al. (2024) examines whether increased academic performance in distance learning is correlated with active participation in online forums and educational platforms.
However, much of the existing literature focuses on fully online or distance learning contexts. Fewer studies address blended learning environments, where face-to-face instruction is supplemented with LMS-based resources, quizzes, and assignments. In such contexts, LMS engagement functions as a complementary learning support rather than the primary instructional medium.

Alam et al. (2021) developed a holistic e-learning service framework using data from 397 higher education students, identifying key determinants of LMS effectiveness, including learner, instructor, system, information, and institutional quality, and demonstrating that perceived usefulness and use of LMS tools explained a substantial proportion of variance in student academic performance (Alam et al. ,2021). This finding suggests that contextual and service-quality factors significantly influence the learning outcomes achieved through digital platforms.

Within the Sri Lankan higher education context, much of the existing research has focused on emergency remote teaching or fully online learning environments, particularly during the COVID-19 pandemic. During this period, online platforms became the primary mode of instruction rather than a supplementary tool. Studies such as Akpen et al. (2024) emphasize that LMS engagement significantly influenced student performance in fully online settings while also highlighting challenges including inconsistent participation, limited digital readiness, unstable internet access, and reduced face-to-face academic support. In such contexts, LMS interaction represented the central mechanism for content delivery, communication, and assessment.

Naveed et al, (2019) paper specifically focuses on identifying and ranking the critical success factors (CSFs) necessary for the successful implementation of cloud-based e-learning systems. While Alam et al (2021) focuses on the outcome (academic performance and sustainable learning), this 2019 paper specifically investigates the structural validity of the factors themselves.

The study by Elfaki et al. (2020) mainly focuses on comparing e-learning and traditional learning methods in terms of students' academic performance and attitudes. The findings showed that students in the e-learning group achieved higher examination scores and

ISSN 3121-4347

reported more positive attitudes toward learning than those in traditional classroom settings.

The study by Elfaki et al. (2020) examined the impact of e-learning compared with traditional face-to-face learning on students' academic performance and attitudes. In comparison, the study by Kokoç and Altun (2019) focused specifically on learner interaction with learning dashboards within an e-learning environment. Rather than comparing learning modes, this research analyzed how students' engagement with dashboard tools, progress tracking, and learning analytics influenced academic performance.

Ashrafi et al. (2020) investigates the factors that influence students' continuance intention to use a Learning Management System (LMS) after the initial adoption stage. Unlike studies that focus only on first-time acceptance, this research examines what makes students continue using the LMS over time. The key finding was that perceived usefulness is the strongest predictor of students' intention to continue using the LMS. In other words, when students believe that the LMS genuinely helps their learning and improves performance, they are more likely to keep using it. Interestingly, the study found that student satisfaction and attitude did not significantly affect continuance intention. This research is closely related to the topic of e-learning engagement and academic performance, but the focus is slightly different. This study focuses on continuance intention, which means students' willingness to keep using the LMS. While this paper mainly explains why students continue using e-learning platforms, engagement-performance studies explain how that continued use affects academic results.

In contrast, blended learning environments differ substantially from fully online or pandemic-induced remote instruction. In blended settings, face-to-face lectures remain the primary instructional mode, and the LMS functions as a complementary platform for distributing materials, conducting quizzes, and managing assignments. Student engagement in blended courses therefore reflects reinforcement and extension of in-class learning rather than complete reliance on digital interaction. Despite the growing adoption of blended learning in Sri Lanka, empirical studies utilizing LMS log data to examine engagement–performance relationships in physically conducted undergraduate courses remain limited. This gap underscores the need for context-specific investigations that distinguish between fully online and blended learning dynamics.

ISSN 3121-4347

Therefore, this study contributes to the existing literature by empirically examining LMS engagement and academic performance in a Sri Lankan blended learning context using large-scale LMS activity logs and linear regression analysis.

3. **Methodology**

This study was designed to examine how different forms of e-learning engagement influence student academic performance within a blended learning environment. The research was conducted with a cohort of 37 undergraduate IT students enrolled in a course delivered primarily through face-to-face lectures, supplemented by a Learning Management System (LMS). The LMS was utilized to distribute lecture materials, conduct quizzes, and manage assignment and activity submissions. To encourage continuous engagement, an online quiz related to the lecture content was administered after each physical lecture session, and students were required to attempt these quizzes within the LMS.

A total of more than 14,000 LMS activity log records were collected and analyzed for this study. These logs captured detailed information on student interactions with the LMS throughout the course duration of one semester. For analytical purposes, LMS activities were categorized into four key engagement indicators: LMS quiz attempts, assignment and activity submissions, course page views, and lecture material views. These indicators were further grouped into two engagement dimensions. LMS quiz attempts and submissions were classified as active engagement, as they require direct student participation and cognitive effort. In contrast, course page views and lecture material views were classified as passive engagement, reflecting non-interactive content consumption behaviors.

This categorization enabled a structured analysis of engagement patterns and supported the application of linear regression techniques to evaluate their relationship with academic performance.

Academic performance in this study was evaluated using students' assignment marks, which served as the primary performance indicator. Assignment assessments were selected as they reflect students' understanding and application of course concepts over an extended period rather than momentary performance. To examine the relationship between academic performance and LMS engagement, a regression analysis was conducted between

ISSN 3121-4347

assignment marks and the number of LMS views recorded for each student.

The graph presented illustrates the relationship between assignment marks and LMS view counts, representing students' passive engagement behaviors within the learning platform. Passive engagement, in this context, includes activities such as course page visits and lecture material views, where students access learning resources without direct interaction or assessment submission. The observed trend in the graph provides insight into how increased exposure to course content through LMS navigation relates to student academic performance. This analysis enables an understanding of whether passive engagement alone contributes to improved learning outcomes and supports the differentiation between passive content consumption and more active forms of engagement within blended learning environments.

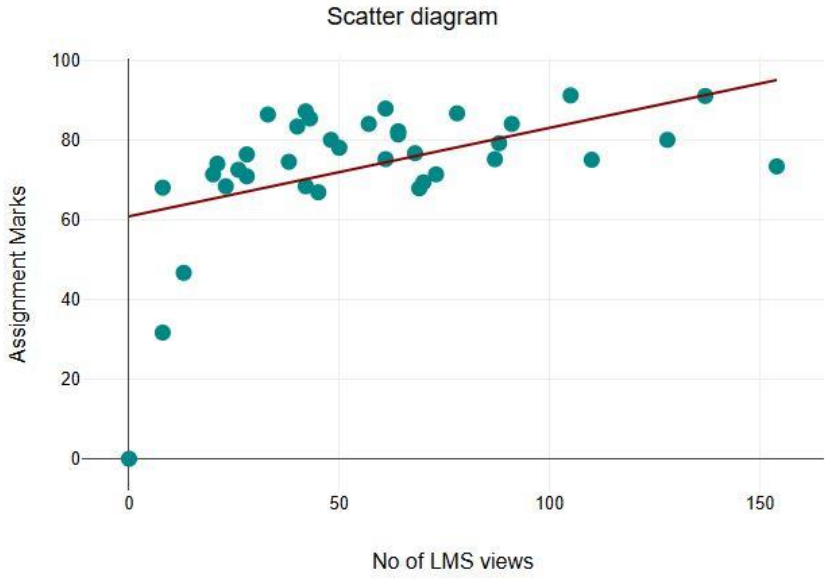

Figure 1
Relationship between Assignment marks and LMS View counts

Based on the linear regression analysis presented in figure 1, a statistically meaningful relationship was observed between LMS course views and student academic performance. Specifically, the analysis revealed a moderate positive correlation between the number of

ISSN 3121-4347

LMS course views and assignment marks (r = 0.49). The regression model examining LMS views as a predictor of assignment marks was statistically significant, β = 0.49, t(35) = 3.33, p = 0.002, explaining 24% of the variance in assignment performance (R^2 = 0.24, 95% CI [0.20, 0.70]). This indicates that students who accessed course pages and lecture materials more frequently tended to achieve higher assignment scores. Although LMS course views represent a form of passive engagement, the observed correlation suggests that regular exposure to learning resources contributes positively to academic performance. However, the strength of the relationship also indicates that LMS views alone do not fully explain variations in assignment outcomes, implying the influence of additional factors such as active engagement and individual learning behaviors. This finding supports the notion that passive engagement plays a supportive, though not exclusive, role in enhancing student academic performance within blended learning environments.

Figure 2 illustrates the relationship between students' assignment marks and the number of quiz attempts, representing active engagement within the Learning Management System. Active engagement activities require direct student participation and cognitive involvement, as students must apply their understanding of lecture content to complete quizzes and assessments. In this study, quiz attempts were administered after each lecture and were designed to reinforce learning and encourage continuous academic involvement.

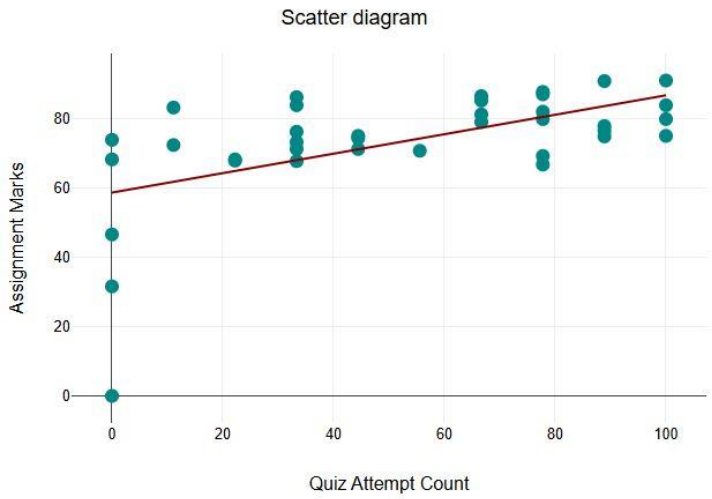

Figure 2
Relationship between Students' Assignment marks and the Number of Quiz attempts

ISSN 3121-4347

The regression analysis further revealed a stronger positive relationship between students' average quiz marks and assignment marks, with a correlation coefficient of $r = 0.57$. The regression model using quiz performance as the predictor demonstrated stronger explanatory power, $\beta = 0.57$, $t(35) = 4.10$, $p < 0.001$, accounting for approximately 33% of the variance in assignment marks ($R^2 = 0.33$, 95% CI [0.30, 0.75]). This indicates that students who performed well in quizzes were more likely to achieve higher assignment scores. Compared to passive LMS engagement indicators, quiz performance representing active engagement demonstrated a more substantial association with academic outcomes.

To quantitatively examine the relationship between active engagement and academic performance, linear regression analysis was employed. The regression model was used to assess the extent to which quiz attempts predict assignment marks and to determine the strength and direction of the relationship between these variables. By training the regression model on the observed data, the analysis provides insights into how increased active participation through quizzes contributes to improved academic outcomes. This approach enables a clear comparison between active engagement indicators and performance metrics, supporting data-driven conclusions regarding the effectiveness of assessment-based engagement strategies in blended learning environments.

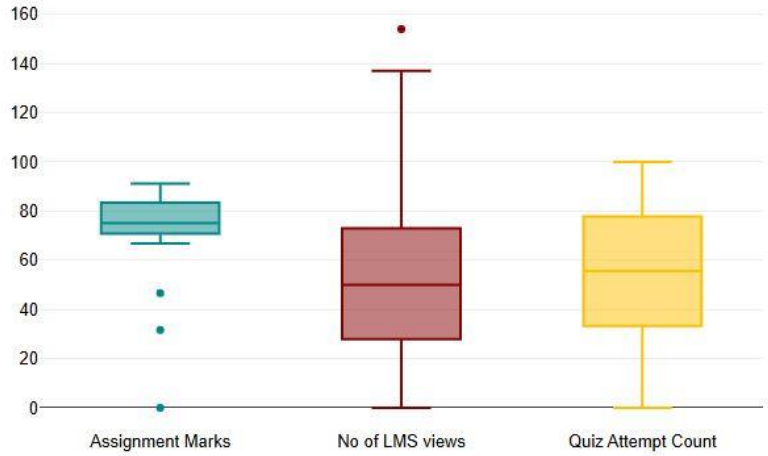

Figure 3

Box Plot; distribution of Assignment marks, Number of LMS views, and Quiz attempt

ISSN 3121-4347

counts

Figure 3: The box plot above illustrates the distribution of three key variables used in this study: assignment marks, number of LMS views, and quiz attempt counts. Assignment marks show a relatively narrow interquartile range, indicating that most students achieved scores within a consistent performance band. The median assignment mark is positioned toward the upper range of the scale, suggesting generally satisfactory academic performance across the cohort. A small number of lower-end outliers are observed, representing students with comparatively weaker performance.

In contrast, the distribution of LMS views exhibits a wider spread, with a large interquartile range and the presence of high-end outliers. This indicates substantial variability in passive engagement behaviors among students. While some students accessed course pages and lecture materials frequently, others demonstrated minimal LMS interaction. The median LMS view count lies near the center of the distribution, reflecting moderate overall usage, but the variability suggests that passive engagement alone differs significantly across learners.

Quiz attempt counts also display moderate dispersion, with a higher median compared to LMS views, indicating consistent participation in assessment-based activities. The presence of fewer extreme outliers in quiz attempts suggests more uniform active engagement among students. Overall, the box plot highlights that while passive engagement behaviors vary widely, active engagement through quizzes is more consistently distributed and aligns more closely with assignment performance. This visual evidence supports the analytical findings that active engagement indicators serve as stronger predictors of academic success than passive LMS usage alone.

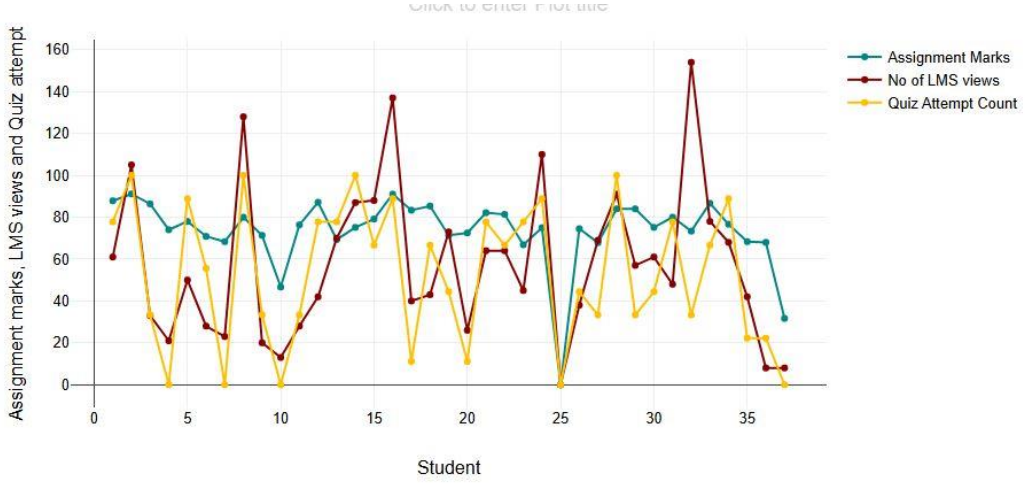

Figure 4
Variation of Assignment marks, LMS view counts, and Quiz attempt frequencies

The graph, figure 4 illustrates the distribution and variation of assignment marks, LMS view counts, and quiz attempt frequencies across the student cohort. This combined visualization provides insight into the relationship between students' e-learning engagement behaviors and their academic performance within a blended learning environment.

A general trend observed in the graph is that students with higher assignment marks tend to exhibit greater levels of active engagement, as reflected by higher quiz attempt counts. In several instances, peaks in assignment performance align closely with increased quiz participation, indicating a strong association between assessment-based engagement and improved academic outcomes. This pattern supports the premise that active engagement activities, which require direct cognitive involvement, play a significant role in reinforcing learning and enhancing performance.

Conversely, LMS view counts demonstrate considerable variability across students. While some high-performing students also recorded high numbers of LMS views, there are multiple cases where elevated LMS activity did not correspond to proportionally higher assignment marks. This suggests that passive engagement behaviors, such as accessing course pages and viewing lecture materials, contribute to learning support but are insufficient on their own to ensure higher academic achievement.

Furthermore, students with consistently low quiz attempt frequencies generally exhibit lower assignment marks, regardless of their LMS viewing activity. This observation further underscores the importance of active participation in formative assessments. Overall, the graph reinforces the study's findings that active e-learning engagement serves as a more reliable indicator of academic performance than passive LMS usage alone.

ISSN 3121-4347

4. Results

This study examined the relationship between e-learning engagement and academic performance in a blended learning environment using LMS activity data and assignment marks of 37 undergraduate IT students. More than 14,000 LMS log records were analyzed, capturing both passive and active engagement behaviors. Academic performance was measured using assignment marks, while engagement indicators included LMS course views, lecture material views, quiz attempts, and quiz marks.

Linear regression analyses revealed meaningful relationships between LMS engagement and academic performance. A moderate positive regression was identified between the independent variable LMS course views and the dependent assignment marks $\beta = 0.49$, $t(35) = 3.33$, $p = 0.002$, variance in assignment performance ($R^2 = 0.24$, 95% CI [0.20, 0.70]), indicating that students who accessed course pages and lecture materials more frequently tended to achieve higher assignment scores.

A stronger positive regression was observed between the quiz marks and the dependent variable assignment marks $\beta = 0.57$, $t(35) = 4.10$, $p < 0.001$, the variance in assignment marks ($R^2 = 0.33$, 95% CI [0.30, 0.75]). Linear regression analysis further demonstrated that quiz performance served as a stronger predictor of academic success, with higher quiz averages consistently associated with improved assignment results. These findings indicate that assessment-based engagement has a greater influence on academic performance than passive LMS usage.

The linear regression analysis reinforced this observation by quantifying the predictive strength of each engagement variable. The regression model using LMS views as the independent variable explained approximately 24% of the variance in assignment marks, suggesting a moderate explanatory capacity. In contrast, the model incorporating average quiz marks exhibited a stronger predictive relationship, with higher quiz scores consistently associated with improved assignment performance. These findings highlight that while passive engagement through LMS content access contributes positively to learning, active assessment-based engagement plays a more decisive role in predicting student academic success. This emphasizes the importance of integrating frequent formative assessments within blended learning environments to enhance student performance.

ISSN 3121-4347

5. Discussion

The comparative visualization of assignment marks, LMS view counts, and quiz attempt frequencies across students provides additional insight into engagement–performance relationships. The graph illustrates that students achieving higher assignment marks generally exhibit greater levels of active engagement, as reflected by increased quiz attempts. In several cases, peaks in assignment performance align closely with higher quiz participation, reinforcing the statistical findings that active engagement plays a critical role in enhancing learning outcomes.

In contrast, LMS view counts show substantial variability across students. While some high-performing students demonstrate frequent LMS access, there are multiple instances where high levels of passive engagement do not correspond to proportionally higher assignment marks. This suggests that accessing learning materials alone, without active participation in assessments, is insufficient to ensure academic success.
Furthermore, students with consistently low quiz attempt frequencies tend to achieve lower assignment marks regardless of their LMS viewing behavior. This observation further emphasizes the importance of active, assessment-driven engagement in reinforcing learning and supporting academic achievement. These results align with prior studies that highlight the stronger predictive power of active e-learning engagement compared to passive content consumption.

Overall, the findings support the research objective of identifying engagement behaviors that influence academic performance in a blended learning context. The study confirms that while passive LMS usage contributes to learning support, active engagement through quizzes and assessments serves as a more reliable indicator of student academic success.

6. References

Akpen, C. N., Asaolu, S., Atobatele, S., Okagbue, H., & Sampson, S. (2024). Impact of online learning on student's performance and engagement: a systematic review. Discover Education, 3(1). https://doi.org/10.1007/s44217-024-00253-0

Alam, M. M., Ahmad, N., Naveed, Q. N., Patel, A., Abohashrh, M., & Khaleel, M. A.

(2021). E-Learning Services to Achieve Sustainable Learning and Academic Performance: An Empirical Study. Sustainability, 13(5), 2653. https://doi.org/10.3390/su13052653

Rasheed, H. M. W., He, Y., Khalid, J., Khizar, H. M. U., & Sharif, S. (2020). The relationship between e-learning and academic performance of students. Journal of Public Affairs, 22(3). https://doi.org/10.1002/pa.2492

Tolika, M., Karapiperis, D., Tsoni, R., & Verykios, V. S. (2024). Unpacking the Impact of E-Learning Participation on Academic Performance: An experimental study. Διεθνές Συνέδριο Για Την Ανοικτή & Εξ Αποστάσεως Εκπαίδευση, 12(3), 168–179. https://doi.org/10.12681/icodl.5712

Naveed, Q. N., Qureshi, M. R. N. M., Shaikh, A., Alsayed, A., Sanober, S., & Mohiuddin, K. (2019). Evaluating and ranking cloud-based e-learning critical success factors (CSFs) using combinatorial approach. IEEE Access, 7, 157145–157157. https://doi.org/10.1109/ACCESS.2019.2949044

Elfaki, N. (2020). Impact of e-learning vs traditional learning on students' performance and attitude. ResearchGate. https://doi.org/10.13140/RG.2.2.14811.59681

Kokoç, M., & Altun, A. (2019). Effects of learner interaction with learning dashboards on academic performance in an e-learning environment. Behaviour and Information Technology, 40(2), 161–175. https://doi.org/10.1080/0144929x.2019.1680731

Ashrafi, A., Zareravasan, A., Savoji, S. R., & Amani, M. (2020). Exploring factors influencing students' continuance intention to use the learning management system (LMS): a multi-perspective framework. Interactive Learning Environments, 30(8), 1475–1497. https://doi.org/10.1080/10494820.2020.1734028

Bansah, A. K., & Agyei, D. D. (2022). Perceived convenience, usefulness, effectiveness and user acceptance of information technology: evaluating students' experiences of a Learning Management System. Technology Pedagogy and Education, 31(4), 431–449. https://doi.org/10.1080/1475939x.2022.2027267

Shafiei Sarvestani, M., Mohammadi, M., Afshin, J., Raeisy, L., & Department of Educational Administration and Planning, Faculty of Education & Psychology, Shiraz University, Shiraz, Iran. (2019). Students' Experiences of E-Learning Challenges; A phenomenological study. In Interdisciplinary Journal of Virtual Learning in Medical Sciences (Vols. 10–10, Issue 3, pp. 1–10). https://doi.org/10.30476/IJVLMS.2019.45841

Shurygin, V. Y., & Sabirova, F. M. (2017, August 31). Particularities of blended learning implementation in teaching physics by means of LMS Moodle.

ISSN 3121-4347

http://www.revistaespacios.com/a17v38n40/17384039.html

Xue, S., & Churchill, D. (2019). A review of empirical studies of affordances and development of a framework for educational adoption of mobile social media. Educational Technology Research and Development, 67(5), 1231–1257. https://doi.org/10.1007/s11423-019-09679-y

ISSN 3121-4347

JASPER Research Journal
Volume 1 Issue 1
March 2026

LANKA NIPPON BIZTECH INSTITUTE

Japanese IT University

278, HIGH LEVEL ROAD,
MAHARAGAMA, SRI LANKA.

011 7 370 620

ISSN 3121-4347