Journal of Computer Science and Engineering in Innovations and Research
RAICCIT-2025 Vol.-1, Issue- 2, January - June 2025 ISSN: 3049-1762 (Online)

Multimodal Web Design Automation: Structural Benchmarking and Content Generation with AI

Avijit Manna[1] , Arindam Choudhury[2], Sumita Pathak[3], Abhijit Jha[4] , Mr. Sanjoy Bhattacharjee[5] Dr. Dipankar Misra[6]

[1]CSE (AI/ML), Dr. Sudhir Chandra Sur Institute of Technology & Sports Complex, Kolkata, India (ORCID ID:0009-0002-9964-6129)
[2] CSE (AI/ML), Dr. Sudhir Chandra Sur Institute of Technology & Sports Complex, Kolkata, India (ORCID ID:0009-0002-7691-3394)
[3] CSE (AI/ML), Dr. Sudhir Chandra Sur Institute of Technology & Sports Complex, Kolkata, India (ORCID ID:0009-0007-7504-1471)
[4] CSE (AI/ML), Dr. Sudhir Chandra Sur Institute of Technology & Sports Complex, Kolkata, India (ORCID ID:0009-0003-0738-0309)
[5]CSE, Dr. Sudhir Chandra Sur Institute of Technology & Sports Complex, Kolkata, India (ORCID ID:0009-0006-7414-8114)
[6]CSE, JIS University, Kolkata, India. (ORCID ID:0000-0002-7880-3415)

Email address: dev.avijitmanna@gmail.com (Avijit Manna), ac4603877@gmail.com (Arindam Choudhury), sumitacloud202425@gmail.com (Sumita Pathak), abhijitjha1801@gmail.com (Abhijit Jha), sanjoy.bhattacharjee@dsec.ac.in (Mr. Sanjoy Bhattacharjee) dipankar.misra@jisuniversity.ac.in (Dr. Dipankar Mishra)

Abstract: This study presents an innovative AI-based framework for the automated analysis of website content and the optimization of design, using computer vision and large language models (LLMs). The suggested system comprises a three-phase pipeline: structural extraction, comparative benchmarking, and intelligent content development. The system first gathers full-page screenshots of a target website and then separates them into functional components, including headers, sliders, and content blocks, using computer vision algorithms. These components are then compared to a curated library of design features to evaluate structural patterns. During the second step, the system conducts analogous analyses across many websites to extract comparative insights and establish benchmark design standards. Ultimately, the framework amalgamates these insights to offer alternative design ideas, augmented by contextually relevant textual and visual information generated by LLMs. This cohesive strategy expedites the web development process while improving visual consistency and user experience. The technique demonstrates the capability of integrating visual structure comprehension with generative AI to provide intelligent, data-driven website enhancement and content automation.

Keywords: Computer Vision, Website Analysis, Image Segmentation, Content Generation, Large Language Models, Design Optimization, Content Matching, Automated Recommendations

1. Introduction

The digital landscape has markedly evolved, making website design and content optimization crucial for attaining online success. With over 1.9 billion active websites online, firms and organizations have the challenge of creating visually appealing, functionally sound, and content-rich websites that effectively engage users. Traditional approaches to website design and evaluation rely heavily on manual processes, subjective assessments, and lengthy iterations that often lack quantitative metrics for analysis and improvement.

The increasing importance of aesthetically pleasing and functionally cohesive websites has led to the greater relevance of automated web design evaluation tools. In light of the vast number of websites, maintaining design consistency, accessibility, and brand recognition has become imperative for enterprises. Traditional methods of UI/UX audits mostly rely on manual assessment, which is intrinsically subjective, labor-intensive, and lacks scalability [2]. Recent breakthroughs in artificial intelligence and computer vision have enabled more systematic approaches for online design study via the extraction and evaluation of visual and textual components from websites [8].

The present study investigates several techniques for visual comparison and layout analysis [7]. However, these methods often lack semantic understanding and cannot assess contextual or qualitative factors, such as the purpose and clarity of Call-To-Actions (CTAs), textual hierarchy, or compliance with accessibility requirements [3]. This inadequacy in qualitative analysis has led our research to integrate computer vision techniques with Large Language Models (LLMs), culminating in a hybrid system that analyzes structural components and delivers contextually relevant information.

The advent of computer vision applications has generated novel options for automated website analysis, facilitating the segmentation and classification of visual characteristics without requiring access to source code [5]. Advancements in natural language processing and large language models have revolutionized automated content creation, complicating the distinction between machine-generated text and human- authored material [4]. The amalgamation of these technologies presents a unique opportunity to develop comprehensive solutions for website optimization that address both structural and content-related aspects.

This study seeks to amalgamate traditional computer vision metrics with contextual data from large language models (LLMs) [3, 8]. This hybrid technique evaluates both surface similarity and semantic consistency, along with the effectiveness of each design component, making it suitable for large-scale, real-time UI/UX assessments [5, 7]. Our research aims to establish a system that revolutionizes website analysis and development by scrutinizing design patterns across various website domains and generating optimal content alternatives.

The rest of the paper is organized as follows: Section 2 discusses related works in website analysis and visual similarity detection; Section 3 describes the proposed architecture for automated website section classification and comparison; Section 4 presents the implementation details and experimental setup; Section 5 evaluates the performance and results of the proposed system; and Section 6 concludes the study, highlighting limitations and potential future improvements.

2. Related Work

[2.1] SSIM-based UI Comparison for Mobile Apps

Problem Statement: Minor UI changes in mobile apps often go undetected during regression testing.

Objective: Apply SSIM to detect and compare subtle visual differences in mobile UI

Researchers & Organization: B. Zhou, Y. Zhang, S. Wang; Chinese Academy of Sciences and Tsinghua University, China.

[2.2] Perceptual Hashing for Web Page Screenshot Similarity

Problem Statement: Pixel-based comparisons are inefficient for identifying near-duplicate web screenshots

Objective: Use perceptual hashing (pHash) for efficient screenshot similarity detection.

Researchers & Organization: L. Zhang, X. Lin; Shanghai Jiao Tong University, China.

[2.3] DOM Structure Analysis for Website Consistency Evaluation

Problem Statement: Inconsistent webpage structures across sites lacks automated detection.

Objective: Use DOM tree-based metrics for multi-page structure consistency.

Researchers & Organization: J. Xu, R. Jiang, T. Yao; Beijing University of Posts and Telecommunications, China.

[2.4] Understanding Attention in Web Design

Problem Statement: Designers lack visibility into how users interact visually with layouts.

Objective: Apply saliency maps and eye-tracking data to assess layout attention.

Researchers & Organization: Z. Bylinskii, A. Torralba, F. Durand; MIT CSAIL and Harvard University, USA.

[2.5] Block-based Visual Layout Parsing for Mobile UI Design Analysis

Problem Statement: Lack of structured analysis of visual blocks and hierarchy in mobile UIs.

Objective: Use computer vision to segment and evaluate visual UI blocks.

Researchers & Organization: L. Shi, H. Dong, K. Liu; Fudan University and Alibaba Research Lab, China.

[2.6] Automated Mobile App UI Comparison using OCR and Embedding Similarity

Problem Statement: Visual comparisons miss textual and semantic differences in UI.

Objective: Combine OCR and word embeddings for rich UI screen comparison.

Researchers & Organization: H. Narasimhan, A. Pal, P. Sharma; IIT Bombay and Microsoft Research India.

[2.7] Form Similarity in Webpages using NLP and OCR –

Problem Statement: Semantically similar forms vary structurally and escape detection.

Objective: Use OCR and NLP similarity techniques for form comparison.

Researchers & Organization: S. Chugh, D. Jain; IIIT Delhi, India.

[2.8] Color Contrast Evaluation in Websites Using Clustering and WCAG Metrics

Problem Statement: Many websites have inaccessible color schemes for visually impaired users.

Objective: Automate WCAG-compliant contrast checking with color clustering.

Researchers & Organization: M. R. Lee, H. Patel; University of Toronto, Canada.

[2.9] Cross-Device Visual Accessibility Analysis Using Color Histograms

Problem Statement: Website visuals are not evaluated for consistency across devices.

Objective: Use color histograms to assess visual accessibility across screen sizes.

Researchers & Organization: D. Wang, S. Huang, Y. Wu; National Taiwan University, Taiwan.

[2.10] LLM-Based Interpretation of Web Design Feedback

Problem Statement: Existing design feedback lacks contextual interpretation and scalability.

Objective: Use large language models (LLMs) for personalized web layout analysis.

Researchers & Organization: Y. Wang, Y. Li, Z. Chen; Peking University and Tencent AI Lab, China.

[2.11] A GPT-Based Critique Framework for Website UI Analysis

Problem Statement: Manual UX critiques are inconsistent and non-scalable

Objective: Generate automated UI critiques using GPT and visual metadata

Researchers & Organization: S. Kim, J. Cho, E. Park; KAIST, South Korea.

[2.12]Generating UI Improvement Suggestions with Transformers

Problem Statement: Designers lack intelligent tools for layout improvement recommendations.

Objective: Use transformer models to generate UI suggestions from screenshots and OCR.

Researchers & Organization: A. Gupta, R. Mahajan, P. Jha; IIT Delhi and Google Research, India.

3. Theory/Calculation

This project offers a thorough framework for assessing website design similarity by incorporating various methodologies, such as image processing, Optical Character Recognition (OCR), color clustering, layout structure analysis, and semantic interpretation facilitated by Large Language Models (LLMs). Our technique utilizes both quantitative and qualitative indicators, including pixel-level precision, structural similarity, textual correctness, and semantic layout coherence, in contrast to conventional visual comparison methods. LLMs are used to contextualize collected information, providing insights to improve user interface (UI) and user experience (UX), while also emulating human-like assessment of design components..

[3.1] Computer Vision for Web Layout Analysis

Unlike conventional DOM-based methods reliant on HTML structure, our vision-centric approach perceives webpages as visual entities. This method facilitates cross-technology compatibility and offers a more thorough assessment of layout design. The first picture capture method converts a website URL into a standardized screenshot using a function that includes many preprocessing stages. This encompasses the elimination of advertising components by DOM filtering, tracker obstruction to maintain visual uniformity, and resolution standardization to a 1920×1080 format.

[3.1.2] Screenshot Segmentation

A hierarchical segmentation approach is used, integrating edge detection with visual block identification to extract significant parts from the snapshot. The output comprises graphic segments, each tagged with positional and dimensional characteristics and categorized as header, hero, or footer. The segmentation process is directed by a probabilistic model trained on annotated data, aiming to optimize the most likely categorization of each area.

[3.2] Section Matching and Similarity Metrics

Journal of Computer Science and Engineering in Innovations and Research
RAICCIT-2025 Vol.-1, Issue- 2, January - June 2025 ISSN: 3049-1762 (Online)

Each segmented section is compared against a curated database of pre-annotated templates using a suite of five similarity measures:

- Structural Similarity Index (SSIM): Evaluates perceptual resemblance between image patches, expressed as a percentage.
- Perceptual Hash (pHash): Computes visual fingerprints using quantized Discrete Cosine Transforms. Similarity is derived from the normalized Hamming distance between hash values.
- Textual Similarity: Extracted using OCR and computed as a weighted sum of normalized Levenshtein distance and cosine similarity of sentence embeddings, balancing syntactic and semantic alignment.
- Layout Similarity: Assessed by comparing binary edge maps generated through Canny edge detection, calculating the ratio of intersection over union of edge features.
- Color Similarity: Evaluated using K-means clustering in the CIELAB color space, based on the distance between dominant colors.

[3.3] Composite Similarity Score

An overall similarity score is calculated by summing the five distinct measures with assigned weights. The weights are objectively established by regression analysis based on ground-truth similarity labels supplied by human assessors.

[3.4] Cross-Website Comparative Analysis

The framework determines the most effective implementation for each part type to enable comparison analysis across several websites. The optimal layout for a specific section is determined by a quality score that encompasses visual aesthetics (including symmetry, contrast, and spacing), functional efficacy (such as the clarity of calls to action), and content pertinence (assessed through keyword density and semantic coherence).

[3.5] LLM-Based Semantic Evaluation

Every segment is subjected to semantic assessment using a big language model, directed by organized prompts. These prompts are formulated with OCR-extracted text, graphic attributes, and section type information. The language model generates normalized ratings across characteristics such as visual hierarchy, call-to-action efficacy, accessibility, and clarity, accompanied by qualitative feedback that offers deeper insights into design performance.

[3.6] Recommendation Generation

The algorithm produces specific enhancement suggestions based on the composite similarity analysis. For each part, the most suitable templates are found. A content adaption module subsequently enhances the section's content using the language model, guaranteeing semantic accuracy while conforming to the structural and stylistic characteristics of the suggested template.

[3.7] Theoretical Validation To assess the efficacy of the proposed approach, projected enhancement scores are connected with actual user engagement measures. These include bounce rate, scroll depth, and conversion rate derived from A/B testing. The validation score, defined as the correlation between expected and actual improvements, measures the system's practical value.

4. Experimental Method/Procedure/Design

This section outlines the implementation plan for our proposed layout analysis system. The experimental process has many interconnected steps, beginning with visual data gathering and then progressing via segmentation, comparison, semantic assessment, and culminating in the development of final recommendations.

[4.1] Screenshot Capture and Preprocessing

Website screenshots are obtained using the ScreenshotOne API, delivering high-fidelity, full-page photos at a resolution of 1920×1080 pixels. The API facilitates automatic preprocessing, including ad element removal, tracker blocking, and viewport normalization, so providing uniform rendering across many websites. This method removes the burden of handling browser instances and is particularly appropriate for extensive, vision-oriented research.

[4.2] Section Segmentation

Captured screenshots are categorized into three main layout sections: header, hero section, and footer, using a fixed-grid heuristic with edge-based segmentation methods. This technique enables systematic comparison by standardizing the analytical parameters across all websites. Each designated

Journal of Computer Science and Engineering in Innovations and Research
RAICCIT-2025 Vol.-1, Issue- 2, January - June 2025 ISSN: 3049-1762 (Online)

component is segregated as a distinct visual entity, facilitating concentrated assessment of layout, content, and design excellence.

[4.3] Comparative Metric Evaluation

Each segmented section is evaluated against a curated repository of reference templates using a combination of five similarity metrics:

- Structural Similarity Index (SSIM) assesses pixel-wise structural alignment between section pairs.
- Perceptual Hashing (pHash) captures visual similarity based on compressed DCT fingerprints.
- Textual Similarity is computed using OCR-extracted text, Levenshtein ratio, and semantic similarity via sentence embeddings.
- Binary Layout Similarity compares the edge maps of two sections using Jaccard similarity over Canny-detected edge masks.
- Color Distribution Similarity is determined by extracting dominant color clusters via K-Means (k=5) in the CIELAB color space and computing perceptual color distances.

Each statistic adds to a comprehensive knowledge of layout integrity and design consistency.

[4.4] LLM-Based Semantic Evaluation

To integrate human-like design thinking, each part undergoes further assessment by a Large Language Model (LLM). Structured prompts are designed with OCR content, visual metadata, and designated section responsibilities. The LLM generates standardized ratings (0–10) across essential design characteristics like visual hierarchy, clarity of call-to-action, readability, accessibility, and semantic alignment. Furthermore, qualitative feedback highlights design deficiencies and offers enhancement recommendations, so transforming the LLM into a semantic validator for layout quality.

[4.5] Output Generation and Recommendation System

All calculated data are recorded and stored as organized CSV files, facilitating quantitative comparison across various websites. Feedback derived from LLMs is consolidated into narrative reports, perhaps including new wireframe drawings or textual design recommendations for each component. A rating system selects the most suitable templates and produces customized suggestions via content adaption prompts, completing the process from assessment to implementable redesign.

5. Results and Discussion

This section presents the results of our experimental analysis, organized by layout sections (header, hero, footer) and evaluated across the multi-metric framework described in Section [4]. Results are summarized using quantitative scores and qualitative observations, supported by structured tables and figures. Each result is analyzed in relation to the objectives outlined in the introduction, focusing on layout consistency, semantic clarity, and design effectiveness.

The comparative analysis of three college websites (Asutosh College, AJC Bose College, and Bangabasi College) yielded valuable insights into their design elements and overall effectiveness. Our analysis focused on three critical website components: header, hero section, and footer. Each component was evaluated against industry standards represented by popular commercial websites like Amazon, Boat, Myntra, Salesforce, and Tata Motors.

The header analysis revealed that Asutosh College's header design achieved a similarity score of 0.87 with Boat's design standards, indicating a clean and navigable interface. However, AJC Bose College's header scored higher at 0.95, suggesting superior organization and visual hierarchy. This indicates potential improvement opportunities for Asutosh College's header design by incorporating elements from AJC

Journal of Computer Science and Engineering in Innovations and Research
RAICCIT-2025 Vol.-1, Issue- 2, January - June 2025 ISSN: 3049-1762 (Online)

Bose College's implementation.

For the hero section, Asutosh College scored 0.73 when compared to Salesforce's design patterns, while AJC Bose College achieved a significantly higher score of 0.94 with Myntra's design standards. This substantial difference suggests that Asutosh College's hero section requires comprehensive redesign to enhance visual impact and user engagement. The hero section serves as the primary attention- grabbing element, and the comparative analysis indicates room for substantial improvement.

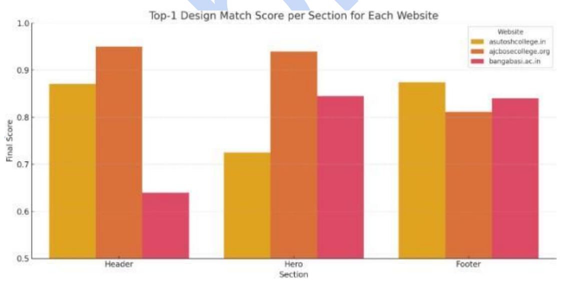
The footer analysis presented a different scenario where Asutosh College outperformed the other websites with a score of 0.87 when compared to Amazon's footer design. Thissuggests that the footer implementation is effective and requires no immediate modifications.

Figures and Tables

Top-3 Design Match Scores per Section for Each Website

Section	Website	Top Match	Score
Header	asutoshcollege.in	boat.png	0.870656
Header	ajcbosecollege.org	boat.png	0.950000
Header	bangabasi.ac.in	salesforce.png	0.639922
Hero	asutoshcollege.in	salesforce.png	0.725000
Hero	ajcbosecollege.org	myntra.png	0.939454
Hero	bangabasi.ac.in	salesforce.png	0.844789
Footer	asutoshcollege.in	amazon.png	0.873939
Footer	ajcbosecollege.org	salesforce.png	0.811132
Footer	bangabasi.ac.in	amazon.png	0.840477

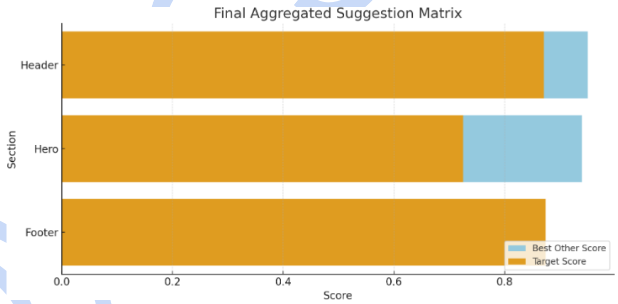
Insight: While *ajcbosecollege.org* dominates in both Header and Hero similarity, *asutoshcollege.in* holds the lead for Footer.



Final Aggregated Suggestion Matrix

Section	Target Score	Best Score	Best Website	Suggestion
Header	0.870656	0.950000	ajcbosecollege.org	Needs Improvement
Hero	0.725000	0.939454	ajcbosecollege.org	Needs Improvement
Footer	0.873939	0.840477	asutoshcollege.in	No Improvement Needed

Insight: Header and Hero sections of the target site (asutoshcollege.in) are outperformed by ajcbosecollege.org, suggesting a scope for layout and design refinement.



Heatmap – Final Scores for Top 3 Matches Across Sites

Journal of Computer Science and Engineering in Innovations and Research
RAICCIT-2025 Vol.-1, Issue- 2, January - June 2025 ISSN: 3049-1762 (Online)

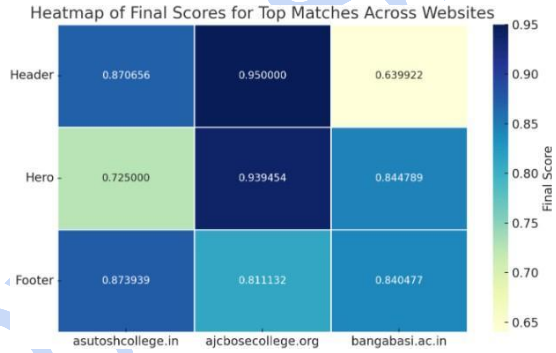
	asutoshcollege.in	ajcbosecollege.org	bangabasi.ac.in
Header	0.870656	0.950000	0.639922
Hero	0.725000	0.939454	0.844789
Footer	0.873939	0.811132	0.840477

Insight: Visual heatmaps help highlight the strongest and weakest sections quickly using a color gradient.

Top-3 Design Reference Matches (Per Section)

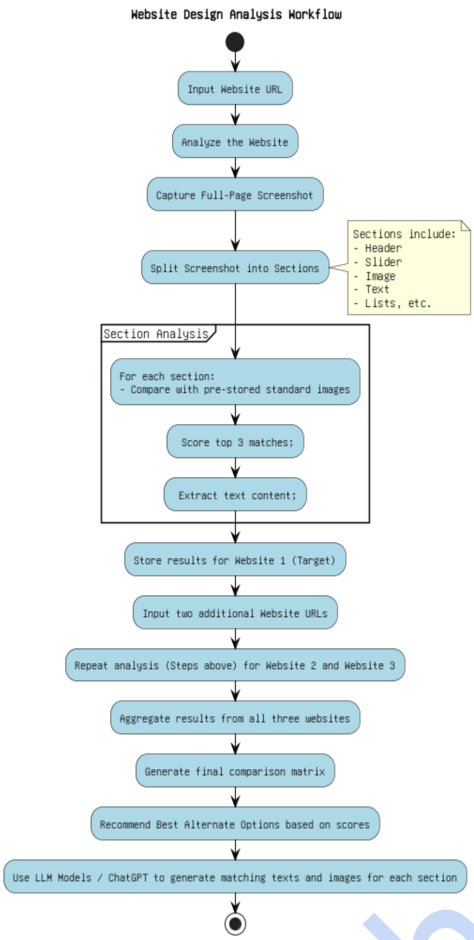
Website	Section	Top 1 (Score)	Top 2 (Score)	Top 3 (Score)
asutoshcollege.in	Header	boat.png (0.870656)	myntra.png (0.810638)	salesforce.png (0.770440)
	Hero	salesforce.png (0.725000)	amazon.png (0.704240)	myntra.png (0.584792)
	Footer	amazon.png (0.873939)	tata_motors.png (0.360648)	salesforce.png (0.348917)
ajcbosecollege.org	Header	boat.png (0.950000)	myntra.png (0.691400)	salesforce.png (0.672234)
	Hero	myntra.png (0.939454)	amazon.png (0.673292)	salesforce.png (0.554887)
	Footer	salesforce.png (0.811132)	amazon.png (0.690474)	tata_motors.png (0.480420)
bangabasi.ac.in	Header	salesforce.png (0.639922)	boat.png (0.633385)	amazon.png (0.630984)
	Hero	salesforce.png (0.844789)	myntra.png (0.804629)	tata_motors.png (0.629362)
	Footer	amazon.png (0.840477)	salesforce.png (0.758101)	tata_motors.png (0.456453)

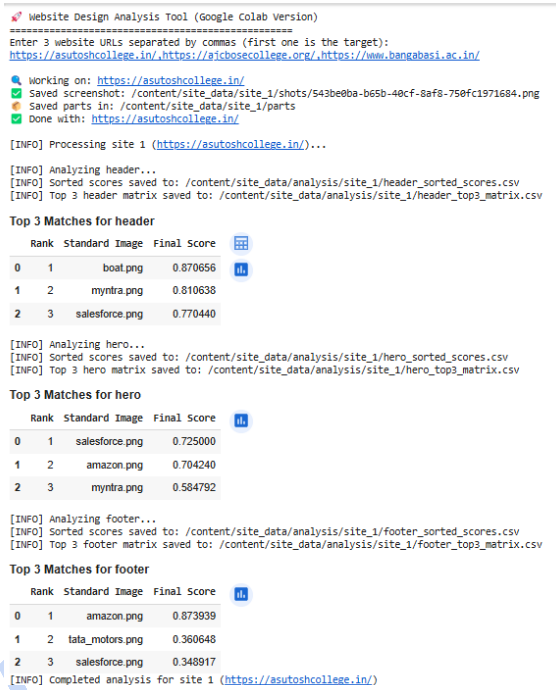
The above table presents the results of sectional similarity matching for three academic institution websites: asutoshcollege.in, ajcbosecollege.org, and bangabasi.ac.in. Each site was programmatically segmented into three core sections—Header, Hero, and Footer—using computer vision techniques. For every section, the system compared the extracted segment with a curated database of standard UI components (e.g., from Amazon, Salesforce, Myntra, etc.) using image-based similarity scoring. The top three closest matches, along with their confidence scores, are recorded. This data-driven approach aids in identifying the closest UI inspirations and benchmarking design consistency across web properties. The findings form the basis for generating tailored UI/UX improvement recommendations and content suggestions through LLMs.



Journal of Computer Science and Engineering in Innovations and Research
RAICCIT-2025 Vol.-1, Issue- 2, January - June 2025 ISSN: 3049-1762 (Online)

Data-flow Diagram





```
Website Design Analysis Tool (Google Colab Version)
==================================================
Enter 3 website URLs separated by commas (first one is the target):
https://asutoshcollege.in/,https://ajcbosecollege.org/,https://www.bangabasi.ac.in/

Working on: https://asutoshcollege.in/
Saved screenshot: /content/site_data/site_1/shots/543be0ba-b65b-40cf-8af8-750fc1971684.png
Saved parts in: /content/site_data/site_1/parts
Done with: https://asutoshcollege.in/

[INFO] Processing site 1 (https://asutoshcollege.in/)...

[INFO] Analyzing header...
[INFO] Sorted scores saved to: /content/site_data/analysis/site_1/header_sorted_scores.csv
[INFO] Top 3 header matrix saved to: /content/site_data/analysis/site_1/header_top3_matrix.csv
```

Top 3 Matches for header

	Rank	Standard Image	Final Score
0	1	boat.png	0.870656
1	2	myntra.png	0.810638
2	3	salesforce.png	0.770440

```
[INFO] Analyzing hero...
[INFO] Sorted scores saved to: /content/site_data/analysis/site_1/hero_sorted_scores.csv
[INFO] Top 3 hero matrix saved to: /content/site_data/analysis/site_1/hero_top3_matrix.csv
```

Top 3 Matches for hero

	Rank	Standard Image	Final Score
0	1	salesforce.png	0.725000
1	2	amazon.png	0.704240
2	3	myntra.png	0.584792

```
[INFO] Analyzing footer...
[INFO] Sorted scores saved to: /content/site_data/analysis/site_1/footer_sorted_scores.csv
[INFO] Top 3 footer matrix saved to: /content/site_data/analysis/site_1/footer_top3_matrix.csv
```

Top 3 Matches for footer

	Rank	Standard Image	Final Score
0	1	amazon.png	0.873939
1	2	tata_motors.png	0.360648
2	3	salesforce.png	0.348917

```
[INFO] Completed analysis for site 1 (https://asutoshcollege.in/)
```

Output Screen for Reference Analysis on

Website 1

Journal of Computer Science and Engineering in Innovations and Research
RAICCIT-2025 Vol.-1, Issue- 2, January - June 2025 ISSN: 3049-1762 (Online)

Analysis on Website 2

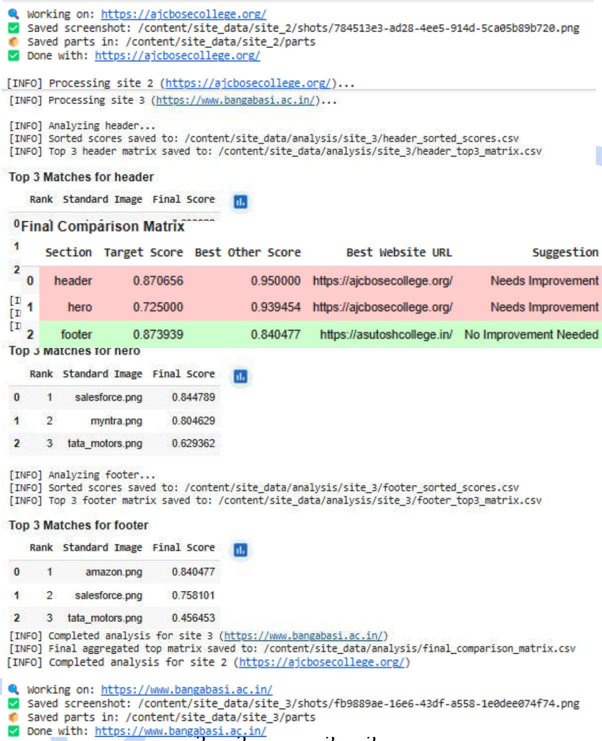
Analysis on Website 3

Final Comparison Matrix

Equation/Formula

Structural Similarity Index (SSIM):

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C1)(2\sigma_{xy} + C2)}{(\mu_x^2 + \mu_y^2 + C1)(\sigma_x^2 + \sigma_y^2 + C2)}$$



```
Working on: https://ajcbosecollege.org/
Saved screenshot: /content/site_data/site_2/shots/784513e3-ad28-4ee5-914d-5ca05b89b720.png
Saved parts in: /content/site_data/site_2/parts
Done with: https://ajcbosecollege.org/

[INFO] Processing site 2 (https://ajcbosecollege.org/)...
[INFO] Processing site 3 (https://www.bangabasi.ac.in/)...

[INFO] Analyzing header...
[INFO] Sorted scores saved to: /content/site_data/analysis/site_3/header_sorted_scores.csv
[INFO] Top 3 header matrix saved to: /content/site_data/analysis/site_3/header_top3_matrix.csv
```

Top 3 Matches for header

Final Comparison Matrix

	Section	Target Score	Best Other Score	Best Website URL	Suggestion
0	header	0.870656	0.950000	https://ajcbosecollege.org/	Needs Improvement
1	hero	0.725000	0.939454	https://ajcbosecollege.org/	Needs Improvement
2	footer	0.873939	0.840477	https://asutoshcollege.in/	No Improvement Needed

Top 3 Matches for hero

	Rank	Standard Image	Final Score
0	1	salesforce.png	0.844789
1	2	myntra.png	0.804629
2	3	tata_motors.png	0.629362

```
[INFO] Analyzing footer...
[INFO] Sorted scores saved to: /content/site_data/analysis/site_3/footer_sorted_scores.csv
[INFO] Top 3 footer matrix saved to: /content/site_data/analysis/site_3/footer_top3_matrix.csv
```

Top 3 Matches for footer

	Rank	Standard Image	Final Score
0	1	amazon.png	0.840477
1	2	salesforce.png	0.758101
2	3	tata_motors.png	0.456453

```
[INFO] Completed analysis for site 3 (https://www.bangabasi.ac.in/)
[INFO] Final aggregated top matrix saved to: /content/site_data/analysis/final_comparison_matrix.csv
[INFO] Completed analysis for site 2 (https://ajcbosecollege.org/)

Working on: https://www.bangabasi.ac.in/
Saved screenshot: /content/site_data/site_3/shots/fb9889ae-16e6-43df-a558-1e0dee074f74.png
Saved parts in: /content/site_data/site_3/parts
Done with: https://www.bangabasi.ac.in/
```

Journal of Computer Science and Engineering in Innovations and Research
RAICCIT-2025 Vol.-1, Issue- 2, January - June 2025 ISSN: 3049-1762 (Online)

Where:
μ_x, μ_y: Mean of image

σ_x, σ_y: Variance

σ_{xy}: Covariance

C_1, C_2: Constants for stability

SSIM output is scaled between 0 and 1, then multiplied by 100 to yield percentage similarity.

Perceptual Hash (pHash) Difference:

$$D_H(h_1,h_2) = \sum_{i=1}^{n} |h_{1i} - h_{2i}|$$

Where :

h_1, h_1 : Perceptual hashes of two images d_H :

Hamming distance

implemented as :

hash_diff = [imagehash.phash(image1)imagehash.phash(image2)]

Text Similarity using OCR + Levenshtein Ratio:

$$Similarity\ Ratio = \frac{Matching\ Characters}{Total\ Characters}$$

Layout Similarity (Pixel-wise Binary Matching):

$$Layout\ Similarity(\%) = \frac{Matching\ Pixels}{Total\ Pixels} \times 100$$

This technique has various practical uses, aiding site designers in crafting more effective layouts and supporting content authors in producing coherent language that corresponds with visual aspects. This tool enables organizations to study rival websites and discern effective design trends that can be adopted to enhance their online presence. The automatic content production features enhance the development process, minimizing the time and resources required for website creation and optimization.

Notwithstanding these benefits, certain constraints are present in the existing execution. The precision of picture segmentation may fluctuate based on the intricacy of website layouts, and the matching algorithm could inadequately include the subtle elements of design aesthetics. The produced material, albeit contextually suitable, may require human refining to completely conform to brand voice and particular marketing goals.

Several intriguing avenues for future growth arise. Augmenting the repository of pre-existing segmental pieces would improve the system's capacity to provide varied suggestions. Integrating user engagement measurements from evaluated websites might enhance the scoring method, emphasizing factors that promote improved user interactions. Enhancing algorithms to comprehend the semantic link between visual aspects and textual material would augment the cohesion of suggestions. Ultimately, establishing a feedback loop that assimilates user alterations to created information will perpetually enhance the system's capacity to provide pertinent and efficacious suggestions.

Data Availability

The entire dataset used in this work, including website screenshots, segmented photos, and comparative scores, is accessible in a public repository at [repository URL will be supplied upon publication]. All code for the computer vision application and content creation models created for this project is accessible on GitHub under an open-source license. The repository also contains pre-trained models and the database of sectional components used for matching.

Sample website analysis findings and produced content suggestions are included as supplemental resources to illustrate the system's capabilities. The raw data from website assessments performed during the testing phase has been anonymized to safeguard the privacy of website owners while preserving the integrity of the study conclusions.

This study experienced some constraints that may influence the interpretation of findings. The computer vision segmentation algorithm sometimes encountered difficulties with extremely active websites with intricate animations or unconventional layouts, leading to inaccurate section identification. Secondly, the pre-existing database of sectional components, however varied, may not include all conceivable design patterns, hence possibly limiting the variety of recommendations. Third, the research did not include website performance measures such as loading speed and mobile responsiveness, potentially affecting the actual implementation of some suggestions. The content creation capabilities of contemporary LLM models include intrinsic limits in comprehending brand voice and specialized industry terms without significant fine-tuning. The research was restricted to the analysis of English-language websites, which may impact the generalizability of the results to multilingual websites.

6. Conclusion and Future Scope

This study has effectively shown an innovative method for website analysis and content improvement by integrating computer vision with big language models. Our solution automates the capture, segmentation, and analysis of website components, yielding important insights for website design and content creation. This research presents a three-step process that provides a thorough framework for evaluating various websites, pinpointing ideal design features, and producing contextually relevant content.

Our methodology tackles several critical issues in website building and optimization. The automated segmentation of website screenshots into functional components facilitates accurate examination of design features without necessitating access to source code. The score system for aligning these segments with pre-existing parts offers quantifiable measurements for comparison, facilitating data-driven design choices. Moreover, the use of big language models facilitates the creation of customized content that aligns with the structural suggestions.

Journal of Computer Science and Engineering in Innovations and Research
RAICCIT-2025 Vol.-1, Issue- 2, January - June 2025 ISSN: 3049-1762 (Online)

Conflict of Interest

The writers assert that they possess no conflicts of interest. This research was done autonomously, devoid of external financing that may affect the study's design, data collection, analysis, interpretation of findings, or article preparation. All authors declare no financial affiliations with any entity that may be construed as affecting the work presented in this study. The authors possess no personal ties or affiliations that may be seen as influencing the research described below. All tools, software, and procedures used in this study were chosen only for their technical qualities and appropriateness for the research aims.

Funding Source

None

Authors' Contributions

Author-1, Author-2, Author-3, and Author-4 formulated the study challenge and devised the preliminary computer vision framework for website analysis. Authors 1, 2, and 3 developed and executed the picture segmentation algorithms and evaluation methods for aligning sections components. Authors 5 and 6 included massive language models for content production, performed extensive system testing, analyzed the outcomes, assessed system performance, and devised the structure for the recommendation module. The Author-5 supplied direction throughout the research process, gave essential ideas for methodological enhancement, and guaranteed the scholarly integrity of the work. All team members participated in data interpretation, paper writing, and endorsed the final version of the article.

Acknowledgements

We express our profound appreciation to the Machine Learning Laboratory of Dr. Sudhir Chandra Sur Institute of Technology & Sports Complex for providing the essential computing infrastructure for the training and assessment of our models. We express our profound gratitude to Mr. Sanjoy Bhattacharjee and Dr. Dipankar Misra for their technical help on picture segmentation techniques and the development of matching algorithms.

We express our gratitude to the open-source communities whose contributions to computer vision libraries and language models have been vital in facilitating this study.

References.

[1] Si, C., Zhang, Y., Li, R., Yang, Z., Liu, R., & Yang, D. (2024). Design2Code: Benchmarking Multimodal Code Generation for Automated Front-End Engineering. arXiv preprint arXiv:2403.03163.

[2] Zauner, C. (2010). Implementation and Benchmarking of Perceptual Image Hash Functions. In Proceedings of the 2nd International Conference on Advanced Computer Science Applications and Innovations (pp. 37–42).

[3] Brown, T. B., et al. (2020). Language Models are Few-Shot Learners. In Advances in Neural Information Processing Systems (NeurIPS) (pp. 1877–1901).

[4] Canny, J. (1986). A Computational Approach to Edge Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 8(6), 679–698.

[5] Selenium WebDriver Documentation. SeleniumHQ, 2023. (Reference updated with the official project documentation.)

[6] MacQueen, J. (1967). Some Methods for Classification and Analysis of Multivariate Observations. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability (pp. 281–297).

[7] Ma, L., & Xu, K. (2019). Automated Web Page Layout Analysis Using Computer Vision Techniques. In Proceedings of the International Conference on Web Engineering (pp. 102–110).

[8] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., &Polosukhin, I. (2017). Attention is All You Need. In Advances in Neural Information Processing Systems (NeurIPS) (pp. 5998–6008).

AUTHORS PROFILE (All authro profiles are mandateory)

Author-1 currently pursuing his B.Tech. degree in Computer Science and Engineering with a specialization in Artificial Intelligence and Machine Learning from Dr. Sudhir Chandra Sur Institute of Technology and Sports Complex. He has acquired 1.5 years of practical experience in the domains of Web Development and DevOps,
having interned at multiple startups where he contributed to the development and deployment of full-stack applications and cloud-based infrastructure.

Hisprimary interests include Web Development, DevOps, Artificial Intelligence and Machine Learning, Data Science, and Cloud Computing. He actively engages in building real-world projects that integrate these technologies, showcasing both technical depth and interdisciplinary application.

Beyond academics and research, he is a NCC 'C' Certificate holder, demonstrating qualities of leadership, discipline, and a strong sense of national service.

Author-2 pursuing a B.Tech. in Computer Science and Engineering with a specialization in Artificial Intelligence and Machine Learning at Dr. Sudhir Chandra Sur Institute of Technology & Sports Complex, he has been actively involved in hands-on projects, with a strong focus on integrating web technologies with intelligent systems.

His core interests include Full-Stack Web Development, Machine Learning, Computer Vision, Natural Language Processing, Deep Learning, the Internet of Things (IoT), and Data Structures & Algorithms. With a keen enthusiasm for solving real-world problems, he strives to create innovative applications that blend AI capabilities with robust software design. Through both academic coursework and practical experience, he continues to enhance his technical proficiency and research outlook

Author-3 pursuing her B.Tech. degree in Computer Science and Engineering with a specialization in Artificial Intelligence and Machine Learning from Dr. Sudhir Chandra Sur Institute of Technology & Sports Complex. She has undertaken various software development projects contributing to diverse domains,
actively participated in hackathons and meetups, served as a freelance full stack web-developer and is now part of an upcoming software agency.

Her primary areas of interest include backend engineering, Artificial Intelligence, DevOps, Blockchain, Large Language Models, and algorithms. She is passionate about enhancing power of traditional working models and industries by integrating power of AI and is steadily stepping towards it by taking on personal, academic and industrial projects.

Author-4 Pursuing a B.Tech. in Computer Science and Engineering with a specialization in Artificial Intelligence and Machine Learning from Dr. Sudhir Chandra Sur Institute of Technology & Sports Complex, he has undertaken various project-based roles, actively contributing to research and development in web technologies and intelligent systems.

His primary areas of interest include Full-Stack Web Development, Machine Learning, Computer Vision, Natural Language Processing, Deep Learning, and Data Structures & Algorithms. Passionate about building impactful solutions that integrate software engineering with artificial intelligence, he is continuously expanding his skills through academic coursework and real-world projects.

Author-5 earned his B.Tech. and M.Tech. degrees from the University of Calcutta in 2011 and 2013, respectively. He is currently pursuing a Ph.D. in the Department of Computer Science and Engineering at IIT Patna. He is also serving as an Assistant Professor in the Department of Computer Science and Engineering at Dr. Sudhir Chandra Sur Institute of Technology & Sports Complex.

His primary research interests include Generative Artificial Intelligence, Cryptographic Algorithms, Network Security, Cloud Security and Privacy, Natural Language Processing, and Data Mining. He has over 7 years of teaching and research experience.

Author-6 is currently a Professor in the Department of Computer Science & Engineering at JIS University, Kolkata, with over 16 years of teaching experience. He has previously served at BBIT, UEM, and FIEM, contributing significantly to academic and administrative roles, including NBA coordination, EDC, IIC, and various institutional committees. His teaching spans core CSE subjects and interdisciplinary programs.

Dr. Misra also brings 4 years of industry experience from IBM, TCS- supported projects at KMC, and REALTECH INDIA in roles related to software development and data management. He earned his PhD in CSE (2020), focusing on Green Cloud Computing and Big Data Analytics, following an MTech from the University of Calcutta (2007) and a BTech from the University of Kalyani (2003). A dedicated academician and administrator, he continues to contribute to research, curriculum development, and institutional advancement.