

Enhancing visual testing through AI and CV innovations



Contents

Abbreviations	3
Abstract	3
Introduction	4
Market trends	5
Problem statement	5
Solution	5
Benefits	11
Conclusion	11
References	12
Author info	13

Abbreviations

Abbreviation	Definition
OCR	Optical Character Recognition
CV	Computer Vision
DL	Deep Learning
JPG	Joint Photographic Experts Group
Bmp	Bit map
PNG	Portable Network Graphics
WPF	Web Picture Format
RoI	Region of Interest
UI	User Interface
URL	Uniform Resource Locator

Abstract

In the digitalization arena, the products/application UI layer undergoes frequent changes in the content, images and screens to enhance the customer experience. The frequent design modifications require testers to validate the contents of both the previous and new UI designs, including text labels such as username, password, and login, as well as elements like text boxes, icons, and buttons. This results in recurring changes in the web application's layout, leading to the need for developers to ensure the availability of critical text labels and UI elements.

Frequent design changes in application development create challenges in areas like product lifecycle management, engineering drawing management, and quality management. To solve these challenges, we need to compare current products or drawings with new designs. This process or solution helps us spot the differences between the two versions. This solution aims to compare UI screens, drawings, webpages, or any images and highlights the differences with very high accuracy in a meaningful form.

In this whitepaper the proposed solution will provide detailed reports for each highlighted difference with great precision, enabling testers to easily track and understand the nature of the variations. This solution also provides an option X-factor that helps in identifying a portion of the image on the target image. By automating the comparison process, the invention eliminates the need for manual efforts, reducing the time and effort required for testing. In these scenarios, the solution will help testers quickly and accurately identify the variations, reducing the time and effort required for testing. The meaningful callouts provided with high precision on text and image differences help testers validate images seamlessly.

Introduction

This solution streamlines and optimizes the process of validating software application UI design changes by automating the comparison of baseline and production images. Baseline images are the golden standard images that represent the existing or previous version of the UI screen, while production images represent the modified version under development. The solution may provision an option to save these images with versioning for future reference. There are tools available in the market that perform image comparisons and may only highlight the pixel-level differences as a final report. This report will be challenging for users to get the details on each highlighted difference and will require more manual effort to generate the report. The proposed next-generation Gen AI-based solution efficiently highlights pixel-level differences between the images by providing detailed callouts on reported differences. This will help testers quickly identify the changes based on the meaningful callouts with precision and accuracy. The meaningful output of the differences is highlighted with numbering, enabling users to quickly understand the significance of the variations observed. This solution eliminates the need for manual efforts, simplifies the process of image validation in quick turn-around time, and supports quality assurance teams, web developers, and designers.

The proposed next-generation Gen AI-based solution efficiently highlights pixel-level differences between the images by providing detailed callouts on reported differences. This will help testers quickly identify the changes based on the meaningful callouts with precision and accuracy. The meaningful output of the differences is highlighted with numbering, enabling users to quickly understand the significance of the variations observed. This solution eliminates the need for manual efforts, simplifies the process of image validation in quick turn-around time, and supports quality assurance teams, web developers, and designers.

Automating the image validation process helps to overcome the challenges encountered by testers in performing validations, ensuring that applications are of high quality to meet the desired specifications.

The image validation use cases encompass various digital assets, including application UI, sketches, and webpage screens. These screenshots are captured and saved as image files in different formats (PNG, JPG, JPEG, BMP) for future reference and version control. This solution offers flexible options for image validation, including headless, normal mode with multi-browser and multi-dimension support. Additionally, it accommodates the X-factor, where a portion of the image is segmented from the baseline and overlaid onto the production image for thorough validations. This provision enables testers to cover a wide range of combinations for performing validations in testing cycles, ensuring seamless execution of tasks. The following are some of the supported input combinations which can be further extended:

1. Baseline and live screen images in various formats.
2. Multi-browser support, desktop apps, and devices for normal mode validation.
3. Headless validation for automated testing
4. Image segmentation with X-factor for detailed validations.

Market trends

- By 2033, machine learning and pattern recognition are projected to dominate the global industry, boasting a significant market share. The market size is being driven by the growing need to implement testing tools with AI capabilities. Over the course of the forecast period, it is anticipated that testing tools with AI capabilities coupled with machine learning (ML) and pattern recognition technologies will continue to hold the top revenue share. The market for AI-enabled testing tools is expected to reach US\$ 2,030.75 million by 2033, having been valued at US\$ 426.1 million in 2023. The market is expected to grow at a rate of 16.9% CAGR over the projected period.
- Market Size & Trends for AI-enabled Testing: The market for AI-enabled testing was estimated to be worth USD 414.7 million in 2022, and between 2023 and 2030, it is projected to expand at a compound annual growth rate (CAGR) of 18.4%. Numerous prominent companies are providing AI-powered testing solutions. Google, for example, has developed Deep Test, an AI-powered testing tool that leverages Machine Learning (ML) to automatically generate test cases and run them on web applications. Its goals are to decrease the amount of work needed to build test cases and increase test coverage. Furthermore, notable advancements and improvements in testing tools provided by AI are driving market expansion.

Problem Statement

- During application development/testing, the baseline image is to be validated with a production (UI/screenshot) image, which will be a challenging task for manual tester to identify the differences and report precisely.
- The standard image comparison tools only highlight the identified differences in reports, and that may not provide detailed insights, callouts, or meaningful reports for a tester.
- There should be a way to classify the type of the highlighted difference with meaningful callouts.

Solution

The proposed solution provides a method to validate the images in a testing environment seamlessly without any manual effort. This solution provides an advantage over the other tools present in the market for image comparison tasks by eliminating the tedious job of manually reporting the identified changes in the screens. It will also help provide the cumulative report of changes with maximum accuracy in one place for the tester, with insights as a pop-up on output, along with an Excel report with detailed callouts.

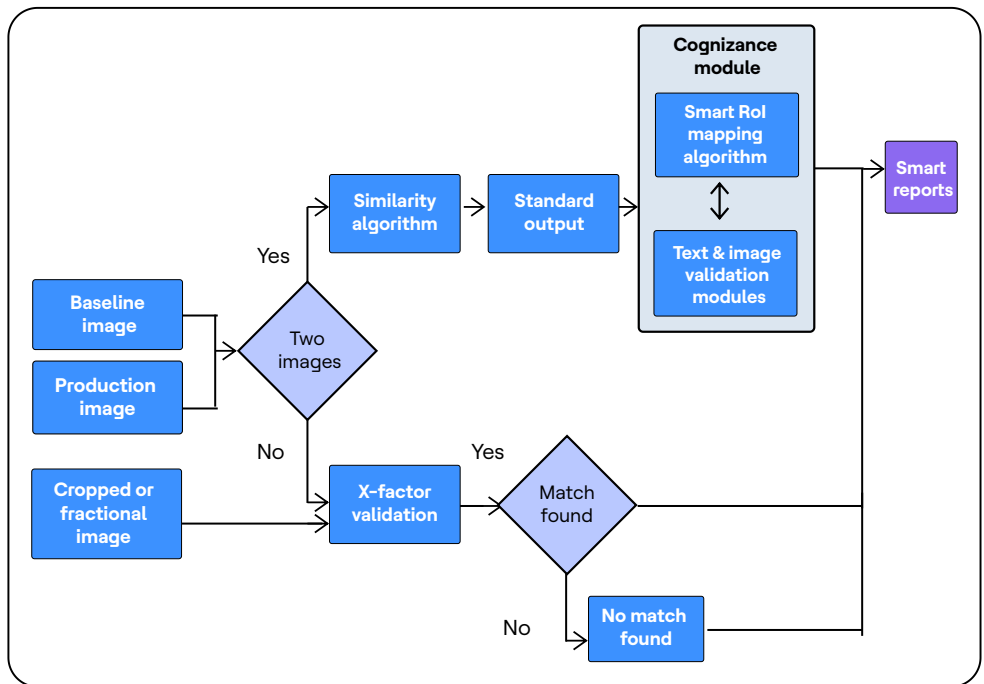


Figure 1 – Workflow diagram

Figure 1 refers to the workflow of the solution. The input to the solution consists of a baseline (Fig 2) and production image (Fig 3) that is passed through a decision block, which selects the appropriate algorithm based on input arguments. The two images are provided as input for combinations like Image vs Image, URL vs URL, Image vs URL and vice versa. These images are passed to the Similarity algorithm for identification of differences.

The pre-processing steps are applied to ensure the image format type, convert both images to a suitable format for validations, and improve the image quality as needed. The validation is performed on the pre-processed image at the pixel level to detect all deviations between the images under test and these variations, along with the region of interest (ROI). This similarity algorithm is capable of drawing boxes and highlighting the identified differences in the production image, assigning a unique number to each difference based on ROI information. The output from this process is a standard output that many comparison tools in the market offer.

The cognizance module plays a crucial role in this solution by processing all coordinate information received from the standard output, utilizing a Smart ROI mapping algorithm that analyses the data in conjunction with appropriate ROI numbers. This algorithm detects overlapping regions and other characteristics within the image and groups them accordingly. The enhanced regions are then subjected to intelligent merging/mapping, allowing for meaningful ROI assignments. The module compares the coordinates of the production image with those of the baseline for all validations and sends the data to the text/image validation module.

The text validation module performs extraction only when the data contains text information. This module uses a deep learning-based OCR model to check for text (missing, new addition, size, colour, shift, etc.). This module can perform font classification by detecting the changes in font style, such as Times New Roman versus Calibri. The image validation module processes the information and can identify variations like background colour, image shift, dimension, image missing, distinct pictures, etc.

The observations with callouts from the aforementioned blocks are provided to the smart ROI algorithm, which generates smart reports containing ROI coordinate markings, numberings, and callouts plotted on the production image (Fig 4). An Excel report is generated with the information shown in (Fig 6), along with a summary report shown in (Fig 7). The users are provided with an option to mouse hover and view the callout information (Fig 5) on the output image. The smart report generation facilitates detailed insights for users on all differences detected in a production image.

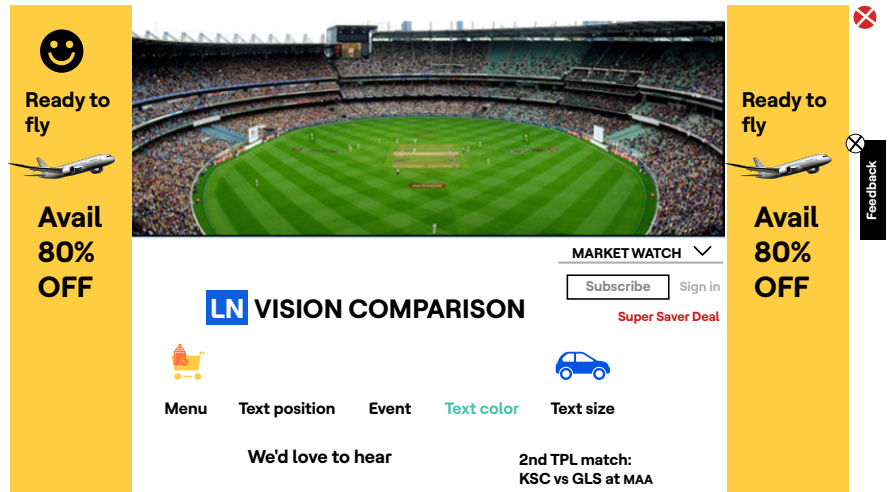


Figure 2: Baseline or golden standard image

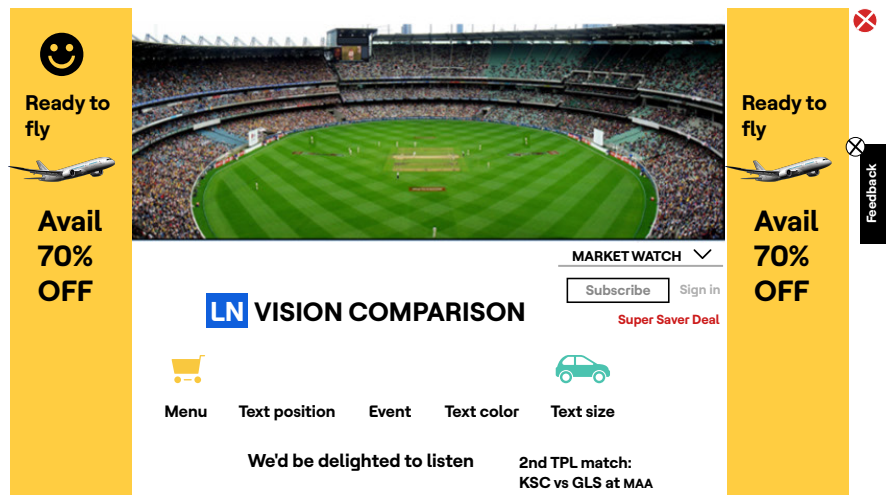


Figure 3: Production image



Figure 4: standard output

Figure 4 is the "standard output" generally comparison tools may provide. There is a difference observed in the Text "We'd be delighted to listen" the standard output gives the output shown in Figure 4, which is difficult to report; this solution smartly converges ROIs and marks all the differences into single ROI, which can be seen in ROI 7 in Figure 5.

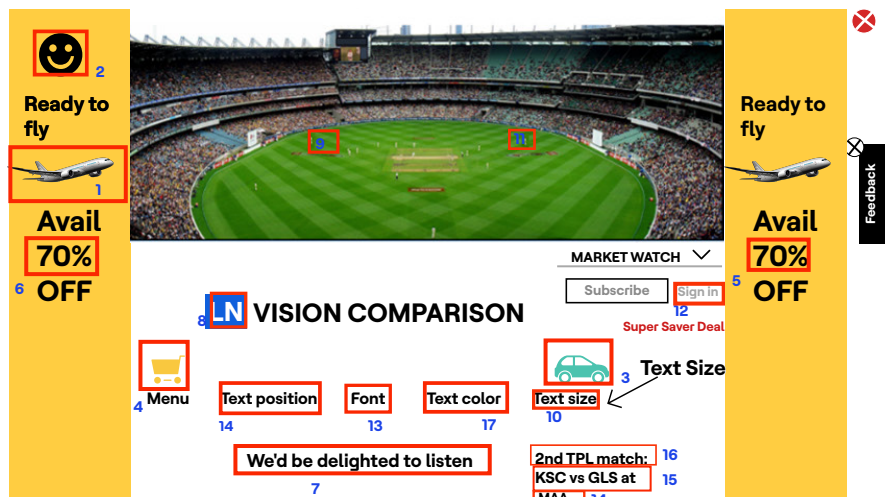


Figure 5: Meaningful output with call-out option

From a user perspective, the output in Figure 5 is more readable and simple to understand than in Figure 4. The following report is generated for user reference, and the coordinates of the difference can be provided by a configurable parameter.

ROI	Callout	Remarks
1	Image shift	Down-shift
2	Distinct image	Different image
3	Distinct image	Different image
4	Distinct image	Different image
5	Text difference	Change in text
6	Text difference	Change in text
7	Text difference	Change in text
8	Font style changed to Calibri	Change in font style
9	Distinct image	Different image
10	Text size	Change in text size
11	Distinct image	Different image
12	Background change	Background color change
13	Text difference	Change in text
14	Text alignment	Shift in text - Right
15	Text difference	Change in text
16	Text difference	Change in text
17	Font color change	Change in font color
18	Text difference	Change in text

Figure 6: The final report generated with the call-out information.

Summary Report:

Difference	ROI number
Image shift	1
Distinct image	2, 3, 4, 9, 11
Text difference	5, 6, 7, 13, 15, 16, 18
Font style changed to Calibri	8
Text size	10
Background change	12
Text alignment	14
Font color change	17

Figure 7 is the summary report of the final report, which groups the differences based on the call-out to which they belong.

In X-factor validation, the inputs are a portion of the image (Figure 8a and Figure 9a) and a real-time image Figures 8 and 9). There are use cases where a cropped/fractional part of the image is to be validated with the target (Baseline/production) image. The X-factor validation block will perform validation on the cropped image with the target image, and if a match is found, it will be shared with the reports with details (Figure 8.1 and 9.1), and when there is a match not found, it is marked directly as "No Match Found". In X-factor validation, users are provided with an option to change the match percentage as needed to meet their matching requirements.

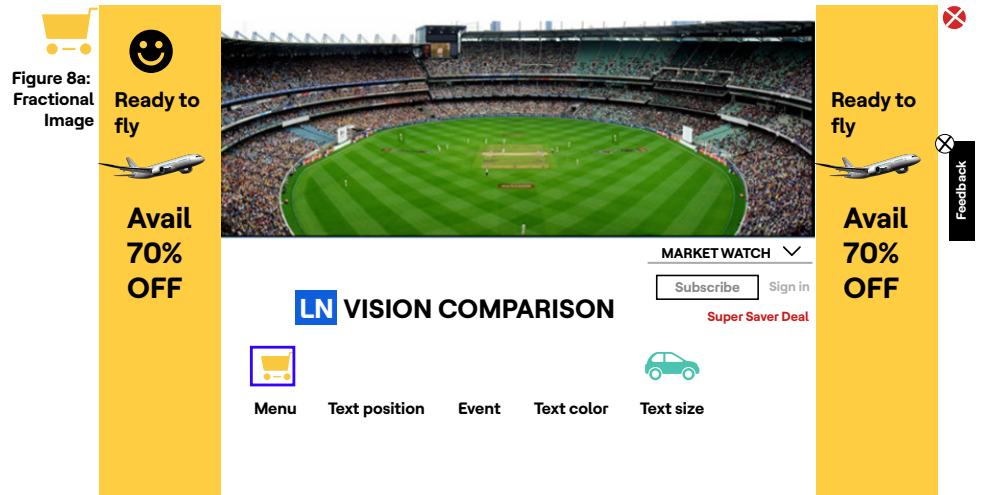


Figure 8b: Exact-Match

Figure 8a depicts the input fractional portion of the image in Figure 8b.

Match number	Quadrant
1	Q3

Figure 8c: Report

Figure 8c provides the coordinates of the marking along with the position on the image.



Figure 9b: Exact-Match

Figure 9a depicts the input fractional portion of the image in Figure 9b.

Match Number	Quadrant
1	Q3
2	Q4

Figure 9c: Report

Figure 9c provides the coordinates of the marking along with the position on the image.

Benefits

1. This solution works as a standalone or can plug into any existing test automation infrastructure.
2. To compare any two images or screens irrespective of dimension and resolution.
3. This solution marks the differences in images and classifies them to provide the type of difference they belong to.
4. Testing cycle turn-around time may be reduced by 50-60%.
5. Meaningful callouts with greater precision for easy reporting.
6. Various input and configurable options for validations.
7. The solution functionality can be extended to multiple domains.

Conclusion

This white paper focuses on validating images and providing meaningful callouts for the identified differences with a detailed report. This will help the tester to quickly understand the type of difference and comes in handy when a developer/tester wants to check all the changes in the UI/design once implemented. This solution provides X-factor validation where a fractional image is identified on a target image; it will help when the user wants to target a specific portion of the image rather than the entire image.

The key differentiator in this solution is that it will smartly converge the ROIs to provide meaningful outputs with detailed callouts, whereas the market tools may provide pixel-to-pixel differences. This whitepaper offers a powerful solution for developers and testers seeking to ensure greater accuracy and precision in visual testing. Enabling a comprehensive comparison of two images/screens streamlines the process of identifying differences, which is crucial for maintaining consistency and quality in development environments.

References

- <https://www.geeksforgeeks.org/introduction-to-opencv/?ref=lbp>
- <https://stackabuse.com/pytesseract-simple-python-optical-character-recognition/>
- <https://www.geeksforgeeks.org/python-opencv-getting-and-setting-pixels/>
- <https://nanonets.com/blog/deep-learning-ocr/>
- <https://www.futuremarketinsights.com/reports/ai-enabled-testing-tools-market>
- <https://www.grandviewresearch.com/industry-analysis/ai-enabled-testing-market-report>

Author info



Narender S

Narender is a Senior Solution Architect with 20+ years of experience. Experienced in product engineering and sustenance engineering and testing for a couple of years in various domains, including product, platform and enterprise customers. Strong experience in creating AI/ML-based automation solutions to address the current market challenges and improve the efficiency of digital customers.



Kiran Ayyagari

Kiran is a Solution Architect with 19+ years of experience in IT, including Test Automation, Framework Design, Tool Development and Application Testing. He spent 8+ Years at multiple Geo locations in Europe and the USA, working with top IT customers. He has extensive knowledge of UI, API, and Mobile test automation, using both open-source and commercial automation tools/frameworks.



Srihari V

Srihari is a Solution Architect with 20+ years of experience. He has managed various testing teams and has been creating next-gen solutions as value adds for leading OEM clients. He is currently part of the Solutions team and generates AI-based solutions to support business needs.

HCLTech | Supercharging Progress™

HCLTech is a global technology company, home to 222,000+ people across 60 countries, delivering industry-leading capabilities centered around Digital, Engineering and Cloud powered by a broad portfolio of technology services and software. The company generated consolidated revenues of \$12.3 billion over the 12 months ended December 2022. To learn how we can supercharge progress for you, visit hcltech.com.

hcltech.com

