Technical Application of Dynamic Visual Recognition System in Textile Industry Advertising and Brand Communication

Bin Zhang, Qian Zhang
Organization and Human Resources Department, Weifang Vocational College, Weifang 262737, Shandong, China
Corresponding author: Bin Zhang (e-mail: angie0111@163.com)

ABSTRACT With the rapid development of intelligent multimedia communication and immersive information delivery, dynamic visual perception technologies are becoming increasingly important for digital advertising systems and next-generation interactive communication applications. To overcome the limited engagement and weak brand memorability of conventional static textile advertisements, this study proposes a dynamic visual recognition system integrating YOLOv8-based object detection, convolutional neural network (CNN) feature extraction, and augmented reality (AR) rendering. The framework first detects fabric textures, garment outlines, and brand logos in real time, and subsequently constructs multidimensional representations of brand elements through hierarchical feature modeling and fusion. The extracted semantic features are then incorporated into an AR-driven interactive presentation mechanism to generate immersive advertising experiences. Experimental results demonstrate that the proposed system significantly improves user attention, brand recall, and interactive engagement, with brand logo gaze duration reaching 2450 ms, free recall increasing to 78.5%, and participation in fabric-detail exploration achieving 92.5%. By combining intelligent visual recognition with dynamic information presentation, the proposed approach provides an effective solution for enhancing advertising communication and offers valuable technical insights for vision-assisted multimedia transmission and intelligent electromagnetic information interaction systems.

INDEX TERMS dynamic visual recognition, YOLOv8, convolutional neural network, augmented reality, intelligent information interaction

I. INTRODUCTION

As an important part of the manufacturing industry, the textile industry faces increasingly diversified brand communication needs in the context of intensified global competition. As consumers' visual experience requirements increase, advertising not only conveys information but also plays a role in brand image construction and consumer psychology guidance [1,2]. Traditional static advertising relies on flat images and text, but it is easy to cause content stacking, distraction, and visual fatigue, which weakens the communication effect [3,4]. The development of digital and intelligent technology provides an opportunity to break through the limitations of static advertising and enhance brand competitiveness. In the research related to advertising and visual communication, scholars have conducted multi-dimensional discussions on the composition and recognition of visual elements. Ali et al. proposed an enhanced artificial intelligence virtual influencer framework to analyze the impact of emotional display on user engagement [5]; Yang et al. reviewed the application of computer vision in design optimization and interaction improvement, providing technical support for dynamic visual advertising [6]. In the field of textile recognition, Akram et al. proposed an improved method of convolutional neural network for fabric texture classification, which improved the recognition accuracy of complex textures and provided a basis for the system to realize brand element recognition and product feature recall [7]. Nalbant and Aydin discussed the transformation of brand marketing and digital marketing in the metaverse environment based on artificial intelligence and digital technology, revealing the important role of virtual interaction and immersive experience in enhancing user engagement and brand memory [8]. In addition, Halepoto et al. summarized the research trends of artificial intelligence in the textile industry through bibliometric analysis, pointing out that intelligent recognition and real-time monitoring technologies have broad prospects in textile inspection and advertising applications [9]. Li and Zhu proposed a high-precision, real-time convolutional neural network for fabric defect detection. Their method improves detection accuracy while ensuring speed [10].

Overall, although existing research provides theoretical support for the improvement of textile advertising communication, it has not yet formed a systematic solution path in terms of dynamic, intelligent and interactive applications. To further promote the application of visual recognition technology in advertising communication, some researchers have attempted to use deep learning models to identify target elements in advertisements and apply convolutional neural networks to extract texture and color features to improve the consistency between advertising content and brand identity. In the field of textile quality inspection and production automation, Jung et al. proposed StitchingNet and deep transfer learning methods for sewing stitch defect detection, which improved detection accuracy under small sample conditions [11]. Nguyen et al. explored visual inspection and synchronous control methods in the context of sustainable production transformation, emphasizing the important role of computer vision-based textile surface automation inspection systems in improving production efficiency and quality monitoring [12]. Glogar et al. summarized the application of digital technology in the sustainable design and development of textiles and clothing through a literature review, pointing out that intelligent design and data-driven production processes can effectively reduce resource consumption and optimize user experience [13]. In addition, Rainy proposed an artificial intelligence-driven marketing analysis method, systematically reviewed data-based retail advertising optimization strategies, and emphasized the use of real-time data and intelligent analysis to improve user engagement and brand memory [14].

While research has demonstrated that deep learning and augmented reality methods can improve advertising communication, the integration of visual object detection, deep feature modeling, and dynamic interactive presentation remains a research gap. To address this issue, this paper proposes a dynamic visual recognition system that integrates YOLOv8 object detection, convolutional neural network feature extraction, and augmented reality rendering technology to overcome the bottleneck in the brand communication effectiveness of textile industry advertising. This paper explores the application of dynamic visual recognition systems in textile industry advertising and brand communication, developing a solution that integrates real-time detection, feature extraction, and interactive presentation. The study first utilizes YOLOv8 to efficiently detect key visual elements, and then extracts multidimensional brand features through a convolutional neural network. Finally, an augmented reality engine is used to generate immersive interactive advertising scenes. This research, centered on dynamic visual recognition, integrates computer vision with advertising communication to provide innovative communication tools for textile brands.

II. IMPLEMENTATION PATH OF A DYNAMIC VISUAL RECOGNITION SYSTEM

A. VISUAL OBJECT DETECTION IN ADVERTISING SCENES

1) Data Acquisition and Preprocessing

To enable real-time detection of core elements in textile advertisements, a multi-source image dataset covering fabric texture, garment outlines, and brand logos is constructed. The dataset includes advertisement video frames and high-resolution promotional images, ensuring representative samples in terms of texture, color, and morphology. The samples are uniformly scaled to 640×640, converted to RGB, and normalized for brightness and contrast. Local features are enhanced using Gaussian filtering and adaptive histogram equalization. All samples are annotated with bounding boxes for categories including fabric texture, garment outline, and brand logos. The dataset is split into training, validation, and test sets with a 7:2:1 ratio for model training and evaluation. The dataset constructed in this study contains 12,480 samples, including 8,736 static high-resolution advertising images and 3,744 keyframes from advertising videos (sampled at 5 frames per second from 72 brand promotional videos). The distribution of the three types of labeled objects is as follows: 5,620 fabric texture samples, 4,310 clothing outline samples, and 2,550 brand logo samples. This distribution reflects the reality that textures and outlines occupy a larger visual area in advertising content, while ensuring that brand logos are sufficiently representative during training to support high-precision recognition.

2) Model Construction and Detection Implementation

In terms of model selection, the YOLOv8 deep learning object detection framework is used as the backbone network, and multi-scale feature extraction is achieved through convolution and feature pyramid structures. The input image first passes through a combination of convolutional layers and batch normalization layers to extract the underlying texture information, and then forms a multi-resolution feature map through a feature fusion module, thereby taking into account both fabric details and overall structure. The detection head adopts a decoupled structure, and optimizes the classification and bounding box regression tasks separately to reduce the gradient interference between different tasks [15]. The loss function adopts a composite form as shown in Formula 1:

$$L = L_{\text{cls}} + \lambda_1 L_{\text{box}} + \lambda_2 L_{\text{obj}} \tag{1}$$

L_{cls} represents the category classification loss, which is calculated using the cross entropy function; L_{box} represents the bounding box regression loss, which uses the CIoU (Complete IoU) method to improve positioning accuracy; L_{obj} represents the target existence loss, which measures the confidence of the target in the detection area; λ_1 and λ_2 are weight parameters used to balance the contribution of different losses. During the training process, the AdamW optimizer is used to update the parameters with a dynamic learning rate. The initial learning rate is set to 0.001, and the cosine annealing strategy is used to gradually decrease the learning rate to improve the convergence stability of the model. In the final model training, the weight parameters are set to $\lambda_1 = 0.8$ and $\lambda_2 = 0.4$. This configuration is optimized on the validation set through grid search, achieving the best balance between localization accuracy and target existence confidence, while avoiding the dominance of bounding box regression in class loss.

Figure 1 shows the performance of a dynamic visual recognition system for object detection and model training in textile advertisements. In Figure 1(a), the horizontal axis represents training rounds, and the vertical axis represents loss. The training curves for both YOLOv8 and Faster R-CNN (Faster Region-based Convolutional Neural Network) show a downward trend. YOLOv8 converges quickly in the first 50 rounds, ultimately stabilizing its loss at a low level, demonstrating its high training efficiency for detecting fabric textures, garment outlines, and brand logos. Figure 1(b) shows detection confidence on the horizontal axis and frequency on the vertical axis. YOLOv8's confidence distribution is concentrated in the 0.7-1.0 range, while Faster R-CNN's is more dispersed, with a higher proportion of medium and low confidence levels. This indicates that YOLOv8 is more reliable and stable in identifying key visual elements in advertising images. These data demonstrate the advantages of dynamic visual recognition systems in improving the accuracy and stability of advertising brand symbol recognition.

During the detection implementation phase, the trained model is deployed to the advertising image processing system, performing real-time detection on the input image or video stream. The non-maximum suppression (NMS) algorithm removes highly overlapping candidate boxes, retaining the target regions with the highest confidence. The output is presented as a category label and bounding box coordinates, and is fed back in real time to the subsequent feature extraction module. This process effectively ensures that fabric textures, clothing outlines, and brand logos in advertisements can be quickly identified even against complex backgrounds.

B. MULTIDIMENSIONAL FEATURE EXTRACTION OF BRAND ELEMENTS

1) Feature Representation Construction

After detecting the target area in the advertisement image, the extracted fabric texture, garment outline, and brand logo area are fed into a convolutional neural network for multidimensional feature modeling. The input image is first stacked through three convolutional layers to extract low-level local edge and texture information. Each convolutional layer uses a 3×3 convolution kernel. A batch normalization layer and ReLU activation function are added after the convolution operation to enhance training stability and accelerate convergence. After the feature map is generated, spatial downsampling is performed through the maximum pooling layer to reduce the amount of computation and retain key features. In view of the fine-grained differences in textile textures, a deformable convolution module is additionally introduced to adapt to irregular texture distribution, thereby obtaining higher local expression capabilities. For color distribution features, a multi-channel convolution structure is used to independently calculate weights on the three RGB channels, so that the color information is fully expressed in the network layer. The capture of morphological parameters is achieved by setting a convolution kernel with a larger receptive field in the high-level convolution layer to cover the overall structure of the clothing in the advertising picture. The feature tensor output by the convolutional neural network is expressed as Formula 2:

$$F = f_\theta(X) \tag{2}$$

X represents the input target area image; f_θ is the convolutional neural network mapping function; θ is the network parameter; F is the resulting multidimensional feature representation. This process ensures that brand elements in the advertisement are systematically modeled across the three dimensions of texture, color, and form.

Figure 2 demonstrates the results of multidimensional feature extraction of brand elements in textile industry advertisements. In Figure 2(a), the horizontal axis represents feature ranking; the left side of the vertical axis represents feature importance; the right side of the vertical axis represents cumulative importance. The data shows that the importance of the top ten features is significantly higher than that of other features, and the cumulative importance exceeds 50%, indicating that some key features in the advertising image play a dominant role in brand recognition. Figure 2(b) shows the two-dimensional feature distribution after dimensionality reduction using t-SNE. Most brands form clearly separated clusters in space, indicating that the extracted multidimensional features have good discriminative ability. It is worth noting that the brands visually overlap to some extent, mainly due to the use of highly similar color schemes and classic plaid patterns in their advertising designs. Nevertheless, Euclidean distance analysis in the original high-dimensional feature space shows that the average inter-class distance of each cluster is still significantly greater than the intra-class variance ($p < 0.01$, t-test), indicating that the system can still effectively distinguish them in actual recognition tasks. This overlap phenomenon precisely reflects the challenge of visual similarity of brand elements in real advertising scenarios, while this system can still maintain a high brand discrimination by integrating texture, color, and shape cues.

2) Multi-Level Feature Fusion

After obtaining features of different dimensions, a layered fusion strategy is used to integrate low-level detail features with high-level semantic features. Specifically, the outputs of different convolutional layers are stacked layer by layer using a residual structure to prevent gradient vanishing and maintain feature continuity. Subsequently, the texture features and color features are spliced at the channel level;

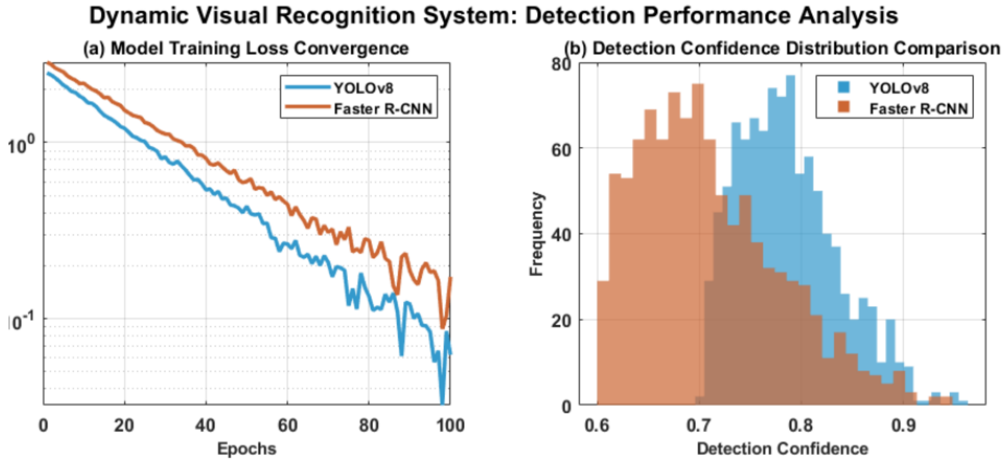

FIGURE 1. Dynamic visual recognition system training and confidence analysis. Figure 1(a). Model training loss convergence curve. Figure 1(b). Comparison of detection confidence distributions

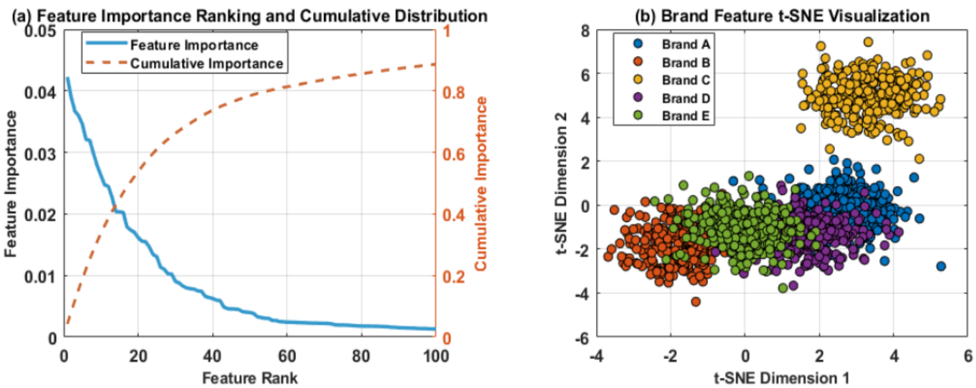

FIGURE 2. Analysis of Multidimensional Feature Extraction of Brand Elements. Figure 2(a). Feature Importance Ranking and Cumulative Distribution. Figure 2(b). t-SNE Dimensionality Reduction Visualization of Brand Features

the number of channels is compressed using the convolution kernel; cross-channel information interaction is achieved. For the high-level semantic features of the morphological parameters, the attention mechanism is used for weighted processing. The weights are assigned by calculating the correlation matrix of the feature channels to improve the sensitivity to the brand logo and clothing contour areas [16]. The final fused feature vector is mapped to a unified feature space through a fully connected layer, expressed as Formula 3:

$$Z = W \cdot \mathrm{Concat}(F_t, F_c, F_s) + b \tag{3}$$

F_t, F_c, and F_s represent texture, color, and morphological features, respectively. Concat is the channel concatenation operation. W and b are mapping parameters. Z is the final fused representation of brand element features. This representation serves as input for subsequent dynamic ad rendering, ensuring that brand information remains complete and distinct across multiple layers of features.

To achieve structured separation of the three types of features, branches at layers 3, 5, and 7 of the CNN backbone are introduced: the output of layer 3 (feature map size 64×64, 128 channels) is compressed into 32-dimensional texture features through global average pooling and 1×1 convolution F_t; layer 5 (32×32, 256 channels) is compressed into 32-dimensional color features after passing through a color attention module (weighting the RGB channels) F_c; layer 7 (16×16, 512 channels) captures multi-scale morphological information through spatial pyramid pooling (SPP) to generate 64-dimensional shape features F_s. These three are concatenated into a 128-dimensional vector through a concat operation, and then mapped to the final vector by a fully connected layer Z. Through the above process, a convolutional neural network extracts multidimensional features of brand elements in textile advertisements, preserving details of fabric texture and color while capturing the overall structural characteristics of the garment. Furthermore, a fusion mechanism ensures the effective expression of multiple features within a unified space.

C. DYNAMIC INTERACTIVE VISUAL PRESENTATION

After extracting brand element features, the resulting fused features are used as input and loaded into the augmented reality rendering engine to construct the virtual scene. First, a 3D coordinate system is created based on the background environment of the textile advertisement. Using the advertisement image as a reference plane, the spatial positioning of the product display area is determined. The brand logo and textile outline are mapped to the 3D model surface using the numerical parameters of the eigenvector Z, ensuring that the spatial distribution of visual elements is consistent with the actual advertising layout. During the texture loading process, a material modeling method based on physical rendering (PBR) is used to convert the color distribution and texture details extracted by the convolutional neural network into material mapping parameters, including diffuse channel, normal channel, and specular channel [17]. This process ensures that the brand symbol and the textile surface maintain a realistic and consistent appearance under changes in lighting and viewing angle. Feature vectors Z are not only used for spatial localization, but also directly drive the parametric generation of materials and dynamic behaviors. Specifically, the texture sub-vector F_t is mapped to the normal and roughness channel weights of the PBR material, determining the microscopic details of the fabric under different lighting conditions; the color sub-vector F_c controls diffuse reflection and ambient light occlusion intensity; the morphology sub-vector F_s is converted into initial deformation parameters for skeleton binding and elastic coefficients for fabric simulation through a regression network, thereby ensuring that the dynamic wrinkles of the simulated clothing are consistent with the physical properties of the real product.

Therefore, AR rendering is not an independent module, but a parametric process driven by semantic vectors output from the recognition-feature extraction chain. It is worth emphasizing that AR interaction does not solely rely on click coordinates or generic trigger areas; it is driven by semantic information output from a dynamic visual recognition module. When YOLOv8 detects a brand's unique logo (such as the double-C pattern) and its CNN confirms it belongs to the "brand identity" category, the system activates a brand-specific AR script, including specific animation sequences, historical story overlays, or limited-edition toggle options. Similarly, if the "jacquard fabric" texture is recognized, the system uses high-precision normal mapping and dynamic gloss rendering, rather than generic fabric materials. This "recognition-context-response" mechanism ensures that every interaction is based on a semantic understanding of brand elements, rather than simple location mapping.

After the model and symbol mapping is complete, a spatial calibration algorithm is used to correct the position and proportions of the brand logo. By jointly calculating the camera's intrinsic parameter matrix and the projection matrix, the virtual elements are accurately embedded in the real advertising image, avoiding the misalignment of the symbol and the textile outline caused by perspective distortion [18,19]. For multi-view advertising video frames, a combined method of feature point matching and perspective transformation is used to maintain the continuity and stability of the brand symbol from different perspectives, thereby ensuring the complete expression of the visual elements in the advertising communication. After completing the mapping of the 3D scene and symbols, a real-time rendering module is introduced to generate dynamic visual effects. During the rendering process, a hybrid strategy of rasterization and ray tracing is used to calculate the dynamic wrinkles and lighting changes of the textiles, enhancing the realism of the advertising images. Through skeletal rigging and cloth simulation technology, the motion trajectory of the clothing model in the advertising scene is combined with the physical properties of the fabric, resulting in a natural dynamic effect when the fabric is blown by the wind or when walking [20,21]. To maintain real-time performance, a hierarchical subdivision strategy is employed during the rendering process, using a high-precision mesh for the center of the user's field of view and reducing resolution at the edges to balance image quality and rendering speed.

Interactive features are integrated through sensors on the user's device. Data from the mobile device's gyroscope and accelerometer is used to update the viewing angle of the ad scene in real time, allowing viewers to view the textiles and brand logo from different angles. For touch input, a raycast-based interaction detection mechanism is established. When a user taps a brand logo or textile area on the screen, the system captures the touchpoint coordinates and locates the corresponding object in three-dimensional space, triggering the corresponding interaction event [22,23]. Interaction events include brand information display, product style switching, and dynamic color replacement. The scene is updated instantly by calling the rendering engine's event-driven interface. This design ensures that advertising communication is not limited to visual presentation but also has real-time interactive properties.

Figure 3 shows the rendering performance and user interaction behavior of the dynamic visual recognition system for the textile industry during advertising display. Figure 3(a) shows scene complexity (from simple to complex) on the x-axis, rendering quality (from low to high) on the y-axis, and frame rate (from low to high) on the z-axis. The surface shows that the frame rate decreases nonlinearly with increasing scene complexity and rendering quality. With medium complexity and medium rendering quality, the frame rate reaches a maximum of approximately 85 FPS, while with high complexity and high quality, the frame rate decreases, reflecting the performance distribution of the system in different advertising scenarios. The X-axis and Y-axis in Figure 3(b) represent the horizontal and vertical positions of user interactions, respectively, and the colors represent the probability density of interactions. The primary hotspot is concentrated in the lower right corner of the image, with secondary hotspots located in the upper left corner. The maximum value of the heat map is normalized to 1, indicating that users primarily focus on the central area of the ad. Meanwhile, a small number of interactions are distributed around the periphery, reflecting the system's visual focus and the effectiveness of ad element guidance. Overall, Figure 3 demonstrates that the dynamic visual recognition system ensures high-performance rendering under different rendering conditions while also attracting users' primary attention to the core elements of the ad, achieving immersive brand communication.

In Figure 3(a), "Scene Complexity" is defined as levels 1 to 5, comprehensively evaluated based on the number of polygons (0.5M–4M), the number of dynamic light sources (1–5), and the activation status of the particle system; "Rendering Quality" is quantified by the number of PBR material channels enabled (2–6 channels), shadow resolution (512–4096), and anti-aliasing level. At medium settings (complexity 3, quality 3), the system maintains 85 FPS, meeting the real-time interaction requirements of mobile devices. Finally, to further optimize interactive effects, a deferred rendering pipeline is introduced into the rendering process, separating lighting and shadow calculations to reduce latency during complex interactions. The resulting ad scene maintains an immersive feel both visually and interactively, enhancing the visual recognition of the brand symbol and fostering deeper brand memory in the audience during the interactive process.

III. VERIFICATION MECHANISM FOR SYSTEM COMMUNICATION EFFECTIVENESS

In the system communication effect verification experiment, the datasets used were derived from two types of channels. The first type was dynamic advertising samples collected by the laboratory itself, mainly including textile promotion videos and brand promotion images, which were obtained through high-definition camera equipment and processed through format unification and resolution standardization. The second type was a public textile and clothing visual material library, covering advertising images of multiple brands, styles, and textures to ensure the representativeness of the test data in terms of product diversity and brand element expression [24]. The dataset was divided according to experimental requirements, with samples divided into three groups: fixation assessment, memory assessment, and interaction assessment. This ensured consistency and comparability of materials across the different experimental stages.

A. AD FIXATION METRICS

To assess ad fixation, a representative user group was selected, and their visual behavior while viewing the ads was recorded using an eye tracker. During the test, participants viewed dynamic ads in an undisturbed environment, and the system collected data such as fixation location, fixation duration, and scan trajectory in real time. Before the experiment, interest areas were set up, with fabric texture, clothing outline, and brand logos designated as key areas to ensure that the analysis results could reflect the appeal of the core elements of the advertisement. After data collection, the results were visualized using heat map distribution and gaze trajectory maps, and the average gaze duration and first gaze delay of each area were calculated [25]. This method can accurately determine the visual priority of different elements in the advertisement, thereby evaluating the effectiveness of the dynamic visual recognition system in guiding user attention.

Figure 4 shows an analysis of user gaze behavior in textile industry advertisements. Figure 4(a) compares the average fixation duration across eight interest areas in a traditional static ad and a dynamic visual recognition system. The dynamic system significantly increased fixation durations on the "brand logo" and "fabric texture," reaching 2450ms and 2250ms, respectively, demonstrating that dynamic ads are more effective in attracting users' attention to core elements. Figure 4(b) shows a scatter plot of first fixation latency and total fixation duration, with each point (A–H) representing a single ad sample. The X-axis represents first fixation latency (ms), and the Y-axis represents total fixation duration (ms). The data showed a negative correlation, with a correlation coefficient of $r \approx -0.801$. This indicates that faster first fixations are associated with longer total fixation durations. This suggests that ads with effective visual guidance can increase user retention and focus. Overall, this demonstrates the effectiveness of dynamic visual recognition technology in enhancing user attention during ad presentation. The strong negative correlation between first-gaze delay and total gaze duration ($r \approx -0.801$) indicates that the system effectively reduces user cognitive load through YOLOv8's rapid detection (average inference time < 15ms) and AR-guided animations, enabling key elements to be captured early in visual scanning. This "rapid anchoring" mechanism prolongs subsequent deep browsing behavior, thereby increasing total gaze duration. This finding supports the effectiveness of this system in attention guidance design, rather than solely relying on the attractiveness of the content itself.

B. BRAND RECALL METRICS

Brand recall was measured immediately after users finish viewing an ad to prevent external factors from interfering with recall. The study used a questionnaire-based recall test, categorizing questions into free recall and cued recall. Free recall required participants to write down, without prompting, the brand logo, textile style, or prominent product features from an advertisement; cued recall provided visual cues to assess user recognition of brand elements. All questionnaire data was statistically analyzed to calculate the correct recognition rate, and differences across user groups were compared.

Table 1 shows the impact of different advertising formats and systems on brand recall. From static ads to the optimized dynamic recognition system, all evaluation metrics

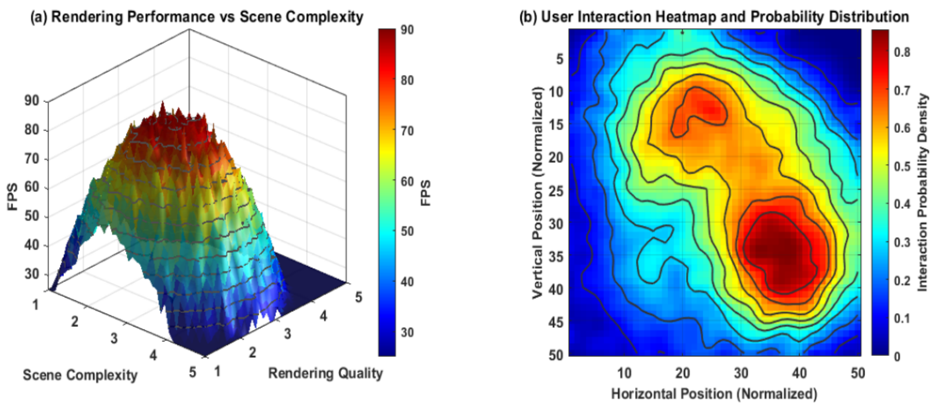

FIGURE 3. Rendering Performance and User Interaction Analysis. Figure 3(a). Relationship between Rendering Performance and Scene Complexity. Figure 3(b). User Interaction Heatmap and Probability Distribution

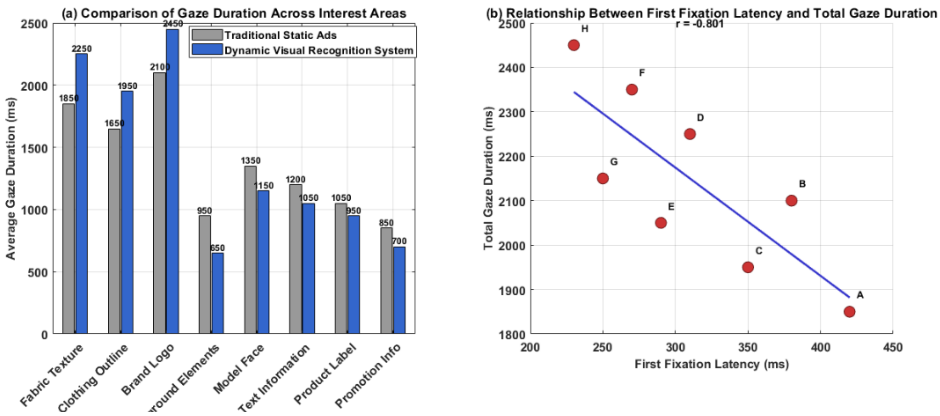

FIGURE 4. Advertisement Attention Analysis. Figure 4(a). Comparison of Attention Duration in Different Interest Areas. Figure 4(b). Relationship between First Fixation Latency and Total Fixation Duration

showed a significant improvement. Free recall increased from 32.5% for traditional static ads to 78.5% for the optimized dynamic recognition system, and cued recall also improved from 58.3% to 89.6%, demonstrating that dynamic interaction and visual recognition technology significantly enhance users' memory for brand elements. Similar trends were observed in brand logo recognition and product feature recall. The former increased from 72.1% to 94.3%, while the latter rose from 28.4% to 73.6%. This demonstrates that the dynamic recognition system not only improves overall recall but also strengthens the ability to identify core brand information and product details. Slogan recall also increased from 35.2% to 64.8%, further demonstrating the system's effectiveness in enhancing brand communication impressions across multiple dimensions. Overall, the dynamic visual recognition system enhances user engagement and interactive experience, creating a clearer and more lasting impression of the brand in the minds of audiences. This is particularly evident in core metrics such as free recall and brand recognition.

Note: ^ indicates statistically significant differences compared to static ads (control group) ($p < 0.05$).

To clarify the contribution of each technical component to the final effect, this study set up a tiered control group: static advertisement (no dynamics), video advertisement (time dimension only), basic AR interaction (no intelligent recognition, only fixed hotspots), dynamic recognition system (basic version, including YOLOv8+CNN but without optimized interaction), and optimized version (complete system). The results showed that simply introducing video or basic AR could improve the metrics, but the free recall rate increased from 32.5% in static to 45.8% in video and 58.3% in basic AR. However, after adding YOLOv8 and CNN feature modeling (basic version), the recall rate jumped to 68.9%, indicating that intelligent recognition and feature modeling themselves brought an incremental gain of 10.6 percentage points, far exceeding the marginal effect of the AR interface itself. This confirms that the dynamic visual recognition module has an independent and significant contribution to brand memory.

C. INTERACTIVE ENGAGEMENT METRICS

Interactive engagement was assessed by recording and analyzing user behavior within augmented reality advertising

TABLE 1. Comprehensive Brand Recall Assessment Results

Evaluation Dimension	Static Ad (Control)	Video Ad	Basic AR Interaction	Dynamic Recognition System (Basic)	Dynamic Recognition System (Optimized)
Free Recall Rate	32.5 ± 5.2	45.8 ± 4.7	58.3 ± 5.5	68.9 ± 4.1^	78.5 ± 3.2^
Cued Recall Rate	58.3 ± 6.7	68.4 ± 5.8	75.2 ± 4.9	82.7 ± 3.8^	89.6 ± 2.7^
Brand Logo Recognition Rate	72.1 ± 7.5	78.5 ± 6.2	82.4 ± 5.1	88.9 ± 3.5^	94.3 ± 2.1^
Product Feature Recall Rate	28.4 ± 6.3	38.7 ± 5.9	52.1 ± 5.8	62.8 ± 4.7^	73.6 ± 3.9^
Slogan Recall Rate	35.2 ± 7.1	42.6 ± 6.5	48.3 ± 6.0	55.4 ± 5.2^	64.8 ± 4.5^

scenarios. In the experimental design, participants interacted with the ads via mobile devices, and the system automatically collected data such as click count and interaction duration. To ensure data representativeness, interactive tasks were categorized into five categories, covering the main interactive forms found in advertising. After testing, the occurrence and duration of each type of action were statistically analyzed, and subjective user experience was analyzed using user feedback questionnaires. All eligible participants signed an informed consent form. This method not only reflects audience engagement in interactive advertising but also verifies the effectiveness of the dynamic visual recognition system in enhancing user immersion and deepening brand communication.

Table 2 shows user engagement and experience in different interactive tasks within augmented reality ads, including statistical analysis of engagement rate, interaction duration, number of actions, task completion rate, and user satisfaction. The data showed that the task of viewing fabric details had the highest engagement rate of 92.5%, with an average interaction duration of 18.4 seconds, 3.8 actions, a task completion rate of 95.8%, and a user satisfaction score of 4.6, indicating that users are most actively engaged in product details. The product color switching task had a participation rate of 85.3% and 5.2 actions, indicating frequent user interaction with color changes but short durations. The 360° product rotation task, while only receiving a 68.4% participation rate, had an average interaction duration of 22.5 seconds and a high satisfaction score of 4.7, reflecting that despite relatively small participation, the experience was highly immersive. Sharing ad content and saving product information had lower participation rates of 45.2% and 52.7%, respectively, but completion rates for both tasks were above 95%, indicating that while active participation was limited, there was a clear willingness to take action. Overall, Table 2 demonstrates the significant effectiveness of dynamic visual recognition systems in enhancing user interaction experience, increasing immersion, and brand engagement. It also reveals the differences in appeal and user behavior across different interactive tasks.

TABLE 2. Comprehensive Evaluation Results of Interactive Engagement

Interaction Task Type	Participation Rate (%)	Avg. Interaction Duration (s)	Avg. Interaction Count (times)	Task Completion Rate (%)	User Satisfaction (1–5)
View Fabric Details	92.5	18.4 ± 4.2	3.8 ± 1.2	95.8	4.6
Switch Product Color	85.3	12.7 ± 3.5	5.2 ± 1.8	88.7	4.3
Click Brand Logo	78.6	8.9 ± 2.8	2.5 ± 0.9	92.4	4.1
Rotate Product 360°	68.4	22.5 ± 5.1	1.8 ± 0.7	82.6	4.7
Share Ad Content	45.2	6.3 ± 2.1	1.2 ± 0.5	98.5	3.8
Save Product Information	52.7	7.8 ± 2.4	1.5 ± 0.6	96.3	4.2

IV. CONCLUSION

This paper addresses the limitations of textile industry advertising communication and proposes a technical application study based on dynamic visual recognition systems. The research develops an object detection module based on YOLOv8, which is used to extract fabric textures, clothing outlines, and brand logos. This module, combined with a convolutional neural network, achieves multi-dimensional feature modeling. It also generates interactive advertising scenes through augmented reality rendering, forming a complete technical path from detection and modeling to dynamic presentation. Experimental design demonstrates that this system offers significant advantages in increasing ad attention, enhancing brand recall, and promoting interactive engagement, effectively overcoming the limitations of traditional static advertising, which suffer from limited information expression and inefficient dissemination. However, current research still faces challenges in real-time performance optimization and cross-platform application. Future work can combine lightweight models with cloud-based rendering technology to promote the large-scale application of dynamic visual recognition in textile industry advertising, providing a more efficient and scalable solution for brand communication.

AUTHOR CONTRIBUTIONS

Conceptualization – Bin Zhang and Qian Zhang; methodology – Bin Zhang and Qian Zhang; investigation – Bin Zhang and Qian Zhang; writing-original draft preparation – Bin Zhang and Qian Zhang. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research was supported by, This paper represents one of the research achievements of the 2025 horizontal scientific and technological project of Weifang Vocational College, namely “Research on the Application of Dynamic and Interactive VI (Visual Identity System) Design” (Project No.: WZHX2025011).

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] H. Li and N. Zhang, “Computer Vision Models for Image Analysis in Advertising Research,” Journal of Advertising, vol. 53, no. 5, pp. 771-790, 2024, doi: 10.1080/00913367.2024.2407644.

[2] P. Alipour, E. E. Gallegos, and S. Sridhar, “AI-driven marketing personalization: Deploying convolutional neural networks to decode consumer behavior,” International Journal of Human-Computer Interaction, vol. 41, no. 16, pp. 10266-10284, 2025, doi: 10.1080/10447318.2024.2432455.

[3] M. A. Moreno-Armendariz, H. Calvo, J. Faustinos, and C. A. Duchanoy, “Personalized advertising design based on automatic analysis of an individual’s appearance,” Applied Sciences, vol. 13, no. 17, p. 9765, 2023, doi: 10.3390/app13179765.

[4] J. Xu, Y. Zhou, L. Zhang, J. Wang, and D. Lefloch, “Sportswear retailing forecast model based on the combination of multi-layer perceptron and convolutional neural network,” Textile Research Journal, vol. 91, no. 23-24, pp. 2980-2994, 2021, doi: 10.1177/00405175211020518.

[5] I. Ali, M. U. Saleem, A. A. Khan, A. Naz, M. Nawaz, and H. Khan, “An Enhanced Artificial Intelligence Generated Virtual Influencer Framework: Examining the Effects of Emotional Display on User Engagement based on Convolutional Neural Networks (CNNs),” Annual Methodological Archive Research Review, vol. 3, no. 4, pp. 184-209, 2025, doi: 10.1016/j.jretconser.2023.103560.

[6] L. Yang, R. Kumar, R. Kaur, A. Babbar, G. S. Makhanshahi, A. Singh, et al., “Exploring the role of computer vision in product design and development: a comprehensive review,” International Journal on Interactive Design and Manufacturing (IJIDeM), vol. 18, no. 6, pp. 3633-3680, 2024, doi: 10.1007/s12008-024-01765-7.

[7] N. Akram, R. A. Butt, and M. Amir, “Modification of a convolutional neural network for the weave pattern classification,” Mehran University Research Journal of Engineering & Technology, vol. 43, no. 2, pp. 79-90, 2024, doi: 10.22581/muet1982.2998.

[8] K. G. Nalbant and S. Aydin, “Development and transformation in digital marketing and branding with artificial intelligence and digital technologies dynamics in the Metaverse universe,” Journal of Metaverse, vol. 3, no. 1, pp. 9-18, 2023, doi: 10.57019/jmv.1148015.

[9] H. Halepoto, T. Gong, S. Noor, and H. Memon, “Bibliometric analysis of artificial intelligence in textiles,” Materials, vol. 15, no. 8, p. 2910, 2022, doi: 10.3390/ma15082910.

[10] X. Li and Y. Zhu, “A real-time and accurate convolutional neural network for fabric defect detection,” Complex & Intelligent Systems, vol. 10, no. 3, pp. 3371-3387, 2024, doi: 10.1007/s40747-023-01317-8.

[11] W. K. Jung, J. Kang, W. Kwon, and H. Kim, “StitchingNet and deep transfer learning method for sewing stitch defect detection,” Journal of Computational Design and Engineering, vol. 12, no. 4, pp. 140-154, 2025, doi: 10.1093/jcde/qwaf037.

[12] T. P. Nguyen, H. Nguyen, and H. Q. T. Ngo, “Toward a sustainable transition in automated production: enabling the vision-based approach and synchronous control for textile surface inspection systems,” Textile Research Journal, vol. 93, no. 23-24, pp. 5391-5415, 2023, doi: 10.1177/00405175231199257.

[13] M. Glogar, S. Petrak, and M. Mahnic Naglic, “Digital technologies in the sustainable design and development of textiles and clothing-a literature review,” Sustainability, vol. 17, no. 4, p. 1371, 2025, doi: 10.3390/su17041371.

[14] T. A. Rainy, “AI-driven marketing analytics for retail strategy: a systematic review of data-backed campaign optimization,” International Journal of Scientific Interdisciplinary Research, vol. 6, no. 1, pp. 28-59, doi: 10.63125/0k4k5585.

[15] N. Ingle and W. J. Jasper, “A review of deep learning and artificial intelligence in dyeing, printing and finishing,” Textile Research Journal, vol. 95, no. 5-6, pp. 625-657, 2025, doi: 10.1177/00405175241268619.

[16] B. Wei, B. Xu, K. Hao, and L. Gao, “Textile defect detection using multilevel and attentional deep learning network (MLMA-Net),” Textile Research Journal, vol. 92, no. 19-20, pp. 3462-3477, doi: 10.1177/00405175211073773.

[17] S. Singh, “Artificial intelligence in the fashion and apparel industry,” Tekstilec, vol. 67, no. 3, pp. 225-240, 2024.

[18] Y. Li and R. C. Mat, “The Impact of Visual Image on Brand Communication: A Multi-Dimensional Interaction Model Analysis,” Journal of Logistics, Informatics and Service Science, vol. 10, no. 4, pp. 250-266, 2023.

[19] N. Skliarenko, I. Gryshchenko, and M. Kolosnichenko, “Symmetry in the visual communication design: methods of dynamic image construction,” Art and design, vol. 3, no. 4, pp. 9-20, 2021, doi: 10.30857/2617-0272.2021.3.1.

[20] M. Mohiuddin Babu, S. Akter, M. Rahman, M. M. Billah, and D. Hack-Polay, “The role of artificial intelligence in shaping the future of Agile fashion industry,” Production Planning & Control, vol. 35, no. 15, pp. 2084-2098, doi: 10.1080/09537287.2022.2060858.

[21] A. T. Abdirahmonovich, Z. Hakimov, B. Tursunov, and A. Oqboyev, “Evaluation of competitiveness of brands of local sewing and knitting enterprises,” Revista Geintec-gestao Inovacao E Tecnologias, vol. 11, no. 2, pp. 716-739, 2021, doi: 10.47059/REVISTAGEINTEC.V11I2.1707.

[22] R. Wagner and A. Kabalska, “Sustainable value in the fashion industry: A case study of value construction/destruction using digital twins,” Sustainable Development, vol. 31, no. 3, pp. 1652-1667, doi: 10.1002/sd.2474.

[23] Z. Wang, Y. Liu, Z. Zhou, P. Chen, and H. Peng, “Towards integrated textile display systems,” Nature Reviews Electrical Engineering, vol. 1, no. 7, pp. 466-477, 2024, doi: 10.1038/s44287-024-00063-4.

[24] J. Nascimento and S. M. C. Loureiro, “Mapping the sustainability branding field: emerging trends and future directions,” Journal of Product & Brand Management, vol. 33, no. 2, pp. 234-257, doi: 10.1108/jpbm-02-2023-4349.

[25] A. Spry, B. Figueiredo, L. Gurrieri, J. A. Kemper, and J. Vredenburg, “Transformative branding: A dynamic capability to challenge the dominant social paradigm,” Journal of Macromarketing, vol. 41, no. 4, pp. 531-546, doi: 10.1177/02761467211043074.