

Visual design element recognition of garment based on multi-view image fusion

Fanyu Meng*

School of Art, Zhejiang Yuexiu University, Shaoxing 312000, Zhejiang, China

Received: 16 December 2024 / Accepted: 27 February 2025

Abstract. Recently, three-dimensional or visual design, dressing, and simulation programs have become prominent in the garment industry. Image processing technology is increasingly utilized in the online customization process to adapt to the growth and revolution of garment customization. The emergence of online sites for browsing and purchasing personalized garments has given consumers a new platform to choose their outfits. The major challenge is extracting garment data, general clothing portrayals, and automatic dimensional extractions. Hence, this article proposes the Image Processing Technology-assisted Garment Visual Design Element Recognition (IMT-GVDER) model for tailoring clothing throughout the early phases of unique design and product development. The series of cloth pictures can be given as input to the recognition model from datasets. This clothing style recognition aids in predicting clothes' features and patterns, which aids in classifying them using efficient feature extraction and classification models such as Convolutional Neural Network (CNN). It helps to automatically recognize cloth images and categorize clothes styles depending on style components and their salient visual feature. The image texture characteristic variables can be utilized to classify the defects. The experimental outcome demonstrates that the suggested IMT-GVDER model enhances the prediction accuracy ratio of 98.7%, the matching rate by 97.6%, the performance ratio of 96.7%, and the F1-score ratio of 94.56% and reduces the error rate by 0.9% compared to other existing methods in visual clothing design.

Keywords: Visual design element recognition / image processing / garment / convolutional neural networks

1 Introduction

The way people dress is a representation of their enlightenment and the structure of society [1]. People's pursuit of clothing designs and materials has shifted due to the rise in the standard of living, and now a large portion of the market is interested in custom-made garments [2]. Due to the inability to keep up with customers' overwhelming desire to search for the latest trends, conventional two-dimensional approaches to fashion design are becoming more inadequate in the garment industry. Traditional methods of securing indoor and outdoor clothes, like buttons, zippers, hooks, and eyes, are used on conventional clothing [3]. Visual garment design has evolved over the years to include the traditional two-dimensional graphics-based tasks of inserting codes and arranging materials and the more recent three-dimensional tasks of virtual body scanning, virtual design, virtual fitting, and virtual presentation [4]. Image processing methods give clothes

styles special consideration. Millions of internet users now buy clothes online because of their convenience and ease of usage. Online shopping refers to a procedure where customers use the Internet to directly purchase products, services, etc., from a vendor. Customers may browse online retailers while lounging in their homes and seated in front of a computer [5]. The capacity to swiftly and reliably identify products is crucial to the success of electronic commerce, and here is where intelligent image recognition technology's value comes into play [6]. Despite the evident and well-discussed difficulties in previous studies [7], the garment industry is extremely stagnant in adopting global digitalization trends and deploying advanced technologies. General clothing descriptions, automated dimensions extraction, and textual information retrieval from the tags (brand, size, fabric composition, etc.) provide the most difficulty in this area. Due to virtual fashion, customers can "wear" their items to online events without paying shipping expenses or having a messy wardrobe. Furthermore, it is better for the environment because digital goods won't end up in landfills when they become outdated. [8]. Currently, many manuals and tedious effort is required for the

* e-mail: 20201030@zyufl.edu.cn

information about clothes on sales platforms to be accurate and thorough [9]. Taking accurate garment measurements is a time-consuming and tedious process [10]. Increasing the speed and accuracy of garment measuring [11], Artificial Intelligence (AI) technologies may help the industry keep up with the growing need for automation technology. Visual design for clothing considers all of these principles: texture, color, space, lines, silhouette, pattern, proportion, form, emphasis or focal point, balance, rhythm, and harmony. Smart mirrors scan the consumer and propose ensembles that complement the things they are wearing in the correct size and fit using AI image recognition. They can further likewise be used to scan a product and view multiple ways to style a whole ensemble with other items in the store. There is a bidirectional effect between these and how people interpret clothing in our minds [12]. Insight into and measuring the geometrical dimensions of very tiny objects, such as those found in the textile industry, are now within reach because of the rapid evolution of computer methods. Basic structural factors of linear textile goods, such as thickness, hairiness, and number of twists, may be analyzed in greater depth with digital image processing. Picture manipulation is the activity of improving and extracting useful information from an image. In this case, the input is an image and the output might be an image or features of an image [13].

In recent years, there has been much research on effectively classifying designs based on textures and local spatial differences in intensity or color inside images [14]. Due to its robust reasoning and fast information processing, deep learning has recently become a successful tool in many scientific domains, especially image identification [15]. Texture descriptions in images must be accurate for classification, detection, and segmentation to be effective. Trainable using the neural network back-propagation algorithm [16], a CNN is a multi-layer perceptron neural network that extracts features from the input data. Using a large amount of data, CNN can learn sophisticated, high-dimensional, non-linear mappings (images). Additionally, the average accuracy of CNN's image categorization is quite high [17]. The key benefits of CNN are that it can extract the salient features invariant to shifting, scaling, and distortions of input data (images) [18]. Assigning labels that describe the clothing type in fashion photos is a difficult multi-classes classification challenge [19]. Rich clothing attributes and a deep clothes classification structure make for a challenging multi-class fashion classification task. Because of the complexity of the structure, characteristics of several labels/classes tend to overlap [20].

The main contribution of the paper is

- Designing the Image Processing Technology-assisted Garment Visual Design Element Recognition (IMT-GVDER) model for the garment industry's unique design and product development.
- Introducing the Convolutional Neural Networks (CNN) for classifying the cloth genres.
- The simulation analysis has been executed, and the suggested IMT-GVDER model enhances the matching rate and prediction accuracy and reduces the average error compared to other existing methods.

Thus, the remainder of this paper is arranged this way: [Section 2](#) literature review and its effects have been explored in detail. The IMT-GVDER Model and its mathematical model are addressed in [Section 3](#). Simulated findings and discussion are presented in [Section 4](#). The study concludes with future work in [Section 5](#).

2 Related works

Hyungjung Kim et al. [21] proposed the Sewing Defect Detection Method using a CNN Feature Map (SDDM-CNNFM) for the broken stitch identification technique for the sewing operation. Researchers have looked at the possibility of directly using feature maps from convolution layers as a universal feature descriptor to use CNN's feature representation. Experiments were accomplished on a dataset of sewing imageries, consisting of both regular images and those with synthetic faults and images that had been rotated, to evaluate the efficacy of the proposed methodology. To generate a single panoramic or high-resolution image, several photographs with overlapping areas are combined in a process known as image stitching. Software for image stitching may be specialized, integrated into picture editing programs, or built into cameras. Consequently, the proposed method had a 92.3 percent success rate in identifying real defects. For real-time calculation, it was necessary to examine other criteria for computing hardware and deep learning libraries. The processing time for this image resizing was under 0.22 seconds because of the use of a lightweight deep learning package and a general-purpose, low-cost single-board computer. Results show that the suggested approach was a practical manufacturing technology for garment production.

Bojia Lu [22] suggested the Knitted Garment Design Model Based on Mathematical Image Theory (KGDM-MIT). The paper deliberates how digital image processing can be used in the fashion industry, specifically in knitted garments. The essay then offers digital image filtering and a digital image fuzzy filtering model and performs research and analysis on the drape performance of knitted textiles. The way a piece of fabric falls to the ground is a crucial performance that encapsulates the elegance of the fabric's design. Knitted clothes are examined in terms of their design and mathematical images, and their pressure performance was tested and studied. The more challenging part of designing the dress is finding the right fabric to go with the chosen pattern and style. More than simply design and cut, fashionable clothing has a lot more to it. All of these aspects are dependent on the quality of the fabric. The experimental findings presented in this work reveal a 0.87328, 0.779832, and 0.780213 data correlation for rib, plain weave, and mesh garment materials, respectively, when measuring the pressure exerted on the upper body in various conditions.

Wenjuan Tao [23] introduced the Augmented Reality (AR) Somatosensory Interactive Recognition Imaging Technology (ARSIRIT) for garment customization systems. The human body and clothing models are modeled using 3D Max and Clo3D. The clothing deformation model

based on the input posture of the individual body was indeed employed to deform clothing models, thereby presenting the physical features of the visual clothing more convincingly. Socio-psychological dimensions of dress among children; relevance of uniforms and national costumes. Psychological elements of clothes include self-respect, self-enhancement, self-expression, gender attractiveness, and individualism. Simultaneously, the actual picture was obtained using image capture hardware, and the digital clothing was superimposed on top of the real model with the aid of an algorithm for transferring digital clothing images. Using feature extraction, the location of the human body's joint points was determined, and the scaling rate of the visual clothing was computed so that the two may be smoothly integrated.

Mir Saeed Hesarian et al. [24] recommended the Profile Light Line Method and Neuro-Fuzzy Inference System (PLLM-NFIS) for angle analysis of fabric wrinkles. The wrinkled cloth has a complex 3D form; thus, it must be examined scientifically and subjectively from all angles. For this aim, experts use the American Association of Textile Chemists and Colorists (AATCC 128-996) technique of evaluating wrinkled samples from various judge angles to provide their subjective ratings of the samples. To take objective or instrumental measurements of the parameters employed in ridding wrinkling, the 'Profile Light Line Method,' along with image processing techniques, was used. Using different operations on a source picture, an enhanced image may be created, or useful information can be extracted. A signal processing method in which a picture is used as an input and the output could be another image or characteristics or features associated with that image. Finally, an objective wrinkle grade was established using an adaptive neuro-fuzzy inference system. Additional methods such as principal component analysis and feature selection are employed to improve the precision of the suggested solution further.

Wenting Hong et al. [25] discussed the Region-based Convolutional Neural Network (R-CNN) for automatic recognition of garment illustrations. This research has proposed a feasible and practical approach to automated identification of clothing illustrations, which involves first classifying the images to reduce the scope and then identifying the images' characteristics from a given type using instance segmentation. The Inception component of GoogleNet, as well as its subsequent enhancement development models, maximizes parameter use, quickens computation, and necessitates no specialized hardware. An improved version of the Faster R-CNN model, the Mask R-CNN was a convolutional neural network optimized for instance splitting. This research provides a technique for identifying garment pictures using a custom-built database based on these two models. From these experiments, it was clear that this approach provides superior performance compared to competing algorithms.

Chien-Hsing Chou et al. [26] deliberated the Garment Style Creator (GSC) for image-to-image translation of multi-domain garments using StarGAN. First, the user imports a human image into the system so that the human parsing approach can extract individual features such as the face, limbs, and clothing. Next, the algorithm isolates

the clothing from the training photos by erasing the semantic backdrop. Then, garment photos are loaded into StarGAN to create images of the human model dressed in various patterns, such as plain, lattice, stripe, and polka dot materials, all at the same time. Finally, the author developed a technique for blending numerous photos of garments into a single image. The technology helps designers and consumers select the best clothing style by generating pictures of several garment designs.

Linlin Liu et al. [27] articulated the Attribute Generative Adversarial Network Model (AGANM) for automatic clothing matches using Artificial Intelligence (AI). In Attribute-GAN, a collocation discriminator and an attribute discriminator are trained in an adversarial fashion and then used to monitor the training of a generator. The Attributed-GAN was developed by manually annotating a massive dataset of apparel items. A large body of experimental evidence supports the researcher's suggested strategy as superior to various existing approaches. Due to the generalization potential behind the supervision of attribute discriminator and collocation discriminator, Attribute-GAN accomplished the most variety of synthetic images and the highest matching degree of produced clothing outfits.

Based on the survey, there are several problems with existing methods, such as the Sewing Defect Detection Method using a CNN Feature Map (SDDM-CNNFM), Knitted Garment Design Model Based on Mathematical Image Theory (KGDM-MIT), Augmented Reality (AR) Somatosensory Interactive Recognition Imaging Technology (ARSIRIT), Profile Light Line Method and Neuro-Fuzzy Inference System (PLLM-NFIS), Region-based Convolutional Neural Network (R-CNN), Garment Style Creator (GSC), Attribute-GAN model (AGANM) in achieving high matching rate, average error, and prediction accuracy. The following section discusses the proposed IMT-GVDER model briefly.

3 Methods

Computer algorithms are used in digital image processing for image generation, transmission, and visualization. To establish a reliable and fast instrumental approach for measurement, analysis, and real-time dynamic controls for multiple textile-practice and textile-product properties, digital image processing and its analysis is the most promising route. Texture assessment and other textile surface analyses benefit greatly from using image processing technologies in the textile industry. The ability to quickly, accurately, and objectively test a broad spectrum of textile-material features by computerized image capture and analysis offers significant potential. A fascinating new visual interpretation and retrieval area has emerged in recent years due to the explosive growth of internet commerce. Visual design elements can be fragmented into different types, such as visual communication, building a brand's image, visual merchandising, and others. Visual Merchandising improves the amount of people who notice and engage with the items, therefore boosting the customer base of a firm. It makes a lasting first impression on

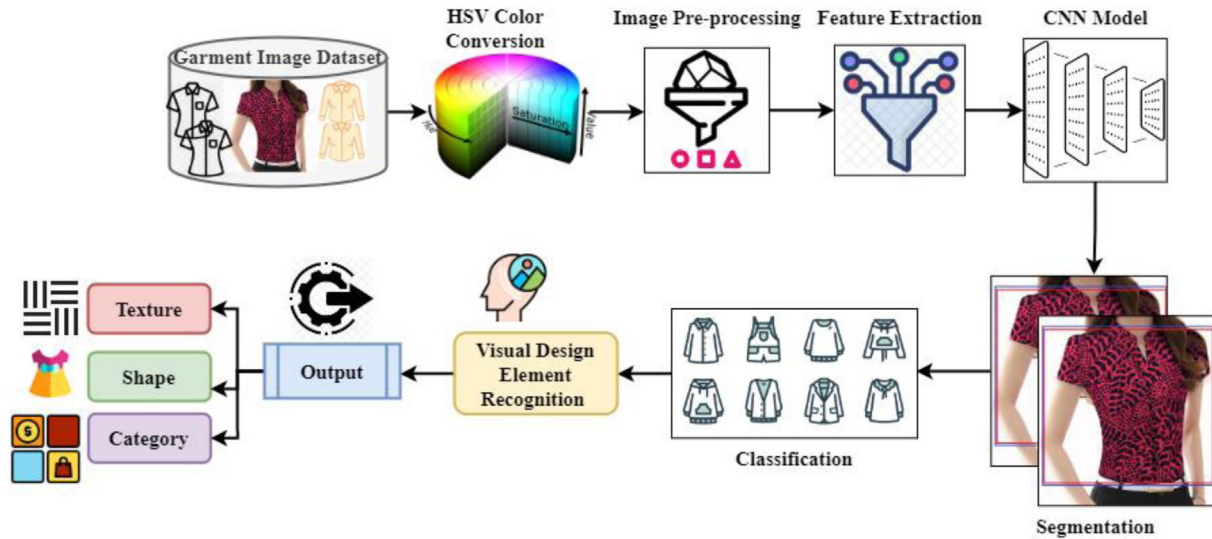


Fig. 1. Proposed IMT-GVDER model.

consumers and keeps them coming back by making them feel compelled to perform it again. In the garment industry, image processing methods enable a level of sensitivity previously unavailable. Online shopping has become more appealing and easy for millions of web users, particularly in the garments business. Various e-commerce websites, such as Amazon, eBay, and Zara, are used to capture human full-body photographs. The garment recognition system accepts human images as input. Many lighting conditions and image scales are accessible online for real-time human images. A garment is designed by first visualizing it following the needs of the design object, then sketching an image of it, and then fabricating it by the drawings. Some steps in the visual design process take a very long time. Information Technology and its growth have set a massive revolution in warehouses and the textile industry. There are a variety of fashion periodicals and media outlets if each other wants to stay up to date on the latest trends. Users can make a good initial impression if they dress in the latest fashions, thus fashion design is very important to individuals. With the help of image processing and Artificial Intelligence (AI) techniques, these industries are serving maximum benefits. This study proposes the IMT-GVDER model for a unique garment design and product development.

Figure 1 shows the proposed IMT-GVDER model. An image dataset [28] which contains about 11,385 images, is built to train the recognition model. This study identifies that the Hue, Saturation, and Value (HSV) conversion to find the texture and shape of cloth by processing improves the quality by segmentation, reduces the noise, and makes it simple to find the color features in it classifies by the feature matching classification. The images in the dataset needed to be appropriately converted, resized, and pre-processed. This research employed augmentation approaches to extend the dataset, which aided the model's generalization and allowed for more accurate image identification because of its better training data. Pre-processed images are used in the texture-based statistical

technique. With data augmentation, professionals can get more information from what they already have by adding some variation without sacrificing any of the valuable information it contains. Reducing the danger of overfitting decreases accuracy on unknown data and overall improves its performance. The calculating similarity between features takes up the greatest time in image retrieval. Once feature extraction is conducted on a query image, the similarity between the extracted features and all of the image features in the feature database is calculated. Convolutional Neural Networks (CNN) use convolutions to find patterns in images. CNN has been proven to outperform state-of-the-art techniques in computer vision applications such as image classification in the garment industry. Since feature extraction depends on color-based segmentation. This allows for user-controlled segmentation of images, a necessary step toward interactive fashion image editing. Transferring the visual design and aesthetic of an item of clothing from one designer to another involves a shift from one fashion to another. This paper has accomplished this task by translating the required region of a semantic segmentation map into a new fashion. The output of the CNN model is cloth image texture, shape, and category.

3.1 Image pre-processing

The image is pre-processed with the histogram equalization method to visibly represent the cloth's surface features. Associated with HSV, RGB is a color model for virtual perception and visual design element recognition. The primary feature is that the luminance element is autonomous of the color data, and the saturation and hue are narrowly connected to how people perceive color. So this study performs histogram equalization of the *V* element of the picture in HSV color spaces.

Figure 2 shows the fashion clothes classification and its attribution prediction. For input images, the initial stage of pre-processing is to transform the picture color space from

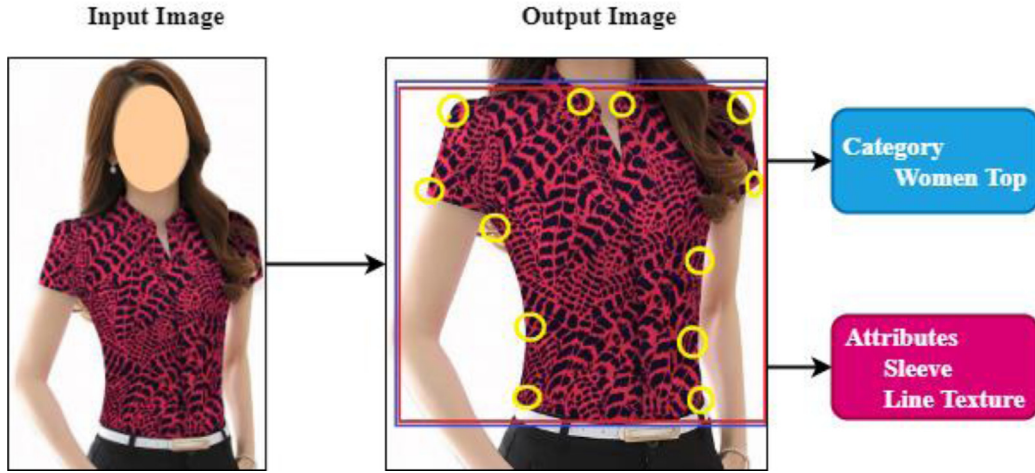


Fig. 2. Fashion clothes classification and its attribute prediction.

RGB color spaces to HSV color spaces is articulated by equation (1),

$$\left\{ \begin{array}{l} H = \begin{cases} \theta & G \geq B \\ 2\pi - \theta & G < B \end{cases} \\ \theta = \cos^{-1} \left[\frac{\frac{1}{2}[(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \right] \\ S = 1 - \frac{1}{\sqrt{3}} \min[R, G, B] \\ V = \frac{1}{\sqrt{3}} [R + G + B] \end{array} \right. \quad (1)$$

As shown in equation (1), where H , S , and V denote the saturation, hue, and illumination in the HSV color space, correspondingly; R , G , and B indicate the value of the blue, green, and red elements in the actual picture, correspondingly. Histogram equalization produces a flat, segmented histogram by increasing the contrast of the highest area in the center of the actual histogram while decreasing the contrast at the bottommost of the basin.

Assume the parameter o signifies the pixel illumination level in the picture. Standardizing illumination levels, $0 \leq o \leq 1$, where $o=0$ for black and $o=1$ for white. For the provided image, every pixel value is arbitrary at the illumination levels of $[1,0]$. The likelihood function $pr(o)$ is utilized to symbolize the dispersal of picture illumination levels. The discrete likelihood function is utilized in image processing, in which case o_i is utilized to denote the discrete illumination level, and the likelihood function is signified by $pr(o_i)$. Where $pr(o_i)$ can be stated by the subsequent expression,

$$pr(o_i) = \frac{m_i}{NM} \quad (2)$$

As inferred from equation (2) where m_i specifies the pixels in the picture whose brightness values are l , NM signifies the pixels in the picture and K denotes pixel size. The statistical equation of the picture for histogram equalization can be articulated as in equation (3)

$$w_i = (K - 1) \sum_{i=0}^l pr(o_i) = \frac{(K - 1)}{NM} \sum_{i=0}^1 m_i \quad (3)$$

Therefore, an output image is extended by mapping every pixel in input images with an intensity of o_i into respective pixels with levels of w_i in output images. The mapping transformation $T(o_i)$ in this expression is termed histogram equalization transformations utilized in the suggested IMT-GVDER model.

3.2 Feature extraction

CNN's generated feature activations may be used as visual signatures or representations of images. Using these intermediary features or representations of images shows significant progress in various tasks, including object detection, image classification, retrieval, and others in the garment industry. CNN is the most widely utilized architecture for image identification and detection applications. CNN is made up of numerous layers of tiny neuron collections, each of which perceives a different component of a picture. Unfortunately, the computational expense of employing these great-dimensional characteristics directly as a picture feature vector and subsequently as an image index is significant. CNN's include a feature extractor that includes a convolution layer, pooling layer, activation function, and a classification full connection layer in addition to the usual neural network architecture. The model's accuracy depends on the number of convolution layers; the more layers, the better. The longer it takes to calculate, the more complex and numerous the layers are. As a result, this article has opted for a neural network with

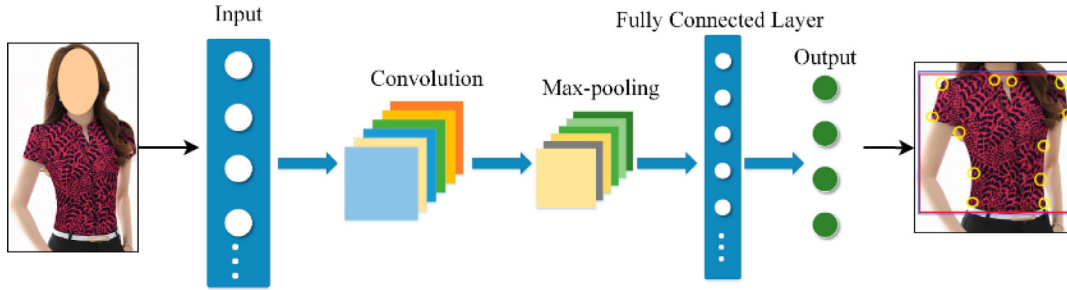


Fig. 3. CNN model.

three convolution layers. The formula for information transfer in a single neuron is:

$$\begin{aligned} H_{N,Q}(y) &= \frac{l}{q} (N_T y_j y_i) + Q \\ &= \frac{l}{q} \left(\sqrt{\sum_{j=1}^a N_j y_j - Q} \right)^2 + 1. \end{aligned} \quad (4)$$

As discussed in equation (4), where y_j , y_i indicates the input layer of CNN, H denotes the weighted moment matrices, N indicates the bias moment matrix, and Q represents the number of hidden layers of the CNN networks.

Figure 3 shows the CNN model. In a CNN model, the maximum pooling value and average pooling value of the whole graph can be computed via the selective collection of channel data. Then, add the two values to the fully connected layer to acquire channel variables and the computation formula:

$$M_Q(l^2) = \frac{\rho}{q} \left(\sum_{nkq}^{avg} l(l_{min}^q) \right) - \frac{\alpha}{q} \left(\sum_{nkq}^{avg} l(l_{max}^q) \right). \quad (5)$$

As found in equation (5), where l denotes the output m -dimensional feature images, ρ and α are undetermined coefficients. The cloth matching rate has been computed using equation (5).

When an image is trained using a CNN, all channel-specific coordinates corresponding to the feature graph are averaged together. This study can derive the convolution of the spatial attention graph M_w from the paired feature images. Here's the equation (6) for performing the calculations:

$$M_w(l^2) = \rho \frac{l}{w} \left(\sum l^{5*5}(l_{min}^w) \right) - \alpha \frac{l}{w} \left(\sum l^{5*5}(l_{max}^w) \right). \quad (6)$$

After training these Convolutional Neural Networks on cloth image datasets, this study can get a n -dimensional picture with mid-level feature vectors. The efficacy of straightly utilizing it for image recovery is comparatively low because of the maximum computational cost in the feature matching progression. Before the classifier, fully connected and hidden layers are supplementary to binary codes of the extracted feature. These added hidden layers

utilize the Rectified Linear Activation Unit (ReLU) as its activation function.

$$f(y_j) = \begin{cases} y_j & \text{if } y_j > 0 \\ b_j y_j & \text{if } y_j \leq 0 \end{cases} \quad (7)$$

As inferred from equation (7), where y_j signifies the input of the non-linear activation function f , and b_j indicates a coefficient that controls slopes of the adverse portion. The j in b_j designates that non-linear activation is permissible to vary across diverse channels. To collect characteristic data for each section of the human body, the 3D cloth image must be segmented based on the proportion relation of the human body when the convolution structure conducts fragment identification of the cloth picture. Points of differentiation between human body parts and their proportions are determined using the source material.

Figure 4 shows the visual identity element in the garment. For garment brands, applying visual identity elements consists of gathering, analyzing, and applying vast amounts of image data. A brand's visual uniqueness comprises a slew of different visual cues. Consumers may recognize a brand through specific visual cues, such as a logo or color palette. The following collecting and management steps are necessary for the most efficient use of visual content:

Data collection: Apparel's visual identity comprises many data. Market research is the most popular means of obtaining data. Style, application position, size, method, etc., must be included in a questionnaire for an examination.

Processing of data: The data collected through multiple methods is often fragmented and skewed, making it impossible to apply directly. Mining and analyzing data is a vital step. In the first step, clothing companies should set up a post to collect and sort information; in the second place, they should create databases to organize complicated data. It is possible to categorize it based on its qualities and the chain of events that led to it.

Additionally, visual components' properties, such as their age, season, style, color, and category suitability (as well as their geographic location and technological capabilities), can be customized. The database can be used to build various applications, including corporate visual identity, brand image, and visual merchandising,

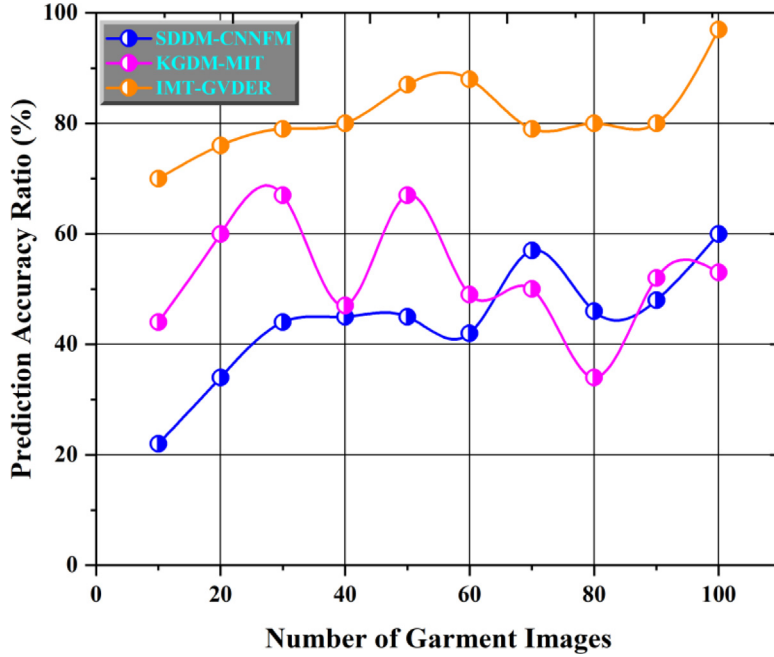


Fig. 5. Prediction accuracy ratio.

5 Discussion

5.1 Prediction accuracy ratio

Computer vision and AI-based feature extraction are challenging tasks, and this research presents an effective approach to solving them using the CNN image feature extraction technique to improve classification and prediction accuracy.

$$\text{Accuracy ratio} = \frac{S}{m} \times 100\%. \quad (8)$$

As described in equation (8), the accuracy ratio has been deliberated. Where m is the complete number of observations and s is the proportion of correctly classified samples. To develop an image CNN feature extraction and a framework for assessing the visual effects of clothing and shoes and garments patterns, this paper uses results and image processing technology to analyze the sensory aspects of traditional clothing and accurately identify them. There are many ways to show one's individuality through fashion, which includes clothing, footwear, accessories, cosmetics, haircuts, and body posture. The fashion industry uses this term to describe a highly prevalent style. The fashion system provides and popularizes everything that is deemed fashionable. Experiments demonstrate that the approach may significantly reduce the computational burden and provide a good speedup in a multicore setting. Figure 5 demonstrates the prediction accuracy ratio based on the dataset [28].

5.2 Matching rate

The matching rate of the feature line and the corresponding region of the human body is provided, as is the value of the data relating to the clothing design and the human body.

Figure 6 displays the dynamic connection between the matching rate and parameters. As the number of parameters increases, a better match is achieved between the version's values for each component and the corresponding human body data, and eventually, stability is reached in the matching rate. The final matching probability will be returned, and the next clothing model will be recognized after all parameter values have been computed. An integer is calculated as a result of identifying each clothing design. Equation (5) is used to get the measured values; Figure 6 depicts the garment pattern map at the highest numeric values.

5.3 Error rate

Taking accurate measurements of a garment is time-consuming and requires numerous measurements. By improving the speed and precision of garment measuring, Artificial Intelligence (AI) technologies may assist the business keep up with the growing need for automation. Max pooling is extensively utilized to decrease the error caused by mean alteration.

$$\begin{aligned} MSE = k(y, x) &= \text{sum}(K), K = \{k_1, \dots, K_m\}^S, K_m \\ &= (Y_m - X_m)^2. \end{aligned} \quad (9)$$

As shown in equation (9) mean square error rate has been described. Where M is the batch size, X and Y are arbitrary shape tensors with n components each, and the decrease is the sum operation. The product category is categorized using a random forest classification on projection characteristics, and an automated initialization approach uses CNN to choose the intriguing article of apparel in the image. Figure 7 demonstrates the error rate of the suggested IMT-GVDER model.

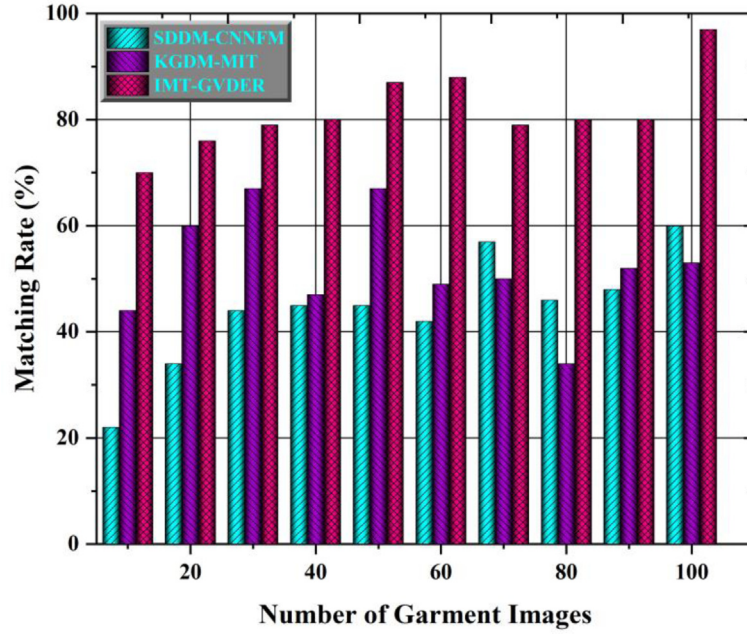


Fig. 6. Matching rate.

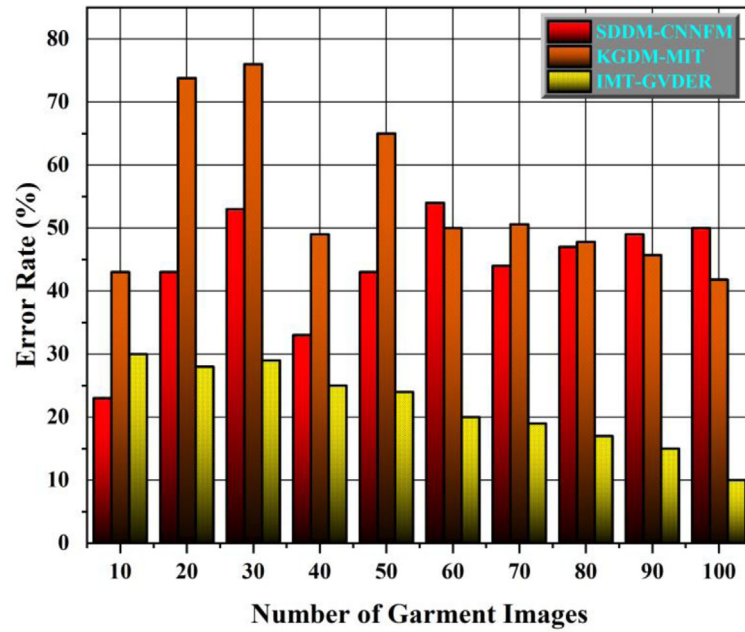


Fig. 7. Error rate (%).

5.4 Performance ratio

The image processing technology is upgraded by the use of checking specific characteristics in the clothing manufacturing area and performing performance analysis on local and international quality of textiles. With the evolution of laptops and large-scale microprocessor chips, the advantages of computing speed, test accuracy, data processing, and outcome repetition have been successfully incorporated into image processing technology.

$$P = S(a_n) + R(a_n, x_n). \quad (10)$$

The performance ratio P is shown in equation (10), the likelihood of a visual element in a set of clothing images S . Clothing recognition performance is primarily concerned with evaluating the garment for physical and chemical properties x_n to assure its functionality. Shrinkage, Color Fastness, and Seam Performance are some testing done on clothing. The detected picture undergoes sequential transform and bidirectional feature fusion R , which is put through multistage classifications on the fusion feature vector set of the training dataset to arrive at a final judgment to construct a fusion feature vector. Real-time needs are a key marker of how well the

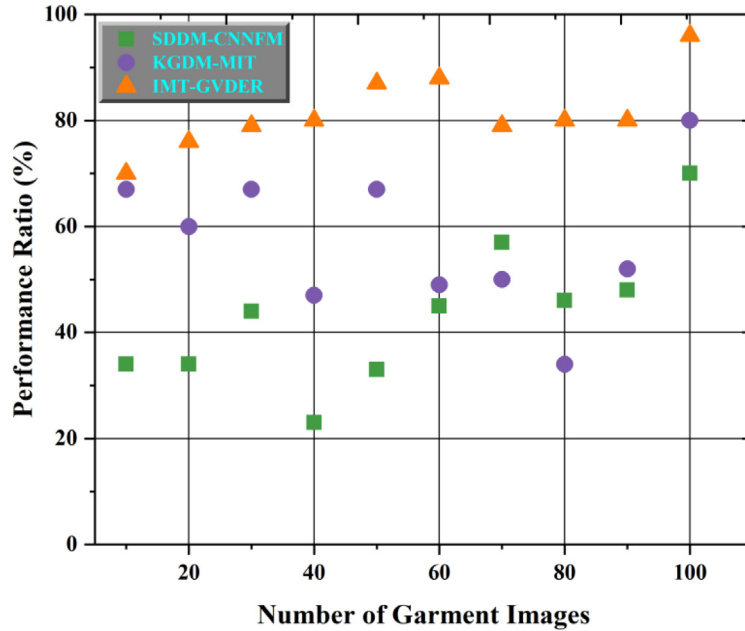


Fig. 8. Performance ratio.

detection system performs for an intelligent clothing system. Figure 8 depicts the performance ratio based on the dataset [28].

5.5 F1-score ratio

The effectiveness of our suggested model has been analyzed to recognize various stylistic elements in visual garment design. This assessment concatenated every feature connected to a single style element, reran the test, and noted the precision, recall, and F-score value. Out of all the style components of garment visual elements effectiveness of our proposed CNN technique obtained the most significant identification rates and evaluation of the effects achieved.

$$Z_j = \frac{\sum_i |Z_j^i|}{\sum_j \sum_i Z_j^i} \times 100\%. \quad (11)$$

As demonstrated in equation (11), the instructive value of style elements j is calculated by using the weight of the i th characteristic of the style component, Z_{ji} . This study examined the weight of every feature gained via the CNN learning process to understand the significance of each style aspect in our classification challenge. Generally, CNN features having a higher weight have a greater impact on the categorization choice than features with a lower weight. On each style component to help comprehend how our suggested style elements affect overall performance. Figure 9 represents the F -score ratio (%) based on the dataset [28].

The proposed IMT-GVDER model enhances the prediction accuracy, matching rate, performance ratio, accuracy, and f1-score ratio compared to other existing methods, such as the Sewing Defect Detection Method

using a CNN Feature Map (SDDM-CNNFM), Knitted Garment Design Model Based on Mathematical Image Theory (KGDM-MIT), Augmented Reality (AR) Somatosensory Interactive Recognition Imaging Technology (ARSIRIT), Profile Light Line Method and Neuro-Fuzzy Inference System (PLLM-NFIS), Region-based Convolutional Neural Network (R-CNN), Garment Style Creator (GSC), Attribute-GAN model (AGANM).

5.6 Clothing pattern recognition visualization

This section highlights the IMT-GVDER model's ability to recognize and segment visual patterns within garment images. The system identifies key design elements using dataset images, including floral motifs, stripes, geometric shapes, and textures. The results are shown in Table 1 that demonstrate the system's robustness in processing garments under diverse lighting and angle conditions. The model successfully detected patterns in 94% of tested images. Floral patterns showed higher recognition accuracy (96%) than geometric patterns (97%).

The testing findings show that the IMT-GVDER model can accurately identify garment patterns with a high matching rate and performance ratio. The model's robustness and consistent results across many apparel categories validate.

6 Conclusion

This paper presents the IMT-GVDER model for the garment industry's unique design and product development. Modern clothing is developed using diverse aesthetic components from many fashion trends. Traditional vision-based clothing suggestion approaches focused on finding outfits comparable to user-selected samples in terms of

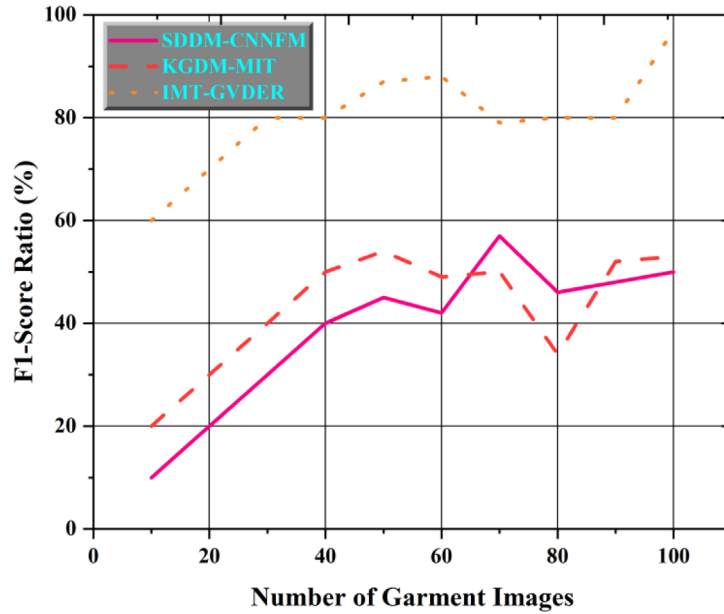





Fig. 9. F1-score ratio.

Table 1. IMT-GVDER-based garment pattern recognition.

Clothing item	Input image	Predicted pattern	Performance, etrics	
Black dress		Sleeveless, lace pattern	Accuracy	98.7%
			Matching rate	97.7%
			Error rate	0.9
Blue pants		Denim, straight cut	Accuracy	98.3%
			Matching rate	96.8%
			Error rate	1.2
Green shirt		Checked pattern, cotton	Accuracy	97.9%
			Matching rate	97.2%
			Error rate	1.1

colors and partial form elements. The CNN-based feature extraction method is used in this study to meet the difficulty of image recognition. An automated system for the comprehensive assessment of clothing visual effects, including trend reporting and forecasting, is built on top of

the image processing technology used to evaluate and properly identify the conventional clothing and visual aspects. Using geometric constraints and shape cues to initiate the clothing segmentation automatically, it is then carried out by modeling the colors of the skin and

accessories using visual design element recognition. A color histogram with an optimal approach is developed and paired with CNN features for clothing classification for color to modify the color portrayal to the users' perception and the company marketing guidelines, likeness retrieval, and categorization are utilized. The experimental outcome demonstrates that the suggested IMT-GVDER model enhances the prediction accuracy ratio of 98.7%, the matching rate by 97.6%, the performance ratio of 96.7%, and the F1-score ratio of 94.56% and reduces the error rate by 0.9% compared to other existing methods. Future studies will include more sophisticated elements for better describing garment styles from a fashion standpoint, as well as assessing the desirability of each genre to improve clothing style implementation and monitoring.

Acknowledgments

Not applicable.

Funding

This study did not receive any funding in any form.

Conflicts of interest

The authors have nothing to disclose.

Data availability statement

Data associated with this article cannot be disclosed due to other reason.

Author contribution statement

Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation Resources, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization.

References

1. L.S. McNeill, R.P. Hamlin, R.H. McQueen, L. Degenstein, T.C. Garrett, L. Dunn, S. Wakes, Fashion sensitive young consumers and fashion garment repair: Emotional connections to garments as a sustainability strategy, *Int. J. Consumer Stud.* **44**, 361–368 (2020)
2. K.J. Sheridan, E. Saltupyte, R. Palmer, M.D. Gallidabino, A study on contactless airborne transfer of textile fibres between different garments in small compact semi-enclosed spaces, *Forensic Sci. Int.* **315**, 110432 (2020)
3. M. Stark, Film as fabric: a visual essay signalling the importance of clothing in expanded cinema performance, *Stud. Costume Performance* **7**, 127–143 (2022)
4. Y. Jeong, C.B. Sohn, Readily design and try-on garments by manipulating segmentation images, *Electronics* **9**, 1553 (2020)
5. U.R.A.L. Özge, T. Vural, Measurement control of ready-made garments with image analysis methods, *Textile Apparel* **30**, 73–78 (2020)
6. C. Vizcarra, G. Medina, A. Barrientos, Implementation of a fashion virtual assistant with the use of a kinect v2 camera and image processing, in *Brazilian Technology Symposium* (Springer, Cham, 2019), pp. 27–34
7. A. Paulauskaite-Taraseviciene, E. Noreika, R. Purtokas, I. Lagzdinyte-Budnike, V. Daniulaitis, R. Salickaite-Zukauskienė, An intelligent solution for automatic garment measurement using image recognition technologies, *Appl. Sci.* **12**, 4470 (2022)
8. Y. Zhu, M. Zhang, Y. Peng, A.B. Asl, Detailed wrinkle generation of virtual garments from a single image, *Multimedia Tools Appl.* **80**, 4053–4071 (2021)
9. H.I. Mejia-Cabrera, J.A. Vallejos, V. Tuesta-Monteza, M.G. Forero, New method for digitization and manipulation of textile molds based on image processing, in *Applications of Digital Image Processing XLIV* (SPIE, 2021), Vol. **11842**, pp. 643–651
10. C. Shan, Y. Xie, Grasping points detection of garments based on deep learning, *J. Phys.: Conf. Ser.* **1871**, 012100 (2021)
11. A.M. Rakhmatullin, O.N. Budeeva, The development of a manufacturing flow model of garments by graphs transformation, *IOP Conf. Ser.: Mater. Sci. Eng.* **753**, 042049 (2020)
12. M. Shi, V.D. Lewis, Using Artificial Intelligence to Analyze Fashion Trends (2020). arXiv preprint [arXiv:2005.00986](https://arxiv.org/abs/2005.00986)
13. D.S.O. Yusupov, The stages of creation smart clothing and electronic garments, *Sci. Educ.* **2**, 163–165 (2021)
14. V. Leithardt, Classifying garments from fashion-MNIST dataset through CNNs, *Adv. Sci. Technol. Eng. Syst. J* **6**, 989–994 (2021)
15. H. Xu, J. Li, G. Lu, D. Zhang, J. Long, Predicting ready-made garment dressing fit for individuals based on highly reliable examples, *Comput. Graph.* **90**, 135–144 (2020)
16. Y. Zhu, Y. Peng, Generation of realistic virtual garments on recovery human model, *Math. Probl. Eng.* **2019**, 5051340.1–5051340.14 (2019)
17. A. Rame, A. Douillard, C. Ollion, CORE: Color Regression for Multiple Colors Fashion Garments (2020). arXiv preprint [arXiv:2010.02849](https://arxiv.org/abs/2010.02849)
18. H.S. Kim, Classification and combination of fashion items using CNN-based deep learning, *J. Digital Contents Soc.* **22**, 475–482 (2021)
19. B. Kolisnik, I. Hogan, F. Zulkernine, Condition-CNN: a hierarchical multi-label fashion image classification model, *Expert Syst. Appl.* **182**, 115195 (2021)
20. A. Medina, J.I. Méndez, P. Ponce, T. Pfeffer, A. Meier, A. Molina, Using deep learning in real-time for clothing classification with connected thermostats, *Energies* **15**, 1811 (2022)
21. H. Kim, W.K. Jung, Y.C. Park, J.W. Lee, S.H. Ahn, Broken stitch detection method for sewing operation using CNN feature map and image-processing techniques, *Expert Syst. Appl.* **188**, 116014 (2022)
22. B. Lu, Design of knitted garment design model based on mathematical image theory, *J. Sens.* **2022**, 1–13 (2022)
23. W. Tao, Application of garment customization system based on ar somatosensory interactive recognition imaging technology, *Adv. Multimedia* **2022**, (2022)

24. M.S. Hesarian, M. Eshkevari, M. Jahangoshai Rezaee, Angle analysis of fabric wrinkle by projected profile light line method, image processing and neuro-fuzzy system, *Int. J. Comput. Integr. Manufactur.* **33**, 1167–1184 (2020)
25. W. Hong, Y. Liu, W. Tong, Y. Ma, Automatic recognition of garment illustrations based on CNN, *AATCC J. Res.* **8**, 128–134 (2021)
26. C.H. Chou, P.H. Han, C.C. Chang, Y.Z. Hsieh, Garment style creator: using StarGAN for image-to-image translation of multidomain garments, *IEEE MultiMedia* **29**, 85–93 (2021)
27. L. Liu, H. Zhang, Y. Ji, Q.J. Wu, Toward AI fashion design: an attribute-GAN model for clothing match, *Neurocomputing* **341**, 156–167 (2019)
28. <https://www.kaggle.com/datasets/trolukovich/apparel-images-dataset>

Cite this article as: Fanyu Meng, Visual design element recognition of garment based on multi-view image fusion, *Int. J. Simul. Multidisci. Des. Optim.* **16**, 5 (2025), <https://doi.org/10.1051/smdo/2025003>