INTERNATIONAL REVIEW OF ARTIFICIAL INTELLIGENCE IN MARKETING

NAWALA

Volume 2, Number 1, 2023

Leveraging Computer Vision for Visual Merchandising: AI Applications in Online Retail

Syakira Gavrila Haerus [1]

[1] Telkom University, Bandung, Indonesia

Abstract

Article history:
Received: January 21, 2023
Revised: February 13, 2023
Accepted: April 7, 2023
Published: June 30, 2023

Keywords:
Computer Vision, Electronic Commerce, Online Retail, Visual Merchandising, Virtual Try On.

Identifier:
Nawala
Page: 50-62
https://nawala.io/index.php/iraim

This conceptual article examines how computer vision can be leveraged as a strategic resource for visual merchandising in online retail. It first reviews marketing and neuroscience evidence showing that digital visual design elements such as image quality, layout, and aesthetic coherence strongly shape engagement, perceived value, and purchase intentions in electronic commerce interfaces. The study then synthesises prior research on image based recommendation, visual search, and virtual try on, which use deep learning models to curate visually consistent assortments, enable "shop by image" journeys, and create interactive fitting room experiences on consumer devices. Using a systematic literature review approach, the article maps these applications onto core visual merchandising objectives related to product presentation, assortment framing, and cross selling. The resulting framework highlights that the greatest potential of computer vision lies in integrating these tools into coordinated visual journeys that enhance customer experience and commercial outcomes, while opening avenues for future empirical and design oriented research in diverse electronic commerce categories.

*Corresponding author:
(Syakira Gavrila Haerus)

1. Introduction

Visual merchandising has traditionally referred to the strategic design of window displays, store layouts, and in store atmospherics that guide attention and shape shoppers' emotions. In online retail environments, these functions are increasingly performed by digital visual elements such as product imagery, page layout, color schemes, and interactive presentation tools. Empirical work on online visual merchandising shows that cues like image quality, richness of product presentation, and on-site promotional visuals significantly influence flow experience, satisfaction, and approach behaviour, highlighting that visual design is a primary carrier of value in e-commerce interfaces (Sina & Wu, 2018). At the same time, consumers are moving from text-oriented browsing toward highly visual journeys, where discovery, evaluation, and even trust are constructed through images and videos rather than long descriptions.

Parallel advances in artificial intelligence, especially computer vision, have transformed how these visual elements can be created, organised, and personalised. Deep convolutional neural networks now allow retailers to extract rich feature representations from product images and use them for image based recommendation systems that mitigate information overload and improve match quality between products and customer preferences (Malekpour Alamdari et al., 2022). Large platforms such as JD.com and Alibaba have deployed real time visual search engines that retrieve similar items from catalogues containing hundreds of millions of images within sub second response times, demonstrating the technical

feasibility of visual merchandising that is dynamically assembled around each query image rather than statically designed (Li et al., 2018; Dagan et al., 2021).

These capabilities translate into a growing family of applications at the marketing and merchandising interface. Visual search and "shop by image" features let customers begin their journey with a photo instead of a keyword, while image based recommender systems populate "similar items" and "complete the look" carousels based on visual compatibility rather than only collaborative or textual signals (Malekpour Alamdari et al., 2022; Zhong et al., 2022). Virtual try on technologies extend visual merchandising further by overlaying garments, accessories, or cosmetics on consumer images, increasing perceived usefulness and enjoyment and thereby strengthening purchase intentions in online fashion contexts (Zhang et al., 2019). Recent large scale log analyses of visual e commerce search show that image based interaction patterns differ systematically from text search, which has implications for how retailers should design visual funnels and allocate display real estate on product listing pages (Dagan et al., 2021).

Despite these developments, much of the existing literature either focuses on the psychological effects of static online visual merchandising or on technical performance metrics of computer vision systems, with relatively little integrative work that treats computer vision as a strategic merchandising resource in its own right. There is still limited understanding of how specific vision based applications, such as visual search, image based recommendation, and virtual try on, reconfigure core merchandising decisions related to product presentation, assortment framing, and cross selling in online retail environments. By examining how computer vision

can be leveraged systematically in online retail, the discussion seeks to address this gap in the literature. It reviews and synthesises recent advances in visual search and image based recommendation, maps them onto established visual merchandising objectives, and develops a conceptual framework that links computer vision enabled merchandising practices to customer experience and commercial outcomes, providing a foundation for future empirical and design oriented research.

2. Literature Review

Existing studies on visual merchandising increasingly emphasize that visual stimuli are central to how consumers process and evaluate retail environments. Neuroscientific work shows that creative and story driven visual merchandising elicits stronger prefrontal activation and more positive emotional responses than more conventional displays, suggesting that richer visual compositions can deepen engagement and perceived value (Kim et al., 2020). Although this evidence is grounded in physical fashion stores, it underscores that visual complexity, coherence, and aesthetic fit are not merely decorative, but core drivers of attention, affect, and downstream purchase-related responses. Translated into online contexts, these insights imply that the design of product galleries, hero images, and on page visual arrangements should be considered a strategic merchandising lever rather than a purely operational design task.

At the same time, computer vision has reshaped how visual merchandising decisions can be automated and personalized in digital channels. Deep learning based recommendation models extract high level visual features from product

images and use them to suggest visually similar or stylistically coherent items, thereby aligning online assortments with users' style preferences at scale (Tuinhof et al., 2019). In fashion e-commerce, aesthetic-aware models go further by explicitly modelling aesthetic attributes, showing that recommendations that incorporate aesthetic features better capture individual taste and improve ranking performance compared with systems that rely only on generic convolutional features (Yu et al., 2018). This body of work suggests that computer vision can move beyond simple similarity matching toward curating visually consistent collections that function as dynamic, data driven shop windows.

Visual search extends this logic by allowing consumers to initiate their journey with an image rather than a textual query. Large scale implementations demonstrate that convolutional visual embeddings can support fast and accurate retrieval across very large catalogues, enabling "shop the look" and "find similar" journeys that are tightly aligned with a user-supplied reference image. Compared with keyword search, image based retrieval reduces the friction of articulating product attributes in language and may better capture tacit style preferences, which is particularly relevant in visually driven categories such as fashion and home décor. From a merchandising perspective, visual search effectively reconfigures category navigation and shelf exposure, because the system can assemble context specific assortments around each query image instead of relying solely on fixed taxonomies and filters.

Finally, virtual try on and augmented reality applications illustrate how computer vision can transform visual merchandising from static presentation into interactive, embodied experience. Empirical evidence from beauty and cosmetics

contexts shows that augmented reality try-on experiences enhance perceived control, stimulate cognitive and behavioural engagement, and significantly raise purchase intentions, with consumer control acting as a key mediator (Whang et al., 2021). These findings indicate that vision-based overlays which allow users to "see" products on their own faces or bodies do not only reduce uncertainty, but also create immersive visual journeys that blur the line between merchandising and product usage. Taken together, prior research demonstrates the potential of computer vision for recommendation, search, and virtual try on, but it has rarely been synthesised through the lens of visual merchandising objectives. The present study therefore builds on this literature to conceptualise how these disparate applications can be integrated into a coherent, computer vision enabled visual merchandising strategy in online retail.

3. Methods

This study employed a systematic literature review to synthesise existing knowledge on how computer vision is used for visual merchandising in online retail. Relevant publications were identified through structured searches in major scholarly databases such as Scopus, Web of Science, and Google Scholar using combinations of keywords related to visual merchandising, computer vision, online retailing, visual search, recommendation systems, and virtual try-on or augmented reality. The search was restricted to peer-reviewed journal articles and high-quality conference proceedings in marketing, retailing, information systems, and computer science. After removing duplicates, the records were screened in several stages: titles and

abstracts were first examined to exclude studies that focused solely on traditional offline visual merchandising, generic artificial intelligence in marketing without visual components, or purely technical computer vision work without an explicit retail or merchandising context. The remaining articles were assessed in full text against predefined inclusion criteria, namely that they addressed applications such as visual search, image-based recommendation, aesthetic or style modelling, or virtual try-on in an online retail setting and reported either conceptual frameworks or empirical findings related to customer experience or commercial outcomes. For each included study, data were extracted on research purpose, methodological approach, computer vision technique, type of visual merchandising application, and key outcomes such as engagement, perceived value, purchase intention, or sales effects. The evidence was then synthesised using a narrative and thematic approach, mapping computer vision applications onto core visual merchandising objectives and deriving a conceptual framework that highlights prevailing design patterns, performance implications, and unresolved research gaps.

4. Results and Discussion

The systematic review reveals a consistent pattern in which computer-vision applications enhance the traditional objectives of visual merchandising by making product presentation more immersive, personalised, and responsive. Studies on online visual merchandising show that high image quality, rich product displays, and well designed promotional visuals increase flow experience, satisfaction, and approach behaviour, confirming that visual design is a central carrier of value in

digital interfaces (Sina & Wu, 2018). When these visual elements are combined with neuroscientific evidence that creative, story-driven displays trigger stronger emotional responses and perceived value, visual merchandising emerges as a core driver of affect and attention rather than a purely aesthetic layer (Kim et al., 2020). Together, these findings support the view that any computer-vision system deployed in online retail should be evaluated not only on technical accuracy but also on its contribution to emotionally engaging and coherent visual narratives.

Across the reviewed studies, computer vision mainly supports three clusters of applications: image-based recommendation, visual search, and virtual try on or augmented reality. Deep learning recommendation models that extract high level visual features from product images enable retailers to curate assortments that are visually coherent and aligned with individual style preferences at scale (Tuinhof et al., 2019; Malekpour Alamdari et al., 2022). Aesthetic aware models further show that explicitly modelling aesthetic attributes improves the match with individual taste compared with relying only on generic visual features (Yu et al., 2018). These results indicate that computer vision driven recommendation systems can function as dynamic, data driven shop windows that directly operationalise visual merchandising principles such as harmony, balance, and style consistency in an automated way.

The evidence on visual search complements this picture by demonstrating how image based journeys restructure online navigation. Large scale implementations of visual search and "shop by image" features on major platforms illustrate that convolutional visual embeddings support fast and accurate retrieval from very large catalogues, thereby allowing consumers to begin their journey with

a reference image rather than a keyword (Li et al., 2018; Dagan et al., 2021). Compared with text search, these interactions reduce the cognitive effort of verbalising product attributes and better capture tacit style preferences, particularly in visually driven categories such as fashion and home décor (Zhong et al., 2022). From a merchandising standpoint, this means that category exposure and cross selling are no longer constrained by static taxonomies and filters, since assortments can be assembled contextually around each query image, which strengthens the alignment between visual intent and displayed products.

Virtual try on and augmented reality applications add a further layer by transforming visual merchandising from static display into interactive, embodied experience. Empirical findings in beauty and cosmetics indicate that augmented reality try-on enhances perceived control, stimulates cognitive and behavioural engagement, and increases purchase intentions, with perceived control playing a central mediating role (Zhang et al., 2019; Whang et al., 2021). When viewed through a visual merchandising lens, these systems do not only reduce uncertainty about fit or appearance; they extend the "fitting room" into the interface and embed the product directly into the consumer's own image. This shifts merchandising decisions from arranging items on a digital shelf toward designing immersive visual journeys in which the consumer's body or face becomes part of the display.

Overall, the review shows that existing studies support a positive link between computer-vision enabled applications and key outcomes such as engagement, perceived value, and purchase intention across recommendation, search, and virtual try-on contexts. However, most contributions still evaluate isolated tools in terms of

technical metrics or immediate behavioural outcomes, while giving limited attention to how these tools jointly reconfigure core merchandising decisions about product presentation, assortment framing, and cross selling. The conceptual framework developed in this study therefore synthesises these strands by positioning computer vision as a strategic visual merchandising resource. It highlights that the greatest potential lies in integrating image based recommendation, visual search, and virtual try on into coordinated visual scenarios that tell a consistent story across touchpoints, an area that remains underexplored and offers promising directions for future empirical and design oriented research.

5. Conclusion

This article concludes that computer vision can no longer be viewed as a purely technical add on to online retail systems, but as a strategic resource for visual merchandising. The review demonstrates that image-based recommendation, visual search, and virtual try on collectively extend classic merchandising objectives into digital environments by making product presentation more immersive, personalised, and context sensitive. Evidence from prior work shows that when these applications are aligned with established principles of visual complexity, coherence, and aesthetic fit, they enhance engagement, perceived value, and purchase intentions, confirming that visual design remains a central driver of customer response in online retail.

At the same time, the synthesis highlights that current research is still fragmented. Most studies examine individual tools or algorithms in isolation, focusing on accuracy and short term behavioural metrics rather than on how

computer vision reshapes core merchandising decisions about product presentation, assortment framing, and cross selling. By mapping different computer-vision applications onto visual merchandising objectives, this study proposes a conceptual framework that positions image-based recommendation, visual search, and virtual try on as complementary components of a coordinated visual strategy. This perspective underscores that the greatest potential lies not in any single technology, but in integrating these tools into coherent visual journeys that are consistent across touchpoints.

For practitioners, the findings suggest that investments in computer vision should be guided by merchandising logic, not only by technical capability. Retailers need to consider how visual algorithms support narrative consistency, style coherence, and perceived control, and how they can design interfaces where recommendation, search, and try on experiences reinforce each other as parts of a unified visual environment. For researchers, the framework points to several promising directions, including multi tool evaluations, longitudinal studies of customer experience across visual touchpoints, and investigations of ethical and fairness issues in visually driven personalisation. Addressing these questions will be essential to fully understand and govern the role of computer vision in the future of visual merchandising.

References

Alamdari, P. M., Navimipour, N. J., Hosseinzadeh, M., Safaei, A. A., & Darwesh, A. (2022). An image-based product recommendation for E-commerce applications using convolutional neural networks. *Acta Informatica Pragensia, 11*(1), 15–35.

Dagan, A., Guy, I., & Novgorodov, S. (2021, July). An image is worth a thousand terms? Analysis of visual e-commerce search. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 102–112).

Kim, H. S., Lee, J. H., & Yoo, S. H. (2020). Is consumer neural response to visual merchandising types different depending on their fashion involvement? *PLOS ONE, 15*(12), e0241578.

Li, J., Liu, H., Gui, C., Chen, J., Ni, Z., Wang, N., & Chen, Y. (2018, December). The design and implementation of a real time visual search system on JD E-commerce platform. In *Proceedings of the 19th International Middleware Conference Industry* (pp. 9–16).

Sina, A. S., & Wu, J. (2018, January). The effect of online visual merchandising cues on consumers' flow experience, satisfaction, and approach/avoidance behavior. In *International Textile and Apparel Association Annual Conference Proceedings* (Vol. 75, No. 1). Iowa State University Digital Press.

Tuinhof, H., Pirker, C., & Haltmeier, M. (2018, September). Image-based fashion product recommendation with deep learning. In *International Conference on*

Machine Learning, Optimization, and Data Science (pp. 472–481). Cham: Springer International Publishing.

Whang, J. B., Song, J. H., Choi, B., & Lee, J. H. (2021). The effect of augmented reality on purchase intention of beauty products: The roles of consumers' control. *Journal of Business Research, 133*, 275–284.

Yu, W., Zhang, H., He, X., Chen, X., Xiong, L., & Qin, Z. (2018, April). Aesthetic-based clothing recommendation. In *Proceedings of the 2018 World Wide Web Conference* (pp. 649–658).

Zhang, T., Wang, W. Y. C., Cao, L., & Wang, Y. (2019). The role of virtual try-on technology in online purchase decision from consumers' aspect. *Internet Research, 29*(3), 529–551.

Zhong, H., Wang, Y., & Yue, W. (2022). An e-commerce product recommendation method based on visual search and customer satisfaction. *International Journal of Knowledge and Systems Science, 13*(1), 1–14.