International Journal of Research Publication and Reviews

Conversational Image Recognition Chatbot for Fashion

Harshavardhika K[1], Nejiya K S[2], Sujithra R[3], Reshma R[4], Mr. S. Janarthanan[5]

[1,2,3,4] Department: AIML Organization: Raise Smart City, Coimbatore, India
[5] Mentor, Raise Smart City, Coimbatore, India
Email address: [1]harshavardhikak@gmail.com, [2]ksnejiya@gmail.com, [3]sujithra675@gmail.com, [4]reshma6761@gmail.com

ABSTRACT—

The project develops an advanced conversational image recognition chatbot customized for the fashion industry, which is expected to revolutionize the way fashion-related content is accessible for consumers online. The core function of the chatbot features the fusion of cutting-edge computer vision and NLP techniques enabling users to upload images of different fashion pieces, such as clothing, accessories, and different outfits, and even communicate in real-time to gather information on those items. The recognition function provides key insights into minutiae in the uploaded photographs via image recognition algorithms, basically demographics—types of fabric, clothing category, colours, patterns, and even brand logos. Based on this information, users get instantaneous insight on their uploaded objects. Besides, this application can give personalized recommendations based on the likes and dislikes set by a user on the fly during the course of their conversation. This makes the whole application widely variable in any place. When users inquire about similar products or look for styling advice, the chatbot receives both the visual and text inputs and can use them to recommend alternative options or suitable complementary pieces, thus providing a good-customized and personalized shopping experience. This proves especially useful in the fashion industry where personalization plays a decisive role in customer satisfaction. Here, the users inquire about the latest fashion trends, seek styling tips, or request recommendations on affordable alternatives to high-end items based on the image inputs.

*Keywords—*chatbot, image recognition, fashion industry, natural language processing, personalized recommendations and styling tips

I. Introduction

Today, the fashion sector is undergoing a significant digital transformation with the rise of e-commerce and prominence of personalized shopping experiences among the users globally. The increasing consumer reliance on online platforms for their fashion needs, makes the brands and retailers propel toward creating a more interactive and engaging shopping journey which transcends the experience of ordinary traditional browsing. Outdated online shopping relies excessively on text-based searching and static product descriptions; thus, this makes online shopping less engaging to entice the users. Instead, consumers seek personalized, smart, and dynamic ways of engaging with clothing to reach a purchasing decision. This shift of transformations encompasses the development of a model of generative systems and newer technologies to address the needs of the ever-evolving digital consumer. AI and machine learning are opening a world of possibilities that would revolutionize e-shopping, especially visual recognition and conversational interfaces. With advancements in computer vision and NLP, artificial intelligence can now intertwine visual data with textual data to make way for even more interactive and personalized experiences. Arguably one of the most exciting applications of these technologies are the conversational image-recognition chatbots, which provide users with a novel manner of interacting with fashion goods.
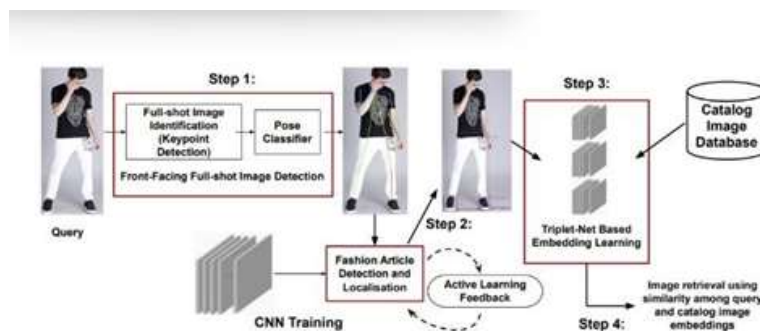
II. LITERATURE SURVEY

Image Recognition in Fashion

Computer vision has seen significant advancements in recent years, making it an essential tool for fashion applications. One prominent development is the use of deep learning models, particularly convolutional neural networks (CNNs), for image classification and object recognition. Fashion-specific datasets like DeepFashion, Fashion-MNIST, and the Clothing1M dataset have been instrumental in training these models to identify a wide range of fashion items, from clothing to accessories. In a study by Liu et al. (2016), the authors introduced DeepFashion, a large-scale dataset specifically designed to assist in understanding fashion- related images. The dataset enabled the creation of models that could recognize clothing types, detect landmarks (such as buttons or collars), and classify garments by style, which are crucial capabilities for a fashion-focused image recognition system. Another key study by understanding fashion-related images. The dataset enabled the creation of models that could recognize clothing types, detect landmarks (such as buttons or collars), and classify garments by style, which are crucial capabilities for a fashion- focused image recognition system. Another key study by Yang et al. (2020) utilized deep learning techniques to develop a fashion recommendation system based on images. Their work focused on enhancing the accuracy of image recognition by incorporating both visual and textual data, providing better context for the item in the image. This dual input approach laid the groundwork for the integration of NLP with computer vision in fashion e-commerce, offering more robust and meaningful recommendations.
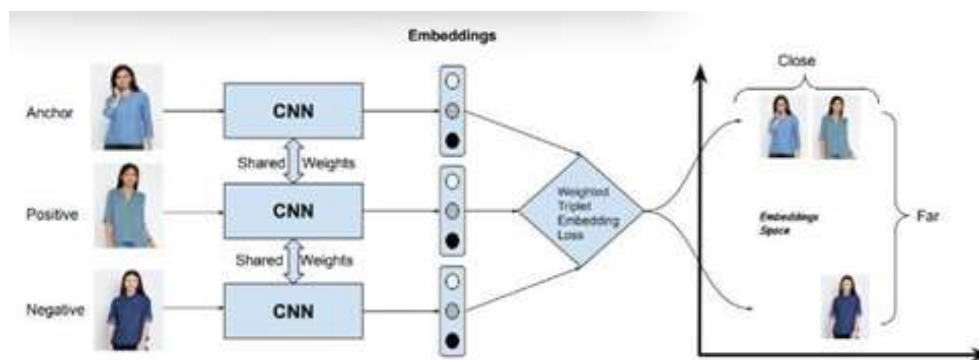
Conversational AI and Fashion

While image recognition has made strides, the integration of conversational AI is what sets these systems apart in providing a more engaging user experience. Early advancements in chatbots were primarily focused on text-based interfaces, which lacked the ability to handle multimodal inputs, such as images. However, over the past decade, significant progress has been made in conversational agents, especially with the advent of models like GPT-3 and BERT from OpenAI and Google, respectively. These NLP models excel in understanding and generating human- like dialogue, enabling more fluid and context- aware conversations. In the fashion domain, conversational AI systems have been applied to create virtual stylists and shopping assistants. For instance, Cunha et al. (2019) developed a conversational recommender system that leverages NLP to interact with users about their fashion preferences and

suggest items based on user profiles. However, such systems are often limited to textual input, requiring users to explicitly describe their fashion needs, which may not always be intuitive or effective.

III. PROBLEM STATEMENT

The online fashion shopping experience often lacks the personalized, interactive, and intuitive elements that consumers expect in today's digital landscape. Traditional search-based e-commerce systems, which rely primarily on textual queries and filters, do not fully capture the nuanced preferences of users, especially when it comes to visual products like clothing and accessories. Users may struggle to express their fashion preferences through words alone, making it difficult to find items that match their exact needs or style. Additionally, while fashion images are a core element of online shopping, current platforms fail to integrate image recognition with personalized, conversational experiences, limiting user engagement and satisfaction.

Furthermore, despite the growth of AI-powered recommendation systems in e-commerce, these systems often rely on historical user data or broad category-based suggestions, rather than incorporating real-time visual feedback. While image recognition has made strides in identifying fashion items, existing systems lack the ability to provide meaningful context around these items, such as styling tips, brand recommendations, or personalized advice based on the user's unique preferences and past interactions.

IV. PROPOSED WORK

The proposed work aims to develop a conversational image recognition chatbot designed for the fashion industry, combining computer vision and natural language processing (NLP) to create a more interactive and personalized shopping experience. The system will allow users to upload images of fashion items, such as clothing or accessories, which will be processed by the image recognition module. Using deep learning techniques, particularly Convolutional Neural Networks (CNNs), the system will identify the items, classify them by category (e.g., dress, shoes, accessories), and extract visual attributes like color, style, and fabric.

Once the image is processed, the chatbot will engage users through natural language, responding to text-based queries with context- aware, personalized suggestions. Leveraging advanced NLP models such as GPT-4 or BERT, the chatbot will understand user inputs, offer fashion advice, recommend

complementary items, and suggest alternative styles based on user preferences and past interactions. For example, if a user uploads an image of a dress and asks for styling tips, the chatbot might recommend pairing it with accessories or other outfits based on current fashion trends.

V. RESULT ANALYSIS

Image Recognition Accuracy: One of the primary indicators of success will be the accuracy with which the system can recognize and classify fashion items from user-uploaded images. The performance of the image recognition module will be assessed using standard metrics such as precision, recall, and F1 score. The system's ability to correctly identify different clothing types, patterns, and colors will be compared to state-of-the-art fashion recognition models. A high level of accuracy will indicate that the system can reliably understand user inputs, forming the foundation for personalized recommendations.

Natural Language Processing (NLP) Effectiveness: The chatbot's ability to engage in meaningful, context-aware conversations will be evaluated based on its understanding of user queries and the relevance of its responses. The effectiveness of the NLP module will be assessed through metrics such as response relevance (i.e., how well the chatbot's suggestions match the user's intent), conversation fluency, and user satisfaction. User interactions will be logged and analyzed to measure the chatbot's ability to handle multi-turn conversations and adapt to changing user needs.

Recommendation Relevance: The recommendation engine's effectiveness will be evaluated based on how well it provides personalized suggestions that align with the user's preferences. Metrics such as click-through rate (CTR), conversion rate (the number of users who make a purchase after receiving a recommendation), and user satisfaction surveys will be used to assess the quality of recommendations. The system's ability to adapt to user feedback over time will also be analyzed to measure the improvement in the relevance of recommendations with repeated interactions.

User Engagement and Satisfaction: User engagement will be measured through metrics like session length, frequency of interactions, and repeat usage, providing insight into how engaging and useful the chatbot is in the long term. User satisfaction surveys and feedback will provide direct insights into the chatbot's perceived usefulness, ease of use, and overall value in enhancing the shopping experience. These insights will help identify areas for improvement, such as refining the conversational abilities or enhancing the diversity of fashion recommendations.

Comparative Analysis: To assess the performance of the chatbot relative to existing fashion recommendation systems, the proposed system will be compared to traditional text-based recommendation systems and other multimodal systems. Key performance indicators (KPIs), such as user retention, purchase behavior, and overall satisfaction, will be compared to understand the added value of integrating image recognition and conversation.

VI. CONCLUSION

In conclusion, the proposed conversational image recognition chatbot aims to enhance the online fashion shopping experience by integrating advanced image recognition and natural language processing. By enabling users to upload images and engage in personalized, context-aware conversations, the system provides tailored fashion recommendations and styling advice. The combination of computer vision and conversational AI offers a more intuitive, interactive, and engaging shopping journey, leading to improved user satisfaction and increased sales. This approach has the potential to redefine e-commerce in the fashion industry, offering a scalable solution for personalized, dynamic, and seamless customer interactions.

VII. References

1) Liu, Z., Luo, P., Wang, L., & Tang, X. (2016). "DeepFashion: Powering Robust Clothes Recognition and Retrieval with Rich Annotations." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1096- 1104.

2) Cunha, J., & Rodrigues, P. (2019).

3) "Conversational Fashion Recommender Systems: A Survey and Future Directions." Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM), 1089-1098

4) Das, A., Kottur, S., Gupta, K., Singh, A., Yadav, D., Moura, J. M. F., Parikh, D., & Batra, D. (2017). "Visual Dialog".Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).Retrieved from visualdialog.org.

5) Sriram G., P., & Venkatesh S., P. (2021). "Image Classifying AI Chatbot". International Journal of Advanced Research in Innovative Ideas and Technology (IJARIIT), 7(5), V7I5-1378. ISSN: 2454-132X. Impact Factor: 6.078. Available at: https://www.ijariit.com.

6) Levy, M., Ben-Ari, R., Darshan, N., & Lischinski, D. (Year). "Chatting Makes Perfect: Chat-based Image Retrieval". The Hebrew University of Jerusalem & OriginAI, Israel. Available at https://ar5iv.labs.arxiv.org/html/2305.20 062

7) Zhang, Q., Lu, J., & Jin, Y. (2020)."Artificial intelligence in recommender systems". Published online on 1 November 2020.

8) Ayemowa, M. O., Ibrahim, R., & Khan, M. M. (2023). Analysis of Recommender System Using Generative Artificial Intelligence: A Systematic Literature Review. *IEEE Xplore*. Available at: https://ieeexplore.ieee.org/stamp/stamp.j sp?tp=&arnumber=10565860.

9) T. Liu, Q. Wu, L. Chang, and T. Gu, ''A review of deep learningbased recommender system in e-learning environments,'' Artif. Intell. Rev., vol. 55, no. 8, pp. 5953–5980, Dec. 2022, doi: 10.1007/s10462-022- 10135-2.