Review

Fashion Recommendation Systems, Models and Methods: A Review

Samit Chakraborty [1,2,*] , **Md. Saiful Hoque** [2,3] , **Naimur Rahman Jeem** [4], **Manik Chandra Biswas** [1] , **Deepayan Bardhan** [5] and **Edgar Lobaton** [5]

1   Wilson College of Textiles, North Carolina State University, Raleigh, NC 27695, USA; mbiswas2@ncsu.edu
2   Department of Textile Engineering, Daffodil International University, Dhaka 1207, Bangladesh; hoque1@ualberta.ca
3   Department of Human Ecology, University of Alberta, Edmonton, AB T6G 2R3, Canada
4   Department of Computing Science, University of Alberta, Edmonton, AB T6G 2R3, Canada; rahmanje@ualberta.ca
5   Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC 27695, USA; dbardha@ncsu.edu (D.B.); ejlobato@ncsu.edu (E.L.)
*   Correspondence: schakr22@ncsu.edu

Abstract: In recent years, the textile and fashion industries have witnessed an enormous amount of growth in fast fashion. On e-commerce platforms, where numerous choices are available, an efficient recommendation system is required to sort, order, and efficiently convey relevant product content or information to users. Image-based fashion recommendation systems (FRSs) have attracted a huge amount of attention from fast fashion retailers as they provide a personalized shopping experience to consumers. With the technological advancements, this branch of artificial intelligence exhibits a tremendous amount of potential in image processing, parsing, classification, and segmentation. Despite its huge potential, the number of academic articles on this topic is limited. The available studies do not provide a rigorous review of fashion recommendation systems and the corresponding filtering techniques. To the best of the authors' knowledge, this is the first scholarly article to review the state-of-the-art fashion recommendation systems and the corresponding filtering techniques. In addition, this review also explores various potential models that could be implemented to develop fashion recommendation systems in the future. This paper will help researchers, academics, and practitioners who are interested in machine learning, computer vision, and fashion retailing to understand the characteristics of the different fashion recommendation systems.

Keywords: fashion recommendation system; e-commerce; filtering techniques; algorithmic models; performance

1. Introduction

Clothing is a kind of symbol that represents people's internal perceptions through their outer appearance. It conveys information about their choices, faith, personality, profession, social status, and attitude towards life. Therefore, clothing is believed to be a nonverbal way of communicating and a major part of people's outer appearance [1]. Recent technological advancements have enabled consumers to track current fashion trends around the globe, which influence their choices [2,3]. The fashion choices of consumers depend on many factors, such as demographics, geographic location, individual preferences, interpersonal influences, age, gender, season, and culture [4–8]. Moreover, previous fashion recommendation research shows that fashion preferences vary not only from country to country but also from city to city [9]. The combination of fashion preferences and the abovementioned factors associated with clothing choices could transmit the image features for a better understanding of consumers' preferences [7]. Therefore, analyzing consumers' choices and recommendations is valuable to fashion designers and retailers [9–11]. Additionally, consumers' clothing choices and product preference data have become available on the Internet

in the form of text or opinions and images or pictures. Since these images contain information about people from all around the world, both online and offline fashion retailers are using these platforms to reach billions of users who are active on the Internet [10,12,13]. Therefore, e-commerce has become the predominant channel for shopping in recent years. The ability of recommendation systems to provide personalized recommendations and respond quickly to the consumer's choices has contributed significantly to the expansion of e-commerce sales [14].

According to different studies, e-commerce retailers, such as Amazon, eBay, and Shopstyle, and social networking sites, such as Pinterest, Snapchat, Instagram, Facebook, Chictopia, and Lookbook, are now regarded as the most popular media for fashion advice and recommendations [15–22]. Research on textual content, such as posts and comments [23], emotion and information diffusion [24], and images has attracted the attention of modern-day researchers, as it can help to predict fashion trends and facilitate the development of effective recommendation systems [5,25–27]. An effective recommendation system is a crucial tool for successfully conducting an e-commerce business. Fashion recommendation systems (FRSs) generally provide specific recommendations to the consumer based on their browsing and previous purchase history. Social-network-based FRSs consider the user's social circle, fashion product attributes, image parsing, fashion trends, and consistency in fashion styles as important factors since they impact upon the user's purchasing decisions [28–38]. FRSs have the ability to reduce transaction costs for consumers and increase revenue for retailers. With the exception of a single study from 2016 that focuses only on apparel recommendation systems [10], no current research presents recent advances in research on fashion recommendation systems. Therefore, the purpose of this paper is to present an integrative review of the research related to fashion recommendation systems. Moreover, Guan et al. cited research published until 2015. Therefore, the first objective of this paper is to review the most recent research published on this topic from 2010 to 2020. The previous study did not provide an in-depth analysis of the computational methods or algorithms corresponding to the fashion recommendation systems. This review study aims to fulfill this research gap and rigorously study the principles underlying, the methods used by, and the performance of the state-of-the-art fashion recommendation systems. To the best of our knowledge, this in-depth study is first of its kind. It includes research articles related to image parsing, clothing and body shape identification, and fashion attribute recognition, which are critical parts of fashion recommendation systems (FRSs). This review paper also provides a guideline for a research methodology to be used by future researchers in this field. The first section of this review discusses the history and background of FRSs. The second section presents a concise history and overview of recommendation systems. The third section aims to integrate the scholarly articles related to FRSs published in the last decade. The fourth section defines the metrics that are used by researchers to present and discuss recommendation results. The fifth section forms the major part of this review and focuses on various FRSs followed by different computational algorithmic models and recommendation filtering techniques used in fashion recommendation research. It will help researchers to understand these crucial parts of a FRS. The final section highlighted the existing challenges of using state-of-the-art recommendation systems followed by providing recommendations to overcome them and proposing a novel FRS based on the research findings discussed in section five. The study of the existing literature revealed that fashion recommendation systems have a huge impact on consumers' buying decisions. Hence, fashion retailers and researchers are exploring and developing state-of-the-art recommendation models to improve the accessibility, navigability and consumers' overall purchasing experience. One of the prime elements that has been continuously researched in these articles was the improvement of existing and the development of new algorithms relevant to the filtering techniques [4,15,33,39–51]. This review paper has identified state-of-the art algorithms and filtering techniques that have high potential to become more popular in the future. The sections of this paper are arranged in the order of the important FRS components, so that the reader can gain a substantial understanding of components such

as algorithmic models before moving to other important components such as filtering techniques. This review paper will guide future aspirants to conduct further in-depth and innovative empirical research on fashion recommendation systems. The organizational structure of this article is presented in Figure 1.
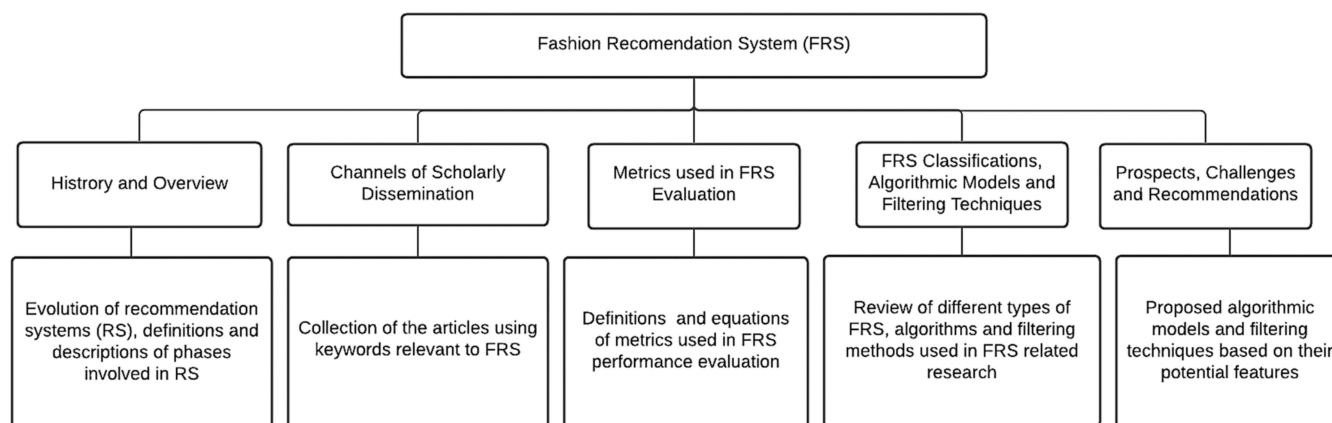
Figure 1. Organizational structure of the article.

2. History and Overview of Recommendation System

The era of recommendation systems originally started in the 1990s based on the widespread research progress in Collective Intelligence. During this period, recommendations were generally provided to consumers based on their rating structure [52]. The first consumer-focused recommendation system was developed and commercialized by Goldberg, Nichols, Oki and Terry in 1992. Tapestry, an electronic messaging system was developed to allow users only to rate messages as either a good or bad product and service [53]. However, now there are plenty of methods to obtain information about the consumer's liking for a product through the Internet. These data can be retrieved in the forms of voting, tagging, reviewing and the number of likes or dislikes the user provides. It may also include reviews written in blogs, videos uploaded on YouTube or messages about a product. Regardless of communication and presentation, medium preferences are expressed in the form of numerical values [52,54]. Table 1 presents the history of the progress of fashion recommendation systems over the last few decades.

Table 1. History of recommendation systems; produced by the authors based on [52,55,56].

Year	Recommendation System Approach	Properties
Before 1992	Mafia, developed in 1990	• Content filtering. • Mail filtering agent for providing a cognitive intelligence-based service for document processing.
1992 to 1998	Tapestry, developed in 1992	• Collaborative filtering. • Developed by Palo Alto. • Allowed users only to rate messages as either good or bad product and service.
	Grouplens, first used in 1994	• Rate data to form the recommendation.
	Movielens, proposed in 1997	• Useful to construct a well-known dataset.

Year	Recommendation System Approach	Properties
1999 to 2005	PLSA (Probabilistic Latent Semantic Analysis), proposed in 1999	• Developed by Thomas Hofmann. • Collaborative filtering.
2005 to 2009	Several Latent Factor Models such as Singular Value Decompositions (SVD), Robust Singular Value Decomposition (RSVD), Normalized Singular Value Deviation (NSVD).	• Collaborative filtering approach. • Find out factors from rating patterns.
2010 to onwards	Context-aware-based, instant-personalization-based	• Combined techniques of content and collaborative approach.

E-commerce retailers started implementing fashion recommendation systems in the early 2000s. However, implementation was mostly in the development stage until 2007–2008 [10,52,55,57–59]. As with other products such as electronics and books, fashion products were also recommended based on the user's previous purchase history. With the continuous progress in computer vision algorithms, personalized recommendations utilizing personal factors and user reviews have become more popular today [10,58,60].

2.1. Recommendation System

Recommendation system (RS) is referred to as a decision-making approach for users under a multidimensional information environment [61]. RS has also been defined as an e-commerce tool, which helps consumers search based on knowledge that is related to a consumer's choices and preferences [59]. RS also assists in augmenting social processes by using the recommendations of other users when there is no abundant personal information or knowledge of the alternatives [52]. RS handles the complication of information overload that consumers usually encounter by offering customized service, exclusive content, and personalized recommendations [57].

There are multiple phases involved in the recommendation system that develop the foundation of any state-of-the-art recommendation system. These are defined as the information collection phase, the learning phase, and the recommendation phase. Figure 2 shows the interrelationship of these phases involved in the recommendation process. It shows that information collection is the initial stage of RS, which is followed by the learning phase and the recommendation phase. The recommendation provided in the last phase can be generated based on information gathered during the information collection phase.

2.1.1. Information Collection Phase

In this phase a user's relevant information is collected to develop a user profile or model based on the user's characteristics, behaviors, and the content of the resources they have browsed, which are applicable in prediction phase tasks. The accurate functioning of a recommendation agent depends on the proper construction of a user profile or model. The system can offer a quick yet appropriate recommendation when it has all the required information about the user. Thus, the success of a recommendation or recommender system largely depends on the ability of the model to denote users' current preferences or choices [57,62,63].

The foundation of the recommendation system relies on three types of input such as explicit feedback, implicit feedback, and hybrid feedback. Explicit feedback needs to be of high quality as it encompasses users' explicit input regarding their interest in or choice of a product. The accuracy of the prediction or recommendation relies on user ratings. Therefore, if the users do not provide enough information, it limits the accuracy of the system. Despite this requirement, explicit feedback is still considered a

crucial information input process as it provides more reliable data and builds transparency into the recommendation procedure [57,64,65]. Implicit feedback is also important in understanding users' preferences, which are inferred indirectly through observation of user behavior. Although this method does not require the same effort from the users, it is often seen as less accurate [57,66]. Hybrid feedback is considered a combination of explicit and implicit feedback. It can be accomplished by utilizing the implicit feedback data as a check on the explicit feedback rating or by providing users with the opportunity to give feedback only if they choose to explicitly express their interest.
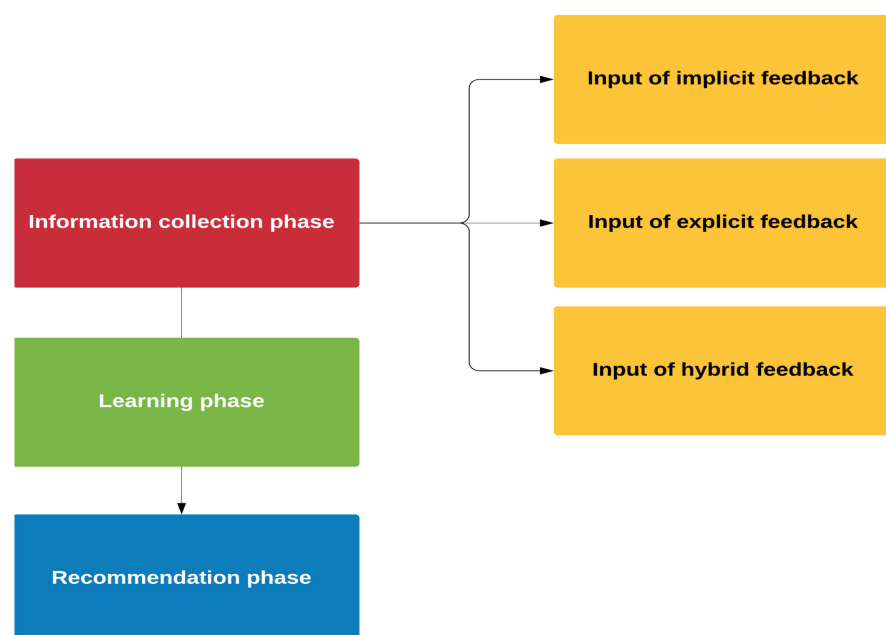
Figure 2. Phases of recommendation process.

2.1.2. Learning Phase

A learning algorithm is applied in this phase to filter and exploit the users' features based on the feedback collected in the information collection phase. The learning algorithms used in this phase are helpful for drawing out the appropriate patterns relevant for application during the recommendation stage [57,62,63].

2.1.3. Recommendation Phase

The recommendation phase recommends the types of items that a user or consumer may prefer. Recommendations can be provided either directly based on the dataset collected during the information collection phase (which might be memory- or model-based) or through the browsing history of users observed by the system [57,62,63]. Recommendations can also be provided by combining the learned information with the rating matrix to recommend learning resources [67]. Researchers reported improved recommendation accuracy using hybrid models in comparison with product content-based or other user-preference-based collaborative models [68].

3. Channels of Scholarly Dissemination Related to Fashion Recommendation System (FRS)

Articles published from January 2010 to June 2020 have been considered for the review purpose of this article. Various online literature resources or databases such as Scopus, Web of Science, Science Direct, and Design and Applied Arts Index (DAAI) have been used to find the literature. Boolean operator techniques i.e., "AND" or "OR" strategies were used to search articles from these sources. Keywords grouped in three categories as listed below were used to conduct the final search.

Group 1: Fashion OR Style OR Apparel OR Clothing.
Group 2: Recommend*.
Group 3: Filtering Technique OR Algorithm OR Model OR Artificial Intelligence OR Neural Network OR Deep Learning OR Meta-Learning OR Fuzzy Techniques OR Model OR Image Processing OR Image Retrieval OR Image Feature extraction.
Final Search = Group 1 AND Group 2, Group 1 AND Group 2 AND Group 3.

Overall, 230 scholarly articles and 9 web sources have been reviewed. Among these, 214 scholarly articles were found containing the required keywords when using the search strategy mentioned above. Among these, 132 articles are indexed in Scopus, 26 in Web of Science, 3 in Science Direct and 1 in the Design and Applied Arts Index (DAAI) database. In addition, 50 articles and 2 patents were found in Google Scholar, published in different peer-reviewed journals and conferences.

4. Metrics Used in Fashion Recommendation System Evaluation

The performance of a recommendation algorithm is evaluated by using some specific metrics that indicate the accuracy of the system. The type of metric used depends on the type of filtering technique. Root Mean Square Error (RMSE), Receiver Operating Characteristics (ROC), Area Under Cover (AUC), Precision, Recall and F1 score is generally used to evaluate the performance or accuracy of the recommendation algorithms.

Root-mean square error (RMSE). RMSE is widely used in evaluating and comparing the performance of a recommendation system model compared to other models. A lower RMSE value indicates higher performance by the recommendation model. RMSE, as mentioned by [69], can be as represented as follows:

$$RMSE = \sqrt{\frac{1}{N_p} \sum_{u,i} (p_{ui} - r_{ui})^2} \tag{1}$$

where, N_p is the total number of predictions, p_{ui} is the predicted rating that a user u will select an item i and r_{ui} is the real rating.

Precision. Precision can be defined as the fraction of correct recommendations or predictions (known as True Positive) to the total number of recommendations provided, which can be as represented as follows:

$$Precision = \frac{True\ Positive\ (TP)}{True\ Positive(TP) + False\ Positive\ (FP)} \tag{2}$$

It is also defined as the ratio of the number of relevant recommended items to the number of recommended items expressed as percentages.

Recall. Recall can be defined as the fraction of correct recommendations or predictions (known as True Positive) to the total number of correct relevant recommendations provided, which can be as represented as follows:

$$Recall = \frac{True\ Positive\ (TP)}{True\ Positive(TP) + False\ Negative\ (FN)} \tag{3}$$

It is also defined as the ratio of the number of relevant recommended items to the total number of relevant items expressed as percentages.

F1 Score. F1 score is an indicator of the accuracy of the model and ranges from 0 to 1, where a value close to 1 represents higher recommendation or prediction accuracy. It represents precision and recall as a single metric and can be as represented as follows:

$$F1\ score = 2 \times \frac{Precision * Recall}{Precision + Recall} \tag{4}$$

Coverage. Coverage is used to measure the percentage of items which are recommended by the algorithm among all of the items.

Accuracy. Accuracy can be defined as the ratio of the number of total correct recommendations to the total recommendations provided, which can be as represented as follows:

$$Accuracy = \frac{TP + FN}{TP + FN + TN + FP} \tag{5}$$

Intersection over union (IoU). It represents the accuracy of an object detector used on a specific dataset [70].

$$IoU = \frac{TP}{TP + FN + FP} \tag{6}$$

ROC. ROC curve is used to conduct a comprehensive assessment of the algorithm's performance [57].

AUC. AUC measures the performance of recommendation and its baselines as well as the quality of the ranking based on pairwise comparisons [5].

Rank aware top-N metrics. The rank aware top-N recommendation metric finds some of the interesting and unknown items that are presumed to be most attractive to a user [71]. Mean reciprocal rank (MRR), mean average precision (MAP) and normalized discounted cumulative gain (NDCG) are three most popular rank aware metrics.

MRR. MRR is calculated as a mean of the reciprocal of the position or rank of first relevant recommendation [72,73]. MRR as mentioned by [72,73] can be expressed as follows:

$$MRR = \frac{1}{N_u} \sum_{u \in N_u} \frac{1}{L_u^n[k] \in R_u} \tag{7}$$

where u, N_u and R_u indicate specific user, total number of users and the set of items rated by the user, respectively. L indicates list of ranking length (n) for user (u) and k represents the position of the item found in the he lists L.

MAP: MAP is calculated by determining the mean of average precision at the points where relevant products or items are found. MAP as mentioned by [73] can be expressed as follows.

$$MAP = \frac{1}{N_u|R_u|} \sum_{k=1}^{n} 1(L_u^n[k] \in R_u) P_u@k \tag{8}$$

where P_u represents precision in selecting relevant item for the user.

NDCG: NDCG is calculated by determining the graded relevance and positional information of the recommended items, which can be expressed as follows [73].

$$NDCG_u = \frac{\sum_{k=1}^{n} G(u, n, k) D(k)}{\sum_{k=1}^{n} G^*(u, n, k) D(k)} \tag{9}$$

where $D(k)$ is a discounting function, $G(u, n, k)$ is the gain obtained recommending an item found at k-th position from the list L and $G^*(u, n, k)$ is the gain related to k-th item in the ideal ranking of n size for u user.

5. Fashion Recommendation System (FRS), Algorithmic Models and Filtering Techniques

FRS can be defined as a means of feature matching between fashion products and users or consumers under specific matching criteria. Different research addressed apparel attributes such as the formulation of colors, clothing shapes, outfit or styles, patterns or prints and fabric structures or textures [10,58,74,75]. Guan et al. studied these features using image recognition, product attribute extraction and feature encoding. Researchers have also considered user features such as facial features, body shapes, personal choice or preference, locations and wearing occasions in predicting users' fashion interests [31,75–78]. A well-defined user profile can differentiate a more personalized or customized recommendation system from a conventional system [28,79]. Various research projects on apparel recommendation systems with personalized styling guideline and intelligent recommendation engines have been conducted based on similarity recommendation and expert advisor

recommendation systems [10,58,61]. Image processing, image parsing, sensory engineering, computational algorithms, and computer vision techniques have been extensively employed to support these systems [32,80–84].

5.1. Classification of Fashion Recommendation System (FRS)

Fashion recommendation systems (FRS) proposed by researchers vary from each other based on the filtering techniques used, information collection and learning procedures, feature extraction methods and types of recommendations provided to users or consumers. The paper has categorized the recommendation systems into five classes such as fashion image retrieval, a personal wardrobe recommendation system, a knowledge-based recommendation system, smart or intelligent recommendation systems and a social-network-based recommendation system based on previous research and academic articles. These recommendation systems or approaches have been discussed briefly in Table 2

A fashion image retrieval system is formulated based on clusters of fashion products and their feature similarity as well as correlation analysis based on individual historical data [85,86]. Personal wardrobe recommendation systems explore similar fashion styles based on wardrobe usage history [10,87]. Fashion pairing recommendation systems, also referred to as fashion coordination systems, are based on the rules of matching different types of clothing items with styling knowledge [4,10,51]. A smart or intelligent recommendation approach uses features or attributes of the clothing and user in terms of users' body shapes, contextual information of wear, outfit type and genre characteristics [88–90]. A social-network-based recommendation approach offers recommendations to many social-media-based information discovery and social collaborations among potential collaborators using social networking features. Sachdeva and Pandey (2020) focused on the analysis of patterns for different consumer groups with finely grained fashion elements using a large-scale fashion trend dataset (FIT) compiled from Instagram reports. The usage details were provided to the Knowledge Enhanced Recurrent Network model (KERN), which takes advantage of the capacity of deep recurrent neural networks to model time series of fashion elements, considering very complex patterns effectively. It can reinforce the prediction of styles. These recommendation systems or approaches have been discussed briefly in Table 2.

Table 2. Classification of fashion recommendation systems (FRS).

Recommendation System	References	Features and Implementation
Fashion image retrieval	[7,10,11,25,34,85, 86,91–99]	• Offers recommendations based on previous sales, clothes purchase records, eye movement records and item click rate. • Provides clothing suggestions using analytical hierarchy procedure of use's choice and interest. • CNN can be used for feature extraction and image classification in conjunction with RNN, which helps in the retrieval of similar image products.
Personal wardrobe recommendation	[10,31,88,90,93, 100–108]	• Offers clothing recommendations by matching wardrobe management profile with explicit input of time, location, weather conditions and typical schedule provided by user. • Smart closet system can suggest appropriate fashion items estimating the information related to weather and events. • Bayesian network can be employed to offer personalized fashion recommendation system developed based on the history of wardrobe items usage.

Recommendation System	References	Features and Implementation
Fashion pairing recommendation system	[4,10,15,22,36,45–49,109–124]	- Adoption of this system helps in the representation of different style genres based on the knowledge of fashion coordination and image recognition. - Implementation of this approach combines both visual and textual information to express a knowledge-based fashion coordination system and use image detection technology for extracting fashion styles with similar features. - It can recommend design scheme via a searching method using genetic algorithms (GA) and artificial neural networks.
Smart or intelligent recommendation	[33,39–44,50,74,88,89,112,123,125–137]	- Its domain expertise knowledge of mixing and matching criteria facilitates exploring the interrelationship between the fashion and the user using intelligent algorithms. - Use of decision tree, analytical hierarchy process, sensory engineering, fuzzy mathematics, genetic algorithms, neural networks and support vector machines to learn the skill of clothing attribute evaluation. - Implementation of expert rules to propose an intelligent fashion recommendation system of expert information collection based on eye gaze tracking and the application of interactive evolutionary algorithms to predict users' style preferences.
Social-network-based recommendation	[7,8,31,43,92,133,138–152]	- Personalized clothing recommendation built using three types of data: (1) user social circles that show the relationships among users; (2) user clothing records that indicate the interest and preferences of users for certain fashion items; and (3) matching of pairs of fashion items that represent style consistency among different items. - Combination with wardrobe recommendations provides more information about users to retailers, which can create an interactive online shopping experience. - Peer recommendations functioning through social shopping sites can increase the accuracy of predictions based on the sharing of lifestyles or experiences with friends, family members and colleagues, who understand the users.

5.2. Algorithmic Models Used in Fashion Recommendation Systems

The models most used in developing fashion recommendation systems are multilayer perceptron (MLP), recurrent neural network (RNN), k-nearest neighbor (kNN), convolutional neural networks (CNN), Bayesian networks, generative adversarial network (GAN) and autoencoder (AE) [8,12,31,51,86,103,153–158]. Researchers modified the algorithms and tuned the hyperparameters to different extents to increase the prediction accuracy. The rest of this section provides an overview of the main methodologies used.

5.2.1. Convolutional Neural Network (CNN)

A convolutional neural network (CNN) is constructed of multiple convolutional layers, where the number of layers is customized based on the desired recommendation system outcome. These layers can vary in terms of convolutional layers, filter size and fully connected layers. Researchers increase or decrease the depth of the network to achieve better results with the highest accuracy. Kernel and batch sizes are fixed depending on the desired input/output of the layer. There is an optional pooling layer to reduce the dimensionality of the data. The most common form of pooling layer is max pooling,

which often ranges between 2×2 and 4×4. Softmax, Sigmoid, ReLU and TanH are the most common activation functions for CNN, which can be used either separately or in stacked form. Adam and stochastic gradient (SGD) are two popular optimizers used in tuning hyperparameters of CNN models. CNN is very popular in recommendation systems for its strong feature extraction and image classification capabilities. Yu et al proposed a combined matrix and tensor factorization model using CNNs for an aesthetic-based clothing recommendation to learn the images and their aesthetic features [159]. Nguyen et al. utilized the convolutional and max-pooling layer to obtain visual features from different patches of images [120].

5.2.2. Recurrent Neural Network (RNN)

Recurrent neural network (RNN) is a generalization of feed forward neural network that has an internal memory. RNN can use the internal state (memory) to process sequences of inputs. There can be one to many input vectors as well as output nodes depending on the type of research and goal, where these are not co-dependent. The dimension of the input vectors can be of any size. The hidden states vary from the number of input vectors to the number of states for the next cell. The most common activation function used in RNN is ReLU. Long Short Term memory (LSTM) networks are a modified version of RNNs, which can remember past data in memory more effectively. The vanishing gradient problem of RNN is resolved in LSTM and so it is highly used to classify, process, and predict time series data. Wu et al. designed a session-based recommendation model for a fashion e-commerce website by utilizing the basic recurrent neural network (RNN) to predict what a user will buy next based on the click history [154]. Quadrana et al. proposed a similar hierarchical RNN for session-based recommendation, which deals with both session-aware recommendations when user identifiers are present [160]. However, to inject context information into input and output layers, Smirnova and Vasil proposed a context-aware and session-based recommendation system based on conditional RNNs [161]. Han et al. studied two distinct fashion recommendation techniques in their research, where they trained a bidirectional LSTM model to predict an item conditioned on previous items to facilitate the learning of compatibility relationships among the items [162]. They used Polyvore dataset to perform extensive experiments to evaluate the performance of their proposed model compared to different state-of-the-art recommendation models.

Li et al. presented an attention-based LSTM (Long Short-term Memory) model for hash tag recommendation, which can capture the sequential property and also recognize the informative words from micro blog posts by taking advantage of both RNNs and attention mechanism [92]. The predicted rating ($\hat{r}$) of item j given by user i at time t is defined as:

$$\hat{r}_{ui|t} = f(u_{ut}, v_{it}, u_u, v_i) \tag{10}$$

Here, in Equation (7), u_u and v_i are the stationary latent attributes of user and item. Besides, u_{ut}, and v_{it} are learned from LSTM, u_u and v_i are learned by standard matrix factorization.

5.2.3. Multilayer Perceptron (MLP)

A multilayer perceptron (MLP) is a form of artificial neural network architecture that contains a series of layers, which are composed of neurons and their connections. The neurons have the ability to calculate the weighted sum of its inputs and apply an activation function to obtain a signal, which is transmitted to the next neuron. The number of layers can be from 2 to infinity as more and more deep neural networks are introduced depending on the project goal. The batch value and neurons vary but not limited from 8 to 64 and 128 to 512, respectively. Usually, the input and output layers have linear activation function, and the hidden layers commonly have Sigmoid, ReLU, Softmax, tanh, etc. The common optimizers used in MLP are Adam, Adadelta, Adagrad, Adamax, Nadam, SGD, RMSprop, etc.

A multilayer perceptron (MLP) model has been applied io different recommendation systems including for fashion clothing [163] and makeup recommendations [164]. Cheng et al. proposed an advanced model of MLP, which can solve both regression and classification problems [165]. Here, the wide learning component is a single layer perceptron, and the deep learning component is a multilayer perceptron. Combining these two (wide learning component: single layer perceptron and the deep learning component: multilayer perceptron) learning techniques enabled the model to capture both memorization and generalization, which helps the model to catch direct features from historic data as well as generalizing it, which can improve both the accuracy and diversity of the recommendation. Chen et al. extended the model to decrease the running time for large-scale industrial-level recommendation tasks by replacing the deep learning component with an efficient locally connected network [15].

5.2.4. Generative Adversarial Network (GAN)

Generative adversarial networks (GAN) are deep-learning-based generative models in which two neural networks (generator and discriminator) compete to become more accurate in their predictions. The goal of the generator is to artificially manufacture (fake) outputs that could easily be mistaken for real data and the discriminator tries to identify which output is not real.

Among the two major components of GAN: the generator network is a simple feed-forward neural network (i.e., five layers) and the discriminator network is a classifier, which is slightly different from the generator network. The discriminator network processes the image and gives a probability of a class for that image. Calculation of a GAN's accuracy is performed using a scoring algorithm such as using pre-trained inception V3 network by comparing extracted features of both the generated and real image.

The objective functions of the GAN's discriminator and generator are defined as:

$$DminV(D) = E_{x,c \sim p_{data}(x,c)}L_{real}(x,c) + E_{c \sim p(c),z \sim p(z)}L_{fake}(G(z,c),c), \ GminV(G)DminV(D) = E_{c \sim p(c),z \sim p(z)}L_{fake} \quad (11)$$

Here, in Equation (11), $L_{real}(x,c) = [D(x,c) - 1]^2$ and $L_{fake}(x,c) = [D(x,c)]^2$, which means the discriminator D tries to predict '1' for the real image and '0' for fake images and the generator G tries to generate 'realistic' images to fool discriminator D until the quality of the generated images is acceptable.

Kang et al. proposed a generative adversarial network: an unsupervised learning framework-based recommendation system that can generate new clothing in order to provide more accurate recommendations to users [5]. They used the Bayesian personalized ranking (BPR) optimization framework for implicit feedback, which optimizes rankings by considering triplets $(u,i,j) \ \epsilon \ D$.

$$D = \left\{ (u,i,j) \mid u \epsilon U \ \wedge \ i \epsilon I_u^+ \ \wedge \ j \epsilon I \backslash I_u^+ \right\} \quad (12)$$

Here, in Equation (12), $i \epsilon I_u^+$ is an item about which the user u has expressed interest, and $j \epsilon I \backslash I_u^+$ is the one about which they have not.

5.2.5. k-Nearest Neighbor (kNN)

The k-nearest neighbor (kNN) algorithm is a simple supervised learning algorithm which can be used to solve both classification and regression problems. It depends on labeled input data to produce output when given new unlabeled data. It is a non-parametric algorithm, so it does not make any assumptions on any underlying data distribution and does not use the training data points to perform any generalization. The output of the algorithm is based on the feature similarity.

In the KNN algorithm, the *k* most similar items are obtained by using different similarity measures such as Cosine, Euclidean, etc. The formula can be derived using simple Euclidean Distance as:

$$d(p,q) = d(q,p) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \ldots + (q_n - p_n)^2} = \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2} \quad (13)$$

Here, in Equation (13), *n* is the number of dimensions or features. The data point located at the minimum distance from the test point is assumed to belong to the same class.

Viriato De Melo et al. proposed a content-based approach for clothing recommendation by combining textual attributes, visual features, and human visual attention in order to compose the clothes' profile, which outperformed their baseline approaches [158]. The kNN algorithm is used for the item rating considering the past behavior of the user and the similarity between items.

5.2.6. Autoencoder (AE)

An autoencoder is an unsupervised learning technique for neural networks that can learn efficient data representations or encoding by training the network to ignore signal noise. It consists of an encoder, code, and decoder where the encoder and decoder are both fully connected to, and feed forward neural networks and the code is a single layer of an artificial neural network user for reducing dimensionality. Here, the input passes through the encoder to produce the code and then output is produced at the decoder using only the code as the input where the dimensionality of the input and output needs to be the same. The number of layers of code determines the rate of compression, the lower the number, the higher the compression ratio. An autoencoder can be as deep as we want and can be any number between 2 and infinity depending on the research goal. Besides, the most common loss functions used in these kinds of models are mean squared error (MSE) and binary cross-entropy.

The two general reasons for using an autoencoder in a recommendation system are either to learn about lower-dimensional feature representations or to fill the blanks in the interaction matrix. Collaborative denoising autoencoder (CDAE) is one of the most highly used autoencoders, which is principally used for ranking prediction. The reconstruction is defined as:

$$h\left(\hat{r}_{pref}^{(u)}\right) = f\left(W_2 \cdot g\left(W_1 \cdot \hat{r}_{pref}^{(u)} + V_u + b_1\right) + b_2\right) \quad (14)$$

Here, in Equation (14), $\hat{r}_{pref}^{(u)}$ is user observed implicit feedback; $V_u \in R^k$ denotes the weight matrix for the user node. Gao et al. developed a F clothes matching scheme based on Siamese network and autoencoder [12]. They used triple autoencoder and Bayesian personalized ranking to map three kinds of features into the same latent space to learn the compatibility between tops and bottoms.

5.2.7. Bayesian Networks

Bayesian networks are a kind of probabilistic graphical model, which are usually used for prediction, anomaly detection, diagnostics, time series prediction, reasoning, and decision-making. This kind of network is a graph, which is made of nodes and directed links. Each node represents a variable, where the number of variables in a node can be one to any number, and in that case, they are called multivariable nodes. The variables can be both discrete and continuous. The required links are automatically determined from data using the structural learning algorithm.

Bayesian networks can be used to model the joint probability distribution of multiple random variables where a random variable is represented as a node and the links represent dependencies between the variables. Ono et al. constructed a user recommendation model using Bayesian networks considering the users' contexts in addition to their personalities [156]. They selected effective variables from many observed attributes and

determined local network structures and estimated conditional probability tables. For this, the pseud product attributes were calculated from impression attributes whose scores were defined as:

$$score = \frac{I(r, \; CID)}{H(r|CID)} \tag{15}$$

Here, in Equation (15), $I(r, \; CID)$ is the mutual information between the attribute value r and the content ID, $H(r|CID)$ is the conditional entropy of r.

Researchers also used the naïve Bayes algorithm to develop recommendation system [166,167]. It is a classification algorithm, which uses Bayes' theorem for classification. This algorithm performs equally well with the problems above irrespective of their linear or non-linear separation [166]. Wei, et. al. (2020) used a naïve Bayes classifier to predict and investigate users' emotions followed by the determination of users' sentiments toward specific items and calculating the product-to-product similarity based on collaborative filtering [167].

Table 3 presents the machine-learning algorithms that are most used in fashion recommendation system research. It exemplifies the research that used these algorithmic models to develop recommendation systems and highlights the performance of these models for the benefit of the researchers and retailers.

Table 3. Popular algorithmic models used in fashion recommendation systems.

Algorithm/Model	Recommendation System Used	Performance
Convolutional Neural Networks (CNN)	• Guan, et al. and Liu, et al. used CNN to develop content-based filtering technique [10,168]. • The recommendation system showed weather-oriented clothing pairing results as output based on the image attributes.	The proposed CNN model achieved a maximum of Normalized Discounted Cumulative Gain (NDCG) ranking score of 0.50, which outperformed support vector machine (SVM), because SVM achieved an NDCG score of 0.45.
Recurrent Neural Network (RNN)	• Heinz et al. used RNN to build a recommendation system utilizing dynamic collaborative filtering technique [50]. • The RNN-based recommendation system recognized individual style preferences from a modest number of purchases by combining sales events.	The proposed RNN model achieved a higher AUC value of 88.5% compared to the AUC value of 80.2% achieved by a popularity ranking baseline approach.
Multilayer Perceptron (MLP)	• Alashkar et al. used MLP to build a fully automatic makeup recommendation system utilizing content-based filtering techniques [164]. • The model recommended homogeneous makeup style according to its automatically classified facial traits and synthesized the makeup style as well.	The proposed MLP model achieved a minimal squared loss function value, which was 48% lower than distance-based similarity recommendation model.

Table 3. *Cont.*

Algorithm/Model	Recommendation System Used	Performance
Generative adversarial network (GAN)	• Kang et al. used GAN to develop a personalized recommendation system utilizing collaborative filtering techniques [5]. • The models learnt the distribution of fashion images and generated novel fashion items, which maximized users' preferences.	The proposed method outperformed the strongest content unaware method (Bayesian Personalized Ranking) substantially by around 5.13% in terms of accuracy and achieved a 6.8% improvement over a retrieval-based method in terms of preference score.
kNN (k-nearest neighbor)	• Leininger et al. proposed an advanced retail recommendation system using kNN and collaborative filtering techniques [169]. • The model computed the distance to similar items by using cosine similarity followed by individual clustering of the products.	The model achieved a higher accuracy in terms of AUC (91%) than that of the AUC (85%) of the baseline model.
Autoencoder	• Gao et al. developed clothes matching scheme based on Siamese network and autoencoder utilizing content-based filtering techniques [12]. • The model extracted visual and textual features followed by recommending clothing based on the input image (top/bottom part clothing) of the user.	The proposed model achieved an AUC value 0.884 compared to the AUC value of 0.762 achieved by the probabilistic knowledge distillation (PKD) method.
Bayesian Networks	• Yu-Chu et al. used a Bayesian network to develop a personalized clothing-recommendation system utilizing content-based filtering techniques [103]. • The model recommended a clothing combination or selected personal items based on personal preferences rather than other users' behavior.	The proposed model outperformed the basic Bayesian model by 50% in terms of frequency of selection (of the same cloth) and by 90% in terms of recommended combinations

5.2.8. Other Methodologies

The study of algorithmic models revealed that researchers achieved better recommendation accuracy when combining multiple algorithms and techniques together rather than using a single algorithm-based baseline model. Apart from the abovementioned algorithmic models, researchers have also proposed various other potential algorithms, such as the cosine similarity-based algorithm [51,169] and fuzzy logic techniques [33,132,170] and multicriteria-based methods [77,137,171–173] have gained popularity in recent times.

Hu et al. proposed three cosine-based similarity algorithms for aesthetic-based collaborative clothing recommendation systems to provide recommendations to users using frequency similarity [51]. These were: cosine-based similarity (COS); adjusted cosine-based similarity (ACOS); and COS-based inverse user frequency (IUF).

$$S^{COS}(i,j) = \frac{|N(i) \cap N(j)|}{\sqrt{|N(i)| \cdot |N(j)|}} \tag{16}$$

$$S^{ACOS}(i,j) = \frac{\sum_{u \in U}(C(u,i) - \overline{C}(i)) \cdot (C(u,j) - \overline{C}(j))}{\sqrt{\sum_{u \in U}(C(u,i) - \overline{C}(i))} \cdot \sqrt{\sum_{u \in U}(C(u,j) - \overline{C}(j))}} \tag{17}$$

$$S^{IUF}(i,j) = \frac{|N(i) \cap N(j)| \cdot \sum_{u \in N(i) \cap N(j)} \frac{1}{\log(1 + |\Theta(u)|^\lambda)}}{\sqrt{|N(i)| \cdot |N(j)|}} \tag{18}$$

Here, in Equations (16)–(18), $N(i)$ and $N(j)$ are two sets of consumers who purchased goods i and j individually. $|N|$ is the size of the set, $C(u,i)$ is the co-occurrence matrix, $\overline{C}(j)$ is used to represent the average times consumers purchase goods j. Further, $\frac{1}{loglog(1 + |\Theta(u)|^\lambda)}$ is used to control the impact of the active consumers.

Researchers also used fuzzy theory and the analytic hierarchy process (AHP) to develop a consumer-oriented fashion recommendation system to facilitate using online shopping experience as a virtual sales advisor [33,132,170]. The initial stage of fuzzy logic application establishes the fuzzy membership functions of fuzzy sets. These membership functions are important for reflecting the features of fuzzy concepts as well as attaining mathematical tasks and processing [170]. Researchers have also used fuzzy comprehensive evaluation to assess objects influenced by multifactor criteria [174].

Researchers have frequently used multicriteria decision-based models to develop fashion recommendation systems [77,100,130–132,137,171–173,175]. Adewumi et al. used unified modeling language (UML) to develop a unified framework for outfit recommendations, where users specified the date as an input parameter through a weather API that was afterwards passed through an inference engine, a multicriteria decision making module and hybrid artificial intelligence techniques to provide a suitable recommendation [171]. Zeng et al. recommended a perception-based recommender system comprised of two distinct models, where these two models worked together to provide recommendations by characterizing human body measurements and their perceptions of different body shapes [135].
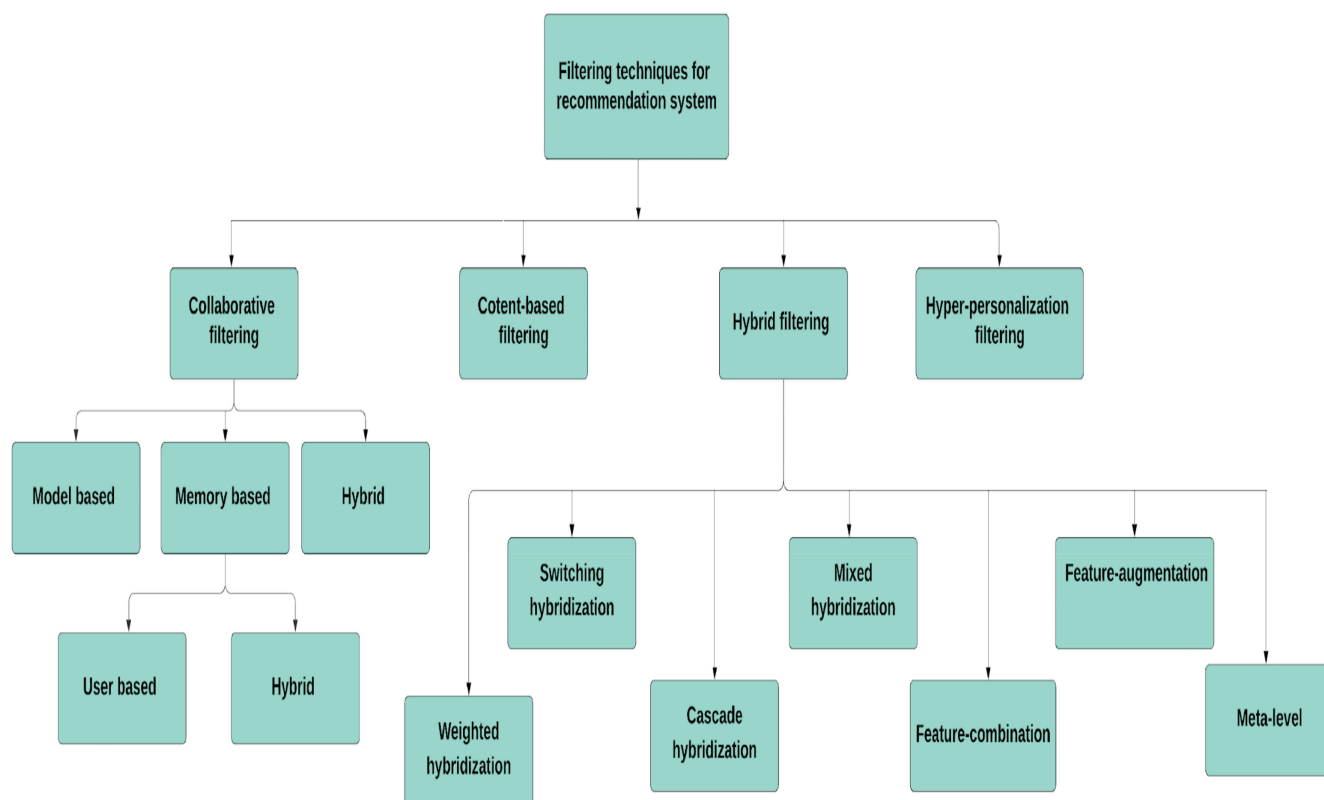
5.3. Recommendation Filtering Techniques

The selection of an effective and accurate filtering technique is crucial for developing a successful recommendation system. Therefore, an elaborate understanding of these techniques is required before implementing them in a commercial platform. Figure 3 presents a classification tree containing four widely used recommendation-filtering techniques.

5.3.1. Content-Based Filtering (CBF) Technique

The content-based filtering (CBF) technique examines the features of a recommended item by classifying users' (or consumers') and products' profile data based on the products' features [10]. The use of domain-dependent algorithms emphasizes the analysis of the products' features, which are utilized to generate predictions. Although the applications of content-based filtering techniques have been more successful in recommending web pages, publications and news articles, researchers have implemented this technique to develop fashion recommendation system as well [93,111,119,123,143,176–178]. In this technique, user profiles are matched with the features extracted from the product content, which provides the recommendation where the user has evaluated a specific product in the past [56]. The products that have the highest relation with the positively scored or rated items are generally recommended to users. The content-based technique uses different kinds of models to explore the similarity between items to generate a meaningful recommendation, which is the main distinctive feature between content-based and collaborative filtering techniques [179,180]. These machine-learning techniques propose recommendations by learning the core or foundations of the underlying model. In this type of filtering the

rating of an item is calculated based on the other ratings. Figure 4 shows a bipartite graph generated from the user interaction, where the orange shirt will receive a rating for user 1 because the other 2 reviews about that shirt both gave it 5 stars. This method of filtering is used when the target user is not known and much about the apparel to be sold is known. Here, the directed edges from the users to the items represent users' interaction with the items through likes, comments, retweets, etc.

Figure 3. Classification of filtering techniques for recommendation systems.
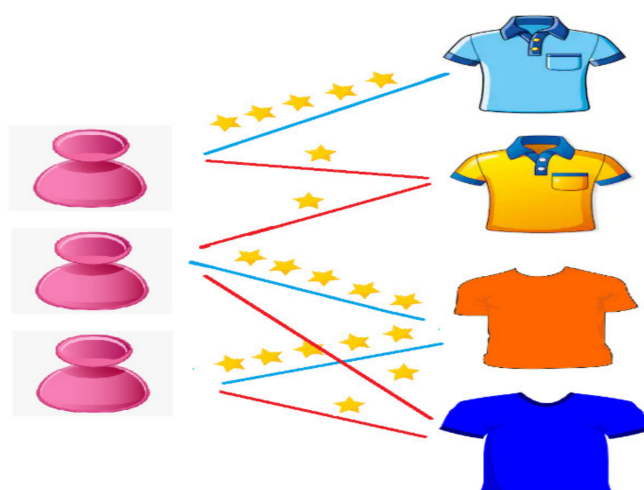
Figure 4. Content-based filtering process.

Researchers have used probabilistic models such as the Bayesian classifier [176], decision tree [181] and neural network model [123] to develop content-based recommendation systems. CBF does not require profiles of other users as it can adjust its recommendations

within a short period even if the user's profile changes. Viriato de developed a recommendation system based on a combination of textual features, visual attributes and visual attention using a content-based filtering technique [158]. Their proposed model, named CRESA (Clothing Recommendation System developed using Attributes such as textual attributes, visual features, and visual attention), outperformed standard models such as the k-nearest neighbor (kNN) model. It achieved an average precision of 74.8%, which was better than the other standard models. Wu et al. also adopted a similar approach of providing fashion recommendations based on the visual and textual information provided by the users [153].

5.3.2. Collaborative Filtering (CF) Technique

The collaborative filtering (CF) algorithm is one of the most successful techniques among all of the filtering techniques available for the recommendation system [182]. CF is a domain-independent prediction technique for analyzing hard-to-describe content by observing metadata [97,159,183]. This filtering technique is formed by using a dataset of the preferences of a group of users to make a recommendation to another group of users who show similar types of behavior. The fundamental assumption of CF is based on the similarities of users, which build a neighborhood group. Therefore, this technique is called user-based collaborative filtering [159,179,184,185]. In collaborative filtering, automatic predictions are made based on the reviews given by other people. Therefore, the major assumption is that if two people have similar interests in a common dataset then their interests would be similar for the rest as well [159,185]. Figure 5 represents an interaction matrix, where each row represents users, and each column represents product or item. This utility matrix contains partial data, where likeliness or interaction needs to be predicted based on the rating (i.e., 1 to 5) given by other users to a specific item. In this figure, rating 5 represents the highest interaction and rating 1 represents the lowest interaction. When a new item or user is added to the platform, the cosine similarity algorithm or k-nearest neighbor can be used while calculating the predicted value [186]. The figure also shows that as user 2 and user 3 have similar choices for item 3 and item 4, the CF method would predict that for item 2 (the yellow shirt) user 3 will have a similar interaction as user 2, i.e., 1 star. Additionally, that is how the rest of the selections would be filled in.

User 1	5	1	?	?
User 2	?	1	5	1
User 3	?	?	5	1

Figure 5. Collaborative filtering process.

Although the CF technique is critical and has some issues, such as data sparseness and the cold-start problem, recommendation systems based on CF techniques have successfully worked for many renowned business stores and services [179,184,187,188]. Yu et al. proposed a collaborative clothing recommendation system that overcomes the problem of capturing the aesthetic preferences of users by using a novel tensor factorization model [159]. They used the Amazon dataset and the Aesthetic Visual Analysis (AVA) dataset to train the recommendation models and the aesthetic network, respectively. The Amazon dataset

contains records of 39,371 users and 23,022 items. The AVA dataset contains over 250,000 images with aesthetic ratings from 1 to 10 and 14 photographic styles representing complementary colors, duotones, light on white, long exposure, high dynamic range, motion blur, negative image, silhouettes, soft focus, vanishing point and image grain. They proposed a dynamic collaborative filtering model using both aesthetic features and CNN features (DCFA) and compared it with baseline models such as the matrix factorization (MF) method, state-of-the-art visual-based recommendation method (VBPR) and state-of-the-art context-aware recommendation method (CMTF). DCFA and VBPR performed better on the test dataset compared to other models. However, the proposed DCFA model outperforms VBPR by 8.53% in terms of higher recall and 8.73% in terms of higher normalized discounted cumulative gain. Song et al. developed a personalized compatibility-based recommendation model (GP-BPR) using collaborative filtering [189]. The model is comprised of two key modeling elements: general compatibility and personal preference, which illustrate the interaction between items as well as the interaction between user and item. Their proposed personal preference modeling technique can facilitate delivering vital clues regarding user's personal preference. They also developed a large-scale dataset, named IQON3000, using the images available in the online fashion community IQON for the performance evaluation of the recommendation model. De Divitiis et al. also adopted a similar approach to propose a garment recommender system by combining memory augmented neural network (MANN) and matrix factorization (MF) techniques [190]. They considered personalized suggestions as an additional element to user preferences and purchase histories. They also used IQON3000 for their experiment and reported better performance compared to GP-BPR. The MANN+MF and GP-BPR obtain mean average precision of 0.15 and 0.13, respectively, while retrieving the same number of items (~20 items). Additionally, Sagar et al. introduced PAI-BPR (Personalized Attributewise Interpretable—BPR) as an outfit compatibility model that can capture user–item interaction along with general item–item interaction based on the user's personal preferences and identifying the discordant and harmonious attributes between fashion items [191]. They used multilayer perceptron (MLP) to learn the non-linear interactions and leverage both the textual and visual modalities in the context of item description and image, respectively. They also used matrix factorization to incorporate the latent content-based preference factors for the personal preference modeling of an item. They also used IQON3000 for their experiment and reported better performance in terms of AUC (0.8502) compared to Bi-LSTM (0.66110, BPR-MF (0.7867), VBPR (0.8088) and GP-BPR (0.8321).

Model-Based Collaborative Filtering Technique

The model-based CF algorithm works by constructing a model for the prediction of ratings on the unseen items of users based on the past ratings of the users [97,183]. Machine learning or data mining approaches can be used to build the model-based CF technique. To do so, this model may categorize users into single or multiple clusters. However, single cluster categorization is often problematic for prediction or recommendation as the user may have a variation of tastes with the different items [42,192]. Therefore, most of the model-based CF algorithms categorize the user into multiple clusters [184]. With the evolution of the use of learning algorithms, model-based recommendation systems have begun to use some algorithms such as association rules, clustering, decision tree, artificial neural network, link analysis, regression and Bayesian classifiers [57,193,194].

Memory-Based Collaborative Filtering Technique

Unlike the model-based CF filtering technique, a memory-based CF algorithm predicts the user's seen items based on the users' past ratings. This technique of CF filtering is simple and straightforward, and that is why it has been broadly accepted for real-life application [195]. This model can quickly incorporate the most up-to-date information for the prediction, which is considered one of the advantages of this model. However, making the memory-based algorithm scalable is one of the biggest challenges of this

technique [196]. A memory-based CF recommendation system can be built on either a user-based or an item-based technique. In the case of a user-based technique, this involves calculating the similarities of the user ratings on the same items forms the model. On the contrary, the item-based technique is constructed by calculating the similarities between the items [57,187,197].

Hybrid Collaborative Filtering Technique

The hybrid collaborative filtering (CF) technique is a combination of the memory- and model-based CF filtering techniques. This technique has been developed to utilize the advantages of memory and model-based CF techniques and to mitigate the issues CF technique has, such as sparsity and diversity [188]. Wang et al. proposed a hybrid collaborative filtering technique-based recommendation model, which combines a fashion theme-based model with user's body attributes as well as clothing features [173]. They developed the dataset using both real world data and a 3D scanned body image dataset. Their model achieved 80% accuracy, which is higher than existing fashion recommendation systems used in the research conducted by [41,94,131,198].

5.3.3. Hybrid Filtering Technique

The hybrid filtering (HF) technique combines multiple recommendation techniques to achieve better system optimization and avoid different limitations and challenges of a basic recommendation system. The concept behind implementing the hybrid technique is that the combination of algorithms would provide more appropriate and effective recommendations to users than a single algorithm. Hence, this is the disadvantage of using one algorithm-based recommendation system [179,199]. This construction is beneficial when the dataset lacks user preferences; information about such preferences builds the foundation of collaborative recommendations. By assuming the result of content-based filtering (R1) and result of collaborative filtering (R2), the hybrid filtering technique calculates the weights of these results as R3 and then, depending on the weights, it combines the results by influencing the higher weighted result and recommends the final product R4, which resembles the results R1 and R2, as shown in Figure 6.

Qian et al. proposed a hybrid visual recommendation system by combining conditional random fields with deep lab multiscale (MSc) and large field-of-view (large FOV)-based neural networks [143]. Their recommendation system has two properties. Firstly, it is knowledge-based, which helps it learn a pairwise preference or occurrence matrix based on the knowledge learnt from examples such as images uploaded to fashion blogs. Secondly, it has features of content-based filtering as it uses a deep learning network for learning the feature representation. They used 10,000 street-style images for image segmentation, 45,645 street-style images for product localization and 14,000 online fashion images for texture classification. Their proposed DeepLab-MSc-LargeFOV + CRF for image segmentation outperformed other baseline models such as fully convolutional networks (FCN), combination of convolutional networks (FCN) and the conditional random field (CRF) network model. The proposed model achieved 73.99% mean intersection over union (IoU), which was higher than the other baseline models. Their proposed recurrent fully convolutional networks (R-FCN) achieved an average mean precision (m-AP) of 83.4%, which was higher than that of the baseline models single shot multibox detector and recurrent convolutional neural network (R-CNN).

Hybrid filtering techniques can be classified into seven categories: weighted hybridization; switching hybridization; cascade hybridization mixed hybrids; feature-combination; feature-augmentation; and metalevel hybridization [200]. Weighted hybridization generates a recommendation or prediction list by combining the results of multiple recommendation systems based on the integration of the scores derived from all of the filtering techniques using a linear formula [201]. Switching hybridization switches to one of the used recommendation techniques based on a heuristic, which reflects the system's ability to generate a good rating [202]. Cascade hybridization employs an iterative refinement

process to construct an order of preferences or choices of different items. A mixed hybrid recommendation combines the prediction results of different recommendation techniques simultaneously instead of having a single recommendation per item [201]. In the feature-combination hybrid technique the features generated by a particular recommendation technique are supplied to another recommendation technique [200]. The feature-augmentation technique uses the ratings and other relevant information produced by a previous recommendation system as an input for another recommender, which results in the generation of a model that is always richer in terms of information usage in comparison with a single rating [57]. The metalevel hybrid technique uses an entire model as an input for a second recommendation algorithm, which was previously learned by a first algorithm [202].
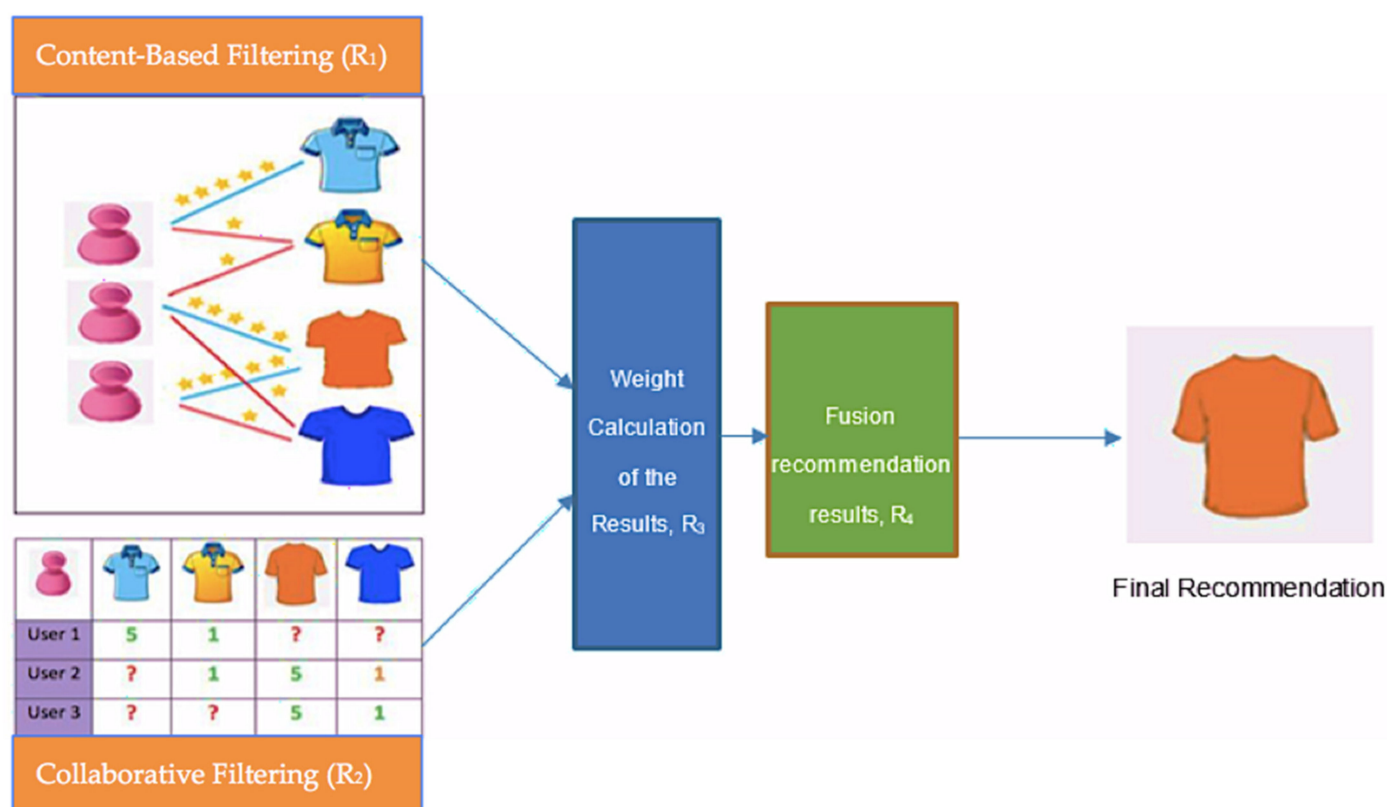

Figure 6. Hybrid filtering process.

5.3.4. Hyperpersonalization Filtering Technique

Personalization is a system that uses the profiling of customers to make certain assumptions about the users. These assumptions are based on certain specific features and traits gathered from the profiling. For example, suggesting ads to buyers since they have ordered or searched for a similar product online is a very common strategy used these days. This technique of personalization can bring a huge boom in sales for companies according to their sales reports. Hyperpersonalization uses the same strategy and works more on it. Hyperpersonalization is an advanced technique built over the concept of personalization, in which the model not only investigates the item or product that was bought, but also looks into other details such as location of purchase, mode of purchase, cost of purchase, keywords inserted during purchase, demographics of the person who purchased, etc. [34,124,130,135,139,203,204]. Hyperpersonalization delves into the intricate details and thereby produces much better and effective personalization, which has made it popular in recent times [5,29,146,205,206]. The implementation of this filtering technique with virtual try-on facilities can develop a recommendation system that offers instant image generation of a user wearing the selected fashion item [207,208]. The use of a gen-

erative adversarial network (GAN) in developing a state-of-the-art hyperpersonalization recommendation system can facilitate an apparel type recommendation or clothing fit recommendation with a real time image generation [157,206,209]. Kang et al. proposed a visually aware deep Bayesian personalized ranking method (DVBPR) recommendation system that can generate new clothing given a user profile and a product category, which is designed with more customization in order to provide more accurate recommendations to users or consumers [5]. Their system achieves an area-under-curve or AUC value of 79.64% on provided dataset, which is higher than that of baseline models such as random ranking, popularity ranking (PopRank), weighted approximated ranking pairwise (WARP), Bayesian personalized ranking combined with matrix factorization method (BPR-MF), visual similarity-based raking method (VisRank), factorization machines (FM) and visually aware personalized ranking (VBPR). There has not been rigorous research on hyperpersonalized filtering-based recommendation models, as it is comparatively new compared to the other three filtering techniques mentioned above. Therefore, this paper proposes a novel hyperpersonalized filtering-based recommendation model in Section 6 below, which can be used by future researchers in this field.

5.4. Strengths and Weakness of Filtering Techniques

The successful outcome of the recommendation system depends on the relevance of the filtering technique and its compatibility with the proposed model. Therefore, researchers should consider the strengths and weaknesses of the corresponding filtering techniques while conducting research on fashion recommendation systems. Table 4 presents the strengths and weakness of the each of the recommendation filtering techniques discussed above.

Table 4. Strengths and weakness of recommendation filtering techniques.

Filtering Techniques	Strength	Weakness
Content-based	• Products recommended based on the evaluation of experienced users. • CBF does not need any information from other users, which makes this technique more feasible and less time consuming. • CBF can attain the specific interest of a user and make recommendations accordingly. • Provides a valuable explanation, which motivates users to make decisions.	• As it is CBF domain-dependent, rigorous domain knowledge is required to make precise recommendations. • The model only recommends products based on an existing database of previous users' interest, which restricts its expansion. • Due to cold start problem, cannot be applied to make recommendations to new users. • This method suffers limited content analysis issues, meaning users are restricted to the items already recommended.
Collaborative	• CF does not depend on domain knowledge. • It does not require contextual attributes. This technique can be applied to one of the multiple users' generators. • This method can allow users to discover new interests despite the absence of content in the user's profile.	• Not applicable for new users, similar to content-based methods. • Difficult to include side features for query/items. • Single-rating CF was successful whereas multicriteria rating is still under optimization.

Table 4. *Cont.*

Filtering Techniques	Strength	Weakness
Hybrid	• Proposed multicriteria rating over CF-based recommendation. • Solved cold start problem. • Hybrid algorithm overcame the single algorithms' shortcomings. • Cascade hybridization, one of the HF methods, exhibited high sensitivity, resulting in efficient recommendations. • HF does not exclusively depend on collaborative data. • Solved sparsity issue of CF method using metalevel technique.	• Hybrid nature made this method complex because of the necessity of applying numerous recommendation parameters for analysis.
Hyperpersonalization	• Yields better results when it comes to customer satisfaction and needs. • Enhanced customer experiences. • Higher return on investment (ROI). • Highly engaging social campaigns.	• Involves taking additional data, which makes the process a bit more expensive. • Privacy invasion can be a concern, as when more than the required data are collected it can lead to a privacy issue.

6. Prospects, Challenges and Recommendations for Future Research

There has been significant progress recently in fashion recommendation system research, which will benefit both consumers and retailers soon. The use of product and user images, textual content, demographic history, and cultural information is crucial in developing recommendation frameworks. Product attributes and clothing style matching are common features of collaborative and content-based filtering techniques. Researchers can develop more sophisticated hyperpersonalized filtering techniques considering the correlation between consumers' clothing styles and personalities. The methods based on employing a scoring system for quantifying each product attribute will be helpful in increasing the precision of the model. The use of virtual sales advisers in an online shopping portal would provide consumers with a real time offline shopping experience. Retailers can collect the data on users' purchase history and product reviews from the recommendation system and subsequently use them in style prediction for the upcoming seasons. The integration of different domain information strengthens the deep learning paradigm by enabling the detection of design component variation, which improves the performance of the recommendation system in the long run. Deep learning approaches should be more frequently used to quickly explore fashion items from different online databases to provide prompt recommendations to users or consumers.

Image quality has always been a critical issue for recommendation systems. Product images taken under controlled environments give higher accuracy in product retrieval and prediction. However, photos taken in a random environment, such as selfies and street style photos, create challenges for the model and lead to inaccurate predictions. Therefore, there should be more research on image parsing, as it is crucial to understand product attributes and human postures, which are applied to predict consumers' fashion preferences. Besides, the development of new state-of-the-art algorithms to analyze randomly captured social media or street photos would be helpful to overcome different obstacles related to image resolution, background, and other technical features. Database generation is always challenging for researchers, particularly when the model is designed to identify new factors and product contents from images that were not identified earlier. The annotation or labeling of such a database is a tedious, time consuming and costly process. The

combination of different databases such as runway images, street photos, look-book images, photos from photo sharing sites and social media images will make it easier to train the model on various fashion categories. Hence, it will increase the robustness of the model. The integration of product images available in online shops with street snapshots will create a large dataset that can be used to parse body and clothing images and distinguish attributes of clothes such as textures and clothing types. There has been limited in-depth research on developing recommendation systems using text (review and comments), product images and user photos together. Therefore, there should be more novel research on developing recommendation models by combining sentiment analysis with user images to provide intelligent and social-network-based hyperpersonalized recommendations. This can be achieved by using hybrid and hyperpersonalized filtering techniques together to develop the recommendation system. The use of social media is rapidly increasing around the world. Nevertheless, retailers and researchers have not widely explored the potentiality of using social media images for clothing recommendation. Moreover, there is still limited research on using image analysis for online fashion recommendation. Therefore, future research on social media should include a holistic analysis of users' images, texts and facial expressions to make the recommendation system more effective.

Researchers should also explore the potential of some widely used statistical tests, such as the sign test and the Friedman test, as a metric of testing the significance of performance evaluation or recommendation accuracy. The sign test is a simple test of significance used to measure the performance of one system over another based on the probability distribution [210]. The Friedman test is a non-parametric test identical to the repeated-measures ANOVA. It ranks the algorithms individually for each dataset, where the best performing algorithm is assigned the 1st rank, the second-best algorithm is assigned the 2nd rank, etc. [211].

6.1. Potential Algorithmic Models for the Future

6.1.1. Multi View Deep Neural Network

Multi view deep neural network (MV-DNN) is a cross-domain recommendation system which treats users as the pivot view and each domain as an auxiliary view. The primary model is based on the hypothesis that if the users have similar tastes in one domain, they should have similar tastes in other domains as well [212]. Therefore, this model can be inefficient in some cases for which it must have some preliminary knowledge on the correlations across different domains. A MV-DNN or MV-CNN can be used in a FRS along with a MLP, which could potentially learn from features of items from cross-domains and user features as well as map users and items to a latent space where the similarity between users and their preferred items can be maximized. Moreover, it can be a great addition to a highly scalable fashion recommendation system to attain rich, feature-based user representation by reducing the dimension of the inputs and the amount of training data required.

6.1.2. Neural Collaborative Filtering

There was an idea to develop a two-way interaction between users' preferences and item features to pinpoint the recommendation system. A dual neural network was proposed to model the two-way interaction between users and items. Fusing the neural interpretation of matrix factorization (MF) with multilayer perceptron (MLP) to develop a new generalized model by making use of both linearity of MF and non-linearity of MLP could enhance the recommendation quality. Xue et al. and Zhang et al. proposed replacing one-shot identifiers with columns or rows of the interaction matrix to retain the user–item interaction patterns [213,214]. In general fashion recommendation systems, the user satisfaction is not observed and there is an inherent scarcity of negative feedback. Using a neural collaborative network in the final part of the system can reduce this by using the output of one layer as the input of the next one to generate feedback for the predicted recommendation. Moreover, it can increase the performance and quality of any

fashion recommendation system, for example color and style selection confidence, for any specific user based on accurate feedback related to the product.

6.1.3. Neural Autoregressive-Based Recommendation

The neural autoregressive distribution estimator (NADE) is a tractable distribution estimator, which provides a desirable alternative to the restricted Boltzmann machine (RBM), which is not tractable. Zheng et al. proposed a NADE-based collaborative filtering model (CF-NADE) by modeling the distribution of user ratings [215]. Later, Du et al. improved the model using a user–item co-autoregressive approach and achieved better performance in both rating estimation and personalized ranking tasks, which gives the model high potential to be used in fashion recommendation systems [216]. Moreover, the model can be further extended to a deep model by combining it with MLP or autoencoder and can potentially use the implicit feedback to overcome the sparsity problem of the rating matrix. The increased performance in rating estimation and personalized ranking tasks can make the cross-matching of the products of the fashion recommendation system much more accurate and time efficient.

6.1.4. Neural Graph Filtering

Liu, X. et al. developed a recommendation system applying neural graph filtering, where they used the Polyvore dataset, the Polyvore-D dataset, and the Amazon Fashion dataset for their experiment [217]. Neural graph filtering is based on graph structures with nodes represented as visual embeddings of the apparel images and edges modeled by the interrelationship among the apparel items. The edge information is accumulated and then circulated through forward propagation to evaluate a compatibility score for a given set of apparel or garment item. It describes the suitability of all the products within the set that match each other. While determining the compatibility score, the inter-relationships among all garment items are considered. Then, the garment set with the highest compatibility score is confidently recommended to the user [217].

6.1.5. Hybrid Model

Zhang et al. combined CNNs with autoencoders for image feature extraction leveraging structural content, textual content and visual content with different embedding techniques, which relates the products' attributes with users' preferences [218]. It combines a heterogeneous network embedding method for interpreting structural information, a stacked denoising autoencoder to learn feature representations from textual information and stacked convolutional autoencoders for visual representation and finally completes the recommendation process in a probabilistic form. Therefore, it could add great value to a fashion recommendation system by generating new and high-quality visual clothing recommendations for the users by combining the confidence scores with the structural information. Vasileva et al. combined conditional similarity networks, visual-semantic embedding, SiameseNet and metric approaches to develop a fashion recommendation model [219]. They conducted extensive experiments on the Maryland Polyvore, Polyvore Outfits-D and Polyvore Outfits datasets and found that their approach outperformed other state-of-the-art recommendation techniques.

7. Discussion

This scholarly article has provided a comprehensive review of the methods, algorithmic models and filtering techniques used in the recent fashion recommendation-based research papers. However, this review paper has some limitations too. Primarily, the focus of this comprehensive review paper was to explore fashion recommendation-based articles published in last decade that explicitly described their frameworks, algorithms, and filtering techniques. To achieve this goal, the articles were searched using keywords relevant to the topic title instead of using the PRISMA technique. However, it did not affect the article extraction methodology, because the authors included and studied all the

research papers relevant to the research focus. However, future researchers could conduct a systematic literature review on the same topic. The initial keyword searching did not include "garment" and "outfit"; however, this did not influence the search results because we also studied the fashion recommendation articles that contained these keywords. The future research can also conduct a review of the datasets that have been used in fashion recommendation-based research articles. Additionally, further reviews of fashion recommendation systems can apply our proposed potential algorithms to any of the available fashion image datasets to evaluate the performance of the recommender systems.

8. Conclusions

Recommendation systems have the potential to explore new opportunities for retailers by enabling them to provide customized recommendations to consumers based on information retrieved from the Internet. They help consumers to instantly find the products and services that closely match with their choices. Moreover, different stat-of-the-art algorithms have been developed to recommend products based on users' interactions with their social groups. Therefore, research on embedding social media images within fashion recommendation systems has gained huge popularity in recent times. This paper presented a review of the fashion recommendation systems, algorithmic models and filtering techniques based on the academic articles related to this topic. The technical aspects, strengths and weaknesses of the filtering techniques have been discussed elaborately, which will help future researchers gain an in-depth understanding of fashion recommender systems. However, the proposed prototypes should be tested in commercial applications to understand their feasibility and accuracy in the retail market, because inaccurate recommendations can produce a negative impact on a customer. Moreover, future research should concentrate on including time series analysis and accurate categorization of product images based on the variation in color, trend and clothing style in order to develop an effective recommendation system. The proposed model will follow brand-specific personalization campaigns and hence it will ensure highly curated and tailored offerings for users. Hence, this research will be highly beneficial for researchers interested in using augmented and virtual reality features to develop recommendation systems.

Author Contributions: S.C.: Conceptualization, Methodology and Writing—Original Draft Preparation; M.S.H.: Conceptualization, Methodology and Writing—Original Draft Preparation; N.R.J.: Writing—Original Draft Preparation and Writing—Reviewing and Editing; M.C.B.: Methodology and Writing—Reviewing and Editing; D.B.: Writing—Original Draft Preparation and Writing—Reviewing and Editing, E.L.: Supervision; Writing—Reviewing and Editing. The authors approved the manuscript and agreed with the submission to the *Informatics* journal. All authors have read and agreed to the published version of the manuscript.

References

1. Barnard, M. *Fashion as Communication*, 2nd ed.; Routledge: London, UK, 2008.
2. Chakraborty, S.; Hoque, S.M.A.; Kabir, S.M.F. Predicting fashion trend using runway images: Application of logistic regression in trend forecasting. *Int. J. Fash. Des. Technol. Educ.* **2020**, *13*, 376–386. [CrossRef]