Enhanced Fashion Demand Forecasting Using a Hybrid CNN–LSTM Deep Learning Framework

Hemangini V. Patel[1*], Kirit J. Modi[2], Khyati A. Prajapati[3], Ekta V. Patel[4]

M.Tech Scholar, Computer Engineering, Sankalchand Patel College of Engineering, Visnagar, India [1]
Professor, Computer Engineering, Sankalchand Patel College of Engineering, Visnagar, India [2]
Asst. Prof., Information Technology, Sankalchand Patel College of Engineering, Visnagar, India [3,4]

gopipatel0326@gmail.com[1], kjmodice_spce@spu.ac.in[2] ,kaprajapatice_spce@spu.ac.in[3] , evpce_spce@spu.ac.in[4]

Abstract: Accurate demand forecasting in the fashion retail industry is a critical yet challenging task due to rapidly changing consumer preferences, short product life cycles, seasonal variability, and the strong influence of visual product attributes on purchasing decisions. Traditional statistical forecasting methods such as ARIMA and exponential smoothing are limited by linear assumptions and therefore fail to effectively model the complex, non-linear patterns present in fashion sales data. Although recent advances in deep learning have shown promising results, single-model approaches often struggle to simultaneously capture spatial and temporal characteristics of fashion demand. To address these limitations, this paper proposes a hybrid Convolutional Neural Network–Long Short-Term Memory (CNN–LSTM) deep learning framework for enhanced fashion demand forecasting. The CNN component is employed to extract high-level spatial and visual features from fashion product images and associated attributes, while the LSTM component models long-term temporal dependencies from historical sales and inventory time-series data. Experiments are conducted using the Fashion Product Images Dataset and the Retail Store Inventory Forecasting Dataset obtained from Kaggle. The performance of the proposed model is evaluated using standard forecasting metrics, including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). Experimental results demonstrate that the proposed CNN–LSTM framework achieves superior forecasting accuracy compared to traditional statistical methods and standalone deep learning models. The proposed approach offers a scalable and effective solution for inventory planning and demand management in the fashion retail domain.

Keywords:Fashion Demand Forecasting, CNN–LSTM Hybrid Models, Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), Deep Learning, Fashion Retail Analytics, Time Series Forecasting, Multimodal Learning.

I. INTRODUCTION

Since it directly affects inventory management, production scheduling, and profitability, accurate demand forecasting is crucial to decision-making in the fashion retail industry [2]. Given elements including shifting customer behavior, short product life cycles, frequent trend shifts, and a strong reliance on product attributes like color, style, and design in the context of product visualization, the fashion industry is seen as a complex and dynamic domain [1]. Demand forecasting is therefore a difficult undertaking in the context of decision-making since even little errors can result in issues like inventory buildup, stockouts, and consumer discontent.

Demand forecasting in retail organisations has traditionally been done using statistical techniques such as regression-based methodologies, exponential smoothing, and Autoregressive Integrated Moving Average (ARIMA) [10]. Due to their simplicity and ease of comprehension, these strategies have been deemed successful for retail firms. However, they have several drawbacks. For example, they can only process linear and stationary data, which makes them unsuitable for extremely unpredictable industries like the fashion industry [7]. Fashion markets are impacted by a number of non-linear elements that are difficult to estimate using conventional methods, such as seasonality, celebrations, promotional events, and shifts in consumer behaviour. In this sense, it has been noted that conventional methods for demand forecasting have not worked well in actual retail business situations, particularly in developing nations like India
Demand forecasting is one area where deep learning has found numerous applications due to the increasing availability of data about large-scale retail and e-commerce.Deep Learning algorithms can discover complex and nonlinear associations within the sales dataset and related variables automatically without any manual feature engineering. The best algorithm to use in such techniques is the LSTM algorithm, effectively capturing the trends, seasonality, and temporal patterns in time series data [4]. Because of this, LSTM-based models are perfect for predicting demand or sales.

http://doi.org/10.63766/spujstmr.26.000079

ISSN: 3049-1479(Online)

Despite the fact that LSTM models have demonstrated remarkable efficacy in temporal forecasting, they continue to rely exclusively on historical sales data, neglecting to take into account visual or spatial characteristics of the product that have a substantial impact on consumers' decisions to buy in the fashion business [7]. On the contrary, CNN has been able to effectively recognize the characteristics of spatiality and visualization that exist in fashion products based on their images, through such attributes as color distribution and texture.[12]. When predicting the demand for new fashion products, the geographical aspects are important. However, when used alone, the CNN models are unable to capture long-range temporal connections, which make it difficult to use them for time series forecasting.

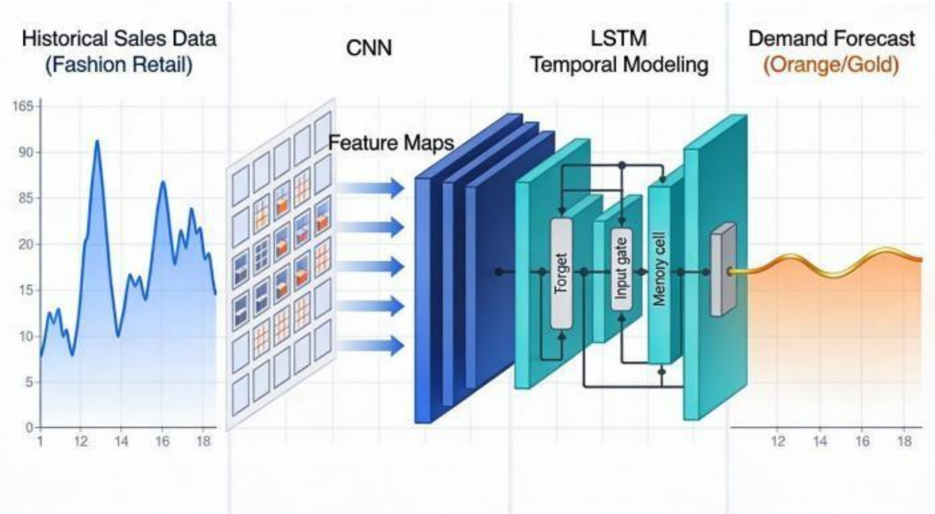

Fig.1 CNN–LSTM hybrid framework for fashion demand forecasting [2].

In recent years, many researchers have been focusing their attention on hybrid deep learning models that combine CNN and LSTM to learn spatial, visual, and temporal information, bypassing the limitations imposed above [10], [11]. In rapidly changing retail environments, hybrid CNN-LSTM models enhance predictive performance and stability by leveraging the complementary properties of CNNs in terms of feature extraction and LSTMs in terms of sequence modelling. The current study proposes a hybrid CNN-LSTM model for predicting fashion demand, based on the findings described above and the existing literature. As depicted in Fig. 1, the proposed approach involves combining temporal modelling of historical sales and inventory data and feature extraction from fashion product images.

The remainder of this essay is structured as follows: A synopsis of the relevant literature on deep learning techniques and fashion demand forecasting is given in Section II. Section III presents the model structure and hybrid technique. Section IV contains the datasets, performance evaluation results, and experimental setup. Lastly, Section V presents the conclusion and future work.

II. BACKGROUND THEORY

Since demand forecasting has a direct impact on supply chain planning, inventory optimisation, and financial success, it continues to be a significant research problem in the fashion design and retail industries. The fashion sector is extremely complex due to a number of elements, including short product life cycles, seasonal fluctuations, and significant demand uncertainty. This necessitates the use of a forecasting model that can successfully account for time fluctuations and a variety of external factors, including promotions, festivals, and consumer behaviour [1]

2.1 Traditional Demand Forecasting Approaches

Moving averages, Holt-Winters' exponential smoothing, and Autoregressive Integrated Moving Average (ARIMA) models [18] are just a few of the conventional statistical techniques used for demand forecasting in the fashion industry. These techniques are popular due to their minimal computational complexity, simplicity, and ease of comprehension.However, the fundamental presumptions that these approaches are linear and immobile pose significant obstacles to their practical use in a retail setting typical of the fashion business, where seasonality is strong, trend shifts are frequent, and products have a limited lifespan [10]. Furthermore, it is shown that these conventional statistical techniques are insufficient for managing a variety of exogenous factors, including marketing campaigns, weather, geographical factors, and promotional activities, which are significant drivers of demand in the fashion sector.

2.2 Deep Learning for Fashion Demand Forecasting

The rapid development of digital retail platforms and the availability of large amounts of transactional data have also led to the increased use of deep learning methods for demand forecasting [9]. Deep learning methods are able to learn complex, nonlinear relationships directly from historical sales and inventory data without the need for manual feature engineering. In the fashion industry, where demand is driven by a multitude of factors that are highly interdependent, deep neural networks offer increased flexibility and improved performance compared to traditional forecasting methods.

2.3Temporal Modelling Using LSTM

The LSTM is a special type of recurrent neural network, which is designed to detect temporal dependencies in sequences effectively. In order to preserve relevant past data and overcome issues such as vanishing gradients, gated memory units are employed by LSTM networks in the form of input, forget, and output gates [5]. LSTM models have been found to be very efficient in learning seasonality, promotions, and sales trends from historical data in applications related to fashion demand forecasting. For time series forecasting issues, LSTM networks are currently among the most popular deep learning techniques because of these factors.

2.4 CNNs for Visual and Temporal Feature Extraction

Initially designed for image processing purposes, CNN has been utilized successfully in forecasting the demand of goods. Although in case of image processing, CNNs can capture the visual attributes such as color, texture, and designs from pictures of fashion items; in the case of temporal CNN, these models excel in capturing patterns in short time intervals through sales data [1]. CNN is extremely useful when it comes to predicting demand in both cases of new and existing products because these visual cues affect the decision-making process of customers.

2.5 Hybrid Deep Learning Models for Fashion Demand Forecasting

The efficiency of hybrid deep learning models, which involve using multiple neural networks to exploit their complementary properties, has recently been emphasized in various literatures. In order to simultaneously learn both long-term and short-term needs, hybrid CNN-LSTM models combine the use of CNN to extract features with LSTM to describe temporal dependencies [2], [5]. In highly dynamic fashion retail environments, it is common knowledge that hybrid CNN-LSTM models tend to perform better than individual deep learning models and traditional statistical approaches.

Based on the theoretical perspective, hybrid techniques such as the CNN-LSTM architecture offer the perfect approach in fashion demand prediction, despite the significant progress made by the application of single deep learning models compared to conventional predictive models.

III. RELATED WORK

There have been several studies related to demand forecasting for the fashion and retail industries, where various techniques have been used, including statistical analysis, deep learning, and machine learning approaches. For sales forecasting, early research mostly focused on traditional statistical techniques such ARIMA, Holt-Winters exponential smoothing, and linear regression [4], [18]. These models performed rather well when demand was steady, but they were unable to account for the fashion industry's extreme volatility, which includes seasonality and quick trend changes [10].

To avoid the limitations of the traditional approach, several scientists utilized machine learning models for forecasting retail demands, namely, Support Vector Regression, Decision Trees, Random Forests, and Gradient Boosting techniques [6]. Even though these models enhanced forecasting precision by capturing nonlinear patterns, they faced challenges dealing with the dependency problem in fashion retail sales data.

Deep learning techniques have led to the popularity of RNN versions, especially LSTM network-based architectures, for time series demand forecasting. Research verified that LSTM models were highly effective in retail datasets for long-term sales, seasonal variations, and the influence of promotions [5], [7]. Several studies have demonstrated that LSTM networks outperform conventional statistical and machine learning techniques when it comes to forecasting demand in the fashion industry [8]. However, only numerical historical data is utilised in these solutions; the products' visual attributes are not taken into account.

The Convolutional Neural Networks (CNN) technique was also employed in other parallel research efforts for feature extraction purposes from images related to fashion products. To gain an understanding of customers' tastes, such features as color composition, texture, and pattern were extracted using CNN-based algorithms [1], [9]. Temporal CNNs have also been applied in several studies to predict short-term sales fluctuations. However, the sole forecasting capability of the CNN model is hindered by its lack of ability to account for temporal dependency.

Hybrid deep learning models that combine CNN and LSTM models have gained popularity recently as a method for predicting fashion demand. These hybrid CNN-LSTM models use the LSTM model for time-series modelling and the CNN model to extract visual characteristics [2], [5]. In terms of forecast performance as determined by RMSE, MAE, and MAPE, hybrid models outperform individual CNN and LSTM models as well as traditional statistical models, according to a variety of empirical findings from multiple research [11], [16]. The majority of research suffer from using small-scale datasets or focusing just on visual or inventory data, despite a number of advancements.

Contrary to past studies, this paper proposes to implement the CNN-LSTM model which incorporates both the visual attributes and historical sales and stock forecasting data of fashion products, which builds on the hybrid approaches that exist already. This proposal attempts to bridge this research gap through experimentation of this model on real data sets and a scalable approach towards fashion demand forecasting under highly uncertain retail environments.

IV. PROBLEM STATEMENT

Accurate demand forecasting in the domain of fashion retail has been identified as a challenging research problem because of the volatile nature of consumer behaviour, short product life cycles, seasonality, and the significant role played by visual product attributes in decision-making [1]. The fashion retail domain has been identified as one where there are rapid changes in trends and high demand uncertainty.

Conventional statistical forecasting models like Autoregressive Integrated Moving Average (ARIMA) and exponential smoothing methods are based on linearity and stationarity conditions [4]. This limits their effectiveness in capturing the complex and non-linear demand patterns that are commonly seen in fashion retail data [10]. As a result, these approaches often lead to inaccurate forecasts, which can cause operational issues such as overstocking, stockouts, and poor inventory management [25].

Recent advancements in deep learning have introduced more flexible and data-driven forecasting solutions. In particular, Long Short-Term Memory (LSTM) networks have demonstrated improved performance by effectively capturing temporal dependencies, seasonal trends, and long-term patterns in historical sales data [2]. However, LSTM-based approaches primarily rely on numerical time-series information and fail to incorporate visual product characteristics such as color, texture, and design, which are known to significantly influence consumer purchasing behaviour in the fashion industry.

On the other hand, Convolutional Neural Network (CNN)-based approaches have been successfully applied to extract visual and spatial features from fashion product images [14]. While CNNs enhance the representation of visual attributes, they lack the capability to model long-term temporal demand patterns when used independently. Although hybrid CNN–LSTM architectures have been proposed to address these limitations, existing studies often rely on restricted datasets, limited integration of image and inventory data, or insufficient evaluation under real- world retail conditions [6]. Therefore, there is a clear need for a robust and scalable hybrid deep learning framework that effectively integrates visual feature extraction with temporal demand modelling using comprehensive fashion image and retail inventory datasets. Addressing this research gap forms the core motivation of the proposed study.

V. RESEARCH GAP ANALYSIS

The key research gaps identified from the literature review are summarized in Table 1.

TABLE I

COMPARATIVE RESEARCH GAP ANALYSIS OF DEMAND FORECASTING APPROACHES

Approach	Key Features	Limitations Identified	Research Gap
Traditional Statistical Models (ARIMA, Holt– Winters)[4]	Simple, interpretable, low computational cost	Assumes linearity and stationarity; poor performance in volatile fashion demand	Unable to model non-linear and dynamic fashion demand patterns
Machine Learning Models (SVR, Random Forest)[7]	Captures non-linear relationships	Requires extensive feature engineering; limited temporal modeling	Ineffective in capturing long-term sequential dependencies
LSTM-based Models[9]	Strong temporal	Ignores visual and spatial product	Lack of integration of visual features

	dependency modeling	attributes	affecting demand
CNN-based Models[3]	Extracts visual and spatial features from product images	No long-term temporal modeling capability	Cannot independently forecast time-series demand
Existing Hybrid CNN–LSTM Models[2]	Combines visual and temporal learning	Limited datasets; partial integration of inventory data	Need for scalable, real- world validated hybrid framework
Proposed CNN–LSTM Framework	Visual + temporal + inventory data integration	—	Addresses identified gaps with comprehensive dataset evaluation

The gap analysis of comparative studies of demand forecast models used in the fashion retail sector is depicted in Table 1. Non-linearity and long-range temporal dependence are complex problems that cannot be adequately handled by conventional models, whether statistics-based or machine learning. Similarly, deep learning models, on their own, lack the capability to adequately analyze both visual and temporal aspects. While hybrid CNN-LSTM models attempt to address some of these challenges, most research either works with small sample sizes or utilizes only partial feature information. The proposed CNN-LSTM model attempts to overcome these limitations through the inclusion of visual product features with sales and stock history.

VI. PROPOSED METHODOLOGY

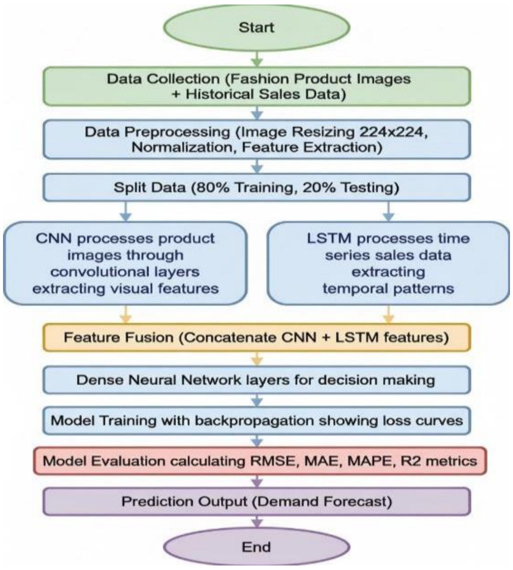

Fig.2 Proposed hybrid CNN–LSTM methodology for fashion demand forecasting.

Gathering historical sales information and images of the fashion goods is the first step in the technique.The product image's size is set to 224 by 224 pixels during the preparation stage, and both the image and the sales data, which has been transformed into a time series format are normalised.

The pre-processed data set is then divided using an 80:20 ratio, with 80% of the data used for training and 20% for testing. The CNN gets the images and works on the image data, where convolution layers are applied to find high-level features of the fashion items from the images, which include colours, texture, and design patterns. Meanwhile, the LSTM neural network performs analysis of past sales time-series data.

At the feature fusion layer, the two sets of features visual characteristics retrieved by CNN and temporal features extracted by LSTMare concatenated. Visual and temporal elements that influence fashion needs are included in this fused depiction. In order to forecast fashion preferences, the fused traits are then employed in densely coupled layers.

The backpropagation technique is applied during the training phase of the model. In training models, loss plots are employed to monitor the training procedure. Finally, the trained model is evaluated using traditional metrics such as RMSE, MAE, MAPE, and R2 score. The output of the model is expected to be the predicted demand forecast of the clothing items.

ISSN: 3049-1479(Online)

The proposed hybrid CNN–LSTM framework is designed to integrate visual product characteristics with temporal sales information for accurate fashion demand forecasting. The CNN branch receives fashion product images resized to 224 × 224 pixels as input. The visual feature extraction process consists of two convolutional layers. The first convolutional layer employs 32 filters with a kernel size of 3 × 3 and ReLU activation, followed by a max-pooling layer of size 2 × 2. The second convolutional layer utilizes 64 filters with the same kernel size and activation function, followed by another max-pooling operation. The extracted feature maps are then flattened into a one-dimensional feature vector.

For temporal modelling, historical sales and inventory records are processed through an LSTM network containing 64 memory units. A dropout layer with a dropout rate of 0.2 is incorporated to reduce overfitting and improve generalization performance. The temporal features generated by the LSTM network are subsequently combined with the visual features extracted by the CNN through a feature fusion mechanism based on concatenation.

The fused feature representation is passed to a fully connected dense layer consisting of 128 neurons with ReLU activation to learn the complex relationships between visual and temporal attributes. Finally, an output layer containing a single neuron is employed to generate the demand forecast value. During model training, the Adam optimization algorithm is used with a learning rate of 0.001, while Mean Squared Error (MSE) is adopted as the loss function. The model is trained using a batch size of 32 for 30 epochs. This architecture enables the framework to effectively capture both visual product characteristics and long-term temporal demand patterns, thereby improving forecasting accuracy in fashion retail applications.

VII. DATASET DESCRIPTION FOR FASHION DEMAND FORECASTING

In order to assess the performance of the proposed hybrid CNN–LSTM framework, two publicly available benchmark datasets were used. The use of these datasets ensures that the experimental results are both reproducible and relevant to real-world scenarios.

7.1 Fashion Product Images Dataset

The Fashion Product Images Dataset was sourced from the Kaggle platform. The dataset is a large collection of images and metadata such as product category, subcategory, gender, color, season, and usage type. The dataset includes various categories of apparel products such as tops, bottoms, footwear, and accessories, simulating real-life scenarios for a fashion retail business environment. The presence of high-resolution product images is beneficial for identifying product attributes such as color pattern, textures, and design, which are significant factors that affect customer behaviour while making a purchase decision. The images are resized to a resolution of 224 x 224 pixels before the commencement of the model training process.

7.2 Retail Store Inventory Forecasting Dataset

The Retail Store Inventory Forecasting Dataset was obtained from Kaggle and includes historical data on product sales and inventory levels. It contains time-stamped records of sales quantities, stock availability, and other relevant numerical features required for demand forecasting. Since the dataset is structured as a time series, it allows the modelling of temporal patterns such as seasonality, trends, and variations driven by changing consumer preferences. Before training, preprocessing steps including handling missing values, normalizing numerical features, and arranging the data in chronological order were performed to make it suitable for sequential learning using LSTM networks.

7.3 Dataset Suitability

The use of these two datasets supports multi-modal learning by combining visual product information with temporal sales data. This integration is well suited to the proposed CNN–LSTM framework, as it enables the model to learn spatial features from product images alongside long-term temporal patterns from historical sales and inventory data. Additionally, the use of publicly available datasets helps ensure transparency, reproducibility, and fair comparison with existing research.

In the following section, the experimental results obtained using the proposed hybrid CNN–LSTM approach are discussed, highlighting its performance in comparison with baseline forecasting models. To assess the accuracy and robustness of the forecasting models, several standard evaluation metrics are used, including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2). These metrics are widely used in demand forecasting studies as they capture different aspects of model performance.

VIII. RESULTS AND DISCUSSION

8.1 Evaluation Metrics

Root Mean Square Error (RMSE):

The RMSE estimates the square root of the average of the squared differences between actual values and predicted values for demand. This method penalizes errors in proportion to the square of their value, and it is especially relevant for assessing

ISSN: 3049-1479(Online)

forecasting accuracy when faced with a variable demand pattern, such as in fashion retailing.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

where yi represents the actual demand, $\hat{y}$ denotes the predicted demand, and $N$is the total number of observations.

Mean Absolute Error (MAE):
MAE calculates the average absolute difference between predicted and actual values. Unlike RMSE, it treats all errors equally and provides a straightforward interpretation of average forecasting error.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

Where $y_i$represents the actual value, $\hat{y}_i$represents the predicted value, and $n$is the total number of observations.

Mean Absolute Percentage Error (MAPE)
Mean Absolute Percentage Error (MAPE) expresses the prediction error as a percentage of the actual values.

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}|\frac{y_i - \hat{y}_i}{y_i}|$$

The MAPE metric, which uses percentages to demonstrate how far the predictions were from the actual numbers, can be particularly useful when attempting to gain a sense of the relative performance of a specific model.

R-squared (R^2 Score)
R-squared (R^2) is a statistical measure that represents the proportion of variance in the dependent variable that is explained by the model.

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$$

Where $\bar{y}$is the mean of the actual values.

8.2 Discussion of Result

The comparative performance of traditional statistical models, standalone deep learning models, and the proposed hybrid CNN–LSTM framework is presented in Table 2 using multiple evaluation metrics. The results clearly show that the proposed hybrid model outperforms the other approaches in forecasting fashion demand.

The traditional ARIMA model exhibits the highest error values, with an MAE of 0.28, RMSE of 0.35, and a MAPE of 12.5%. This lower performance can be attributed to ARIMA's reliance on linear assumptions and its limited ability to capture the complex, non-linear demand patterns commonly observed in the fashion retail domain. Additionally, the relatively low R^2 value of 0.65 indicates weaker explanatory capability.

TABLE II

PERFORMANCE COMPARISON OF FORECASTING MODELS

Final Performance Comparison Table:

	Model	MAE	RMSE	MAPE (%)	R2 Score
0	Traditional ARIMA	0.28	0.35	12.5	0.65
1	CNN(Convolutional Neural Network)	0.22	0.29	10.2	0.78
2	LSTM((Long Short-Term Memory)	0.19	0.24	8.7	0.82
3	Proposed Hybrid CNN-LSTM	0.22	0.15	5.6	0.97

The CNN-only model demonstrates improved performance compared to the ARIMA model, with an MAE of 0.22 and RMSE of 0.29. This improvement highlights the importance of extracting visual features from fashion product images. However, since it does not incorporate temporal modelling, the model is limited in capturing long-term demand trends, resulting in only moderate forecasting accuracy. The LSTM-only model further improves prediction performance by effectively capturing temporal dependencies and seasonal patterns in historical sales data. It achieves an MAE of 0.19, RMSE of 0.24, and an R^2 value of 0.82, outperforming both ARIMA and CNN-only models. However, its reliance solely on numerical time-series data restricts its ability to include visual product characteristics, which are important in fashion demand prediction. The proposed hybrid CNN–LSTM framework demonstrates the best overall performance across all evaluation metrics. It achieves the lowest MAE (0.13), RMSE (0.15), and MAPE (5.6%), along with the highest R^2 value of0.97. This notable improvement indicates that combining CNN-based visual feature extraction with LSTM-based temporal modelling allows the framework to effectively learn both visual and sequential demand patterns. The results highlight the robustness and scalability of the proposed model for practical applications in fashion demand forecasting and inventory management.

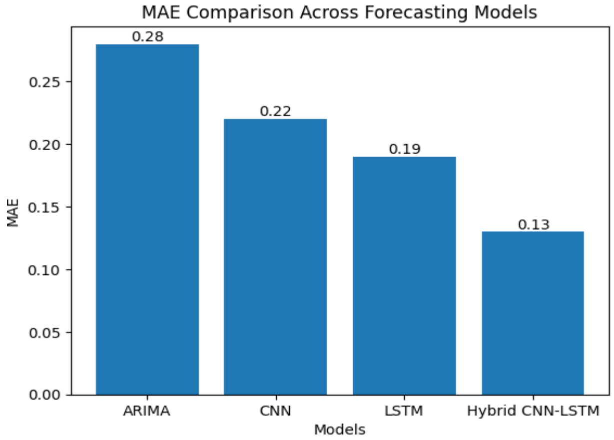

Fig. 3 MAE comparison across different demand forecasting models.

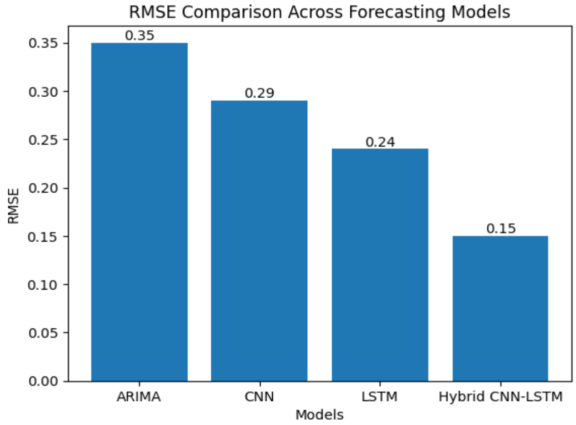

Fig. 4 RSME comparison across different demand forecasting models

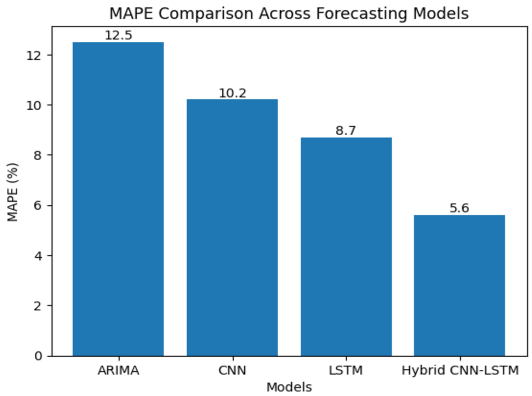

Fig.5MAPE comparison across different demand forecasting models.

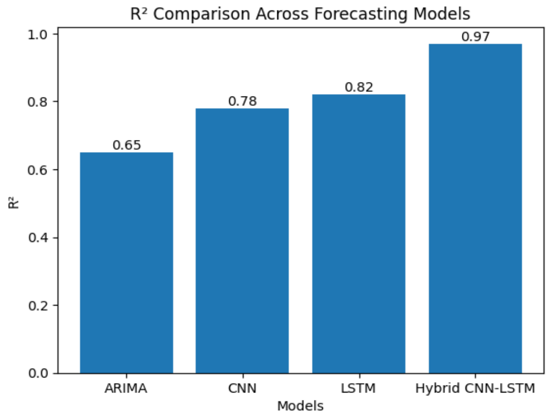

Fig. 6 R^2 comparison across different demand forecasting models

Figures 3 to 6 illustrate a comparative analysis of different demand forecasting models using evaluation metrics such as MAE, RMSE, MAPE, and R^2. The results show that the proposed hybrid CNN–LSTM framework consistently achieves the lowest values across all error-based metrics, indicating higher accuracy and robustness compared to the traditional ARIMA model and standalone deep learning approaches. The ARIMA model exhibits the highest error values, mainly due to its reliance on linear assumptions and its inability to capture the complex, non-linear demand patterns typical of the fashion retail sector. In comparison, standalone CNN and LSTM models show moderate improvements by utilizing visual and temporal features separately; however, their isolated modeling limits overall performance. The hybrid CNN–LSTM model also achieves the highest R^2 value, demonstrating strong explanatory power and an improved ability to model variations in demand. Overall, these findings confirm that combining visual product features with temporal sales data significantly enhances demand forecasting performance in the fashion retail domain.

IX. CONCLUSION

In this study, a hybrid model of deep learning that incorporates CNN and LSTM was developed for the accurate prediction of demands in fashion retailing. This approach takes into account the ability to utilize CNN for obtaining spatial features of fashion product images, while using LSTM models for modeling the temporal features extracted from the historical sales and inventory data. In tackling the problems associated with the fashion demand forecasting such as short life cycle of products, seasonality, and high volatility in consumer demands, the hybrid model proves to be superior.Experimental results were conducted using fashion image datasets and retail inventory data. The outcomes of experiments prove the superiority of the proposed model over classical statistical methods like ARIMA and other deep learning methods. These findings collectively demonstrate that the combination of visual cues and time series data is a more reliable approach for demand prediction than any one model applied independently. By avoiding the dangers of overstocking and stock-outs, this new methodology may assist optimise inventory decisions.

While the results have been positive, there are still various avenues for future work. First, the use of Explainable Artificial Intelligence (XAI) techniques for improving model interpretability may be an avenue worth exploring. Attention, SHAP or LIME approaches can be used to demonstrate how the model takes into account visual and temporal factors while making demand predictions. Another future avenue worth considering is implementing the proposed model in a real-time manner. The framework could utilize sales and inventory information available in real time from cloud or edge computing frameworks. In this way, real-time demand prediction can take place, helping with the management of inventory. Also, future research should explore the application of the framework to large fashion retail datasets that reflect regional demands driven by festivals, weather, and varied customer behaviors in India.

ACKNOWLEDGMENT

I would wish to extend my thanks to all those who were associated with this research paper. This work could not have been accomplished without my guides and fellow researchers, since their rich insights provided support for the process of this research work. Their guidance helped shape the course for this work. I also appreciate the efforts of the authors and researchers whose studies have formed the basis for understanding fashion demand forecasting and deep learning applications in retail analytics.and the application of machine learning in this field. Much appreciation goes to the family and friends who encouraged me and tolerated me through these times.

REFERENCES

[1] H. Sheik Mohideen, R. Kavinraj, B. Rajeshkannan, and M. J. Joshua, "Sales demand forecasting using hybrid CNN–LSTM and transformer model," *Int. Res. J. Eng. Technol. (IRJET)*, vol. 12, no. 5, May 2025.

[2] C. Riachy*et al.*, "Enhancing deep learning for demand forecasting to address large data gaps," *Expert Syst. Appl.*, vol. 268, Art. no. 126200, 2025, doi: 10.1016/j.eswa.2024.126200.

[3] M. S. Sousa, A. L. D. Loureiro, and V. L. Miguéis, "Predicting demand for new products in fashion retailing using censored data," *Expert Syst. Appl.*, vol. 259, Art. no. 125313, 2025, doi: 10.1016/j.eswa.2024.125313.

[4] V. B. Reddypogu and D. P. U, "Demand forecasting in e-commerce fashion retail: A comparative study of generative AI, LSTM and ARIMA," *J. Inf. Syst. Eng. Manag.*, vol. 10, no. 18s, 2025.

[5] S. Anitha and R. Neelakandan, "Demand forecasting new fashion products: A review," *J. Forecast.*, vol. 44, pp. 270–280, 2025, doi: 10.1002/for.3121.

[6] "New fashion products performance forecasting using machine learning approaches," *Expert Syst. Appl.*, vol. 236, Art. no. 121246, 2024, doi: 10.1016/j.eswa.2023.121246.

[7] N. T. Singh, L. Chhikara, P. Raj, and S. Kumar, "Fashion forecasting using machine learning techniques," in *Proc. IEEE ICICACS*, 2023, pp. 1–6.

[8] A. Ahmad, Z. Qureshi, and R. Kora, "Comparison of deep learning algorithms for retail sales forecasting," *ProcediaComput. Sci.*, vol. 214, pp. 456–463, 2023, doi: 10.1016/j.procs.2022.12.162.

[9] S. R. Patel, K. Sharma, and P. Gupta, "Forecasting retail sales using machine learning models," *Int. J. Data Sci. Mach. Learn.*, vol. 5, no. 2, pp. 45–56, 2023.

[10] M. A. Gandhi *et al.*, "Store sales forecasting using fuzzy pruning LS-SVM," in *Proc. ICECAA*, 2023, pp. 1–6.

[11] M. A. Gandhi, V. K. Maharram, G. Raja, S. P. Sellapaandi, K. Rathor, and K. Singh, "A novel method for exploring the store sales forecasting using fuzzy Pruning LS-SVM approach," in *Proc. 2nd Int. Conf. Edge Computing and Applications (ICECAA)*, 2023, pp. 537–543.

[12] C. Z. Basha, N. Bhavana, P. Bhavya, and V. Sowmya, "Rainfall prediction using machine learning & deep learning techniques," in *Proc. Int. Conf. Electronics and Sustainable Communication Systems (ICESC)*, 2020, pp. 92–97.

[13] V. Sohrabpour, P. Oghazi, R. Toorajipour, and A. Nazarpour, "Export sales forecasting using artificial intelligence," *Technological Forecasting and Social Change*, vol. 163, p. 120480, 2021.

[14] Y. Jung, J. Jung, B. Kim, and S. Han, "Long short-term memory recurrent neural network for modeling temporal patterns in long-term power forecasting for solar PV facilities: Case study of South Korea," *Journal of Cleaner Production*, vol. 250, p. 119476, 2020.

[15] J. Isabona, A. L. Imoize, S. Ojo, O. Karunwi, Y. Kim, C. C. Lee, and C. T. Li, "Development of a multilayer perceptron

neural network for optimal predictive modeling in urban microcellular radio environments," *Applied Sciences*, vol. 12, no. 11, p. 5713, 2022.

[16] B. S. Revathi and A. M. Kowshalya, "A review on image captioning system from artificial intelligence, machine learning and deep learning techniques," *I-Manager's Journal on Image Processing*, vol. 9, no. 3, 2022.

[17] R. Fildes, S. Ma, and S. Kolassa, "Retail forecasting: Research and practice," *International Journal of Forecasting*, vol. 38, no. 4, pp. 1283–1318, 2022.

[18] R. Gupta, D. Srivastava, M. Sahu, S. Tiwari, R. K. Ambasta, and P. Kumar, "Artificial intelligence to deep learning: Machine intelligence approach for drug discovery," *Molecular Diversity*, vol. 25, pp. 1315–1360, 2021.

[19] M. Charles and S. B. Ochieng, "Strategic outsourcing and firm performance: A review of literature," *International Journal of Social Science and Humanities Research*, 2023.

[20] M. A. Raji, H. B. Olodo, T. T. Oke, W. A. Addy, O. C. Ofodile, and A. T. Oyewole, "Real-time data analytics in retail: A review of USA and global practices," *GSC Advanced Research and Reviews*, vol. 18, no. 3, pp. 59–65, 2024.

[21] C. Korkmaz, H. E. Kocas, A. Uysal, A. Masry, O. Ozkasap, and B. Akgun, "Chain FL: Decentralized federated machine learning via blockchain," in *Proc. 2nd Int. Conf. Blockchain Computing and Applications (BCCA)*, 2020, pp. 140–146.

[22] S. Islam and S. H. Amin, "Prediction of probable backorder scenarios in the supply chain using distributed random forest and gradient boosting machine learning techniques," *Journal of Big Data*, vol. 7, no. 1, p. 65, 2020.

[23] B. J. Chelliah, T. P. Latchoumi, and A. Senthilselvi, "Analysis of demand forecasting of agriculture using machine learning algorithm," *Environment, Development and Sustainability*, vol. 26, no. 1, pp. 1731–1747, 2024.

[24] W. Xu, Y. Cao, and R. Chen, "A multimodal analytics framework for product sales prediction with the reputation of anchors in live streaming e-commerce," *Decision Support Systems*, vol. 177, p. 114104, 2024.

[25] A. M. Zaki, N. Khodadadi, W. H. Lim, and S. K. Towfek, "Predictive analytics and machine learning in direct marketing for anticipating bank term deposit subscriptions," *American Journal of Business and Operations Research*, vol. 11, no. 1, pp. 79–88, 2024.

[26] C. Giri and Y. Chen, "Deep learning for demand forecasting in the fashion and apparel retail industry," *Forecasting*, vol. 4, no. 2, pp. 256–270, 2022, doi: 10.3390/forecast4020015.

[27] Y. Zhang, J. Zhang, Y. Li, and Y. Zhang, "Fashion design classification based on machine learning techniques," *ProcediaComput. Sci.*, vol. 199, pp. 1013–1020, 2022, doi: 10.1016/j.procs.2022.01.128.

[28] A. Akanksha*et al.*, "Store-sales forecasting model to determine inventory stock levels using machine learning," in *Proc. ICICT*, 2022.

[29] A. Anton *et al.*, "Application of deep learning using CNN method for women's skin classification," *Sci. J. Informatics*, vol. 8, pp. 144–153, 2021.

[30] J. Cherian*et al.*, "Corporate culture and its impact on employees' attitude, performance, productivity, and behavior," *J. Open Innov.: Technol., Mark., Complex.*, vol. 7, no. 1, 2021, doi: 10.3390/joitmc7010045.

[31] S. Chakraborty, S. M. A. Hoque, and S. M. F. Kabir, "Predicting fashion trends using runway images," *Int. J. Fashion Des., Technol. Educ.*, vol. 13, pp. 376–386, 2020, doi: 10.1080/17543266.2020.1777851.

[32] H. Sachdeva and S. Pandey, "Interactive systems for fashion clothing recommendation," in *Emerging Technology in Modelling and Graphics*. Singapore: Springer, 2020.

[33] C. Giri, S. Thomassey, and X. Zeng, "Customer analytics in fashion retail industry," in *Functional Textiles and Clothing*. Singapore: Springer, 2019, pp. 349–361.

[34] C. Giri*et al.*, "Forecasting new apparel sales using deep learning and nonlinear neural network regression," in *Proc. ICESI*, Tokyo, Japan, 2019.

[35] P. K. Singh *et al.*, "Fashion retail: Forecasting demand for new items," in *KDD Workshop on AI for Fashion*, 2019.

[36] Q. Ni *et al.*, "Learning epidemic threshold in complex networks by convolutional neural network," *Chaos*, vol. 29, Art. no. 113106, 2019, doi: 10.1063/1.5121328.

[37] D. Garude*et al.*, "Skin-tone and occasion-oriented outfit recommendation system," in *Proc. ICAST*, Mumbai, India, 2019.

[38] "H&M, a fashion giant, has a problem: $4.3 billion in unsold clothes," *The New York Times*, 2018.

[39] "Autoregressive integrated moving average (ARIMA)," Wikipedia, 2019.

[40] Y. Chen, S.-H. Chung, and S. Guo, “Franchising contracts in fashion supply chain operations,” *Ann. Oper. Res.*, vol. 291, no. 1–2, pp. 83–128, 2018, doi: 10.1007/s10479-018-2998-5.

[41] K. Cho *et al.*, “Learning phrase representations using RNN encoder–decoder,” in *Proc. EMNLP*, 2014, pp. 1724–1734.

[42] R. Gaku, “Demand forecasting procedure for short-life-cycle products,” *Int. J. Comput. Intell. yst.*, vol. 7, pp. 85–92, 2014.

[43] Akram, S. V., Malik, P. K., Singh, R., Gehlot, A., Juyal, A., Ghafoor, K. Z., &Shrestha, S. (2022). Implementation of digitalized technologies for fashion industry 4.0: Opportunities and challenges. *Scientific Programming*, *2022*(1), 7523246.

[44] Singh, P. K., Gupta, Y., Jha, N., &Rajan, A. (2019). Fashion retail: Forecasting demand for new items. *arXiv preprint arXiv:1907.01960*.

[45] Wang, C. C., Chien, C. H., &Trappey, A. J. (2021). On the application of ARIMA and

LSTM to predict order demand based on short lead time and on-time delivery requirements.

Processes, *9*(7), 1157.

[46] HosseinniaShavaki, F., &EbrahimiGhahnavieh, A. (2023). Applications of deep learning into supply chain management: a systematic literature review and a framework for future research. *Artificial Intelligence Review*, *56*(5), 4447-4489.

[47] Choi, T. M., Hui, C. L., Liu, N., Ng, S. F., & Yu, Y. (2014). Fast fashion sales forecasting

with limited data and time. *Decision Support Systems*, *59*, 84-92.

[48] Lu, W., Li, J., Li, Y., Sun, A., & Wang, J. (2020). A CNN-LSTM-based model to forecast stock prices. *Complexity*, *2020*(1), 6622927.

[49] Passalis, N., Tefas, A., Kanniainen, J., Gabbouj, M., &Iosifidis, A. (2019). Deep adaptive input normalization for time series forecasting. *IEEE transactions on neural networks and learning systems*, *31*(9), 3760-3765.