International Journal of Engineering and Technology Innovation, vol. 14, no. 2, 2024, pp. 216-230

Personalized Clothing Prediction Algorithm Based on Multi-modal Feature Fusion

Rong Liu[1,2], Annie Anak Joseph[1,*], Miaomiao Xin[2], Hongyan Zang[2],
Wanzhen Wang[2], Shengqun Zhang[2]

[1]Faculty of Engineering, Universiti Malaysia Sarawak, Kota Samarahan, Sarawak, Malaysia

[2]Computer and Information Engineering, Qilu Institute of Technology, Jinan, China

Received 18 February 2024; received in revised form 05 March 2024; accepted 06 March 2024

DOI: https://doi.org/10.46604/ijeti.2024.13394

Abstract

With the popularization of information technology and the improvement of material living standards, fashion consumers are faced with the daunting challenge of making informed choices from massive amounts of data. This study aims to propose deep learning technology and sales data to analyze the personalized preference characteristics of fashion consumers and predict fashion clothing categories, thus empowering consumers to make well-informed decisions. The Visuelle's dataset includes 5,355 apparel products and 45 MB of sales data, and it encompasses image data, text attributes, and time series data. The paper proposes a novel 1DCNN-2DCNN deep convolutional neural network model for the multi-modal fusion of clothing images and sales text data. The experimental findings exhibit the remarkable performance of the proposed model, with accuracy, recall, F1 score, macro average, and weighted average metrics achieving 99.59%, 99.60%, 98.01%, 98.04%, and 98.00%, respectively. Analysis of four hybrid models highlights the superiority of this model in addressing personalized preferences.

Keywords: fashion consumers, image, text data, personalized, multi-modal fusion

1. Introduction

Personalized fashion classification forecasting technology has garnered considerable attention in the e-commerce and fashion industries. Employing deep learning technology to tackle this challenge is a crucial issue that demands attention. Previous studies have delved into the challenges and potential solutions in this field. For instance, Wu and Zhu [1] proposed deep learning technology and the anatomically constrained neural networks (ACNN) method to enhance the accuracy and efficiency of garment product shape recognition, thus introducing innovations and breakthroughs to fashion prediction. Yuan et al. [2] designed and implemented a clothing matching and recommendation system based on clothing pictures and customer historical behavior data using deep learning technology and data mining to meet consumer demand and boost sales. The target detection technology and the deep residual network (ResNet) extract comprehensive clothing features, addressing the issue of interfering factors in clothing image recognition through multi-depth feature fusion [3]. This method fully leverages global, primary, and local area attributes, directing the recognition process towards the clothing itself, thereby significantly improving the accuracy rate of clothing image recognition.

The multi-feature fusion algorithm is applied to accurately retrieve multi-scale clothing images [4]. Retrieval accuracy is considerably enhanced by thoroughly addressing interference factors in clothing image recognition through the comprehensive exploration of global-to-local multi-scale features of clothing images in conjunction with deep learning and traditional feature

* Corresponding author. E-mail address: jannie@unimas.my

English language proofreader: Yen-Chun Hsieh

extraction methods. The algorithm offers an excellent and workable approach for apparel picture retrieval by combining multi-scale convolutional neural network (CNN) features with conventional characteristics to optimize the ordering of search results. This study proposes a CNN model for identifying categories of clothing design, enhancing classification accuracy by effectively learning image attributes [5]. The model surpasses the performance of traditional hand-designed feature coding methods and existing CNN models in category recognition of clothing design. These findings suggest that deep learning and multi-feature fusion algorithms are well-positioned to enhance the accuracy of personalized fashion classification predictions significantly. Consequently, to address the challenges in customized clothing forecasting, this study proposes introducing a multi-modal fusion algorithm to meet the market demand for customized fashion classification forecasting.

In personalized clothing classification prediction, current research grapples with two primary challenges. Firstly, there is a need to effectively represent and consolidate multi-modal data in a manner that exploits the complementary nature of diverse modalities. This entails developing methodologies for efficiently fusing information from distinct modalities to enable models to understand clothing characteristics better. For instance, while text descriptions provide insights into clothing seasons and types, images offer detailed visual characteristics and labels. Integrating these sources provides a more comprehensive garment description, enhancing personalized classification accuracy.

Secondly, research must address the challenge of data transformation between different modalities. In personalized clothing classification prediction, converting text descriptions into image features or vice versa is essential. This study aims to tackle these challenges by developing practical algorithms capable of learning mapping relationships between modalities and seamlessly integrating information. This transformation process considers the heterogeneity of data and the subjective relationships between modalities, presenting a formidable challenge. However, overcoming this obstacle can lead to more effective utilization of multi-modal data, thereby enhancing the comprehensiveness of personalized classification. Building upon these challenges, the optimization of parameter classification is proposed, with the subsequent steps constructed for this purpose.

The significant contributions of this paper include:

(1) Propose a multi-modal weighted combination method primarily utilizing visual information to achieve a more comprehensive representation of fashion items through weighted fusion of multi-modal features encompassing graphical and textual data.

(2) Propose a personalized clothing prediction algorithm based on multi-modal feature fusion, comprising four main components: feature extraction, feature joint representation, feature fusion, and customized clothing category prediction.

(3) A hybrid model that combines a 1D convolutional neural network (1DCNN) and a 2D convolutional neural network (2DCNN) is introduced to enhance the performance of the fashion prediction model. The proposed model is compared with the same dataset's mainstream models such as ResNet and the temporal convolutional network (TCN).

The structure of the remainder of the paper is as follows. Section 2 delineates the preceding endeavors undertaken by researchers within this domain. Section 3 outlines the pertinent methodologies proposed to tackle the challenges of multi-modal personalized fusion research. The outcomes of experimental and comparative analyses are elucidated in Section 4. Section 5 encapsulates the empirical findings and underscores the contributions of this paper.

2. Related Worked

As societal values evolve and individual preferences become increasingly nuanced, the imperative of personalization grows more pronounced. The interactive bisection algorithm proposed by certain studies achieves attribute differentiation through a bottom-up information genetic algorithm (GA) method. It defines an average classification fuzzy subset to facilitate