Article

AI-Driven Precision Clothing Classification: Revolutionizing Online Fashion Retailing with Hybrid Two-Objective Learning

Waseem Abbas [1], Zuping Zhang [1,*], Muhammad Asim [2,3,*], Junhong Chen [3,4] and Sadique Ahmad [2]

1   School of Computer Science and Engineering, Central South University, Changsha 410003, China; wabbas@csu.edu.cn
2   EIAS Data Science and BlockChain Labortary, College of Computer and Information Sciences, Prince Sultan University, Riyadh 11586, Saudi Arabia; saahmad@psu.edu.sa
3   School of Computer Science and Technology, Guangdong University of Technology, Guangzhou 510006, China; junhong.chen@uhasselt.be
4   Expertise Centre for Digital Media, Flanders Make, Hasselt University, 3500 Hasselt, Belgium
*   Correspondence: zpzhang@csu.edu.cn (Z.Z.); asimpk@gdut.edu.cn (M.A.)

Abstract: In the ever-expanding online fashion market, businesses in the clothing sales sector are presented with substantial growth opportunities. To utilize this potential, it is crucial to implement effective methods for accurately identifying clothing items. This entails a deep understanding of customer preferences, niche markets, tailored sales strategies, and an improved user experience. Artificial intelligence (AI) systems that can recognize and categorize clothing items play a crucial role in achieving these objectives, empowering businesses to boost sales and gain valuable customer insights. However, the challenge lies in accurately classifying diverse attire items in a rapidly evolving fashion landscape. Variations in styles, colors, and patterns make it difficult to consistently categorize clothing. Additionally, the quality of images provided by users varies widely, and background clutter can further complicate the task of accurate classification. Existing systems may struggle to provide the level of accuracy needed to meet customer expectations. To address these challenges, a meticulous dataset preparation process is essential. This includes careful data organization, the application of background removal techniques such as the GrabCut Algorithm, and resizing images for uniformity. The proposed solution involves a hybrid approach, combining the strengths of the ResNet152 and EfficientNetB7 architectures. This fusion of techniques aims to create a classification system capable of reliably distinguishing between various clothing items. The key innovation in this study is the development of a Two-Objective Learning model that leverages the capabilities of both ResNet152 and EfficientNetB7 architectures. This fusion approach enhances the accuracy of clothing item classification. The meticulously prepared dataset serves as the foundation for this model, ensuring that it can handle diverse clothing items effectively. The proposed methodology promises a novel approach to image identification and feature extraction, leading to impressive classification accuracy of 94%, coupled with stability and robustness.

Keywords: ResNet152; EfficientNetB7; attire detection; hybrid learning

1. Introduction

With the rapid advancement of e-commerce, online shopping has surged in popularity. In this evolving landscape, consumers often find themselves inundated with abundant information regarding available clothing options. Simultaneously, retailers assist customers in efficiently selecting the items that best suit their needs. Furthermore, customers exhibit individual preferences, and may adapt their purchases to align with their clothing requirements [1].

Two industries that stand to benefit significantly from using image data, as opposed to an overreliance on manual labeling, are the fashion and online retail sectors. Instead of relying on human annotation for images, it would be advantageous to employ this

technique for the automatic labeling of categories, attributes, and virtually any other data associated with photos used to showcase digital inventory. This shift eliminates individuals' [2] need to engage in the labor-intensive task of physically and mentally categorizing images.

The significance of fashion and clothing in contemporary society has grown considerably, influencing people's daily lives. The fashion industry has become a significant contributor to economic growth, mainly due to the Internet serving as the primary platform for a significant portion of the industry's transactions. Technological advancements have led to the proliferation of new online fashion websites. As mentioned, consumers are increasingly drawn to online shopping because it eliminates spending hours traveling to physical stores searching for desired clothing. With the convenience of Internet shopping platforms, customers can effortlessly browse and purchase their preferred attire from anywhere, using a mobile device or their preferred platform. Online retailers must ensure user-friendly search functionalities and provide high-quality results that meet customers' requirements [3]. However, clothing suppliers need help providing online marketplaces with precise, systematic, and comprehensive product descriptions [4], as it necessitates automated equipment they may not currently possess. Moreover, the online retail landscape is highly diverse, with numerous retailers offering various clothing types and subcategories, each characterized by unique attributes.

In the fashion industry, the concept of "fast fashion", characterized by the efficient production of a diverse range of clothing at lower costs and faster turnaround times, has become increasingly prevalent. To meet this objective, fashion companies strive to offer a broad selection of trendy products promptly. Furthermore, these businesses are leveraging technology to tailor their products to meet the specific preferences of their customers and stay attuned to the latest fashion trends. Customers can easily explore clothing items from their favorite brands, including exclusive offerings. By visualizing these brand-aligned apparel products, consumers can gain a better understanding of how new items might complement their wardrobe, thereby reducing the likelihood of unsatisfactory online purchases.

Additionally, fashion brand recommendation systems have the potential to streamline the shopping experience by minimizing the number of clothing items a shopper needs to browse before making a decision. With the guidance provided, customers can try on a select few outfits, simplifying the final selection process. Image classification is a classic challenge in computer vision, machine learning, and image processing. Deep learning algorithms [5], particularly those used for image and video recognition, have significantly improved accuracy across various domains. Interestingly, deep learning techniques are also being applied in information recommendation, with numerous recommendation systems designed to offer tailored suggestions for information retrieval [6].

Our research endeavors encompass the utilization of a clothing dataset [7] for classification purposes, the formulation of an enhanced hybrid model for accurate attire recognition, the fine-tuning of hyperparameters to optimize model performance, and the subsequent selection of the most effective model for ensemble integration. Recommendation systems [8] have employed a diverse array of methods, including deep learning, machine learning, and image processing, over the years. The primary objective of this study is to initially strategically apply fine-tuning techniques to hyperparameters and subsequently identify the optimal configuration for achieving superior performance in attire recognition tasks.

The contributions of this study are:

- Our methodology presents an improved procedure for developing hybrid models, specifically focused on Two-Objective Learning, which illuminates crucial connections within deep feature correlations through the utilization of transfer learning techniques. Additionally, incorporating ensemble learning in our approach offers the advantage of enhancing model robustness and performance compared to relying on a single model, thereby providing more reliable predictions and insights.

- After initial testing, ResNet152 and EfficientNetB7 outperform other models, making them the preferred backbone models in our proposed methodology for accurately detecting attire items. This superiority is attributed to the combined strengths of ResNet152 and EfficientNetB7 deep features, which elevate performance levels, even without the use of data augmentation techniques.
- In the realm of clothing product dataset recognition, this direct classification approach demonstrates superior performance compared to traditional CNN architectures in terms of accuracy and dependability. However, attire prediction poses challenges due to variations in style, texture, and color, making it difficult for a single model to accurately classify all items. The proposed hybrid learning method offers the advantage of combining the strengths of backbone models to enhance prediction accuracy and robustness in attire recognition tasks.

2. Related Work

With recent technological advancements, customers can now effortlessly stay updated with global fashion trends, significantly impacting their shopping preferences. The fashion industry has increasingly harnessed the power of deep learning and other image analysis technologies for various purposes and applications, as elaborated below. Leading online retailers like Amazon [9] and eBay [10] continuously refine their fashion recommendation engines to better cater to consumer preferences. A multitude of approaches can be employed to study customer purchasing patterns. This article concentrates on deep learning-based Fashion Recommendation System (FRS) [11] initiatives developed between 2016 and 2020. Researchers have constructed recommendation systems using either exclusively deep learning models [12,13] or in combination with other machine learning models. The article provides a concise overview of the most noteworthy deep learning models currently employed in recommendation systems (RS).

In this study [14], four distinct neural network models were employed to classify and distinguish images within the Fashion-MNIST dataset: Fully CNN, CNN, MobileNetV1, and MobileNetV2. The experimental findings indicate that MobileNet outperforms FCNN and CNN by achieving a remarkable accuracy rate of 91% in recognizing apparel images. Remarkably, this updated MobileNet model surpasses its predecessor, which attained a 93% success rate. Despite a comparatively slower training process, the MobileNet model can yield superior accuracy compared to FCNN or CNN.

Author [15] propose an enhanced method for advancing fashion style detection algorithms by utilizing component-dependent Convolutional Neural Networks (CD-CNNs). Our approach begins by gathering images of various body parts and postures, leveraging a combination of pose estimation and semantic segmentation techniques. Subsequently, study employ CD-CNNs to model fashion trends, followed by employing a Support Vector Machine (SVM) in conjunction with CD-CNN outputs to perform data classification. Our methodology effectively enhances classifier accuracy, as evidenced by the results obtained on the Hipster Wars and FashionStyle14 datasets. Specifically, our approach achieves a 76.7% and 85.3% increase in precision compared to conventional techniques, underscoring its suitability for fashion style detection improvement.

In their research, ref. [16] developed a clothing recommendation system that utilizes a programmable embedded device and a single-user photo to streamline the process, thereby saving time and resources. This innovative approach opens up new opportunities for clothing companies and advertising. The study showcases the system's ability to provide clothing recommendations independently of existing customer data, achieved through embedded devices and machine learning techniques. The system employs a feed-forward neural network for recommendations and two inception-based CNNs for predictive tasks. Impressively, the system demonstrates a 98% accuracy rate in color forecasting using AI, an 86% accuracy rate in determining gender and textile patterns, and a 75% accuracy rate in generating outfit recommendations.

Machine learning systems have played a pivotal role in enabling the convenience of online clothing shopping [17]. The effectiveness of these systems heavily relies on the accuracy of clothing data and the proficiency of the machine learning algorithms employed. In our research, we introduce an innovative deep-learning approach incorporating principles from fashion communication theories to enhance our understanding of apparel-related data. Our method employs a two-step training process, using multitask CNN models to predict clothing characteristics from visual inputs. In the subsequent stage, the models delve deeper into the semantic interpretations linked to the identified attributes of garments, employing classifiers such as SVM and LKF. Our experimental findings reveal impressive prediction accuracies across eleven distinct categories of meanings, ranging from 80.1% to 93.50%. By adopting this two-step clothing comprehension technique, our study opens doors to developing novel recommendation algorithms within fashion e-commerce. In this research, the authors introduce an innovative method for classifying fashion product photographs as part of a creative investigation [18].

This research [19] proposes an innovative approach to tackle the human activity recognition (HAR) challenge, employing hybrid learning models. Achieving a classification accuracy of 96.03%, the proposed system utilizes a pretrained XGbBoost model as the classifier, and integrates the Convolutional Variational Autoencoder (CVAE) model as the generator. The methodology [20] of the present study utilized deep learning for feature extraction coupled with machine learning for classification to develop hybrid models. Among the various independent deep-learning models trained on the two datasets examined in the experiments, the transfer-learning model achieved the highest performance on the MNIST dataset, boasting an accuracy rate of 0.9967.

The study employs convolutional neural networks (CNNs) based on deep learning to categorize images sourced from the Fashion-MNIST dataset. To enhance the learning process, the researchers utilized three CNN architectures integrated with batch normalization and residual skip connections. Table 1 presents the literature for detecting attire with machine learning and other computer vision base methodologies.

Table 1. Literature on machine learning techniques to detect attire on different areas.

Reference	Publication Year	Purpose	Method	Accuracy	Precision
Agarwal et al. [1]	2023	Automatic system to save time via manual dress code reading	CNN + YOLOv4	-	81.82%
Yousuf et al. [2]	2019	Address the difficulty of using clothes detection in the actual world.	Modified YOLO and SSD	-	91.14%
Ma et al. [3]	2023	Balance between accuracy and speed	YOLO-FL	91.14%	-

3. Materials and Methods

This study seeks to employ a hybrid learning model based on Transfer Learning to predict attire products, with a specific emphasis on cloth recognition techniques. This section furnishes a thorough overview of the tools and methodologies utilized in our research. It offers a detailed elucidation of the crucial facets of our research methodology. Following this, we delve into the particulars of the preprocessing and augmentation techniques that were utilized. Lastly, we explore the architecture and hyperparameters of the proposed hybrid learning approach under scrutiny.

3.1. Dataset Details

The clothing dataset utilized in our study holds a pivotal role, serving as the foundation for the development and assessment of our hybrid learning model [21] for attire recognition and classification tasks. This dataset comprises an array of clothing items, such as shirts, pants, dresses, shoes, and accessories, reflecting a diverse spectrum of styles, colors, and patterns observed in real-world fashion. Each item in the dataset has been meticulously annotated with its respective category, facilitating supervised learning and precise

model training. Furthermore, it addresses potential challenges encountered in clothing recognition, including variations in pose, lighting conditions, and occlusions. The dataset's substantial size and diversity empower our models to grasp robust and distinctive features, thereby significantly enhancing their accuracy in classifying clothing items. Moreover, our utilization of this dataset is in line with our aim of setting benchmarks and advancing clothing recognition algorithms, thereby contributing to the broader domains of computer vision and fashion-related applications. For our study, attire images were obtained from Github's clothing dataset [22], which encompasses a wide range of clothing products. This publicly accessible dataset served as the cornerstone for constructing our proposed hybrid learning approach. A visual representation of this dataset is depicted in Figure 1, while Table 2 provides the total number of samples for each product category.
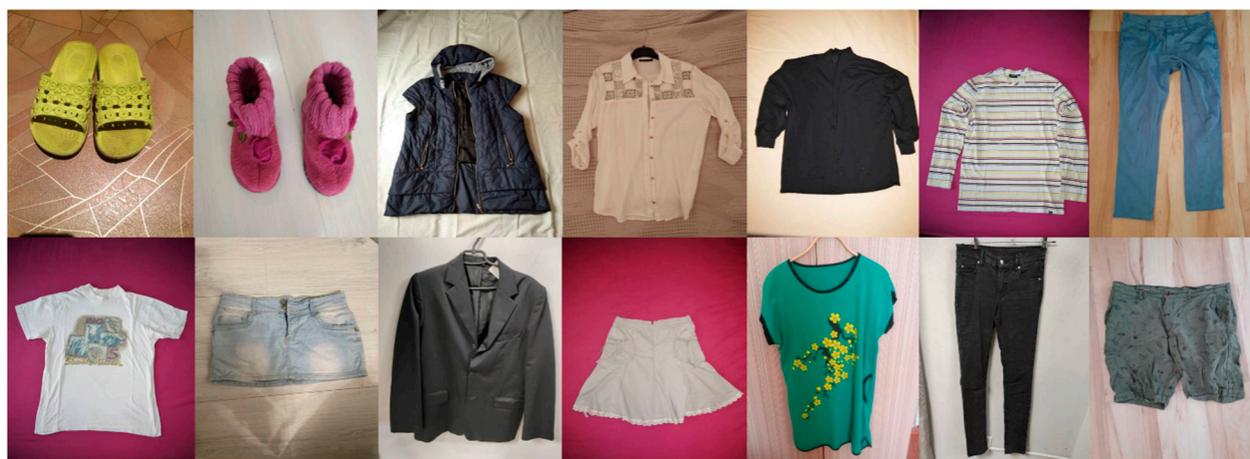

Figure 1. Clothing products sample images.

Table 2. Dataset details.

Product	Samples	Product	Samples
Drees	288	Hat	149
Long Sleeve	576	Out Wear	246
Pants	559	Shirt	345
Shoes	297	Shorts	257
Skirt	136	T-Shirt	928

During the preprocessing and augmentation phase of our study, we implemented a series of crucial steps to prepare our image data for analysis with deep learning. Initially, we standardized the dimensions of all images by resizing them to 224×224 pixels, ensuring consistency and alignment with the input requirements of our deep learning models. Subsequently, to diversify our training dataset and mitigate overfitting, we employed data augmentation techniques. Specifically, we applied horizontal and vertical flips, along with 90-degree rotations, to the images. These augmentation methods introduced variations in the dataset, thereby enhancing the model's capacity to generalize and identify patterns effectively. Such robustness is fundamental for accurate classification tasks. Figure 2 illustrates the details of our preprocessing and augmentation methodology.

In the next stage of our research, following preprocessing and data augmentation procedures, a critical step involved eliminating image backgrounds to extract vital and informative features. To accomplish this, our proposed model utilized the GrabCut Algorithm [23], a robust image segmentation technique. This approach allowed us to accurately separate foreground objects from the background, thereby isolating the regions of interest and retaining the valuable, distinguishing features necessary for our task. Employing the GrabCut Algorithm proved pivotal in improving the quality of our dataset and streamlining

the subsequent phases of our analysis, ultimately contributing to the overall success of our research efforts. Sample images demonstrating the application of the GrabCut algorithm can be found in Figure 3, sourced from Google.

Figure 2. Preprocessing and Augmented Images.

3.2. Proposed Framework

Our research encompassed two separate investigations aimed at identifying the most effective hyperparameters for our proposed model, Two-Objective Learning. Study one and Study Two were dedicated to multiclassification tasks, while Study Three involved reassessing hyperparameters postoptimization. Study one delved into diverse learning rates across nine pretrained models over ten epochs, whereas Study Two investigated varying batch sizes drawn from prior studies on the same nine models. Study Three, in turn, entailed conducting experiments with the best-performing combination of learning rate and batch size.

Once optimal hyperparameters are identified, the model training optimizer becomes crucial. It meticulously analyzes prediction probabilities, pinpointing areas of low accuracy. This analytical insight informs subsequent refinements, ultimately improving the model's performance. A key component of this approach involves crafting a hybrid learning model that combines the strengths of ResNet152 and EfficientNetB7 architectures. By merging these powerful models, the goal is to create a classification system with outstanding precision and recall, adept at distinguishing between images depicting autism and those representing typical cases. A comprehensive evaluation of the model's performance has been conducted, encompassing not just the accuracy of attire prediction but also offering valuable insights for guiding future research directions. The findings from this analysis will guide improvements in dataset preprocessing techniques and the exploration of feature maps, aiming for continual enhancement in predictive accuracy. Embracing the Transfer Learning paradigm for attire prediction, this study seeks to achieve unprecedented accuracy

levels. The results of this research have the potential to significantly advance this critical field. Moreover, the gained insights will pave the way for further enhancements in dataset preprocessing methods and the incorporation of feature maps, ultimately leading to improved accuracy and reliability. Figure 4 presents the proposed hybrid learning architecture.
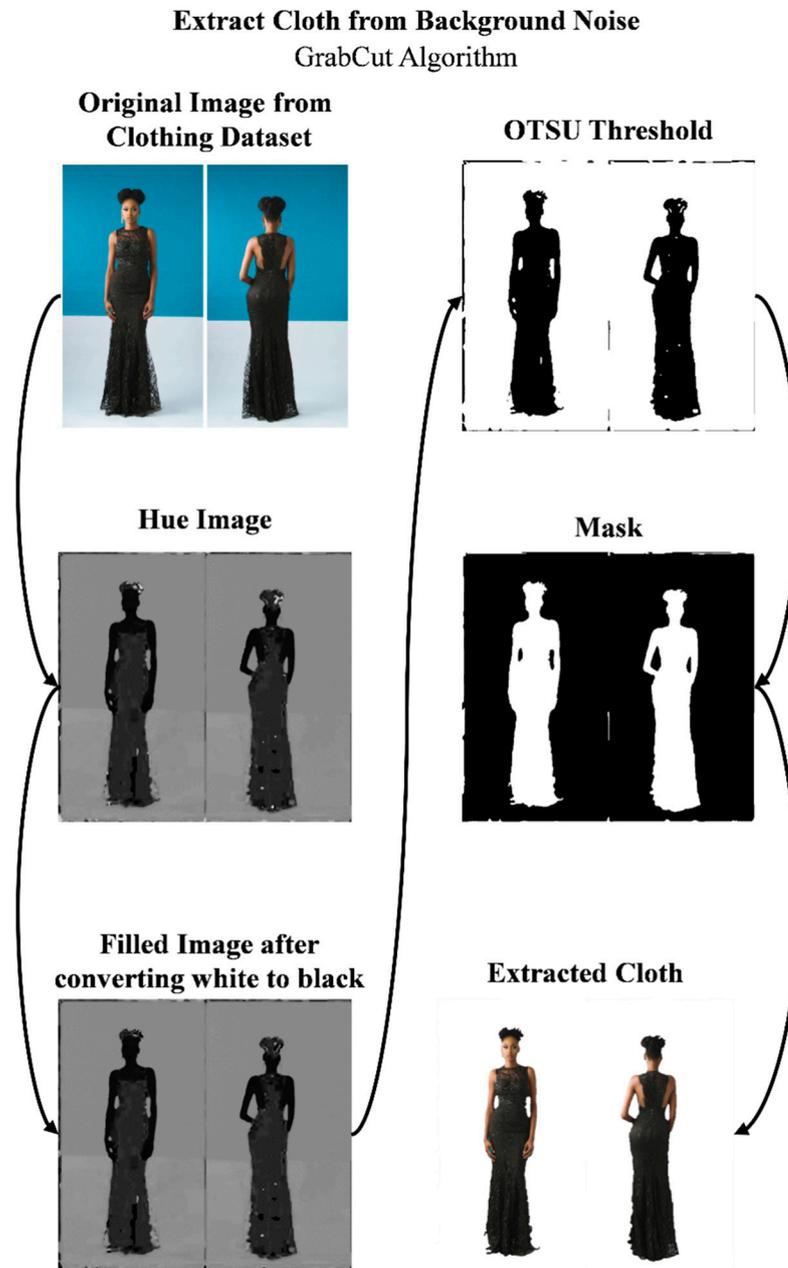
Figure 3. Background remove via GrabCut algorithm.

3.3. Transfer Learning

A pretrained [24] model refers to a neural network that has already undergone training on a large dataset tailored for a particular machine learning task, image classification (IM), or natural language processing (NLP). These models have been trained using substantial computational resources, enabling them to acquire a deep understanding of intricate data patterns and features. The significant advantage of pretrained models lies in their capacity to grasp a comprehensive comprehension of the underlying data patterns, thereby facilitating fine-tuning or utilization as feature extractors for a wide range of subsequent tasks.

Overview of Proposed Methodology

Figure 4. Proposed two-goal learning architecture.

Pretrained models are invaluable as initial building blocks for numerous machine-learning projects. They offer a substantial advantage by diminishing the requirement for training models from the ground up, which can be both computationally intensive and data-demand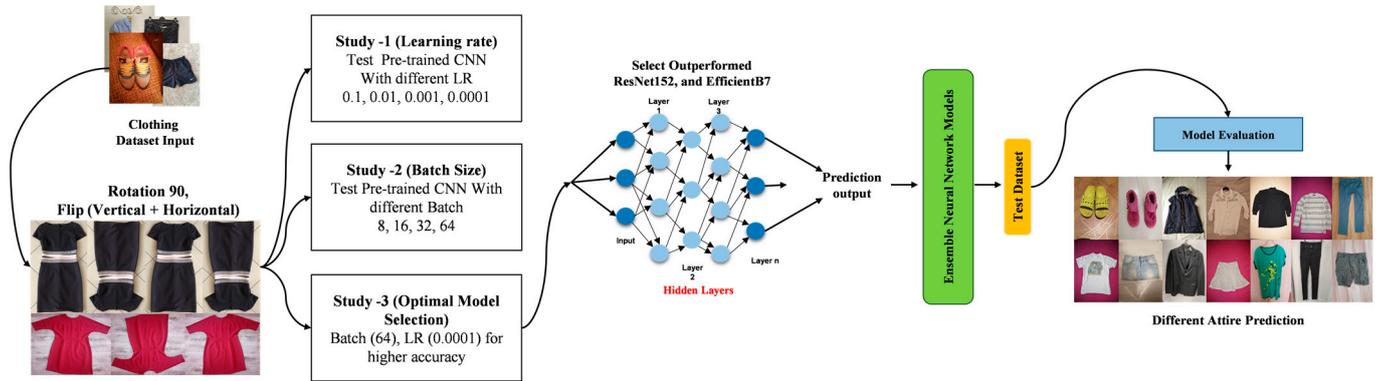ing. Utilizing the insights ingrained within pretrained models, developers and researchers can craft more precise and efficient models tailored to their specific tasks. This knowledge transfer from pretrained models to novel applications is called transfer learning (TL). Transfer learning has become a fundamental pillar of machine learning, facilitating progress across various fields and applications.

3.3.1. ResNet

The ResNet152 [25–27] model is a critical deep-CNN architecture in CV-computer vision. Its inception can be traced back to the seminal research paper where its innovative design first took center stage. ResNet152's significance lies in its ingenious use of skip connections, a strategic choice aimed at effectively tackling the notorious vanishing gradient problem. This issue has long hindered the training of DNN-deep neural networks.

The vanishing gradient problem arises as gradients diminish in magnitude during backpropagation through the layers of deep networks, impeding their convergence during training. ResNet152's distinctive feature is its ability to overcome this challenge. By incorporating skip connections, which allow information to flow more efficiently, it has enabled the training of profound networks. This architectural breakthrough has led to ResNet50 consistently delivering state-of-the-art performance in various image-related applications, solidifying its reputation as a pioneering force in CV. Figure 5 presents basic architecture diagram of ResNet152.

Figure 5. Basic Architecture of ResNet152.
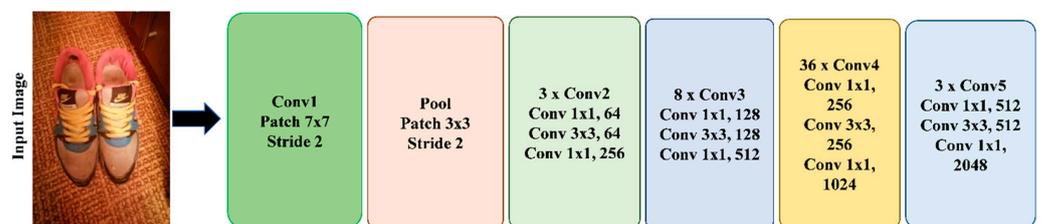
3.3.2. EfficientNetB7

The development of EfficientNet [28], an architectural breakthrough in deep learning, is a remarkable achievement. It is renowned for its unparalleled efficiency and accuracy, showcasing a novel approach to scalability. At its core, EfficientNet leverages compound scaling, a meticulously designed method that optimizes three crucial dimensions of neu-

ral networks: depth, width, and resolution. The architectural blueprint of EfficientNet incorporates a CNN backbone with essential components, including depthwise separable convolutions, squeeze-and-excitation blocks, and efficient scaling techniques. This ingenuity empowers EfficientNet models to excel in various computer vision tasks while maintaining commendable computational efficiency. Thus, EfficientNet seamlessly marries accuracy and computational effectiveness in neural network architectures. EfficientNet stands out as a leading contender in deep learning owing to its remarkable qualities. Its scalability, resulting from a sophisticated interplay between model depth, width, and resolution, highlights its versatility in accommodating diverse computational limitations. Using depthwise separable convolutions, squeeze-and-excitation blocks, and innovative scaling techniques, EfficientNet has showcased its prowess in delivering top-tier performance across various computer vision tasks. It strikes a harmonious balance between achieving high accuracy and maintaining computational efficiency, establishing itself as a valuable asset in contemporary machine-learning endeavors. Figure 6 depicts the fundamental architectural diagram of the EfficientNet model.

Figure 6. Basic Architecture of EfficientNetB7.

4. Results and Implementation

In measuring our models' proficiency in accurately classifying attire products, the models exploit a range of different performance metrics: Accuracy (1), Precision (2), Recall (3), and F1-score (4). These metrics collectively offer a comprehensive evaluation of the model's ability to recognize and categorize attire items effectively.

$$\text{Accuracy} \ = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \tag{1}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

$$\text{F1 Score} \ = 2 \times \frac{\text{Precion} \ \times \ \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

Apart from conventional metrics, employed a confusion matrix to visually illustrate the prediction results for clothing images generated using our proposed approach. A confusion matrix offers valuable insights into the model's performance, enabling a clear

differentiation between true positives, true negatives, false positives, and false negatives in the classifications.

4.1. Experimental Setup

Stacking ensemble technique [13], a sophisticated approach in the realm of machine learning, amalgamates diverse backbase models to enhance predictive performance. The proposed methodology involves training a metalearner, typically a classifier, on the predictions generated by backbone base ResNet152 and EfficientNetB7 models. These base models, each employing distinct feature representations, contribute their individual insights into the dataset, thereby capturing a broader spectrum of patterns and relationships. Through this amalgamation, the stack ensemble model effectively leverages the strengths of each base model while mitigating their weaknesses, resulting in superior predictive accuracy and robustness. By fostering collaboration among diverse models, stacking enables the ensemble to excel in handling complex datasets, adapt to varied data distributions, and ultimately deliver more reliable predictions across a multitude of real-world scenarios.

The hybridization of ResNet152 and EfficientNetB7 models was employed to detect clothing in a dataset. The entire computational process, involving the development of attire product classification, was executed within a Jupiter Notebook environment, utilizing Python. Various methods were explored to accomplish attire detection. Individual classification models were trained using 80% of the dataset as the training set, and their performance was assessed using the remaining 20% as the test set. Both sets were utilized in conjunction with the training data. Several evaluation metrics, including Confusion Matrix visualization, accuracy, sensitivity, specificity, precision, and f1-score, were computed to assess the effectiveness of these classifiers.

4.2. Hyper Parameters Details and Implementation

Table 3 provides a detailed overview of the specifications and configurations of the neural network models utilized in this research study. Model have meticulously tailored two prominent architectures, ResNet152 and EfficientNetB7, for image analysis tasks. These models have undergone training with images resized to a uniform 224 × 224-pixel dimension, ensuring consistency in dataset input. The training process spans ten epochs, reflecting the depth of learning and model refinement achieved during the training phase.

Table 3. Hyper parameter details for Base and proposed models.

Model	Image Size	Epochs	Loss Function	Optimizer	Activation	Learning Rate	Batch Size
ResNet152 (Base)	224 × 224	10	Categorical	Adam	Softmax	0.0001	16
EfficientNetB7 (Base)	224 × 224	10	Categorical	Adam	Softmax	0.0001	16
Two-Goal Learning Model	224 × 224	10	Categorical	Adam	Softmax	0.0001	16

For robust classification performance, the chosen loss function is categorical crossentropy, which quantifies the disparity between predicted and actual category probabilities. Model parameter optimization is accomplished through the highly efficient Adam optimizer. Softmax activation functions are employed to enable the models to make categorical predictions. The learning rate, a crucial hyperparameter governing the rate of parameter updates during training, has been set at 0.0001. This meticulous selection guarantees that the models converge toward optimal solutions without encountering excessive oscillations or divergence during training. Moreover, the batch size, representing the number of input samples processed in each training iteration, is 16. This choice balances computational efficiency and effective gradient-based updates, facilitating efficient parallel processing during training. These carefully chosen configurations constitute an integral part of our

research methodology, ultimately shaping the neural networks' performance in accurately categorizing images and emphasizing their pivotal role in our research endeavor.

4.3. Performance Analysis of Proposed Model

To gauge the overall performance of our trained model on test data, a set of equations that underscore the effectiveness of our innovative Hybrid Learning model. These equations serve as a tangible demonstration of how our proposed model was applied in assessing overall performance. Our approach amalgamated conventional performance metrics, the visual representation provided by the confusion matrix, and the unique attributes of our model. This comprehensive evaluation granted our architecture a profound understanding of the Clothing dataset, revealing the model's remarkable skill in distinguishing between similar attire products. This multifaceted assessment methodology ensures a comprehensive evaluation of recognition accuracy and offers invaluable insights for further refinement and development.

4.4. Classification Performance

In the constant pursuit of enhancing the accuracy and reliability of attire prediction, a rigorous evaluation of model performance is an imperative endeavor. Within Table 4, the capabilities of three models are exploited to assess their efficacy. The baseline models, EfficientNetB7 and ResNet152, are pivotal criteria for comparison. The EfficientNetB7 model, with an accuracy of 0.91, demonstrates remarkable performance, boasting precision, recall, and F1 scores ranging from 0.91 to 0.92. Conversely, the ResNet152 model excels with an accuracy of 0.86, distinguished by notable precision (0.86) and recall (0.86), peaking in an impressive F1 score of 0.86. These baseline models lay a solid foundation for evaluating the proposed hybrid Learning Model. The Hybrid Learning Model emerges as the unequivocal, with an accuracy soaring to 0.94. Beyond its accuracy, this model exhibits improved precision (0.94) and recall (0.94), achieving an equilibrium culminating in a commendable F1 score of 0.94. This outstanding performance underscores the potential of the hybrid learning model to recognize attire prediction, promising a future where recognition accuracy reaches new heights. These meticulous model evaluations corroborate the effectiveness of the proposed model and shed light on its capacity to drive unparalleled advancements in attire prediction.

Table 4. Classification performance of baselines and proposed model.

Models	Evaluations			
	Accuracy	Precision	Recall	F1-Score
EfficientNetB7 (Base)	0.91	0.92	0.91	0.91
ResNet152 (Base)	0.86	0.86	0.86	0.86
Two-Goal Learning Model	0.94	0.94	0.94	0.94

Figure 7 presents the actual and prediction class evaluation on fuse dataset. Our proposed model (Two-Objective Learning) has only 24 incorrect labels out 372 samples. The comparative analysis of learning curves between the baseline model and the proposed (Two-Objective Learning) model, as presented in Figure 8, underscores the enhanced efficiency of the proposed approach. In the case of the baseline model, the training accuracy experiences swift increases after each epoch, yet the loss values exhibit minimal fluctuations. This observation strongly implies that the baseline model is susceptible to overfitting, a phenomenon evident in Figure 9. In contrast, the proposed (Two-Objective Learning) model showcases a commendable convergence pattern, striking a well-balanced equilibrium that mitigates the risks of under and overfitting, highlighting its effectiveness in achieving optimal performance.

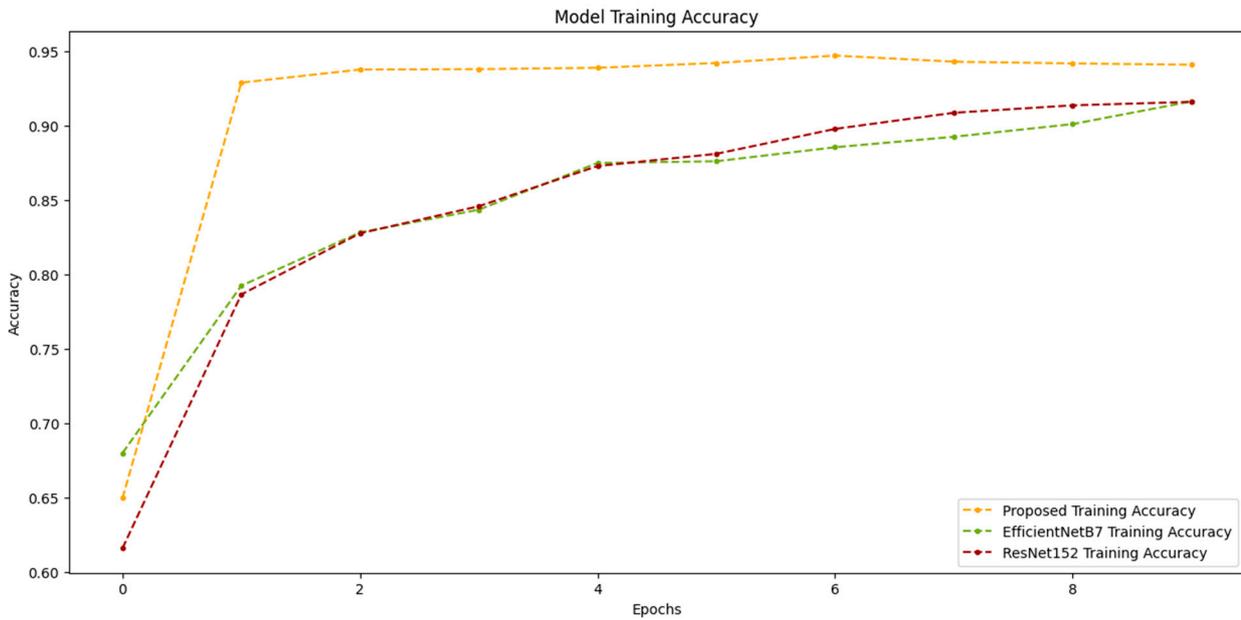**Figure 7.** Proposed model (Two-Objective Learning) confusion matrix analysis.

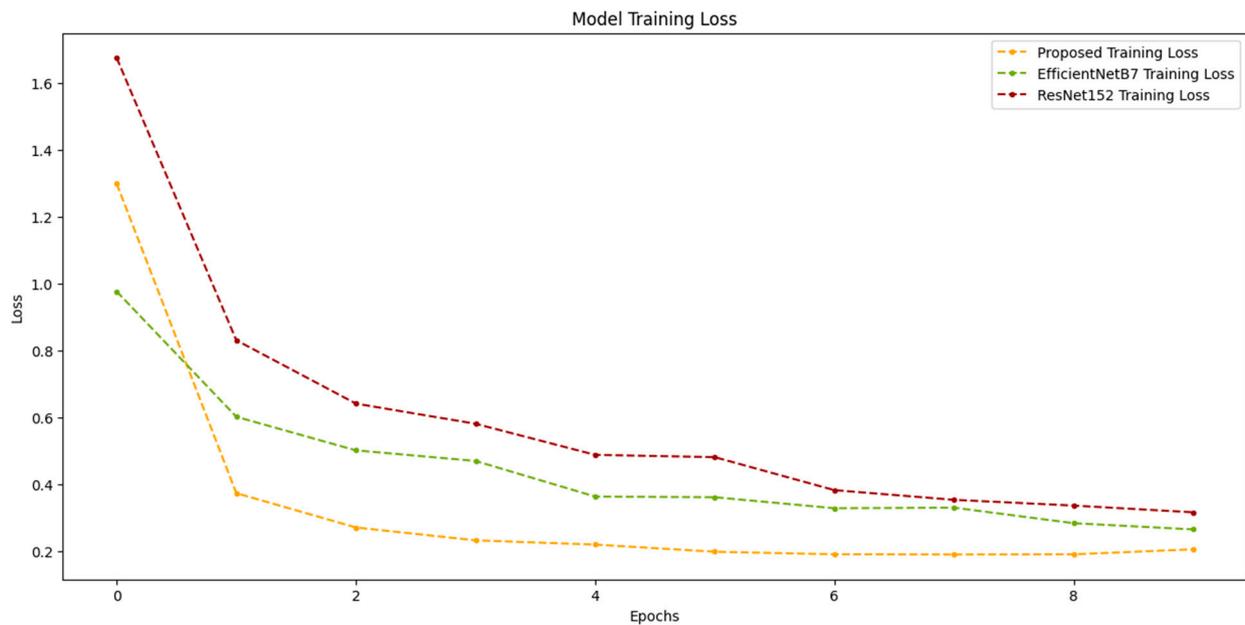**Figure 8.** Baselines model and proposed (Two-Objective Learning) model performance comparison via training accuracy.

Figure 9. Baselines model and proposed (Two-Objective Learning) model performance comparison via training loss.

4.5. Performance Comparison with Some Standard TL Models

The proposed (Two-Objective Learning) learning outperformed all other models on all four-evaluation metrics, with an accuracy of 94%, precision of 94%, recall of 94%, and F1 score of 94%. This indicates that the proposed (Two-Objective Learning) model is able to accurately predict the AD images category in the test set, with very few false positives or false negatives. The EfficientNetB7, and ResNet152 model performed well, with accuracies of 91.3% and 85.7%, respectively. However, they did not achieve the same level of precision, recall, and F1 score as the proposed model. The MobileNet, MobileNetV2, ResNet50, VGG16, and VGG19 models performed less well, with accuracies ranging from 60% to 84%. This suggests that these models are not as well-suited for the classification task at hand. Overall, the results in the Table 5 show that the proposed model is a promising new approach for AD classification. It is able to achieve state-of-the-art results on a challenging test set, outperforming a variety of other popular algorithms.

Table 5. Performance comparison of pretrained model with proposed Two-Objective Learning.

Algorithm	Testing Set			
	A_c	P_r	R_e	F_s
EfficientNetB7 (Base)	91.3%	0.92	0.91	0.91
MobileNet (Base)	66.3%	0.72	0.66	0.68
MobileNetV2 (Base)	60.2%	0.64	0.60	0.61
ResNet50 (Base)	84.9%	0.87	0.85	0.85
ResNet152 (Base)	85.7%	0.86	0.86	0.86
VGG16 (Base)	79.3%	0.81	0.79	0.80
VGG19 (Base)	75.2%	0.81	0.75	0.76
Proposed (Proposed Two-Objective Learning) Model	94%	0.94	0.94	0.94

5. Ablation Study

It is important to clarify that our selection of ResNet152 and EfficientNetB7 as backbone models was not arbitrary but based on extensive experimental results. Through rigorous experimentation and evaluation, we found that these two models consistently outperformed others in terms of various performance metrics, including accuracy, F1-Score,

Precision, and Recall. Therefore, our decision to utilize ResNet152 and EfficientNetB7 as backbone models was grounded in empirical evidence rather than mere preference. This is the purpose of Ablation study.

5.1. Learning Rate Variation—Study One

Study conducted a comprehensive evaluation of various deep learning models Tables 6 and 7, including EfficientNetB7, MobileNet, MobileNetV2, ResNet50, ResNet152, VGG16, and VGG19, to assess their performance on four different testing sets with different learning rates (LR-0.1, LR-0.01, LR-0.001, and LR-0.0001). The evaluation was based on multiple key metrics, including accuracy (A_c), precision (P_r), recall (R_e), and the F1-score (F_s). Notably, EfficientNet Model (B7) exhibited consistently strong performance across learning rates, achieving an accuracy of 87.9% and highest 86.2% on LR-0.0001, respectively. It also demonstrated high precision, recall, and F1-scores, indicating its robustness in classification tasks. In contrast, MobileNet and MobileNetV2 struggled with the lowest accuracy and F1-scores, particularly on the LR-0.01 testing set, underscoring its limitations in this context. These findings provide valuable insights into the suitability of these models for specific learning rates, aiding researchers and practitioners in selecting the most appropriate model for their particular deep learning tasks.

Table 6. Model's performance using Adam optimizer for Study Various Learning Rate (0.1, 0.01), along with ten epochs (Batch Size = 64). A_c—accuracy, P_r—precision, R_e—recall, F_s—F1-score.

Algorithm	Testing Set—LR (0.1)				Testing Set—LR (0.01)			
	A_c	P_r	R_e	F_s	A_c	P_r	R_e	F_s
EfficientNetB7 (Base)	87.9%	0.88	0.88	0.88	82.2%	0.86	0.82	0.82
MobileNet (Base)	67.0%	0.72	0.67	0.69	84.6%	0.85	0.85	0.83
MobileNetV2 (Base)	52.0%	0.59	0.52	0.50	48.1%	0.57	0.48	0.47
ResNet50 (Base)	82.2%	0.84	0.82	0.82	81.9%	0.85	0.82	0.82
ResNet152 (Base)	84.4%	0.86	0.84	0.84	83%	0.85	0.83	0.83
VGG16 (Base)	79.8%	0.84	0.80	0.80	59.6%	0.61	0.60	0.57
VGG19 (Base)	85.2%	0.86	0.85	0.85	76.3%	0.81	0.76	0.76

Table 7. Model's performance using Adam optimizer for Study Various Learning Rate-LR (0.001, 0.0001), along with ten epochs (Batch Size = 64). A_c—accuracy, P_r—precision, R_e—recall, F_s—F1-score.

Algorithm	Testing Set—LR (0.001)				Testing Set—LR (0.0001)			
	A_c	P_r	R_e	F_s	A_c	P_r	R_e	F_s
EfficientNetB7 (Base)	83.6%	0.88	0.84	0.84	86.2%	0.90	0.86	0.87
MobileNet (Base)	65.5%	0.71	0.66	0.67	70.6%	0.72	0.71	0.71
MobileNetV2 (Base)	48.1%	0.57	0.48	0.50	60.2%	0.61	0.60	0.58
ResNet50 (Base)	84.1%	0.85	0.84	0.84	83%	0.84	0.83	0.83
ResNet152 (Base)	84.9%	0.87	0.85	0.85	87%	0.88	0.87	0.87
VGG16 (Base)	81.1%	0.85	0.81	0.82	79.5%	0.81	0.80	0.80
VGG19 (Base)	76.6%	0.80	0.77	0.77	80.3%	0.83	0.80	0.80

5.2. Batch Size Variation—Study Two

In our study next step, thoroughly evaluated the performance of various deep learning models [29–36] on four distinct testing sets Tables 8 and 9: one with a batch size of 8, two with a batch size of 16, three with a batch size of 32 and the other with a batch size of 64 along best learning rate (0.0001). The models included in this assessment were EfficientNetB7, MobileNet, MobileNetV2, ResNet50, ResNet152, VGG16, and VGG19, all serving as base models. Across both testing sets, EfficientNetB0 demonstrated robust classification accuracy, with 89.7% accuracy on BS-8 and 92.2% on BS-16. It exhibited balanced precision, recall, and F1-scores of 0.91, 0.90, and 0.90 on BS-8, and 0.93, 0.92, and 0.92 on BS-16, respectively.

EfficientNetB1 and EfficientNetB7 also displayed competitive results. Conversely, models like MobileNet and MobileNetV2 demonstrated lower accuracy and F1-scores. ResNet50, ResNet152, VGG16, and VGG19 achieved moderate accuracy, with some variation in precision, recall, and F1-scores. These detailed performance metrics provide valuable insights for selecting the most appropriate base model depending on specific batch size requirements and desired classification outcomes in practical applications.

Table 8. Model's performance using Adam optimizer for Study Various Batch Size (8, 16), along with ten epochs (LR = 0.001). A_c—accuracy, P_r—precision, R_e—recall, F_s—F1-score.

Algorithm	Testing Set—Batch Size (8)				Testing Set—Batch Size (16)			
	A_c	P_r	R_e	F_s	A_c	P_r	R_e	F_s
EfficientNetB7 (Base)	87.6%	0.90	0.88	0.88	91.1%	0.92	0.91	0.91
MobileNet (Base)	69.6%	0.71	0.70	0.70	66.3%	0.72	0.66	0.68
MobileNetV2 (Base)	57.2%	0.61	0.57	0.56	60.2%	0.64	0.60	0.61
ResNet50 (Base)	86.2%	0.87	0.86	0.86	84.9%	0.87	0.85	0.85
ResNet152 (Base)	80.9%	0.84	0.81	0.80	85.7%	0.86	0.86	0.86
VGG16 (Base)	79.3%	0.80	0.79	0.79	79.3%	0.81	0.79	0.80
VGG19 (Base)	74.7%	0.78	0.75	0.75	75.2%	0.81	0.75	0.76

Table 9. Model's performance using Adam optimizer for Study Various Batch Size (32, 64), along with ten epochs (LR = 0.0001). A_c—accuracy, P_r—precision, R_e—recall, F_s—F1-score.

Algorithm	Testing Set—Batch Size (32)				Testing Set—Batch Size (64)			
	A_c	P_r	R_e	F_s	A_c	P_r	R_e	F_s
EfficientNetB7 (Base)	87.9%	0.90	0.88	0.88	86.2%	0.90	0.86	0.87
MobileNet (Base)	62.9%	0.67	0.63	0.63	70.6%	0.72	0.71	0.71
MobileNetV2 (Base)	56.9%	0.62	0.57	0.57	60.2%	0.61	0.60	0.58
ResNet50 (Base)	83.8%	0.86	0.84	0.84	83%	0.84	0.83	0.83
ResNet152 (Base)	83%	0.83	0.83	0.82	87%	0.88	0.87	0.87
VGG16 (Base)	81.4%	0.83	0.81	0.82	79.5%	0.81	0.80	0.80
VGG19 (Base)	76.6%	0.82	0.77	0.77	80.3%	0.83	0.80	0.80

In the subsequent stages of our research methodology, meticulously fine-tuned the hyperparameters to optimize model performance. After conducting a series of rigorous experiments, settled on a learning rate of 0.0001, paired with a batch size of 64. This selection was made after carefully considering the trade-offs between convergence speed and model stability. By choosing this specific configuration, study aimed to strike a balance that would facilitate efficient training while preventing issues such as overfitting. These hyperparameter choices reflect our commitment to ensuring the robustness and effectiveness of our approach, a critical aspect of our research in achieving accurate and reliable results.

5.3. Optimal Hyper Parameter—Study Three

Table 10 shows the performance of several deep learning algorithms on a test set with optimal hyper parameter systems, where the performance is evaluated using four metrics: accuracy, precision, recall, and F1 score. Table 10 shows that EfficientNetB7 (Base), which achieve accuracies of 91.3% and 91.1%, respectively. The MobileNet and MobileNetV2 architectures achieve the lowest performance on the test set, with accuracies of 66.3% and 60.2%, respectively. This is likely due to the fact that these architectures are designed to be lightweight and efficient, which comes at the cost of some accuracy. The ResNet50 (Base) and ResNet152 (Base) architectures achieve accuracies of 84.9% and 85.7%, respectively. These architectures are more accurate than the MobileNet and MobileNetV2 architectures, but they are also more computationally expensive. The VGG16 (Base) and VGG19 (Base)

architectures achieve accuracies of 79.3% and 75.2%, respectively. These architectures are the least accurate of the models in the table, but they are also the oldest and most computationally expensive.

Table 10. Model's performance using Adam optimizer for Study Optimal Hyper Parameter Batch Size (16), along with ten epochs (LR = 0.0001). A_c—accuracy, P_r—precision, R_e—recall, F_s—F1-score.

Algorithm	Testing Set			
	A_c	P_r	R_e	F_s
EfficientNetB7 (Base)	91.1%	0.92	0.91	0.91
MobileNet (Base)	66.3%	0.72	0.66	0.68
MobileNetV2 (Base)	60.2%	0.64	0.60	0.61
ResNet50 (Base)	84.9%	0.87	0.85	0.85
ResNet152 (Base)	85.7%	0.86	0.86	0.86
VGG16 (Base)	79.3%	0.81	0.79	0.80
VGG19 (Base)	75.2%	0.81	0.75	0.76

Overall, the Table 10 shows that EfficientNetB7 (Base) is the best performing algorithm on the test set. The MobileNet and MobileNetV2 architectures achieve the lowest performance, likely due to their lightweight design. The ResNet50 (Base) and ResNet152 (Base) architectures are more accurate than the MobileNet and MobileNetV2 architectures, but they are also more computationally expensive. The VGG16 (Base) and VGG19 (Base) architectures are the least accurate of the models in the table, but they are also the oldest and most computationally expensive.

6. Discussion

The research represents a significant advancement in the domain of attire product prediction with CNN models to explore the advantages of clothing recognition techniques. The hybrid approach employed is particularly noteworthy in term of higher performance. This dual-goal learning model aims to enhance precision and recall, critical areas in clothing recognition where traditional methods may fall short. By integrating two architectures, the research aims to create a model that can accurately distinguish between various clothing images. Furthermore, this research serves as a foundation for future improvements and adaptations is crucial. The comprehensive performance analysis conducted during the investigation provides valuable insights for refining dataset preparation methods and exploring clothing features. This iterative process ensures that the dependability and accuracy of the clothing recognition model will progressively improve over time. Overall, this study of ongoing progress in the field of clothing identification research. By embracing CNN paradigms, the findings of this study are poised to significantly advance the field and inspire further developments in manufacturing and AI-driven recognition technologies. As time progresses, the potential for swift intervention and support in clothing recognition becomes increasingly feasible, offering user-friendly solutions across various production industries.

Limitation in this research stemming from this study could delve deeper into the integration of real-time data streams and adaptive learning algorithms to enhance the model's performance and scalability. Additionally, exploring the application of this technology in diverse contexts via question answering platform [33], such as retail, fashion design, and personalized styling services, could unlock new avenues for innovation and practical implementation.

7. Conclusions

This research signifies a substantial step forward in pursuing precise attire product prediction. Leveraging Transfer Learning (TL) and Deep Learning (DL) models, the study embarked on a journey to unlock the potential of clothing recognition techniques in attire products. The results and insights obtained throughout this study hold the promise of sav-

ing considerable time and expediting the attire recognition process. A fundamental aspect of our research lies in the meticulous dataset preprocessing. Model meticulously organized and resized images to establish consistent dimensions, ensuring the reliability and uniformity of the dataset. This preprocessing phase not only facilitated an efficient model training, but also set the stage for the subsequent development of a robust classification system.

The dual-goal learning model being developed within this project represents a hybrid approach, combining elements from the ResNet152 and EfficientNetB7 architectures. Our primary focus lies in enhancing precision and recall, critical areas where this novel approach aims to excel. By integrating these architectures, our objective was to create a model that accurately distinguishes between various clothing photos.

It is important to note that our research serves as a foundation for future improvements and adaptations rather than a conclusion. A comprehensive performance analysis conducted during the investigation has provided invaluable insights, guiding the continuous enhancement of dataset preparation methods and the exploration of clothing features. The dependability and accuracy of our clothing recognition model will progressively improve with each iteration. In its entirety, this study encapsulates the spirit of ongoing progress, paving the way for unprecedented enhancements in prediction accuracy. By embracing the TL paradigm to predict clothing for individuals of all ages and genders, anticipate that the findings of this study will significantly advance the field of clothing identification research and inspire further developments in manufacturing and AI-driven recognition technologies. As time progresses, the prospect of swift intervention and support in clothing recognition becomes increasingly feasible, offering a user-friendly and straightforward solution across various production industries.

Author Contributions: Conceptualization, W.A., Z.Z. and M.A.; methodology, W.A. and M.A.; software, W.A.; validation, W.A., M.A., J.C. and S.A.; formal analysis, W.A. and M.A.; investigation, J.C. and S.A.; resources, Z.Z., J.C. and S.A.; data curation, W.A., Z.Z. and M.A.; writing—original draft preparation, W.A. and M.A.; writing—review and editing, Z.Z., M.A., J.C. and S.A.; visualization, W.A. and M.A.; supervision, Z.Z.; project administration, Z.Z.; funding acquisition, M.A. and S.A. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data can be shared upon reasonable request from corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Agarwal, A.; Das, A. Facial Gestures-Based Recommender System for Evaluating Online Classes. In *Recommender Systems*; CRC Press: Boca Raton, FL, USA, 2023; pp. 173–189.
2. Yousuf, S.B.; Sajid, H.; Poon, S.; Khushi, M. IMDB-Attire: A Novel Dataset for Attire Detection and Localization. In Proceedings of the Neural Information Processing: 26th International Conference, ICONIP 2019, Sydney, NSW, Australia, 12–15 December 2019; Proceedings, Part II 26; Springer International Publishing: New York, NY, USA, 2019.
3. Ma, W.; Guan, Z.; Wang, X.; Yang, C.; Cao, J. YOLO-FL: A target detection algorithm for reflective clothing wearing inspection. *Displays* **2023**, *80*, 102561. [CrossRef]
4. Park, Y.E. Research evidence for reshaping global energy strategy based on trend-based approach of big data analytics in the corona era. *Energy Strategy Rev.* **2022**, *41*, 100835. [CrossRef]
5. Amin, M.S.; Wang, C.; Jabeen, S. Fashion sub-categories and attributes prediction model using deep learning. *Vis. Comput.* **2023**, *39*, 3851–3864. [CrossRef]
6. Cosma, A.C.; Simha, R. Machine learning method for real-time non-invasive prediction of individual thermal preference in transient conditions. *Build. Environ.* **2019**, *148*, 372–383. [CrossRef]
7. Liu, K.H.; Chen, T.Y.; Chen, C.S. Mvc: A dataset for view-invariant clothing retrieval and attribute prediction. In Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval, New York, NY, USA, 6–9 June 2016.

8. Gharaei, N.Y.; Dadkhah, C.; Daryoush, L. Content-based clothing recommender system using deep neural network. In Proceedings of the 2021 26th International Computer Conference, Computer Society of Iran (CSICC), Tehran, Iran, 3–4 March 2021.

9. Guan, C.; Qin, S.; Ling, W.; Ding, G. Apparel recommendation system evolution: An empirical review. *Int. J. Cloth. Sci. Technol.* **2016**, *28*, 854–879. [CrossRef]

10. Schafer, J.B.; Konstan, J.; Riedl, J. Recommender systems in e-commerce. In Proceedings of the 1st ACM conference on Electronic Commerce, Denver, CO, USA, 3–5 November 1999.

11. Sulthana, R. A review on the literature of fashion recommender system using deep learning. *Int. J. Perform. Eng.* **2021**, *17*, 695.

12. Raza, A.; Meeran, M.T.; Bilhaj, U. Enhancing Breast Cancer Detection through Thermal Imaging and Customized 2D CNN Classifiers. *VFAST Trans. Softw. Eng.* **2023**, *11*, 80–92.

13. Khan, S.U.R.; Asif, S.; Bilal, O.; Ali, S. Deep hybrid model for Mpox disease diagnosis from skin lesion images. *Int. J. Imaging Syst. Technol.* **2024**, *34*, e23044. [CrossRef]

14. Jain, K. Transfer Learning-Based Machine Learning Approach to Solve Problems of E-commerce: Image Search. In *International Joint Conference on Advances in Computational Intelligence*; Springer: Singapore, 2022.

15. Yamamoto, T.; Nakazawa, A. Fashion style recognition using component-dependent convolutional neural networks. In Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 22–25 September 2019.

16. Villar-Martinez, A.; Rodriguez-Gil, L.; Angulo, I.; Orduña, P.; García-Zubía, J.; López-De-Ipiña, D. Improving the scalability and replicability of embedded systems remote laboratories through a cost-effective architecture. *IEEE Access* **2019**, *7*, 164164–164185. [CrossRef]

17. Pereira, A.M.; Moura JA, B.; Costa ED, B.; Vieira, T.; Landim, A.R.; Bazaki, E.; Wanick, V. Customer models for artificial intelligence-based decision support in fashion online retail supply chains. *Decis. Support Syst.* **2022**, *158*, 113795. [CrossRef]

18. Xhaferra, E.; Cina, E.; Toti, L. Classification of Standard FASHION MNIST Dataset Using Deep Learning Based CNN Algorithms. In Proceedings of the 2022 International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Turkey, 20–22 October 2022.

19. Tseng, Y.H.; Wen, C.Y. Hybrid Learning Models for IMU-Based HAR with Feature Analysis and Data Correction. *Sensors* **2023**, *23*, 7802. [CrossRef]

20. Albattah, W.; Albahli, S. Intelligent arabic handwriting recognition using different standalone and hybrid CNN architectures. *Appl. Sci.* **2022**, *12*, 10155. [CrossRef]

21. Sakib, S.; Fahad, N.M.; Raiaan, M.A.K.; Rahman, M.A.; Al Mamun, A.; Islam, S.; Mukta, M.S.H. Predicting gender from human or non-human social media profile photos by using transfer learning. In Proceedings of the 2023 International Conference on Computer, Electrical & Communication Engineering (ICCECE), Kolkata, India, 20–21 January 2023.

22. Ye, S.; Wang, S.; Chen, N.; Xu, A.; Shi, X. OMNet: Outfit Memory Net for clothing parsing. *Int. J. Cloth. Sci. Technol.* **2023**, *35*, 493–505. [CrossRef]

23. Shi, H.; Zhao, D. License Plate Localization in Complex Environments Based on Improved GrabCut Algorithm. *IEEE Access* **2022**, *10*, 88495–88503. [CrossRef]

24. Wang, H.; Li, J.; Wu, H.; Hovy, E.; Sun, Y. Pre-trained language models and their applications. *Engineering* **2022**, *25*, 51–65. [CrossRef]

25. Zhang, L.; Li, H.; Zhu, R.; Du, P. An infrared and visible image fusion algorithm based on ResNet-152. *Multimed. Tools Appl.* **2022**, *81*, 9277–9287. [CrossRef]

26. Khan, S.U.R.; Zhao, M.; Asif, S.; Chen, X.; Zhu, Y. GLNET: Global–local CNN's-based informed model for detection of breast cancer categories from histopathological slides. *J. Supercomput.* **2023**, *80*, 7316–7348. [CrossRef]

27. Khan, S.U.R.; Zhao, M.; Asif, S.; Chen, X. Hybrid-NET: A fusion of DenseNet169 and advanced machine learning classifiers for enhanced brain tumor diagnosis. *Int. J. Imaging Syst. Technol.* **2024**, *34*, e22975. [CrossRef]

28. Tan, M.; Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In Proceedings of the International Conference on Machine Learning, Long Beach, CA, USA, 9–15 June 2019.

29. Torrey, L.; Shavlik, J. Transfer learning. In *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*; IGI Global: Hershey, PA, USA, 2010; pp. 242–264.

30. Zhuang, F.; Qi, Z.; Duan, K.; Xi, D.; Zhu, Y.; Zhu, H.; Xiong, H.; He, Q. A comprehensive survey on transfer learning. *Proc. IEEE* **2020**, *109*, 43–76. [CrossRef]

31. Niu, S.; Liu, Y.; Wang, J.; Song, H. A decade survey of transfer learning (2010–2020). *IEEE Trans. Artif. Intell.* **2020**, *1*, 151–166. [CrossRef]

32. Pan, S.J.; Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **2009**, *22*, 1345–1359. [CrossRef]

33. Farooq, M.U.; Beg, M.O. Bigdata analysis of stack overflow for energy consumption of android framework. In Proceedings of the 2019 International Conference on Innovative Computing (ICIC), Lahore, Pakistan, 1–2 November 2019; pp. 1–9.

34. Khan, S.U.R.; Raza, A.; Waqas, M.; Zia, M.A.R. Efficient and Accurate Image Classification Via Spatial Pyramid Matching and SURF Sparse Coding. *Lahore Garrison Univ. Res. J. Comput. Sci. Inf. Technol.* **2023**, *7*, 10–23.

35. Mohsen, S.; Ali, A.M.; Emam, A. Automatic modulation recognition using CNN deep learning models. *Multimed. Tools Appl.* **2023**, *83*, 7035–7056. [CrossRef]

36. Boulila, W.; Khlifi, M.K.; Ammar, A.; Koubaa, A.; Benjdira, B.; Farah, I.R. A Hybrid Privacy-Preserving Deep Learning Approach for Object Classification in Very High-Resolution Satellite Images. *Remote Sens.* **2022**, *14*, 4631. [CrossRef]