Improving DeepFashion Dataset Classification Accuracy with StyleGAN2-ADA: Addressing Imbalanced Data in Fashion Image Recognition*

Agnė Moleikaitytė[1,*,†] and Jonas Paulauskas[2,3,†]

[1] *Kaunas University of Technology, Faculty of Informatics*
[2] *Kaunas University of Technology, Artificial Intelligence Excellence Centre*
[3] *ISM University of Management and Economics*

Abstract

Fashion classification remains challenging due to the high variability in apparel styles and the significant class imbalance in fashion datasets. This study utilizes the DeepFashion dataset, which consists of 289,222 images across 46 clothing categories, to evaluate the impact of data preprocessing, augmentation, and synthetic data generation on classification performance. To address class imbalance, underrepresented categories were merged with similar classes or supplemented with additional training data. Two dataset expansion techniques were explored: traditional augmentation (flipping, rotation, color jittering, and Gaussian blur) and synthetic image generation using StyleGAN2-ADA. The YOLO11 model was employed for classification, and its performance was assessed using top-k accuracy metrics. Experimental results show that preprocessing alone improved classification accuracy, while adding synthetic data further enhanced model performance, achieving a top-3 accuracy of 93.44% and a top-5 accuracy of 97.34%, surpassing previous state-of-the-art results. Analysis of the confusion matrix revealed that while synthetic data helped mitigate class imbalance, misclassifications persisted among visually similar categories. These findings highlight the potential of generative models in enriching training datasets and improving classification performance. However, further feature extraction and inter-class discrimination refinements are necessary for optimal results.

Keywords

DeepFashion, StyleGAN2-ADA, Data Augmentation, Imbalanced Dataset, Fashion Image Recognition

1. Introduction

Fashion image classification is a complex task in computer vision due to the high variability of clothing styles, textures, patterns, and visually similar apparel categories. Deep learning-based models have advanced the field by leveraging large-scale annotated datasets such as DeepFashion [1], which has served as a benchmark for apparel classification, attribute recognition, and retrieval. However, a fundamental challenge persists – class imbalance in fashion datasets. Many real-world fashion datasets exhibit a long-tailed distribution where dominant categories contain thousands of samples, while rare categories have minimal examples [1-2] . This imbalance can lead to biased learning, where the model becomes overconfident in well-represented classes while struggling to generalize to underrepresented categories [3-6]. Improving classification accuracy for such datasets can enhance the robustness of clothing digitalization tasks, including inventory management, automated uploading of clothing data to e-commerce platforms, personalization of product recommendations for end users.

Previous research efforts have focused on improving feature extraction [7], leveraging landmark

* *IVUS2025: Information Society and University Studies 2025, May 15, Kaunas, Lithuania*
* Corresponding author.
† These authors contributed equally.
agne.moleikaityte@ktu.edu (A. Moleikaitytė); jonas.paulauskas@ktu.lt (J. Paulauskas)
0009-0008-2279-0847 (A. Moleikaitytė);

detection [8], and integrating attention mechanisms [9] to enhance classification accuracy. However, relatively less emphasis has been placed on tackling class imbalance directly. Traditional solutions such as data augmentation and re-weighting loss functions [10] partially mitigate the issue, but they often fail to generate sufficient diversity for minority classes [6]. To overcome these limitations, generative adversarial networks (GANs) have emerged as a promising solution for synthetic data generation [11-14]. In particular, StyleGAN2-ADA [15], an advanced GAN model with adaptive discriminator augmentation, has demonstrated the ability to synthesize high-fidelity images even when trained on limited data [16-18].

This study proposes a novel approach to address class imbalance in fashion classification by generating synthetic clothing images using StyleGAN2-ADA. Unlike traditional augmentation techniques, GAN-generated images provide entirely new training samples rather than simple transformations of existing data [19]. In this paper, we focus on investigating whether the expansion of under-represented classes using GAN methods improves classification accuracy more than traditional expansion methods.

The key contributions of this work include a comparative analysis of traditional augmentation and StyleGAN2-ADA-based synthetic data generation for class imbalance, quantitative evaluation on DeepFashion with accuracy results and confusion matrix insights and qualitative analysis of GAN image quality and intra-class confusion redistribution.

2. Related Work

Fashion classification has been a widely studied topic in computer vision, with advancements driven by the availability of large-scale datasets and deep-learning architectures. This section reviews key contributions that have influenced fashion image recognition, including early dataset-driven methods, attention-based models, and landmark-independent approaches.

The DeepFashion dataset has been instrumental in advancing fashion classification and retrieval tasks. It provides 289,222 images across 46 clothing categories, annotated with fine-grained attributes and landmark information. Liu *et al.* introduced FashionNet, a deep learning model which simultaneously predicts clothing attributes and landmark positions, improving recognition accuracy by incorporating spatial relationships between apparel components [1]. This dataset enabled future research by enabling robust feature learning for apparel recognition. Subsequent works leveraged attention mechanisms and landmark detection to refine classification performance. Wang *et al.* [8] Introduced an Attentive Fashion Grammar Network that incorporates two grammar models: a dependency grammar to capture relationships between body parts and a symmetry grammar to enforce bilateral symmetry in garment layout. By using these fashion grammars to guide the model's attention, their approach improved the accuracy of landmark detection and category classification. In contrast, Lee *et al.* [10] developed a landmark-free approach using a two-branch feature selective network. This model performed multi-task learning, extracting global and fine-grained attributes simultaneously. Their model reduced errors caused by inaccurate landmark detection by eliminating explicit landmark dependence and improving generalization across diverse clothing styles. Zhang *et al.* [7] addressed the bias of convolutional networks toward texture over shape by proposing a two-stream network. Their model incorporated a texture-biased branch leveraging pre-trained CNNs and a shape-biased branch using landmark-based attention to capture both fine textures and garment structures. This hybrid approach achieved better performance in clothing attribute recognition by balancing texture- and shape-based feature learning.

More recent studies have moved towards semi-supervised learning and landmark-free models to improve classification accuracy while reducing reliance on expensive annotations. Shajini *et al.* [9] introduced a landmark-driven attention network, integrating spatial and channel-wise attention mechanisms to emphasize discriminative clothing regions. Their approach significantly improved category recognition by highlighting key apparel features, reducing misclassifications among visually similar garments. Shajini *et. al.* [20] proposed a semi-supervised approach, introducing a multi-stage feature-attentive network to enable knowledge sharing between labeled and unlabeled images. Their model utilized a teacher-student framework, reducing dependence on labeled data while maintaining high classification accuracy. Despite these advancements, class imbalance remains an issue in fashion datasets. Most prior works focused on architectural improvements, such as attention modules, grammar-based constraints, and feature selection techniques, but few directly address the imbalanced nature of datasets. While traditional augmentation strategies [10] have been applied to increase sample diversity, they merely alter existing images rather than introduce new training examples. Generative approaches, particularly GANs offer a promising alternative by synthesizing novel samples [21]. StyleGAN2-ADA [22] has been successfully applied in other domains, such as face generation and medical imaging, demonstrating high image quality even with limited training data [15-18]. The aim of this study is to investigate StyleGAN2-ADA [15] as a solution for class imbalance in fashion datasets, evaluating its impact on classification accuracy and comparing it against conventional augmentation techniques, establishing a new benchmark for DeepFashion dataset classification accuracy.

3. Methodology

3.1. Dataset

The DeepFashion dataset has been selected for classification tasks in this study. It consists of 289,222 images across 50 clothing categories. However, only 46 categories have associated images, while four categories lack related images. The dataset is split into 209,222 images for training (72.4%), 40,000 images for validation (13.8%), and 40,000 images for testing (13.8%). Each image in the dataset is annotated with a bounding box that defines the clothing region. Although the dataset provides additional metadata such as landmarks, clothing attributes, and pairwise relationships, these annotations were not utilized in this study. Instead, the dataset was only used for category-level classification, where each image is assigned to one of the 46 available clothing classes. Despite its large scale, this dataset presents class imbalance challenges – some categories contain tens of thousands of samples, whereas others have only a few hundred or fewer. To mitigate class imbalance, data augmentation, and synthetic image generation were applied in this research, creating a more balanced and robust training dataset.

3.2. Dataset pre-processing

To standardize the dataset, each image was cropped using the provided clothing bounding box details and resized to 256 × 256 pixels, ensuring that the longer side was resized to 256 pixels, with padding applied to the shorter side to maintain the aspect ratio.

A dataset analysis was performed to identify severely underrepresented classes. The dataset was preprocessed by removing or merging extremely rare categories into larger ones where possible. Before processing, class sizes ranged from Dress (72,158 images) at 24.95% to Coverup (17 images) at 0.006%. After processing, the Dress class accounted for 25.12%, while the smallest class, Sweatshorts, accounted for 0.38% **Figure 1.**

To further balance the dataset, two new training datasets have been created. These expansions targeted categories with fewer than 10,000 images (3.48% of the dataset). The two datasets are:

- **Augmented Dataset**: expanded using traditional augmentation techniques, including: a) random vertical and horizontal flips; b) rotation; c) color jittering and c) Gaussian blur.
- **Generated dataset**: expanded using StyleGAN2-ADA, synthesizing additional images for underrepresented categories.

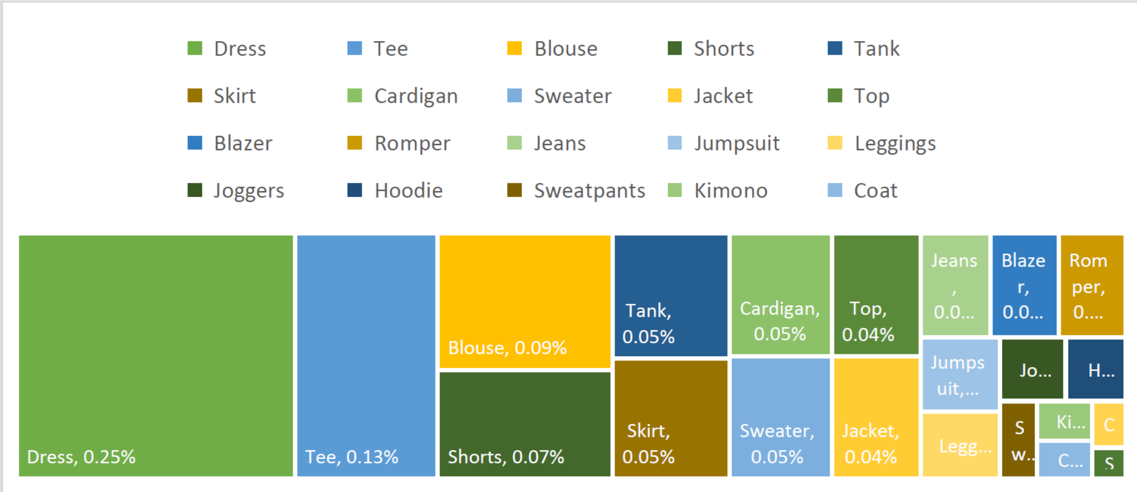


Figure 1: Distribution of training dataset classes after preprocessing.

Following the dataset expansion through augmentation and synthetic data generation, the distribution of training dataset classes became more balanced, as illustrated in Figure 2. The previously underrepresented categories, which constituted less than 3.48% of the dataset, now exhibit a more uniform distribution, reducing the dominance of majority classes like Dress (previously over 24%) and Tee, which previously had a higher representation. The highest category now accounts for 20.38%, with several smaller classes maintaining a more equitable presence around 2.82%. This redistribution ensures that minority categories, such as Sweatshorts, Kimono, and Cutoffs, receive sufficient representation during training, preventing model bias toward majority classes. The new dataset structure enhances generalization, improves classification accuracy across all categories, and mitigates the risk of overfitting to dominant apparel types.

The StyleGAN2-ADA model was fine-tuned using 5,000,000 images per class. During training, stabilization was observed after approximately 2 million images were processed per class, as seen in Figure 3. This suggests that early-stage training introduced sufficient diversity, capturing essential variations in clothing attributes. Further training iterations primarily refined fine details without introducing significant new variations. This aligns with findings that GAN-generated diversity tends to plateau after a certain number of iterations, necessitating a balance between quality and novelty [23-24].

Visual inspections were conducted to validate the quality and realism of synthetic images. The generated clothing items were compared against real images to assess structural consistency and attribute retention, ensuring that minor classes were accurately synthesized.

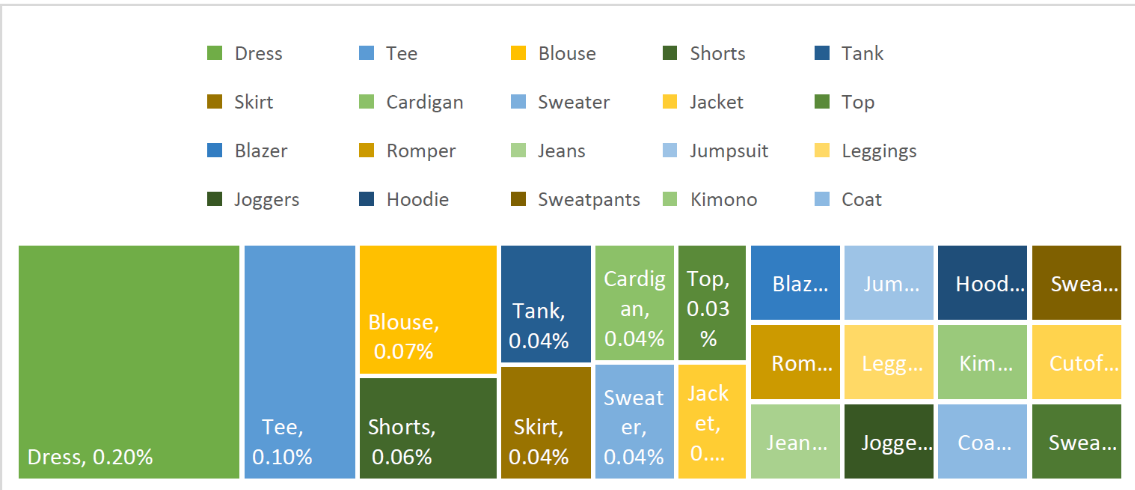


Figure 2: Distribution of training dataset classes after balancing.

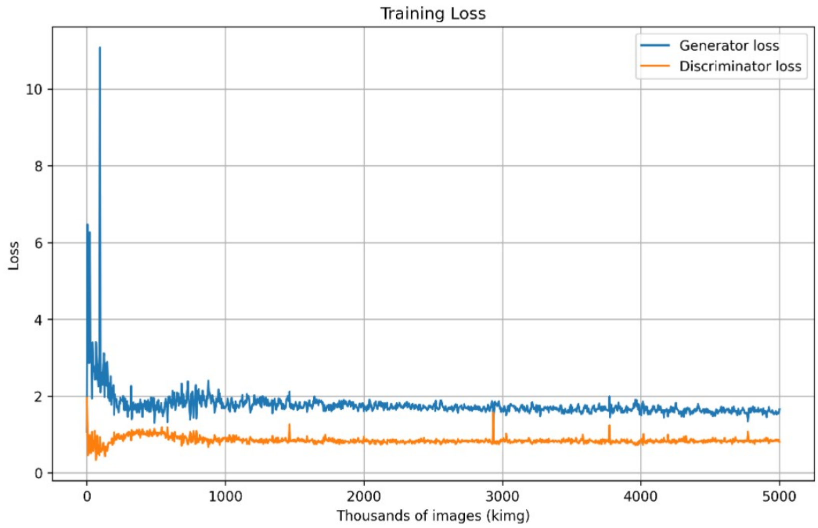


Figure 3: Generator and Discriminator training loss using 5000 images data set. time

4. Experimental Results

4.1. Evaluation Metrics

The best-k metrics were used to evaluate classification accuracy because they provide a standardized measure of model performance and are widely used in fashion classification benchmarks. This approach is particularly useful in multi-class classification tasks, where models must distinguish between visually similar clothing categories.

Several prior works in fashion image recognition, including Liu *et al.* [1], Lee *et al.* [10], Wang *et al.* [8], Shajini *et al.* [9,20], Zhang *et al.* [7], have also adopted top-k accuracy as their primary evaluation metric, ensuring comparability between models. These studies demonstrate that fine-grained clothing classification benefits from top-k evaluation, as many apparel items share overlapping characteristics, making a strict top-1 accuracy metric insufficient. Our results remain directly comparable to existing state-of-the-art methods by aligning our evaluation with these prior works. This accuracy is defined as

$$Acc = \frac{1}{N} \sum_{i=1}^{N} 1\left(y_i \in \hat{Y}_i^K\right), \tag{1}$$

Where N is the total number of samples, y_i is the true label for i-th sample, $\hat{Y}_i^K$ the set of top K predicted labels for the i-th sample and $1(\cdot)$ is the indicator function, which is 1 if the true label is within the top K predicted labels, and 0 otherwise.

4.2. Results

Multiple experiments were conducted using YOLO11 [25] model to evaluate with which data classification model performs best. Comparative results are displayed in the Table 1. YOLO11 model performs better than other methods in top-3 and top-5 accuracy. The best results are obtained when generated training data is incorporated. Our proposed solution surpassed previous state-of-the-art methods.

The models were trained on NVIDIA A100-PCIE-40GB GPU with a batch size of 256, which resulted in 25GB of GPU memory usage during the training. Training the model on original and processed classes data took around 11 hours, while training the model on expanded dataset (with generated or augmented data) took around 13 hours, resulting in 2 hours increase in training time. Despite the increased training time, the model maintained efficient performance while classifying the images – it takes around 280 milliseconds for the model to classify one image.

Initially, when YOLO11 was evaluated across all classes without modifications, it achieved a top-3 accuracy of 92.59% and a top-5 accuracy of 96.60%. By refining the dataset through class merging and removal, the model saw an increase in accuracy, reaching 93.36% (top-3) and 97.26% (top-5).

Further enhancing the training process, generated data is introduced, resulting in a slight boost to 93.44% (top-3) and 97.34% (top-5). This indicates that synthetic data can positively impact the learning process, offering more diverse examples and reinforcing model generalization. However, when training with augmented data, the model performed slightly worse than with generated data but still outperformed most prior methods, achieving 92.72% (top-3) and 96.94% (top-5). This suggests that while augmentation provides variations of existing data, generated data might introduce more meaningful new patterns.

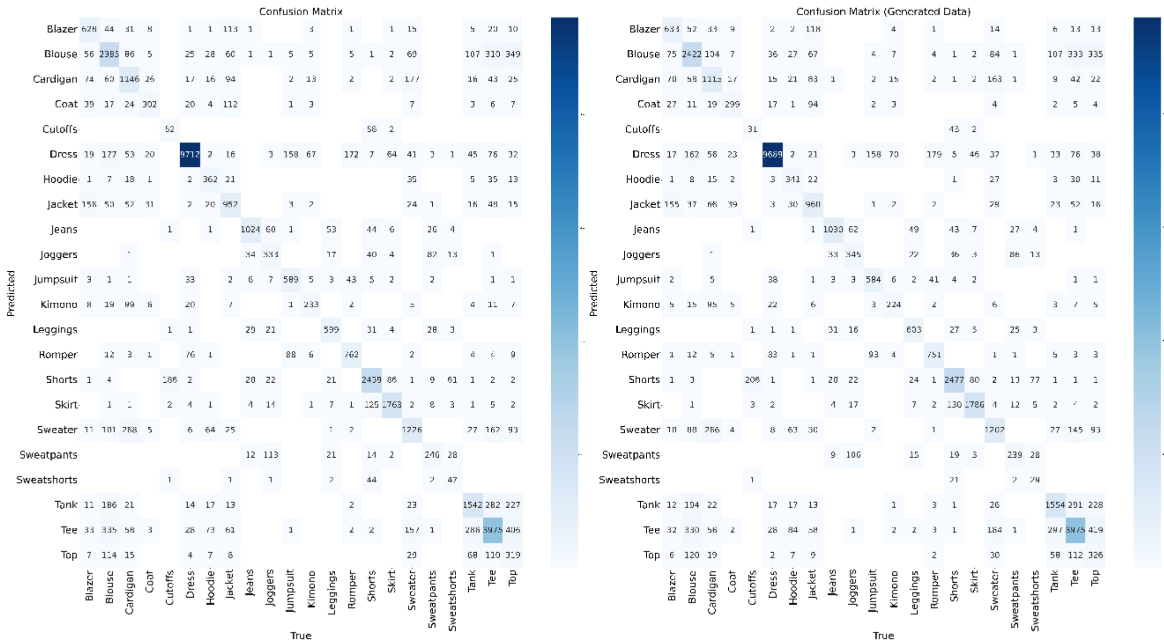
Table 1
Results of category classification using top-k accuracy

Method	Top-3 (%)	Top-5 (%)
Liu *et al.* [1]	82.58	90.17
Lee *et al.* [10]	91.37	95.26
Wang *et al.* [8]	90.99	95.78
Shajini *et al.* [9]	91.02	96.2
Shajini *et al.* [20]	91.06	96.35
Zhang *et al.* [7]	91.99	96.44
YOLO11 all classes	92.59	96.60
YOLO11 processed classes	93.36	97.26
YOLO11 with generated data	**93.44**	**97.34**
YOLO11 with augmented data	92.72	96.94

When comparing the confusion matrix of a model trained using a processed classes dataset and the matrix of a model trained using generated data (Figure 4), we observe notable shifts in classification performance, particularly in the distribution of misclassifications across visually similar categories. While the overall classification accuracy of dominant classes, such as Dress and Tee, remained relatively stable, there were observable redistributions in misclassified instances among semantically related categories. Notably, Blouse exhibited an increased positive detections (+37). Similarly, Jacket, Shorts, Skirt, Tank and Top classes exhibited increase in their positive detections.

A key trend was the reallocation of correct classifications in smaller classes, with Cutoffs experiencing a substantial reduction in true positive detections (-21), predominantly shifting toward Shorts (+20). Likewise, Cardigan exhibited a decline in true positive classifications (-31), with slight increases in misclassifications into Jacket (+14), implying an increased intra-class confusion within layered outerwear. Upon further analysis of the minority classes, additional notable shifts were identified. Blazer experienced a minor improvement in true positive detections (+5), but with an increased misclassification into Blouse (+19). Romper saw a drastic decline in true positive classifications (-11), with increased confusion into Dress (+7). Jeans exhibited a minor increase in correct classifications (+6), while Joggers displayed a significant gain in true positives (+12), reducing their confusion with Jumpsuit, Leggings and Sweatpants. Jumpsuit and Sweatpants saw reductions in correct classifications (-5 and -7, respectively). Leggings showed a slight increase in true positive detections (+4), but also had increased misclassification into Joggers (+5). Jumpsuit's drop in accuracy (-5) came with an increased misclassification into Romper. Hoodie experienced a substantial drop in true positive detections (-21), with misclassifications primarily shifting to Tee (+11) and Jacket (+10). Sweatpants also declined in classification accuracy (-7), while Sweatshorts exhibited a major decrease in true positives (-18), gaining some reallocation into Shorts (+16). Coat displayed a minor reduction in true positive detections (-3), with some misclassifications shifting toward Jacket.

These shifts in performance across minority classes indicate that the updated model exhibits increased confusion within similar visual categories while improving classification in certain structured items. The redistribution of misclassifications suggests that while some classes benefited from the revised training approach, others, particularly those with highly overlapping visual features, suffered from greater inter-class ambiguity. Future refinements may require targeted augmentation strategies or more distinct feature representations to mitigate these confusions and improve classification robustness across all categories, particularly for underrepresented classes.



(a) (b)

Figure 4: (a) confusion matrix of the model that is trained with processed dataset, (b) confusion matrix of the model that is trained with expanded generated images dataset.

Incorporating generated data into the model training improved top-k accuracy metrics, with top-3 accuracy increasing from 93.36% to 93.44% and top-5 accuracy improving from 97.26% to 97.34%. These results suggest that including synthetic data contributes positively to the model's overall generalization, with minimal impact on major classification trends. The stability of large-class accuracy combined with the observed shifts in misclassifications within smaller classes indicates that while synthetic data may enhance general robustness, it does not fully resolve category ambiguities, particularly among visually similar apparel items.

5. Discussion: practical applications and real-world impact

The improved classification system developed in this study provides practical benefits for several use cases in the fashion and digital retail industry. E-commerce platforms often rely on automatic tagging and categorisation of garments to simplify product uploading and inventory management. However, rare or niche categories of clothing are often misclassified due to class imbalances, reducing the accuracy of product metadata and affecting search and recommendation performance. By improving the classification accuracy of under-represented classes through GAN-based synthetic data generation, our approach improves the system's ability to accurately tag a wider range of garment types. As a result, users benefit from more robust filtering and sorting capabilities, better product search capabilities and higher quality recommendations. In addition, visual search systems used in mobile shopping apps or digital wardrobes can benefit from more reliable category detection, especially for items that are not well represented in training datasets.

In addition to retail, such classification systems can also support fashion analysis, trend forecasting and digital fashion design platforms that require accurate input categorisation for further modelling.

6. Conclusion

This study explored the impact of dataset preprocessing, augmentation, and synthetic data generation on fashion classification performance using the DeepFashion dataset. By refining the dataset through class merging and removing underrepresented categories, the YOLO11 model demonstrated increased classification accuracy. Incorporating generated data further improved performance, with top-3 accuracy reaching 93.44% and top-5 accuracy at 97.34%. These findings highlight the effectiveness of synthetic data in enhancing model generalization, offering diverse training examples. The results confirm the hypothesis that synthetic data generation with StyleGAN2-ADA improves classification accuracy in imbalanced dataset. However, while the overall accuracy improved, confusion matrix analysis revealed persistent misclassifications among visually similar categories. This suggests that synthetic data strengthens overall robustness but does not completely resolve inter-class ambiguities. Future research could explore alternative generative approaches or additional fine-tuning strategies to differentiate closely related apparel categories better and further optimize model performance.

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

[1] Z. Liu, P. Luo, S. Qiu, X. Wang, and X. Tang, "DeepFashion: Powering Robust Clothes Recognition and Retrieval with Rich Annotations," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, Dec. 2016, pp. 1096–1104. doi: 10.1109/CVPR.2016.124.

[2] S. Sadegharmaki, M. A. Kastner, and S. Shin'ichi, "FashionGraph: Understanding fashion data using scene graph generation," *Proceedings - International Conference on Pattern Recognition*, pp. 7923–7929, 2020, doi: 10.1109/ICPR48806.2021.9412662.

[3] C. Huang, Y. Li, C. C. Loy, and X. Tang, "Learning deep representation for imbalanced classification," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, Dec. 2016, pp. 5375–5384. doi: 10.1109/CVPR.2016.580.

[4] W. Ouyang, X. Wang, C. Zhang, and X. Yang, "Factors in Finetuning Deep Model for object detection," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 2, pp. 864–873, 2016, doi: 10.48550/arXiv.1601.05150 Focus to learn more.

[5] G. Van Horn and P. Perona, "The Devil is in the Tails: Fine-grained Classification in the Wild," Sep. 2017, doi: 10.48550/arXiv.1709.01450 Focus to learn more.

[6] J. Kozerawski, V. Fragoso, N. Karianakis, G. Mittal, M. Turk, and M. Chen, "BLT: Balancing Long-Tailed Datasets with Adversarially-Perturbed Images," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Springer Science and Business Media Deutschland GmbH, 2021, pp. 338–355. doi: 10.1007/978-3-030-69535-4_21.

[7] Y. Zhang, P. Zhang, C. Yuan, and Z. Wang, "Texture and shape biased two-stream networks for clothing classification and attribute recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, 2020, pp. 13535–13544. doi: 10.1109/CVPR42600.2020.01355.

[8] W. Wang, Y. Xu, J. Shen, and S. C. Zhu, "Attentive Fashion Grammar Network for Fashion

Landmark Detection and Clothing Category Classification," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, Dec. 2018, pp. 4271–4280. doi: 10.1109/CVPR.2018.00449.

[9] M. Shajini and A. Ramanan, "An improved landmark-driven and spatial–channel attentive convolutional neural network for fashion clothes classification," *Visual Computer*, vol. 37, pp. 1517–1526, Jun. 2021, doi: 10.1007/s00371-020-01885-7.

[10] S. Lee, H. Eun, S. Oh, W. Kim, C. Jung, and C. Kim, "Landmark-free clothes recognition with a two-branch feature selective network," *Electron Lett*, vol. 55, pp. 745–747, Jun. 2019, doi: 10.1049/el.2019.0660.

[11] R. Sauber-Cole and T. M. Khoshgoftaar, "The use of generative adversarial networks to alleviate class imbalance in tabular data: a survey," *J Big Data*, vol. 9, Dec. 2022, doi: 10.1186/s40537-022-00648-6.

[12] S. Cusworth, G. V. Gkoutos, and A. Acharjee, "A novel generative adversarial networks modelling for the class imbalance problem in high dimensional omics data," *BMC Med Inform Decis Mak*, vol. 24, Dec. 2024, doi: 10.1186/s12911-024-02487-2.

[13] N. Chereddy and B. K. Bolla, "Evaluating the Utility of GAN Generated Synthetic Tabular Data for Class Balancing and Low Resource Settings," Jun. 2023, doi: 10.1007/978-3-031-36402-0_4.

[14] P. Yadav, G. Sihag, and V. Vijay, "Rebalancing the Scales: A Systematic Mapping Study of Generative Adversarial Networks (GANs) in Addressing Data Imbalance," Feb. 2025.

[15] T. Karras, M. Aittala, J. Hellsten, S. Laine, J. Lehtinen, and T. Aila, "Training generative adversarial networks with limited data," in *Advances in Neural Information Processing Systems*, Neural information processing systems foundation, 2020.

[16] J. Joo *et al.*, "Enhancing automated strabismus classification with limited data: Data augmentation using StyleGAN2-ADA," *PLoS One*, vol. 19, May 2024, doi: 10.1371/journal.pone.0303355.

[17] A. Mohanty, A. Sutherland, M. Bezbradica, and H. Javidnia, "High-Fidelity Synthetic Face Generation for Rosacea Skin Condition from Limited Data," *Electronics (Switzerland)*, vol. 13, Jan. 2024, doi: 10.3390/electronics13020395.

[18] M. Woodland *et al.*, "Evaluating the Performance of StyleGAN2-ADA on Medical Images," Oct. 2022, doi: 10.1007/978-3-031-16980-9_14.

[19] R. Chen *et al.*, "Decoupled Data Augmentation for Improving Image Classification," Oct. 2024, doi: 10.48550/arXiv.2411.02592.

[20] M. Shajini and A. Ramanan, "A knowledge-sharing semi-supervised approach for fashion clothes classification and attribute prediction," *Visual Computer*, vol. 38, pp. 3551–3561, Nov. 2022, doi: 10.1007/s00371-021-02178-3.

[21] S. Ayoub, Y. Gulzar, J. Rustamov, A. Jabbari, F. A. Reegu, and S. Turaev, "Adversarial Approaches to Tackle Imbalanced Data in Machine Learning," *Sustainability 2023, Vol. 15, Page 7097*, vol. 15, no. 9, p. 7097, Apr. 2023, doi: 10.3390/SU15097097.

[22] "GitHub - NVlabs/stylegan2-ada-pytorch: StyleGAN2-ADA - Official PyTorch implementation," https://github.com/NVlabs/stylegan2-ada-pytorch.

[23] J. Dubiński, K. Deja, S. Wenzel, P. Rokita, and T. Trzcinski, "Selectively Increasing the Diversity of GAN-Generated Samples," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Springer Science and Business Media Deutschland GmbH, 2023, pp. 260–270. doi: 10.1007/978-3-031-30105-6_22.

[24] S. Gurumurthy, R. K. Sarvadevabhatla, and R. V. Babu, "DeLiGAN : Generative adversarial networks for diverse and limited data," in *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, Institute of Electrical and Electronics Engineers Inc., Nov. 2017, pp. 4941–4949. doi: 10.1109/CVPR.2017.525.

[25] R. Khanam and M. Hussain, "YOLOv11: An Overview of the Key Architectural Enhancements," Oct. 2024, doi: 10.48550/arXiv.2410.17725.