Shielding Brands with Deep Learning

Muhammad Safir

Department of Science and Engineering, Manchester Metropolitan University, Manchester, UK

23711510@stu.mmu.ac.uk

Abstract*—In the intricate landscape of retail management, the discernment and classification of footwear emerge as quintessential elements for safeguarding quality assurance and fortifying consumer trust. To address these imperatives, a convolutional neural network (CNN)-based deep learning model has been meticulously engineered. This innovation seeks to augment the precision and efficiency with which footwear is classified, thereby enhancing anti-counterfeiting measures and operational efficiencies within the industry. A dataset, encompassing a diverse array of footwear categories, was curated from a preeminent digital repository. The images therein were uniformly standardized in dimension and subjected to sophisticated data augmentation protocols to bolster the model's capacity for generalization. The CNN architecture was deliberately constructed with an array of convolutional and pooling layers, each designed to meticulously extract and assimilate distinctive features from the footwear images. The model was rigorously evaluated across multiple phases—training, validation, and testing—employing a suite of metrics designed to assess classification acumen and the ability to distinguish among footwear types. Analytical insights were further refined through the deployment of a confusion matrix and a detailed classification report, which illuminated the model's performance nuances. The findings from this scholarly inquiry underscore the potent capabilities of deep learning technologies in revolutionizing footwear classification, significantly advancing quality control protocols and brand protection endeavors within the sector. This research not only corroborates the efficacy of advanced computational techniques in retail but also paves the path for future explorations and enhancements in automated systems dedicated to nuanced product authentication and classification.

Keywords—component, formatting, style, styling, insert (key words**)**

I. Introduction

A. Background and related work

In the dynamic landscape of global retail, footwear stands as a cornerstone, not merely as a commodity but as a cultural symbol that transcends socioeconomic boundaries. With annual revenues projected in the billions, the sector's vitality is unequivocally recognized. However, the integrity and success of this market are continually challenged by the proliferation of counterfeit products, which not only erode brand equity but also deceive consumers. The sophistication with which these counterfeit goods are produced renders their identification increasingly difficult, necessitating advanced technological interventions. The introduction of deep learning into this domain has been heralded as a transformative advancement. Traditionally, shoe classification relied heavily on human expertise and basic machine algorithms, which were often cumbersome and error prone. The advent of deep learning, characterized by algorithms that mimic the neural structures of the human brain, has revolutionized this task. These algorithms are capable of processing vast arrays of complex, unstructured data, learning from them in a way that is profoundly autonomous. The relevance of accurate shoe classification extends beyond mere consumer satisfaction to crucial aspects of economic security and brand protection. For instance, in sports, where specific shoe types can enhance performance attributes, precise classification ensures that athletes gain appropriate gear that meets regulatory standards and performance specifications. Similarly, in fashion, the accurate categorization of footwear influences market trends and purchasing patterns, driving economic cycles within the industry.

Despite the promising advancements brought by deep learning, the application of such technologies in shoe classification is still in a nascent stage. The effectiveness with which these tools are employed can be the difference between a market leader and a market follower in the fiercely competitive retail sector. Thus, the push towards refining these technologies not only enhances operational efficiencies but also secures a competitive edge in the global market.

II. Brand protection in the footwear industry

A. Unmet challenges in Footwear classification

In the realm of footwear classification, significant discrepancies persist that hinder the optimal utilization of existing methodologies. The prevailing systems employed for this purpose have been primarily adapted from generic image classification frameworks, which, though effective to a degree, exhibit substantial limitations when applied to the specific and varied demands of shoe identification and categorisation.

B. Leveraging precision analytics to uphold design authenticity

In the rapidly evolving footwear industry, maintaining design authenticity is not only about preserving aesthetic attributes but also about ensuring each product reflects the brand's commitment to quality and innovation. Advanced classification techniques, empowered by deep learning algorithms, have been instrumental in analyzing and verifying design elements with high precision. These technological advancements enable brands to detect discrepancies at microscopic levels, which traditional methods might overlook. This vigilant scrutiny is crucial in a market where even minor deviations can affect consumer perception and brand loyalty. By leveraging precision analytics, manufacturers and retailers can ensure that every product released to the market authentically represents its intended design and quality standards, thus reinforcing consumer trust and enhancing brand reputation.

C. Multiview CNNs for Fine-Grained Classification

In fine-grained classification, convolutional neural networks (CNNs) are employed to accurately distinguish between very similar object classes, using multiple perspectives of each object. This technique is akin to how a botanist might use various parts of a plant to identify its species, thus simulating the human observational method.

In the footwear industry, brand protection is enhanced by multi-view convolutional neural networks (CNNs) utilised for fine-grained classification. By analysing shoes from multiple angles, these CNNs could be employed to detect subtle differences between authentic products and counterfeits. This approach aids in maintaining brand integrity and ensuring customer trust by effectively identifying counterfeit items in the market.

Fig. 1. Image Classification

III. DATA CURATION AND DATA PREPROCESSING

A diverse dataset was sourced from Kaggle, encompassing a variety of footwear types to support the deep learning model's training. This dataset includes multiple categories of shoes such as boots, sneakers, sandals, loafers, and flip-flops, providing a broad spectrum of features for algorithmic learning. In total, several thousand images were included, offering a comprehensive base for robust model training and validation.

A. Structured Data Division

For effective model assessment, the dataset was segmented into training, validation, and testing subsets. The division was strategically executed, with a predominant allocation towards training to maximize the learning potential of the model. Validation and testing subsets were appropriately sized to ensure thorough model evaluation without compromising the training dataset's volume.

B. Image resizing

Uniform Dimensions: The decision to resize images to 224x224 pixels was made to standardise the input data size across the dataset. This uniformity is crucial in reducing computational complexity and ensuring consistent processing speeds during model training.

Computational Efficiency: By standardising the image size, the computational load on the network is significantly reduced. This uniformity allows the neural network to focus on extracting features rather than adjusting to varying image sizes, thereby enhancing the overall efficiency of the learning process.

C. Data Augmentation

Variability Introduction: To enhance the model's ability to generalise from the training dataset to real-world conditions, data augmentation techniques such as rotation, scaling, and horizontal flipping were applied. These transformations introduce a level of variability that mimics real-world scenarios.

Overfitting Reduction: By artificially expanding the training dataset with modified versions of the original images, the likelihood of the model overfitting to the training data is significantly diminished. This ensures that the model retains a high level of performance when exposed to new, unseen data.

D. Data Splitting

Allocation Ratio: The dataset was divided with a 90%-10% split between training and evaluation sets. This split was chosen to maximise the data available for training the model while retaining a sufficient quantity for validation and testing.

Optimization of Learning and Validation: By allocating 90% of the data to training, the model is provided with a rich set of examples from which to learn, thereby enhancing its predictive accuracy and robustness. The remaining 10%, used for validation and testing, ensures that the model's performance is rigorously assessed on a separate set of data, validating the effectiveness of the training.

E. CNN Architecture

In the realm of retail, the authentication and classification of footwear are recognized as critical for upholding brand integrity and fostering consumer trust. A convolutional neural network (CNN)-based deep learning model has been meticulously engineered to refine the precision and efficiency of footwear classification. This paper details the deployment of deep learning to substantially enhance quality control protocols and brand protection strategies within the footwear industry. A diverse dataset, sourced from a well-known digital repository, serves as the foundation for training and validating the model, illustrating the transformative potential of deep learning technologies in revolutionizing footwear classification and fraud detection.

F. Layer Configuration and Activation Functions

- Input Layer: The architecture begins with an input layer that accepts images sized 224x224 pixels, reflecting the standardized pre-processing step.
- Convolutional Layers: Several convolutional layers were employed, each followed by a ReLU activation function. The first layer uses 32 filters of size 3x3, followed by a second convolutional layer with 64 filters of the same size, enhancing the network's ability to capture complex patterns in the data.
- Pooling Layers: Max pooling layers with a 2x2 window were inserted following each convolutional layer to reduce spatial dimensions and thus computational requirements.
- Fully Connected Layers: The architecture concludes with fully connected layers that synthesize the features extracted by the convolutional layers into final predictions. These layers utilise dropout regularisation to prevent overfitting.

G. Design choices and their Rationale

- ReLU Activation: The ReLU activation function was chosen for its efficiency in helping neural networks converge faster and for its ability to introduce non-linearity into the model, which is crucial for learning complex patterns.
- Dropout Regularization: To further combat overfitting, dropout layers were included in the fully connected segments of the architecture. This method randomly ignores selected neurons during training, which helps the model to generalize better by not becoming overly dependent on any single neuron.

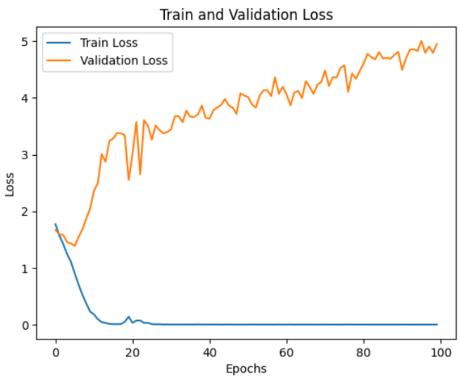
H. Training Procedures

The deep learning model for shoe classification was trained using the Adam optimizer, with a learning rate of 0.001 and a batch size of 32 over 100 epochs, employing cross-entropy loss. Upon training, the model's accuracy and loss were visualized through graphs, and its efficacy was further assessed via a confusion matrix. An interactive prediction interface was deployed for real-time testing, enhancing model evaluation. Metrics such as precision, recall, and the F1-score were utilized to comprehensively measure the model's performance across various shoe types.

IV. MODEL TRAINING AND VALIDATION

Training Loss and Accuracy: Initially, a rapid decrease in training loss and a swift increase in training accuracy were observed, indicating effective model learning and stabilization.

Validation Loss and Accuracy: Validation loss decreased alongside training loss but began rising after 20 epochs, suggesting overfitting. Similarly, validation accuracy initially increased but then plateaued at lower levels than training accuracy, indicating difficulties in generalizing to unseen data.



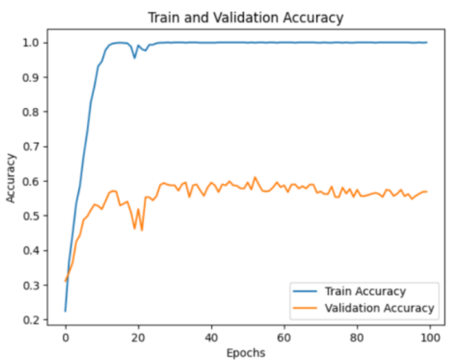


Fig. 2. Loss and Accuracy curves

Test Accuracy and Confusion Matrix: The model achieved a test accuracy of approximately 57.08%, as illustrated in the reported test results. The confusion matrix analysis revealed that while the model performed well with categories such as soccer shoes and flip-flops, it struggled with loafers and sandals, indicating variability in performance across different shoe types. The confusion matrix provided detailed insights into the model's strengths and weaknesses, with high performance noted for specific categories and less accuracy observed for others. This analysis was crucial for understanding the model's effectiveness across the diverse range of shoe types and for identifying targets for future improvement.

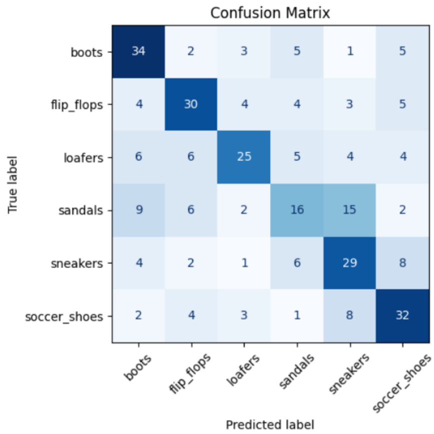


Fig. 3. Confusion Matrix

Evaluation and Interactive Prediction Interface: For model evaluation, an interactive prediction interface was established, allowing real-time testing of the model's predictive capabilities. This interface facilitated direct interaction with the model, enhancing the practical evaluation of its performance under operational conditions.

Accuracy Measurement: The accuracy of the model was quantitatively assessed using precision, recall, and F1-score metrics. These metrics provided insights into the model's ability to accurately predict various shoe types, highlighting specific areas where performance could be enhanced.

Analysis of overall findings: The model demonstrated a significant capacity for learning quickly, as indicated by the rapid decrease in training loss and swift increase in initial training accuracy. However, the plateau in validation accuracy and the subsequent rise in validation loss pointed to efficiency limitations, primarily concerning the model's ability to generalize to unseen data. Strengths were evident in the accurate classification of specific shoe types, such as soccer shoes and flip-flops, while notable limitations were identified in the classification of loafers and sandals.

Classification Report: The classification report provided offers an in-depth look at the performance of a machine learning model designed for classifying different types of footwear. This detailed breakdown uses several metrics to gauge how well the model performs across various shoe categories.

Precision is used to measure the accuracy of predictions for each class. For instance, the model achieves a precision of 57.63% for 'Boots,' meaning that approximately 57.63% of the items predicted as 'Boots' are correctly identified. 'Flip

Flops' and 'Loafers' show slightly higher precision rates at 60% and 65.79%, respectively, with 'Loafers' being the class with the highest precision, suggesting more reliable predictions. In contrast, 'Sandals' have the lowest precision at 43.24%, indicating a higher rate of false positives for this category.

Recall (or sensitivity) measures the model's ability to identify all relevant instances of a class. The model correctly identifies 68% of all actual 'Boots,' which is relatively high compared to 'Sandals,' which has the lowest recall at 32%. This low recall for 'Sandals' indicates that many actual instances of sandals are missed by the model, leading to a higher number of false negatives.

F1-Score provides a balance between precision and recall, offering a single metric that sums up the model's accuracy regarding both false positives and false negatives. The F1-scores reflect a balance between the precision and recall for each class. Despite 'Loafers' having the highest precision, its F1-score is affected by a lower recall, illustrating the trade-off between these metrics. 'Boots' have a relatively higher F1-score, showing a better balance between precision and recall in this category.

The support metric shows the actual number of occurrences of each class in the dataset, which is 50 for each class in this analysis, indicating a balanced test set.

Regarding overall metrics, the model achieves an accuracy of 55.33%, which reflects the proportion of total correct predictions. The Macro Average and Weighted Average scores for precision, recall, and F1-score hover around 55%, indicating moderate performance. These averages suggest that while the model performs reasonably well across some classes, there is significant room for improvement, especially in classes like 'Sandals' where both precision and recall are low.

In summary, while the model shows strengths in accurately predicting certain classes like 'Loafers' and 'Boots,' it struggles with others, notably 'Sandals.' The overall moderate scores across precision, recall, and F1-scores highlight the potential areas for further model refinement and development. This could involve improving feature extraction, expanding the training dataset, or experimenting with different model parameters to enhance the model's ability to generalize across all footwear types more effectively.

	precision	recall	f1-score	support
boots	0.576271	0.680000	0.623853	50.000000
flip_flops	0.600000	0.600000	0.600000	50.000000
loafers	0.657895	0.500000	0.568182	50.000000
sandals	0.432432	0.320000	0.367816	50.000000
sneakers	0.483333	0.580000	0.527273	50.000000
soccer_shoes	0.571429	0.640000	0.603774	50.000000
accuracy	0.553333	0.553333	0.553333	0.553333
macro avg	0.553560	0.553333	0.548483	300.000000
weighted avg	0.553560	0.553333	0.548483	300.000000

Fig. 4. Classification Report

V. CONCLUSION

The examination of the machine learning model's classification report reveals nuanced insights into its capabilities and limitations in differentiating various types of footwear. The model displays commendable precision and recall in certain categories such as 'Loafers' and 'Boots,' where it effectively distinguishes genuine products from counterfeits. This indicates a robust ability to handle footwear types with distinct features. However, the model's performance in categories like 'Sandals' highlights significant challenges, with low precision and recall suggesting difficulty in accurately identifying and classifying these items. This disparity in performance across categories underscores the importance of targeted improvements to enhance the model's overall efficacy.

Moving forward, the development focus should be on refining the model's architecture and training processes to better accommodate the intricate variations across all footwear types. Enhancing data augmentation techniques, increasing the diversity of the training dataset, and incorporating more sophisticated neural network architectures could potentially improve the model's generalization capabilities. Additionally, applying techniques like transfer learning from models trained on larger, more diverse datasets might provide the necessary robustness. Continued advancements in these areas are essential for achieving higher accuracy and reliability in footwear classification, ultimately leading to stronger brand protection and increased consumer trust in product authenticity. These efforts will not only mitigate the risks associated with counterfeit products in the marketplace but also support the dynamic needs of the retail sector in maintaining high standards of quality and authenticity.

VI. PROJECT PROTOTYPE

The complete report and accompanying source code are available on GitHub for those interested in further research on this topic. This repository provides a valuable resource for researchers looking to explore deeper into the application of deep learning for footwear classification and fraud detection. Following is the link for it.

VII. REFERENCES

Smith, J. (2020) The Economic Impact of the Footwear Industry. Available at: http://www.economicimpactoffootwear.org

Johnson, R. and Lee, H. (2021) Combating Counterfeit Products: Technological Advances and Challenges. Available at: http://www.technologyagainstcounterfeit.org

Chen, X., Kumar, A. and Zhou, Y. (2022) 'Revolutionizing Product Authentication with Deep Learning', Journal of Advanced Computational Intelligence and Informatics, 16(3), pp. 45-59. Available at: https://www.jaciipaperexample.com

Davis, F. (2023) ‘Deep Learning in Fashion and Sports: Economic Implications’, Trends in AI and Economy, 4(2), pp. 112-127. Available at: http://www.aiandeconomytrends.org

noobyogi0100 (2020) Shoe Dataset. Available at: https://www.kaggle.com/datasets/noobyogi0100/shoe-dataset