AUTOMATED EXTRACTION AND FUNCTIONAL CLASSIFICATION OF SOLE PATTERNS BASED ON DEEP LEARNING

Yangxue LUO[1], Rui YUAN[2], Yuhan HUANG[3], Jie ZENG[4], Jin ZHOU[4*]

[1]School of Finance and Commerce, Chengdu Vocational & Technical College of Industry, Chengdu, 610000, China

[2]School of Computer and Software Engineering, Chengdu Neusoft University, Chengdu, 610000, China

[3]Faculty of Intelligent Manufacturing and Automobile, Chengdu Vocational & Technical College of Industry, Chengdu, 610000, China

[4]National Engineering Laboratory for Clean Technology of Leather Manufacture, Sichuan University, Chengdu, 610065, China

Received: 30.03.2026 Accepted: 25.06.2026 https://doi.org/10.24264/lfj.26.2.5

AUTOMATED EXTRACTION AND FUNCTIONAL CLASSIFICATION OF SOLE PATTERNS BASED ON DEEP LEARNING

ABSTRACT. This study presented a computer vision-based method for the automatic extraction and functional classification of Outsole-Pattern design, which currently relies significantly on expertise and lacks digital support. An instance segmentation network based on the Mask-RCNN was utilised to segment the Outsole-Pattern from the image, thereby removing background interference during the pattern recognition. Subsequently, an existing deep-learning-based edge detector, DexiNed, was adopted to extract outsole-pattern edges, compute the greyscale map of the Outsole-Pattern lines. Finally, the EfficientNet served as the backbone network to classify the varied patterns according to their functional characteristics. Results demonstrated that the proposed framework effectively performed outsole's image segmentation and precise edge pattern extraction, five classes of functional pattern were categorized: the lateral, the longitudinal, the zigzag, the block-shaped, and the circular suction cups—based on foot functional areas. Furthermore, classification accuracy was generally over 90% on our custom datasets. This method establishes a structured pattern database that transforms design knowledge into reusable digital templates, enabling the transition in footwear design from experience-based to data-driven approaches, hence enhancing design efficiency and aligning with trends of industry digitalisation.

KEYWORDS: Outsole-Pattern, computer vision, edge detection, image segmentation, digital design of footwear

EXTRAGEREA AUTOMATĂ ȘI CLASIFICAREA FUNCȚIONALĂ A MODELELOR DE TALPĂ EXTERIOARĂ PE BAZA ÎNVĂȚĂRII PROFUNDE

REZUMAT. Acest studiu a prezentat o metodă bazată pe viziunea computerizată pentru extragerea automată și clasificarea funcțională a modelelor de talpă exterioară, proces care, în prezent, se bazează în mare măsură pe expertiză și nu beneficiază de suport digital. S-a utilizat o rețea de segmentare de instanță pe baza Mask-RCNN pentru a segmenta modelul tălpii exterioare din imagine, eliminând astfel interferențele de fundal în timpul procesului de recunoaștere a modelelor. Ulterior, s-a utilizat un detector de margini existent, bazat pe învățare profundă, DexiNed, pentru a extrage marginile modelului tălpii exterioare și pentru a calcula harta în tonuri de gri a liniilor modelului tălpii exterioare. În final, rețeaua EfficientNet a servit drept rețea principală pentru clasificarea diverselor modele în funcție de caracteristicile lor funcționale. Rezultatele au demonstrat că acest cadru propus a realizat în mod eficient segmentarea imaginii tălpii exterioare și extragerea precisă a modelelor marginilor, fiind clasificate cinci categorii de modele funcționale: modelul lateral, cel longitudinal, cel în zigzag, cel în formă de bloc și cel cu ventuze circulare — pe baza zonelor funcționale ale piciorului. În plus, precizia clasificării a depășit, în general, 90% în seturile noastre de date personalizate. Această metodă creează o bază de date structurată a modelelor care transformă cunoștințele de proiectare în șabloane digitale reutilizabile, facilitând tranziția în proiectarea încălțămintei de la abordări bazate pe experiență la cele bazate pe date, sporind astfel eficiența proiectării și aliniindu-se la tendințele de digitalizare din industrie.

CUVINTE CHEIE: modelul tălpii exterioare, viziune computerizată, detectarea marginilor, segmentarea imaginilor, proiectarea digitală a încălțămintei

EXTRACTION AUTOMATIQUE ET CLASSIFICATION FONCTIONNELLE DES MOTIFS DE SEMELLES EXTÉRIEURES À L'AIDE DE L'APPRENTISSAGE PROFOND

RÉSUMÉ. Cette étude présente une méthode basée sur la vision par ordinateur pour l'extraction automatique et la classification fonctionnelle des motifs de semelles extérieures, un domaine qui repose actuellement en grande partie sur l'expertise humaine et qui manque de support numérique. Un réseau de segmentation d'instances basé sur le Mask-RCNN a été utilisé pour segmenter le motif de semelle extérieure à partir de l'image, éliminant ainsi les interférences de l'arrière-plan lors de la reconnaissance des motifs. Par la suite, un détecteur de contours existant basé sur l'apprentissage profond, DexiNed, a été utilisé pour extraire les contours des motifs de la semelle extérieure et calculer la carte en niveaux de gris des lignes de ces motifs. Enfin, le réseau EfficientNet a servi de réseau central pour classer les différents motifs en fonction de leurs caractéristiques fonctionnelles. Les résultats ont démontré que le cadre proposé permettait de réaliser efficacement la segmentation des images de la semelle extérieure et l'extraction précise des contours des motifs, cinq classes de motifs fonctionnels ayant été identifiées : les ventouses latérales, longitudinales, en zigzag, en forme de bloc et circulaires — selon des zones fonctionnelles du pied. De plus, la précision de la classification dépasse généralement les 90 % sur nos ensembles de données personnalisés. Cette méthode permet

* Correspondence to: Assoc. Prof. Dr. Jin ZHOU, National Engineering Laboratory for Clean Technology of Leather Manufacture, Sichuan University, Chengdu, 610065, China, zj_scu@scu.edu.cn

de constituer une base de données structurée de motifs qui transforme les connaissances en matière de conception en modèles numériques réutilisables, facilitant ainsi la transition dans la conception de chaussures d'une approche fondée sur l'expérience vers une approche fondée sur les données, ce qui améliore l'efficacité de la conception et s'inscrit dans les tendances de la numérisation du secteur.
MOTS-CLÉS : motif de semelle extérieure, vision par ordinateur, détection des contours, segmentation d'images, conception numérique de chaussures

INTRODUCTION

From a functional design perspective, outsoles' tread patterns that satisfy these functions can be categorized into the five functional categories defined in Section 1 (Fig. 1). They provide enhanced longitudinal flexibility and grip and are typically incorporated into footwear emphasizing flexibility (Fig. 1a). The longitudinal patterns (parallel to the direction of movement) offer superior longitudinal rigidity, smooth motion and lateral grip, and are suitable for shoes requiring fluid rolling motion (Fig. 1b). The zigzag or hybrid patterns combine the advantages of both transverse and longitudinal designs, providing excellent slip resistance (Fig. 1c). The block patterns use protruding lugs to grip the ground, providing excellent traction, and are frequently used in racing shoes (Fig. 1d). The circular suction cup patterns are specifically designed for agile manoeuvring and are commonly found in basketball shoes and similar footwear (Fig. 1e). Collectively, these structural designs represent the tangible mechanisms through which the functional properties of the outsole are realized.

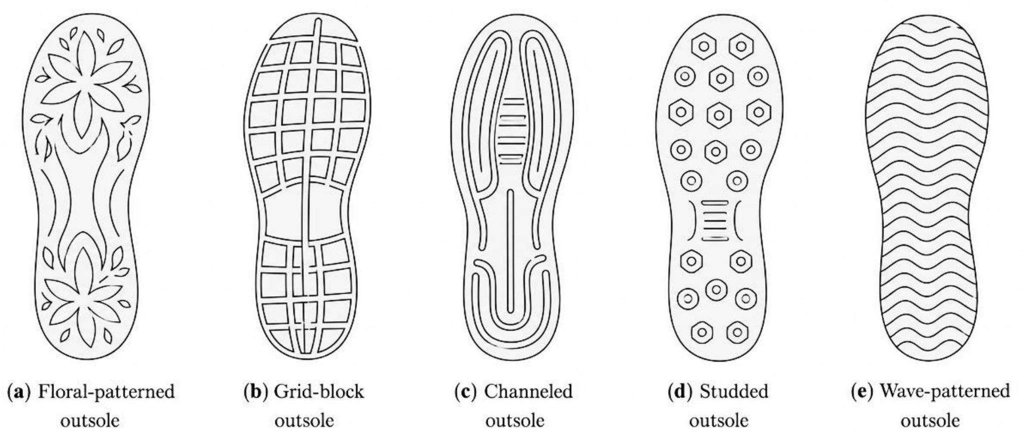

Figure 1. Representative examples of outsole pattern categories used for conceptual illustration

Recent research has recorded recognizing technologies related to footwear patterns. Richetelli *et al*. [1] exploited the Fourier transform method in criminal investigations, achieving an 88.5% accuracy rate in pattern classification for matching suspects' footprints; nevertheless, this technique has not been applied in footwear design. Traditional image recognition techniques are generally inadequate for supporting large-scale, independent design generation, necessitating the integration of advanced computer vision algorithms. The advent of Convolutional Neural Networks (CNNs) in image processing has facilitated their widespread adoption across classification and detection tasks due to their high accuracy and generalizability. Mask Region-based Convolutional Neural Network (R-CNN), introduced by He *et al*. [2], enhances CNN designs by segmenting distinct areas, thus enabling the identification of contours for various categories. This technique has been extensively employed in domains such as facial identification, online virtual try-on, and virtual makeup applications.

Furthermore, instance segmentation and edge detection techniques in computer vision provided an efficient means of generating databases of shoe sole patterns with diverse functional characteristics. Conventional edge detection methods, such as the Canny operator, can produce satisfactory results for greyscale images containing basic structural elements. However, deep learning-based approaches, particularly those employing CNN architectures, have demonstrated superior performance on benchmark datasets such as Berkeley Segmentation datasets 500 (BSDS500). The Dense Extreme Inception Network for Edge Detection (DexiNed), developed by Soria *et al*. [3], demonstrated exceptional edge detection proficiency and the ability to generate fine edges, making it particularly adept at recognizing complex patterns on shoes' outsoles.

Despite the structural simplicity of shoe outsoles' tread patterns, their myriad components, diminutive sizes, and irregular distribution render intelligent patterns notably difficult to recognize, necessitating the continual incorporation of new patterns into the database to maintain the currency of machine learning algorithms. The categorization of outsole patterns is not solely a visual differentiation, but a systematic classification grounded in their physical attributes and functional necessities for locomotion. Liu *et al*.'s edge detection technique [4] enabled the extraction of complex patterns, while Chituc [5] demonstrated enhanced design efficiency through structured databases in the digital transformation of the footwear industry. However, existing research has not performed thorough training on sole models, and the effectiveness of the resulting outsole pattern analysis requires improvement.

This work introduced a computer vision methodology for the automated extraction and functional categorization of patterns to address these challenges. Rather than proposing a new backbone network, the study organized established deep learning models into a task-oriented workflow for outsole-pattern analysis, providing a structured analytical framework for investigating the relationship between outsole-pattern morphology and functional performance.

METHOD

Datasets

This study utilized web crawler technologies to autonomously gather raw footwear photos from publicly accessible fashion platforms, including brand websites, e-commerce product pages, and runway image repositories, to create a representative training dataset. The crawler was developed using the Python-based Scrapy framework; it incorporated keyword filtering (e.g., 'trainers', 'high heels', 'boots') and anti-crawling measures to ensure data diversity and authenticity. Following deduplication, resolution filtering (≥512×512 pixels), and manual review, a final assemblage of 5850 high-quality footwear images was obtained, covering eight principal categories: sports, casual, formal, and outdoor footwear. A total of 2461 images with clearly identifiable outsole regions were utilized for outsole pattern analysis tasks. All data were exclusively intended for non-commercial academic research and had been anonymized to remove facial images and trademark details.

This study employed an internal footwear image dataset in conjunction with two notable public edge detection benchmark datasets—BIPED and BSDS500 for the evaluation and selection of edge detection methodologies. The BIPED dataset, presented by Soria *et al*. [3], consisted of approximately 25,000 high-resolution (1280×720) natural images. The annotations were independently performed by several observers, focusing on intricate edge structures discernible to the human eye, and were particularly suitable for texture-rich, detail-dense environments (such as textiles, masonry, and shoe outsole patterns). In contrast, BSDS500 was a traditional general-purpose benchmark for edge detection, developed by Arbeláez *et al*. [6], comprising 500 images of indoor and outdoor scenes. The annotations highlighted the primary outlines and semantic boundaries of objects, while generally overlooking micro-textural specifics.

Design Framework for the Outsole-Pattern Analysis

Diverse gaits or postures exerted pressure on specific areas of the outsole during locomotion. This study delineated the outsole into three separate regions: the forefoot, midfoot, and heel (Figure 3). The technological process was organized as follows: the outsole pattern was first segmented to acquire the sole contour using the Mask-RCNN, followed by accurate pattern edge detection applying DexiNed. Evaluated using three standard metrics: Optimal Datasets Scale (ODS), Optimal Image Scale (OIS), and Average Precision (AP). The ODS measured the model's overall edge detection effectiveness across the datasets with a uniform threshold; the OIS evaluated the peak performance attainable with a separately selected optimal threshold for each image; while the AP assessed the model's sensitivity to subtle edges by examining the area under the precision-recall curve across multiple thresholds.

Segmentation Algorithm

This study primarily used the Mask-RCNN technique, which was capable of simultaneously performing object detection, classification, and semantic segmentation tasks. The principal workflow of Mask-RCNN was outlined as follows: initially, a Residual Network (ResNet) functioned as the backbone network; through multi-scale feature extraction from the input image, features from layers C1, C2, C3, C4, and C5 were obtained; these feature layers were utilized to construct a feature pyramid, enabling multi-scale fusion to generate effective feature layers. In the region extraction network, candidate bounding boxes of diverse proportions and sizes were generated according to the dimensions of the valid feature layers. Ultimately, redundant candidate boxes were eliminated through non-maximum suppression.

The ROI Pooling was performed on the valid feature layers of the selected candidate boxes to identify regions of interest, which were then fed into the object detection and semantic segmentation networks. In the object detection network, regression was performed for localization, followed by non-maximum suppression to obtain the final predicted bounding box, while simultaneously detecting and classifying objects within the candidate boxes. The resulting bounding box was fed into the semantic segmentation network, which employed it to extract the pertinent feature layer and segment its pixels. This process ultimately generated extensive semantic segmentation data.

Pattern Detection Algorithm

DexiNed was adopted in this study as a pre-trained edge-detection model for extracting outsole-pattern edges. Since the purpose of this work was not to redesign the edge detector itself, detailed architectural settings are referred to in the original publication.

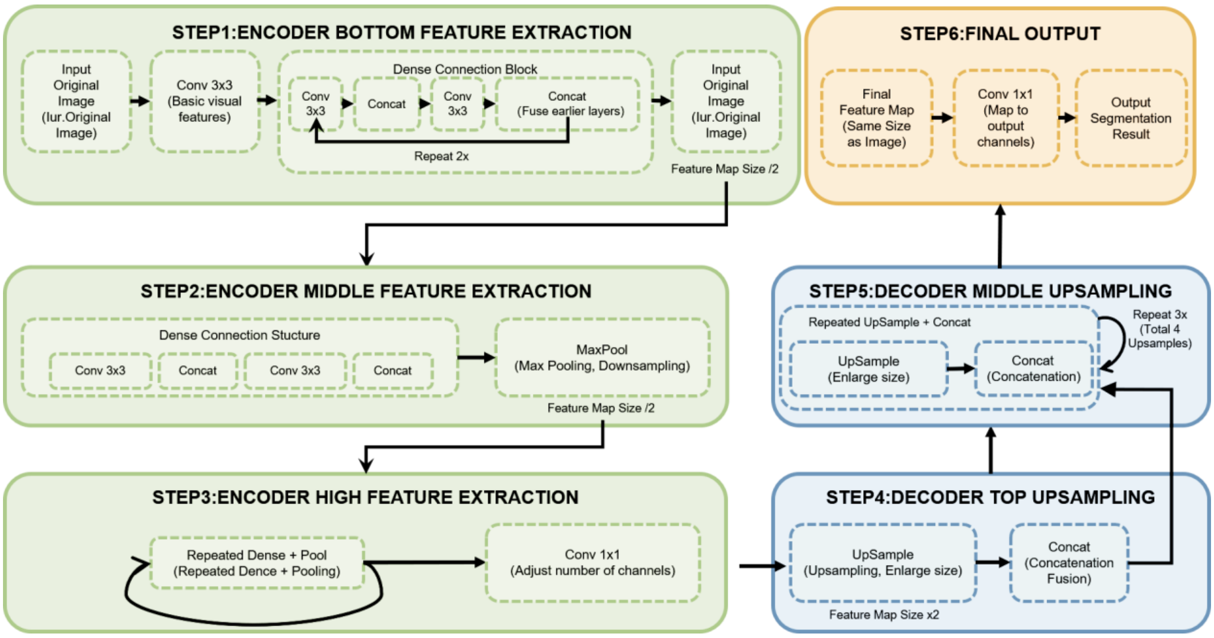

Figure 2. The DexiNet network architecture

The DexiNet network consisted of two main components: the Dexi network and the upsampling module (shown by the blue portion). Upon inputting an RGB image into the Dexi network, feature extraction occurred (Dexi excelled at preserving edge information); the resultant feature maps were further processed by the upsampling module to create edge images, which were ultimately combined to yield the final edge image.

The green block had two stacked convolutional layers, each featuring a kernel size of 3×3, followed by batch normalization, and employed ReLU as the activation function (in the final sub-block, only the last convolutional kernel was active). The max-pooling layer was set with a 3×3 kernel and a stride of 2. The design employed a multi-scale learning approach akin to HED, followed by an upsampling procedure. Due to the potential loss of significant edge features in deeper convolutional layers (as demonstrated in DeepEdge), from the third block onwards, the outputs of the sub-blocks employed edge-connected averaging (indicated by the green parts).

The upsampling module was composed of conditionally stacked sub-blocks, each containing a convolutional layer and a deconvolutional layer. Two categories of sub-blocks existed: Sub-block 1 was activated by either Dexi or Sub-block 2 and was utilized exclusively when the scale disparity between the feature map and the reference standard was exactly 2; Sub-block 2 was employed only when the scale disparity exceeded 2 and was applied iteratively until the scale disparity was reduced to 2. Sub-block 1 consisted of a 1×1 convolution, followed by a ReLU activation function, and culminated in a deconvolution of size S×S, with S representing the dimensions of the input feature map. The final convolutional layer was devoid of an activation function. Sub-block 2 was configured identically to Sub-block 1, with the only difference being the number of filters. The upsampling process in the second layer of the sub-block could be performed with bilinear interpolation, subpixel convolution, or transposed convolution.

DexiNet could be articulated as a regression function, as delineated in Equation 1, where X represented the input image and Y was the corresponding reference standard.

$$Y' = ð(X, Y) \tag{1.}$$

The loss function was articulated by Equation 2:

$$L = -(1-\beta)\sum_{j\in Y^-} \log \sigma\left(y_j = 0 | X; W, W^n\right) \tag{2.}$$

In this context, W signified the collection of all network parameters, comprising n associated parameters, whereas δ specified the weight assigned to each scale level. β = Y/(Y + Y), 1 − β = (Y + Y)/Y, where X and Y represented the edges and non-edges in the reference standard, respectively. We intended to analyse the three parameters—ODS, OIS, and AP—in accordance with the technique.

To assess the efficacy of this approach, we integrated RCF [4] and PidiNet [7] while examining the individual patterns in the BIPED datasets (https://github.com/xavysp/BIPED) and the BSDS500 datasets (https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/resources.html). The selection of these two datasets was predicated on their complementary annotation attributes: BIPED emphasised fine edges discernible to the human eye in high-resolution images (including textures, fabrics, and microstructures), rendering it appropriate for assessing a model's capacity to replicate intricate details in shoe soles; conversely, BSDS500 focused on object contours and semantic boundaries, highlighting overall structural integrity, and was frequently employed for benchmarking general-purpose edge detection algorithms. By juxtaposing results from these two datasets, we were able to thoroughly assess how various approaches reconciled detail fidelity with contour coherence.

Design Framework for the EfficientNet

For information concerning EfficientNet architecture, reference was made to the study by Mingxing Tan *et al*. [8]. Classification of footwear was conducted according to silhouette style: skate shoes—low-top, mid-top, and high-top; trainers—running shoes, football boots, and trail shoes; boots—high-top and low-top.

Presently, three assessment measures were commonly utilised for edge detection: fixed contour threshold (ODS), per-image best threshold (OIS), and average precision (AP). This work utilised the F-measure of AP, ODS, and OIS to evaluate the efficacy of several prominent edge detectors from recent years, with F defined by Equation 3.

$$F = \frac{2 \times precision \times recall}{Precision + recall} \quad (3.)$$

Figure 3 illustrates the processing workflow used for outsole-pattern analysis in this study. The original footwear image was first processed by Mask R-CNN to segment the outsole region; subsequently, pre-trained DexiNed was employed to extract outsole-pattern edges; finally, EfficientNet was used to classify the extracted pattern representation into functional categories.

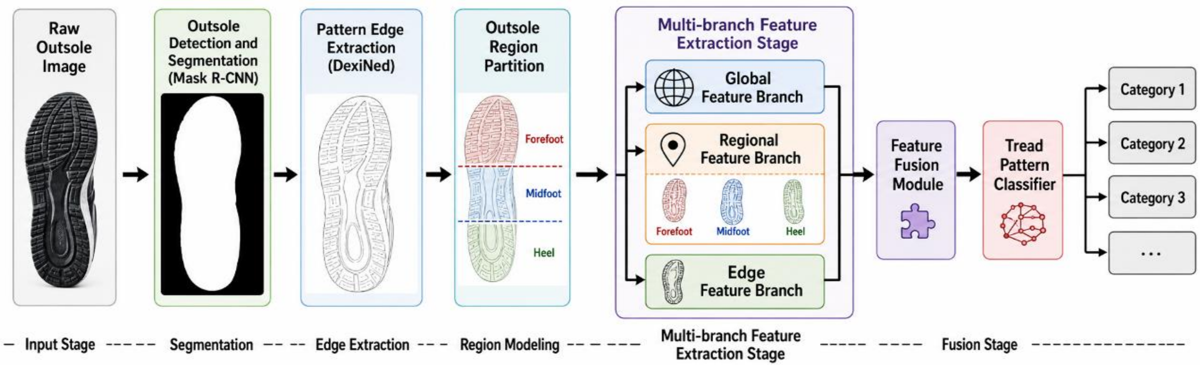

Figure 3. Framework for Outsole Pattern Detection

RESULTS

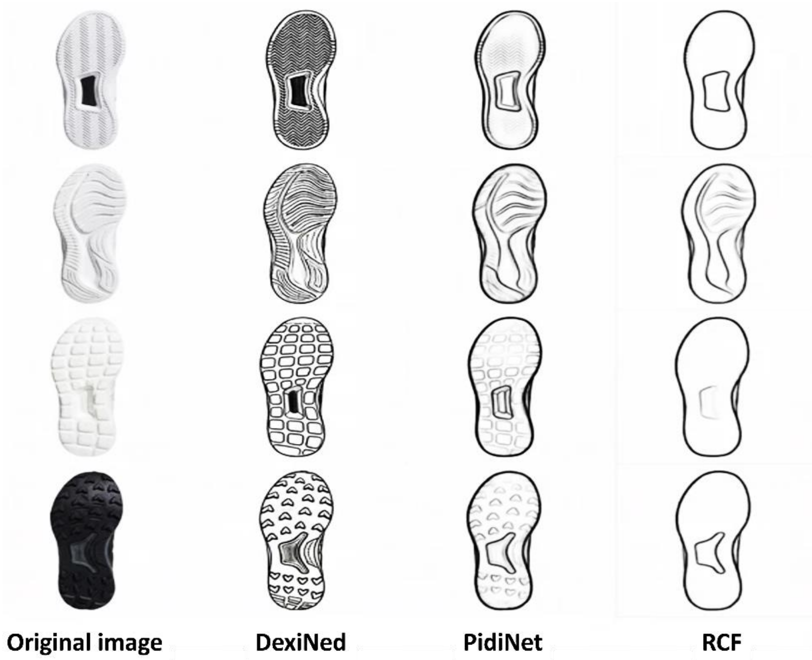

Figure 4. Results of Outsole Segmentation and Edge Detection

Figure 4 illustrated the effectiveness of edge detection for the three algorithms—the DexiNed, PidiNet, and RCF—on individual patterns. From a visual and qualitative standpoint, only the DexiNed was able to accurately reconstruct the fine textural details of the shoe sole. In contrast, the PidiNet and the RCF, due to their multi-pooling network architectures, failed to capture a significant portion of the subtle textures during feature propagation. Consequently, their output edges appeared blurred and coarse, making them unsuitable for the development of a high-precision shoe outsole pattern database.

This phenomenon could be attributed to the underlying principles of algorithmic design. The DexiNed, introduced by Soria *et al*. [3] and motivated by HED and Xception networks, generated visually distinct thin edge maps through dense extreme initial modules, thereby obviating the need for task-specific fine-tuning. It demonstrated superior ODS/OIS performance on benchmark datasets such as BIPED and excelled in tasks requiring precise structural reconstruction, including the extraction of micro-patterns from shoe soles. PidiNet, developed by Su *et al*. [7], aimed to integrate conventional edge operators (e.g., the Canny and the Sobel) into convolutional processes to facilitate backbone edge extraction (ODS = 0.807 on BSDS500 at 100 FPS); however, its lightweight architecture prioritized contour continuity and inference efficiency, resulting in compromised micro-texture fidelity. In contrast, RCF [4] employed the hierarchical pooling mechanisms characteristic of deep networks such as VGG, which inevitably caused the loss of fine-grained spatial information during deep feature abstraction. Therefore, within the context of developing a high-fidelity, single-database framework for functional categorization in this study, DexiNed's ability to preserve delicate edge structures rendered it the most suitable choice.

To further evaluate the efficacy of DexiNed, we performed a quantitative study on the BSDS500 and BIPED datasets. Table 1 illustrated that on the BIPED datasets, which closely mimicked the intricate texture of shoe soles, the DexiNed attained ODS, OIS, and AP scores of 0.859, 0.867, and 0.905, respectively. These scores markedly surpassed those of SED, HED, and the recently high-performing edge detection network RCF (ODS: 0.843; OIS: 0.859; AP: 0.882), thereby affirming its dominance in fine-grained edge detection at the data level. On the general-purpose natural scene dataset BSDS500, DexiNed's scores (ODS: 0.729; OIS: 0.745; AP: 0.689) were markedly inferior to those of RCF (ODS: 0.811; OIS: 0.830; AP: 0.846) and PidiNet (ODS: 0.807; OIS: 0.823; AP: 0.834). This mismatch arose because BSDS500 was not explicitly intended for edge detection tasks; its evaluation metrics penalized the delicate edges identified by DexiNed, which were essential traits to preserve for creating a high-precision solitary pattern database.

Table 1: Performance Comparison of Edge Detection Models

Datasets	Model	ODS	OIS	AP
BIPED	SED	0.717	0.731	0.756
BIPED	HED	0.829	0.847	0.869
BIPED	RCF	0.843	0.859	0.882
BIPED	DexiNed	0.859	0.867	0.905
BSDS500	SED	0.71	0.74	0.74
BSDS500	HED	0.79	0.808	0.811
BSDS500	RCF	0.811	0.83	0.846
BSDS500	PidiNet	0.807	0.823	0.834
BSDS500	DexiNed	0.729	0.745	0.689

To assess the algorithm's precision, our datasets with 2461 available outsole figures were assessed. Following instance segmentation and edge detection, the images were partitioned into forefoot, midfoot, and heel regions and subsequently input into the EfficientNet network for classification; the results were presented in Figure 5. Circular suction cups were exclusively located in the forefoot region, resulting in the absence of corresponding classification data for the remaining regions. This pattern achieved the highest classification accuracy (0.98) within the forefoot region, attributable to its distinctive morphological characteristics. In contrast, the zigzag pattern, characterized by ambiguous textural features and a propensity for confusion with horizontal or vertical patterns, exhibited reduced overall classification performance, with the lowest accuracy observed in the midfoot region (0.82), followed by the forefoot (0.89) and heel (0.88). Block patterns demonstrated consistently high classification accuracy across all foot regions, with rates exceeding 0.90 (0.91 for the forefoot, 0.95 for the midfoot, and 0.91 for the heel). The classification accuracies of horizontal and vertical patterns were comparable across regions, consistently ranging from 0.89 to 0.91 with minimal variation, reflecting uniform performance.

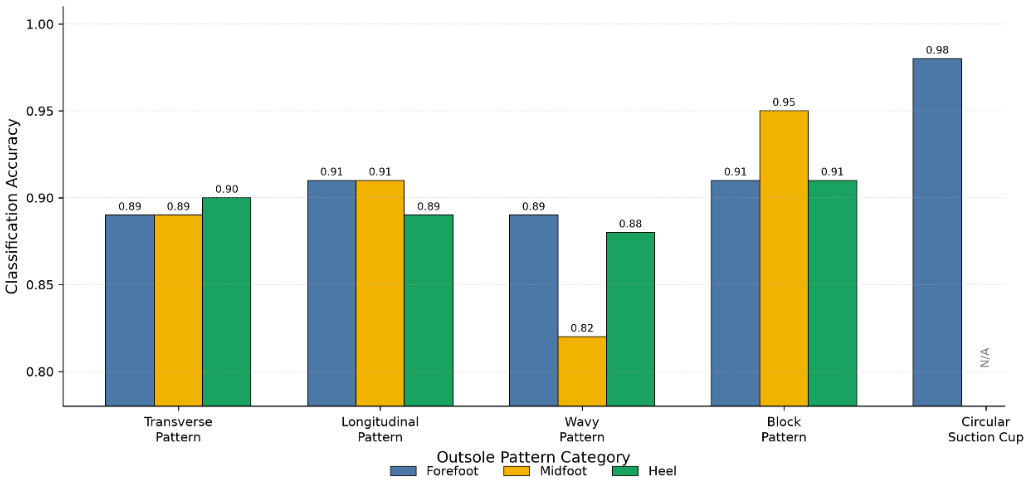

Figure 5. Comparative Analysis of Pattern Classification Accuracy Across Various Foot Regions

DISCUSSION

The present study should be understood as an application-oriented and task-specific methodological framework rather than an attempt to propose new backbone architecture for edge detection or segmentation. The main contribution does not lie in redesigning DexiNed, Mask R-CNN, or EfficientNet at the network level, but in organizing these established models into a coherent pipeline for outsole-pattern analysis. In particular, the methodological value of this work is reflected in three aspects: (1) the construction of a complete workflow linking sole segmentation, pattern-edge extraction, and function-oriented classification; (2) the introduction of outsole region partitioning into forefoot, midfoot, and heel to better align visual pattern analysis with functional interpretation; and (3) the task-specific adaptation of model selection, preprocessing, and category definition for the outsole-design scenario.

From this perspective, the contribution of the study is a form of workflow-level and domain-oriented innovation. Although the underlying backbone models are established methods, their direct use does not automatically yield a usable analytical framework for outsole-pattern design. The outsole domain presents distinctive challenges, including dense micro-textures, irregular local structures, and functionally meaningful regional variation. Therefore, integrating existing models into a structured and interpretable pipeline is itself a meaningful step toward digital design support in footwear research.

At the same time, this study has several limitations. It does not propose a newly designed edge-detection architecture, nor does it investigate architecture-level optimization specifically for outsole data. Future work may focus on building a larger outsole-specific benchmark, improving region-aware feature modeling, and developing more specialized edge-extraction and classification strategies for fine-grained outsole-pattern analysis.

CONCLUSION

This study presented and validated a comprehensive computer vision-based framework for the automated analysis and functional classification of shoe sole patterns. The proposed method first employed the Mask R-CNN model to segment the sole region from raw images. The framework integrates established deep learning models for sole segmentation, pattern-edge extraction, and downstream classification into a unified workflow tailored to outsole-pattern analysis. Results demonstrated that the framework achieved high identification accuracy and robust performance, even in complex image backgrounds. The above findings highlight the potential of transforming traditional, experience-based design methods into data-driven intelligent design approaches, which can improve design efficiency and align with contemporary trends of digitalization and rapid iteration in the footwear manufacturing industry.

REFERENCES

1. Richetelli, N., Lee, M.C., Lasky, C.A., Gump, M.E., Speir, J.A., Classification of Footwear Outsole Patterns Using Fourier Transform and Local Interest Points, *Forensic Sci Int*, **2017**, 275, 102-109, Epub 2017 Mar 4, PMID: 28343023, https://doi.org/10.1016/j.forsciint.2017.02.030.
2. He, K., Gkioxari, G., Dollar, P., Girshick, R., Mask R-CNN, *IEEE Trans Pattern Anal Mach Intell*, **2020**, 42, 2, 386-397, Epub 2018 Jun 5, PMID: 29994331, https://doi.org/10.1109/TPAMI.2018.2844175.
3. Soria, X., Riba, E., Sappa, A., Dense Extreme Inception Network: Towards a Robust CNN Model for Edge Detection, 2020 IEEE Winter Conference on Applications of Computer Vision (WACV), Snowmass, CO, USA, **2020**, pp. 1912-1921, https://doi.org/10.1109/WACV45572.2020.9093290.
4. Liu, Y., Cheng, M.M., Hu, X., Bian, J.W., Zhang, L., Bai, X., Tang, J., Richer Convolutional Features for Edge Detection, *IEEE Trans Pattern Anal Mach Intell*, **2019**, 41, 8, 1939-1946, https://doi.org/10.1109/TPAMI.2018.2878849.
5. Chituc, C.M., Interoperability in the Footwear Manufacturing Networks and Enablers for Digital Transformation, in: Cherfi, S., Perini, A., Nurcan, S. (eds.), *Research Challenges in Information*

Science. RCIS 2021. Lecture Notes in Business Information Processing, vol 415, Springer Cham, **2021**, https://doi.org/10.1007/978-3-030-75018-3_39.

6. Arbeláez, P., Maire, M., Fowlkes, C., Malik, J., Contour Detection and Hierarchical Image Segmentation, *IEEE Trans Pattern Anal Mach Intell*, **2011**, 33, 5, 898-916, https://doi.org/10.1109/TPAMI.2010.161.
7. Su, Z., Liu, W., Yu, Z., Hu, D., Liao, Q., Tian, Q., Pietikäinen, M., Liu, L., Pixel Difference Networks for Efficient Edge Detection, 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, **2021**, pp. 5097-5107, https://doi.org/10.1109/ICCV48922.2021.00507.
8. Tan, M., Le, Q., Efficientnet: Rethinking Model Scaling for Convolutional Neural Networks, Proceedings of the 36th International Conference on Machine Learning, Long Beach, California, PMLR 97, **2019**, 6105-6114.