

Automated program using convolutional neural networks for objective and reproducible selection of corneal confocal microscopy images

Qincheng Qiao^{1,2} , Wen Xue^{1,3} , Jinzhe Li³ , Wenwen Zheng^{1,2},
Yongkai Yuan^{1,2}, Chen Li^{1,4,5,6}, Fuqiang Liu^{1,4,5,6} and Xinguo Hou^{1,4,5,6}

Abstract

Objective: Diabetic peripheral neuropathy (DPN) is a common complication of diabetes, posing a significant risk for foot ulcers and amputation. Corneal confocal microscopy (CCM) is a rapid, noninvasive method to assess DPN by analysing corneal nerve fibre morphology. However, selecting high-quality representative images remains a critical challenge.

Methods: In this study, we propose a fully automated CCM image-selection algorithm based on deep learning feature extraction using ResNet-18 and unsupervised clustering. The algorithm consistently identifies representative images by balancing non-redundancy and representativeness, ensuring objectivity and reproducibility.

Results: When validated against manual selection by researchers with varying expertise levels, the algorithm demonstrated superior performance in distinguishing DPN and reduced inter-observer variability. It completed the analysis of hundreds of images within 1 s, significantly enhancing diagnostic efficiency. Compared with traditional manual selection, the proposed method achieved higher diagnostic accuracy for key morphological parameters, including corneal nerve fibre density, length, and branch density.

Conclusion: The algorithm is open source and compatible with standard CCM workflows, offering researchers and clinicians a robust and efficient tool for DPN diagnosis. Further, multicentre studies are needed to validate these findings in diverse populations.

Keywords

Deep learning, confocal microscopy, diabetic neuropathy

Submission date: 1 December 2024; Acceptance date: 20 February 2025

Introduction

The global prevalence of diabetes has been steadily increasing worldwide. According to the 2021 Diabetes Atlas published by the International Diabetes Federation (IDF), approximately 10.5% of adults worldwide have diabetes,

a figure projected to rise to 12.2% by 2045.¹ Diabetic peripheral neuropathy (DPN) is among the most common chronic complications of diabetes, with distal symmetric polyneuropathy (DSPN) being the most prevalent subtype. Approximately 50% of patients with diabetes develop DSPN.² DPN is a leading cause of foot ulcers,

¹Department of Endocrinology and Metabolism, Qilu Hospital, Shandong University, Jinan, China

²The First Clinical Medical College, Cheeloo College of Medicine, Shandong University, Jinan, China

³College of Intelligence Science and Technology, National University of Defense Technology, Changsha, Hunan, China

⁴Institute of Endocrine and Metabolic Diseases of Shandong University, Jinan, China

⁵Key Laboratory of Endocrine and Metabolic Diseases, Shandong Province Medicine & Health, Jinan, China

⁶Jinan Clinical Research Center for Endocrine and Metabolic Diseases, Jinan, China

Corresponding authors:

Fuqiang Liu, Qilu Hospital of Shandong University, Jinan, Shandong, 250012, P.R. China.
Email: liufuqiang@sdu.edu.cn

Xinguo Hou, Qilu Hospital of Shandong University, Jinan, Shandong, 250012, P.R. China.
Email: houxinguo@sdu.edu.cn



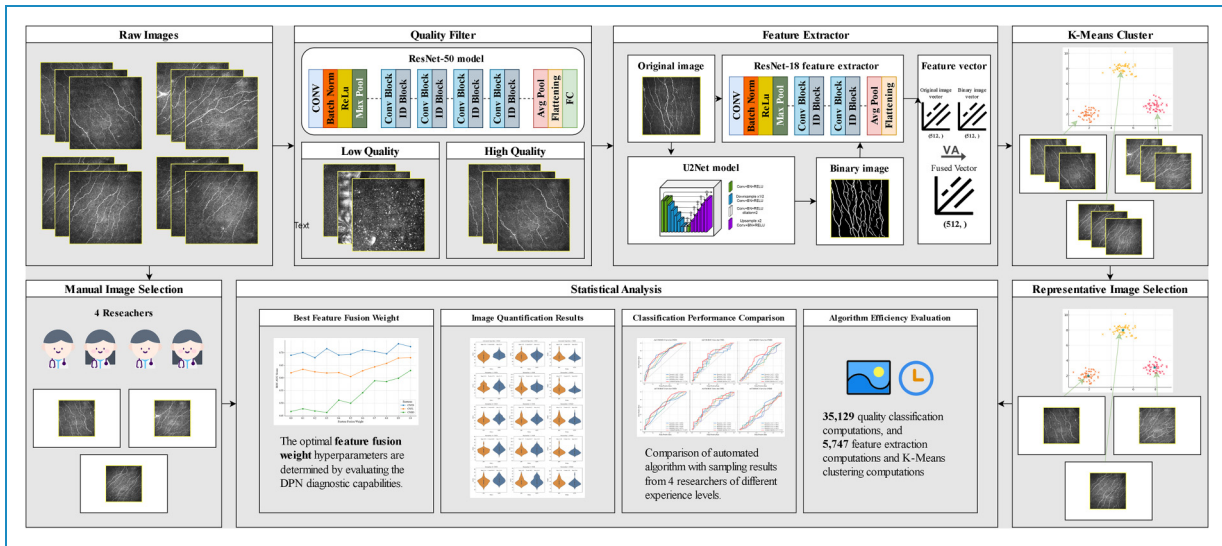


Figure 1. The overall design process and graphic summary of this study.

Charcot joints, and lower limb amputations in patients with diabetes. Therefore, early identification and treatment of DPN are crucial for improving patient outcomes.³

Recent studies have increasingly focused on the relationship between changes in corneal nerve fibres (CNFs) and DPN.^{4,5} Research has demonstrated that morphological alterations in CNF are closely associated with DPN severity and may appear even in the early stages of the condition.^{6–8} The advent of corneal confocal microscopy (CCM) and advanced image analysis software has facilitated the investigation of CNF morphological changes.⁹ CCM offers several advantages as a diagnostic tool for DPN, including being rapid, noninvasive, quantitative, highly reproducible, and sensitive. These features make CCM an ideal method for assessing DPN, with significant potential for clinical application.¹⁰

During CCM examinations, more than 100 consecutive images are typically captured per patient, but only a small subset is selected for research and diagnostic purposes.¹¹ To date, no standardised consensus exists regarding image-selection methods.¹² A multicentre study that established age-adjusted normative values for CNF parameters globally selected three non-overlapping high-quality images from each eye to quantify various morphological parameters.¹³ However, the manual selection of images has significant drawbacks, including high subjectivity, poor reproducibility, and labour-intensive procedures.¹⁴ For an emerging diagnostic method like CCM-based DPN assessment, uncertainty in image selection introduces subjectivity into an otherwise objective approach.

In this study, convolutional neural networks (CNNs) were utilised for quality filtering and feature vector extraction from CCM images. An unsupervised clustering

algorithm was subsequently applied to develop an objective and reproducible image-selection protocol. Its effectiveness was validated by comparing the results with those obtained through a manual selection conducted by researchers with varying expertise levels. In summary, our research aimed to complement the existing CCM-based DPN diagnostic workflows, contributing to improved standardisation and reproducibility.

Methods

Ethical approval and informed consent

This study was approved by the Ethics Committee of Qilu Hospital of Shandong University (Institutional Review Board [IRB] approval number: KYLL-2017–231). Written informed consent was obtained from all participants before the study commenced and imaging procedures were conducted. The study adhered to the principles of the Declaration of Helsinki.

Study design and population

The overall research design is shown in Figure 1. This study was methodological in nature, and diagnostic tests were included to verify the effectiveness of the method.

The study participants were individuals with diabetes admitted to the Department of Endocrinology and Metabolic Diseases at Qilu Hospital of Shandong University, between November 2023 and June 2024. Inclusion criteria were (a) age between 18 and 80 years; (b) diagnosis of diabetes based on the 2024 standards of the American Diabetes Association (ADA)¹⁵; and (c) voluntary provision of written informed consent. Exclusion

criteria included: (a) acute complications or stress states of diabetes, (b) diabetic foot or non-healing ulcers, (c) neuropathy due to central nervous system disorders, chronic alcoholism, tumours, infectious diseases, connective tissue diseases, malnutrition, or deficiencies in folic acid or vitamin B12, (d) chronic liver disease ($\text{AST/ALT} \geq 2$ times the upper limit of normal), (e) chronic kidney disease ($\text{eGFR} \leq 60 \text{ ml/min/m}^2$), (f) active eye diseases, moderate to severe dry eye, history of eye surgery, glaucoma, corneal diseases, systemic conditions affecting the cornea, or prolonged use of contact lenses, and (g) inability to fixate, cooperate with, or tolerate CCM examination. A total of 111 individuals with diabetes participated in this study.

In addition, two external datasets were also utilised: (a) **External Dataset 1**: This dataset comprised 220 participants who visited the Qilu Hospital of Shandong University's (Qingdao and Jinan branches), between December 2016 and June 2020, including 48 healthy controls and 172 patients with type 2 diabetes (100 with DPN, 37 with diabetic nephropathy, and 89 with diabetic retinopathy)¹⁶; (b) **External Dataset 2**: This dataset consisted of 34 healthy participants who underwent physical examinations at the Health Check-up Center of Qilu Hospital of Shandong University between November 2023 and June 2024.¹⁷ All participants underwent CCM examination, and all imaging results were stored. Detailed inclusion and exclusion criteria for the external datasets can be found in the corresponding literature.

Data acquisition

A corneal confocal microscope (HRT-II or HRT-III; Heidelberg Rostock Cornea Module, Heidelberg Engineering Inc., Germany) was used for confocal laser scanning. The main component was a Rostock corneal microscope objective lens equipped with a helium-neon laser source operating at a wavelength of 670 nm. The horizontal and vertical resolution of the microscope was 1 μm , with a magnification of 800 $\times$ and an observation field of 400 $\mu\text{m} \times 400 \mu\text{m}$. Images were acquired and stored in a fully digitalised system. Carbomer eye gel (Berlin, Germany) was used as a coupling agent between the microscope objective and the corneal cap. The eyes were topically anaesthetised with 0.4% proparacaine hydrochloride. During the CCM examination, the participant's head was fixed in position, and their gaze was directed at a central point. The objective lens was moved forward until it contacted the central cornea. The focus was fine-tuned to obtain images of the corneal endothelial cells. The lens was then advanced to capture images of the sub-basal nerve plexus layer, and multiple images were acquired and stored.

Each participant underwent a CCM examination of both eyes. Trained professionals performed electrophysiological tests on patients with diabetes to measure the motor nerve conduction velocity of both the median and tibial nerves.

The Neuropathy Disability Score (NDS) was also used to assess DPN symptoms. The modified NDS evaluated signs of neuropathy, including bilateral ankle reflexes, vibration perception of the great toe (using a 128 Hz tuning fork), pinprick sensation on the dorsal foot, and temperature sensation. DPN was diagnosed according to the Toronto Consensus criteria, which require the presence of symptoms (abnormal Neuropathy Symptom Profile, NSP) or signs of neuropathy ($\text{NDS} > 2$) and abnormal nerve conduction velocity (median nerve $< 50 \text{ m/s}$ in the upper limbs or tibial nerve $< 40 \text{ m/s}$ in the lower limbs).¹⁸

A total of 35,129 CCM images were collected from 111 participants. Among these, 49 patients with diabetes were diagnosed with DPN, while 62 were classified as having NDPN. The dataset composed of these images is referred to as the '**internal dataset**' in this study.

Dataset selection and annotation

This study utilised various selection and annotation processes based on the aforementioned **internal** and **external datasets**. The newly created datasets were non-overlapping and used for subsequent downstream tasks.

Dataset A: Two experienced senior and two less experienced junior researchers selected and annotated the dataset. The senior researchers (Researchers 1 and 2, one male and one female) had at least 1 year of experience with CCM and had published at least one research paper on the subject. The junior researchers (Researchers 3 and 4, one male and one female) were healthcare professionals with no prior CCM experience; however, they received comprehensive training before undertaking the task. All researchers followed a standardised protocol¹⁰: from the original CCM images of each participant, six images of the sub-basal nerve plexus were selected (three images from each eye). These images were required to be of high quality, non-overlapping, and contain varying numbers of major nerve fibres. The selection process covered the entire **internal dataset** and was performed independently by each researcher without interference.

Dataset B: This dataset was annotated by senior researchers (Researchers 1 and 2). High-quality CCM images were defined as those with clear nerve fibres and high contrast against the background. Low-quality images were identified based on the following criteria: (a) parts or the entirety of the image were out of focus and did not belong to the sub-basal nerve plexus; (b) distortion due to rapid eye movement during image acquisition; (c) indentations in the image resulting in truncation of nerve fibres; and (d) any other reasons that resulted in incomplete nerve fibre imaging (Figure 2). High- or low-quality labels were annotated interactively by two researchers using a Web UI developed with the Vue3 framework. Specifically, Researcher 1 randomly selected a CCM image from **External Dataset 1** and initially labelled it as high or low

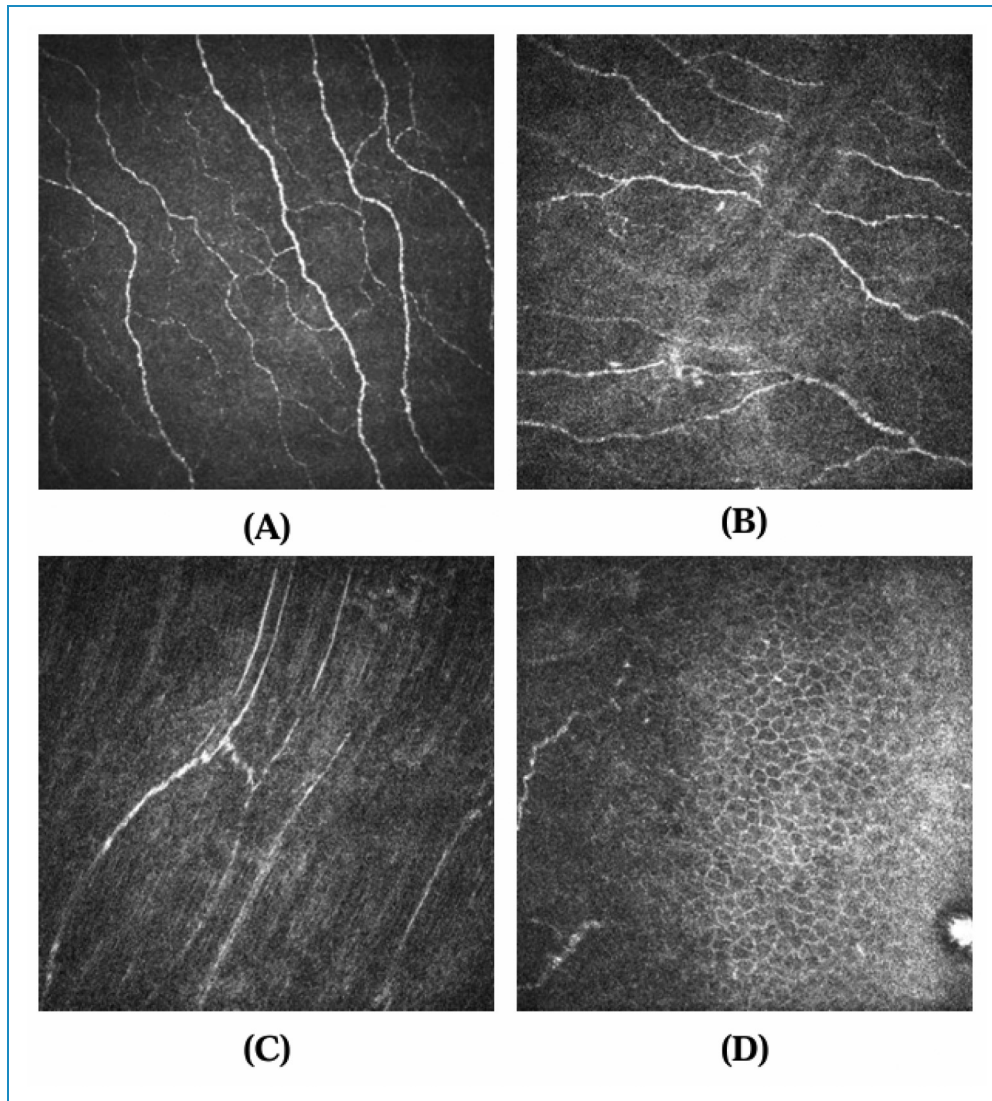


Figure 2. Example of CCM image quality classification. Image A is considered a high-quality image, while the remaining images, B, C, and D, are regarded as low-quality. (A) The nerve fibres are clear, with a distinct contrast against the background. (B) The centre of the image shows an indentation caused by excessive contact of the lens with the cornea, leading to discontinuity of the nerve fibres. (C) The rapid movement of the eyeball causes distortion in the image, preventing an accurate representation of the nerve fibre morphology. (D) The majority of the image is focused on the corneal endothelium, failing to display the nerve plexus beneath the base. CCM: corneal confocal microscopy.

quality. The image was then added to a queue, and Researcher 2 performed a secondary annotation without knowledge of the initial label. Only images with consistent labels from both researchers were stored in the results. The labelling process was terminated once a sufficient number of labelled images were collected. To ensure sample balance, 260 high-quality and 260 low-quality images were randomly selected from the labelled set for inclusion in the dataset.

Dataset C: A total of 210 CCM images were selected from **External Dataset 2**. Researcher 2 used the Labelme tool to outline the boundaries between the nerve fibres and the background in each CCM image, completing the

initial image segmentation labelling task. Subsequently, Researcher 1 reviewed and modified the annotation. The final labels were determined through consensus between the two researchers.

Algorithm design

The complete image-selection algorithm comprises three main steps:

First, a binary classification model was developed to categorise raw CCM images into high- and low-quality groups. ResNet-50 was used as the backbone network, with the output channels of the final fully connected layer

set to two. The model was initialised with pre-trained weights from ImageNet. A total of 220 high- and 220 low-quality CCM images were randomly selected from **Dataset B** as the training set, with 40 high- and 40 low-quality images as the validation set. Transfer learning was applied to train the model on the training set, and classification accuracy was assessed on the validation set after each epoch. Once trained, the model filtered the raw CCM images, excluding those classified as low-quality images.

Second, a model was designed to extract and integrate features from the images. This model comprised two parallel ResNet-18 networks, each modified by removing the fully connected layer and initialised with pre-trained weights from ImageNet. The model received two input types: the original CCM image and the binary nerve fibre image. The U2Net model, trained on **Dataset C**, generated binary nerve fibre images through semantic segmentation. Detailed training parameters and procedures are documented in prior studies.¹⁷ Instead of concatenation, features extracted from both modalities were fused using a weighted feature fusion method, with weights corresponding to pre-defined fusion ratios (ranging from 0.0 to 1.0). This fused feature representation balanced the abstraction of information from both modalities.

Finally, the 512-dimensional feature vectors obtained in the second step were subjected to k-means clustering and grouped into predefined clusters. For each cluster, the distance between the cluster centre and each feature vector was computed. The image(s) closest to the cluster centre were selected as representative images.

An image-selection operation was performed on the **internal dataset** using this algorithm. To ensure comparability, the same rule was applied, selecting three images per eye, totalling six images per participant. Multiple subset datasets were generated by varying the fusion ratios of the original and binarised image features. The fusion ratio ranged from 0.0 (binarised image features only) to 1.0 (original image features only) in intervals of 0.1, resulting in 11 subset datasets with different hyperparameter settings. As the fusion ratio increased, the contribution of features from binarised images decreased, while the contribution of features from original images increased.

Statistical analysis

The morphological parameters of CNFs in the CCM images were calculated using two automated tools: ACCMetrics (ACCM), a well-established automated calculation tool,¹⁹ and AiCCMetrics (AiCCM), a recently developed deep learning-based tool.¹⁷ The parameters included corneal nerve fibre length (CNFL), corneal nerve fibre density (CNFD), and corneal nerve branch density (CNBD). Since each participant selected six representative images by manual or automatic methods, the parameters of each

image were calculated and their average values were obtained as the final values for diagnosis and evaluation.

All statistical tests for difference analyses were two-tailed, with the significance level set at $p < 0.05$. Normality was assessed using the Shapiro–Wilk test. For continuous variables with normal distributions, *t*-tests were conducted, whereas the Mann–Whitney U test was applied for non-normally distributed variables.

The diagnostic and classification performances of each parameter for the outcome were evaluated using receiver operating characteristic (ROC) curves. The area under the curve (AUC) was calculated. For scenarios involving multiple parameters, a logistic regression model was used to construct a combined prediction model. The overall classification and diagnostic performance of the model were analysed using ROC curves. The truth label used to calculate the ROC curve, which is the gold standard for diagnosing DPN, was determined using the Toronto Consensus criteria.

The minimum sample size required for the diagnostic test was calculated using the PASS 11 software, with type I and type II error probabilities set at 0.05 (one-sided) and 0.1 (i.e., 90% statistical power).

All analyses, statistics, and code implementation were performed in Python 3.11.5, and executed on hardware equipped with an Intel Core i5–11300H processor (11th generation) and an NVIDIA GeForce RTX 3050 laptop GPU.

Results

Sample size and statistical efficiency calculation

Based on a review of previously published studies,²⁰ the expected AUC was set at 0.75. When the sample size of the disease and control groups was 21, the statistical power was 0.908. With a sample size of 40 for both groups, the statistical power increased to 0.994. These results confirm that the sample size in this study was sufficient.

Best feature fusion weight

A quantitative analysis was performed on the subset datasets generated using 11 different feature fusion weight hyperparameters with AiCCM. ROC curve analysis was used to determine AUC values and evaluate the ability to distinguish DPN. The results showed that as the hyperparameter increased from 0.0 to 1.0, the overall classification performance of the three metrics—CNFL, CNFD, and CNBD—for DPN improved.

The highest AUC for CNFD (0.7351) was observed at a hyperparameter of 0.9. For CNFL and CNBD, the highest AUCs were 0.6797 and 0.6289, respectively, both achieved at a hyperparameter of 1.0 (Figure 3).

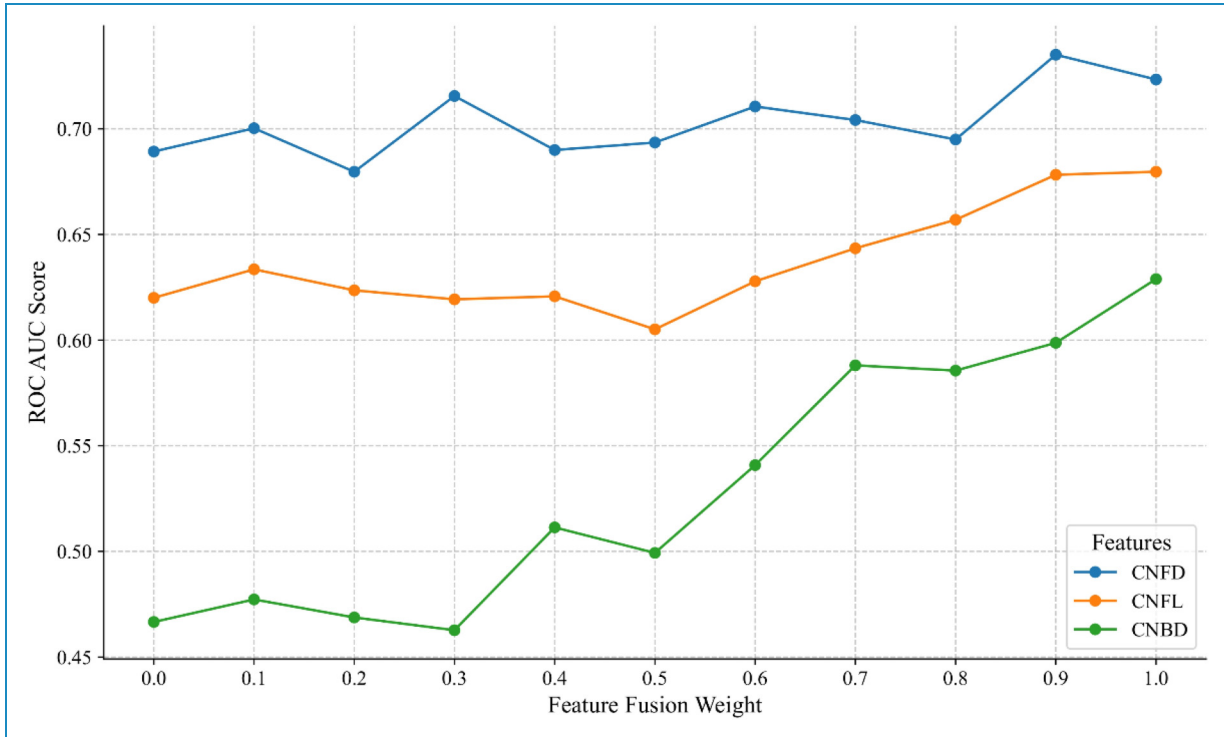


Figure 3. ROC curve analysis of diagnostic performance for DPN using different feature fusion weights. DPN: diabetic peripheral neuropathy; ROC: receiver operating characteristic.

Image quantification results

A subset with a feature fusion weight of 1.0, where the features were entirely derived from the original images, was compared with the manual selection results (**Dataset A**). Parameter quantification was conducted using both the ACCM and AiCCM.

For CNFD using AiCCM, the automated algorithm yielded average values of 20.70 (DPN group) and 24.20 (NDPN group), closely matching the average values from the four researchers (20.95 and 24.33). However, the researchers' measurements showed considerable variability, ranging from 19.40 to 21.91 in the DPN group and 23.66 to 25.73 in the NDPN group.

For CNFL using AiCCM, the automated algorithm measured nerve fibre lengths as 17.12 (DPN) and 18.44 (NDPN). These values were comparable to the researchers' averages of 16.72 and 17.52. The researchers' measurements showed wider variability, ranging from 15.82 to 21.91 in the DPN group and 17.07 to 18.74 in the NDPN group.

For CNBD, using AiCCM, the automated algorithm measured nerve fibre branch density as 38.64 (DPN) and 40.84 (NDPN). In comparison, the researchers' averages were 37.71 and 37.31, with ranges of 35.48 to 43.07 in the DPN group and 34.30 to 41.57 in the NDPN group.

The comparative results obtained using ACCM were similar to those from AiCCM, although the average values for CNFL and CNBD were lower. Overall, the

automated algorithm produced results close to the researchers' averages, significantly reducing variability and individual differences (Table 1, Figures 4, and 5).

Classification performance comparison

Using AiCCM, the automated algorithm demonstrated comparable performance to Researcher 1 in distinguishing DPN for CNFD and outperformed the other researchers. For CNFL and CNBD, the automated algorithm outperformed all other methods in its ability to distinguish DPN.

Similarly, when using ACCM, the automated algorithm outperformed all researchers in distinguishing DPN for CNFD and CNFL. However, for CNBD, all the results showed relatively poor performance, possibly due to inherent limitations in the parameter quantification tool.

Moreover, when applying logistic regression to combine the predictions from the three parameters, the performance was as follows: With AiCCM, the automated algorithm performed comparably to Researchers 1 and 2 and outperformed the rest. With ACCM, this automated algorithm using logistic regression performed slightly worse than Researcher 2 but still outdid the other researchers (Figure 6).

Algorithm efficiency evaluation

These tasks were executed on an NVIDIA GeForce RTX 3050 laptop GPU, involving 35,129 quality-classification

Table 1. The mean values of CNFD, CNFL and CNBD of different sampled datasets were calculated using two automatic quantification tools for morphological parameters of nerve fibres.

Method/Researcher	Source	CNFD (DPN)	CNFD (NDPN)	CNFL (DPN)	CNFL (NDPN)	CNBD (DPN)	CNBD (NDPN)
Automated Algorithm	AiCCM	20.70	24.20	17.12	18.44	38.64	40.84
Researcher 1	AiCCM	19.40	23.66	15.82	17.07	35.48	34.30
Researcher 2	AiCCM	20.37	24.94	17.14	18.65	36.24	40.18
Researcher 3	AiCCM	21.91	25.73	16.35	17.62	36.04	35.18
Researcher 4	AiCCM	22.10	24.98	17.58	18.74	43.07	41.57
Average (Researchers)	AiCCM	20.95	24.33	16.72	17.52	37.71	37.31
Automated Algorithm	ACCM	19.99	22.39	13.33	14.12	31.30	29.40
Researcher 1	ACCM	19.69	21.15	12.34	13.02	27.57	24.09
Researcher 2	ACCM	20.64	23.46	13.41	14.42	33.14	30.01
Researcher 3	ACCM	20.00	23.78	12.71	13.68	26.50	24.12
Researcher 4	ACCM	21.61	24.00	14.02	14.43	34.24	29.66
Average (Researchers)	ACCM	20.49	23.10	13.12	13.89	30.36	27.97

DPN: diabetic peripheral neuropathy; CNFD: corneal nerve fibre density; CNBD: corneal nerve branch density; CNFL: corneal nerve fibre length; ACCM: ACCMetrics; AiCCM: AiCCMetrics.

computations, 5747 feature-extraction computations, and k-means clustering computations, with a total runtime of 377.33 s. Given that the time complexity of the algorithm is $O(n)$, the average time per task was calculated. The results showed that each selection task required an average of 3.39 s, while the average time per 100 images was 1.07 s.

Discussion

This study introduces a CCM image-selection algorithm that utilises deep learning for feature extraction and unsupervised clustering. The proposed algorithm outperformed manual selection by four researchers with varying expertise levels. It consistently produced results with identical inputs, ensuring reproducibility in the CCM image selection. Additionally, the algorithm operates at high speed on standard personal computers, facilitating its integration into existing CCM-based DPN diagnostic workflows.

Diagnosing DPN with CCM involves two steps: selecting representative images from raw data and quantifying nerve fibre morphological parameters either manually or automatically. The parameters are then averaged and compared with population standard values. While most current research focuses on the latter step, with the validity of automated quantitative parameter analysis well established^{21,22} studies on the image-selection process remain limited.

Kalteniece et al. (sample size: $n = 35$) proposed a standardised manual selection protocol, demonstrating inter-observer consistency and reproducibility.¹⁰ Similarly, Schaldemose et al. (sample size: $n = 23$ and $n = 62$) introduced a simple automated selection method based on systematic selection along nerve fibre orientation.^{12,23} However, the first study did not evaluate diagnostic performance differences between automated and manual selection, while the second reported slightly lower DPN classification performance with automated methods compared to manual selection. This discrepancy may stem from the limited consideration of image features in automated approaches.

Interestingly, previous studies have shown that CNNs mimic human visual processing mechanisms through pooling, residual connections, and activation functions.²⁴ In this study, the feature extraction layers of ResNet-18 were used to abstract image features into high-dimensional vectors. ImageNet pre-trained weights enabled the model to capture multilevel features, ranging from low-level details (e.g., edges and textures) to high-level patterns (e.g., shapes and structures). These features are versatile and applicable across various domains. An unsupervised clustering algorithm grouped similar features, identifying cluster centres as representative features and embodying the principles of ‘non-redundancy’ and ‘representativeness’. Being unsupervised and free from pretraining

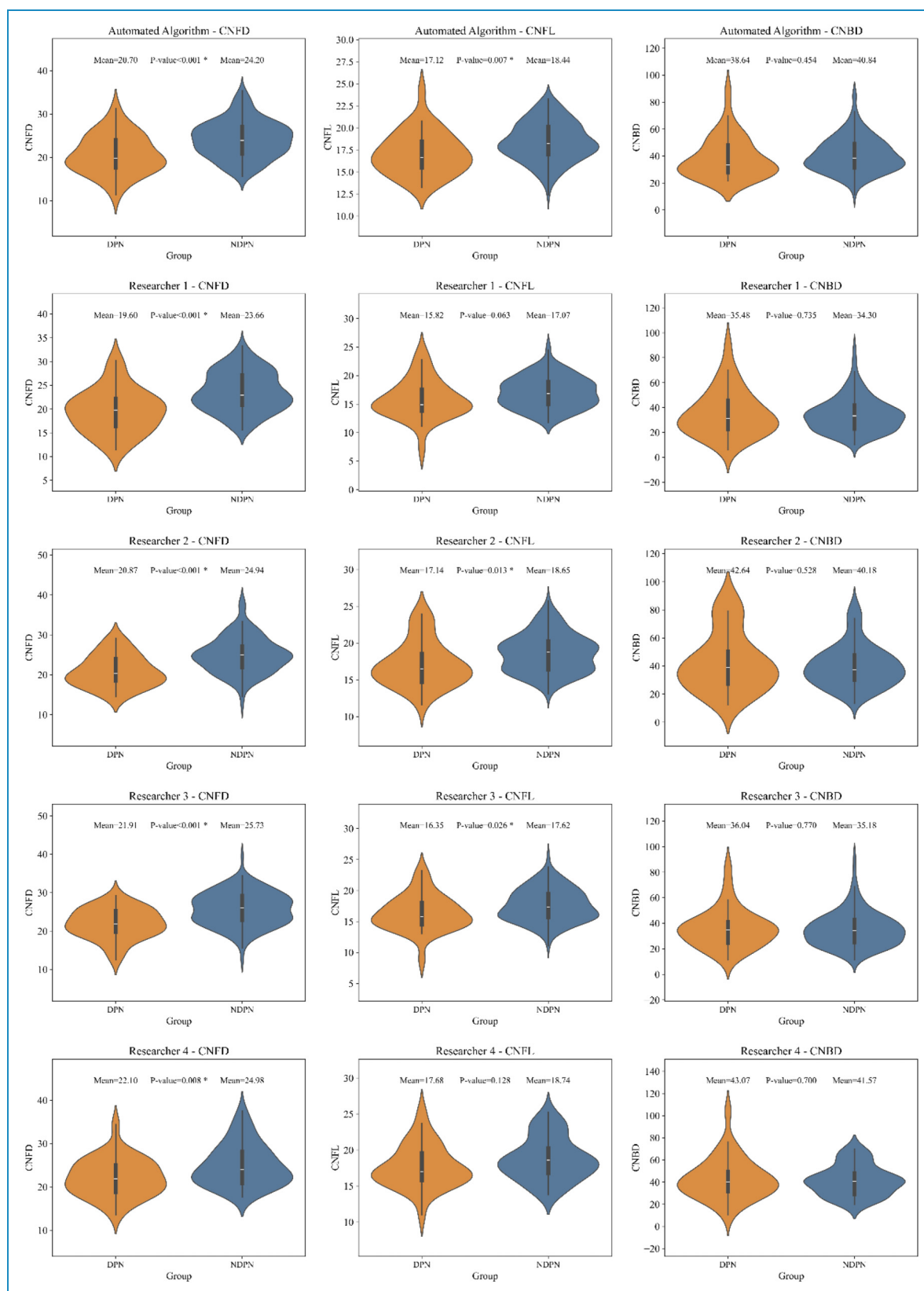


Figure 4. Diagnostic performance of automated selecting using AiCCM for DPN. DPN: diabetic peripheral neuropathy; AiCCM: AiCCMetrics.

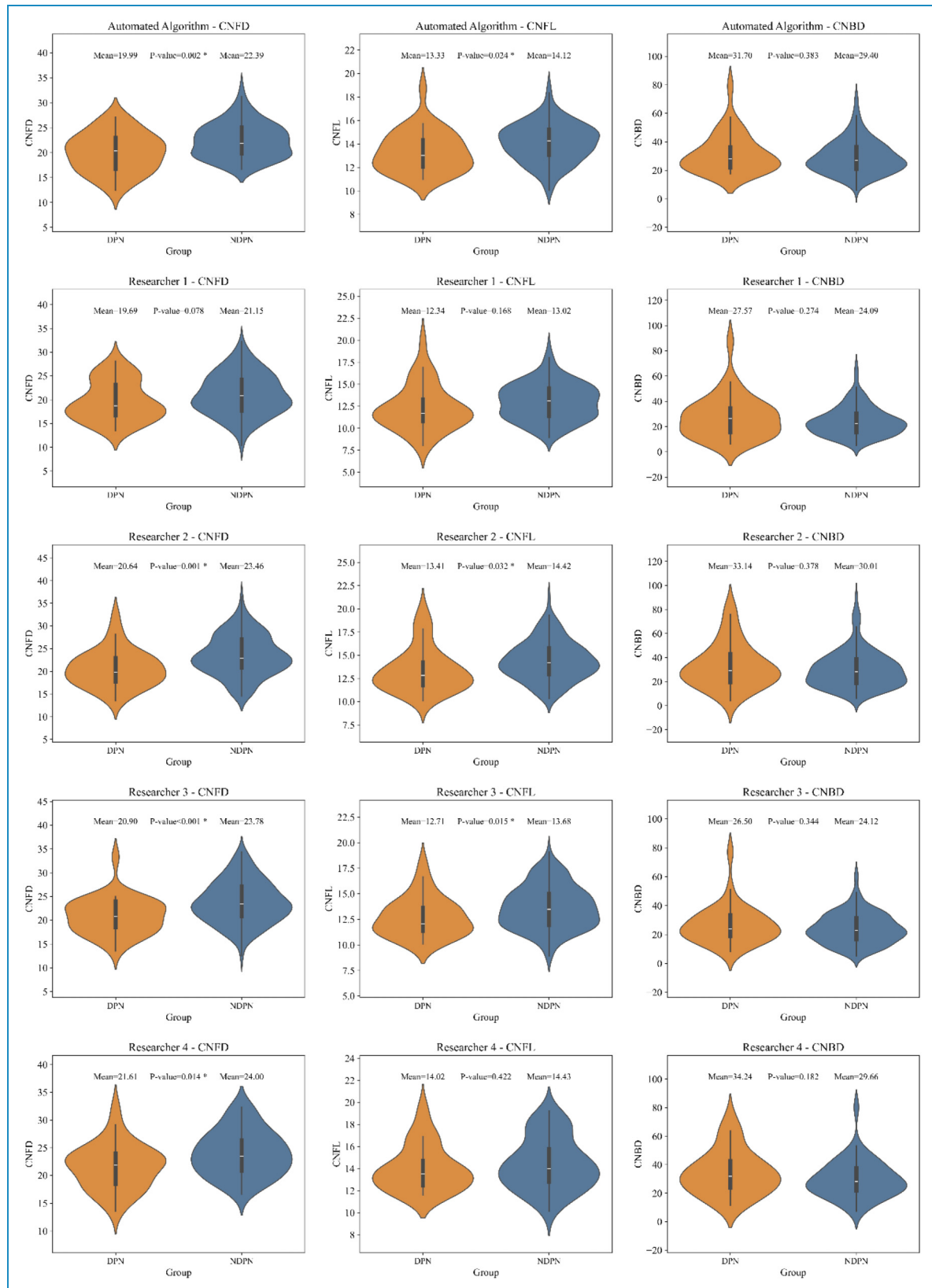


Figure 5. Diagnostic performance of automated selecting using ACCM for DPN. DPN: diabetic peripheral neuropathy; ACCM: ACCMetrics.

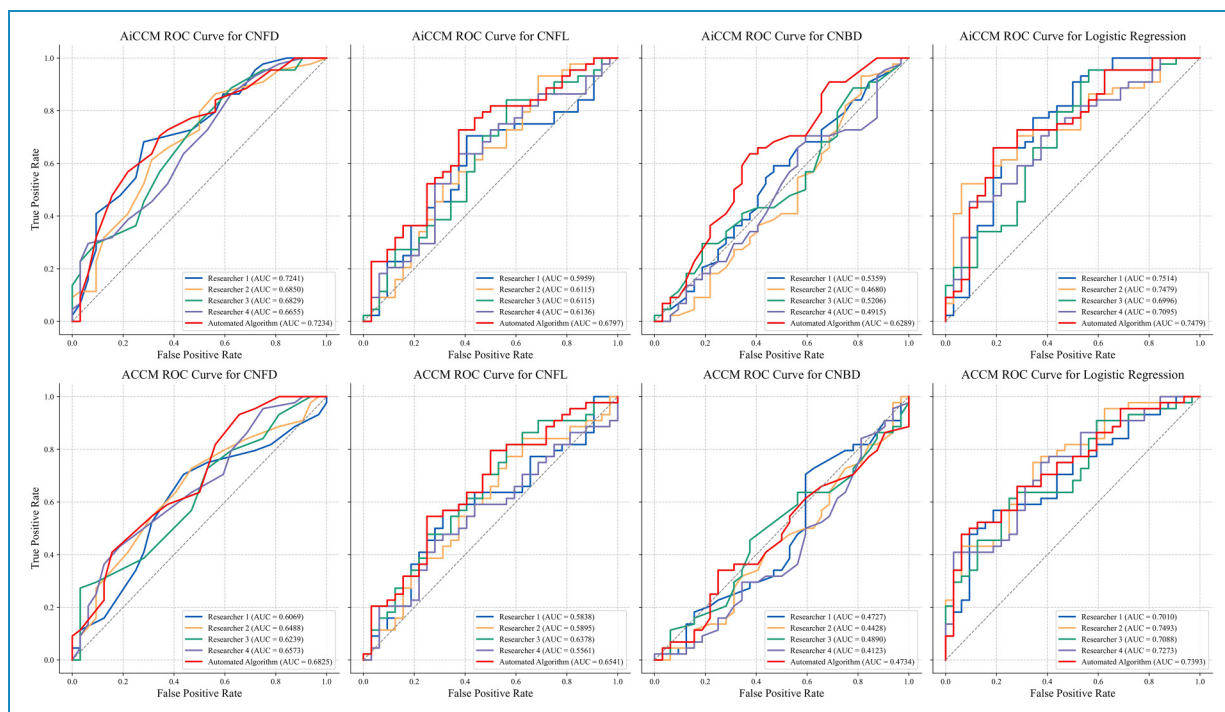


Figure 6. Comparison of classification performance between automated and manual selecting for DPN diagnosis. DPN: diabetic peripheral neuropathy.

biases or randomness, the algorithm ensures objective and reproducible results. Our study, based on larger sample size ($n=111$), demonstrated that the proposed automated selection algorithm achieved superior diagnostic performance for DPN compared to manual selection by multiple researchers. This significantly reduces the workload of researchers and enhances diagnostic efficiency for DPNs. On average, the algorithm processed 100 CCM images in less than 1 s.

Notably, when the feature fusion weight was set to 0, relying solely on binarised nerve fibre image features, the diagnostic performance for DPN matched that of manual selection. In this scenario, the features of interest (nerve fibre skeletons) aligned with those selected manually. As the feature fusion weight increased, incorporating more features from the original images steadily improved the diagnostic performance of the three morphological parameters. This suggests that CNNs can identify subtle features beyond those captured in manual selection, thereby enhancing diagnostic efficacy.

An emerging alternative involves stitching multiple local images into a panoramic view, either manually or automatically. Preliminary studies have developed automated stitching algorithms,^{25,26} but these methods require high-quality original images and lack the accuracy required for clinical applications. Moreover, current tools for quantifying nerve fibre parameters and reference normal values are incompatible with stitched images, hindering their integration into existing CCM workflows for DPN diagnosis.

While this approach holds significant potential, a suitable selection algorithm is currently more practical and urgently needed.

This study has some limitations. First, it was conducted at a single centre, and its findings need validation through multicentre studies with larger and more diverse populations. Additionally, all participants were Chinese, which may limit the generalisability of the findings to other populations.

In conclusion, this study is the first to propose a fully automated CCM image-selection algorithm using deep learning, demonstrating superior performance in DPN diagnosis compared to manual selection.

Conclusion

This study proposes the first fully automated CCM image-selection algorithm using deep learning, which demonstrates superior diagnostic performance for DPN compared to manual selection. This algorithm ensures reproducibility, reduces workload, and integrates seamlessly into existing workflows, making it a practical tool for clinical and research applications. However, further multicentre studies with diverse populations are required to validate the generalisability.

Acknowledgements: This study was supported by the Jinan Clinical Research Centre for Endocrine and Metabolic Diseases. This study was funded by the National Key Research and Development Program of China (2023YFA1801100,

2023YFA1801104, and 2022YFA1004800) and the Taishan Scholars Program of Shandong Province (Grant No. tstp20231250). The funders played no role in the study design, data collection, data analysis and interpretation, or the writing of the manuscript.

Authors' contributions: X.H. and F.L. were responsible for the overall project conceptualisation, supervision, and funding and reviewed the manuscript. Q.Q. was responsible for the research, algorithm design, statistical analysis, figure preparation, and manuscript writing and was also one of the researchers involved in the manual selection mentioned in the manuscript. J.L. provided suggestions for algorithm design. W.X. was responsible for data collection and was one of the researchers involved in the manual selection mentioned in the manuscript. W.Z. and Y.Y. were involved in the manual selection mentioned in the manuscript. L.C. provided the computational resources. All the authors have read and approved the final version of the manuscript.

Code availability: All the codes for this project are open source under the GNU General Public Licence v3.0 (GPL-3.0) and can be freely accessed at <https://github.com/SummerColdWind/AiCCMainstay>

Data availability: Data supporting the analyses in this study are available from the corresponding author upon request.

Declaration of conflicting interests: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethics approval and consent to participate: This study was approved by the Ethics Committee of the Qilu Hospital of Shandong University (Institutional Review Board [IRB] approval number: KYLL-2017–231), and all participants provided informed consent at their respective institutions. This study was conducted in accordance with the principles of the Declaration of Helsinki. All participants provided informed consent prior to imaging.

Funding: The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Key Research and Development Program of China, Taishan Scholars Program of Shandong Province, (grant number 2022YFA1004800, 2023YFA1801100, 2023YFA1801104, tstp20231250).

ORCID iDs: Qincheng Qiao  <https://orcid.org/0009-0009-7900-2984>

Wen Xue  <https://orcid.org/0009-0004-7688-8226>

Jinze Li  <https://orcid.org/0009-0007-8366-5110>

References

1. Sun H, Saeedi P, Karuranga S, et al. IDF Diabetes atlas: global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabetes Res Clin Pract* 2022; 183: 109119.
2. Feldman EL, Callaghan BC, Pop-Busui R, et al. Diabetic neuropathy. *Nat Rev Dis Primers* 2019; 5: 41.
3. Yang H, Sloan G, Ye Y, et al. New perspective in diabetic neuropathy: from the periphery to the brain, a call for early detection, and precision medicine. *Front Endocrinol (Lausanne)* 2019; 10: 929.
4. Alam U, Anson M, Meng Y, et al. Artificial intelligence and corneal confocal microscopy: the start of a beautiful relationship. *J Clin Med* 2022; 11: 6199.
5. Ferdousi M, Kalteniece A, Azmi S, et al. Diagnosis of neuropathy and risk factors for corneal nerve loss in type 1 and type 2 diabetes: a corneal confocal microscopy study. *Diabetes Care* 2021; 44: 150–156.
6. Kalteniece A, Ferdousi M, Petropoulos I, et al. Greater corneal nerve loss at the inferior whorl is related to the presence of diabetic neuropathy and painful diabetic neuropathy. *Sci Rep* 2018; 8: 3283.
7. Malik RA, Kallinikos P, Abbott CA, et al. Corneal confocal microscopy: a non-invasive surrogate of nerve fibre damage and repair in diabetic patients. *Diabetologia* 2003; 46: 683–688.
8. Ziegler D, Papanas N, Zhivov A, et al. Early detection of nerve fiber loss by corneal confocal microscopy and skin biopsy in recently diagnosed type 2 diabetes. *Diabetes* 2014; 63: 2454–2463.
9. Petropoulos IN, Alam U, Fadavi H, et al. Rapid automated diagnosis of diabetic peripheral neuropathy with in vivo corneal confocal microscopy. *Invest Ophthalmol Visual Sci* 2014; 55: 2071–2078.
10. Kalteniece A, Ferdousi M, Adam S, et al. Corneal confocal microscopy is a rapid reproducible ophthalmic technique for quantifying corneal nerve abnormalities. *PloS one* 2017; 12: e0183040.
11. Jin Y, Wang W, Chen W, et al. Corneal confocal microscopy: a useful tool for diagnosis of small fiber neuropathy in type 2 diabetes. *J Diabetes Investig* 2021; 12: 2183–2189.
12. Schaldemose EL, Hammer RE, Ferdousi M, et al. An unbiased stereological method for corneal confocal microscopy in patients with diabetic polyneuropathy. *Sci Rep* 2020; 10: 12550.
13. Tavakoli M, Ferdousi M, Petropoulos IN, et al. Normative values for corneal nerve morphology assessed using corneal confocal microscopy: a multinational normative data set. *Diabetes Care* 2015; 38: 838–843.
14. Badian RA, Ekman L, Pripp AH, et al. Comparison of novel wide-field in vivo corneal confocal microscopy with skin biopsy for assessing peripheral neuropathy in type 2 diabetes. *Diabetes* 2023; 72: 908–917.
15. American Diabetes Association Professional Practice Committee. 2. Diagnosis and classification of diabetes: standards of care in diabetes-2024. *Diabetes Care* 2024; 47: S20–S42.
16. Wang M, Zhang C, Zuo A, et al. Diagnostic utility of corneal confocal microscopy in type 2 diabetic peripheral neuropathy. *J Diabetes Investig* 2021; 12: 574–582.
17. Qiao Q, Cao J, Xue W, et al. Deep learning-based automated tool for diagnosing diabetic peripheral neuropathy. *Digital Health* 2024; 10: 20552076241307573.
18. Tesfaye S, Boulton AJ, Dyck PJ, et al. Diabetic neuropathies: update on definitions, diagnostic criteria, estimation of severity, and treatments. *Diabetes Care* 2010; 33: 2285–2293.

19. Chen X, Graham J, Dabbah MA, et al. An automatic tool for quantification of nerve fibers in corneal confocal microscopy images. *IEEE Trans Biomed Eng* 2017; 64: 786–794.
 20. Perkins BA, Lovblom LE, Bril V, et al. Corneal confocal microscopy for identification of diabetic sensorimotor polyneuropathy: a pooled multinational consortium study. *Diabetologia* 2018; 61: 1856–1861.
 21. Tummanapalli SS, Issar T, Kwai N, et al. Association of corneal nerve loss with markers of axonal ion channel dysfunction in type 1 diabetes. *Clinical Neurophys* 2020; 131: 145–154.
 22. Giannaccare G, Bernabei F, Pellegrini M, et al. Bilateral morphometric analysis of corneal sub-basal nerve plexus in patients undergoing unilateral cataract surgery: a preliminary in vivo confocal microscopy study. *Br J Ophthalmol* 2021; 105: 174–179.
 23. Schaldemose EL, Fontain FI, Karlsson P, et al. Improved sampling and analysis of images in corneal confocal microscopy. *J Microsc* 2017; 268: 3–12.
 24. LeCun Y, Bengio Y and Hinton G. Deep learning. *Jn* 2015; 521: 436–444.
 25. Badian RA, Allgeier S, Scarpa F, et al. Wide-field mosaics of the corneal subbasal nerve plexus in Parkinson's disease using in vivo confocal microscopy. *Sci Data* 2021; 8: 306.
 26. Li G, Li T, Li F, et al. Nervestitcher: corneal confocal microscope images stitching with neural networks. *Comput Biol Med* 2022; 151: 106303.
-