

Automated Segmentation and Quantitative Analysis of Cotton Fiber Cross Sections Using a Deep Learning-Based Workflow

Shalini P. Etukuri

Clemson University

Chaney L. Courtney

Clemson University

Conner Rhoden

Clemson University

Matthew Laufer

Clemson University

Riley Main

Clemson University

Ryan A. Smith

Clemson University

Damon Kenny

Clemson University

Khushi Patel

Clemson University

Jacob Johnson

Clemson University

Helan Baby Thomas

Clemson University

Neha Kothari

Cotton Incorporated

Brooke Shumate

Cotton Incorporated

Vikki B. Martin

Cotton Incorporated

Trevor W. Rife

Clemson University

Christopher A. Saski

`saski@clemson.edu`

Clemson University

Method Article

Keywords: Cotton fiber, Automated image analysis, Deep learning, Instance segmentation, YOLO11, Reproducible workflow, Plant phenotyping

Posted Date: August 15th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10594498/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Cross-sectional analysis is considered the reference method for measuring cotton fiber fineness and maturity; however, its widespread use has been limited by labor-intensive sample preparation and manual image analysis. The objective of this study was to develop and validate a reproducible deep-learning-based workflow for automated segmentation and quantitative analysis of cotton fiber cross-sections. A total of 249 composite light microscopy images of cotton fiber cross-sections were collected and manually annotated to generate training and validation datasets. A YOLO11m instance segmentation model was developed to identify cotton fiber and lumen regions and automatically extract quantitative traits, including fiber area, lumen area, fiber perimeter, and lumen perimeter. The workflow integrates automated image segmentation, post-processing, quantitative trait extraction, and data export to facilitate reproducible cotton fiber phenotyping. Model performance was evaluated using mean Average Precision (mAP), and workflow outputs were validated against Adobe Photoshop using descriptive comparisons of six cross-sectional traits. The model achieved Box mAP50 scores of 0.984 for cotton fiber regions and 0.789 for lumen regions, demonstrating high segmentation accuracy. The automated workflow substantially reduced manual analysis time while producing measurements with central tendencies comparable to those obtained using Adobe Photoshop. To facilitate reproducibility and adoption, the workflow, trained model weights, and supporting documentation are publicly available through GitHub and a Hugging Face web application. The workflow substantially increases analytical throughput while providing a reproducible and publicly accessible method for automated cotton fiber cross-sectional phenotyping, facilitating quantitative analysis for cotton genetics and breeding research.

Plain English Summary

Cotton fiber quality affects how cotton performs during textile manufacturing. Two important fiber traits, fineness and maturity, are most accurately measured from fiber cross-sections, but traditional image analysis is slow and requires extensive manual effort. In this study, we developed an automated deep-learning-based workflow to analyze microscope images of cotton fiber cross-sections. The workflow identifies the outer fiber boundary and the inner hollow region (lumen), then automatically extracts quantitative measurements such as fiber area, lumen area, and perimeter. Compared with manual image analysis, the workflow substantially reduces analysis time while providing standardized measurements. The software is publicly available through a web application and GitHub repository, making automated cotton fiber phenotyping more accessible for genetics and breeding research.

1 Background

Improving cotton fiber fineness remains a significant challenge in fiber quality research, in part because precise, high-throughput phenotyping methods are still limited (1). Among cultivated cotton species, *Gossypium barbadense* (Pima cotton) produces some of the finest fibers and is valued for textiles because of its soft, silky, and smooth feel. In contrast, *Gossypium hirsutum* (Upland cotton) generally has shorter, coarser fibers and a texture distinct from Pima (2). These intrinsic differences in fiber quality translate into economic impacts, with Pima-based fabrics typically commanding higher prices than those made from

Upland cotton (3). While Pima cotton is known for its superior quality, it accounts for only 3% of U.S. cotton output because production is largely limited to semi-arid and arid environments (3). In contrast, Upland cotton accounts for approximately 97% of U.S. production, as it is broadly adapted to diverse growing regions and has a shorter growing season. Therefore, improving fiber quality in this species is a priority for stakeholders across the entire value chain. (4).

However, improving fiber quality is challenging because key traits are highly polygenic and strongly influenced by the environment, limiting selection efficiency in breeding programs (5). A persistent bottleneck for genetic improvement has been phenotyping, as fiber fineness is difficult to measure at the scale and throughput needed for breeding. Micronaire is the most widely used industry metric for cotton fiber quality and is measured using air permeability through a compressed sample (6). Because it indirectly measures both fineness and maturity, low micronaire values can indicate either immature or fine fibers, while high micronaire values may reflect either mature or coarse fibers (7–9). This ambiguity can lead to misclassification and economic penalties, including cases where producers receive discounted prices for cotton that is fine and mature but categorized as low micronaire (10).

Cross-sectional analysis is the most accurate technique to directly measure cotton fiber fineness and maturity, and can serve as a reference method for calibrating other indirect fiber quality metrics (11, 12). Despite these advantages, cross-sectional methods are not widely adopted because they require labor intensive sample preparation, time consuming image analysis, and a high level of expertise to ensure accurate sectioning, imaging, and data interpretation (12). Existing workflows rely on a combination of computer algorithms and manual or semi-manual image processing tools such as ImageJ, Adobe Photoshop, and Fiber Image Analysis System (FIAS) (12–14). ImageJ- and Photoshop-based measurements are particularly slow and can introduce operator bias and inconsistency, leading to highly variable results. FIAS can improve accuracy, but the workflow remains tedious and is not well suited for rapid, high-throughput assessment across large breeding populations (13).

To address these limitations, we developed an automated deep-learning-based workflow for segmentation and quantitative analysis of cotton fiber cross-sections. The workflow identifies cotton fiber and lumen regions from light microscopy images and automatically extracts quantitative traits including fiber area, fiber perimeter, lumen area, and lumen perimeter. In addition to image segmentation, the workflow performs automated trait extraction and exports results in a standardized format for downstream analysis. The workflow, trained model weights, and documentation are publicly available to support reproducible cotton fiber phenotyping and facilitate genetic and breeding studies involving fiber fineness and maturity.

2 Methods

2.1 Imaging, Annotation, and Ground Truth

Cotton fiber samples used in this study were obtained from upland cotton (*Gossypium hirsutum* L.) genotypes U1 and Siokra-250. These parental lines were selected because they are being used to develop a recombinant inbred line (RIL) population for genetic studies of fiber quality traits. Plants were grown under greenhouse conditions, and mature fibers were harvested for cross-sectional imaging and analysis.

Samples were prepared following (13). Briefly, 0.1 g of mature cotton fiber was embedded into a gelatin capsule using uncatalyzed London Resin (LR) White with accelerator. The Leica RM2265 glass microtome was used to cut the capsule containing fiber samples into 6 μm sections. Sections were transferred to a microscope slide for imaging. Cross-sectional images of cotton fibers were collected using a LEXT OLS5100 Olympus optical profiler at the Clemson Light Imaging Facility (CLIF). Fiber sections were imaged at 100x magnification (Fig. 1), and four images were stitched together to generate a larger high-resolution image (Supplementary Fig. 1).

Figure 1. Preparation and imaging workflow for cotton fiber cross-sections. (A) Cotton fiber sample and embedded resin capsule. (B) Sectioning of embedded samples using a Leica RM2265 glass microtome. (C) Image acquisition using a LEXT OLS5100 optical profiler. (D) Representative light microscopy image of cotton fiber cross-sections used for annotation, model training, and automated trait extraction.

A total of 249 composite images were collected and divided into training and validation datasets. Of these, 192 images were used for model training, and 57 images were used for validation. Two semantic classes were defined for annotation: cotton (fiber tissue) and lumen (internal cavity) (Fig. 2A). Manual annotations were performed using the VGG Image Annotator (VIA), an open-source image annotation tool that enables users to define regions in an image and add textual descriptions of those regions (15). In the cotton fiber cross-sectional images, polygonal regions were manually drawn around the perimeter of both the cotton fiber and the lumen (13). The annotated cotton and lumen regions were labeled into two separate classes and assigned a unique identifier for model training (Fig. 2B). Polygon coordinates were exported for each image and converted into contours for training and for reference measurements during validation. These manually annotated regions served as the ground truth for model training and validation and enabled extraction of quantitative measurements from fiber cross-sections.

Figure 2. Ground-truth annotation workflow for cotton fiber cross-sections used in model training and validation. (A) Representative light microscopy image of cotton fiber cross-sections. The inset highlights an individual fiber cross-section, with the cotton fiber boundary shown in cyan and the lumen boundary shown in orange. (B) Manual annotation of cotton fiber and lumen boundaries using the VGG Image Annotator (VIA). Annotated regions were assigned unique identifiers and used as ground truth for model training and validation.

2.2 Model Development

The cotton-lumen analysis pipeline was developed using the Ultralytics YOLO11m instance segmentation model for automated detection and segmentation of cotton fiber and lumen regions (16, 17). The model was initialized with pretrained weights and trained on manually annotated cotton fiber cross-sectional images. The YOLO11m architecture consists of approximately 464 layers with 56.8 million trainable parameters. Training was performed using 640×640 -pixel input images with a batch size of 51 and required approximately 47.3 GB of GPU memory on an NVIDIA A100 GPU. The final serialized model weight file was approximately 114.4 MB. The model was used to simultaneously perform object detection and instance segmentation, enabling identification of individual cotton fiber and lumen regions for subsequent quantitative analysis. The model was trained using MuSGD optimizer with an initial learning rate of 0.01 and

momentum 0.9. Training used default data augmentation and early stopping was applied with a patience of 100 epochs.

2.3 Model Evaluation and Post-processing

Model performance was evaluated using Box mean Average Precision at an intersection-over-union threshold of 0.50 (Box mAP50) to assess detection accuracy (18). A total of 57 composite images were used to evaluate model performance. Class-level recall and precision were summarized using a normalized confusion matrix. The model generated bounding boxes, class confidence scores, and segmentation masks for detected regions. Segmentation masks were converted into contour coordinates for subsequent measurements. To reduce spurious detections, predicted contours were filtered by size: objects larger than one-quarter of the total image area were excluded; cotton contours were required to have an area greater than 100 pixels; and lumen contours were required to have an area greater than 50 pixels.

2.4 Geometric Measurements

For each detected object, the segmentation mask was converted to a binary image, and contour complexity was reduced using OpenCV's `approxPolyDP` function. Two image analysis approaches adapted from previous studies were used to estimate the primary (length) and secondary (width) axes, and to robustly estimate the minor axis of irregular fiber cross-sections (19, 20). This approach was used because cotton fiber cross-sections frequently deviate from ideal circular or elliptical geometries. Fiber morphology varies with maturity. Immature fibers typically have more rounded cross-sections, whereas mature fibers often develop collapsed or kidney-shaped cross-sections. The major axis was defined as the maximum distance between any two contour vertices. Minor axis length was determined from the widest segment oriented approximately perpendicular to the major axis. Segments within $1-2^\circ$ of orthogonality were considered candidates, and the widest candidate segment was selected as the minor axis.

2.5 Pixel to Micrometer Unit Conversion

To convert pixel-based measurements to metric units, the pipeline automatically identified the yellow 40 μm scale bar present in each image. The source image was converted to HSV color space, and a threshold corresponding to the yellow hue range was applied to generate a binary mask. Yellow contours were detected, and the largest yellow object was assumed to correspond to the scale bar. The pixel length of the detected scale bar was measured, and a conversion factor was calculated from the known scale bar length (40 μm). This factor was applied to convert all pixel-based measurements to micrometers. This method assumes that a consistent yellow scale bar is present in every image and that the largest yellow contour corresponds to it.

2.6 Lumen–Cotton Pairing

A deterministic containment check was used to pair lumen contours with their corresponding cotton contours. The algorithm iterated over each cotton contour and checked whether any lumen polygon was spatially contained within it. A bounding-rectangle containment heuristic was used to improve computational efficiency. If no lumen was detected within a cotton contour, the fiber was recorded as lacking an associated lumen. If no lumen was found, the cotton contour was recorded as having an absent

lumen. This pairing strategy enabled calculation of lumen-to-cotton ratios and ensured that paired measurements were reported together in the output files.

2.7 Output and Reporting

For each analyzed image, the workflow generated a comma-separated value (CSV) file containing measurements for each detected object, including object class (cotton or lumen), area, perimeter, bounding-box dimensions, and spatial coordinates. In addition to tabular outputs, segmented overlays and individual PNG files of detected fiber cross-sections were generated to facilitate quality control and visual verification. Overlay visualizations displayed cotton and lumen regions in distinct colors to aid interpretation of segmentation results.

The workflow exported primary geometric measurements, including cotton area, lumen area, cotton perimeter, and lumen perimeter, for each detected fiber. During subsequent data analysis, true fiber area was calculated as the difference between cotton area and lumen area, representing the cross-sectional area of the fiber cell wall. Fiber fineness and maturity ratio were then calculated from the exported measurements using the equations reported in (21).

2.8 Web Application Deployment

The cotton-lumen pipeline was deployed as a web application hosted on Hugging Face Spaces (22). The application allows users to upload light microscopy images, perform automated segmentation and measurement, and export results as CSV files without requiring machine learning expertise (Supplementary Fig. 2). The web interface improves accessibility and facilitates use of the workflow by researchers without specialized computational training. Under the tested conditions, image analysis was completed in less than 30 seconds per image.

2.9 Photoshop-Based Measurements

To validate measurements generated by the automated workflow, approximately 400 cotton fiber cross-sections from the parental genotypes U1 and Siokra-250 were manually analyzed in Adobe Photoshop following the protocol described by (13). Briefly, fiber and lumen regions were manually delineated using layer-based masks, and measurements were extracted using the Photoshop measurement tool after calibration with the image scale bar. Separate masks were generated for the fiber and lumen of each cross-section, allowing calculation of area- and perimeter-based traits. Measurements were exported and compiled for comparison with output generated by the automated workflow.

3 Results

3.1 Model Performance and Evaluation

Model performance was evaluated using mean Average Precision (mAP), a standard object detection or segmentation metric used in computer vision tasks (Table 1). We report Box mAP at an intersection-over-union (IoU) threshold of 0.50 (Box mAP50), which quantifies how accurately the model detects and

classifies, and localizes objects by measuring the average precision of predicted bounding boxes that overlap the ground truth by at least 50%. The maximum possible score is 1.0, indicating that all objects in the validation dataset were correctly detected, classified, and localized.

The model achieved a Box mAP50 score of 0.984 for cotton and 0.789 for lumen. The lower performance observed for lumen regions is likely attributable to their smaller size, irregular morphology, and lower contrast relative to surrounding fiber tissue. In addition, lumen boundaries can be difficult to distinguish in immature fibers or partially collapsed cross-sections, increasing annotation and segmentation uncertainty.

Table 1
Summary of standard object detection for validation data.

Class	Images	Instances	Box				Mask			
			P	R	mAP50	mAP50-95	P	R	mAP50	mAP50-95
All	192	7499	0.88	0.85	0.887	0.64	0.80	0.76	0.77	0.44
Cotton		5110	0.95	0.96	0.984	0.86	0.94	0.95	0.97	0.78
Lumen		2389	0.81	0.73	0.789	0.42	0.66	0.58	0.56	0.16

P = precision; R = recall; mAP50 = mean Average Precision at an intersection-over-union threshold of 0.50; mAP50–95 = mean Average Precision averaged across IoU thresholds from 0.50 to 0.95.

Further evaluation using a normalized confusion matrix on the manually annotated validation dataset revealed class-level performance across cotton, lumen, and background classes (Fig. 4). The confusion matrix showed that most classification errors occurred between lumen and background regions, whereas cotton regions were rarely confused with other classes. Misclassification was most common between lumen and background regions, likely because lumen structures are substantially smaller than cotton fiber regions and often exhibit low contrast against surrounding tissue. In some cases, lumen regions exhibited grayscale intensity values similar to the background, making boundary detection more difficult. Additional training data and targeted augmentation of lumen examples may improve lumen detection performance in future versions of the workflow.

3.2 Comparison of Automated Workflow and Adobe Photoshop Measurements:

To compare measurements generated by the automated workflow with manual measurements, descriptive statistics were calculated for six cotton fiber cross-sectional traits in the parental genotypes Siokra-250 and U1 (Supplementary Table 1). Adobe Photoshop was used to measure a representative subset of approximately 400 fibers per genotype, whereas the automated workflow analyzed all detectable fibers from the same 18 composite images (845 fibers for Siokra-250 and 1,019 fibers for U1). Because the datasets

were not paired, comparisons focused on descriptive statistics and distributional characteristics rather than paired agreement analyses.

Across both genotypes, measurements generated by the automated workflow showed central tendencies comparable to those obtained using Adobe Photoshop (Supplementary Table 1; Fig. 5). Mean values and median estimates were generally similar across all six traits, while the automated workflow exhibited greater variability for several measurements, particularly cotton area, true fiber area, and fiber fineness. These broader distributions likely reflect the larger number of fibers analyzed by the automated workflow compared with the representative manually measured subsets.

In addition to producing measurements comparable to those obtained using Adobe Photoshop, the automated workflow substantially reduced the time required for image analysis. Adobe Photoshop required more than 20 h to manually measure approximately 400 fiber cross-sections per genotype. Under the tested conditions, the YOLO11m model completed inference in approximately 8.8 ms per composite image. Following segmentation, automated measurement extraction required, on average, less than 0.1 s per detected fiber, substantially reducing the manual effort required for quantitative image analysis and trait extraction.

4 Discussion

4.1 Automated Cotton Fiber Phenotyping:

Cross-sectional analysis is considered one of the most accurate methods for measuring cotton fiber fineness and maturity; however, its widespread use has been limited by labor-intensive sample preparation and time-consuming manual image analysis (12, 13). Traditional Adobe Photoshop-based measurements require manual delineation of individual fiber and lumen boundaries, making large-scale phenotyping impractical and increasing the potential for operator-dependent variability. The automated workflow developed in this study substantially reduces manual effort while providing standardized quantitative measurements of cotton fiber cross-sections. By integrating image segmentation, quantitative trait extraction, and automated data export into a single workflow, the method minimizes user intervention and improves the reproducibility of cross-sectional phenotyping. The comparable central tendencies observed between automated and manual measurements, together with the substantial increase in analytical throughput, demonstrate that the workflow can efficiently generate biologically consistent quantitative measurements while reducing the labor associated with traditional image analysis. Standardized automated measurements also facilitate consistent phenotypic evaluation across experiments and research groups. Collectively, these improvements make detailed cotton fiber characterization more practical for genetic studies and breeding programs requiring measurements from hundreds or thousands of fiber cross-sections.

4.2 Challenges in Lumen Segmentation:

Lumen segmentation has been one of the most challenging aspects of this project. In comparison, cotton fiber boundaries were detected more accurately because they are larger and more distinct than lumen

regions. Lumen segmentation is more difficult because the lumen occupies only a small portion of the fiber cross-section and, in some cases, exhibits grayscale intensity similar to the surrounding background, making it difficult for the model to distinguish the lumen boundary. Fiber maturity also influences lumen detection. Immature fibers typically have a larger and more visible lumen, whereas mature fibers develop thicker secondary cell walls that reduce the lumen to a narrow or partially collapsed central cavity (23, 24). These biological differences likely contributed to the lower lumen detection performance observed in this study. Future improvements may be achieved by incorporating additional lumen-focused annotations and augmentation strategies during model training.

4.3 Applications for Cotton Phenotyping and Breeding:

Accurate phenotypic data is essential for genetic studies of cotton fiber quality. The automated workflow developed in this study provides quantitative measurements of individual fiber cross-sections that can be readily used for quantitative trait locus (QTL) mapping and other genetic analyses of fiber fineness and maturity (2). Because the workflow substantially reduces the time required for cross-sectional image analysis, it is well suited for phenotyping recombinant inbred line (RIL) populations and other breeding materials where large numbers of samples must be evaluated. In addition, cross-sectional measurements generated by this workflow can serve as reference data for calibration and validation of widely used fiber quality assessment systems, including the High Volume Instrument (HVI) and the Advanced Fiber Information System (AFIS) (6, 7). Collectively, these applications may improve the accuracy and consistency of phenotypic data used in cotton genetics and breeding research.

4.4 Future Improvements:

Although the proposed workflow substantially improves the efficiency of cotton fiber cross-sectional analysis, opportunities remain for further refinement. Future work should focus on expanding the training dataset to include additional cotton genotypes, fiber maturity levels, and imaging conditions to improve model robustness. Increasing the diversity of annotated lumen examples may further improve lumen segmentation accuracy, particularly for mature fibers with narrow or partially collapsed lumens. Refinement of post-processing procedures may also reduce edge-related detections and improve measurement accuracy for challenging fiber cross-sections. Continued development and validation of this publicly available workflow will further increase its utility for cotton phenotyping, genetic studies, and breeding applications.

5 Conclusion

We developed and validated a reproducible deep-learning-based workflow for automated segmentation and quantitative analysis of cotton fiber cross-sections. The workflow accurately identified cotton fiber and lumen regions while substantially reducing the time required for manual image analysis. By integrating image segmentation, quantitative trait extraction, and standardized data export into a single pipeline, the proposed method provides a practical approach for reproducible cotton fiber phenotyping. Public availability of the workflow through a web application and GitHub repository facilitates its adoption by the broader research community. This workflow provides a foundation for future genetic studies, cotton breeding

applications, and continued development of automated image analysis methods for fiber quality assessment.

Declarations

Acknowledgments

The authors gratefully acknowledge the Clemson Light Imaging Facility for assistance with microscope usage, the Histology Core Facility at Clemson University for support with the Leica glass microtome, and the Clemson University Palmetto Cluster for training and user support.

Authors' contributions

C.A.S. conceived the study, supervised the research, acquired project funding, and substantially revised the manuscript. **S.P.E.** prepared the cotton fiber samples, collected the microscopic images, contributed to the annotation of the training dataset using the VGG Image Annotator (VIA), curated the dataset, analyzed the data, and drafted the manuscript. **C.L.C.** developed and trained the deep learning model, implemented the software workflow, and substantially revised the manuscript. **C.R., M.L., R.M., R.S., D.K., K.P., J.J., and H.B.T.** contributed to the annotation of the training dataset using the VGG Image Annotator (VIA). **T.W.R.** supervised the research, acquired project funding, and substantially revised the manuscript. **N.K.** conceived the study, participated in study design, and critically reviewed the manuscript. **V.B.M.** and **B.S.** provided scientific guidance and critically reviewed the manuscript. All authors read and approved the final manuscript.

Availability of data and materials

The cotton fiber image analysis workflow is publicly available through a web-based application hosted on Hugging Face at: https://huggingface.co/spaces/chaneylc/cotton_fiber_microscopy_measure

The complete source code, training scripts, dataset configuration, example input images, example outputs, and supporting documentation are publicly available through the GitHub repository:

<https://github.com/RifeLab/cotton-lumen-micro>

Funding

This research was funded by Cotton Incorporated, grant number 22-273, the USDA AFRI program, grant number SC-2023-07687, and FFAR Fellows Program.

Conflict of interest

The authors declare no conflicts of interest.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Orcid

Shalini Etukuri: 0009-0002-5290-7280

Chaney Courtney: 0009-0004-7273-1280

Conner Rhoden: 0009-0007-5854-4747

Matthew Laufer: 0009-0009-4219-7546

Ryan Smith: 0009-0001-9474-9746

Damon Kenny: 0009-0002-0828-2623

Jacob Johnson: 0009-0006-1734-9147

Helan Thomas: 0000-0002-4870-1442

Brooke Shumate: 0009-0004-2720-1328

Vikki Martin: 0009-0005-9683-3553

Trevor Rife: 0000-0002-5974-6523

Chrsitopher Saski: 0000-0002-2780-4274

References

1. 3D Microscopy for high-throughput phenotyping of cotton fibers. National Institute of Food and Agriculture; University of North Texas 2024.
2. Lacape J-M, Llewellyn D, Jacobs J, Arioli T, Becker D, Calhoun S, et al. Meta-analysis of cotton fiber quality QTLs across diverse environments in a *Gossypium hirsutum* x *G. barbadense* RIL population. *BMC Plant Biol.* 2010;10(1):132.
3. Cotton. and Wool – Cotton Sector at a Glance. United States Department of Agriculture Economic Research Service; Meyer and Dew; 2025.
4. Han JM, Flexas J, Xiong DL, Galmés J, Zhang YL. Regulation of photosynthesis and water-use efficiency in pima and upland cotton species subjected to drought and recovery. *Photosynthetica.* 2024;62(1):6–15.
5. Waghmare VN. Cotton Breeding. In: Yadava DK, Dikshit HK, Mishra GP, Tripathi S, editors. *Fundamentals of Field Crop Breeding.* Singapore: Springer Nature Singapore; 2022. pp. 609–76.

6. Rodgers J, Zumba J, Fortier C. Measurement comparison of cotton fiber micronaire and its components by portable near infrared spectroscopy instruments. *Text Res J.* 2017;87(1):57–69.
7. Zumba J, Rodgers J, Indest M. Fiber micronaire, fineness, and maturity predictions using NIR spectroscopy instruments on seed cotton and cotton fiber, in and outside the laboratory. 2017.
8. Hsieh Y-L. Structural development of cotton fibers and linkages to fiber quality. *Cotton fibers: CRC;* 2024. pp. 137–65.
9. May OL. Genetic variation in fiber quality. *Cotton fibers: CRC;* 2024. pp. 183–229.
10. Gordon S. Cotton fibre quality. *Cotton: Elsevier;* 2007. pp. 68–100.
11. Adedoyin AA, Li C, Toews MD. Characterization of single cotton fibers using a laser diffraction system. *Text Res J.* 2011;81(4):355–67.
12. Xu B, Huang Y. Image analysis for cotton fibers part II: cross-sectional measurements. *Text Res J.* 2004;74(5):409–16.
13. LaFave Q, Etukuri SP, Courtney CL, Kothari N, Rife TW, Saski CA. A Simplified Microscopy Technique to Rapidly Characterize Individual Fiber Traits in Cotton. *Methods Protocols.* 2023;6(5):92.
14. Lamichhane S. An evaluation of cotton fiber cross-sections with an image analysis software 2016.
15. Dutta A, Zisserman A, editors. The VIA annotation software for images, audio and video. *Proceedings of the 27th ACM international conference on multimedia;* 2019.
16. Baldovino RG, Vidad AJP, Abastillas RPB, Bugtai NT, Dadios EP, Vicerra RRP, et al. Comprehensive analysis on Ultralytics-supported YOLO models for detection and recognition of large office objects for indoor navigation. *Procedia Comput Sci.* 2024;246:3851–8.
17. Khanam R, Hussain M. Yolov11: An overview of the key architectural enhancements. *arXiv preprint arXiv:241017725.* 2024.
18. Everingham M, Van Gool L, Williams CKI, Winn J, Zisserman A. The Pascal Visual Object Classes (VOC) Challenge. *Int J Comput Vision.* 2010;88(2):303–38.
19. Brainard SH, Bustamante JA, Dawson JC, Spalding EP, Goldman IL. A digital image-based phenotyping platform for analyzing root shape attributes in carrot. *Front Plant Sci.* 2021;12:690031.
20. Tanabata T, Shibaya T, Hori K, Ebana K, Yano M, SmartGrain. High-Throughput Phenotyping Software for Measuring Seed Shape through Image Analysis. *Plant Physiol.* 2012;160(4):1871–80.
21. Hequet EF, Wyatt B, Abidi N, Thibodeaux DP. Creation of a set of reference material for cotton fiber maturity measurements. *Text Res J.* 2006;76(7):576–86.
22. Wolf T, Debut L, Sanh V, Chaumond J, Delangue C, Moi A, et al. editors. *Transformers: State-of-the-Art Natural Language Processing* 2020 October; Online: Association for Computational Linguistics.
23. Smith B. A review of the relationship of cotton maturity and dyeability. *Text Res J.* 1991;61(3):137–45.
24. Long RL, Bange MP. Consequences of immature fiber on the processing performance of Upland cotton. *Field Crops Res.* 2011;121(3):401–7.

Figures



Figure 1A - Cotton fiber sample and embedded capsule



Figure 1B - Leica RM2265 glass microtome

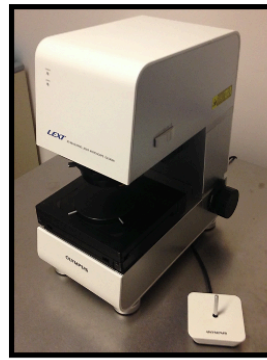


Figure 1C - LEXT OLS5100 optical profiler

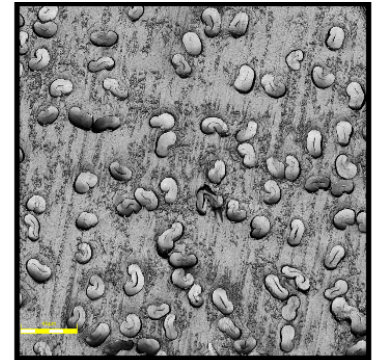


Figure 1D - Representative cotton fiber cross-sectional image

Figure 1

Preparation and imaging workflow for cotton fiber cross-sections. (A) Cotton fiber sample and embedded resin capsule. (B) Sectioning of embedded samples using a Leica RM2265 glass microtome. (C) Image acquisition using a LEXT OLS5100 optical profiler. (D) Representative light microscopy image of cotton fiber cross-sections used for annotation, model training, and automated trait extraction.

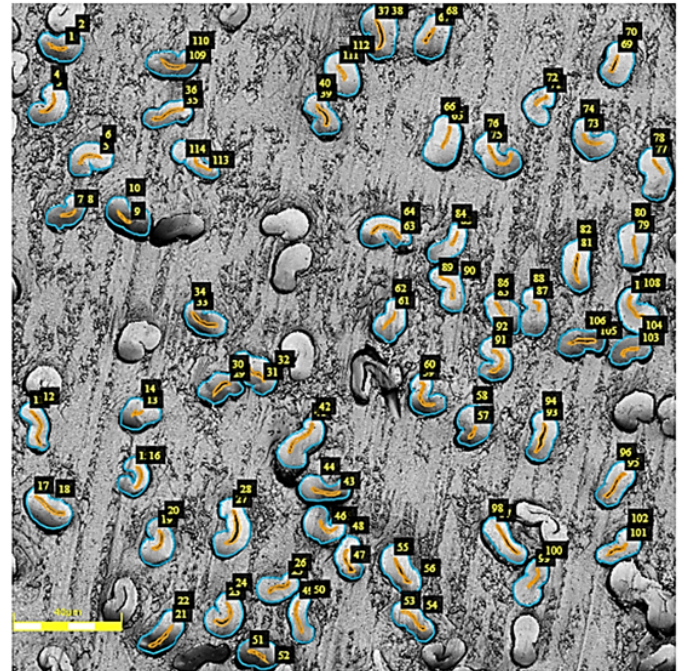
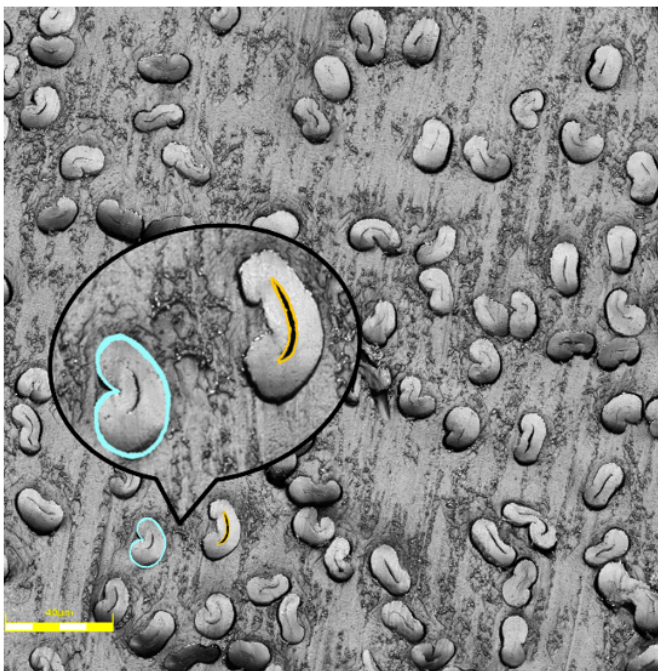


Figure 2

Ground-truth annotation workflow for cotton fiber cross-sections used in model training and validation. (A) Representative light microscopy image of cotton fiber cross-sections. The inset highlights an individual fiber cross-section, with the cotton fiber boundary shown in cyan and the lumen boundary shown in orange. (B)

Manual annotation of cotton fiber and lumen boundaries using the VGG Image Annotator (VIA). Annotated regions were assigned unique identifiers and used as ground truth for model training and validation.

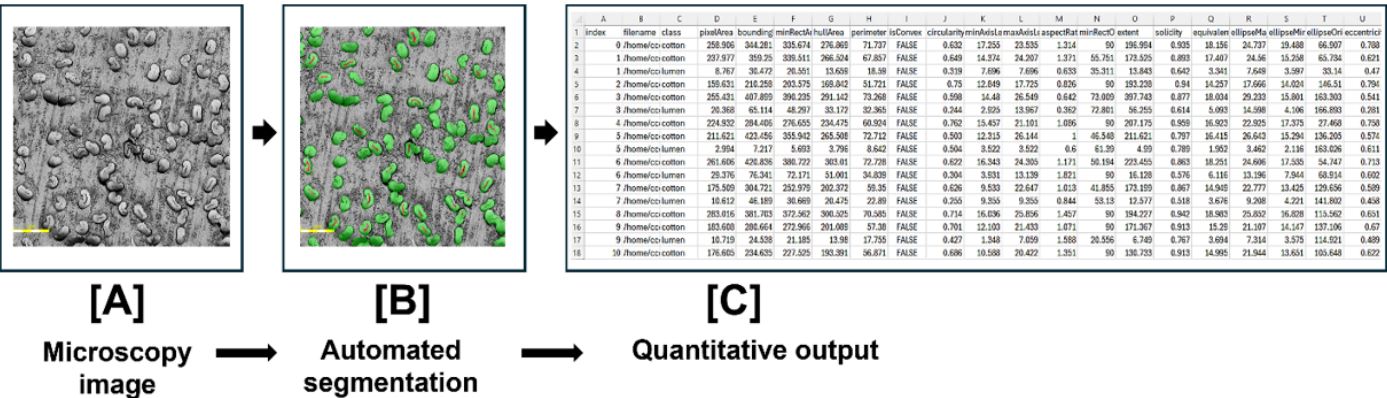


Figure 3

Computational workflow for automated segmentation and quantitative analysis of cotton fiber cross-sections. (A) Representative microscopy image used as input to the workflow. (B) Automated segmentation of cotton fiber and lumen regions. (C) Example CSV output containing object-level measurements generated by the pipeline.

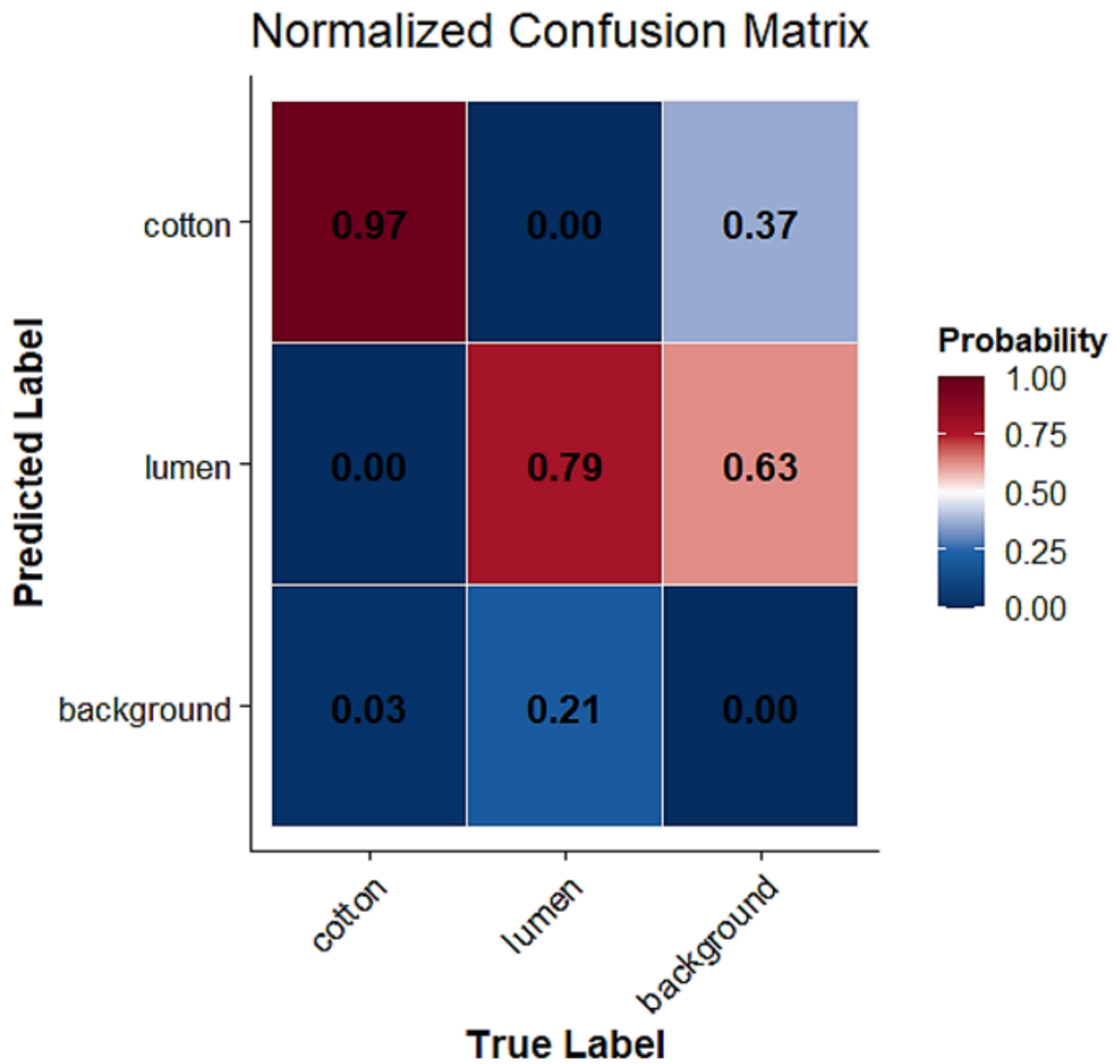


Figure 4

Normalized confusion matrix for the validation dataset showing classification performance across cotton, lumen, and background classes.

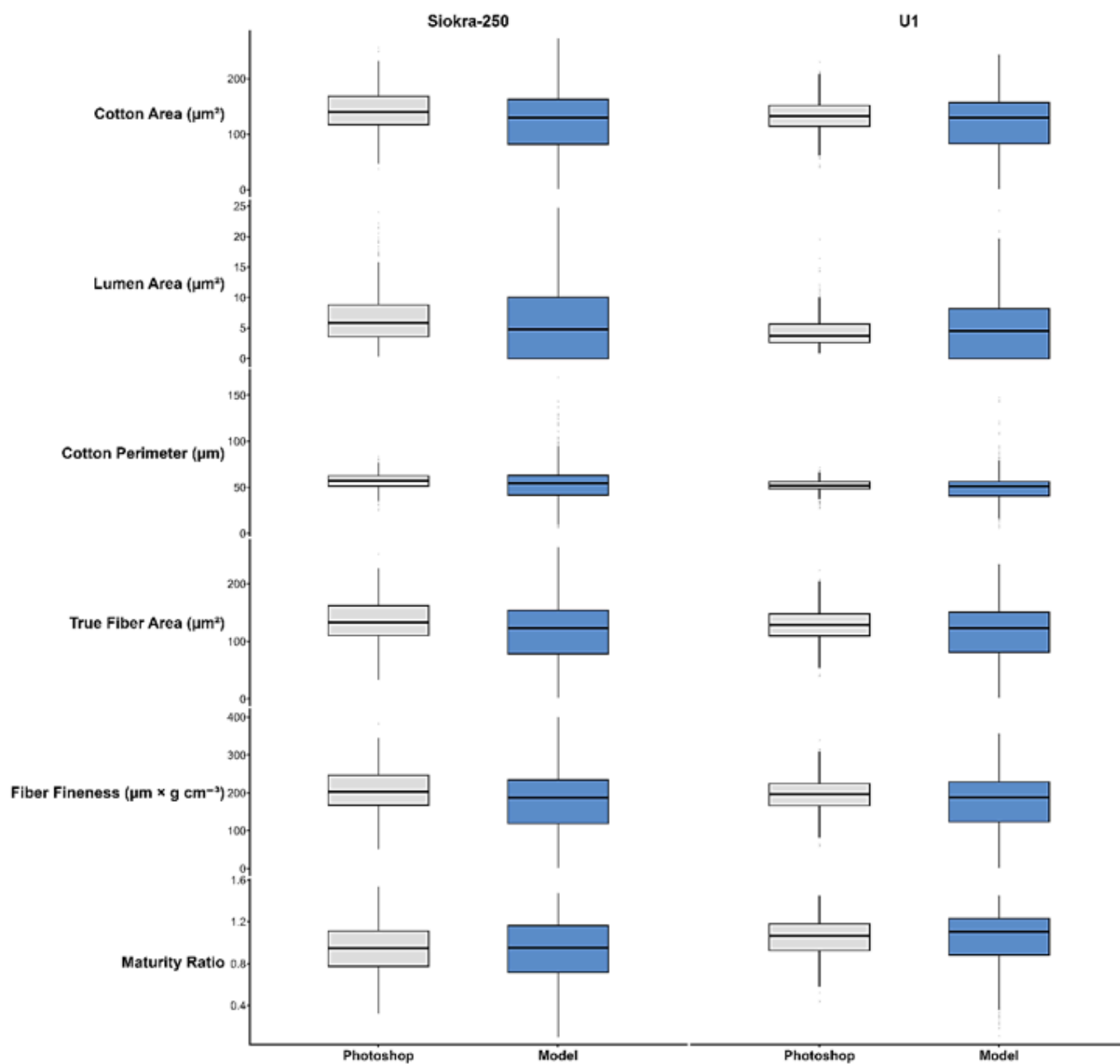


Figure 5

Distribution of cotton fiber cross-sectional traits measured using Adobe Photoshop (gray) and the automated YOLO11m workflow (blue) for the Siokra-250 and U1 genotypes. Boxplots display the median, interquartile range, whiskers, and outliers, illustrating comparable central tendencies while highlighting the broader distributions associated with the larger number of fibers analyzed by the automated workflow.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryData.docx](#)