International Journal of Laboratory Hematology

WILEY

ISLH International Journal of Laboratory Hematology

ORIGINAL ARTICLE

Automated Von Willebrand Factor Multimer Image Analysis for Improved Diagnosis and Classification of Von Willebrand Disease

Karthik Anand[1] | Vincent Olteanu[1] | Chi Zhang[2] | Katelynn Nelton[2] | Erin Aakre[2] | Juliana Perez Botero[2] | Rajiv Pruthi[2] | Dong Chen[2] | Jansen N. Seheult[2]

[1]Mayo High School, Rochester, Minnesota, USA | [2]Special Coagulation Laboratory, Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota, USA

Correspondence: Jansen N. Seheult (seheult.jansen@mayo.edu)

Received: 26 October 2024 | **Revised:** 9 January 2025 | **Accepted:** 14 February 2025

Funding: The authors received no specific funding for this work.

Keywords: laboratory automation | von Willebrand disease | von Willebrand factor

ABSTRACT

Introduction: Von Willebrand factor (VWF) multimer analysis is essential for diagnosing and classifying von Willebrand disease (VWD) but requires expert interpretation and is subject to inter-rater variability. We developed an automated image analysis pipeline using deep learning to improve the reproducibility and efficiency of VWF multimer pattern classification.

Methods: We trained a YOLOv8 deep learning model on 514 gel images (6168 labeled instances) to classify VWF multimer patterns into 12 classes. The model was validated on 192 images (2304 instances) and tested on an independent set of 94 images (1128 instances). Images underwent preprocessing, including histogram equalization, contrast enhancement, and gamma correction. Two expert raters provided ground truth classifications.

Results: The model achieved 91% accuracy compared to Expert 1 (macro-averaged precision = 0.851, recall = 0.757, F1-score = 0.786) and 87% accuracy compared to Expert 2 (macro-averaged precision = 0.653, recall = 0.653, F1-score = 0.641). Inter-rater agreement was very high between experts ($\kappa = 0.883$), with strong agreement between the model and Expert 1 ($\kappa = 0.845$) and good agreement with Expert 2 ($\kappa = 0.773$). The model performed exceptionally well on common patterns (F1 > 0.93) but showed lower performance on rare subtypes.

Conclusion: Automated VWF multimer analysis using deep learning demonstrates high accuracy in pattern classification and could standardize the interpretation of VWF multimer patterns. While not replacing expert analysis, this approach could improve the efficiency of expert human review, potentially streamlining laboratory workflow and expanding access to VWF multimer testing.

1 | Introduction

Von Willebrand factor (VWF) is a large multimeric glycoprotein with essential roles in primary hemostasis [1]. VWF is synthesized by endothelial cells and megakaryocytes as ultra-large molecular weight multimers and is cleaved by the plasma protease ADAMTS13 into low, intermediate, and high molecular weight multimers [2, 3]. A normal individual has a uniform distribution of all three sizes in the plasma [4].

Karthik Anand and Vincent Olteanu contributed equally.

VWF multimer analysis is essential in the diagnosis of both congenital von Willebrand disease (VWD) and acquired von Willebrand syndrome (AVWS), which can both present with mucocutaneous bleeding [5, 6]. VWF multimer analysis by either manual or semi-automated electrophoresis and immunoblotting can aid in discriminating among type 1, type 2, and type 3 VWD, as well as among subtypes of type 2 VWD [7, 8].

Despite diverse methodologies, VWF multimer analysis generally consists of five steps: (1) electrophoresis of a sample in an agarose gel, (2) in-gel fixation or transfer of the electrophoretic protein product to a membrane, (3) immunodetection of the protein using a labeled anti-VWF antibody, (4) visualization of the VWF multimer bands in the gel or membrane, and (5) interpretation or quantification of the VWF multimer bands [9, 10].

Unfortunately, due to the high level of expertise required, long manual analysis time, and high complexity, VWF multimer analysis has historically only been performed in coagulation reference laboratories [11]. There is substantial inter-rater variability in the interpretation of some VWF multimer patterns, especially the subtle loss of high molecular weight multimers in some cases of AVWS [12]. For this reason, current laboratory workflows often require the evaluation of all multimer gel lanes by two experts [13, 14].

Computer vision methods have the potential to improve the reproducibility of VWF multimer gel interpretation and improve efficiency and turnaround times in the clinical laboratory [15]. In this study, we developed an automated image analysis pipeline to detect and classify sample and normal control lanes in electrophoresis gel images.

2 | Methods

Our laboratory utilizes a multi-day protocol for von Willebrand Factor (VWF) multimer analysis using a 2% agarose separating gel. Gel images are captured on the Odyssey infrared imaging system (LI-COR Biotech, Lincoln, NE, USA). Our protocol has been optimized through validation studies to handle samples across the VWF antigen range without requiring pre-dilution normalization. A validation study specifically examining samples with elevated VWF:Ag levels showed no additional value in pre-diluting samples with the 2% SDS sample prep buffer. For standard samples, we use 10 μL sample with 5 μL bromophenol blue and 85 μL buffer, and for low VWF:Ag samples (< 30%), we use 20 μL sample with 5 μL bromophenol blue and 75 μL buffer. Post-electrophoresis, our methodology allows for adjustment of brightness and contrast in gel images to optimize band visualization across the range of VWF concentrations. Each gel image typically contains 10 patient sample lanes and two normal control lanes (positions 3 and 9); occasionally, one or more lanes may be left blank or empty. The normal control consists of pooled plasma from normal donors, prepared from our standard citrated plasma aliquots that are frozen at −70 °C shortly after processing. To ensure consistency and stability, we evaluate and establish a new normal pool every 18 months. This standardized approach to control material preparation and pooling minimizes potential variability that could affect the algorithm's performance across different control lots.

The YOLOv8 model was chosen for its state-of-the-art performance in object detection and instance segmentation tasks [16]. The model was trained to classify patient sample lanes into one of 12 classes or patterns and to identify the normal control lanes and whether the controls were acceptable or unacceptable (failed). The training set comprised 514 gel images, containing a total of 6168 labeled instances across 12 classes (Table 1). The validation set consisted of 192 gel images, containing 2304 labeled instances across 12 classes (Table 1). Training and validation gel images were collected from 2021–2022 and 2024; a subset of 120 images was hand-selected to include atypical multimer patterns to improve class balance. The independent test set of 94 images with 1128 instances was hand-selected from 2023 to include at least one abnormal pattern on each gel image. The date range for the training, validation, and test sets spanned multiple lots of critical reagents (i.e., agarose, primary and secondary antibodies and normal control) used in the analytical method (Table S1). The original classifications of the gel images reported by the laboratory (called Expert 2) were verified or corrected by another expert (called Expert 1) to establish ground truth for model training. Expert 2 represented one of up to 10 coagulation consultants reviewing VWF multimer gels for clinical cases, such that the same expert did not adjudicate all gel images. Performance in the test dataset was evaluated against Expert 1 "ground truth" and also against the Expert 2 clinically reported classifications. Examples of multimer patterns are shown in Figure 1.

The original gel images were 443 × 709 pixels in size and stored as 16-bit grayscale TIFF files. To prepare these images for model training, we developed a custom preprocessing pipeline. This pipeline included several enhancement steps to improve image quality and contrast. First, we applied histogram equalization to normalize the intensity distribution across the image. This was followed by Contrast Limited Adaptive Histogram Equalization (CLAHE) with a clip limit of 50.0 and a tile grid size of 5 × 5, which helped to enhance local contrast. We then applied gamma correction with a gamma value of 0.7 to adjust image brightness. Finally, the images were converted from 16-bit to 8-bit depth, scaling the pixel intensities appropriately. This conversion was necessary to make the images compatible with the YOLOv8 model input requirements and to replicate the image characteristics used by trained experts when visually analyzing gel patterns. The preprocessing steps were implemented using OpenCV and applied to all images in the training, validation, and testing sets.

Training was performed for 500 epochs using a batch size of 36 and an image size of 736 × 736 pixels. The YOLO model was trained using a set of hyperparameters optimized for the specific task. The learning rate started at 0.00381 and decreased by a learning rate factor of 0.00497 using cosine annealing. A momentum of 0.86125 was applied with a weight decay of 0.00042. The model used a warm-up period of 4.03571 epochs with a warm-up momentum of 0.46311. Loss coefficients were set to 7.30202 for bounding box regression, 0.61146 for classification, and 1.5547 for distribution focal loss. Data augmentation techniques included subtle color adjustments (HSV shifts), scaling (up to 29.048%), horizontal flipping (28.949% probability), and mosaic augmentation (77.523% probability). No vertical flipping, shearing, or perspective transformations were applied. These

TABLE 1 | Class distribution in labeled training and validation sets.

Class	Training set	Training %	Validation set	Validation %	Test set	Test %
1A (Normal distribution)	3916	63.5%	1436	62.3%	681	60.4%
NORMAL CTRL (Normal Control Passed)	987	16.0%	363	15.8%	185	16.4%
2H (Acquired loss of highest MWM)	537	8.7%	233	10.1%	78	6.9%
CMA (Suboptimal resolution–repeat)	298	4.8%	134	5.8%	37	3.3%
2B (Loss of high and intermediate MWM)	143	2.3%	34	1.5%	42	3.7%
2D (Loss of high MWM)	77	1.2%	19	0.8%	20	1.8%
3B (Absence of low, intermediate and high MWM)	65	1.1%	25	1.1%	37	3.3%
4A (Presence of ultra-large MWM)	61	1.0%	23	1.0%	19	1.7%
CTRL FAILED (Normal Control Failed)	41	0.7%	21	0.9%	3	0.3%
6D (Interfering substance – cannot interpret)	23	0.4%	14	0.6%	10	0.9%
1F (Abnormal triplet band infrastructure)	12	0.2%	2	0.1%	12	1.1%
BLANK (No sample loaded in lane)	8	0.1%	0	0%	4	0.4%
Total	6168	100%	2304	100%	1128	100%

Abbreviation: MWM, molecular weight multimers.

parameters were carefully tuned to balance model performance and generalization for the VWF multimer classification task. The model was trained on an NVIDIA A100-SXM4-40GB GPU.

Model performance was evaluated using the class-specific and macro-averaged precision, recall, and F1 metrics. Briefly, class-specific scores measure these metrics for each class in a classification task—precision (positive predictive value) shows the proportion of correct positive predictions out of all positive predictions for that class, while recall (sensitivity) shows the proportion of actual positive cases that were correctly identified for that class. The F1 score is the harmonic mean of precision and recall, which means it heavily penalizes cases where either precision or recall is very low, ensuring that a good F1 score reflects balanced performance on both metrics. The macro-averaged value is calculated by first computing the metric for each class individually and then taking the arithmetic mean across all classes. This approach is particularly useful when class imbalance exists in the dataset, as it gives equal weight to minority and majority classes, preventing the overall performance assessment from being dominated by the more frequent classes. However, this equal weighting means macro-averaging may not reflect performance on the raw distribution of classes in real-world applications where class frequencies differ significantly.

A Python-based web application was developed to automatically process machine-generated outputs using the deep learning model and visualize the results for expert review and approval. Built with Dash Open Source, a Python framework commonly used in data science and machine learning, the application allows users to upload batch outputs for automated analysis. After processing, the results are stored in a local database for future analysis, and both the gel images and the model's outputs are visualized for expert review. The application also enables reviewers to manually override the model's inference with their own classification, which is stored in the database to inform potential future model optimizations.

3 | Results

The model's learning progress in the training and validation sets, as well as the Precision-Recall curves and confusion matrix for the validation set, are shown in the Supporting Information (Figures S1, S2 and S3). In the test image set, the deep learning model demonstrated high performance in classifying VWF multimer patterns and identified all but one lane in the 94 test images. When compared to Expert 1, the model achieved an overall accuracy of 91%, with macro-averaged precision, recall, and F1-score of 0.851, 0.757, and 0.786, respectively (Table 2). Performance against Expert 2 was slightly lower, with an overall accuracy of 87% and macro-averaged precision, recall, and F1-score of 0.653, 0.653, and 0.641, respectively (Table 2).

Inter-rater agreement was assessed using Cohen's Kappa coefficient. The agreement between Expert 1 and Expert 2 was very high ($\kappa = 0.883$). The model showed strong agreement with Expert 1 ($\kappa = 0.845$) and good agreement with Expert 2

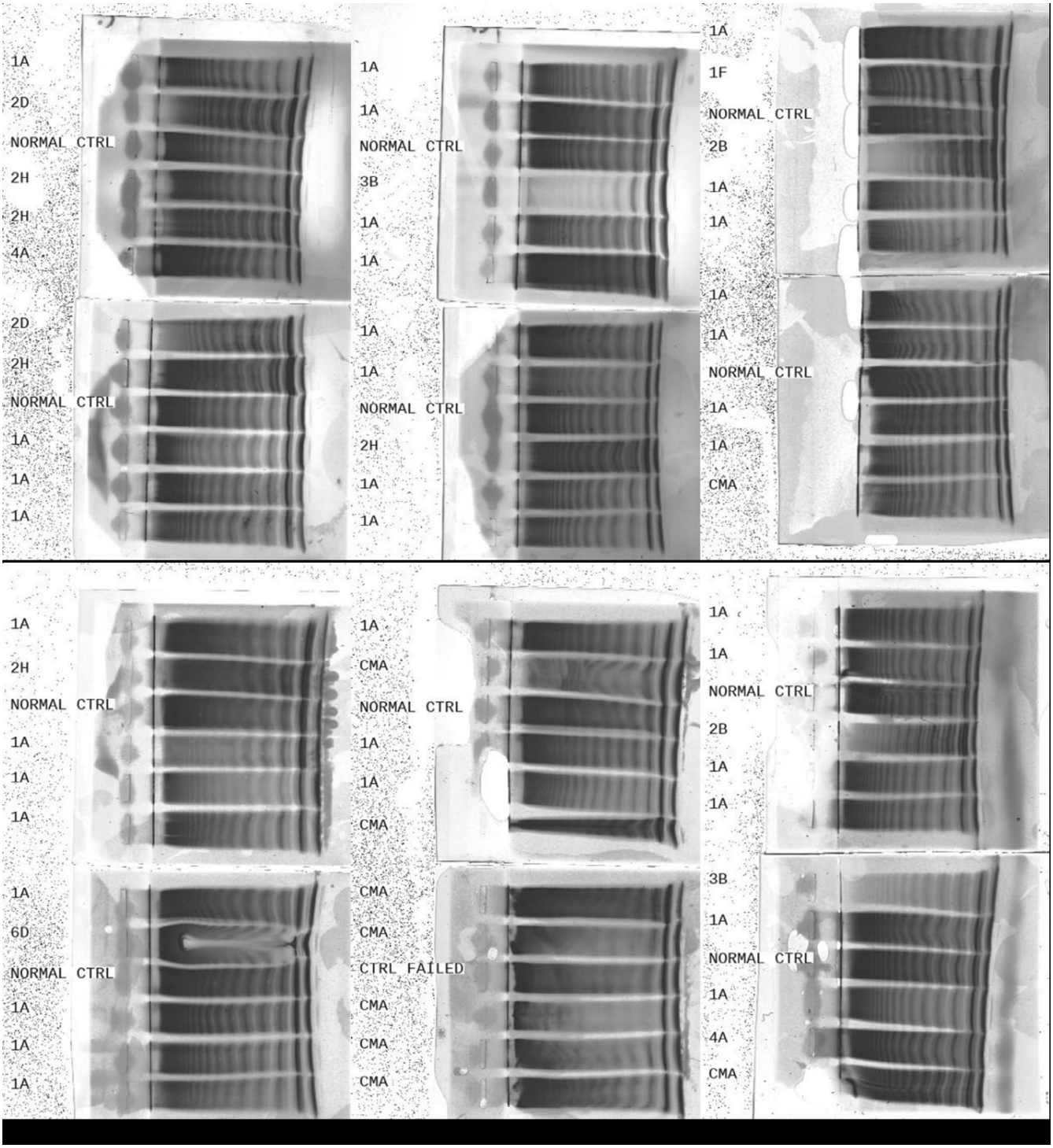

FIGURE 1 | Enhanced gel images showing Normal control lanes in positions 3 and 9 (NORMAL CTRL), failed control lanes (CTRL FAILED), normal VWF multimer distribution (1A), abnormal triplet band infrastructure (1F), loss of high and intermediate MWM (2B), acquired loss of highest MWM (2H), loss of high MWM (2D), absent or barely detectable low, intermediate and high MWM (3B), presence of ultra-large MWM (4A), interfering substance—cannot interpret (6D), suboptimal resolution (CMA).

($\kappa = 0.773$). To assess the significance of differences in classification between the model and experts, a Cochran's Q test was performed for each class (Table S2). Significant differences ($p < 0.05$) were observed for classes 1A, 1F, 2D, and CMA, indicating areas where the model's performance differed from the experts.

The model performed exceptionally well on the most common classes, such as 1A and NORMAL CTRL, with F1-scores above 0.93 for both experts. However, performance was lower for some less common classes, particularly 1F, 2D, and CMA. As shown in the confusion matrices in the Figures S4 and S5, and the Sankey diagram in Figure 2 demonstrating how discrepancies were allocated

TABLE 2 | Class-specific and macro-averaged performance metrics of the deep learning model by VWF multimer pattern class, compared to two expert raters.

	vs Expert 1			vs Expert 2		
Class	**Precision**	**Recall**	**F1-score**	**Precision**	**Recall**	**F1-score**
1A	0.930	0.974	0.951	0.926	0.947	0.936
1F	0.800	0.333	0.471	0.400	0.222	0.286
2B	0.857	0.857	0.857	0.833	0.700	0.761
2D	0.545	0.300	0.387	0.273	0.333	0.300
2H	0.714	0.705	0.710	0.519	0.597	0.556
3B	0.805	0.892	0.846	0.707	0.879	0.784
4A	0.875	0.737	0.800	0.688	0.786	0.733
6D	1.000	0.800	0.889	1.000	0.727	0.842
BLANK	0.800	1.000	0.889	0.800	1.000	0.889
CMA	0.905	0.514	0.655	0.381	0.174	0.239
CTRL FAILED	1.000	1.000	1.000	0.333	0.500	0.400
NORMAL CTRL	0.978	0.978	0.978	0.973	0.968	0.970
Macro-averaged metrics	0.851	0.757	0.786	0.653	0.653	0.641

Note: 1A (Normal distribution), NORMAL CTRL (Normal Control Passed), 2H (Acquired loss of highest MWM), CMA (Suboptimal resolution – repeat), 2B (Loss of high and intermediate MWM), 2D (Loss of high MWM), 3B (Absence of low, intermediate and high MWM), 4A (Presence of ultra-large MWM), CTRL FAILED (Normal Control Failed), 6D (Interfering substance—cannot interpret), 1F (Abnormal triplet band infrastructure), BLANK (No sample loaded in lane).

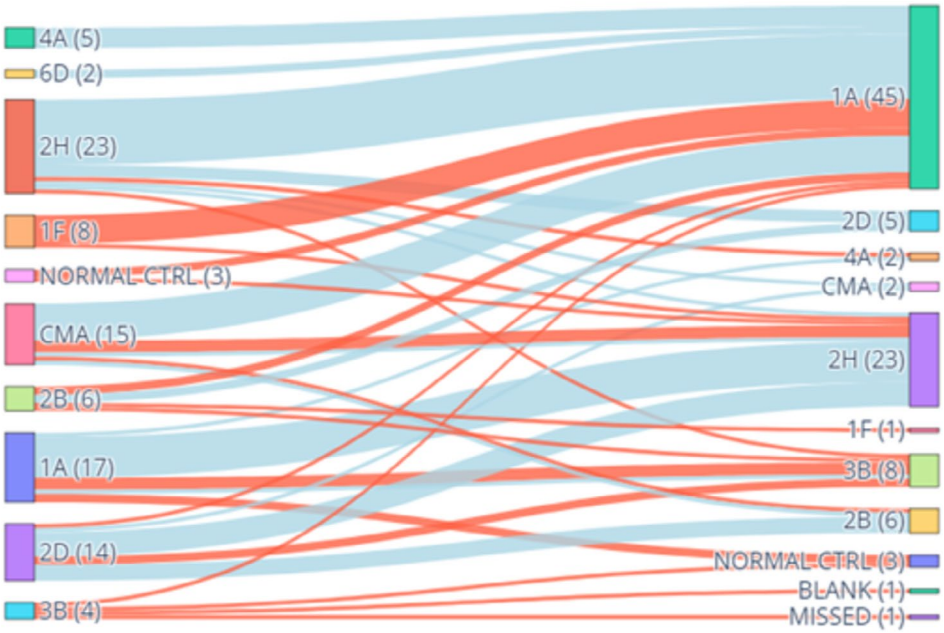

FIGURE 2 | The Sankey diagram visualizes the discrepancies between ground truth classifications (on the left) and model Inferences (on the right) for the test image set. Each label represents a unique VWF multimer pattern class, with the number of occurrences of misclassifications shown in parentheses. Arrows represent transitions between ground truth and model classifications, with the thickness of the arrows corresponding to the frequency of each discrepancy. Arrows are color-coded based on the type of discrepancy: Blue for Minor discrepancies and red for Major discrepancies as assessed by Expert 1. 1A (Normal distribution), NORMAL CTRL (Normal Control Passed), 2H (Acquired loss of highest MWM), CMA (Suboptimal resolution—repeat), 2B (Loss of high and intermediate MWM), 2D (Loss of high MWM), 3B (Absence of low, intermediate and high MWM), 4A (Presence of ultra-large MWM), CTRL FAILED (Normal Control Failed), 6D (Interfering substance—cannot interpret), 1F (Abnormal triplet band infrastructure), BLANK (No sample loaded in lane).

between Expert 1 and the model inferences, there was class confusion between 2H versus 2D versus 2B and between CMA versus 1A versus 2H. The distribution of classes in the test image set for gel positions 1 through 12, excluding the normal control lanes (positions 3 and 9) was relatively balanced (Figure S6). Examples of model inferences on test images are shown in Figure 3 and a screenshot of the Dash web application is shown in the Figure S7.

4 | Discussion

This study demonstrates that automated von Willebrand factor (VWF) multimer image analysis using deep learning techniques can improve the diagnosis and classification of von Willebrand disease (VWD). The high performance of the YOLOv8 model in classifying VWF multimer patterns, with overall accuracies of

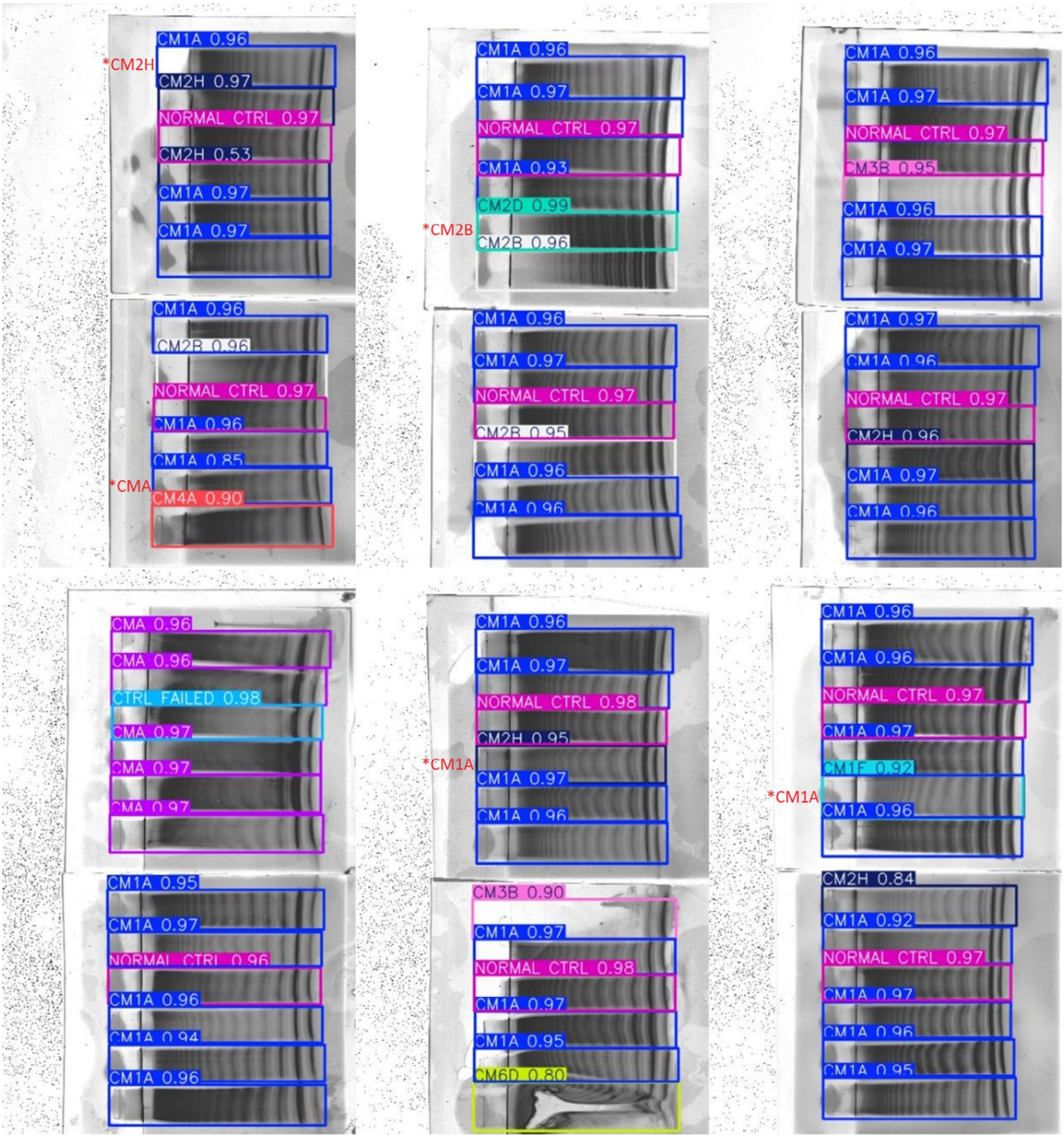

FIGURE 3 | Enhanced gel images showing model inferences in colored boxes for normal control lanes in positions 3 and 9 (NORMAL CTRL), failed control lanes (CTRL FAILED), normal VWF multimer distribution (1A), acquired loss of highest MWM (2H), loss of high MWM (2D), loss of high and intermediate MWM (2B), absence of low, intermediate and high MWM (3B), suboptimal resolution (CMA), interfering substances (6D), abnormal triplet band infrastructure (1F). Discordances between the model inferences and ground truth classifications are shown in red font beside the corresponding lane; 5 of the 72 lanes in these example images were misclassified by the model.

91% and 87% compared to two expert raters, indicates the potential of this approach to standardize and expedite VWF multimer interpretation.

VWF multimer analysis plays a critical role in diagnosing and classifying both congenital VWD and acquired von Willebrand syndrome (AVWS) [5]. As noted by Saadalla et al., VWF multimer patterns can provide key insights into the underlying pathologic mechanisms related to VWF dimerization, multimerization, secretion, and proteolysis that result in different VWD subtypes [14]. However, traditional human interpretation of multimer patterns requires expert training and is complex, time-consuming, and subject to inter-rater variability, especially for subtle abnormalities [13].

The automated deep learning approach described here addresses many of the limitations of manual analysis. By providing rapid, objective classification of multimer patterns, it could improve the efficiency and reproducibility of VWD diagnosis. The model showed particularly high performance for common patterns like type 1 VWD and normal controls. Its ability to detect subtle loss of high molecular weight multimers in AVWS is promising, as this can be challenging even for expert raters [12]. The model's performance in identifying poor resolution lanes for repeat analysis and artifacts due to monoclonal immunoglobulins is also reassuring.

The lower performance for some rare subtypes like type 1F and 2D VWD likely reflects their limited representation in the training dataset. These rare variants have distinct multimer patterns that require careful evaluation [14]. Expanding the training set to include more examples of these uncommon subtypes could further improve the model's diagnostic capabilities. The confusion between closely related patterns like 2D versus 2B VWD reflects the overlapping features of these subtypes, which can be challenging to distinguish for human experts [17]. Additional clinical and laboratory data may be needed to definitively classify these cases.

This automated approach to VWF multimer classification should be viewed as a complement to, rather than a replacement for, expert analysis. The implementation of this AI-driven workflow aims to reduce manual effort while maintaining diagnostic accuracy through strategic expert oversight. This could streamline workflow in coagulation laboratories and potentially enable wider access to VWF multimer testing. VWF multimer interpretation also requires integration with other clinical and laboratory findings [14]. While Figure 2 reveals several instances where the model classified abnormal patterns as normal (1A), many of these discrepancies would be caught by existing laboratory safeguards. For example, abnormal VWF activity/antigen ratios (<0.7) would trigger expert review regardless of the automated classification, which is particularly important for catching misclassifications of types 2B, 2D, and subtle 2H patterns. Some apparent discrepancies, such as CMA to 1A transitions, may actually represent cases where automated analysis outperformed initial human interpretation, potentially reducing unnecessary repeat testing. Other discrepancies, like those involving type 1F or 3B patterns, reflect either known challenges in pattern recognition that affect expert raters as well, or differences in describing intensity rather than true pattern misclassifications. Based on these findings, successful clinical implementation would require integration with other laboratory parameters, maintenance of human oversight in the workflow, and careful monitoring of discrepancy rates during the initial rollout and periodically thereafter.

Automated VWF multimer analysis using computer vision shows considerable promise for improving the classification of VWF multimer patterns. Further validation on larger image sets will be important to ensure the model's generalizability. Integration with other VWD diagnostic tests and clinical data could create a more comprehensive automated classification system. Additionally, the ability to quantify multimer band density, as described for some semi-automated systems [15], could enhance the model's capabilities. By providing rapid, standardized interpretation of multimer patterns, it could enhance both the efficiency and accuracy of VWD evaluation. Further refinement and validation of this technique may help to make VWF multimer analysis more widely available as a diagnostic tool for VWD.

Author Contributions

K.A. developed the deep learning algorithms, performed computational analysis, and assisted with study design. V.O. performed image annotation and contributed to study conceptualization. C.Z. developed and implemented the web-based application for model deployment. K.N. and E.A. contributed to data curation and provided technical expertise on VWF multimer methodology. J.P.B., R.P., and D.C. contributed to study design, results interpretation, discrepancy resolution, and validation of ground truth classifications. J.N.S. conceived and designed the study, developed the deep learning models, performed model evaluation, supervised model deployment, and wrote the manuscript. All authors reviewed and edited the manuscript.

Ethics Statement

The Mayo Clinic Institutional Review Board (IRB) acknowledged, in accordance with the Code of Federal Regulations, 45 CFR 46.102, that this study does not require IRB review.

Consent

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Research data are not shared due to commercial restrictions.

References

1. E. A. Jaffe, L. W. Hoyer, and R. L. Nachman, "Synthesis of von Willebrand Factor by Cultured Human Endothelial Cells," *Proceedings of the National Academy of Sciences of the United States of America* 71, no. 5 (1974): 1906–1909.

2. D. D. Wagner, J. B. Olmsted, and V. J. Marder, "Immunolocalization of von Willebrand Protein in Weibel-Palade Bodies of Human Endothelial Cells," *Journal of Cell Biology* 95, no. 1 (1982): 355–360.

3. J. F. Dong, J. L. Moake, A. Bernardo, et al., "ADAMTS-13 Metalloprotease Interacts With the Endothelial Cell-Derived Ultra-Large von

Willebrand Factor," *Journal of Biological Chemistry* 278, no. 32 (2003): 29633–29639.

4. M. Stockschlaeder, R. Schneppenheim, and U. Budde, "Update on von Willebrand Factor Multimers: Focus on High-Molecular-Weight Multimers and Their Role in Hemostasis," *Blood Coagulation & Fibrinolysis* 25, no. 3 (2014): 206–216.

5. P. D. James, N. T. Connell, B. Ameer, et al., "ASH ISTH NHF WFH 2021 Guidelines on the Diagnosis of von Willebrand Disease," *Blood Advances* 5, no. 1 (2021): 280–300.

6. A. B. Federici, J. H. Rand, P. Bucciarelli, et al., "Acquired von Willebrand Syndrome: Data From an International Registry," *Thrombosis and Haemostasis* 84, no. 2 (2000): 345–349.

7. R. A. Crist, N. M. Heikal, G. M. Rodgers, D. G. Grenache, and K. J. Smock, "Evaluation of a New Commercial Method for von Willebrand Factor Multimeric Analysis," *International Journal of Laboratory Hematology* 40, no. 5 (2018): 586–591.

8. E. J. Favaloro, S. Oliver, S. Mohammed, and R. Vong, "Comparative Assessment of von Willebrand Factor Multimers vs Activity for von Willebrand Disease Using Modern Contemporary Methodologies," *Haemophilia* 26, no. 3 (2020): 503–512.

9. Z. M. Ruggeri and T. S. Zimmerman, "The Complex Multimeric Composition of Factor VIII/von Willebrand Factor," *Blood* 57, no. 6 (1981): 1140–1143.

10. R. K. Pruthi, T. M. Daniels, J. A. Heit, D. Chen, W. G. Owen, and W. L. Nichols, "Plasma von Willebrand Factor Multimer Quantitative Analysis by In-Gel Immunostaining and Infrared Fluorescent Imaging," *Thrombosis Research* 126, no. 6 (2010): 543–549.

11. S. Oliver, K. K. E. Lau, K. Chapman, and E. J. Favaloro, "Laboratory Testing for Von Willebrand Factor Multimers," *Methods in Molecular Biology* 1646 (2017): 495–511.

12. C. O. Austin, D. Chen, C. S. Thomas, et al., "Von Willebrand Factor Multimer Quantitation for Assessment of Cardiac Lesion Severity and Bleeding Risk," *Research and Practice in Thrombosis and Haemostasis* 2, no. 1 (2017): 155–161.

13. J. D. Studt, U. Budde, R. Schneppenheim, et al., "Quantification and Facilitated Comparison of von Willebrand Factor Multimer Patterns by Densitometry," *American Journal of Clinical Pathology* 116, no. 4 (2001): 567–574.

14. A. Saadalla, J. Seheult, R. K. Pruthi, and D. Chen, "Von Willebrand Factor Multimer Analysis and Classification: A Comprehensive Review and Updates," *Seminars in Thrombosis and Hemostasis* 49, no. 6 (2023): 580–591.

15. I. Vangenechten and A. Gadisseur, "Improving Diagnosis of von Willebrand Disease: Reference Ranges for von Willebrand Factor Multimer Distribution," *Research and Practice in Thrombosis and Haemostasis* 4, no. 6 (2020): 1024–1034.

16. G. Jocher, A. Chaurasia, and J. Qiu, "YOLO by Ultralytics (Version 8.0.0) [Computer Software]," 2023, https://github.com/ultralytics/ultralytics.

17. A. Casonato, V. Daidone, E. Galletta, and A. Bertomoro, "Type 2B von Willebrand Disease With or Without Large Multimers: A Distinction of the Two Sides of the Disorder Is Long Overdue," *PLoS One* 12, no. 6 (2017): e0179566.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.