Deep learning for high-throughput quantification of oligodendrocyte ensheathment at single-cell resolution

Yu Kang T. Xu[1], Daryan Chitsaz[1], Robert A. Brown[1], Qiao Ling Cui[1], Matthew A. Dabarno[1], Jack P. Antel[1] & Timothy E. Kennedy[1]

High-throughput quantification of oligodendrocyte myelination is a challenge that, if addressed, would facilitate the development of therapeutics to promote myelin protection and repair. Here, we established a high-throughput method to assess oligodendrocyte ensheathment *in-vitro*, combining nanofiber culture devices and automated imaging with a heuristic approach that informed the development of a deep learning analytic algorithm. The heuristic approach was developed by modeling general characteristics of oligodendrocyte ensheathments, while the deep learning neural network employed a UNet architecture and a single-cell training method to associate ensheathed segments with individual oligodendrocytes. Reliable extraction of multiple morphological parameters from individual cells, without heuristic approximations, allowed the UNet to match the accuracy of expert-human measurements. The capacity of this technology to perform multi-parametric analyses at the level of individual cells, while reducing manual labor and eliminating human variability, permits the detection of nuanced cellular differences to accelerate the discovery of new insights into oligodendrocyte physiology.

[1] McGill Program in Neuroengineering, Department of Neurology and Neurosurgery, Montreal Neurological Institute, McGill University, H3A 2B4 Montreal, QC, Canada. Correspondence and requests for materials should be addressed to T.E.K. (email: timothy.kennedy@mcgill.ca)

In the CNS, oligodendrocyte produce myelin by wrapping specialized multi-lamellar membrane sheaths around axons. In pathological conditions, like multiple sclerosis, leukodystrophies, or CNS trauma, myelin is progressively lost due to inadequate repair of damaged sheaths[1]. Enhancing myelin production and stability are critical clinical objectives, and several in vitro systems have been established to study these dynamic process[2–4]. In one particularly reduced system, oligodendrocytes are cultured on axon-diameter electrospun polymer nanofibers. Immature O4+ oligodendrocytes initially contact fibers, followed by differentiation and expression of mature oligodendrocyte markers like myelin-basic protein (MBP)[5]. The extent of ensheathment around these axon-like structures can be measured in response to different treatments, ranging from substrate coatings to the application of bioactive small molecules[6–9]. Regardless of the culture system, however, oligodendrocyte myelination is typically quantified manually, creating a bottleneck and introducing human subjectivity.

To resolve this quantitative challenge, we developed automated high-throughput methods that combine micro-engineered nanofiber culture devices, automated microscopy, and analytic algorithms to extract detailed morphological properties from individual oligodendrocytes. Quantitative algorithms were developed to assess myelin elaborated on electrospun 2 μm-diameter poly-L-lactic acid fibers[6,7]. Oligodendrocytes cultured on nanofibers recapitulate biologically relevant features of myelin development and the planar parallel fiber arrangement is conducive to automated imaging and segmentation. In this report, we describe optimized protocols for oligodendrocyte nanofiber cultures and automated imaging, along with the development of a heuristic algorithm whose limitations ultimately informed the construction of a deep learning neural network to quantify immunocytochemically labelled markers of oligodendrocyte ensheathment.

Our initial heuristic approach targeted the general morphological characteristics of oligodendrocyte sheaths, such as the presence of thick elongated structures, to segment fluorescently labelled oligodendrocyte segments. Although the approximations encoded in the heuristic MATLAB algorithm resulted in rapid and accurate measurements of individual ensheathed segments, this method revealed that reliably associating myelin-like ensheathments with the corresponding cell bodies presented a substantial challenge. To improve beyond this initial approach, we developed a deep learning neural network with the ability to extract and associate multiple morphological parameters with individual oligodendrocytes.

To develop the convolutional neural network (CNN), we employed a cell-by-cell training method and a UNet architecture[10] (CNN-UNet) that enhanced the algorithm's capacity to perform segmentations on individual oligodendrocytes using spatial and object specific information. Validation demonstrated that the deep-learning approach performs beyond the accuracy of the heuristic algorithm, particularly at single-cell resolution. Further, the CNN-UNet reached a level of analytic accuracy within the range of human experts on multiple parameters, while reducing workloads and providing standardization that eliminates human variability.

We then tested the sensitivity of the heuristic and CNN-UNet algorithms to detect a previously reported difference in the number of sheaths formed per oligodendrocyte on poly-D-lysine (PDL) versus laminin 1 coated nanofibers[7]. Both algorithms successfully detected the difference in number of sheaths and, due to the high number of cells analyzed, also revealed a previously unidentified increase in the length of oligodendrocyte sheaths on laminin 1 compared to PDL coated nanofibers. Our findings highlight the power of this analytic approach to quantify large

numbers of cells, thereby providing sufficient sensitivity to detect nuanced differences in cellular processes.

Overall, the analytic pipeline we describe allows for a 96 or 384 well plate to be imaged overnight and analyzed the following day with a single computing station. By extracting multiple morphological parameters at single-cell resolution, while obtaining analytic quality similar to human segmentations, this high-throughput system has the potential to accelerate the discovery of therapeutics that promote myelin protection and enhance our understanding of oligodendrocyte physiology.

Results

Comparison of heuristic and CNN-UNet segmentations. To extract detailed morphological properties from individual OLs, we combined automated high-throughput methods, micro-engineered nanofiber culture devices, automated microscopy, and analytic algorithms (Fig. 1a). The heuristic algorithm developed to quantify oligodendrocyte ensheathments targeted two primary morphological properties: linearity of ensheathments and their continuity with cell bodies (Fig. 1b, c). In addition to performing qualitatively valid ensheathment segmentations (Fig. 1c), this algorithm revealed the challenge of associating ensheathed segments with specific oligodendrocytes, as morphological approximations of continuity and cell body size were not always viable in images containing crowded cells, faded fluorescence, or certain atypical morphologies. To resolve this, we developed a deep learning approach using a single-masked-nucleus training method (See "Materials and Methods: Training", Fig. 2) which exhibited an enhanced capacity to associate multiple-morphological features with specific oligodendrocytes (Fig. 3). A qualitative visual evaluation of the output of select layers of the trained CNN-UNet confirmed that the network had learned to identify ensheathments belonging to individual oligodendrocytes when prompted with images of cells with masked nuclei (Fig. 3c, d). The capacity for the CNN-UNet to combine spatial and object-specific information to reliably associate specific sheaths with individual cells highlights a key advantage of our masked-nucleus training method over approximating morphological features with heuristic algorithms.

CNN-UNet best matched expert human accuracy. After generating heuristic and deep-learning algorithms, the performance of both systems was evaluated in relation to human researchers. Three humans were employed for this analysis and were ranked based on their experience with oligodendrocyte cultures—human 1 (H1) being a cell biologist, H2 a computer scientist with some experience in cell biology, and H3 a summer medical student with no previous experience working with oligodendrocytes. Initial analysis of one test image containing ~1000 candidate cells revealed that the output of either the heuristic or CNN-UNet algorithms fell within the range of human quantification for multiple parameters (Fig. 4). No statistically significant differences were detected between the log and absolute mean sheath lengths, nor in the mean sheath length per cell (mSLC), between H1–3 and machine algorithms (Fig. 4a, b, d). However, despite predefined segmentation criteria (see "Materials and Methods: Manual Segmentation Criteria"), variability between humans was much higher at a single-cell level (Fig. 4c). Only the CNN-UNet identified a mean number of sheaths per cell that did not differ significantly from H1 (Fig. 4c). The overall number of ensheathing cells identified was also consistent between groups, except for H2 (Fig. 4e).

After evaluating global morphological parameters, we then used Jaccard Index (JI) comparisons, a measure of the overlap similarity between any binarized segmentation and a
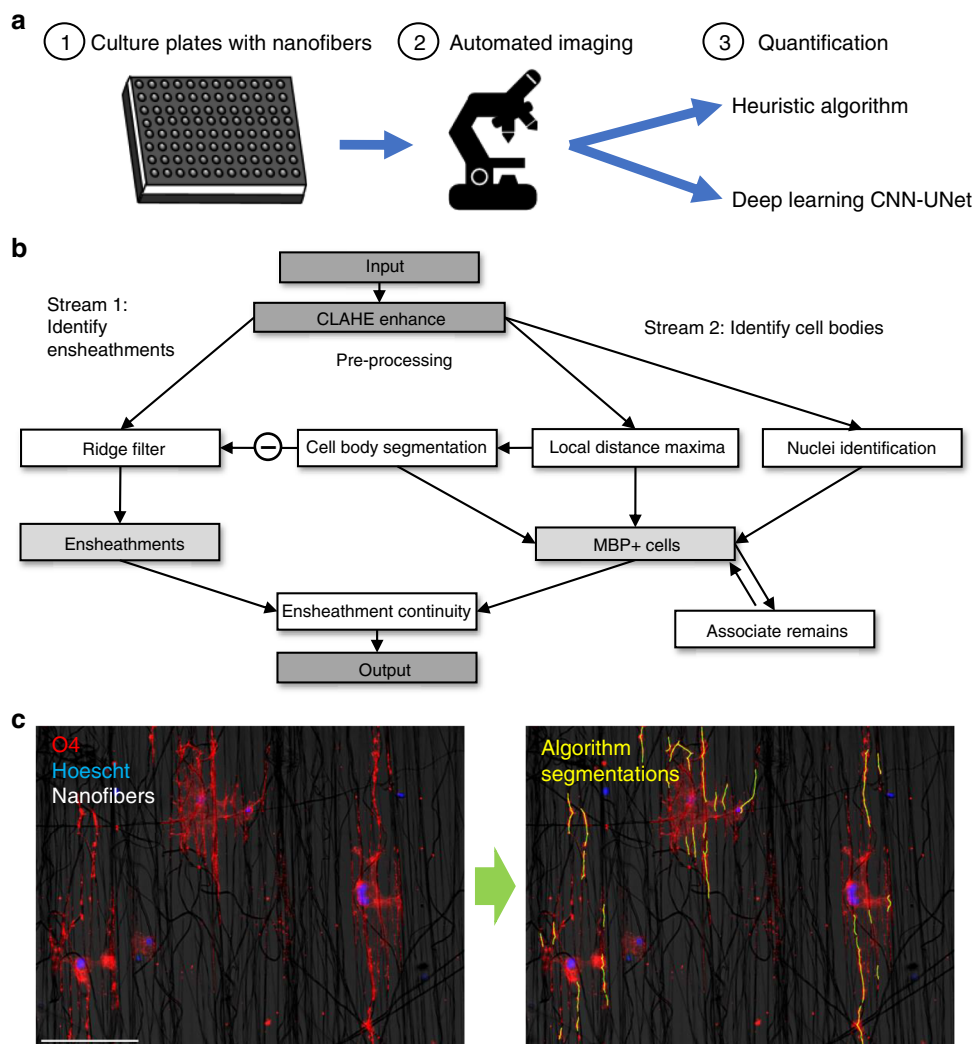
Fig. 1 Overview of analytic pipeline and heuristic MATLAB algorithm. **a** Graphic illustrating automated pipeline using nanofiber culture devices, automated imaging, and analytic algorithms to quantify oligodendrocyte ensheathments. **b** Input images are pre-processed using CLAHE enhancement to increase contrast before identifying ensheathments and oligodendrocyte cell bodies in two separate analytic streams. (Stream 1) a ridge-filter is applied to identify ensheathed nanofiber segments and (Stream 2) cell-bodies are estimated using watershed segmentation with imposed local distance maxima. The middle link (subtraction sign) subtracts the cell bodies from the ensheathments identified by the ridge filter to eliminate segments located within cell bodies. The final outputs of both pathways are combined for ensheathment continuity analysis. **c** Input and output of the heuristic algorithmic approach. Cell nuclei (blue), O4 (red), nanofibers (bright field in background). Scale bar 20 μm. Yellow outlines in the output image indicate the algorithm's segmentations

corresponding ground truth[11], to ensure that the algorithmic segmentations were similar to those of human researchers on a single cell and whole-image level. To assess the similarity of segmentations on individual cells, we selected the segmentations provided by H1 as the ground-truth and computed the JI for every cell that was commonly identified after pairing H1 with each researcher or computer program. We found that the CNN-UNet outperformed human participants and the heuristic algorithm in matching H1 at single-cell resolution (Fig. 4f). We also computed the global JI for the entire whole-well image and found that the CNN-UNet best matched expert H1 (Fig. 4g). Thus, the CNN-UNet outperforms the heuristic algorithm in analysis at single-cell resolution and in global parameters due to its heightened capacity to extract information from individual oligodendrocytes.

CNN-UNet reduces variability and generalizes to experts. To verify that the performance of the algorithms was not biased to

the specific ground-truth used for training, we introduced an external expert (Ex1), from a different laboratory than H1, who was not involved in the development of the analytic models. Segmenting the same whole-well image, Ex1 matched the CNN-UNet on several parameters, including: the mean number of sheaths per cell, mSLC, single-cell JI values, global JI values, and the number of total ensheathing cells (Fig. 4c–g). The heuristic algorithm also matched Ex1 in identifying the mean number of sheaths formed per cell, mSLC, and the number of total ensheathing cells, but performed poorly in associating individual sheaths and in overall segmentation overlap, as indicated by lower single-cell and global JI values (Fig. 4c–g). Additionally, Ex1 reported a slight but significant shift of ~4 μm in mean ensheathment lengths (Fig. 4a, b) and differed from H1 in identifying the mean number of sheaths per cell (Fig. 4c). This difference shows that, while human experts matched each other better than the non-expert (H3), there was still slight variability in specific parameters of the analyses carried out by the two different experts. Conversely, the CNN-UNet matched Ex1 on several
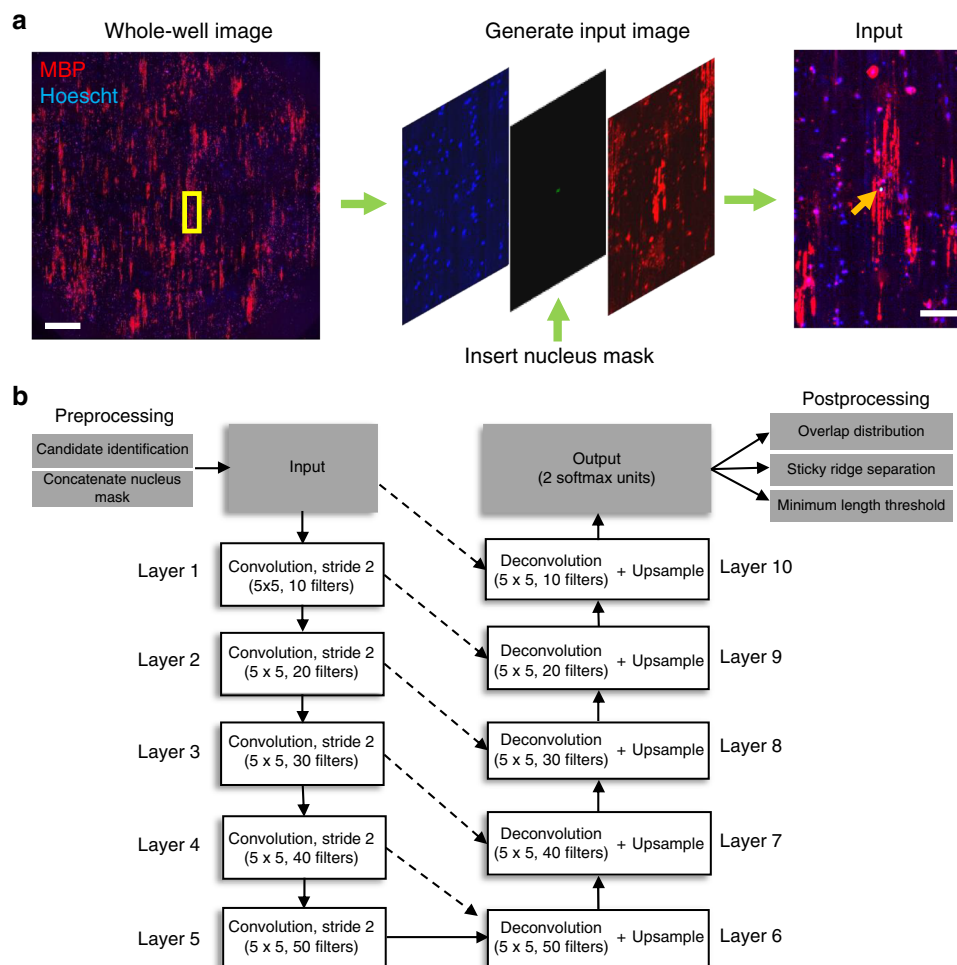
3

Fig. 2 Pre-processing, network architecture and work-flow of CNN-UNet approach. **a** The yellow rectangular mask on the whole-well image (left) indicates the portion of the image cropped around a candidate Hoechst center. That Hoechst center is then used as a mask to replace the green channel (middle) in the "Input" image (right). Arrowhead on "Input" image indicates the nucleus of the masked cell of interest. This method is used to generate both the training images and the input images when testing the algorithm. Scale bar on "Whole-well image" corresponds to 1 mm. Scale bar on "Input" corresponds to 100 μm. **b** Starting from the left, a whole-well image is pre-processed to generate nuclei masks for candidate cells. These candidates are then piped into the network architecture as an "input" image that is down-sampled on the left arm of the network, and then up-sampled on the right-arm of the network. Spatial information from the down-sampling branch is also mixed into the up-sampling to allow the network to learn to associate ensheathments with individual oligodendrocytes. The final "output" segmentation is then subjected to post-processing

parameters, and even identified a mean number of sheaths per cell that was not significantly different from either expert H1 or Ex1 (Fig. 4c). This capacity to match Ex1 and H1 demonstrates that the CNN-UNet is generalizable to external experts beyond the training ground-truth and highlights the capacity for automated techniques to offer an unvarying analytic standard that eliminates inherent human variability in quantification.

CNN-UNet generalizes to different orientations. We also assessed the capacity of the deep learning approach to quantify ensheathments at differing orientations by training a CNN-UNet with images rotated at random angles. Notably, training a CNN-UNet with rotated images dramatically decreased the training rate and resulted in a lower plateau in JI after 16 days (1,000,000 epochs) (Fig. 5a). The performance of the CNN trained with rotated images was then tested with a whole well image ("CNN-UNet – rotated"), or by randomly rotating each "input" image that was cropped around a candidate cell within a whole well image ("CNN-UNet – rand rotated"). Interestingly, the random rotations of cropped "input" images during testing did not drastically affect the performance of the CNN-UNet (Fig. 5).

Regardless of the testing method, however, the mean length of ensheathed segments identified by the CNN-UNet trained with rotated images was significantly lower than that identified by H1 and the non-rotated control CNN-UNet (Fig. 5b, c). The number of sheaths per cell, however, matched between all CNN's and H1 (Fig. 5d). As well, the CNN-UNet trained with rotated images had a slightly lower sensitivity towards ensheathing cells overall (Fig. 5e), and matched H1 on global segmentations to a moderate degree (Fig. 5f). Thus, our results demonstrated that a masked-nucleus training method can allow our CNN-UNet to generalize to both rotated and aligned sheaths without requiring a researcher to model more heuristic features through computer code. Notably, while training the CNN-UNet with rotated images reduced sensitivity, the network still provides a useful quantification that offers standardization to eliminate human variability in analysis.

Subtle differences detected between nanofiber coatings. Having validated the analytic algorithms on both whole-image and cell-to-cell levels, we assessed the capacity of the algorithms to detect a subtle biological effect by examining the difference between PDL
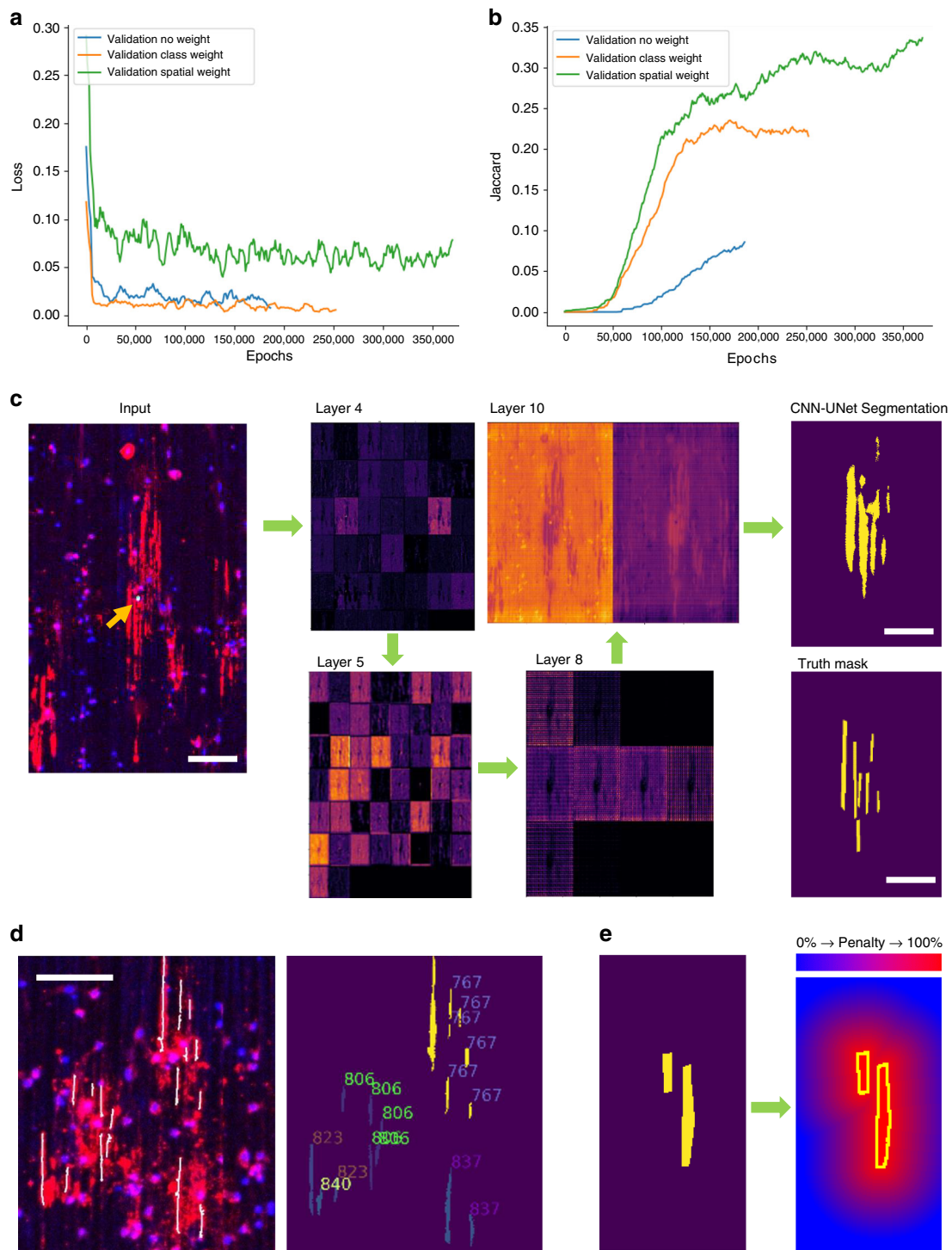
Fig. 3 CNN-UNet training and demonstration of analysis at single-cell resolution. **a** Loss decreased continuously for all validation sets, however, the spatially weighted network did see an increase in validation loss after ~301,000 epochs, indicating over-fitting. Thus, the training was stopped early for testing. **b** Both class and spatial weighting improved learning rates dramatically, as indicated by an earlier and steeper increase in the JI of the validation sets. However, class weighting appears to plateau without an improvement in the JI after ~180,000 epochs. The spatially weighted CNN-UNet was therefore selected as our best model. **c** Sample outputs of specific layers leading to the final output "CNN-UNet segmentation". "Layer 4" exhibited a binary response to the distribution of MBP immunoreactivity, while "Layer 5" appeared to activate for different parts of the MBP image, suggesting the extraction of features corresponding to linear ensheathments. Arrowhead shows nucleus of masked cell of interest. Scale bar "Input" 100 μm. Scale bar "Segmentation" and "Truth mask" 50 μm. (**d**, left) Zoomed inset of the 10x whole-well image used for testing. The ensheathed fibers segmented by the CNN-UNet are overlaid (white). Hoechst is shown in blue and MBP in red. Scale bar 100 μm. (**d**, right) Ensheathed fibers segmented by the CNN-UNet are annotated to indicate the specific cell they belong to, showing that the network learned to associate fibers with specific nuclei, rather than identifying individual segments globally across a whole-image. **e** Illustration of spatial-weighting penalization to draw "attention" to object edges in training
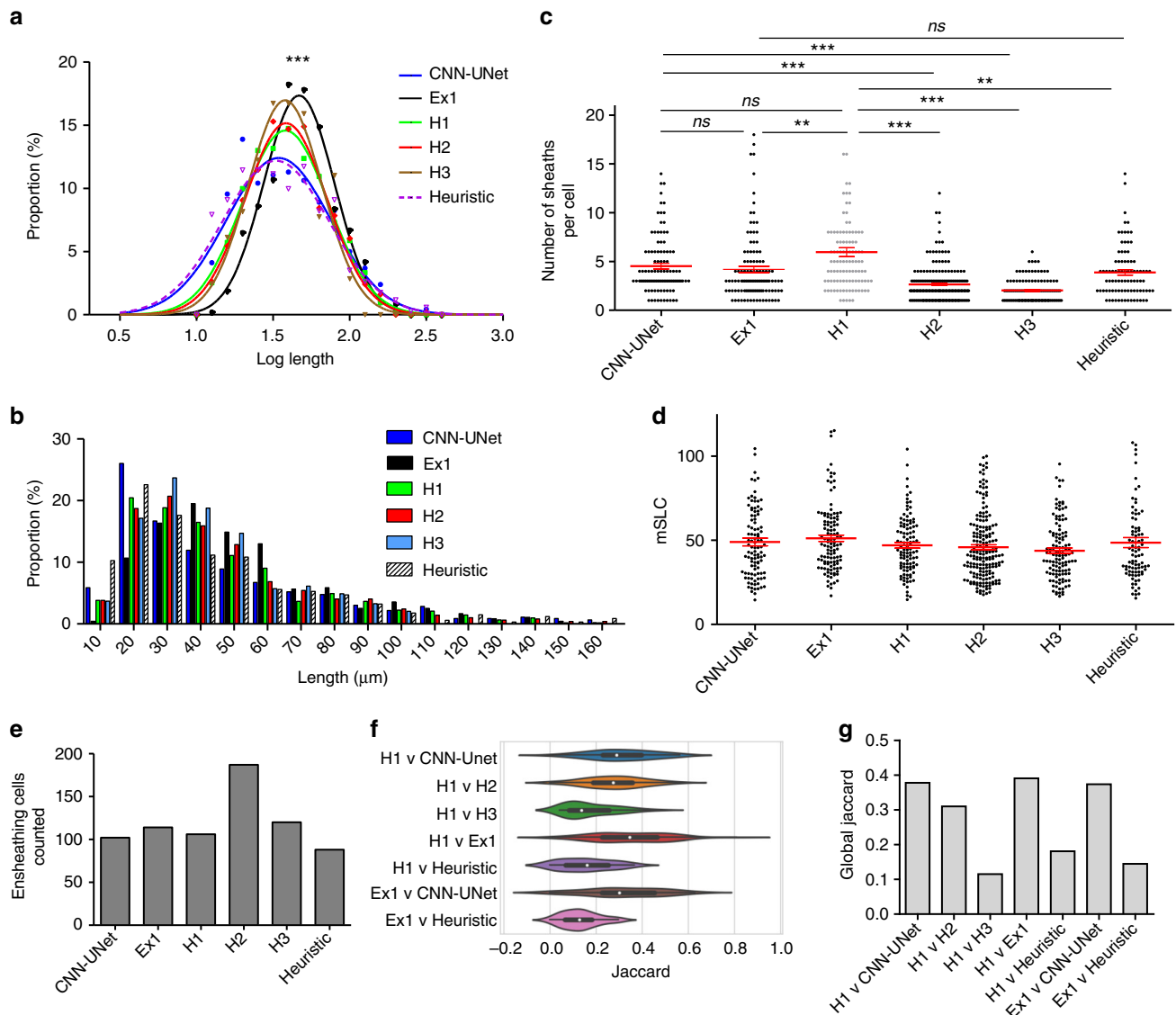
Fig. 4 Comparison of the performance of humans (H1, H2, H3, Ex1) and machines (CNN-UNet, Heuristic) on a whole-well image. **a** Log and **b** absolute distributions of ensheathment lengths are not significantly different between machines and humans (H1–3) except Ex1 (one-way-ANOVA with Tukey's post-test). **c** Only the number of sheaths per cell identified by the CNN-UNet matched that of H1 (the expert). Other humans with less expertise, and the heuristic program, differed significantly from H1 (Kruskal–Wallis test with Dunn's post-test). CNN-UNet also matched human Ex1, showing no bias towards training ground truth data. Error bars and midline show mean ± SEM. **d** No significant difference between mean sheath lengths per cell (mSLC). Error bars and midline show mean ± SEM. **e** The total number of ensheathing cells was identified similarly by all groups, except H2. Such variability in analysis by human non-experts supports the need for standardized analytic systems. **f** Violin plots of the JI similarity measure when comparing the segmentations from Human 1 (H1) with H2, H3, Ex1, CNN-UNet, and the heuristic algorithm on a single-cell level. Overall, the experts best matched one another (Ex1 and H1), but the CNN-UNet also closely matched both Ex1 and H1 on a cell-to-cell basis. Middle dot of violin plots represents the median value of the distribution, and the upper/lower box limits show the upper/lower quartiles. **g** The global JI was also computed across an entire whole-well image between the same comparison groups. The CNN-UNet and Ex1 again had the highest match with H1. See also Supplementary Table 1

and laminin 1 coated nanofibers. Compared to PDL coatings, laminin has been shown to increase the number of sheaths formed per cell without influencing the distribution of ensheathment lengths[7]. Both machine programs recapitulated the finding that a coating of laminin 1 increases the number of sheaths per cell compared to PDL, while also identifying a difference in the distribution of sheath lengths, due to the increased sample size of our automated approach. For the number of sheaths formed per cell, the heuristic algorithm found a difference of $0.74 ± 0.10$ sheaths ($p < 0.0001$, Mann–Whitney U-test) and the CNN-UNet found a difference of $0.30 ± 0.08$ sheaths ($p = 0.0018$, Mann–Whitney U-test) between laminin and PDL coated

conditions (Fig. 6c). For log length distributions, both the heuristic algorithm and the CNN-UNet identified a small but statistically significant difference in means of $0.018 ± 0.0032$ ($p < 0.0001$, Student's t-test) and $0.0137 ± 0.0037$ ($p = 0.0002$, Student's t-test) log units respectively (Fig. 6a, b).

While Bechler et al.[7] previously reported an upward trend in the lower and upper bounds of the log length distribution in laminin coated conditions, the difference was not statistically significant. Our results demonstrate a key advantage of the high-throughput system: by sampling ~43 times more ensheathments in each experimental condition (~13,000 instead of ~300 ensheathments), we detected a nuanced, yet significant, difference in the length of the myelin-like
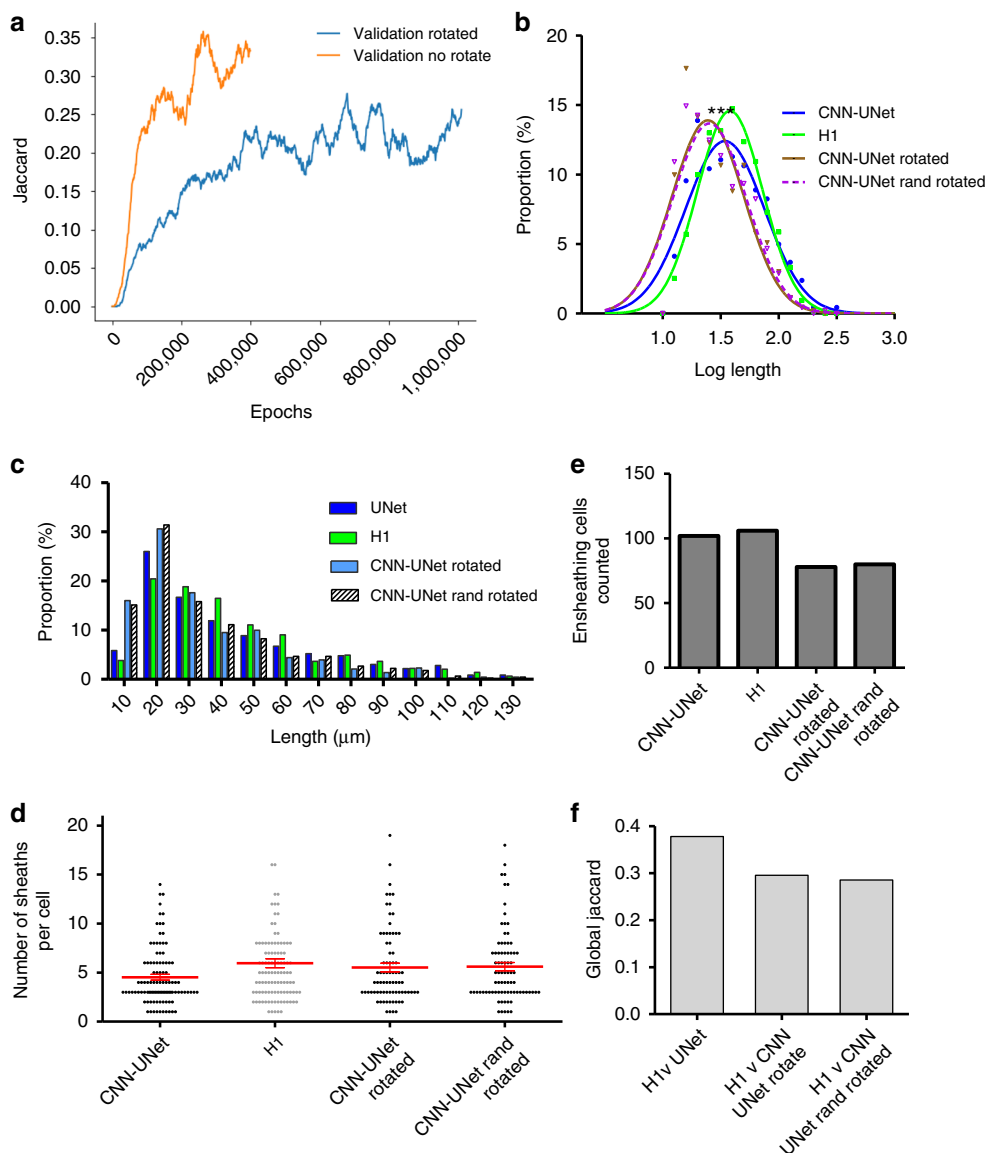
Fig. 5 CNN-UNet performance when trained with randomly rotated images. **a** Performance of CNNs during training as monitored through JI. Compared to the control CNN-UNet ("Validation no rotate"), the JI of the CNN-UNet trained with rotated training images ("Validation rotated") increased at a much slower rate and seemed to plateau after 16 days of training (1,000,000 epochs) at a lower JI. The CNN-UNet was then tested using a whole-well image ("CNN-UNet – rotate") or after rotating each candidate "input" cell within a whole-well image by a random angle ("CNN-UNet – rand rotate"). The random rotation had no significant impact on testing performance. **b**, **c** The length distribution identified by the rotated network differed significantly from H1 and non-rotated control (CNN-UNet), but **d** the number of sheaths per cell matched between all CNN's and H1. Error bars and midline show mean ± SEM. Additionally, **e** the CNN's trained with rotated images also had a slightly lower sensitivity towards ensheathing cells overall, and **f** matched H1 global segmentations to a moderate degree

ensheathments in the laminin 1 coated condition. A power analysis, using Lenth's Power Calculator[12,13], demonstrates that a sample size of 13,000 sheaths (from 10 wells with ~260 ensheathing cells and ~5 sheaths per cell) can detect a difference of 0.01 log lengths with 80% confidence, whereas a smaller sample size of 300 sheaths can only discern a difference of 0.07 log units. Thus, only with the larger sample size were we able to detect the difference of ~0.015 log lengths between PDL and laminin coated nanofibers. This assessment verified the sensitivity of our systems and highlights the importance of using a large sample size to evaluate multiple morphological parameters when assessing differences in forming myelin-like sheaths.

Automated pipeline permits high-throughput assessment. Finally, we evaluated the speed of our analytic algorithms for

practical use. Our external expert (Ex1) took ~2 h to trace all MBP stained ensheathments (corresponding to ~100 ensheathing cells) in a single well. Assuming that a researcher traces for about 8 h per day, it would take ~24 days to analyze an entire 96-well plate. In contrast, the heuristic algorithm and CNN-UNet (using a NVIDIA GTX 1070 graphics card) analyzed the entire surface of a 96-well plate in ~1 day (Table 1). Automated microscopy also reduced the imaging time from approximately 77 manual hours to 15 automated hours for a full plate (Table 2), with imaging performed autonomously overnight after setting up and calibrating the system. Altogether, the automated pipeline dramatically enhanced the speed of analysis while largely eliminating manual-labor, thus providing the capacity to screen hundreds of compounds within a week.
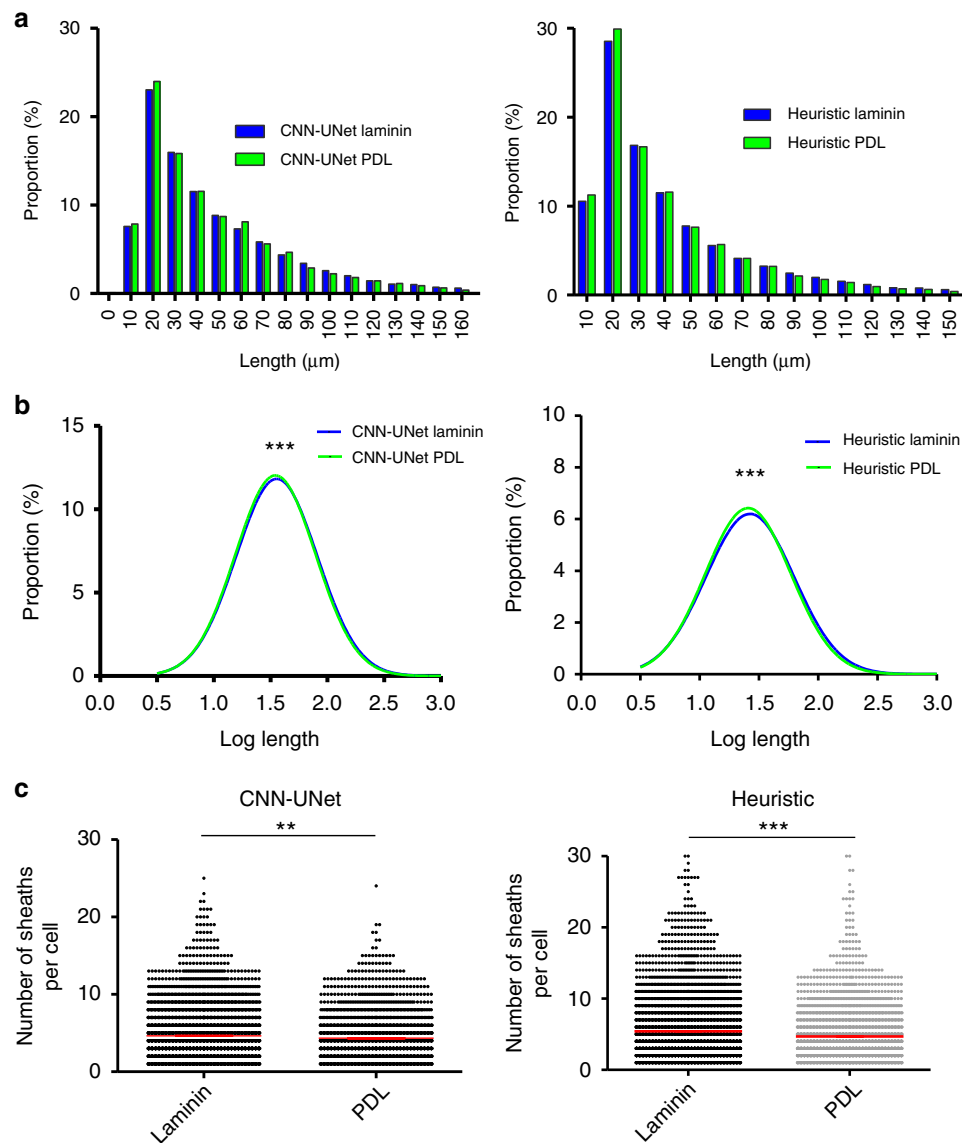
Fig. 6 Detecting biological differences using the CNN-UNet and heuristic algorithm. **a** Absolute and **b** log length distributions had a small but significant difference in means of 0.0137 ± 0.0037 log units ($p = 0.0002$, Student's T-test) as detected by the CNN-UNet, and a difference in means of 0.018 ± 0.0032 log units ($p < 0.0001$) as detected by the heuristic algorithmic approach. **c** Both systems detected increases in the number of sheaths formed per cell between laminin and PDL coated nanofibers. The CNN-UNet found a significant difference in means of 0.30 ± 0.08 sheaths ($p < 0.0018$, Mann–Whitney U-test), while the heuristic algorithm identified a significant difference in means of 0.74 ± 0.10 sheaths ($p < 0.0001$) between the two conditions. Cells were pooled from six rats and cultured as 10 duplicate wells per coating condition. Error bars and midline show mean ± SEM. See also Supplementary Table 2

Discussion

To address the need for high-throughput systems that can quantify multiple morphological properties of oligodendrocytes and associate them with specific cells, we developed an analytic pipeline using nanofiber cultures, automated imaging, and quantitative algorithms. We demonstrate that this analytic paradigm has the capacity to detect finely nuanced biological differences and to match the quality of quantification carried out by human experts.

The biological and computational techniques employed in our experimental system have unique advantages and limitations. The reduced complexity of the nanofiber culture system, which lacks axons and other factors that may influence sheath stability and extension[14–16], is advantageous to isolate oligodendrocyte intrinsic responses while also substantially increasing the reproducibility of quantification for high-throughput screening.

Oligodendrocytes cultured on nanofibers broadly recapitulate the temporal and spatial progression of differentiation seen in vivo[6], forming compact myelin-like MBP positive layered sheaths that wrap nanofibers[7]. Further, the parallel arrangement of the nanofibers facilitated the development of our analytic programs, as it was not necessary to disentangle myelin-like sheaths crossing-over one another, as occurs in typical axon-oligodendrocyte co-cultures. Nanofiber cultures are also highly scalable, allowing many different conditions to be analyzed in parallel.

A final advantage of the nanofiber system is that it facilitates the assessment of multiple biologically relevant parameters, such as the number of sheaths formed per cell, the distribution of sheath lengths and widths, and the proportion of cells that form ensheathments[2,17]. Access to quantitative information for multiple parameters per cell provides key insight into which cellular

Table 1 Comparison of human and machine approaches for quantifying oligodendrocyte ensheathment

	Time per cell (s)	Time per well (min)	Time 96 wells (days)[a]	Benefits	Limitations
Human	80	400[b]	~26.7	Adaptable Fast learner	Variable Slow and costly
MATLAB	–	20[c]	~1.3	Fast Standardization User customizability	Inaccurate due to approximations of cell-by-cell analysis
CNN-UNet	1.5	50[d]	~3.3	Moderate – Fast speed Standardization Adaptable given time to train Analysis at level of single-cells	Uncertain generalizability across variations in image quality when given insufficient training data Training time can be long
CNN-UNet (GPU accelerated)[e]	0.5	15[d]	~1		

[a]Analysis time similar for 384-well plates due to identical plate area
[b]Assume ~300 ensheathed cells per well
[c]Program cuts single-well into 16 sections, taking ~1 min on each section
[d]Assume ~2000 candidates per well
[e]NVIDIA GTX 1070 graphics card

Table 2 Comparison of imaging time

	Time 96 wells (h)	Benefits	Limitations
Manual imaging	~77[a]	Able to selectively focus	Extremely slow if imaging whole wells Biased by human field selection
Automated imaging	~16	Fast Whole-well imaging eliminates field selection bias	Automated focus may not always focus on the best stack

Automated imaging approach reduces manual labor significantly, while also reducing capture time due to automated focus and pre-defined acquisition area
[a]16 fields with 4–5 z-stacks per well. Roughly 3 min/field

process may be affected by any given treatment. For example, changes in the number of ensheathing cells implies an effect on OPC proliferation/differentiation, whereas changes in sheath length suggest altered process extension or stabilization. Acquiring multiple parameters may reveal unexplored correlations, such as between the mean number of sheaths per cell and sheath length per cell. In contrast, other existing high-throughput systems often trade multi-parametric analysis for speed. For example, a multi-well assay that measures the MBP intensity of oligodendrocyte ensheathments around single micropillars obtains a high-throughput level of analysis, but discards information pertaining to sheath length and number[8]. Other automated methods to analyze myelin have focussed on g-ratio quantification[18–24] or the segmentation of myelin globally in images, without associating internodes with specific cells[25–27]. The automated pipeline we developed retains analytic speed while extracting multiple parameters from individual oligodendrocytes, allowing the detection of nuanced physiological processes that can then be further studied in more complex systems.

Both computational approaches presented offer advantages beyond reducing human workloads and extracting multiple morphological features. For example, both algorithms provide standardized reproducible analysis, eliminating the human variability that may result from internal biases, mistakes, and even environmental factors, such as screen brightness, background lighting, and time-constraints. Analytic speed may be enhanced nearly indefinitely by employing parallel computing workstations and faster GPUs. Limiting factors become the time for sample preparation, rather than daunting hours of manual analysis. Finally, the architectural design of the CNN-UNet also offers a substantial advantage. Compared to the original UNet

architecture proposed by Ronneberger et al.,[10] which scaled to 1024 filters in the deepest layers, our reduced network structure is computationally more efficient, scaling to only 50 filters in the deepest layer, allowing our network to run in reasonable time on common CPUs.

Aside from similar advantages, the two programs differed in their accuracy and speed. The heuristic approach is limited as it uses computational approximations to extract global parameters from an image and then attempts to mimic cell-by-cell analysis by associating ensheathments with nearby nuclei. In contrast, the CNN-UNet provides a more human-like assessment of nanofibers by performing high-level decisions using spatial and object-specific information simultaneously to associate specific parameters with individual cells. Although the heuristic algorithm is less accurate in this capacity, it is important to highlight that the approach is still consistent, fast, and responds more readily to user-specified parameters. The heuristic algorithm is particularly appropriate for studies composed of insufficient data to train a neural network, or for analyzing data-sets containing high-variability that require rapid optimization of analytic algorithms.

A final consideration is the generalizability of these programs. While the heuristic algorithm was designed explicitly for aligned nanofiber cultures, it has the immediate capacity to analyze other linear sheath-like objects. The major obstacle to this generalizability is that such an application may require a researcher to have programming experience to alter the program to fully suit their needs. In contrast, the CNN-UNet can match human segmentations on a variety of computational problems without substantially altering the algorithm. For example, we demonstrated that the masked-nucleus training method can generate a CNN-UNet capable of segmenting aligned oligodendrocytes in

multiple orientations (Fig. 5), suggesting that this technique may be applied to more complex oligodendrocyte-neuron co-culture systems. Given that the CNN-UNet performs well with evenly-distributed parallel sheaths in our nanofiber culture system, this quantitative approach could be reasonably applied to tissue sections from CNS regions with an aligned geometry. The major limiting factor for generalizability with the CNN-UNet is the time required to re-train the network and to create ground-truth data for supervised learning. While training is ideally only performed once using a variety images containing individually masked cells, the full capacity of the CNN-UNet to generalize across larger variations in experimental preparations and imaging systems remains to be determined.

In this study, we have developed a powerful paradigm for high-throughput screening of oligodendrocyte ensheathments through the combination of state-of-the-art deep learning technology and readily available imaging and cell culture systems. Our methods extract multiple morphological parameters from individual oligodendrocytes, preserve human analytic quality, remove bias and variability, and facilitate massive increases in sample size and analytic speed. These new developments promise to advance the understanding of fundamental oligodendrocyte physiology and accelerate the discovery of new drugs for the treatment of demyelinating diseases.

Methods

Oligodendrocyte progenitor isolation and differentiation. The study was conducted in accordance with the guidelines of the Canadian Council for Animal Care and approved by the McGill University Animal Care Committees. Oligodendrocyte progenitor cells (OPCs) were isolated from P2 Sprague-Dawley rat brains via a protocol adapted from McCarthy and de Vellis[28]. Cortices of 12 pups from a single litter were pooled and digested in 0.025% trypsin-EDTA and 0.2 mg/mL DNAseI prior to mechanical dissociation via trituration. Dissociated cells were then plated in PDL-coated T75 flasks in DMEM containing 10% FBS and 1% penicillin/streptomycin (Gibco) and cultured for 2–3 weeks at 37 °C with 5% CO_2 and full media changes every 2 days. In these conditions, glial populations rapidly expand to confluency, permitting isolation of OPCs from other cell types based on differential adhesion. Media in the flask was replaced with 37 °C 0.01% trypsin-EDTA in HBSS for 5 min, then swirled and aspirated to remove weakly-adherent microglia. DMEM was then added and the flasks were vigorously hit 20 times with a styrofoam box to detach OPCs while leaving most astrocytes attached. The OPC-containing supernatant was then passed through 70 μm filters and centrifuged for 5 min at $677 \times g$ using a swinging bucket rotor in a Centra CL2 centrifuge (Thermo Fisher), and resuspended in oligodendrocyte defined medium (OLDEM) prior to cell counting and seeding in nanofiber plates. OLDEM consisted of DMEM with 5 μg/ml insulin, 100 μg/ml holotransferrin, 30 nm sodium selenite, 30 nm triiodothyronine, 6.3 ng/ml progesterone, 16 μg/ml putrescine, 100 U/ml penicillin, 100 μg/ml streptomycin, 0.1% fetal bovine serum, and 2% GlutaMAX (Thermo Fisher; all other components from Sigma).

Electrospun nanofiber culturing and immunostaining. 96-well plates with Mimetix aligned ploy-L-lactic acid electrospun nanofibers (AMS.TECL-005-8X, The Electrospinning Company), were pretreated with 20% EtOH in water for 5 min prior to coating with 5 μg/mL PDL in PBS for 1 h at room temperature (RT), followed by 2 h at 37 ºC with 10 μg/μL recombinant laminin 1 (Sigma) in HBSS. OPCs were seeded at 6000 cells per well in OLDEM and cultured at 37 ºC with 5% CO_2 for 7 or 14 days, with full media changes every 2 days or half changes every 4 days, for O4 and MBP-based experiments respectively. After the stated culture period, live cells were incubated with the O4 antibody (1:200; R&D) for 15 min at 37 ºC and then fixed by adding 100 μL of chilled 4% paraformaldehyde in 20% sucrose to each well for 15 min at RT. For MBP staining, cells were fixed the same way and then blocked and permeabilized for 15 min with 5% heat-inactivated normal horse serum in 0.1% Triton X-100 at RT, and then incubated with a primary chicken anti-MBP antibody in 5% HINHS (1:1000; Aves labs) for 4 h at RT. The oligodendrocytes were then washed in PBS before being incubated with IgM Cy3 for O4 stained cells, or Alexa-546 conjugated goat anti-chicken secondary antibody (1:1000; Molecular Probes) for anti-MBP stained cells, in addition to 1 μg/mL of DNA-staining Hoechst 33342 (Sigma) for 2 h at RT in 5% HINHS. Cells were washed again with PBS and kept at 4 ºC until they were imaged.

Automated imaging. Nanofiber plates were imaged using an LSM 880 inverted confocal laser scanning microscope (Carl Zeiss) with a 10X (NA 0.45) Plan-Apochromat objective. Alexa-555-conjugated secondary antibodies (Thermo-Fisher) were used to ensure minimal spectral overlap so that Hoescht and MBP

signals could be acquired on a single track to reduce acquisition time. Acquisitions were automated using Zen Black Systems v2.3 and the Tiles and Positions software module. For each experiment, the user is only required to set the PMT gain to account for differences in staining intensity and to calibrate the motorized stage to target the center of the first well of each plate. In 16 h of acquisition, the system imaged 4×4 frames centered in the middle of each well to cover a ~4×4 mm area with a pixel resolution of 0.5 μm, thereby acquiring >50% of every well with no bias. Focus was performed automatically for each well with the DefiniteFocus.2 module (Carl Zeiss). Since the nanofibers are not perfectly planar, three to five confocal slices covering 10–20 μm in the Z-dimension were acquired for each frame to ensure that all cells were sampled. At lower resolutions, each well took <10 min to acquire, allowing the whole plate to be imaged in <16 h. Z-stacks were stitched and compressed via the Extended Depth of Focus (maximum projection) algorithm in Zen Black prior to exporting the images as TIFF files.

Manual segmentation criteria. A set of criteria was defined for classification of oligodendrocyte ensheathments based on morphological characteristics visualized using Hoechst, O4, and MBP staining: (1) The presence of an identifiable cell body and nucleus; (2) The presence of ensheathing processes, defined as segments of O4 or MBP running parallel to the nanofibers in the culture plate having a length > 12 μm and a thickness > 2 μm; (3) Continuity of the cell body with the ensheathments; (4) Each ensheathed nanofiber may only be associated with 1 cell nucleus. Human researchers were given these criteria and instructed to trace the border of oligodendrocyte ensheathments using the ImageJ polygon tool. These manual segmentations were then used to train and validate the deep learning network.

Heuristic segmentation algorithm. A heuristic algorithm, written in MATLAB, was first constructed to implement the manual segmentation criteria. To fulfill criteria 1, cell bodies were identified through colocalization of Hoechst stained nuclei with O4 or MBP-positive cytoplasm. To separate overlapping and crowded cell bodies, the local maxima of a distance transformed O4 or MBP stained image were used for watershed segmentation. For criteria 2, after convolution with second derivative Gaussian filters, ensheathed nanofibers were segmented using a ridge-filter that extracts the eigenvalues from the Hessian matrix of the O4 or MBP stained image. The sensitivity of segmentation was user specified and the diameter of fibers segmented could be modified by altering the sigma of the initial Gaussian filters. The algorithm then cycled through each cell body to associate segmented fibers with corresponding nuclei, thereby satisfying criteria 3 and 4 (Fig. 1b). An important limitation of this heuristic algorithm is that the identification of ensheathments and the association of segments to individual cells is performed as two separate quantitative steps. The ridge-filter is applied to extract ensheathments globally from an image and then the local distance maxima of cell bodies is used to associate those segments to individual oligodendrocytes based on proximity to mimic a cell-by-cell level of quantification (Fig. 1b). By separating the analysis into two streams, however, the algorithm does not perform relatively high-level decisions required to reliably associate ensheathments to specific cells when faced with over-crowded images or when fluorescent staining is faded. In contrast, the deep learning approach overcomes these challenges by using spatial and object-specific information simultaneously to analyze and associate oligodendrocyte ensheathments with individual cells.

Neural network training data. To train a neural network to extract morphological features from individual oligodendrocytes, we generated 3-channel images for our input data, with the middle channel replaced by a mask over a specific nucleus of interest (Fig. 2a). This mask prompted the network to learn to associate oligodendrocyte ensheathments with individually masked nuclei. The input data was also appended with a ground-truth image, containing the human tracings of oligodendrocyte ensheathments. To create the ground-truth images, human researchers: (1) manually traced a single cell nucleus with its corresponding ensheathments, saved that as an ROI file in ImageJ, and then repeated the process for every cell in a whole-well image; (2) The ROIs were then used to generate two 16-bit images, one containing all manually traced nuclei, and the other containing all manually traced ensheathments; (3) A Python script then looped through the nuclei-containing image, using each nucleus as a center to crop/create input and ground-truth images for every cell in a full-well image. Cropping input data using individual cell nuclei increased the number of training images, generating ~40,000 input examples from a dozen full-well images.

Neural network architecture. Google's open source Tensorflow[29] was used to implement a CNN with a UNet architecture[10]. The U-shaped design was selected for its capacity to learn to associate ensheathments with individual oligodendrocytes, as the down-sampling arm of the network learns features at a variety of scales, and the up-sampling arm combines spatial information from earlier layers to produce a high resolution final segmentation (Fig. 2b). The down-sampling branch contained five two-dimensional convolutional layers with 5×5 pixel kernels and a stride of 2. The number of filters at each subsequent layer increased by 10 sequentially from 10 to 50. The up-sampling branch of the network was symmetric to the down-sampling branch but used fractional strides to up-sample the information

back to the original size. The up-sampling convolutional layers had 5×5 pixel kernels and a decreasing number of filters that dropped sequentially by 10 from 50 to 10. The up-sampling branch also mixed in spatial information by concatenating data from earlier layers with similar resolution prior to each convolution. The output of the up-sampling branch was then passed through a softmax convolutional layer with a 1×1 pixel kernel to generate two mutually exclusive probabilistic output categories corresponding to background and fibers. This probabilistic output was then thresholded at a value of 0.5 to produce a binary segmentation containing the ensheathments identified by the network for a single cell.

Pre-processing data for CNN-UNet analysis. An image first undergoes pre-processing by watershed segmentation to identify Hoechst stained nuclei. These nuclei are then used to generate "input" images (Fig. 2a). To avoid redundancy, we recognized that some cells are clearly not ensheathing—defined as having zero positively-stained cell membrane markers, such as MBP, near the cell nucleus. Such cells are excluded to save analytic time by binarizing the cell membrane image followed by morphological closing, then opening to generate blobs. Only nuclei within cell membrane blobs were considered "candidates" for quantification, thereby avoiding analysis of unstained cells.

Post-processing after CNN-UNet analysis. After the network cycles through each candidate cell within a full-well image, we applied three post-processing steps to assign overlapped ensheathments—associated to more than one cell—while removing sticky connections (see below) and setting a minimum threshold for ensheathment length (Fig. 2b). The first post-processing step is a method of unfair distribution. This method operates under the assumption that it is more likely that the uncertain segment will belong to the larger cell, given the choice to assign an ambiguous overlapped ensheathment to a large cell with multiple ensheathed nanofibers, or a small cell with only a single ensheathment, Next, we applied a sticky-removal method. To separate sticky-segments—adjacent ensheathments that overlap and are wrongly identified as a single object—the method skeletonized the output image and removed all branch points. It then cycled through all the broken segments to eliminate any that were horizontal. Branch points were then re-inserted to restore the vertical connectivity of the ensheathments without their sticky horizontal connections. Lastly, we applied a user-specified min-length threshold, where ensheathments below a minimum length (~12 μm) were eliminated. As well, since we noticed that human non-experts exhibited a slight over-sensitivity toward cells with only one or two ensheathments, we programmed a sensitivity measure to reduce variability across experiments. We adopted a length threshold that is three to four times more strict for single and doubly-ensheathing cells to reduce the chance of identifying false positives, but this value can be adjusted to suit a researcher's needs. Length was determined by fitting the objects with an ellipse and calculating the distance of the major axis, as defined in *regionprops* OpenCV[30].

Training. To train a deep learning model to match the cell-by-cell analysis performed by human researchers, we presented the network with ~40,000 input images that each contained a masked nucleus of interest (Fig. 2a, b). 1111 of these cropped images were utilized for validation during training, allowing the performance of the model to be followed as it learned. The performance of our CNN-UNet in task-learning was monitored through the loss value and JI. Loss functions serve as a numerical representation of the magnitude of *error* for a prediction made by the CNN-UNet. Here, we used the categorical cross-entropy, on a per-pixel basis, as the loss function to measure the difference between the output of the network and the manually traced ensheathments generated by cell biologists (ground truth). In our model, a decrease in loss and increase in JI, calculated on the set of 1111 validation images, served as an indicator of successful task-learning and the generalizability of the network to new data (Fig. 3a, b).

To improve learning speed and accuracy, we also trained two CNNs with differently weighted loss functions. We implemented "class weighting" for the first network, whereby the loss at every pixel corresponding to ensheathments was multiplied by a factor of 10. This placed more emphasis on decisions corresponding to oligodendrocyte sheaths, thereby helping to balance the training dataset by increasing the importance of the small number of images that contained ensheathing cells. For the second network, in addition to class weighting, we also implemented "spatial weighting" to enhance the sensitivity of the network to learn about the edges of objects, weighing the loss function with the following exponential decay formula:

$$w = e^{1 - \frac{D}{f} + b} \tag{1}$$

w refers to the weighted loss at each pixel, f modifies the rate of falloff from the edge, and b prevents the weights from reaching zero. D is the distance from the edge of the oligodendrocyte ensheathments, calculated by computing the Chebyshev distance transform of the inverted ground-truth mask. Overall, this formula weights the training loss such that it decays exponentially with distance from the margins of the ensheathed segments of interest, thus helping the network focus on correctly identifying the edges of objects (Fig. 3e). Although both weighting schemes dramatically improved CNN performance, indicated by an earlier and steeper increase in JI, the spatially weighted network was superior, as its

JI continued to improve while the class weighted network plateaued at ~180,000 training epochs (Fig. 3b). Using these parameters, along with testing at progressive checkpoints, we found the best model was generated at ~301,000 epochs by the spatially weighted network (Fig. 3a, b). Finally, after training was performed for 350,000 epochs (about 6 days) on an NVIDIA Tesla P100-PCIE, one entire whole-well image (containing 1667 input cells traced by a human researcher) was used to test the network for its performance on global parameters and generalizability to new data.

Statistics. Data are presented as mean ± standard error of the mean (SEM). Significance was set at $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Statistical analyses were completed with Graphpad Prism 5 and all Student's t-tests are two-tailed. The log of ensheathment length was used as described[7] previously because sheath length data was log normal (log lengths were Gaussian distributions). Power analysis was performed using Lenth's power calculator software[12,13].

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Data availability
The datasets generated and analyzed in the current study are available from the corresponding author on request.

Code availability
The CNN-UNet resources in this paper will be made available through www. deepdiscovery.org. Early access, useable version of both the heuristic and CNN-UNet algorithms are available at www.github.com/yxu233/Myelin.

References

1. Franklin, R. J. M. Why does remyelination fail in multiple sclerosis? *Nat. Rev. Neurosci.* **3**, 705 (2002).
2. Jarjour, A. A., Zhang, H., Bauer, N., Ffrench-Constant, C. & Williams, A. In vitro modeling of central nervous system myelination and remyelination. *Glia* **60**, 1–12 (2012).
3. Thomson, C. E. et al. Myelinated, synapsing cultures of murine spinal cord--validation as an in vitro model of the central nervous system. *Eur. J. Neurosci.* **28**, 1518–1535 (2008).
4. Stoppini, L., Buchs, P. A. & Muller, D. A simple method for organotypic cultures of nervous tissue. *J. Neurosci. Methods* **37**, 173–182 (1991).
5. Bansal, R., Warrington, A. E., Gard, A. L., Ranscht, B. & Pfeiffer, S. E. Multiple and novel specificities of monoclonal antibodies O1, O4, and R-mAb used in the analysis of oligodendrocyte development. *J. Neurosci. Res.* **24**, 548–557 (1989).
6. Lee, S. et al. A culture system to study oligodendrocyte myelination processes using engineered nanofibers. *Nat. Methods* **9**, 917–922 (2012).
7. Bechler, MarieE., Byrne, L. & ffrench-Constant, C. CNS myelin sheath lengths are an intrinsic property of oligodendrocytes. *Curr. Biol.* **25**, 2411–2416 (2015).
8. Mei, F. et al. Micropillar arrays as a high-throughput screening platform for therapeutics in multiple sclerosis. *Nat. Med.* **20**, 954–960 (2014).
9. Espinosa-Hoyos, D. et al. Engineered 3D-printed artificial axons. *Sci. Rep.* **8**, 478 (2018).
10. Ronneberger, O., Fischer, P. & Brox, T. U-Net: convolutional networks for biomedical image segmentation. *arXiv*:1505.04597v1 (2015).
11. Jaccard, P. The distribution of the flora in the alpine zone. *New Phytol.* **11**, 37–50 (1912).
12. Lenth, R. V. Java Applets for Power and Sample Size. http://www.stat.uiowa.edu/~rlenth/Power/ (2018).
13. Quertemont, E. How to Statistically Show the Absence of an Effect. *Psychol. Belg.* 109–127 (2011).
14. Krasnow, A. M., Ford, M. C., Valdivia, L. E., Wilson, S. W. & Attwell, D. Regulation of developing myelin sheath elongation by oligodendrocyte calcium transients in vivo. *Nat. Neurosci.* **21**, 24–28 (2018).
15. Michailov, G. V. et al. Axonal Neuregulin-1 Regulates Myelin Sheath Thickness. *Science* **304**, 700–703 (2004).
16. Wake, H., Lee, P. R. & Fields, R. D. Control of local protein synthesis and initial events in myelination by action potentials. *Science.* **333**, 1647–1651 (2011).
17. Zhang, H., Jarjour, A. A., Boyd, A. & Williams, A. Central nervous system remyelination in culture--a tool for multiple sclerosis research. *Exp. Neurol.* **230**, 138–148 (2011).

18. Moiseev, D., Hu, B. & Li, J. Morphometric analysis of peripheral myelinated nerve fibers through deep learning. *J. Peripher. Nerv. Syst.* https://doi.org/10.1111/jns.12293. [Epub ahead of print] (Nov 28, 2018).

19. Naito, T. et al. Identification and segmentation of myelinated nerve fibers in a cross-sectional optical microscopic image using a deep learning model. *J. Neurosci. Methods* **291**, 141–149 (2017).

20. Novas, R. B., Fazan, V. P. & Felipe, J. C. A new method for automated identification and morphometry of myelinated fibers through light microscopy image analysis. *J. Digit. Imaging* **29**, 63–72 (2016).

21. Isaacs, J., Mallu, S. & Batchelor, M. Modification of commercially available image analysis software for semi-automated qualitative analysis of axon regeneration and myelination in the rat sciatic nerve. *J. Neurosci. Methods* **233**, 45–49 (2014).

22. Hunter, D. A. et al. Binary imaging analysis for comprehensive quantitative histomorphometry of peripheral nerve. *J. Neurosci. Methods* **166**, 116–124 (2007).

23. Zaimi, A. et al. AxonSeg: open source software for axon and myelin segmentation and morphometric analysis. *Front. Neuroinform.* **10**, 37 (2016).

24. More, H. L., Chen, J., Gibson, E., Donelan, J. M. & Beg, M. F. A semi-automated method for identifying and measuring myelinated nerve fibers in scanning electron microscope images. *J. Neurosci. Methods* **201**, 149–158 (2011).

25. Çimen, S. et al. In 26th European Signal Processing Conference (EUSIPCO) 61–65 (Rome, Italy; 2018).

26. Kerman, B. E. et al. In vitro myelin formation using embryonic stem cells. *Development* **142**, 2213–2225 (2015).

27. Kreshuk, A. et al. Automated tracing of myelinated axons and detection of the nodes of Ranvier in serial images of peripheral nerves. *J. Microsc.* **259**, 143–154 (2015).

28. McCarthy, K. D. & de Vellis, J. Preparation of separate astroglial and oligodendroglial cell cultures from rat cerebral tissue. *J. Cell Biol.* **85**, 890–902 (1980).

29. Abadi, M. et al. TensorFlow: large-scale machine learning on heterogeneous distributed systems. *arXiv*:1603.04467v2 (2016).

30. Bradski, G. The OpenCV Library. *Dr. Dobb's Journal of Software Tools.* http://www.drdobbs.com/open-source/the-opencv-library/184404319. (2000).

Author contributions

Y.K.T.X., D.C., R.A.B., Q.L.Q., J.P.A., and T.E.K. contributed to the experimental design and conceptualization of the project. Y.K.T.X., D.C., M.A.D., and Q.L.Q. carried out the experiments. Y.K.T.X. and D.C. wrote the original draft. All authors reviewed and edited the manuscript. Funding obtained by J.P.A. and T.E.K. supported the project.

Additional information

Supplementary information accompanies this paper at https://doi.org/10.1038/s42003-019-0356-z.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.