Preprint

EntireAxon: Deep learning deciphers four distinct patterns of axonal degeneration

Alex Palumbo[1,2,3,*] (ORCID: 0000-0002-1631-785X), **Philipp Grüning**[4] (ORCID: 0000-0003-2946-4020), **Svenja Kim Landt**[1,2] (ORCID: 0000-0002-4133-7035), **Lara Eleen Heckmann**[1,2] (ORCID: 0000-0003-1264-9715), **Luisa Bartram**[1], **Alessa Pabst**[1,2] (ORCID: 0000-0003-1938-3753), **Charlotte Flory**[1,2] (ORCID: 0000-0001-5049-6380), **Maulana Ikhsan**[1,2,3] (ORCID: 0000-0002-8756-4524), **Sören Pietsch**[1,2,5], **Reinhard Schulz**[6], **Christopher Kren**[7], **Norbert Koop**[7], **Johannes Boltze**[1,2,8] (ORCID: 0000-0003-3956-4164), **Amir Madany Mamlouk**[4] (ORCID: 0000-0001-9709-1620) **and Marietta Zille**[1,2,3,*] (ORCID: 0000-0002-0609-8956)

[1] Fraunhofer Research Institution for Marine Biotechnology and Cell Technology, Lübeck, Germany

[2] Institute for Medical and Marine Biotechnology, University of Lübeck, Lübeck, Germany

[3] Institute for Experimental and Clinical Pharmacology and Toxicology, University of Lübeck, Lübeck, Germany

[4] Institute for Neuro- and Bioinformatics, University of Lübeck, Lübeck, Germany

[5] Department of Neonatology, Universitätsklinikum Leipzig, Leipzig, Germany

[6] Wissenschaftliche Werkstätten, University of Lübeck, Lübeck, Germany

[7] Medical Laser Center Lübeck GmbH, Lübeck, Germany

[8] School of Life Sciences, The University of Warwick, Gibbet Hill Campus, Coventry, United Kingdom

[*] Correspondance to: alex.palumbo@uni-luebeck.de; m.zille@uni-luebeck.de

Abstract

Different axonal degeneration (AxD) patterns have been described depending on the biological condition. Until now, it remains unclear whether they are restricted to one specific condition or can occur concomitantly. Here, we present a novel microfluidic device in combination with a deep learning tool, the EntireAxon, for the high-throughput analysis of AxD. We evaluated the progression of AxD in an *in vitro* model of hemorrhagic stroke and observed that axonal swellings preceded axon fragmentation. We further identified four distinct morphological patterns of AxD (granular, retraction, swelling, and transport degeneration) that occur concomitantly. These findings indicate a morphological heterogeneity of AxD under pathophysiologic conditions. The newly developed microfluidic device along with the EntireAxon deep learning tool enable the systematic analysis of AxD but also unravel a so far unknown intricacy in which AxD can occur in a disease context.

Introduction

Axonal degeneration (AxD) is a process in which axons disintegrate physiologically during nervous system development and aging, or as a pathological element of degenerative nervous system diseases[1–3]. During development and neural circuit assembly, inappropriately grown axons can undergo axonal retraction, axonal shedding or local AxD[4,5]. Axonal retraction is characterized by retraction bulb formation at the distal tip, and subsequent pullback[4]. During axonal shedding, the axon sheds small vesicles, called axosomes[6]. Local AxD is characterized by axon disintegration into separated axonal fragments[5].

Acutely and chronically injured axons may degenerate retrogradely (dying-back mechanism), anterogradely or in a Wallerian degeneration pattern, ultimately resulting in the generation of axonal fragments[7–9]. Another morphological feature that occurs during AxD in development and disease is axon swelling. The swollen axon contains disorganized cytoskeleton and organelles resulting from an interruption of axonal transport, reactive oxygen species exposure or nutrient deprivation[8,10,11]. Both axonal swellings and axonal fragments represent hallmarks of AxD independent of the biological context[2,12,13]. Therefore, it is relevant to detect the different degeneration patterns and to ascertain their underlying morphological changes to better understand their role in AxD during development and pathophysiology.

To assess the spatiotemporal progression of the different AxD patterns, morphological changes of axons must be recorded continuously. However, available software fails to automatically detect and quantify high axon numbers and the morphological features of AxD (swellings and fragments). The reason may be two-fold: 1) Available software relies on image binarization[14,15], which can lead to information loss and low sensitivity as thin axons may not be recognized. 2) The analysis requires subjective and time-consuming manual annotations, e.g., thresholding and defining the region of interest[16–18]. So far, immunostained images were used to investigate morphological changes in AxD as the analysis of phase-contrast images has been limited by the lower target-to-background signal. Immunofluorescence images, however, entail certain

disadvantages such as photobleaching and the requirement for cell fixation, which allows only to observe a single time point. Hence, a software tool for the automatized detection and quantification of the morphological patterns of AxD in long-term live cell imaging is required to improve the sensitivity and throughput, which will overcome current limitations in understanding AxD.

In this study, we hypothesized that axons can undergo different AxD patterns, which rely on different underlying morphological features and depend on the severity of AxD. To systematically study AxD in a high-throughput manner, we fabricated a microfluidic device in the format of a conventional cell culture plate containing 16 independent microfluidic units that separate axons from their somata. We also developed a novel deep learning tool called "EntireAxon" to analyze AxD in phase-contrast time-lapse recordings. Our newly developed detection algorithm can sensitively and specifically recognize axons, axonal swellings and axonal fragments during AxD in an *in vitro* model of hemorrhagic stroke and its performance is equal to human expert rating. EntireAxon enabled the identification of four different morphological patterns of AxD based on a histogram analysis of segmentation changes over time. These patterns occurred concomitantly and in a concentration-dependent manner in hemin-induced AxD.

Results

A novel microfluidic device for the high-throughput analysis of axonal degeneration

Experimental models such as microfluidic devices have been developed to spatially separate axons from their somata[19,20]. Thus, the axons can be treated in an isolated manner to investigate AxD. The major limiting factor of commercially available microfluidic devices is that they are single, individual systems and hence, can only be used to assess one condition, which is time-consuming and precludes high-throughput analyses.

To overcome this limitation, we manufactured a microfluidic device containing 16 individual microfluidic units by using the soft lithography technique replica molding (**Fig. 1, suppl. Fig. 1**). The spatial separation of axons from their somata was confirmed by immunostaining (**Fig. 1c**) using the established dendritic marker MAP2 and axonal marker synaptophysin[21,22].

The EntireAxon convolutional neural network (CNN) recognizes axons and the morphological features of AxD

To enable the systematic high-throughput analysis of AxD over time in phase-contrast microscopy, we trained the EntireAxon CNN to segment all relevant features of AxD, i.e. axons, axonal swellings, and axonal fragments (**Fig. 2**).

Specifically, we adapted a standard u-net with ResNet-50 encoder[23,24] and used a CNN ensemble, which combines predictions from multiple CNNs to generate a final output and is superior to individual CNNs[25–27]. To validate the training success, a separate labeled validation set unknown to the EntireAxon was created (ground truth was labeled by human expert 1). While the EntireAxon CNN recognized the class 'background' better than the three axon classes 'axon', 'axonal swelling', and 'axonal fragment' (mean F1 score: 0.995), axon-specific segmentation revealed the highest mean F1 score for the class 'axon' (0.780), followed by the classes 'axonal swelling' (0.567), and 'axonal fragment' (0.301) (**Fig. 3a**). The comparably

lower performance of the CNN to recognize axonal fragments may be explained by the disproportional distribution of pixels in the training data with a mean of 96.42 % of pixels belonging to the class 'background', while only 2.77% represented the class 'axon', 0.58% 'axonal swelling', and 0.23 % 'axonal fragment'.

Next, we compared the performance of the trained EntireAxon CNN on the ground truth (human expert 1) with the performance of two additional human experts. The EntireAxon CNN reached higher mean F1 scores for all classes, except for the class 'axonal fragment', where human expert 2 outperformed the EntireAxon CNN (**Fig. 3b**). This may have been due to the fact that the EntireAxon CNN was trained on images labeled by the same human expert (1) that labeled the ground truth. To assess whether its performance is more generalizable across the different experts, we compared the EntireAxon CNN to each of the human experts on the consensus labels of the two other human experts (**Fig. 3c-d**). Visual inspection of the labels showed a wide overlap between the different experts, but also that there was considerable uncertainty, especially for the classification of axonal fragments (**Fig. 3c**). When comparing the mean F1 scores for all classes, the EntireAxon reached similar or even higher scores than the other three experts (**Fig. 3d**). Collectively, this suggests that the EntireAxon CNN sensitively and specifically recognizes axons and the morphological features of AxD, i.e. axonal swellings and axonal fragments.

AxD progresses concentration- and time-dependent

We then applied the EntireAxon CNN to assess AxD in an *in vitro* model of hemorrhagic stroke[28]. Accordingly, axons were exposed to the hemolysis product hemin and recorded by time-lapse microscopy for 24 hours. Hemin induced concentration- and time-dependent morphological changes leading to AxD compared to vehicle-treated axons (**Fig. 4 and suppl. Videos 1-4**). Area under the curve (AUC) analyses revealed a significant decrease in axon area in all three hemin concentrations (50 µM vs. 0 µM: p=0.026; 100 µM vs. 0 µM: p=0.018,

200 µM vs. 0 µM: p<0.001). The axonal swelling area also increased in all three concentrations (50 µM vs. 0 µM: p=0.012, 100 µM vs. 0 µM: p=0.005, 200 µM vs. 0 µM: p=0.016), while the axonal fragment area was elevated only for axons treated with 100 and 200 µM hemin (vs. 0 µM: p=0.004, **suppl. Fig. 2**).

Comparing the time course of AxD between hemin- and vehicle-treated axons (0 µM), the axon area decreased starting at 11.5 hours at 200 µM (p=0.020, from 15 hours onwards p<0.001), at 14 hours at 100 µM (p=0.040, from 18.5 hours onwards p<0.001), and at 15 hours at 50 µM (p=0.018, from 19 hours onwards p<0.001). Hemin treatment also an elevated axonal fragment area starting at 9 hours at 200 µM (p=0.037) and at 17 hours at 100 µM hemin (p=0.044). Interestingly, the axonal swelling area increased prior to the changes in axon and axonal fragment area, i.e. starting at 6 hours at 200 µM (p=0.010) and 100 µM (p=0.019), and at 8 hours at 50 µM hemin (p=0.030). For the highest hemin concentration, the increase was only transient (until 18.5 hours), suggesting that axonal swellings preceded the fragmentation of the axon.

The results of the time course analysis were further substantiated by live cell fluorescent staining (calcein AM), which indicated the starting point of AxD after hemin treatment between 8 and 12 hours for 200 µM hemin, between 12 and 16 hours for 100 µM hemin and 16 and 20 hours for 50 µM hemin (**suppl. Fig. 3**). Taken together, the EntireAxon CNN was able to quantify the degree of AxD and to determine the underlying temporal progression.

The EntireAxon recurrent neural network (RNN) recognizes four distinct patterns of AxD

AxD time series data revealed that different morphological patterns of degeneration can occur in the same axons over time (**Fig. 5** and **suppl. Videos 5-8**). We therefore classified the following four morphological patterns of AxD based on the time-lapse recordings:

i) Granular degeneration: AxD resulting in granular separated fragments.

ii) Retraction degeneration: AxD in which the distal part of the axon retracts ultimately resulting in granular degeneration.

iii) Swelling degeneration: AxD with the enlargement of axonal swellings and subsequent granular degeneration.

iv) Transport degeneration: AxD in which axonal swellings of constant size, which do not enlarge, are transported along the axon resulting in granular degeneration.

We hypothesized that the different morphological AxD patterns can occur concomitantly. We therefore trained the EntireAxon RNN to identify changes in class segregation over time using a training dataset of AxD segmentation recordings (**Fig. 6a**). For each time point, the RNN computed the probability of a change in class for each pixel of an image relative to the previous time point.

Based on the 16 different possible class changes, the RNN determined seven clusters (cluster 0-6) that were characterized by an idiosyncratic segmentation pattern over 24 hours (**Suppl. Fig. 4**). All clusters showed a decrease in the class 'axon' and an increase in the class 'background', although to different degrees, temporal patterns, and some of them with concomitant increases in either the class 'axonal swelling' and/or 'axonal fragment'. In cluster 0, there was an early decrease in the class 'axon', which then continued more linearly as well as a later rise in the class 'axonal fragment'. In contrast to cluster 0, cluster 1 showed no increase in the class 'axonal fragment' and a linear decrease in the class 'axon' from the start. In cluster 2, there was a strong increase in the class 'axonal swelling'. Cluster 3 demonstrated an early and lasting high level of the class 'axonal swelling' with a later increase in the class 'axonal fragment'. Cluster 4 showed a rapid decrease in the class 'axon' concomitant with an increase in the classes 'background' and 'axonal swelling'. Cluster 5 was similar to cluster 1, but with an early drop in the class 'axon'. Cluster 6 showed an increase in the class 'axonal swelling' similar to but to a greater extent than cluster 2.

The RNN categorized each cluster to one of the four morphological patterns (**Fig. 6b**): i) Granular degeneration was defined by clusters that describe the degeneration of axons into axonal fragments, i.e. clusters 0, 1, 3, and 5. ii) Retraction degeneration only included the clusters 1 and 5, indicating the retraction of the axon followed by its fragmentation. iii) Swelling degeneration was characterized by the three clusters that included the class 'axonal swelling, i.e. clusters 2, 3, and 6, as well as cluster 5 showing the exchange of the class 'axon' for 'background'. iv) Transport degeneration was the only pattern that relied on cluster 4 and was also characterized partly on clusters 0, 1, 2, and 6. Although some clusters overlap among morphological patterns, the unique combination of the different clusters allows to distinguish all four morphological patterns.

To validate the EntireAxon RNN, a 10-fold cross-validation was performed, which demonstrated its ability to distinguish between the four morphological patterns of AxD (**Fig. 6c**). These data confirm that the combination of the different AxD features as well as their spatiotemporal progression defines distinct morphological AxD patterns.

The morphological patterns of AxD depend on the extent of AxD

We then applied the EntireAxon RNN to quantify the occurrence of the different morphological patterns of AxD (**Fig. 7**). Hemin-treated axons underwent granular, retraction, swelling, and transport degeneration concomitantly as these patterns occur either close to or at the exact same locations (**Fig. 7a**). Hemin concentration-dependently increased all morphological AxD patterns, with granular degeneration being significantly increased for the 100 µM ($p=0.004$) and 200 µM hemin samples ($p=0.001$), while swelling degeneration was significantly increased for 50 µM ($p=0.006$) and 100 µM ($p=0.004$) hemin-treated compared to vehicle-treated axons. Neither retraction nor transport degeneration differed among the investigated concentration (**Fig. 7b**).

Collectively, our data suggest that hemin gradually shifts the morphological pattern of AxD from swelling degeneration to granular degeneration, which is in accordance with our findings that axonal swellings preceded axonal fragmentation.

Discussion

We here describe for the first time the concomitant occurrence of four different morphological patterns of AxD under pathophysiological conditions, i.e. granular, retraction, swelling, and transport degeneration. These rely on time- and concentration-dependent changes of the morphological features of AxD, with axonal swellings preceding axon fragmentation. The herein introduced complementary tools consisting of a novel microfluidic device and the EntireAxon deep learning platform allow to increase the experimental yield and the in-depth high-throughput analysis of AxD.

The extent of AxD has been mainly investigated with a focus on axon fragmentation so far. To quantify axon fragmentation, Sasaki and colleagues introduced the AxD index as the ratio of fragmented axon area versus total axonal area[14]. However, the AxD index did not include axonal swellings, which is another characteristic feature of degenerating axons[11,29]. Hence, the precise spatiotemporal role of axonal swellings in AxD remains elusive.

The herein presented EntireAxon CNN performs an automatic segmentation and quantification of axons and morphological features relevant to AxD, including axonal swellings and fragments, on phase-contrast time-lapse microscopy images (**Fig. 2**). The EntireAxon CNN recognized the four classes ‘background’, ‘axon’, ‘axonal swelling’, and ‘axonal fragment’, with the highest mean F1 score for the class ‘background’ (**Fig. 3a**) as background pixels covered most of the image.

Although the performance of the EntireAxon CNN on the axon-related classes was comparably lower due to the imbalanced prevalence of each class in the images, the comparison with human experts revealed that the EntireAxon CNN reached a similar performance level. As expected, its performance was slightly better than the human experts on the ground truth as both, ground truth and training data, were labeled by the same human expert (**Fig. 3b**). Interestingly, when comparing the EntireAxon CNN with a human expert on the consensus label of the other two

239 human experts, not only was the EntireAxon CNN as good as or even better than the human
240 expert, but the mean F1 scores were also higher than on the ground truth labels (**Fig. 3d**). This
241 may be because pixels that were differentially assigned by the human expert, i.e. more difficult
242 to classify, were excluded from the comparison. Taken together, these findings demonstrate
243 that the EntireAxon CNN is suitable to quantify automatically and in a high-throughput manner
244 AxD and its accompanying morphological changes.
245 Using the newly developed EntireAxon CNN, we were able to investigate AxD in an *in vitro*
246 model of hemorrhagic stroke. To our knowledge, this is the first *in vitro* study to characterize
247 AxD after hemorrhagic stroke. AxD is an active and commonly observed process in
248 intracerebral hemorrhage[30,31], with larger hematomas being associated with a more severe
249 progression of white matter injury resulting in poor motor and functional recovery[30,32]. As the
250 underlying mechanisms of AxD in hemorrhagic stroke remain to be elucidated and represent a
251 research priority in the field[33], we investigated the morphological changes that can occur in
252 axons in response to hemorrhage.
253 We therefore modeled hemorrhagic stroke by exposing outgrown axons from primary cortical
254 neurons to the hemolysis product hemin. In our model, AxD started within 12 to 18 hours after
255 the administration of hemin (**Fig. 4** and **suppl. Fig. 3**). Similar AxD starting time points have
256 been observed *in vitro* in other models. For instance, under circumstances of growth factor
257 withdrawal, axons undergo AxD at 12-24 hours[11,34,35].
258 We further demonstrated that the relative axon area decreased with increasing hemin
259 concentrations, while the axonal fragment area increased. Our results are in accordance with
260 other experimental conditions such as axotomy-mediated or paclitaxel-induced AxD, in which
261 the axonal fragments also increased[14,36]. As the axonal swelling area preceded the increase of
262 axonal fragments and axon area loss, our findings are also in line with results reported in a
263 model of experimental autoimmune encephalomyelitis indicating that axonal swelling
264 anticipates fragmentation[10].

Interestingly, axonal swellings and axonal fragments were related to different morphological patterns of AxD in our model. Specifically, we observed axons that showed signs of axonal retraction, enlarging of axonal swellings and axonal transport before degeneration (**Fig. 5**). These findings led to the hypothesis that different morphological patterns of AxD can occur concomitantly and should therefore be detectable based on our segmentation data. We therefore trained the EntireAxon RNN to quantify the occurrence of granular, retraction, swelling, and transport degeneration based on the clusters of unique changes of classes over time (**Fig. 6** and **suppl. Fig.4**):

i) Granular degeneration has previously been observed in retrograde, anterograde, Wallerian and local AxD[5,7–9]. ii) Retraction degeneration has been described in axonal retraction and shedding[4,6]. iii) Swelling degeneration was previously reported in experimental autoimmune encephalitis and growth factor deprivation[10,11]. iv) In contrast, transport degeneration has not been reported before. However, microtubule breaks have been demonstrated in an *in vitro* model of axonal stretch injury. Those developed into axonal swellings resulting in axonal transport interruption with AxD as a consequence[37]. The underlying molecular mechanisms of the different patterns of AxD need to be further investigated, which could be greatly facilitated by the EntireAxon RNN that is able to automatically detect the different morphological patterns in time-lapse recording due to its capacity to relate each output to previous images in the stacks by its current units.

All four patterns occurred concomitantly in hemin-induced AxD (**Fig. 7**), which may point to a potential synergy of the different patterns of AxD. Interestingly, we also observed a concentration-dependent effect, whereat swelling degeneration was significantly increased at lower hemin concentrations while granular degeneration occurred more frequently at higher hemin concentrations. This is in line with our data suggesting that axonal swellings precede axon fragmentation. To what extent our model of hemin-induced AxD in hemorrhagic stroke

290 is molecularly similar to developmental or pathophysiological AxD needs to be further
291 investigated.
292 To compare different experimental conditions of AxD, we herein propose a novel monolithic
293 microfluidic device consisting of 16 individual microfluidic units that enable the parallel and
294 separated treatment and/or manipulation of axons and somata (**Fig. 1**). The currently available
295 microfluidic devices do not allow high-throughput experiments as they comprise only single
296 microfluidic units[20,38]. Although some devices can harbor multiple experimental conditions,
297 they employ a radial design where a single soma compartment is used, in which influence from
298 one experimental condition to another experimental condition cannot be excluded due to the
299 potential of retrograde signaling[39,40]. Another option is the parallel use of multiple individual
300 devices, which allows to handle up to 12 devices in a conventional 12-well plate[18]. Compared
301 to our microfluidic device, this procedure is time-consuming in both the manufacturing process
302 and the adjustment for recordings.
303
304 **Limitations and outlook**
305 i) Our microfluidic device currently does not allow to investigate AxD at more proximal axonal
306 parts to the soma such as the axonal initial segment. Since intracerebral hemorrhage often
307 occurs in deep brain regions, affecting axons of near-by neurons as well as more distal axonal
308 trajectories, comparing AxD in distal and proximal axonal regions may be of interest.
309 Shortening the length of the microgrooves or including a more proximal compartment, are
310 possible modifications of the current design.
311 ii) Our results are based on unmyelinated axons. Therefore, the time course and morphological
312 changes may be different in co-culture with oligodendrocytes. This warrants possible future
313 investigations.
314 iii) The observed effects of AxD in hemorrhagic stroke within this study were based on hemin
315 toxicity, and we cannot exclude that other hemolysis products such as thrombin or bilirubin

have different effects. Additional studies should investigate differences of hemolysis products to increase our understanding of the mechanisms of AxD in hemorrhagic stroke.

iv) The overall CNN performance may be further improved with more general inputs. For example, the segmentation of fragment pixels cannot be done accurately from a single image at a specific time point as the whole process of AxD, ultimately resulting in the disintegration of the axons (i.e. the generation of axonal fragments), needs to be considered. CNNs using 3D convolutions could, in principle, perform a segmentation over an entire time-lapse recording and model temporal dependencies. However, we decided against the 3D approach, as it severely restricts general applicability due to its greatly increased effort to label suitable time series for training. In this context, the identification of the images that will yield the best results is crucial to effectively reduce labeling costs, which we have previously described using an active learning method[41].

Conclusion

Taken together, our approach enables the systematic investigation of AxD in a high-throughput manner. In combination with the microfluidic device, the EntireAxon expands our possibilities to track AxD by detecting axons, axonal swellings and axonal fragments and identifying the related different morphological patterns of AxD. This will help to tackle the complex processes of AxD and may significantly enhance our understanding of AxD in health and disease. It may also support the development of novel treatment approaches for neurodegenerative diseases.

Online methods

Chemicals and reagents are listed in the **supplementary information (suppl. Tables 4-5)**.

Fabrication of a high-throughput microfluidic device based on soft lithographic replica molding

Thirty-two wells were milled in a polymethyl methacrylate (PMMA) plate of the size of a conventional cell culture plate (**Fig. 1, suppl. Fig. 1**) using a universal milling machine (Mikron WF21C, Mikron Holding AG) with a 1 mm triple tooth cutter (HSS-CO8 Type N, Holex) at a precision of 0.01 mm. During the milling procedure, we applied a half-synthetic cooling lubricant (Opta Cool 600 HS, Wisura GmbH) on a mineral base to reduce the amount of debris. Additionally, we milled screw holes in the intermediate spaces between each microfluidic unit to later detach the PMMA from the negative casting mold. To remove debris, we washed the PMMA plate by sonication (Sonicator Elmasonic S, Elma Schmidbauer GmbH) at room temperature for 30 minutes. Next, we lasered the microgrooves on the PMMA plate to connect both milled compartments of each individual microfluidic unit by using an Excimerlaser (Excistar XS 193 nm, Coherent). The PMMA plate was then washed again by sonication at room temperature for 30 minutes.

Polydimethylsiloxane (PDMS) was prepared in a 1:10 ratio and mixed properly before inducing vacuum at 0.5 Torr in a vacuum desiccator (Jeio Tech VDC-31) for 30 minutes. After the PDMS was poured into an empty aluminum basin to cover the ground, we applied vacuum at 0.5 Torr for 30 minutes to remove air bubbles. The PDMS was cured at room temperature for 48 hours. We put the PMMA plate on top of the PDMS ground with the milled and lasered structures showing upwards. Half of each well of the microfluidic units was filled with PDMS before curing at room temperature for 48 hours. We mixed the epoxy solution in a 1:1 ratio and poured it over the microfluidic device to cover its surface by at least 1 cm. Vacuum was applied at 0.5

Torr for 10 minutes to remove all air bubbles located above the channel side of the microfluidic device. The epoxy was cured at room temperature for a minimum of 2 hours. We subsequently detached the epoxy from the PMMA plate via a metallic block that consisted of screw holes in the intermediate spaces between the individual systems. The epoxy represented a negative casting mold to produce the microfluidic devices using PDMS.

PDMS was prepared as described above. We poured the PDMS into the negative epoxy casting mold and applied vacuum at 0.5 Torr for 30 minutes. The liquid PDMS was cured at 75 °C for 2 hours to induce the polymerization. We peeled the microfluidic devices from the casting mold and punched the wells with an 8 mm biopsy punch (DocCheck Shop GmbH) to ensure a sufficient amount of medium for cell culture. We cleaned customized 115 x 78 x 1 mm glass slides by sonication (Sonicator Elmasonic S, Elma Schmidbauer GmbH) and subsequently cleaned them by ethanol before plasma treatment (High Power Expanded Plasma Cleaner, Harrick Plasma). Plasma was applied at 45 W and 0.5 Torr for 2 minutes to activate the silanol groups of the glass slides and the microfluidic devices enabling firm attachment.

We washed the microfluidic devices with ethanol and then twice with distilled water to remove any debris. After aspirating the distilled water, except from the inside of the compartments, 0.1 mg/mL of poly-d-lysine solution (PDL) in 0.02 M borate buffer (0.25 % (w/v) borate acid, 0.38 % (w/v) sodium tetraborate in distilled water, pH 8.5) was used for coating at 4 °C overnight. We aspirated the PDL the next morning, not removing it from the compartments, and added 50 µg/mL of laminin as a second coating surface for incubation at 4 °C overnight. At the day of neuron isolation, the microfluidic devices were washed twice with pre-warmed medium after aspirating the laminin. Immediately prior to cell seeding, we aspirated the medium from the wells without removing it from the compartments.

Experimental animals

Crl:CD1 (ICR) Swiss outbred mice (Charles River) were used. The animals were kept at 20-22 °C, 30-70 % humidity in a 12-hour/12-hour light/dark cycle and were fed a normal chow diet (Altromin Spezialfutter GmbH) *ad libitum*. Animal experiments followed the protocol of the "NIH Guide for the care and use of laboratory animals" and were approved by the Schleswig-Holstein Ministry for Energy Transition, Agriculture, Environment, Nature and Digitalization (under the prospective contingent animal license number 2017-07-06 Zille).

Isolation and culture of primary cortical neurons

We isolated primary cortical neurons (PCNs) from murine E14 embryos after decapitation as previously described[28]. We seeded the PCNs at a density of 10,000 cells/mm^2 in 5 µL MEM+Glutamax medium into one compartment (soma compartment) of each microfluidic unit of the device. The cells adhered at 37 °C for 30 minutes. In order to promote the growth of the axons in one direction into the other compartment (axonal compartment) by medium microflux, 150 µL of MEM+Glutamax medium were applied to the well of the soma compartment, while 100 µL were added to the well of the axonal compartment. PCNs were cultured at 37 °C in a humidified 5 % CO_2 atmosphere. We changed from MEM+Glutamax medium to Neurobasal Plus Medium containing 2 % B-27 Plus Supplement, 1 mM sodium pyruvate and 1 % penicillin/streptomycin at day-in vitro (DIV) 1. The volume differences among the wells ensured the microflux for the directional axonal growth over the following days.

Immunofluorescence

Soma and axonal compartments in the microfluidic units were fixed at room temperature for 1 hour in 4 % formaldehyde solution in PBS. They were washed twice with PBS and permeabilized with blocking solution (2 % BSA, 0.5% Triton-X-100 and 1x PBS) at room temperature for 1 hour. We incubated the neurons/axons on both compartments with primary antibodies against synaptophysin (1:250) and MAP2 (1:4000) at 4 °C overnight. The next day,

both compartments were washed three times with PBS and incubated with the secondary antibodies anti-mouse Alexa Fluor 546 (1:500) and anti-rabbit Alexa Fluor 488 (1:500) at room temperature for 1 hour. After washing three times with PBS, both compartments were incubated with DAPI (1 µg/mL) for nuclear counterstaining at room temperature for 10 minutes. Both compartments were finally washed three times with PBS prior to fluorescence microscopy.

Selection of microfluidic units for hemin treatment and time-lapse recording

At DIV 6 or 7, microfluidic devices were considered for recording if they met the following inclusion criteria: i) axon growth through at least 80 % of all microgrooves and ii) axon length of at least 150 µm from the end of the microgrooves. Microfluidic units were randomly assigned to experimental conditions.

Time-lapse recording of axonal degeneration

Microfluidic units were randomly selected and axons were treated with 0 (vehicle), 50, 100, and 200 µM hemin. For the treatment, the media was removed from the wells of the microfluidic units; hemin was diluted in the collected media and added back to the respective wells. The media volume between the two wells was equalized during the treatment to prevent any microflux. We started the recordings at 1 hour after treatment to allow for the adjustment of the well plates to the humidity of the incubation chamber of the microscope and the setup of the recording positions. We recorded AxD with a 30-minutes interval for 24 hours using an Olympus IX81 time-lapse microscope (Olympus Deutschland GmbH) with a 10X objective (0.3 NA Ph1) at 37 °C, 5 % CO_2 and 65 % humidity.

Live cell fluorescent staining

To evaluate axonal vitality, we washed the axonal compartment once with PBS and incubated the axonal compartment with calcein AM (4 µM) for 30 minutes at 37 °C at the end of the time-lapse recording or in 4-hour intervals upon hemin treatment.

Training of the EntireAxon CNN for the segmentation of phase-contrast microscopic images

We trained the EntireAxon CNN for the image-wise semantic segmentation of AxD features. To this end, we adapted a standard u-net with ResNet-50 encoder[23] to automatically determine the class probability for each pixel of an input image. Our segmentation aimed to classify each pixel of a microscopic image of a time-lapse recording into one of four classes: 'background', 'axon', 'axonal swelling', and 'axonal fragment'.

For the training dataset, we selected 33 images and created corresponding image labels (masks) using GIMP (v.2.10.14, RRID:SCR_003182). For each image, a label image with the same height and width was created, in which each pixel value denotes a pixel class. Specifically, the classes 'background', 'axon', 'axonal swelling', and 'axonal fragment' had the values 0, 1, 2, and 3, respectively. For each pixel of the input image, we retained 4 values that are the probability distribution of the pixel over the 4 classes. To create a segmentation map, we assigned each pixel the most probable class. During training, the CNN observed an input image, produced an output and compared this output to the label. Via backpropagation, the weights of the network were adapted so that the output better fitted the label. The pixelwise loss function, from which the weight changes were derived, was the cross-entropy loss:

$$Loss(P,Y) = -\sum_{x,y,c} Y(x,y,c)\, log(P(x,y,c));$$

With $P(x,y,c)$ and $Y(x,y,c)$ being the probability of class c at pixel (x,y) for the prediction and ground truth of the network, respectively.

We trained a mean ensemble consisting of eight neural networks for 180 epochs using the Adam optimizer, a batch size of four and a learning rate of 0.001 that decreased by a factor of ten after every 60 epochs. The input images were standardized by the image-net mean and standard deviation [42]. For data augmentation, we used random cropping (size 512 x 512), image flipping along the horizontal axis and rotation by a random angle between -90° and +90°.

Validation of the EntireAxon CNN compared to human experts

To measure how well the EntireAxon CNN segments unknown images, we used a second validation set comprising eight images that were labeled by three human experts. Importantly, the EntireAxon CNN did not update its parameters during training to fit the validation set, but only used the training set.

For each image, the EntireAxon CNN inferred a segmentation. We generated a binary mask from the prediction of the network, where 1 denotes the respective class and 0 all other classes. We computed a binary label mask in the same manner. We counted the true positive (TP), false positive (FP), and false negative (FN) pixels and computed the recall (sensitivity) and precision[43]:

$$Recall = \frac{TP}{TP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

Recall and precision were calculated for each class separately on each validation image. The mean recall and precision over all eight validation images was determined subsequently.

A mean of 96.42 % of pixels in the axonal images were ‘background’ pixels, while only 2.77 % represented the class ‘axon’, 0.58 % ‘axonal swelling’, and 0.23 % ‘axonal fragment’ pixels. This reflects a challenging degree of class imbalance, where the probability of having any positives for a class in a validation image is low. Thus, we did not use the computed recall and precision of the individual images or the mean recall and precision to compute the mean F1

score, i.e. the harmonic mean of recall and precision. This has been shown to lead to bias, especially when a high degree of class imbalance is present in the dataset[43] as it may result in undefined values for an image for recall (due to the absence of TP), precision (in case the CNN does not recognize the few positives), and F1 score (in case either recall or precision are undefined). To avoid bias, we computed the total TP, FP, and FN of all validation images from which we calculated the mean F1 score[43]:

$$mean\ F1\ score = \frac{2\ *\ TP_{total}}{2 * TP_{total}\ +\ FP_{total}\ +\ FN_{total}}$$

In addition, we computed a consensus label between human expert 1 and 2, 1 and 3 as well as 2 and 3 and compared the EntireAxon CNN versus the remaining expert (human expert 3, 2, and 1, respectively) to the consensus labels. Mean F1 scores for all classes were computed as described above.

Image preprocessing

Prior to the analysis of AxD after hemin exposure, we preprocessed the time-lapse recordings in ImageJ (v1.52a, RRID: RRID:SCR_003070) using a custom-written macro. Specifically, each individual recording was converted from a 16-bit into an 8-bit recording to make it compatible with the ImageNet (8-bit) pre-trained ResNet-50. The recording was aligned automatically with the ImageJ plug-in "Linear Stack Alignment with SIFT" as described previously[44]. The following settings were used: initial Gaussian blur of 1.6 pixel, 3 steps per scale octave, minimum image size of 64 pixel, maximum image size of 1024 pixel, feature descriptor size of 4, 8 feature descriptor orientation bins, closest/next closest ratio of 0.92, maximal alignment error of 25 pixel; inlier ratio of 0.05, expected transformation as rigid, "interpolate" and "show info" checked. Black edges appearing on the recording after alignment were cropped.

513 **AxD analysis using the EntireAxon CNN**

514 All recordings of AxD after hemin exposure were automatically analyzed by the trained
515 EntireAxon CNN, which classified each pixel as one of the four different classes 'background',
516 'axon', 'axonal swelling', and 'axonal fragment. For each experimental condition (i.e. hemin
517 concentration), the sum percentage of all pixels per class on all images of that experimental day
518 were added at each time point ('$Axon_{t1.5–24h}$, Axonal $swelling_{t1.5-24h}$, Axonal $fragment_{t1.5-24h}$). To
519 determine the changes for the classes 'axon', 'axonal swelling', and 'axonal fragment' over
520 time, we calculated the sum percentage of pixels for all given time points (t_i with i = 1.5 to
521 24 hours) of the corresponding class over the sum of the pixels of all three classes at baseline:

522 $$normalized\ 'class'area\ (t_i)$$

523 $$= \frac{'Class'_{ti}}{Axon_{t1.5h} + Axonalswelling_{t1.5h} + Axonalfragments_{t1.5h}} * 100$$

524

525 **Classification of the morphological patterns of AxD using attention-based RNN**

526 We used the segmentation videos to identify four morphological patterns of AxD: i) granular,
527 ii) retraction, iii) swelling, and iv) transport degeneration. To reduce the dimensions of the
528 input, each frame was converted into a histogram. To compute a histogram for a frame t_i, we
529 compared the pixels of the frames t_i and t_{i+1}. Each pixel was assigned into one of 16 classes that
530 consisted of the pairwise tuples $(c_1, c_2) \in \{0,1,2,3\}^2$ of the four segmentation classes. For
531 example, the class (background, axon) means that in frame t_i, the pixel was classified as
532 background while in frame t_{i+1}, it was an axon pixel. For T time steps, we therefore computed
533 T-1 histograms. Additionally, we normalized each histogram to sum up to 1:

534 $$H(t_i, (c_1, c_2)) := H_0(t_i, (c_1, c_2)) \,/\, \sum_{a,b} H_0(t_i, (a, b))$$

535 Note that, the histograms were computed over small patches (height and width < 90 pixels)
536 during training and during inference on windows of size 32x32 pixels.

We used an encoder-decoder RNN with attention[45]. The encoder f_{enc} consisted of a gated recurrent unit (GRU) that obtained the histogram time sequence H as input. The encoder computed the hidden representation of the histograms:

$$V = f_{enc}(H);\ V \in \mathbb{R}^{T\ x\ d}, H \in \mathbb{R}^{T\ x\ 16}$$

For our purpose, we used an architecture that was able to base the decision for a degeneration class on the previous class predictions. To this end, the output $\vec{o}_i$ was computed iteratively in C+1 steps as a sum of the previous output and the output of the decoder f_{dec}:

$$\vec{o}_i = \vec{o}_{i-1} + f_{dec}(\sigma(\vec{o}_{i-1}), \vec{s}_{i-1}); \vec{o} \in \mathbb{R}^C, \vec{s}_{i-1} \in \mathbb{R}^d$$

$$f_{dec}(\sigma(\vec{o}_{i-1}), \vec{s}_{i-1}) = W_{out}\vec{z}_i;\ W_{out} \in \mathbb{R}^{C\ x\ d}$$

C is the number of degeneration classes (4) and d is the hidden dimension (we used 256); $i = 1, \ldots, C+1$. σ is the sigmoid function. The decoder employed a GRU that depended on the context vector $\vec{c}_i$ and the hidden state vector $\vec{s}_{i-1}$:

$$\vec{z}_i, \vec{s}_i = GRU(\vec{c}_i, \vec{s}_{i-1});\ \vec{z}_i \in \mathbb{R}^d$$

The context vector is a weighted sum of the encoder representations. At each iteration, these weights can change, enabling the network to focus on different time-steps. We assumed that a specific pattern of degeneration happened only in a limited number of time frames that were smaller than the whole input video. The weights depended on the current state of the decoder and the current output:

$$\vec{c}_i = V^T\vec{\alpha}_i;\ \vec{\alpha}_i \in \mathbb{R}^T$$

$$\vec{\alpha}_i = Softmax(\ W_{att}\ [\vec{s}_{i-1}, ReLU(W_{in}\vec{o}_{i-1})]\ );\ W_{att} \in \mathbb{R}^{T\ x\ 2d},\ W_{in} \in \mathbb{R}^{d\ x\ C}$$

Here, $[\vec{a}, \vec{b}]$ is the concatenation of two vectors. The final output y is normalized by the sigmoid function:

$$y = \sigma(\vec{o}_{C+1}) \in [0,1]^C.$$

Apart from the weights used by the GRUs, W_{in}, W_{att}, and W_{out} are learnable weights.

The EntireAxon RNN was trained with 162 images for 60 epochs using the lamb optimizer[46] with a batch size of 128. We used a learning rate of 0.01 that was reduced by a factor of ten every 15 epochs and an additional weight decay of 0.0001. The two GRUs (encoder and decoder) contained three layers, and we used dropout with a p-value of 0.9. To increase the RNN robustness against varying axon thickness, we also added eroded versions of the segmentation data using a cross-shape as kernel with the sizes three, five, and seven. Accordingly, each image existed six times in the dataset: three eroded versions and three unchanged copies, to keep a 50 % chance of having the original image for training.

Ten-fold cross-validation of the RNN

To validate the RNN, we used ten-fold cross-validation[47]. For the given 162 training images, we determined ten separate test sets (nine sets containing 17 images and one set including nine images) and, each time, trained with the remaining 145 (153) images as described above. Mean recall, precision, and F1 score over the ten sets were determined as described above.

Analysis of morphological pattern of AxD using the EntireAxon RNN

All segmentations of AxD after hemin exposure were automatically analyzed with the trained EntireAxon RNN, which predicted the occurrence of the four morphological patterns of AxD in a pixel-wise manner. To note: A pixel can be predicted to belong to 0, 1 or multiple morphological patterns. Only pixels previously identified as degenerated over time were considered by applying a ‘fragmentation mask’ that included all no-background pixels that changed to either background or fragment during the recording time.

For each experimental condition (i.e. hemin concentration), the percentage of the occurrence of each morphological pattern was calculated as the sum of all pixels per morphological pattern on all images of that experimental day divided by the ‘fragmentation mask’ as follows:

$$'morphological\ pattern'[\%] = \frac{\sum pixel\ of\ morphological\ pattern_i}{\sum pixel\ no\ background \rightarrow background\ or\ fragment} * 100$$

Statistical analysis

Normality was evaluated with the Kolmogorov-Smirnov test, variance homogeneity using the Levené test, and sphericity by the Mauchly test. When the data were normally distributed and variance homogeneity was met, one-way ANOVA followed by the Bonferroni post hoc test was performed. In case the data were not normally distributed, the Kruskal-Wallis test was performed for multiple comparisons of independent groups followed by the post hoc Mann-Whitney U test with α-correction according to Bonferroni to adjust for the inflation of type I error due to multiple testing. For the repeated testing with covariates, a repeated measures ANOVA was performed with Greenhouse-Geisser adjustment if sphericity was not given. Data are represented as mean ± standard deviation (SD) except for the nonparametric data of the AUC for axonal fragments as well as retraction and swelling degeneration, where medians are given. A value of $p<0.05$ was considered statistically significant. For the Kruskal-Wallis test followed by Mann-Whitney U, $p=0.05/k$ was used, with k as the number of single hypotheses. K=3 for AUC analyses (comparison of three different concentrations of hemin vs. 0 µM hemin), thus $\alpha=0.0167$ was considered statistically significant. K=6 for swelling degeneration (comparison of four different concentrations of hemin against each other), thus $\alpha=0.0083$ was considered statistically significant. The detailed statistical analyses can be found in the **suppl. Tables 1-3**. All statistical analyses were performed with IBM SPSS version 23 (RRID:SCR_002865).

Acknowledgments

This work was supported by a Fraunhofer MEF grant (Project number 600199) to M.Z.

Author contributions

M.Z. designed the experiments. A. Palumbo, A. Pabst and M.Z. designed the device. A. Palumbo, A. Pabst, S.P., R.S., C.K., and N.K. carried out the fabrication of the device. S.K.L. performed the immunostaining of the somata and axons and analyzed the respective data. P.G. and A.M.M. developed the deep learning tool, P.G., L.H., and L.B. developed the algorithms to retrieve the output. A. Palumbo, S.K.L. and L.H. labeled the images for the deep learning training and validation. A. Palumbo and C.F. performed the time-lapse recordings of AxD. A. Palumbo conducted the live cell imaging, the determination of the morphological patterns of AxD and analyzed the data for the respective experiments. A. Palumbo, P.G., A.M.M., J.B., and M.Z. discussed and interpreted the data. M.Z. performed the statistical analysis. A. Palumbo, M.I. and M.Z. performed the graphical artwork. A. Palumbo, P.G., and M.Z wrote the manuscript. All authors discussed and commented on the final version of the manuscript.

Competing interest statement

A. Palumbo, P.G., and M.Z. declare that they have filed a patent for the microfluidic device and the EntireAxon deep learning algorithm to quantify axonal degeneration (European Patent Office, file number: 20152016.0, in revision). All other authors declare no conflict of interest.

Data availability

The data and code are available upon reasonable request to the corresponding authors.

References

1. Luo, L. & O'Leary, D. D. M. Axon Retraction and Degeneration in Development and Disease. *Annu. Rev. Neurosci.* **28**, 127–156 (2005).
2. Lingor, P., Koch, J. C., Tönges, L. & Bähr, M. Axonal degeneration as a therapeutic target in the CNS. *Cell Tissue Res* **349**, 289–311 (2012).
3. Salvadores, N., Sanhueza, M., Manque, P. & Court, F. A. Axonal Degeneration during Aging and Its Functional Role in Neurodegenerative Disorders. *Front. Neurosci.* **11**, 451 (2017).
4. Pease, S. E. & Segal, R. A. Preserve and protect: maintaining axons within functional circuits. *Trends in Neurosciences* **37**, 572–582 (2014).
5. Neukomm, L. J. & Freeman, M. R. Diverse cellular and molecular modes of axon degeneration. *Trends in Cell Biology* **24**, 515–523 (2014).
6. Bishop, D. L., Misgeld, T., Walsh, M. K., Gan, W.-B. & Lichtman, J. W. Axon Branch Removal at Developing Synapses by Axosome Shedding. *Neuron* **44**, 651–661 (2004).
7. Cavanagh, J. B. The 'dying back' process. A common denominator in many naturally occurring and toxic neuropathies. *Arch. Pathol. Lab. Med.* **103**, 659–664 (1979).
8. Coleman, M. P. Axon degeneration mechanisms: commonality amid diversity. *Nat Rev Neurosci* **6**, 889–898 (2005).
9. Beirowski, B. *et al.* The progressive nature of Wallerian degeneration in wild-type and slow Wallerian degeneration (WldS) nerves. *BMC Neurosci* **6**, 6 (2005).
10. Nikić, I. *et al.* A reversible form of axon damage in experimental autoimmune encephalomyelitis and multiple sclerosis. *Nat Med* **17**, 495–499 (2011).
11. Yong, Y. *et al.* p75NTR and DR6 Regulate Distinct Phases of Axon Degeneration Demarcated by Spheroid Rupture. *J. Neurosci.* **39**, 9503–9520 (2019).
12. Saxena, S. & Caroni, P. Mechanisms of axon degeneration: From development to disease. *Progress in Neurobiology* **83**, 174–191 (2007).
13. Wang, J. T., Medress, Z. A. & Barres, B. A. Axon degeneration: Molecular mechanisms of a self-destruction pathway. *J Cell Biol* **196**, 7–18 (2012).
14. Sasaki, Y., Vohra, B. P. S., Lund, F. E. & Milbrandt, J. Nicotinamide Mononucleotide Adenylyl Transferase-Mediated Axonal Protection Requires Enzymatic Activity But Not Increased Levels of Neuronal Nicotinamide Adenine Dinucleotide. *Journal of Neuroscience* **29**, 5525–5535 (2009).
15. Becker, T. & Madany, A. Morphology-based Features for Adaptive Mitosis Detection of In Vitro Stem Cell Tracking Data. *Methods Inf Med* **51**, 449–456 (2012).
16. Pool, M., Thiemann, J., Bar-Or, A. & Fournier, A. E. NeuriteTracer: A novel ImageJ plugin for automated quantification of neurite outgrowth. *Journal of Neuroscience Methods* **168**, 134–139 (2008).
17. Ho, S.-Y. *et al.* NeurphologyJ: An automatic neuronal morphology quantification method and its application in pharmacological discovery. *BMC Bioinformatics* **12**, 230 (2011).
18. Li, Y. *et al.* AxonQuant: A Microfluidic Chamber Culture-Coupled Algorithm That Allows High-Throughput Quantification of Axonal Damage. *Neurosignals* **22**, 14–29 (2014).
19. Campenot, R. B. Development of sympathetic neurons in compartmentalized cultures. *Developmental Biology* **93**, 13–21 (1982).
20. Park, J. W., Vahidi, B., Taylor, A. M., Rhee, S. W. & Jeon, N. L. Microfluidic culture platform for neuroscience research. *Nat Protoc* **1**, 2128–2136 (2006).
21. Dehmelt, L. & Halpain, S. The MAP2/Tau family of microtubule-associated proteins. *Genome Biol* **6**, 204 (2004).
22. Wiedenmann, B. & Franke, W. W. Identification and localization of synaptophysin, an

integral membrane glycoprotein of Mr 38,000 characteristic of presynaptic vesicles. *Cell* **41**, 1017–1028 (1985).
23. Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Preprint at https://arxiv.org/abs/1505.04597 (2015).
24. He, K., Zhang, X., Ren, S. & Sun, J. Deep Residual Learning for Image Recognition. Preprint at https://arxiv.org/abs/1512.03385 (2015).
25. Dietterich, T. G. Ensemble Methods in Machine Learning. in *Multiple Classifier Systems* vol. 1857 1–15 (Springer Berlin Heidelberg, 2000).
26. Huang, H.-K. *et al.* Mixture of deep CNN-based ensemble model for image retrieval. in *2016 IEEE 5th Global Conference on Consumer Electronics* 1–2 (IEEE, 2016). doi:10.1109/GCCE.2016.7800375.
27. Vuola, A. O., Akram, S. U. & Kannala, J. Mask-RCNN and U-Net Ensembled for Nuclei Segmentation. in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)* 208–212 (IEEE, 2019). doi:10.1109/ISBI.2019.8759574.
28. Zille, M. *et al.* Neuronal Death After Hemorrhagic Stroke In Vitro and In Vivo Shares Features of Ferroptosis and Necroptosis. *Stroke* **48**, 1033–1043 (2017).
29. Cui, Y. *et al.* Axonal degeneration in an in vitro model of ischemic white matter injury. *Neurobiology of Disease* **134**, 104672 (2020).
30. Venkatasubramanian, C. *et al.* Natural History and Prognostic Value of Corticospinal Tract Wallerian Degeneration in Intracerebral Hemorrhage. *JAHA* **2**, (2013).
31. Tao, C., Hu, X., Li, H. & You, C. White Matter Injury after Intracerebral Hemorrhage: Pathophysiology and Therapeutic Strategies. *Front. Hum. Neurosci.* **11**, 422 (2017).
32. Chen, X. *et al.* The Impact of Intracerebral Hemorrhage on the Progression of White Matter Hyperintensity. *Front. Hum. Neurosci.* **12**, 471 (2018).
33. Hemorrhagic Stroke Academia Industry (HEADS) Roundtable Participants. Basic and Translational Research in Intracerebral Hemorrhage: Limitations, Priorities, and Recommendations. *Stroke* **49**, 1308–1314 (2018).
34. Nikolaev, A., McLaughlin, T., O'Leary, D. D. M. & Tessier-Lavigne, M. APP binds DR6 to trigger axon pruning and neuron death via distinct caspases. *Nature* **457**, 981–989 (2009).
35. Maor-Nof, M. *et al.* Axonal Degeneration Is Regulated by a Transcriptional Program that Coordinates Expression of Pro- and Anti-degenerative Factors. *Neuron* **92**, 991–1006 (2016).
36. Pease-Raissi, S. E. *et al.* Paclitaxel Reduces Axonal Bclw to Initiate IP3R1-Dependent Axon Degeneration. *Neuron* **96**, 373-386.e6 (2017).
37. Tang-Schomer, M. D., Johnson, V. E., Baas, P. W., Stewart, W. & Smith, D. H. Partial interruption of axonal transport due to microtubule breakage accounts for the formation of periodic varicosities after traumatic axonal injury. *Experimental Neurology* **233**, 364–372 (2012).
38. Van Laar, V., Arnold, B. & Berman, S. Primary Embryonic Rat Cortical Neuronal Culture and Chronic Rotenone Treatment in Microfluidic Culture Devices. *BIO-PROTOCOL* **9**, (2019).
39. Hosmane, S., Yang, I. H., Ruffin, A., Thakor, N. & Venkatesan, A. Circular compartmentalized microfluidic platform: Study of axon–glia interactions. *Lab Chip* **10**, 741 (2010).
40. Biffi, E. *Microfluidic and Compartmentalized Platforms for Neurobiological Research*. (2015).
41. Grüning, P., Palumbo, A., Zille, M., Barth, E. & Madany Mamlouk, Amir. A task-dependent active learning method for axon segmentation with CNNs. *Proc AUTOMED* **1**, (2020).
42. Deng, J. *et al.* ImageNet: A large-scale hierarchical image database. in *2009 IEEE*

Conference on Computer Vision and Pattern Recognition 248–255 (IEEE, 2009). doi:10.1109/CVPR.2009.5206848.

43. Forman, G. & Scholz, M. Apples-to-apples in cross-validation studies: pitfalls in classifier performance measurement. *SIGKDD Explor. Newsl.* **12**, 49 (2010).

44. Lowe, D. G. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision* **60**, 91–110 (2004).

45. Bahdanau, D., Cho, K. & Bengio, Y. Neural Machine Translation by Jointly Learning to Align and Translate. Preprint at https://arxiv.org/abs/1409.0473 (2016).

46. You, Y. *et al.* Large Batch Optimization for Deep Learning: Training BERT in 76 minutes. Preprint at https://arxiv.org/abs/1904.00962 (2020).

47. Hastie, T., Tibshirani, R. & Friedman, J. *The Elements of Statistical Learning*. (Springer New York, 2009). doi:10.1007/978-0-387-84858-7.

Figures

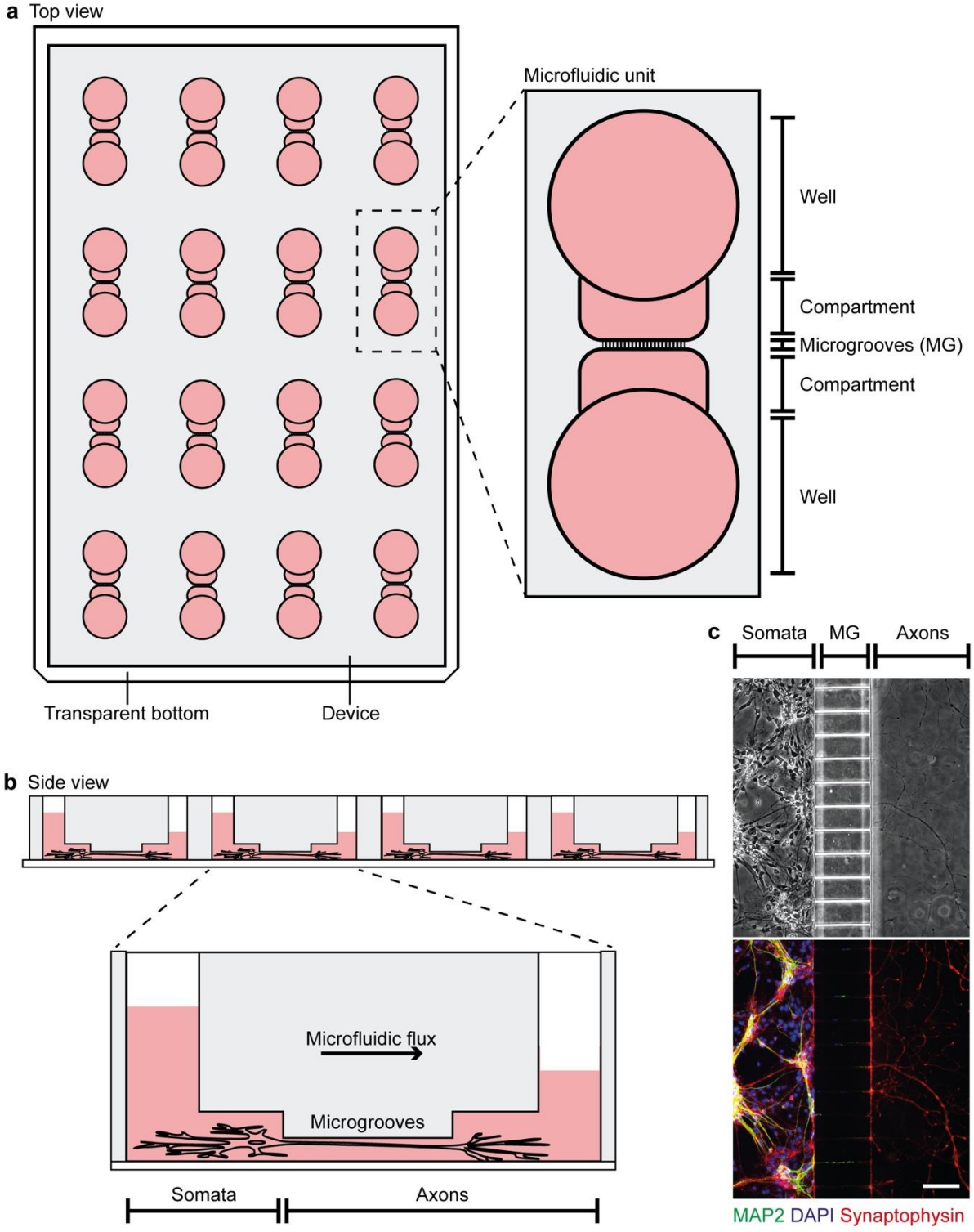


Fig. 1 Microfluidic device for the high-throughput cultivation of axons. a, The microfluidic device incorporates 16 individual microfluidic units for axon cultivation. One microfluidic unit consists of two wells that are connected through compartments and microgrooves (MG). **b**, Primary cortical neurons are seeded into the soma compartment from which their axons grow through the MG into the axon compartment. Directed growth is supported by culture medium microflux due to different medium volumes between the two wells. **c**, Phase-contrast image of axons that were spatially separated from their somata by the MG at *day in vitro* 7, which we confirmed by immunofluorescence staining of dendrites using microtubule-associated protein 2 (MAP2, green) and axons using synaptophysin (red). DAPI (blue) was used for nuclear counterstaining (top). Scalebar, 100 µm.

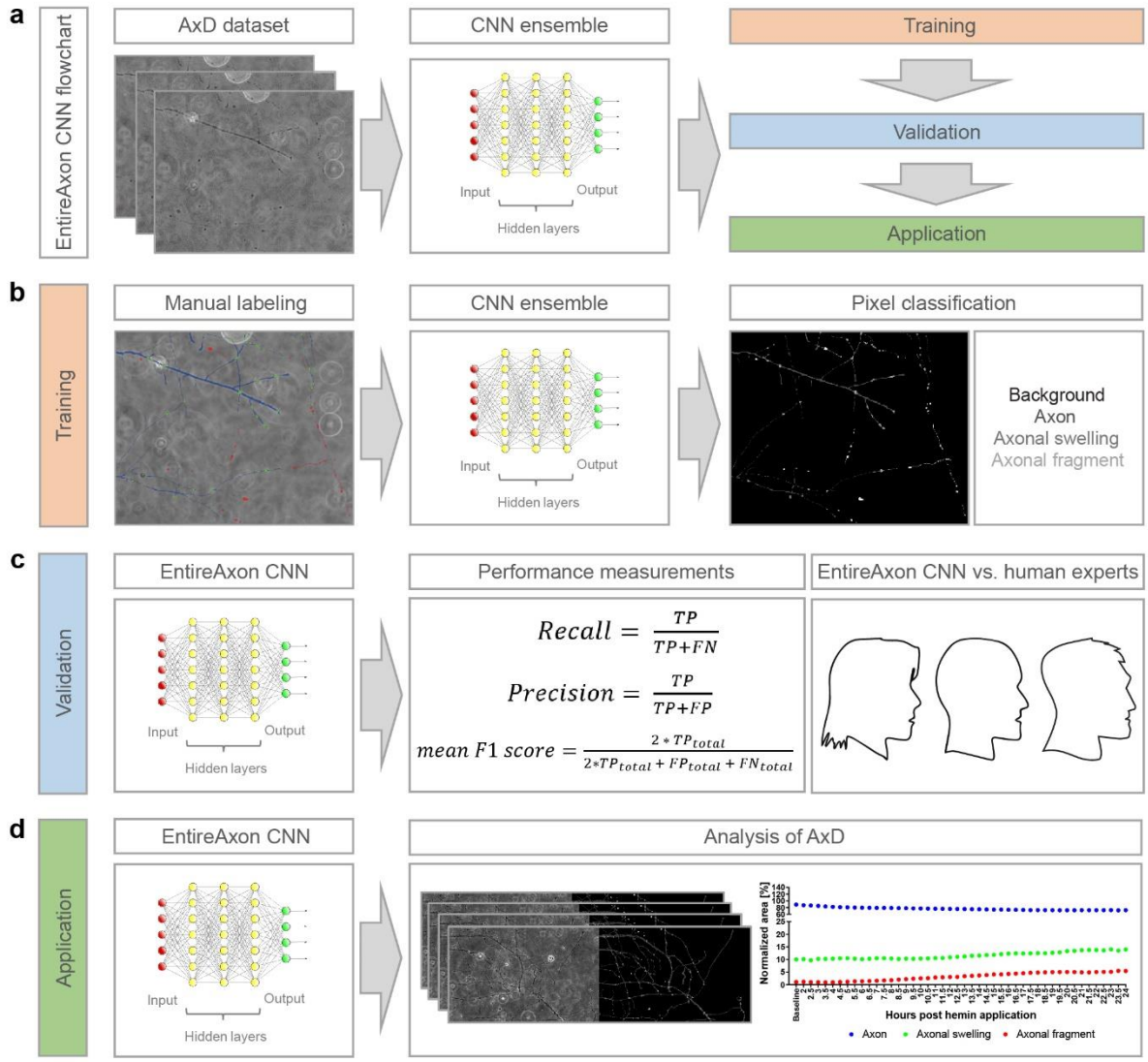


Fig. 2 EntireAxon convolutional neural network (CNN) for the high-throughput analysis of axonal degeneration (AxD). **a,** The flow chart of the EntireAxon CNN. The AxD data was separated into training, validation, and testing data. **b,** We manually labeled the training data to segment each pixel into the four classes 'background', 'axon', 'axonal swelling', and 'axonal fragment', which are displayed in the output image in black, dark grey, intermediate grey, and light grey, respectively. We trained an ensemble comprising 8 CNNs to segment the four classes. **c,** The EntireAxon CNN was validated with a separate validation dataset to assess its performance (recall, precision, and mean F1 score), which was compared to human experts. **d,** The EntireAxon CNN was applied to data on AxD induced by the exposure of hemin, which is used to model of hemorrhagic stroke *in vitro*.

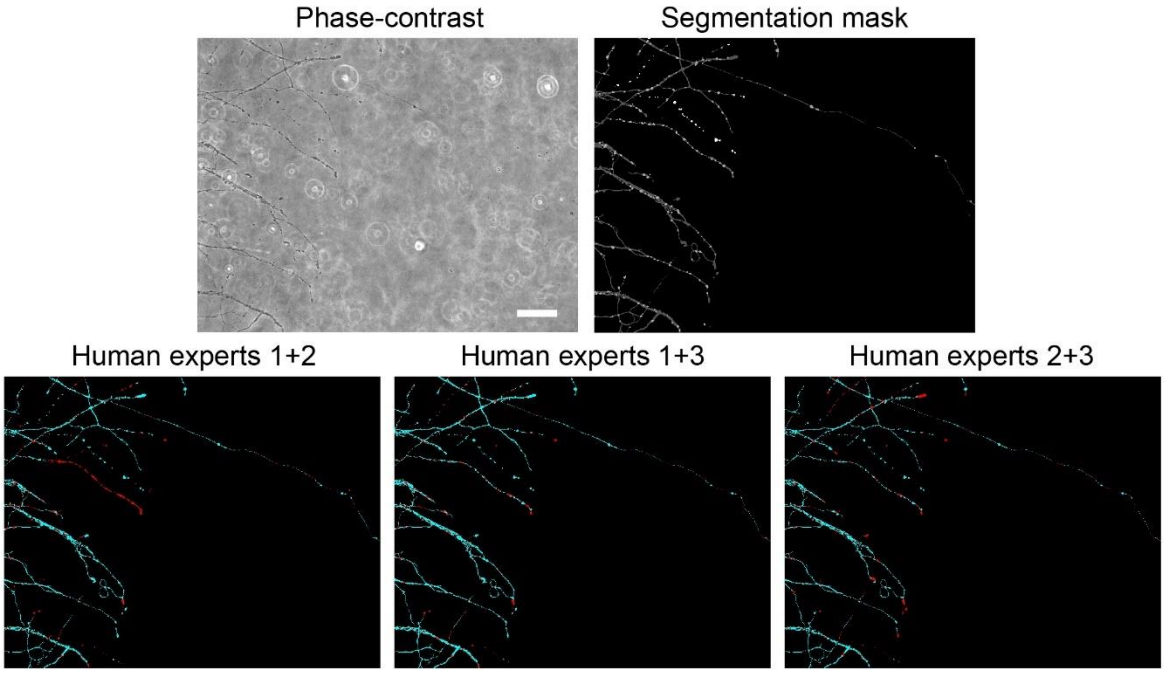
a

Class	Precision	Recall	Mean F1 score
Background	0.993	0.996	0.995
Axon	0.789	0.774	0.780
Axon swelling	0.609	0.534	0.567
Axonal fragment	0.805	0.196	0.301

b

Class	Mean F1 score Background	Mean F1 score Axon	Mean F1 score Axonal swelling	Mean F1 score Axonal fragments
EntireAxon CNN	0.995	0.780	0.567	0.301
Human expert 2	0.991	0.654	0.485	0.548
Human expert 3	0.993	0.704	0.489	0.221

c



d

Consensus	Class	Mean F1 score Background	Mean F1 score Axon	Mean F1 score Axonal swelling	Mean F1 score Axonal fragments
Human expert 1+2	EntireAxon CNN	0.998	0.847	0.667	0.400
	Human expert 3	0.998	0.808	0.647	0.376
Human expert 1+3	EntireAxon CNN	0.998	0.870	0.710	0.674
	Human expert 2	0.996	0.759	0.716	0.564
Human expert 2+3	EntireAxon CNN	0.998	0.781	0.607	0.590
	Human expert 1	0.996	0.747	0.592	0.421

Fig. 3 The EntireAxon convolutional neural network (CNN) sensitively and specifically recognizes features of axonal degeneration (AxD). a, Validation of the EntireAxon CNN performance for all four classes 'background', 'axon', 'axonal swelling' and 'axonal fragment' in before unseen phase-contrast microscopic images. **b**, Comparison of the mean F1 scores between the EntireAxon CNN and two human experts on the ground truth (human expert 1 who also labeled the training images) to recognize background, axon, axonal swelling and axonal fragments. **c**, Phase-contrast validation image, its EntireAxon CNN segmentation mask, and the consensus labeling masks of two human experts that show the overlap (cyan) or difference (red) between the labels. Scalebar, 100 µm. **d**, Comparison of the mean F1 scores between the EntireAxon CNN and the human expert on the consensus labeling of the other two human experts.

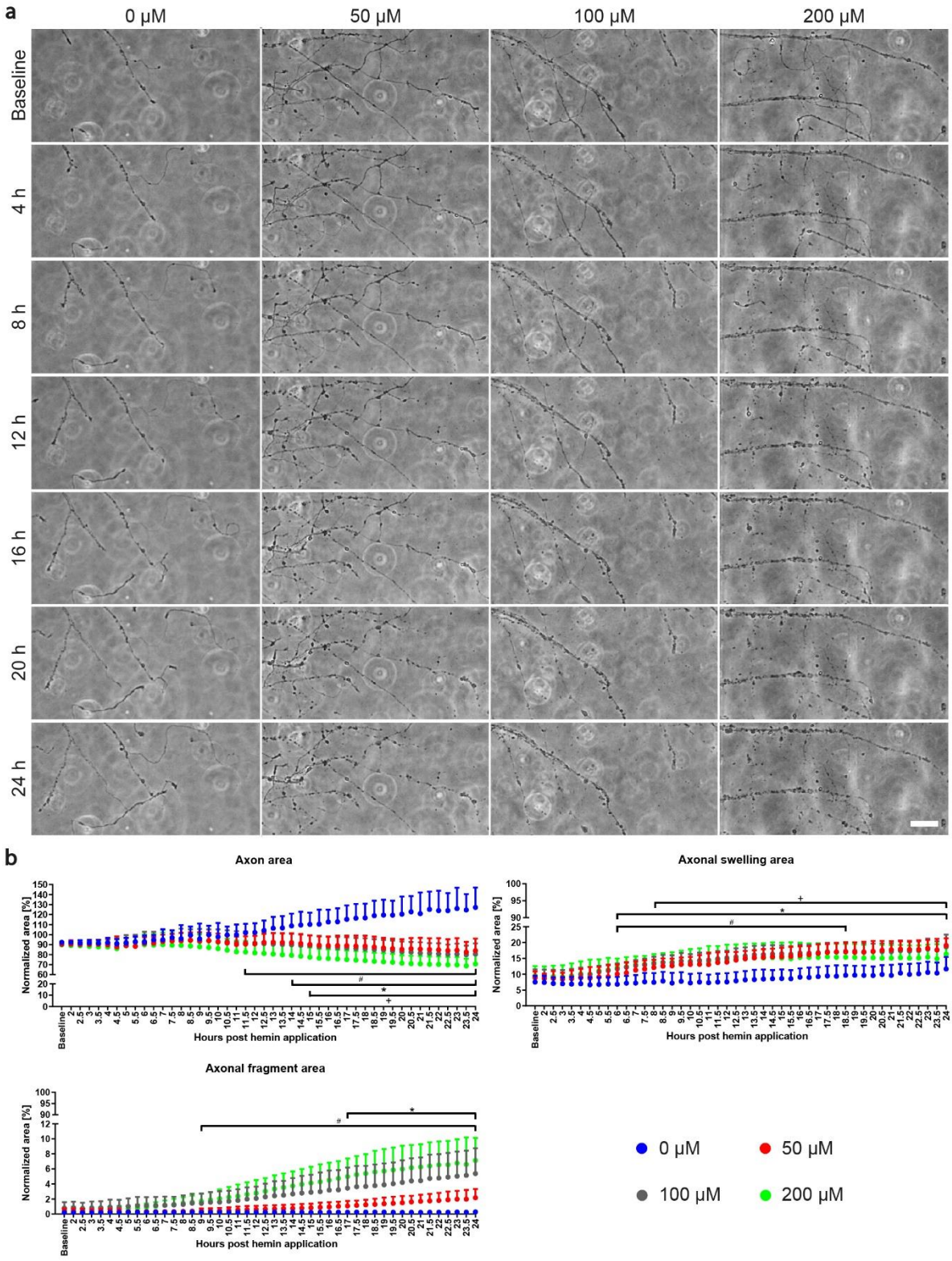


Fig. 4 Hemin-induced axonal degeneration (AxD) progresses time- and concentration-dependently. a, Axons treated with hemin (50, 100, 200 µM) degenerated compared to vehicle-treated axons (0 µM) that continued to grow. Scalebar, 50 µm. For complete time-lapse videos including segmentation, refer to **suppl. Videos 1-4**. **b**, Quantification of AxD over 24 hours in phase-contrast images. For AUC analyses, please refer to **suppl. Fig. 2**. To determine the time course, the sum of pixels in each class and hemin concentration over time was normalized to the baseline of that class and condition. The quantification of the phase-contrast images over 24 hours revealed significantly smaller axon areas starting at 11.5 hours (200 µM), 14 hours (100 µM), and 15 hours (50 µM) after hemin treatment compared to control (0 µM). The axonal fragment area significantly increased from 9.5 hours onwards (200 µM vs. 0 µM) and from 17.5 hours (100 µM vs. 0 µM), while the axonal swelling area increased from 6 hours onwards (100-200 µM) and from 8 hours (50 µM). n=6 independent cultures of primary cortical

788 neurons. Means ± SD are given. +, *, # $p < 0.05$; + = 50 µM vs. 0 µM, * = 100 µM vs. 0 µM, # = 200 µM vs. 0 µM. For exact
789 *p* values, refer to **suppl. Tab. 1**.

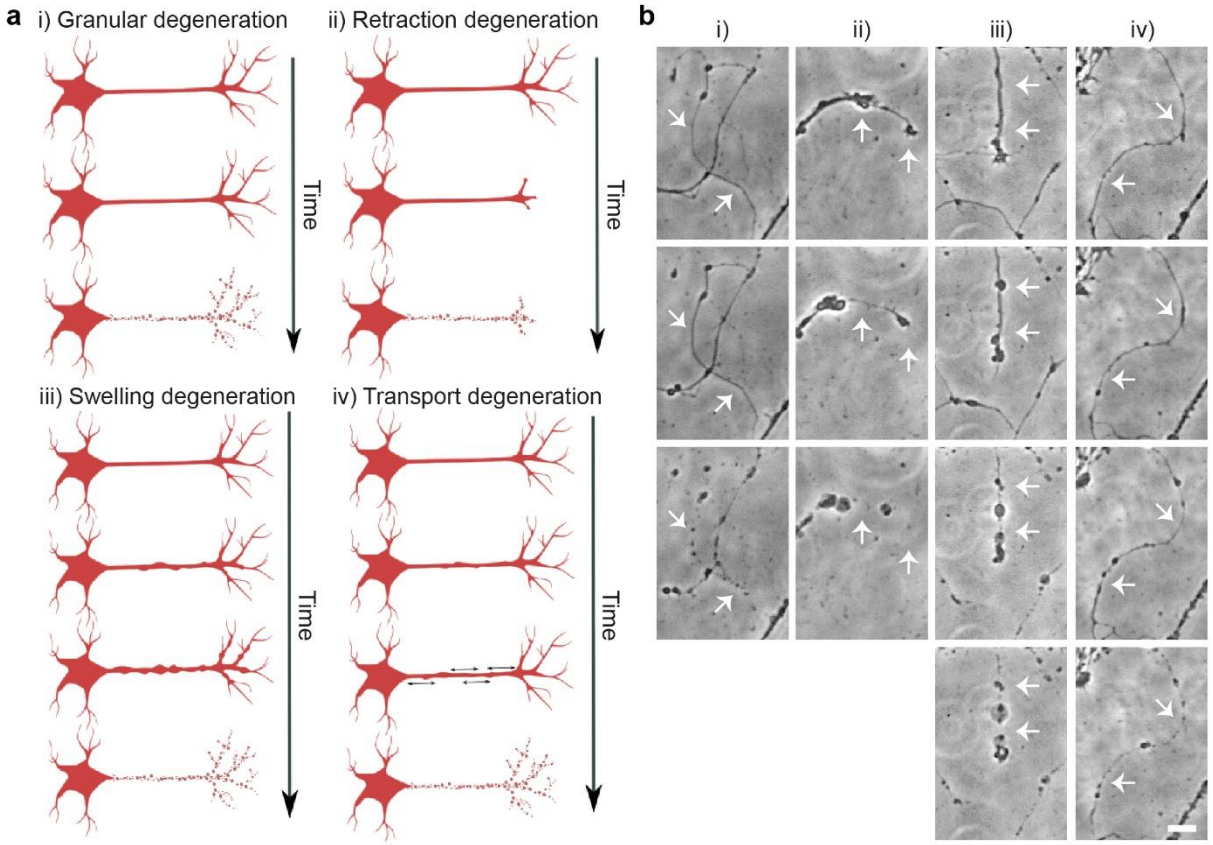
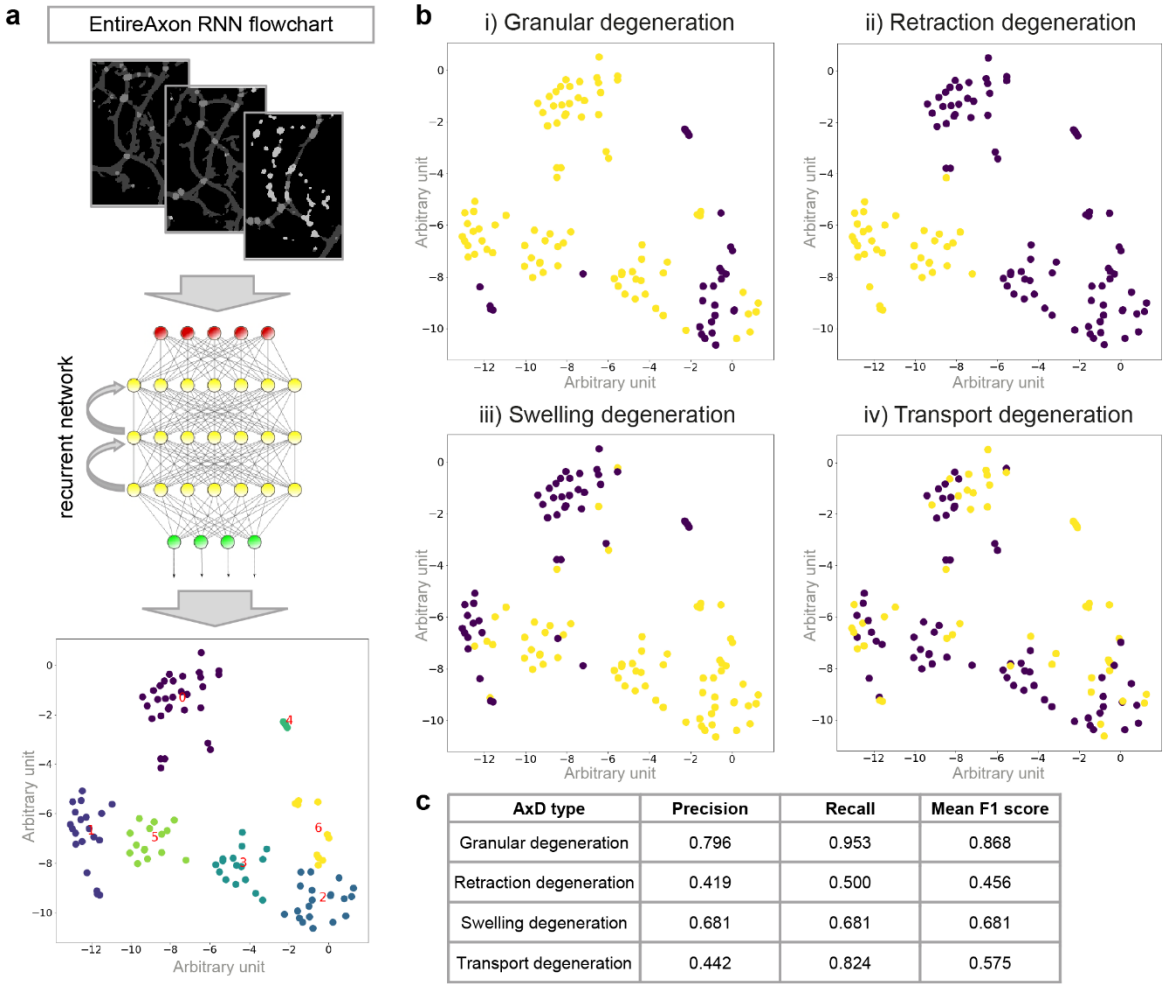


Fig. 5 Axonal degeneration (AxD) is a morphologically heterogeneous process. **a**, Schematic overview of the proposed AxD morphological patterns: i) granular degeneration, ii) retraction degeneration, iii) swelling degeneration, and iv) transport degeneration. **b**, Phase-contrast recordings of the four different morphological patterns of AxD. Granular degeneration (i) is characterized by the fragmentation of the axon (white arrows). During retraction degeneration (ii), the axonal growth cone retracts in the proximal direction and the part of the axon in proximity of the growth cone disintegrates accompanied by axonal swellings (white arrows). During swelling degeneration (iii), many axonal swellings enlarge resulting in axonal fragments (white arrows). During transport degeneration (iv), axonal swellings are transported along the axon prior to the degeneration of the axon (white arrows). Scalebar, 20 µm. For complete time-lapse videos including segmentation, refer to **suppl. Videos 5-8.**



AxD type	Precision	Recall	Mean F1 score
Granular degeneration	0.796	0.953	0.868
Retraction degeneration	0.419	0.500	0.456
Swelling degeneration	0.681	0.681	0.681
Transport degeneration	0.442	0.824	0.575

Fig. 6 The EntireAxon RNN deciphers four distinct patterns of axonal degeneration (AxD). **a,** Schematic workflow of recurrent neural network (RNN) to recognize and quantify morphological patterns of AxD based on the identification of seven clusters. The EntireAxon CNN segmentation masks were used for the RNN training, which determined the change in class over time. Based on the 16 different possible class changes, the RNN determined seven clusters (cluster 0-6). **b**, The clusters classify the four different morphological patterns of AxD with yellow indicating included and purple indicating excluded clusters: i) granular, ii) retraction, iii) swelling, and iv) transport degeneration. Clusters of granular degeneration overlap with recognized clusters of other morphological patterns (retraction, swelling, and transport degeneration). For more details on the morphological changes underlying the cluster analysis, refer to **Suppl. Fig. 4**. **c**, 10-fold cross-validation of the four morphological patterns of AxD.

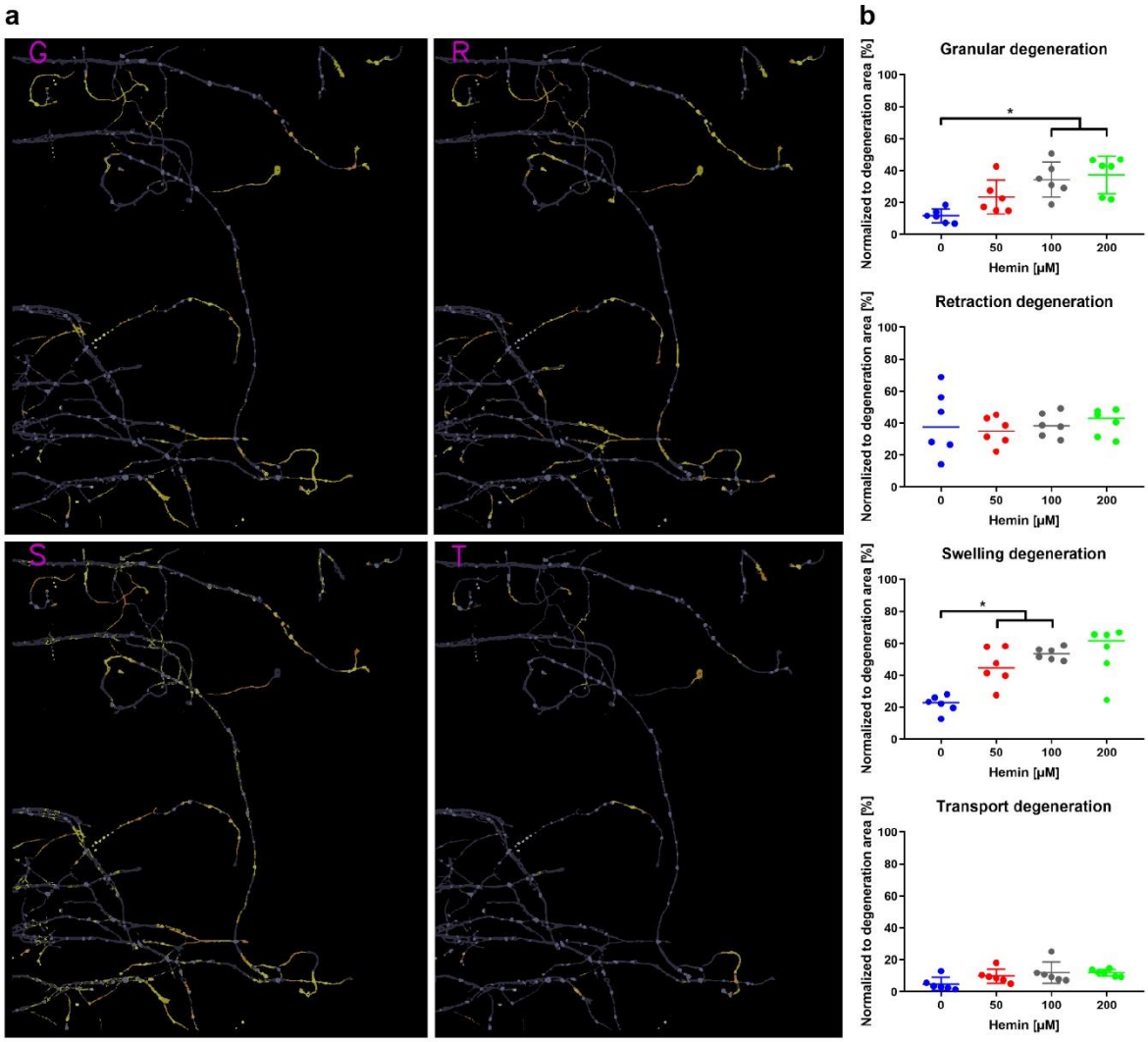


Fig. 7 Hemin induces concentration-dependent differences in the morphological patterns of axonal degeneration (AxD). **a**, The classification of granular (G), retraction (R), swelling (S), and transport degeneration (T) in axons treated with 200 µM hemin. Scalebar, 100 µm. For the complete time-lapse video including segmentation, refer to **suppl. Videos 9. b**, Quantification of the four morphological patterns of AxD in hemin-induced AxD. n=6 independent cultures of primary cortical neurons. Means ± SD are given for granular and transport degeneration and, medians are given for retraction and swelling degeneration. * $p < 0.05$ vs. 0 µM for granular degeneration, * $p < 0.0083$ vs. 0 µM for swelling degeneration due to manual Bonferroni correction for nonparametric data. For exact *p* values, refer to **suppl. Tab. 3**.