Deep learning to decipher the progression and morphology of

axonal degeneration

Running title: EntireAxon deep learning of axonal degeneration

Alex Palumbo[1,2,3,*], Philipp Grüning[4], Svenja Kim Landt[1,2], Lara Eleen Heckmann[1,2], Luisa Bartram[1], Alessa Pabst[1,2], Charlotte Flory[1,2], Maulana Ikhsan[1,2,3], Sören Pietsch[1,2,5], Reinhard Schulz[6], Christopher Kren[7], Norbert Koop[7], Johannes Boltze[1,2,8], Amir Madany Mamlouk[4] and Marietta Zille[1,2,3,*]

Affiliations

[1] Fraunhofer Research and Development Center for Marine and Cellular Biotechnology, Fraunhofer Research Institution for Individualized and Cell-Based Medical Engineering, 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

23   * Corresponding authors:

24   Alex Palumbo

25   Postal address: Fraunhofer Research and Development Center for Marine and Cellular

26   Biotechnology, Fraunhofer Research Institution for Individualized and Cell-Based Medical

27   Engineering, Mönkhofer Weg 239a, 23562, Lübeck, Germany

28   Email: alex.palumbo@uni-luebeck.de

29   Phone: +49 451 3101 7227

30

31   Marietta Zille

32   Postal address: Institute for Experimental and Clinical Pharmacology and Toxicology, University

33   of Lübeck, Ratzeburger Allee 160, 23562, Lübeck, Germany

34   Email: m.zille@uni-luebeck.de

35   Phone: +49 451 3101 7227

36

37   **ORCID**

38   Alex Palumbo ORCID: 0000-0002-1631-785X

39   Philipp Grüning ORCID: 0000-0003-2946-4020

40   Svenja Kim Landt ORCID: 0000-0002-4133-7035

41   Lara Eleen Heckmann ORCID: 0000-0003-1264-9715

42   Alessa Pabst ORCID: 0000-0003-1938-3753

43   Charlotte Flory ORCID: 0000-0001-5049-6380

44   Maulana Ikhsan ORCID: 0000-0002-8756-4524

45   Johannes Boltze ORCID: 0000-0003-3956-4164

46   Amir Madany Mamlouk ORCID: 0000-0001-9709-1620

47   Marietta Zille ORCID: 0000-0002-0609-8956

Abstract

Background: Axonal degeneration (AxD) is a pathological hallmark of many neurodegenerative diseases. Deciphering the morphological patterns of AxD will help to understand the underlying mechanisms and to develop effective therapeutic interventions. Here, we evaluated the progression of AxD in cortical neurons using a novel microfluidic device in combination with a deep learning tool, the EntireAxon, that we developed for the enhanced-throughput analysis of AxD on microscopic images.

Results: The EntireAxon convolutional neural network sensitively and specifically segmented the features of AxD, including axons, axonal swellings, and axonal fragments, and its performance exceeded that of human expert raters. In an *in vitro* model of AxD in hemorrhagic stroke induced by the hemolysis product hemin, we detected the concentration- and time-dependent degeneration of axons leading to a decrease in axon area, while the axonal swelling and axonal fragment area increased. Time course analysis revealed that axonal swellings preceded axon fragmentation, suggesting that swellings may be reliable predictors of AxD. Using a recurrent neural network, we further identified four morphological patterns of AxD (granular, retraction, swelling, and transport degeneration) in cortical axons subjected to hemin.

Conclusions: These findings indicate a morphological heterogeneity of AxD under pathophysiologic conditions. The combination of the microfluidic device 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.

Keywords: axon, brain hemorrhage, cell culture, machine learning, microfluidic, microscopy, stroke, time-lapse

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 (Luo and O'Leary, 2005; Lingor et al., 2012; Salvadores et al., 2017). Apart from axonal fragments, axon swellings (also called axonal beadings, bubblings or spheroids) are a hallmark of degenerating axons (Saxena and Caroni, 2007; Lingor et al., 2012; Wang et al., 2012), containing disorganized cytoskeleton and organelles resulting from an interruption of axonal transport (Coleman, 2005; Nikić et al., 2011; Yong et al., 2019).

It is known that axons disintegrate in different ways depending on the biological context. During development and neural circuit assembly, inappropriately grown axons can undergo axonal retraction, axonal shedding or local AxD (Pease and Segal, 2014; Neukomm and Freeman, 2014). Axonal retraction is characterized by retraction bulb formation at the distal tip, and subsequent pullback (Pease and Segal, 2014). During axonal shedding, the axon retracts leaving behind small pieces of its distal part (axosomes) (Bishop et al., 2004). Local AxD is characterized by axon disintegration into separated axonal fragments (Neukomm and Freeman, 2014). Acutely and chronically injured axons may degenerate retrogradely (distal-to-proximal direction, dying-back), anterogradely (proximal-to-distal direction) or in a Wallerian degeneration pattern (distal part of the axon from injury site), ultimately resulting in the generation of axonal fragments (Cavanagh, 1979; Coleman, 2005; Beirowski et al., 2005). However, AxD patterns have been mainly described in extracerebral axons in models of nutrient deprivation or axotomy.

Not much is known on AxD in cortical neurons subjected to a disease-specific cytotoxic micromilieu. A distinct pathological micromilieu has recently been observed for hemorrhagic stroke, after which the lysis of erythrocytes from the hematoma leads to the release of the cytotoxic product hemin (Robinson et al., 2009; Zille et al., 2017). Patients suffering from hemorrhagic stroke often experience AxD that is associated with worse motor and functional outcome

96   (Venkatasubramanian et al., 2013; Chen et al., 2018). Importantly, AxD occurs in the subacute

97   stages of hemorrhagic stroke. Thus, addressing AxD may not only provide a new therapeutic

98   target, but also a much wider time window for intervention. Since not much is known about the

99   mechanisms, morphological patterns, and the temporal progression of AxD in the context of

100   hemorrhagic stroke, we here sought to examine the progression of AxD and its associated

101   morphological alterations.

102   As the disintegration of the axons endures from minutes to hours (Beirowski et al., 2005;

103   Kerschensteiner et al., 2005), it is necessary to monitor the spatiotemporal progression of AxD

104   and its morphological hallmarks continuously. However, conventional software solutions fail to

105   automatically detect and quantify high axon numbers as well as axonal swellings and fragments in

106   phase-contrast microscopic images. The reason may be two-fold: 1) Conventional software relies

107   on image binarization (Sasaki et al., 2009; Becker and Madany, 2012), which can lead to

108   information loss and low sensitivity as thin axons may not be recognized. 2) The analysis requires

109   subjective and time-consuming manual annotations, e.g., thresholding and defining the region of

110   interest (Pool et al., 2008; Ho et al., 2011; Li et al., 2014). So far, immunostained images were

111   used to investigate morphological changes in AxD as the analysis of phase-contrast images has

112   been limited by the lower target-to-background signal. Immunofluorescence images, however,

113   entail certain disadvantages such as photobleaching and the requirement for cell fixation, which

114   restricts observations to a single time point. Thus, a software tool for the automatized detection

115   and quantification of the morphological patterns of AxD in long-term live cell imaging is required

116   to improve both sensitivity and throughput to overcome current limitations in understanding AxD.

117   In this study, we demonstrate that cortical axons underwent AxD after the exposure to the

118   hemolysis product hemin, with axonal swellings preceding axon fragmentation. Deep learning

119   further detected the occurrence of four AxD patterns being characterized as granular, retraction,

120   swelling, and transport degeneration. This may inform downstream AxD and neurodegeneration

121   research in health and disease. We also provide tools for the enhanced throughput analysis of AxD,

122   including a microfluidic device containing 16 independent experimental units and the deep

123   learning platform "EntireAxon" to analyze AxD, which will help augment our understanding of

124   AxD and may also support the development of novel treatment approaches for neurodegenerative

125   diseases.

Results

An enhanced throughput microfluidic device and the EntireAxon deep learning tool allow the longitudinal study of axonal degeneration

The major limiting factor of commercially available microfluidic devices to study AxD 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 enable the systematic analysis of AxD *in vitro*, we 1) manufactured a microfluidic device containing 16 individual microfluidic units (**Fig. 1 and Supplementary Fig. S1**) that can be investigated in parallel and recorded simultaneously, and 2) trained a convolutional neural network (CNN), the EntireAxon, to segment all relevant features of AxD, i.e., axons, axonal swellings, and axonal fragments (**Fig. 2**).

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**). Next, we compared the performance of the EntireAxon CNN on the ground truth (human expert 1) with 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

151   three experts (**Fig. 3D**). Collectively, this suggests that the EntireAxon CNN sensitively and

152   specifically recognizes axons and the morphological features of AxD.

153

Axonal integrity is lost over time with axonal swellings preceding axon fragmentation

155   We then applied the EntireAxon CNN to assess AxD in the context of hemorrhagic stroke. We

156   applied the hemolysis product hemin, a commonly used agent to mimic hemorrhagic stroke *in*

157   *vitro* (Robinson et al., 2009; Zille et al., 2017; Chen and Regan, 2004), on primary cortical axons.

158   Accordingly, isolated axons were exposed to the hemolysis product hemin and recorded by time-

159   lapse microscopy for 24 hours. Hemin induced concentration- and time-dependent morphological

160   changes leading to AxD compared to vehicle-treated axons (**Fig. 4 and Videos S1-4**). Area under

161   the curve (AUC) analyses revealed a significant decrease in axon area in all three hemin

162   concentrations (50 µM vs. 0 µM: $P = 0.026$; 100 µM vs. 0 µM: $P = 0.018$, 200 µM vs. 0 µM:

163    $P < 0.001$). The axonal swelling area also increased in all three concentrations (50 µM vs. 0 µM:

164    $P = 0.012$, 100 µM vs. 0 µM: $P = 0.005$, 200 µM vs. 0 µM: $P = 0.016$), while the axonal fragment

165   area was elevated only for axons treated with 100 and 200 µM hemin (vs. 0 µM: $P = 0.004$, **Fig. 5**

166   **and Table S2**).

167   Comparing the time course of AxD between hemin- and vehicle-treated axons (0 µM), the axon

168   area decreased starting at 11.5 hours at 200 µM ($P = 0.020$, from 15 hours $P < 0.001$), at 14 hours

169   at 100 µM ($P = 0.040$, from 18.5 hours $P<0.001$), and at 15 hours at 50 µM ($P = 0.018$, from

170   19 hours $P < 0.001$). Hemin treatment also elevated the axonal fragment area starting at 9 hours at

171   200 µM ($P = 0.037$) and at 17 hours at 100 µM hemin ($P = 0.044$). Interestingly, the axonal

172   swelling area increased prior to the changes in axon and axonal fragment area, i.e., starting at

173   6 hours at 200 µM ($P = 0.010$) and 100 µM ($P = 0.019$), and at 8 hours at 50 µM hemin

174   ($P = 0.030$). For the highest hemin concentration, the increase was only transient (until

9

175   18.5 hours), suggesting that axonal swellings preceded the axon fragmentation (**Table S1**), which

176   can also be seen in the time-lapse recordings (**Videos S2-4**).

177   The results of the time course analysis were further substantiated by live cell fluorescent staining

178   (calcein AM), which indicated the starting point of AxD after hemin treatment between 8 and

179   12 hours for 200 μM hemin, between 12 and 16 hours for 100 μM hemin and 16 and 20 hours for

180   50 μM hemin (**Supplementary Fig. S2**). Taken together, AxD progression depends on the severity

181   of the insult and axonal swellings may be reliable predictors of AxD.

182

183   **Deep learning deciphers four patterns of AxD**

184   AxD time-lapse data revealed different morphological patterns of degeneration that can occur in

185   the same axons over time (**Fig. 6** and **Videos S5-8**). We categorized these morphological patterns

186   as:

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

188   ii)   Retraction degeneration: AxD in which the distal part of the axon retracts ultimately

189   resulting in granular degeneration.

190   iii)   Swelling degeneration: AxD in which axonal swellings enlarge, followed by granular

191   degeneration.

192   iv)   Transport degeneration: AxD in which axonal swellings of constant size, which do not

193   enlarge, are transported along the axon resulting in granular degeneration.

194   We trained a recurrent neural network (RNN), the EntireAxon RNN, to identify these

195   morphological patterns based on changes in class segregation over time using a training dataset of

196   AxD segmentation recordings (**Fig. 7A**). Given the four different classes (background, axon,

197   axonal swelling, and axonal fragment), 16 different class pairs can occur between a segmentation

198   at time step *t* and time step *t+1*. For example, a background pixel at *t* can either remain background

199   pixel at *t+1* or change into one of the other three classes, and the same is true for the other classes.

200   Thus, in total, four times four class pairs are possible. We used a window size of 32x32, of which

201   always the probability of a class pair in the central pixel relative to the previous time point was

202   computed for each time point and across the entire image.

203   The RNN determined seven clusters (cluster 0-6) that were characterized by an idiosyncratic

204   pattern of changes in class distribution over 24 hours (**Supplementary Fig. S3**). All clusters

205   showed a decrease in the class 'axon' and an increase in the class 'background'. Depending on the

206   hemin concentration, the changes occurred at a different magnitude and at different time points,

207   and with concomitant increases in either the class 'axonal swelling' and/or 'axonal fragment'. In

208   cluster 0, there was an early decrease in the class 'axon', which then continued more linearly as

209   well as a later rise in the class 'axonal fragment'. In contrast to cluster 0, cluster 1 showed no

210   increase in the class 'axonal fragment' and a linear decrease in the class 'axon' from the start. In

211   cluster 2, there was a strong increase in the class 'axonal swelling'. Cluster 3 demonstrated an

212   early and lasting high level of the class 'axonal swelling' with a later increase in the class 'axonal

213   fragment'. Cluster 4 showed a rapid decrease in the class 'axon' concomitant with an increase in

214   the classes 'background' and 'axonal swelling'. Cluster 5 was similar to cluster 1, but with an early

215   drop in the class 'axon'. Cluster 6 showed an increase in the class 'axonal swelling' similar to but

216   to a greater extent than cluster 2.

217   The RNN categorized each cluster to one of the four morphological patterns (**Fig. 7B**): i) Granular

218   degeneration was defined by clusters that describe the degeneration of axons into axonal

219   fragments, i.e. clusters 0, 1, 3, and 5. ii) Retraction degeneration only included the clusters 1 and

220   5, indicating the retraction of the axon followed by its fragmentation. iii) Swelling degeneration

221   was characterized by the three clusters that included the class 'axonal swelling, i.e., clusters 2, 3,

222   and 6, as well as cluster 5 showing the exchange of the class 'axon' for 'background'. iv) Transport

223   degeneration was the only pattern that relied on cluster 4 and was also characterized partly on

224   clusters 0, 1, 2, and 6. Although some clusters overlap among morphological patterns, the unique

225   combination of the different clusters allows to distinguish all four morphological patterns.

226   To validate the EntireAxon RNN, a 10-fold cross-validation was performed. Therefore, the dataset

227   was randomly divided into 10 datasets and ten models were trained with 9 of the datasets leaving

228   the remaining dataset for validation (not previously seen by the RNN). Based on the combined test

229   samples, the RNN was able to distinguish between the four morphological patterns of AxD

230   (**Fig. 7C**). These data confirm that the combination of the different AxD features as well as their

231   spatiotemporal progression defines distinct morphological AxD patterns.

232

The morphological patterns of AxD depend on the extent of AxD

234   We then applied the EntireAxon RNN to quantify the occurrence of the four morphological

235   patterns of AxD in the context of hemorrhagic stroke (**Fig. 8 and Video S9**). While all AxD

236   patterns were detected (**Fig. 8A**), hemin concentration-dependently increased granular

237   degeneration ($P < 0.001$), swelling degeneration ($P < 0.001$), and transport degeneration

238   ($P = 0.025$, **Fig. 8B**). When comparing the slopes of the different AxD patterns under hemin

239   exposure, granular and swelling degeneration were significantly different from transport

240   degeneration ($P = 0.005$ and $P = 0.004$, respectively, **Table S3**). Collectively, our data suggest that

241   hemin concentration-dependently induces different morphological patterns of AxD in cortical

242   axons.

Discussion

We here describe the occurrence of four morphological patterns of AxD under pathophysiological conditions: 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, a novel microfluidic device and the EntireAxon, allow increasing the experimental yield, the in-depth enhanced throughput analysis of AxD as well as the longitudinal investigation of AxD.

We propose a novel monolithic microfluidic device consisting of 16 individual microfluidic units that enables the parallel and separated treatment and/or manipulation of axons and somata (**Fig. 1**). The currently available devices do not allow enhanced throughput experiments as they comprise only single microfluidic units (Park et al., 2006; Van Laar et al., 2019). Although some devices can harbor multiple experimental conditions, they employ a radial design with a single soma compartment, in which one experimental condition may influence another due to the potential of retrograde signaling (Hosmane et al., 2010; Biffi, 2015). Another option is the parallel use of multiple individual devices, which allows handling up to 12 devices in a conventional 12-well plate (Li et al., 2014). Compared to our device, this procedure is time-consuming in both the manufacturing and adjustment for recordings.

The extent of AxD has so far been mainly investigated with a focus on axon fragmentation as primary readout. To quantify axon fragmentation, Sasaki and colleagues introduced the AxD index as the ratio of fragmented axon area versus total axonal area (Sasaki et al., 2009). However, the AxD index did not include axonal swellings, which are a characteristic feature of degenerating axons (Yong et al., 2019; Cui et al., 2020). Although other analyses considered axonal swellings as a morphological feature of AxD (Nikić et al., 2011; Yong et al., 2019), the approaches were time-consuming and required manual annotations.

13

267   We herein adapted a standard u-net with ResNet-50 encoder (Ronneberger et al., 2015; He et al.,

268   2015) and used a CNN ensemble, which combines predictions from multiple CNNs to generate a

269   final output and is superior to individual CNNs (Dietterich, 2000; Huang et al., 2016; Vuola et al.,

270   2019). The EntireAxon CNN performs an automatic segmentation and quantification of axons and

271   morphological features relevant to AxD, including axonal swellings and fragments, on phase-

272   contrast time-lapse microscopy images (**Fig. 2**). The EntireAxon CNN recognized the four classes

273   'background', 'axon', 'axonal swelling', and 'axonal fragment', with the highest mean F1 score

274   for the class 'background' (**Fig. 3A**). The comparably lower performance of the CNN to recognize

275   axonal fragments may be explained by the disproportional distribution of pixels in the training and

276   validation data ('background' mean of 96.42 % of pixels, 'axon' 2.77%, 'axonal swelling' 0.58%,

277   'axonal fragment' 0.23 %). Hence, every individual segmentation error more strongly affects the

278   false positive or false negative rate in these classes.

279   Comparison with human experts revealed that the EntireAxon CNN reached a similar performance

280   level. As expected, its performance was slightly better than the human experts on the ground truth

281   as both, ground truth and training data, were labeled by the same human expert (**Fig. 3B**).

282   Interestingly, when comparing the EntireAxon CNN with a human expert on the consensus label

283   of the other two human experts, not only was the EntireAxon CNN as good as or even better than

284   the human expert, but the mean F1 scores were also higher than on the ground truth labels

285   (**Fig. 3D**). This may be because pixels that were differentially assigned by the human expert, i.e.

286   more difficult to classify, were excluded from the comparison. Taken together, these findings

287   demonstrate that the EntireAxon CNN is suitable to automatically quantify AxD and its

288   accompanying morphological changes in an enhanced throughput manner.

289   Conventional *in vitro* models of AxD rely mainly on nutrient deprivation or axotomy and focus on

290   axons outside the brain. However, AxD is not only an active and commonly observed process in

291   the brain, but it is also believed to be caused by more complex mechanisms given the different

292   microenvironments in which it may occur. For example, AxD has been demonstrated to occur in

293   intracerebral hemorrhage (Venkatasubramanian et al., 2013; Tao et al., 2017). In this context,

294   cortical axons are exposed to a cytotoxic microenvironment due to hemolysis leading to the release

295   of blood breakdown products, whose effects on axons remain to be elucidated (Hemorrhagic

296   Stroke Academia Industry (HEADS) Roundtable Participants, 2018). We therefore modeled

297   hemorrhagic stroke by exposing axons from primary cortical neurons to the hemolysis product

298   hemin and investigated the progression of AxD. Similar to previous results where 100 µM hemin

299   were sufficient to induce significant neuronal cell death in conventional cultures of somata and

300   axons (Zille et al., 2017), we here observed that 100 µM hemin led to a significant decrease in

301   axon area and an increase in axonal swelling and fragment area (**Fig. 4**).

302   The progression of AxD undergoes a latent phase, during which the structural integrity of the axon

303   is maintained, followed by a catastrophic phase with the rapid disintegration of the axon (Yong et

304   al., 2019). In our model, the catastrophic phase of AxD started within 12 to 18 hours after the

305   administration of hemin (**Fig. 4 and Supplementary Fig. S2**). Similar durations of the latent

306   phases of AxD have been observed in other models. For instance, under circumstances of growth

307   factor withdrawal, the transition to the catastrophic phase occurred at 12-24 hours (Nikolaev et al.,

308   2009; Maor-Nof et al., 2016; Yong et al., 2019).

309   We further demonstrated that the relative axon area decreased at higher hemin concentrations,

310   while the axonal fragment area increased. Our results are in accordance with other experimental

311   conditions such as axotomy-mediated or paclitaxel-induced AxD, in which axonal fragments also

312   increased (Sasaki et al., 2009; Pease-Raissi et al., 2017). As the axonal swelling area preceded the

313   increase of axonal fragments and axon area loss, our findings are also in line with results reported

314   in a model of experimental autoimmune encephalomyelitis indicating that axonal swelling

315   anticipates fragmentation (Nikić et al., 2011). This suggests that axonal swelling may be a reliable

316   predictor of AxD.

317 Interestingly, axonal swellings and axonal fragments were related to different morphological

318 patterns of AxD. Specifically, we observed axons that showed signs of axonal retraction, enlarging

319 of axonal swellings and axonal transport before degeneration (**Fig. 6**). We therefore trained the

320 EntireAxon RNN to quantify the occurrence of four morphological patterns of AxD, i.e. granular,

321 retraction, swelling, and transport degeneration, based on the clusters of unique changes of classes

322 over time (**Fig. 7 and Supplementary Fig. S3**). These patterns have not been described to occur

323 simultaneously in the same biological condition: Granular degeneration has previously been

324 observed in retrograde, anterograde, Wallerian and local AxD after axotomy or trophic factor

325 deprivation (Cavanagh, 1979; Coleman, 2005; Beirowski et al., 2005; Neukomm and Freeman,

326 2014). Retraction degeneration has been described in axonal retraction and shedding in

327 developmental AxD (Bishop et al., 2004; Pease and Segal, 2014). Swelling degeneration was

328 previously reported in experimental autoimmune encephalitis and growth factor deprivation (Nikić

329 et al., 2011; Yong et al., 2019). Transport degeneration has not been reported before. However,

330 microtubule breaks have been demonstrated in a model of axonal stretch injury. Those developed

331 into axonal swellings resulting in axonal transport interruption with AxD as a consequence (Tang-

332 Schomer et al., 2012).

333 Our data demonstrate that all four morphological degeneration patterns can occur along cortical

334 axons (**Fig. 8**). Interestingly, we also observed a concentration-dependent effect in the context of

335 hemorrhagic stroke. Granular, swelling, and transport degeneration were significantly increased

336 with increasing hemin concentrations, with granular and swelling degeneration being more

337 strongly correlated. To what extent our model of hemin-induced AxD in hemorrhagic stroke is

338 molecularly similar to developmental or pathophysiological AxD needs to be further investigated

339 along with the underlying molecular mechanisms of the four patterns of AxD. This could be greatly

340 facilitated by the EntireAxon RNN that is able to automatically detect the morphological patterns

16

341   in time-lapse recording due to its capacity to relate each output to previous images in the stacks by

342   its current units.

343

Limitations and outlook

345   i) Our microfluidic device currently does not allow to investigate AxD at more proximal axonal

346   parts to the soma such as the axonal initial segment. Shortening the length of the microgrooves or

347   including a more proximal compartment, are possible modifications of the current design.

348   ii) Our results are based on unmyelinated axons. Co-culture with glia cells that may play a role in

349   AxD is possible in the presented microfluidic device and the time course and morphological

350   changes may be different under co-culture conditions. These studies are of high relevance to the

351   field, but go beyond the scope of the present study.

352   iii) The observed effects of AxD in hemorrhagic stroke within this study were based on hemin

353   toxicity, and we cannot exclude that other hemolysis products such as thrombin or bilirubin have

354   different effects. Additional studies should investigate differences of hemolysis products to

355   increase our understanding of the mechanisms of AxD in hemorrhagic stroke.

356   iv) The overall CNN performance may be further improved with more general inputs. For example,

357   the segmentation of fragment pixels cannot be conducted accurately based on a single image at a

358   specific time point as the whole process of AxD, ultimately resulting in the disintegration of the

359   axons (i.e., the generation of axonal fragments), needs to be considered. CNNs using 3D

360   convolutions could, in principle, perform a segmentation over an entire time-lapse recording and

361   model temporal dependencies. However, we decided against the 3D approach, as it severely

362   restricts general applicability due to its greatly increased effort to label suitable time series for

363   training. In this context, the identification of the images that will yield the best results is crucial to

17

364   effectively reduce labeling costs, which we have previously described using an active learning

365   method (Grüning, P. et al., 2020).

366

Conclusion

368   In combination with an advanced microfluidic device, the EntireAxon deep learning tool expands

369   our possibilities to track AxD by detecting axons, axonal swellings, and axonal fragments. We

370   further identified four morphological patterns of AxD, i.e., granular, retraction, swelling, and

371   transport degeneration, under pathophysiological conditions in the context of hemorrhagic stroke.

372   This approach will help to tackle the complex processes of AxD and may significantly enhance

373   our understanding of AxD in health and disease to develop novel therapeutic strategies for brain

374   diseases.

375

Methods

Chemicals and reagents are listed in **Tables S4-5**.

Study design

Sample size: Six mice. We did not perform a priori power analysis as this was an exploratory study. We did not change the number of the mice during the course of the study.

Data inclusion/exclusion criteria: Recordings that did not have any technical flaws, such as shifting of the microfluidic device in x, y, or z-axis were included. Recordings with minor x and y-axis shifts that we were able to correct by post-recording alignment (see **Image preprocessing**) were included. All data were processed using the same settings. The training and validation images for the deep learning tool were chosen to represent the testing data as best as possible.

Outliers: No outliers have been excluded in the study.

Selection of endpoints: Endpoints were the area of the axons, axonal swellings, and axonal fragments, respectively.

Replicates: Each individual mouse counted as a biological replicate (N = 6 biological replicate per experiment). Four different microfluidic units have been used for four experimental conditions (0, 50, 100, 200 µM hemin) per biological replicate.

Research objectives: The research objective was to examine the progression of axonal degeneration in primary cortical neurons upon hemin exposure. Therefore, a microfluidic device and deep learning tool to increase the experimental yield and to enable unbiased automatic analysis was developed. Our pre-specified hypothesis was to detect a concentration-dependent effect of hemin on axons. Our suggested hypothesis after conducting time-lapse recording was that there are four morphological patterns of axonal degeneration and that those depend on the severity of axonal degeneration by different hemin concentrations.

400   *Research subjects or units of investigation:* We employed primary cortical neurons from Crl:CD1

401   (ICR) Swiss outbred mice.

402   *Experimental design:* Randomized controlled laboratory experiment with four different

403   concentrations of hemin treatment to induce and record axonal degeneration by time-lapse

404   microscopy and further validation by fluorescence microscopy.

405   *Randomization:* Microfluidic units have randomly been assigned to one of the four experimental

406   conditions (0, 50, 100, 200 µM hemin).

407   *Blinding:* The experimenter was not blinded when axons were treated with different hemin

408   concentrations. The actual analysis was objective as being conducted solely by the deep learning

409   tool.

410

411   **Fabrication of an enhanced throughput microfluidic device based on soft lithographic**

412   **replica molding**

413   Thirty-two wells were milled in a polymethyl methacrylate (PMMA) plate of the size of a

414   conventional cell culture plate (**Fig. 1A and Supplementary Fig. S1**) using a universal milling

415   machine (Mikron WF21C, Mikron Holding AG) with a 1 mm triple tooth cutter (HSS-CO8 Type

416   N, Holex) at a precision of 0.01 mm. During the milling procedure, we applied a half-synthetic

417   cooling lubricant (Opta Cool 600 HS, Wisura GmbH) on a mineral base to reduce the debris.

418   Additionally, we milled screw holes in the intermediate spaces between each microfluidic unit to

419   later detach the PMMA from the negative casting mold. To remove debris, we washed the PMMA

420   plate by sonication (Sonicator Elmasonic S, Elma Schmidbauer GmbH) at room temperature for

421   30 minutes. Next, we lasered the microgrooves on the PMMA plate to connect both milled

422   compartments of each individual microfluidic unit by using an Excimerlaser (Excistar XS 193 nm,

20

423   Coherent). The PMMA plate was then washed again by sonication at room temperature for

424   30 minutes.

425   Polydimethylsiloxane (PDMS) was prepared in a 1:10 ratio and mixed properly before inducing

426   vacuum at 0.5 Torr in a vacuum desiccator (Jeio Tech VDC-31) for 30 minutes. After the PDMS

427   was poured into an empty aluminum basin to cover the ground, we applied vacuum at 0.5 Torr for

428   30 minutes to remove air bubbles. The PDMS was cured at room temperature for 48 hours. We

429   put the PMMA plate on top of the PDMS ground with the milled and lasered structures showing

430   upwards. Half of each well of the microfluidic units was filled with PDMS before curing at room

431   temperature for 48 hours. We mixed the epoxy solution in a 1:1 ratio and poured it over the

432   microfluidic device to cover its surface by at least 1 cm. Vacuum was applied at 0.5 Torr for

433   10 minutes to remove all air bubbles located above the channel side of the microfluidic device.

434   The epoxy was cured at room temperature for a minimum of 2 hours. We subsequently detached

435   the epoxy from the PMMA plate via a metallic block that consisted of screw holes in the

436   intermediate spaces between the individual systems. The epoxy represented a negative casting

437   mold to produce the microfluidic devices using PDMS.

438   PDMS was prepared as described above. We poured the PDMS into the negative epoxy casting

439   mold and applied vacuum at 0.5 Torr for 30 minutes. The liquid PDMS was cured at 75 °C for

440   2 hours to induce the polymerization. We peeled the microfluidic devices from the casting mold

441   and punched the wells with an 8 mm biopsy punch (DocCheck Shop GmbH) to ensure a sufficient

442   amount of medium for cell culture. We cleaned customized 115 x 78 x 1 mm glass slides by

443   sonication (Sonicator Elmasonic S, Elma Schmidbauer GmbH) and subsequently cleaned them by

444   ethanol before plasma treatment (High Power Expanded Plasma Cleaner, Harrick Plasma). Plasma

445   was applied at 45 W and 0.5 Torr for 2 minutes to activate the silanol groups of the glass slides

446   and the microfluidic devices enabling firm attachment.

21

447 We washed the microfluidic devices with ethanol and then twice with distilled water to remove

448 any debris. After aspirating the distilled water, except from the inside of the compartments,

449 0.1 mg/mL of poly-d-lysine solution in 0.02 M borate buffer (0.25 % (w/v) borate acid, 0.38 %

450 (w/v) sodium tetraborate in distilled water, pH 8.5) was used for coating at 4 °C overnight. We

451 aspirated the poly-d-lysine the next morning, not removing it from the compartments, and added

452 50 µg/mL of laminin as a second coating surface for incubation at 4 °C overnight. At the day of

453 neuron isolation, the microfluidic devices were washed twice with pre-warmed medium after

454 aspirating the laminin. Immediately prior to cell seeding, we aspirated the medium from the wells

455 without removing it from the compartments.

456

457 **Experimental animals**

458 Crl:CD1 (ICR) Swiss outbred mice (Charles River) were used. The animals were kept at 20-22 °C,

459 30-70 % humidity in a 12-hour/12-hour light/dark cycle and were fed a standard chow diet

460 (Altromin Spezialfutter GmbH) *ad libitum*. Animal experiments followed the protocol of the "NIH

461 Guide for the care and use of laboratory animals" and were approved by the Schleswig-Holstein

462 Ministry for Energy Transition, Agriculture, Environment, Nature and Digitalization (under the

463 prospective contingent animal license number 2017-07-06 Zille).

464

465 **Isolation and culture of primary cortical neurons**

466 We isolated primary cortical neurons from murine E14 embryos after decapitation as previously

467 described (Zille et al., 2017). We seeded the neurons at a density of 10,000 cells/mm^2 in 5 µL

468 MEM+Glutamax medium into one compartment (soma compartment) of each microfluidic unit of

469 the device. The cells were allowed to adhere at 37 °C for 30 minutes. In order to promote

470 directional axon growth into the other compartment (axonal compartment) by medium microflux,

22

471   150 µL of MEM+Glutamax medium were applied to the well of the soma compartment, while

472   100 µL were added to the well of the axonal compartment (**Fig. 1B**). Neurons were cultured at

473   37 °C in a humidified 5 % CO_2 atmosphere. The next day, we changed from MEM+Glutamax

474   medium to Neurobasal Plus Medium containing 2 % B-27 Plus Supplement, 1 mM sodium

475   pyruvate and 1 % penicillin/streptomycin. The volume differences among the wells ensured the

476   microflux for the directional axonal growth over the following days.

477

478   **Immunofluorescence**

479   Soma and axonal compartments in the microfluidic units were fixed at room temperature for 1 hour

480   in 4 % formaldehyde solution in phosphate buffered saline (PBS). They were washed twice with

481   PBS and permeabilized with blocking solution (2 % BSA, 0.5% Triton-X-100 and 1x PBS) at

482   room temperature for 1 hour. We incubated the neurons/axons on both compartments with primary

483   antibodies against synaptophysin (1:250) and MAP2 (1:4000) at 4 °C overnight. The next day,

484   both compartments were washed three times with PBS and incubated with the secondary

485   antibodies anti-mouse Alexa Fluor 546 (1:500) and anti-rabbit Alexa Fluor 488 (1:500) at room

486   temperature for 1 hour. After washing three times with PBS, both compartments were incubated

487   with DAPI (1 µg/mL) for nuclear counterstaining at room temperature for 10 minutes. Both

488   compartments were washed three times with PBS prior to fluorescence microscopy. An Olympus

489   IX81 time-lapse microscope (Olympus Deutschland GmbH) with a 10X objective (0.3 NA Ph1)

490   and camera F-View soft Imaging system was used at room temperature. Images were acquired

491   with Cell[M] software (Olympus Deutschland GmbH) and further processed via ImageJ (see **Image**

492   **preprocessing**).

493

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

23

495   At six or seven days in culture, microfluidic devices were considered for recording if they met the

496   following inclusion criteria: i) axon growth through at least 80 % of all microgrooves and ii) axon

497   length of at least 150 µm from the end of the microgrooves. All included microfluidic units were

498   randomly assigned to the experimental conditions.

499

500   **Time-lapse recording of axonal degeneration**

501   Axons were treated with 0 (vehicle), 50, 100, and 200 µM hemin. For the treatment, the medium

502   was removed from the wells of the microfluidic units; hemin was diluted in the collected media

503   and added back to the respective wells. The media volume between the two wells was equalized

504   during the treatment to prevent any microflux. All microfluidic units were recorded immediately

505   after each other. We started the recordings at 1 hour after treatment to allow for the adjustment of

506   the well plates to the humidity of the incubation chamber of the microscope and the setup of the

507   recording positions. We recorded AxD in Neurobasal Plus Medium containing 2 % B-27 Plus

508   Supplement, 1 mM sodium pyruvate and 1 % penicillin/streptomycin with a 30-minutes interval

509   for 24 hours using an Olympus IX81 time-lapse microscope (see **Immunofluorescence**) at 37 °C,

510   5 % CO_2 and 65 % humidity.

511

512   **Live cell fluorescent staining**

513   To evaluate axonal vitality, we washed the axonal compartment once with PBS and incubated the

514   axonal compartment with calcein AM (4 µM) in PBS for 30 minutes at 37 °C at the end of the

515   time-lapse recording or in 4-hour intervals upon hemin treatment. An Olympus IX81 time-lapse

516   microscope (see **Immunofluorescence**) was used to record the respective images at 37 °C, 5 %

517   CO_2 and 65 % humidity.

518

24

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

520 We trained the EntireAxon CNN for the image-wise semantic segmentation of AxD features in a

521 supervised manner (**Fig. 2B**). To this end, we adapted a standard u-net with ResNet-50 encoder

522 (Ronneberger et al., 2015) to automatically determine the class probability for each pixel of an

523 input image. Our segmentation aimed to classify each pixel of a microscopic image of a time-lapse

524 recording into one of four classes: 'background', 'axon', 'axonal swelling', and 'axonal fragment'.

525 For the training dataset, we selected 33 images and created corresponding image labels (masks)

526 using GIMP (v.2.10.14, RRID:SCR_003182). For each image, a label image with the same height

527 and width was created, in which each pixel value denotes a pixel class. Specifically, the classes

528 'background', 'axon', 'axonal swelling', and 'axonal fragment' had the values 0, 1, 2, and 3,

529 respectively. For each pixel of the input image, we retained 4 values that reflect the probability

530 distribution of the pixel over the four classes. We assigned each pixel the most probable class to

531 create a segmentation map. During training, the CNN observed an input image, produced an output

532 and compared this output to the label. The weights of the network were adapted via

533 backpropagation so that the output better fitted the label. The weight changes were derived from a

534 pixelwise loss function, i.e. the cross-entropy loss:

535
$$Loss(P, Y) = -\sum_{x,y,c} Y(x,y,c) \, log(P(x,y,c));$$

536 with $P(x, y, c) \; and \; Y(x, y, c)$ being the probability of class c at pixel (x, y) for the prediction and

537 ground truth of the network, respectively.

538

539 We trained a mean ensemble consisting of eight neural networks for 180 epochs using the Adam

540 optimizer, a batch size of four and a learning rate of 0.001 that decreased by a factor of ten after

541 every 60 epochs. The input images were standardized by the image-net mean and standard

25

542   deviation (Deng et al., 2009). For data augmentation, we used random cropping (size 512 x 512),

543   image flipping along the horizontal axis and rotation by a random angle between -90° and +90°.

544

545   **Validation of the EntireAxon CNN compared to human experts**

546   To measure how well the EntireAxon CNN segments unknown images (**Fig. 2C**), we used a second

547   validation set comprising eight images that were labeled by three human experts (A. Palumbo,

548   S.K.L., L.E.H.). Importantly, the EntireAxon CNN did not update its parameters during training

549   to fit the validation set, but only used the training set.

550   For each image, the EntireAxon CNN inferred a segmentation. We generated a binary mask from

551   the prediction of the network, where 1 denotes the respective class and 0 all other classes. We

552   computed a binary label mask in the same manner. We counted the true positive (TP), false positive

553   (FP), and false negative (FN) pixels and computed the recall (sensitivity) and precision (Forman

554   and Scholz, 2010):

555
$$Recall = \frac{TP}{TP + FN}$$

556
$$Precision = \frac{TP}{TP + FP}$$

557   Recall and precision were calculated for each class separately on each validation image. The mean

558   recall and precision over all eight validation images were determined subsequently.

559   A mean of 96.42 % of pixels in the axonal images were 'background' pixels, while only 2.77 %

560   represented the class 'axon', 0.58 % 'axonal swelling', and 0.23 % 'axonal fragment' pixels. This

561   reflects a challenging degree of class imbalance, where the probability of having any positives for

562   a class in a validation image is low. Thus, we did not use the computed recall and precision of the

563   individual images or the mean recall and precision to compute the mean F1 score, i.e., the harmonic

564   mean of recall and precision. This has been shown to lead to bias, especially when a high degree

26

565 of class imbalance is present in the dataset (Forman and Scholz, 2010) as it may result in undefined

566 values for an image for recall (due to the absence of TP), precision (in case the CNN does not

567 recognize the few positives), and F1 score (in case either recall or precision are undefined). To

568 avoid bias, we computed the total TP, FP, and FN of all validation images from which we

569 calculated the mean F1 score (Forman and Scholz, 2010):

570
$$mean\ F1\ score = \frac{2\ *\ TP_{total}}{2\ *\ TP_{total}\ +\ FP_{total}\ +\ FN_{total}}$$

571 In addition, we computed a consensus label between human expert 1 and 2, 1 and 3 as well as 2

572 and 3 and compared the EntireAxon CNN versus the remaining expert (human expert 3, 2, and 1,

573 respectively) to the consensus labels. Mean F1 scores for all classes were computed as described

574 above.

575

576 **Image preprocessing**

577 Prior to the analysis of AxD after hemin exposure, we preprocessed the time-lapse recordings in

578 ImageJ (v1.52a, RRID: RRID:SCR_003070) using a custom-written macro. Specifically, each

579 individual recording was converted from a 16-bit into an 8-bit recording to make it compatible

580 with the ImageNet (8-bit) pre-trained ResNet-50. The recording was aligned automatically with

581 the ImageJ plug-in "Linear Stack Alignment with SIFT" as described previously (Lowe, 2004).

582 The following settings were used: initial Gaussian blur of 1.6 pixel, 3 steps per scale octave,

583 minimum image size of 64 pixel, maximum image size of 1024 pixel, feature descriptor size of 4,

584 8 feature descriptor orientation bins, closest/next closest ratio of 0.92, maximal alignment error of

585 25 pixel; inlier ratio of 0.05, expected transformation as rigid, "interpolate" and "show info"

586 checked. Black edges appearing on the recording after alignment were cropped.

587

588 **AxD analysis using the EntireAxon CNN**

589   All recordings of AxD after hemin exposure were automatically analyzed by the trained

590   EntireAxon CNN, which classified each pixel as one of the four different classes 'background',

591   'axon', 'axonal swelling', and 'axonal fragment. For each experimental condition (i.e. hemin

592   concentration), the sum percentage of all pixels per class on all images of that experimental day

593   were added at each time point ('Axon$_{t1.5-24h}$, Axonal swelling$_{t1.5-24h}$, Axonal fragment$_{t1.5-24h}$). To

594   determine the changes for the classes 'axon', 'axonal swelling', and 'axonal fragment' over time,

595   we calculated the sum percentage of pixels for all given time points (t_i with I = 1.5 to 24 hours) of

596   the corresponding class over the sum of the pixels of all three classes at baseline:

597    $normalized\;'class'\;area\;(t_i)$

598
$$= \frac{'Class'_{ti}}{Axon_{t1.5h} + Axonalswelling_{t1.5h} + Axonalfragments_{t1.5h}} * 100$$

599

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

601   We used the segmentation videos derived from the original microscopic images using the CNN to

602   identify four morphological patterns of AxD: granular, retraction, swelling, and transport

603   degeneration (**Fig 7A**). To reduce the dimensions of the input, the segmentation video was

604   converted into a series of normalized histograms (*H*), one for each (time) frame. Thus, the RNN

605   did not operate on the microscopic images directly, but rather on more efficient representations of

606   the data. To compute a histogram for a frame t_i, we compared the pixels of the frames t_i and t_{i+1}.

607   Each pixel was assigned into one of 16 classes that consisted of pairs $(c_1, c_2) \in \{0,1,2,3\}^2$ of the

608   four segmentation classes (i.e., four times four possible configurations, 16 class pairs). For

609   example, the class (background, axon) means that in frame t_i, the pixel was classified as

610   background, while in frame t_{i+1}, it was an axon pixel. For T time steps, we therefore computed T-

611   1 histograms. $H_0(t_i, (c_1, c_2))$ is the number of pixels that belong to class c_1 at time-frame t_i and

612   that belong to c_2 at time-frame t_{i+1}. Additionally, we normalized each histogram to sum up to 1

613   (i.e. we divided by the sum over all pairs):

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

615   Of note, the histograms were computed over small patches (height and width < 90 pixels) during

616   training and during inference on windows of size 32x32 pixels.

617   We used an encoder-decoder RNN with attention (Bahdanau et al., 2016). The encoder f_{enc}

618   consisted of a gated recurrent unit (GRU) that obtained the histogram time sequence H as input.

619   The encoder computed the hidden representation of the histograms:

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

621   For our purpose, we used an architecture that was able to base the decision for a degeneration class

622   on the previous class predictions. To this end, the output $\vec{o}_i$ was computed iteratively in C+1 steps

623   as a sum of the previous output and the output of the decoder f_{dec}:

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

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

626   C is the number of degeneration classes (4) and d is the hidden dimension (we used 256); $i =$

627    $1, I, C + 1$. σ is the sigmoid function. The decoder employed a GRU that depended on the context

628   vector $\vec{c}_i$ and the hidden state vector $\vec{s}_{i-1}$:

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

630   The entries of the initial hidden vector $\vec{s}_1$ were all zero. The context vector is a weighted sum of

631   the encoder representations. At each iteration, these weights can change, enabling the network to

632   focus on different time-steps. We assumed that a specific pattern of degeneration happened only

29

633 in a limited number of time frames that were fewer than the whole input video. The weights

634 depended on the current state of the decoder and the current output:

635
$$\vec{c}_i \; = \; V^T \vec{\alpha}_i; \; \vec{\alpha}_i \; \in \mathbb{R}^T$$

636
$$\vec{\alpha}_i = \; Softmax(\; W_{att}\, [\vec{s}_{i-1}, ReLU(W_{in}\vec{o}_{i-1})]\;); \; W_{att} \in \mathbb{R}^{T \, x \, 2d}, \; W_{in} \in \mathbb{R}^{d \, x \, C}$$

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

638 function:

639
$$y \; = \; \sigma(\vec{o}_{C+1}) \; \in \; [0,1]^C.$$

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

641 The EntireAxon RNN was trained with 162 images for 60 epochs using the lamb optimizer (You

642 et al., 2020) with a batch size of 128. We used a learning rate of 0.01 that was reduced by a factor

643 of ten every 15 epochs and an additional weight decay of 0.0001. The two GRUs (encoder and

644 decoder) contained three layers, and we used dropout with a p-value of 0.9. To increase the RNN

645 robustness against varying axon thickness, we also added eroded versions of the segmentation data

646 using a cross-shape as kernel with the sizes three, five, and seven. Accordingly, each image existed

647 six times in the dataset: three eroded versions and three unchanged copies, to keep a 50 % chance

648 of having the original image for training.

649

650 **RNN cluster analysis**

651 The unnormalized class output $\vec{o}_{C+1}$ was computed by the matrix-vector product $W_{out}\vec{z}_{C+1}$. Where

652 $\vec{z}_{C+1}$ was a 256-dimensional vector representation of the input sample, computed by the model.

653 For the classes to be linearly separable, the vector representations of each class needed to be close

654 to each other in the 256-dimensional space. To visualize the relationships of the specific samples,

655 we employed t-distributed stochastic neighborhood embedding (T-SNE) to compute a 2-

656 dimensional representation of the high-dimensional data.

657

Ten-fold cross-validation of the RNN

659 To validate the RNN, we used ten-fold cross-validation (Hastie et al., 2009). The dataset S was

660 divided into 10 subsets, ensuring that each subset included at least one sample of each class: $S =$

661 $\bigcup_{i=1}^{10} S_i$; $S_i \cap S_j = \emptyset, i \neq j$. We trained ten models for i=1,..,10 on $Train_i = S / S_i$ and test

662 them on $Test_i = S_i$. Subsequently, we combined and evaluated all test samples $Test =$

663 $\bigcup_{i=1}^{10} Test_i$. Mean recall, precision, and F1 score were determined as described above.

664

Analysis of morphological pattern of AxD using the EntireAxon RNN

666 All AxD segmentations after hemin exposure were automatically analyzed with the trained

667 EntireAxon RNN, which predicted the occurrence of the four morphological patterns of AxD in a

668 pixel-wise manner. Of note, a pixel can be predicted to belong to 0, 1 or multiple morphological

669 patterns. Only pixels previously identified as degenerated over time were considered by applying

670 a 'fragmentation mask' that included all no-background pixels that changed to either background

671 or fragment during the recording time.

672 For each experimental condition (i.e., hemin concentration), the percentage of the occurrence of

673 each morphological pattern was calculated as the sum of all pixels per morphological pattern on

674 all images of that experimental day divided by the 'fragmentation mask' as follows:

675 $'morphological\ pattern'[\%]$

676 $$= \frac{\sum pixel\ of\ morphological\ pattern_i}{\sum pixel\ no\ background \rightarrow background\ or\ fragment} * 100$$

677

31

Statistical analysis

Six biological replicates for each concentration were employed in each experiment to assess hemin-induced AxD. We did not perform a priori power analysis as this was an exploratory study. 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. Linear regressions were performed for AxD patterns. Data are represented as mean ± 95 % confidence interval (CI) except for the nonparametric data of the AUC for axonal fragments, 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 α = 0.0167 was considered statistically significant. K = 6 for the comparison of the linear regression slopes (comparison of the four AxD patterns against each other), thus α = 0.0083 was considered statistically significant. The detailed statistical analyses can be found in **Tables S1-3**. All statistical analyses were performed with IBM SPSS version 23 (RRID:SCR_002865), except linear regressions that were performed with GraphPad Prism version 8 (RRID:SCR_002798).

List of abbreviations

699

700 AxD   axonal degeneration

701 CNN   convolutional neural network

702 FP   false positive

703 FN   false negative

704 GRU   gated recurrent unit

705 PBS   phosphate buffered saline

706 PDMS   polydimethylsiloxane

707 PMMA   polymethyl methacrylate

708 RNN   recurrent neural network

709 TP   true positive

710

711 **Declarations**

712

713 **Acknowledgments**

714 This research was supported by a Fraunhofer MEF grant of the Fraunhofer Society (project number

715 600199) and by the Joachim Herz Stiftung (project number 850022) to M.Z. We would like to

716 thank Sebastian Kärst for his technical assistance regarding the choice of materials for the

717 prototype of the microfluidic device. We would like to thank Dennis Wendt for his help in

718 designing the blueprint of the prototype of the microfluidic device.

719

720 **Authors' contributions**

721 M.Z. designed the experiments. A. Palumbo, A. Pabst and M.Z. designed the device. A. Palumbo,

722 A. Pabst, S.P., R.S., C.K., and N.K. carried out the fabrication of the device. S.K.L. performed the

723 immunostaining of the somata and axons and analyzed the respective data. P.G. and A.M.M.

724 developed the deep learning tool, P.G., L.E.H., and L.B. developed the algorithms to retrieve the

725 output. A. Palumbo, S.K.L. and L.E.H. labeled the images for the deep learning training and

726 validation. A. Palumbo and C.F. performed the time-lapse recordings of AxD. A. Palumbo

727 conducted the live cell imaging, the determination of the morphological patterns of AxD and

728 analyzed the data for the respective experiments. A. Palumbo, P.G., A.M.M., J.B., and M.Z.

729 discussed and interpreted the data. M.Z. performed the statistical analysis. A. Palumbo, M.I. and

730 M.Z. performed the graphical artwork. A. Palumbo, P.G., and M.Z wrote the manuscript. All

731 authors discussed and commented on the final version of the manuscript.

732

733 **Ethics approval**

734 Animal experiments followed the protocol of the "NIH Guide for the care and use of laboratory

735 animals" and were approved by the Schleswig-Holstein Ministry for Energy Transition,

34

736   Agriculture, Environment, Nature and Digitalization (under the prospective contingent animal

737   license number 2017-07-06 Zille).

738

739   **Data availability statement**

740   All data needed to evaluate the conclusions in the paper are present in the paper and/or the

741   **Supporting information**. The time-lapse data and code are available upon reasonable request to

742   the corresponding authors. We plan to launch a website to enable other researchers to use the tool.

743

744   **Competing interests**

745   A. Palumbo, P.G., and M.Z. declare that they have filed a patent for the microfluidic device and

746   the EntireAxon deep learning algorithm to quantify axonal degeneration (European Patent Office,

747   file number: 20152016.0, in revision). All other authors declare that they have no competing

748   interests.

References

Bahdanau, D., K. Cho, and Y. Bengio. 2016. Neural Machine Translation by Jointly Learning to Align and Translate. *arXiv:1409.0473 [cs, stat]*.

Becker, T., and A. Madany. 2012. Morphology-based Features for Adaptive Mitosis Detection of In Vitro Stem Cell Tracking Data. *Methods Inf Med*. 51:449–456. doi:10.3414/ME11-02-0038.

Beirowski, B., R. Adalbert, D. Wagner, D.S. Grumme, K. Addicks, R.R. Ribchester, and M.P. Coleman. 2005. The progressive nature of Wallerian degeneration in wild-type and slow Wallerian degeneration (WldS) nerves. *BMC Neurosci*. 6:6. doi:10.1186/1471-2202-6-6.

Biffi, E. 2015. Microfluidic and Compartmentalized Platforms for Neurobiological Research.

Bishop, D.L., T. Misgeld, M.K. Walsh, W.-B. Gan, and J.W. Lichtman. 2004. Axon Branch Removal at Developing Synapses by Axosome Shedding. *Neuron*. 44:651–661. doi:10.1016/j.neuron.2004.10.026.

Cavanagh, J.B. 1979. The "dying back" process. A common denominator in many naturally occurring and toxic neuropathies. *Arch. Pathol. Lab. Med.* 103:659–664.

Chen, J., and R.F. Regan. 2004. Heme oxygenase-2 gene deletion increases astrocyte vulnerability to hemin. *Biochemical and Biophysical Research Communications*. 318:88–94. doi:10.1016/j.bbrc.2004.03.187.

Chen, X., X. Chen, Y. Chen, M. Xu, T. Yu, and J. Li. 2018. The Impact of Intracerebral Hemorrhage on the Progression of White Matter Hyperintensity. *Front. Hum. Neurosci*. 12:471. doi:10.3389/fnhum.2018.00471.

Coleman, M.P. 2005. Axon degeneration mechanisms: commonality amid diversity. *Nat Rev Neurosci*. 6:889–898. doi:10.1038/nrn1788.

Cui, Y., X. Jin, D.-J. Choi, J.Y. Choi, H.S. Kim, D.H. Hwang, and B.G. Kim. 2020. Axonal degeneration in an in vitro model of ischemic white matter injury. *Neurobiology of Disease*. 134:104672. doi:10.1016/j.nbd.2019.104672.

Deng, J., W. Dong, R. Socher, L.-J. Li, Kai Li, and Li Fei-Fei. 2009. ImageNet: A large-scale hierarchical image database. *In* 2009 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, Miami, FL. 248–255.

Dietterich, T.G. 2000. Ensemble Methods in Machine Learning. *In* Multiple Classifier Systems. Springer Berlin Heidelberg, Berlin, Heidelberg. 1–15.

Forman, G., and M. Scholz. 2010. Apples-to-apples in cross-validation studies: pitfalls in classifier performance measurement. *SIGKDD Explor. Newsl.* 12:49. doi:10.1145/1882471.1882479.

Grüning, P., Palumbo, A., Zille, M., Barth, E., and Madany Mamlouk, Amir. 2020. A task-dependent active learning method for axon segmentation with CNNs. *Proc AUTOMED*. 1. doi:10.18416/AUTOMED.2020.

786   Hastie, T., R. Tibshirani, and J. Friedman. 2009. The Elements of Statistical Learning. Springer
787   New York, New York, NY.

788   He, K., X. Zhang, S. Ren, and J. Sun. 2015. Deep Residual Learning for Image Recognition.
789   *arXiv:1512.03385 [cs]*.

790   Hemorrhagic Stroke Academia Industry (HEADS) Roundtable Participants. 2018. Basic and
791   Translational Research in Intracerebral Hemorrhage: Limitations, Priorities, and
792   Recommendations. *Stroke*. 49:1308–1314. doi:10.1161/STROKEAHA.117.019539.

793   Ho, S.-Y., C.-Y. Chao, H.-L. Huang, T.-W. Chiu, P. Charoenkwan, and E. Hwang. 2011.
794   NeurphologyJ: An automatic neuronal morphology quantification method and its
795   application in pharmacological discovery. *BMC Bioinformatics*. 12:230.
796   doi:10.1186/1471-2105-12-230.

797   Hosmane, S., I.H. Yang, A. Ruffin, N. Thakor, and A. Venkatesan. 2010. Circular
798   compartmentalized microfluidic platform: Study of axon–glia interactions. *Lab Chip*.
799   10:741. doi:10.1039/b918640a.

800   Huang, H.-K., C.-F. Chiu, C.-H. Kuo, Y.-C. Wu, N.N.Y. Chu, and P.-C. Chang. 2016. Mixture
801   of deep CNN-based ensemble model for image retrieval. *In* 2016 IEEE 5th Global
802   Conference on Consumer Electronics. IEEE, Kyoto, Japan. 1–2.

803   Kerschensteiner, M., M.E. Schwab, J.W. Lichtman, and T. Misgeld. 2005. In vivo imaging of
804   axonal degeneration and regeneration in the injured spinal cord. *Nat Med*. 11:572–577.
805   doi:10.1038/nm1229.

806   Li, Y., M. Yang, Z. Huang, X. Chen, M.T. Maloney, L. Zhu, J. Liu, Y. Yang, S. Du, X. Jiang,
807   and J.Y. Wu. 2014. AxonQuant: A Microfluidic Chamber Culture-Coupled Algorithm
808   That Allows High-Throughput Quantification of Axonal Damage. *Neurosignals*. 22:14–
809   29. doi:10.1159/000358092.

810   Lingor, P., J.C. Koch, L. Tönges, and M. Bähr. 2012. Axonal degeneration as a therapeutic target
811   in the CNS. *Cell Tissue Res*. 349:289–311. doi:10.1007/s00441-012-1362-3.

812   Lowe, D.G. 2004. Distinctive Image Features from Scale-Invariant Keypoints. *International*
813   *Journal of Computer Vision*. 60:91–110. doi:10.1023/B:VISI.0000029664.99615.94.

814   Luo, L., and D.D.M. O'Leary. 2005. Axon Retraction and Degeneration in Development and
815   Disease. *Annu. Rev. Neurosci.* 28:127–156.
816   doi:10.1146/annurev.neuro.28.061604.135632.

817   Maor-Nof, M., E. Romi, H. Sar Shalom, V. Ulisse, C. Raanan, A. Nof, D. Leshkowitz, R. Lang,
818   and A. Yaron. 2016. Axonal Degeneration Is Regulated by a Transcriptional Program
819   that Coordinates Expression of Pro- and Anti-degenerative Factors. *Neuron*. 92:991–
820   1006. doi:10.1016/j.neuron.2016.10.061.

821   Neukomm, L.J., and M.R. Freeman. 2014. Diverse cellular and molecular modes of axon
822   degeneration. *Trends in Cell Biology*. 24:515–523. doi:10.1016/j.tcb.2014.04.003.

823   Nikić, I., D. Merkler, C. Sorbara, M. Brinkoetter, M. Kreutzfeldt, F.M. Bareyre, W. Brück, D.
824   Bishop, T. Misgeld, and M. Kerschensteiner. 2011. A reversible form of axon damage in

825   experimental autoimmune encephalomyelitis and multiple sclerosis. *Nat Med*. 17:495–
826   499. doi:10.1038/nm.2324.

827 Nikolaev, A., T. McLaughlin, D.D.M. O'Leary, and M. Tessier-Lavigne. 2009. APP binds DR6
828   to trigger axon pruning and neuron death via distinct caspases. *Nature*. 457:981–989.
829   doi:10.1038/nature07767.

830 Park, J.W., B. Vahidi, A.M. Taylor, S.W. Rhee, and N.L. Jeon. 2006. Microfluidic culture
831   platform for neuroscience research. *Nat Protoc*. 1:2128–2136.
832   doi:10.1038/nprot.2006.316.

833 Pease, S.E., and R.A. Segal. 2014. Preserve and protect: maintaining axons within functional
834   circuits. *Trends in Neurosciences*. 37:572–582. doi:10.1016/j.tins.2014.07.007.

835 Pease-Raissi, S.E., M.F. Pazyra-Murphy, Y. Li, F. Wachter, Y. Fukuda, S.J. Fenstermacher, L.A.
836   Barclay, G.H. Bird, L.D. Walensky, and R.A. Segal. 2017. Paclitaxel Reduces Axonal
837   Bclw to Initiate IP3R1-Dependent Axon Degeneration. *Neuron*. 96:373-386.e6.
838   doi:10.1016/j.neuron.2017.09.034.

839 Pool, M., J. Thiemann, A. Bar-Or, and A.E. Fournier. 2008. NeuriteTracer: A novel ImageJ
840   plugin for automated quantification of neurite outgrowth. *Journal of Neuroscience
841   Methods*. 168:134–139. doi:10.1016/j.jneumeth.2007.08.029.

842 Robinson, S.R., T.N. Dang, R. Dringen, and G.M. Bishop. 2009. Hemin toxicity: a preventable
843   source of brain damage following hemorrhagic stroke. *Redox Report*. 14:228–235.
844   doi:10.1179/135100009X12525712409931.

845 Ronneberger, O., P. Fischer, and T. Brox. 2015. U-Net: Convolutional Networks for Biomedical
846   Image Segmentation. *arXiv:1505.04597 [cs]*.

847 Salvadores, N., M. Sanhueza, P. Manque, and F.A. Court. 2017. Axonal Degeneration during
848   Aging and Its Functional Role in Neurodegenerative Disorders. *Front. Neurosci.* 11:451.
849   doi:10.3389/fnins.2017.00451.

850 Sasaki, Y., B.P.S. Vohra, F.E. Lund, and J. Milbrandt. 2009. Nicotinamide Mononucleotide
851   Adenylyl Transferase-Mediated Axonal Protection Requires Enzymatic Activity But Not
852   Increased Levels of Neuronal Nicotinamide Adenine Dinucleotide. *Journal of
853   Neuroscience*. 29:5525–5535. doi:10.1523/JNEUROSCI.5469-08.2009.

854 Saxena, S., and P. Caroni. 2007. Mechanisms of axon degeneration: From development to
855   disease. *Progress in Neurobiology*. 83:174–191. doi:10.1016/j.pneurobio.2007.07.007.

856 Tang-Schomer, M.D., V.E. Johnson, P.W. Baas, W. Stewart, and D.H. Smith. 2012. Partial
857   interruption of axonal transport due to microtubule breakage accounts for the formation
858   of periodic varicosities after traumatic axonal injury. *Experimental Neurology*. 233:364–
859   372. doi:10.1016/j.expneurol.2011.10.030.

860 Tao, C., X. Hu, H. Li, and C. You. 2017. White Matter Injury after Intracerebral Hemorrhage:
861   Pathophysiology and Therapeutic Strategies. *Front. Hum. Neurosci.* 11:422.
862   doi:10.3389/fnhum.2017.00422.

863   Van Laar, V., B. Arnold, and S. Berman. 2019. Primary Embryonic Rat Cortical Neuronal
864   Culture and Chronic Rotenone Treatment in Microfluidic Culture Devices. *BIO-*
865   *PROTOCOL*. 9. doi:10.21769/BioProtoc.3192.

866   Venkatasubramanian, C., J.T. Kleinman, N.J. Fischbein, J. Olivot, A.D. Gean, I. Eyngorn, R.W.
867   Snider, M. Mlynash, and C.A.C. Wijman. 2013. Natural History and Prognostic Value of
868   Corticospinal Tract Wallerian Degeneration in Intracerebral Hemorrhage. *JAHA*. 2.
869   doi:10.1161/JAHA.113.000090.

870   Vuola, A.O., S.U. Akram, and J. Kannala. 2019. Mask-RCNN and U-Net Ensembled for Nuclei
871   Segmentation. *In* 2019 IEEE 16th International Symposium on Biomedical Imaging
872   (ISBI 2019). IEEE, Venice, Italy. 208–212.

873   Wang, J.T., Z.A. Medress, and B.A. Barres. 2012. Axon degeneration: Molecular mechanisms of
874   a self-destruction pathway. *J Cell Biol*. 196:7–18. doi:10.1083/jcb.201108111.

875   Yong, Y., K. Gamage, I. Cheng, K. Barford, A. Spano, B. Winckler, and C. Deppmann. 2019.
876   p75NTR and DR6 Regulate Distinct Phases of Axon Degeneration Demarcated by
877   Spheroid Rupture. *J. Neurosci.* 39:9503–9520. doi:10.1523/JNEUROSCI.1867-19.2019.

878   You, Y., J. Li, S. Reddi, J. Hseu, S. Kumar, S. Bhojanapalli, X. Song, J. Demmel, K. Keutzer,
879   and C.-J. Hsieh. 2020. Large Batch Optimization for Deep Learning: Training BERT in
880   76 minutes. *arXiv:1904.00962 [cs, stat]*.

881   Zille, M., S.S. Karuppagounder, Y. Chen, P.J. Gough, J. Bertin, J. Finger, T.A. Milner, E.A.
882   Jonas, and R.R. Ratan. 2017. Neuronal Death After Hemorrhagic Stroke In Vitro and In
883   Vivo Shares Features of Ferroptosis and Necroptosis. *Stroke*. 48:1033–1043.
884   doi:10.1161/STROKEAHA.116.015609.
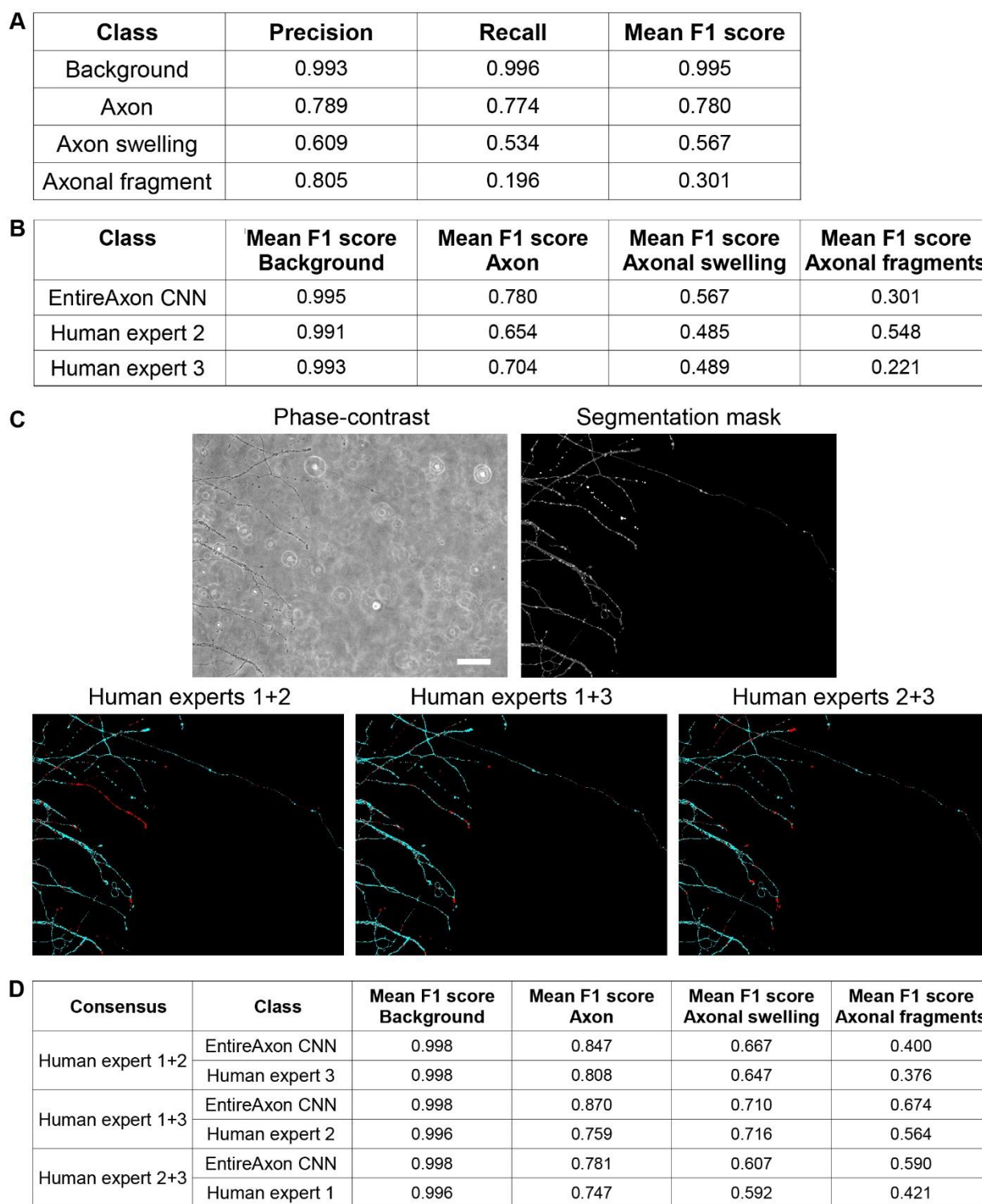
885

886 **Figure legends**
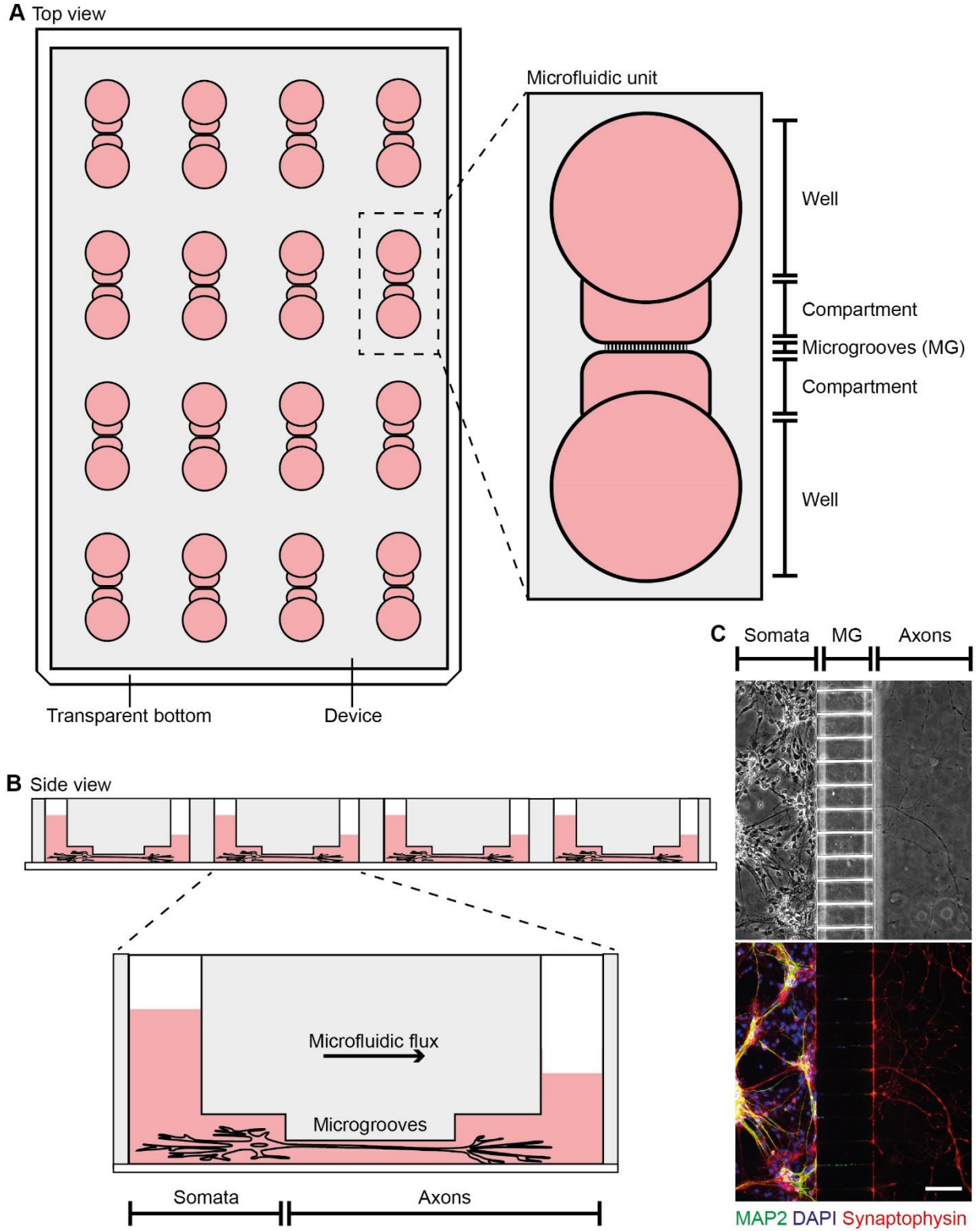
887

888 **Figure 1. Microfluidic device for the enhanced throughput cultivation of axons. (A)** The microfluidic device

889 incorporates 16 individual microfluidic units for axon cultivation. One microfluidic unit consists of two wells that are

890 connected through compartments and microgrooves (MG). **(B)** Primary cortical neurons are seeded into the soma

40

891   compartment from which their axons grow through the MG into the axon compartment. Directed growth is supported

892   by culture medium microflux due to different medium volumes between the two wells. **(C)** Phase-contrast image of

893   primary cortical axons that were spatially separated from their somata by the MG at day *in vitro* 7, which we confirmed

894   by immunofluorescence staining of dendrites using microtubule-associated protein 2 (MAP2, green, 1:4000) and

895   axons using synaptophysin (red, 1:250). DAPI (blue, 1:1000)) was used for nuclear counterstaining (top). Scale

896   bar: 100 µm.

897
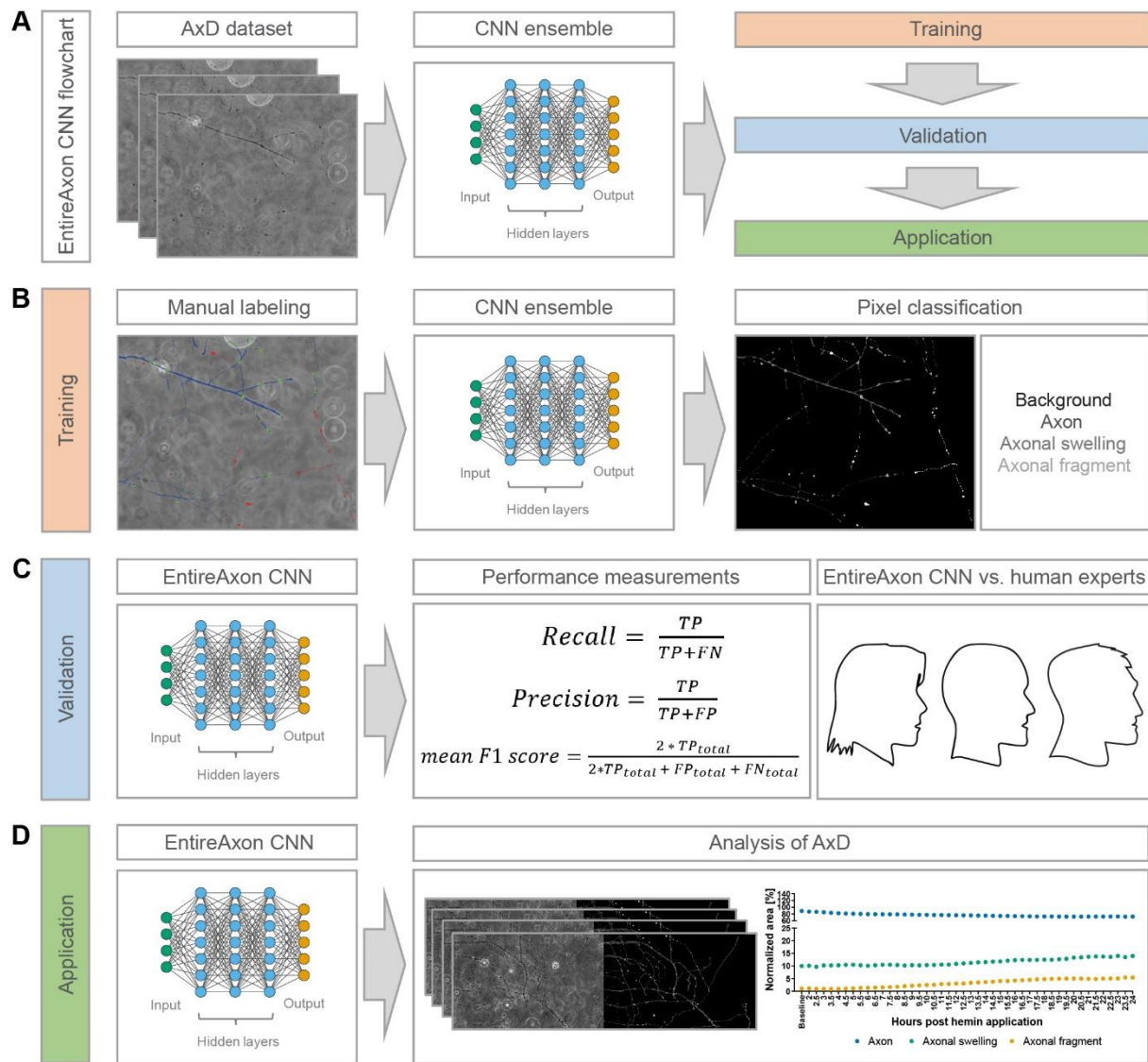
Figure 2. EntireAxon CNN for the enhanced throughput analysis of AxD. (A) The flow chart of the EntireAxon CNN. The AxD data was separated into training, validation, and testing data. We adapted a standard u-net with ResNet-50 encoder (Ronneberger et al., 2015; He et al., 2015) and used a CNN ensemble, which combines predictions from multiple CNNs to generate a final output and is superior to individual CNNs (Dietterich, 2000; Huang et al., 2016; Vuola et al., 2019). **(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 (ground truth was labeled by human expert 1). **(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*.

42

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

Phase-contrast   Segmentation mask

Human experts 1+2   Human experts 1+3   Human experts 2+3

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

909

Figure 3. Performance of the EntireAxon CNN compared to human experts. (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

43

915   mask, and the consensus labeling masks of two human experts that show the segmentation overlap (cyan) or difference

916   (red) between the labels. Scale bar: 100 µm. **(D)** Comparison of the mean F1 scores between the EntireAxon CNN

917   and the human expert on the consensus labeling of the other two human experts.
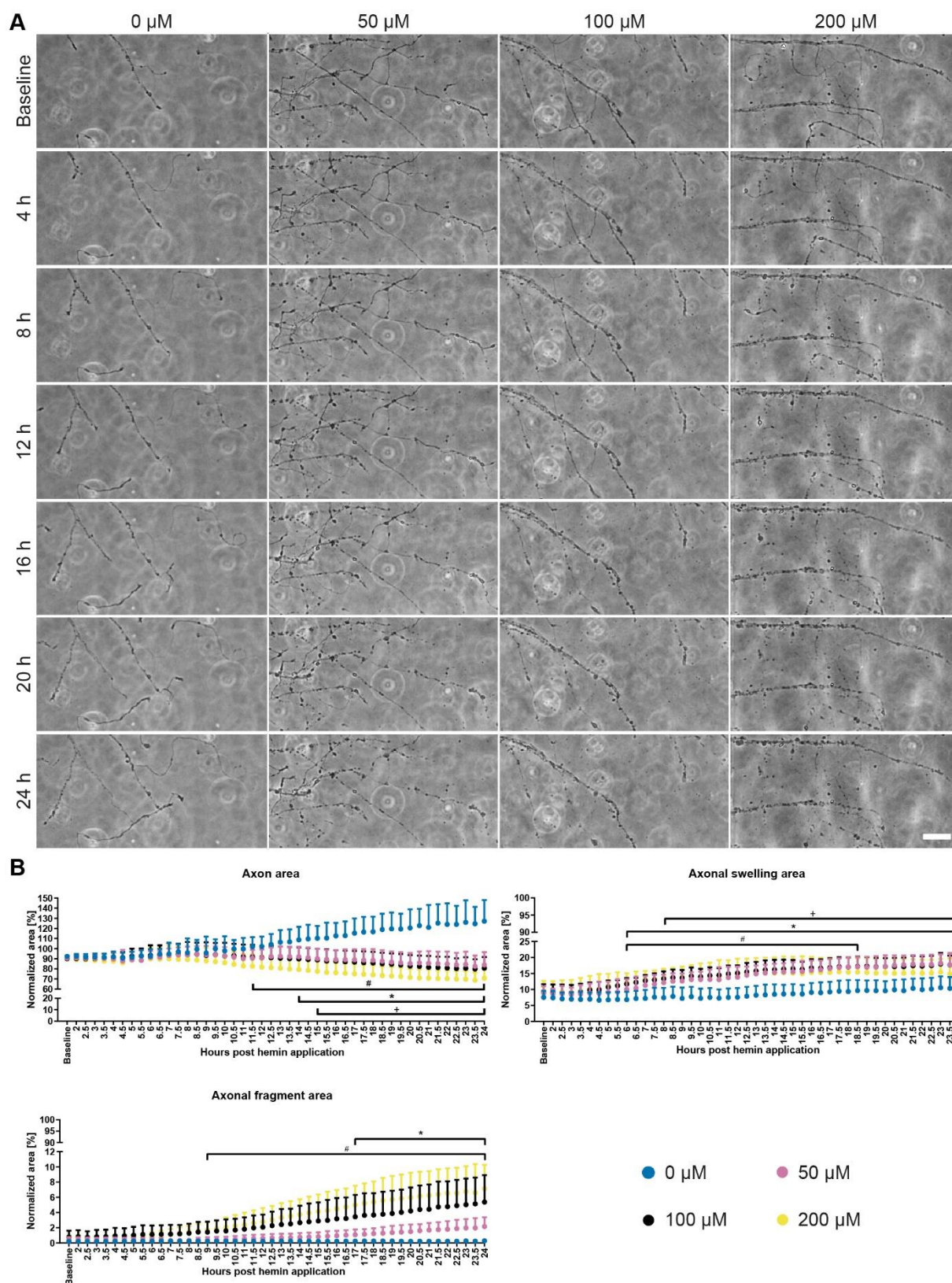
Figure 4. Time- and concentration-dependent hemin-induced AxD. (A) Primary cortical axons treated with hemin

(50, 100, 200 µM) degenerated compared to vehicle-treated axons (0 µM) that continued to grow. Scale bar: 50 µm.

For complete time-lapse videos including segmentation, refer to **Video S1-4**. **(B)** Quantification of AxD over 24 hours

922   in phase-contrast images. To determine the time course, the sum of pixels in each class and hemin concentration over

923   time was normalized to the baseline of that class and condition. The quantification of the phase-contrast images over

924   24 hours revealed significantly smaller axon areas starting at 11.5 hours after 200 µM ($P = 0.020$), at 14 hours after

925   100 µM ($P = 0.040$), and at 15 hours after 50 µM ($P = 0.018$) hemin treatment compared to control (0 µM). The axonal

926   fragment area significantly increased from 9.5 hours onwards in 200 µM hemin ($P = 0.037$) and from 17.5 hours in

927   100 µM hemin ($P = 0.044$), while the axonal swelling area increased from 6 hours onwards in 100 µM hemin

928   ($P = 0.019$) and 200 µM hemin ($P = 0.010$) and from 8 hours in 50 µM hemin ($P = 0.030$). N = 6 independent cultures

929   of primary cortical neurons. Means + 95 % CI are given. One-way ANOVA with Greenhouse-Geisser correction. +,

930   *, # $P < 0.05$; + = 50 µM vs. 0 µM, * = 100 µM vs. 0 µM, # = 200 µM vs. 0 µM. For detailed statistical information,
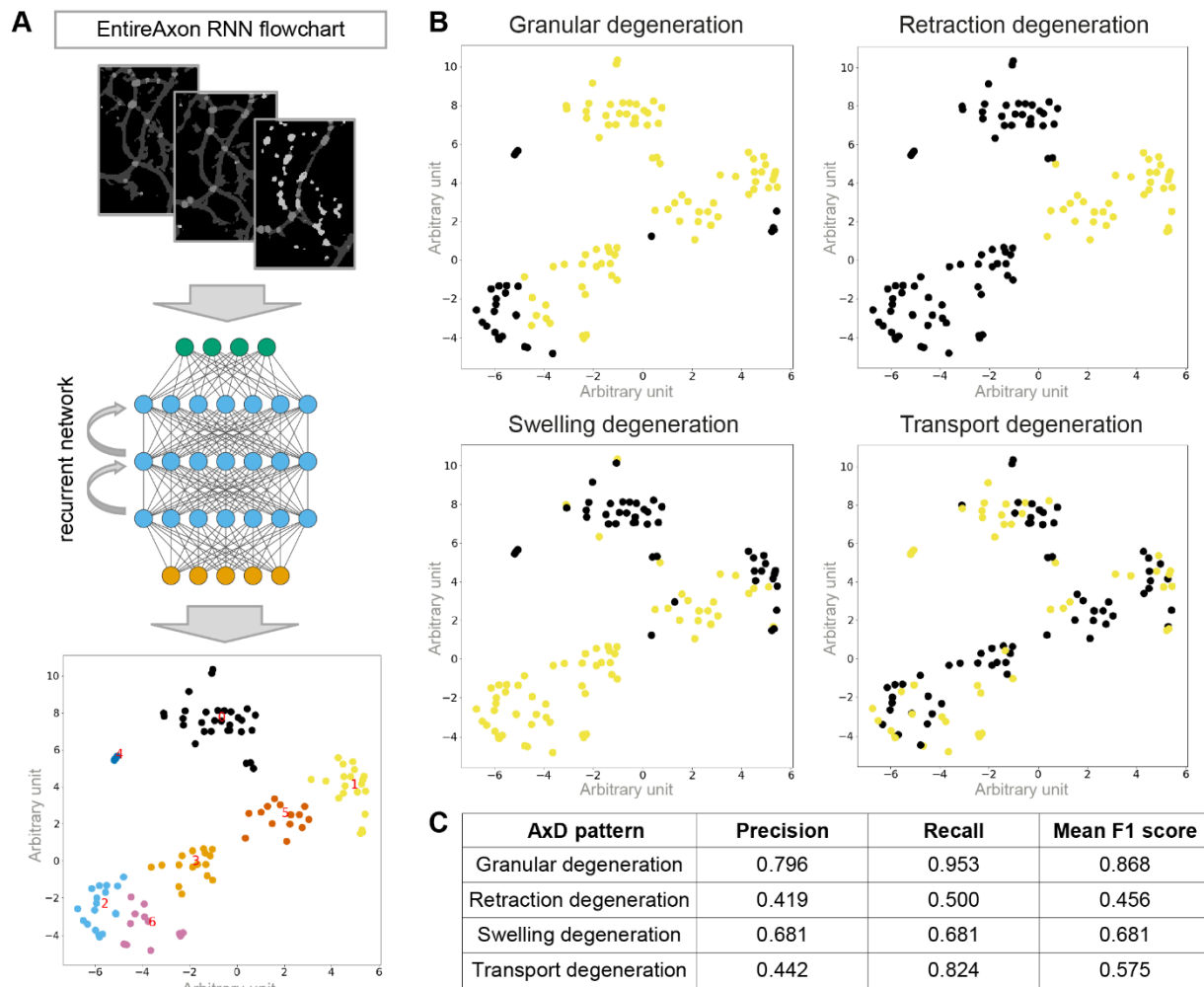
931   refer to **Table S1**.

46
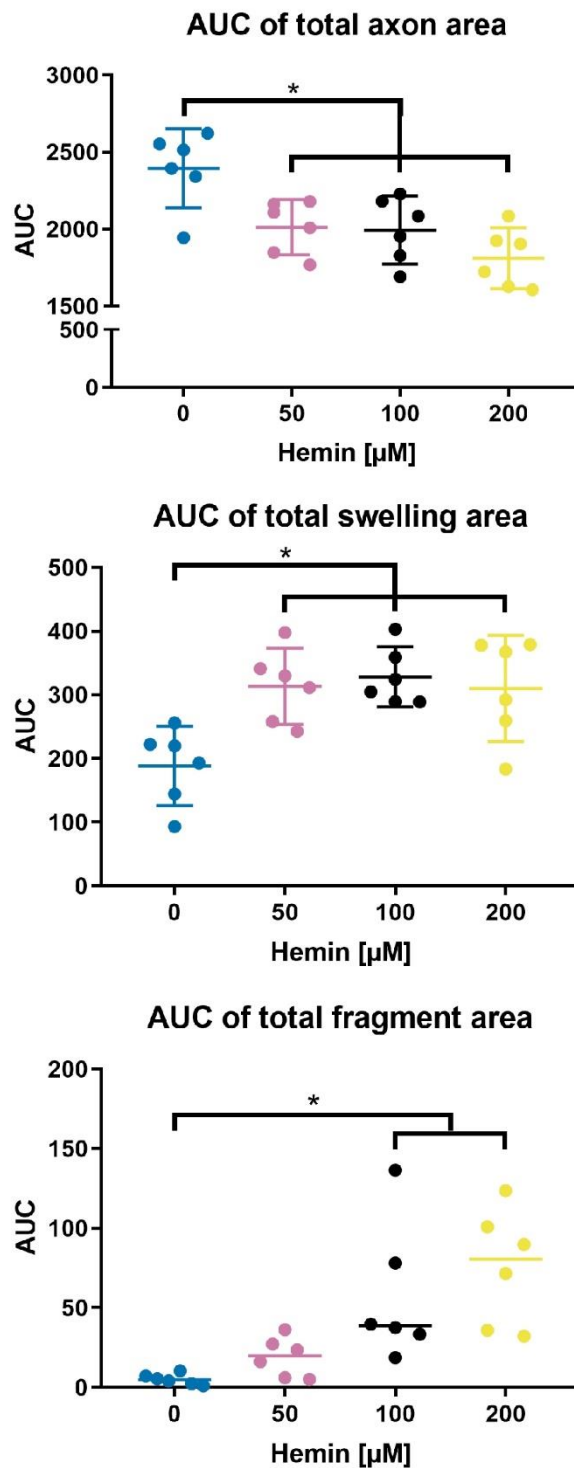
Figure 5. Area under the curve (AUC) analysis of hemin-induced AxD. While axons exposed to hemin showed a decline in axon area, axonal swelling and axonal fragment area increased. N = 6 independent cultures of primary cortical neurons. Means ± 95 % CI are given for axon and axonal swelling area, medians for fragment area. * $P < 0.05$ vs. 0 µM for axon and swelling area, * $P < 0.0167$ for fragment area due to manual Bonferroni correction for nonparametric data. For exact p values, refer to **Table S2**.

938
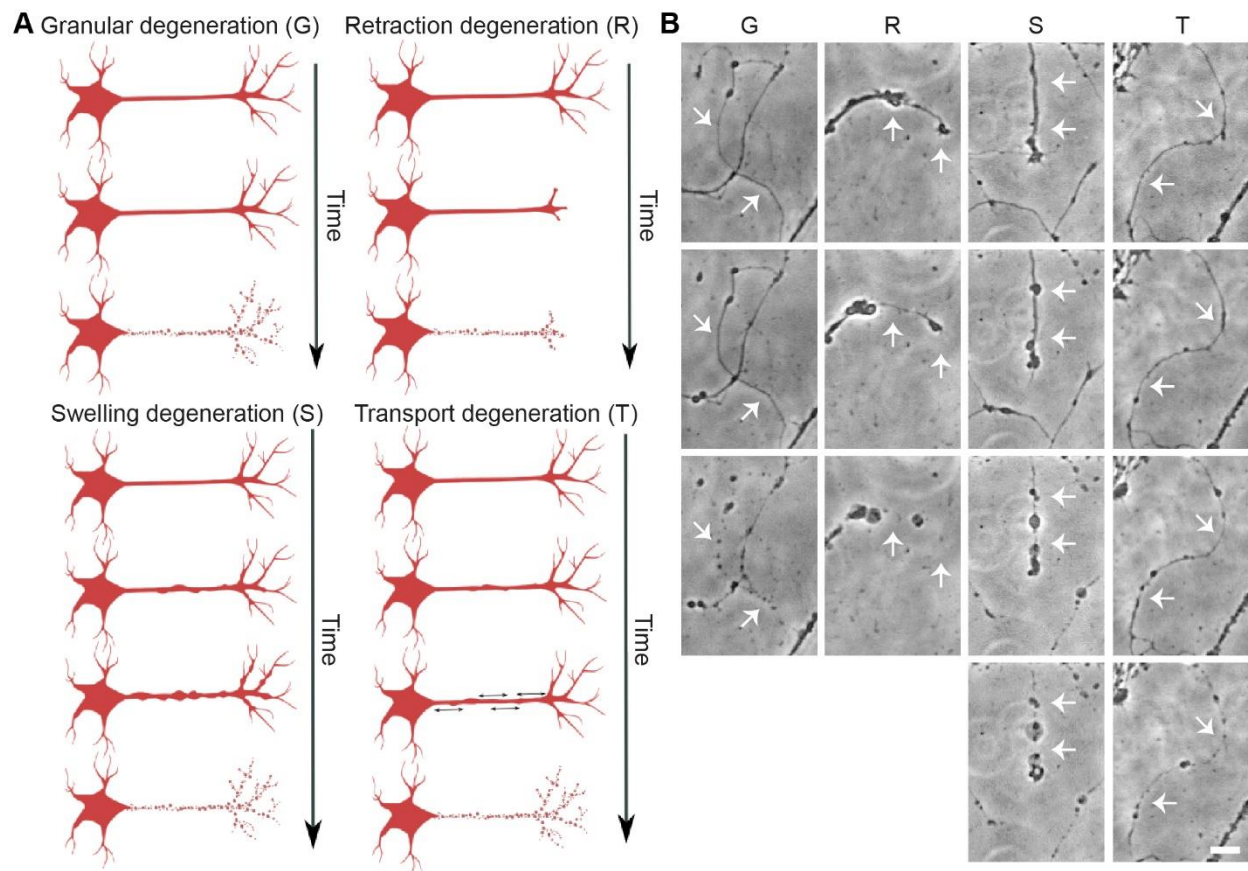
Figure 6. Four morphological patterns of AxD. **(A)** Schematic overview of the proposed AxD morphological patterns: granular degeneration, retraction degeneration, swelling degeneration, and transport degeneration. **(B)** Phase-contrast recordings of the four morphological patterns of AxD in primary cortical axons. Granular degeneration (G) is characterized by the fragmentation of the axon (white arrows). During retraction degeneration (R), 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 (S), many axonal swellings enlarge resulting in axonal fragments (white arrows). During transport degeneration (T), axonal swellings are transported along the axon prior to the degeneration of the axon (white arrows). Scale bar: 20 µm. For complete time-lapse videos including segmentation, refer to **Videos S5-8.**

AxD pattern	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

948

949 **Figure 7. Recognition of four morphological patterns of AxD by the EntireAxon RNN**. **(A)** Schematic workflow

950 of the RNN to recognize and quantify morphological patterns of AxD based on the identification of seven clusters.

951 The EntireAxon CNN segmentation masks were used for the RNN training, which determined the change in class

952 over time. Based on the 16 different possible class pairs, the RNN determined seven clusters (cluster 0-6). To visualize

953 the relationships of the specific samples, we employed t-distributed stochastic neighborhood embedding (T-SNE) to

954 compute a 2-dimensional representation of the high-dimensional data. **(B)** The clusters classify the four morphological

955 patterns of AxD with yellow indicating included and purple indicating excluded clusters: granular (G), retraction (R),

956 swelling (S), and transport degeneration (T). Clusters of granular degeneration overlap with recognized clusters of

957 other morphological patterns (retraction, swelling, and transport degeneration). For more details on the morphological

958 changes underlying the cluster analysis, refer to **Supplementary Fig. S3**. **(C)** 10-fold cross-validation of the four

959 morphological patterns of AxD.
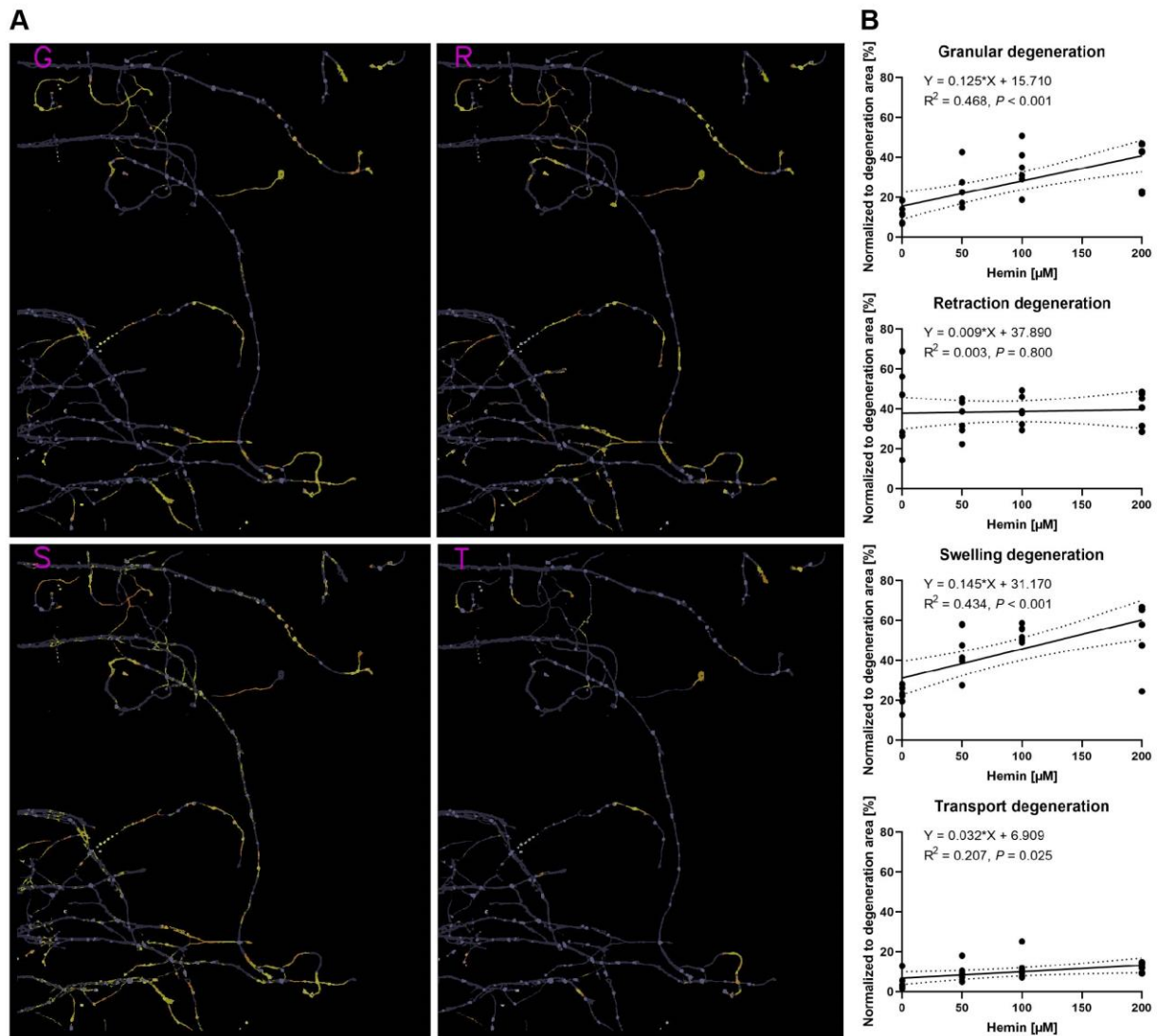
Figure 8. Concentration-dependent differences in the morphological patterns of hemin-induced AxD. (A) The classification of granular (G), retraction (R), swelling (S), and transport degeneration (T) in primary cortical axons treated with 200 µM hemin. For the complete time-lapse video including segmentation, refer to **Video S9. (B)** Linear regressions of the four morphological patterns of AxD in hemin-induced AxD. The area classified for each AxD pattern was normalized to the total degeneration area. Dotted lines show 95 % confidence bands. N = 6 independent cultures of primary cortical neurons. Granular degeneration: $F_{(1,22)} = 19.330$, $P < 0.001$. Retraction degeneration: $F_{(1,22)} = 0.066$, $P = 0.800$. Swelling degeneration: $F_{(1,22)} = 16.900$, $P < 0.001$. Transport degeneration: $F_{(1,22)} = 5.757$, $P = 0.025$. For the comparison of the slopes between the different AxD patterns, refer to **Table S3**.