Atlas-based imaging data analysis tool for quantitative mouse brain histology (AIDAhisto)

Niklas Pallast[a], Frederique Wieters[a], Gereon R. Fink[a,b], Markus Aswendt[a]

[a] *Department of Neurology, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany*
[b] *Cognitive Neuroscience, Institute of Neuroscience and Medicine (INM-3), Research Center Juelich, Juelich, Germany*

Journal of Neuroscience Methods
Published 12 August 2019
https://doi.org/10.1016/j.jneumeth.2019.108394

Abstract

Cell counting in neuroscience is a routine method of utmost importance to support descriptive in vivo findings with quantitative data on the cellular level. Although known to be error- and bias-prone, manual cell counting of histological stained brain slices remains the gold standard in the field. While the manual approach is limited to small regions-of-interest in the brain, automated tools are needed to up-scale translational approaches and generate whole mouse brain counts in an atlas framework. Our goal was to develop an algorithm which requires no pre-training such as machine learning algorithms, only minimal user input, and adjustable variables to obtain reliable cell counting results for stitched mouse brain slices registered to a common atlas such as the Allen Mouse Brain atlas. We adapted filter banks to extract the maxima from round-shaped cell nuclei and various cell structures. In a qualitative as well as quantitative comparison to other tools and two expert raters, AIDAhisto provides accurate and fast results for cell nuclei as well as immunohistochemical stainings of various types of cells in the mouse brain.

1 Introduction

Understanding brain diseases on a cellular level requires a detailed mapping of histopathology to brain function. Fundamental to this correlation is the analysis of multiple cell types and cell markers by immunohistochemistry ("immunostainings"). Engineered fluorescence-coupled antibodies to detect specific proteins are still the gold standard to visualize and quantify processes such as brain inflammation, degeneration, neurogenesis, and axonal remodeling. Immunostainings are performed using ultra-thin (10-50 μm) brain tissue slices and imaged with fluorescence

microscopes (5). In order to acquire a whole brain coverage, most commercial microscopes to-date are equipped with a motorized stage for x,y,z-movements as well as software and stitching algorithms ((6)). Other systems are available for automated mouse brain processing such as serial two-photon tomography (STPT) ((24)) or brain clearing and whole brain lightsheet microscopy ((10)). A key element of that approach is the quantitative cellular analysis within a reference atlas to increase the comparability and transferability of data obtained from different studies.
Despite the recent progress in digital atlasing ((14)) and related automated cell counting tools ((17)), most studies still rely on the analysis of a minimal number of brain slices where a small, manually-defined brain region is imaged and cell counting is performed in error-prone and tedious manual interventions. Both steps, the registration of the experimental tissue sample with a standardized tissue atlas and the quantification of tissue markers are non-trivial and prone to user-errors ((11)). Thus, for correlations and for revealing causal relationships between a pathology in a particular brain region and a measured behaviour, the manual approach is inappropriate. Although software and basic image processing algorithms are available to solve mathematically the cell detection and atlas registration (see Supplement Table 1), there are several practical challenges: The registration of whole brain microscopy data with a 2D segmented mouse brain atlas should account for the deformations induced by tissue preparation and (in part) image acquisition. Secondly, cell counting on histological sections, where a certain protein is targeted with fluorophore-coupled antibodies (immunostaining) requires an algorithm robust enough to differentiate background (usually from unspecific antibody binding) and real signal. Simple threshold-based approaches to identify local maxima, as it would be sufficient for cell nuclei, will not work in such a scenario. Furthermore, conventional linear and non-linear atlas registration approaches cannot merge such diverse images with a simple metric in terms of information content, size, and shape without losing mapping precision. So far, other tools have been optimized only for healthy mouse brain tissue but not for disease models, such as stroke, in which the analysis is challenged by additional tissue changes due to atrophy, swelling and an increased autofluorescence due to dead cells ((12)). Consequently, an efficient, unbiased, scalable, and automated analysis, which can be applied with minimal user input, experience in programming, and image post-processing, is mandatory. Here, we present an atlas-based imaging data analysis tool for mouse brain histology (AIDAhisto) that features a simple protocol for landmark-correspondence registration of selected whole-brain slice microscopy with the Allen Mouse Bran Atlas, ARA, ((18) and (22)) and a two-step cell detection algorithm using the cell nuclei segmentation as a seed for cell counting. We developed AIDAhisto as an open-source script for Python and Matlab and evaluated the cell detection accuracy for cell nuclei and immunostainings of astrocytes and immune cells against other established tools.

2 Material & Methods

2.1 Experimental setup

To preserve brain tissue for imaging and subsequent analysis, adult C57/BL6 mice (The Jackson Laboratory) were intracardially perfused with 20 mL of phosphate buffered saline

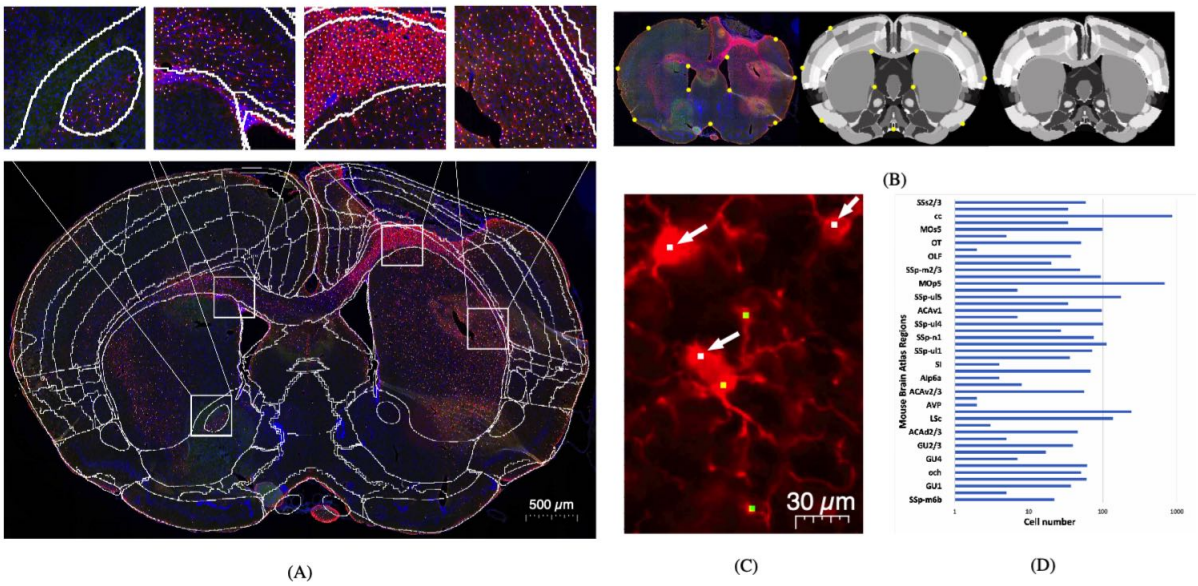


Figure 1: AIDAhisto enables an atlas-based quantitative analysis of immunohistochemistry on whole brain tissue slices. (A) Representative microscopy of Iba1 (red), autofluorescence (green) and DAPI (blue) with registered ARA plate (white lines). Inserts show accuracy of registration for selected white matter areas. Segmented cells are shown as white dots. (B) The atlas slice which best corresponds to the experimental section was registered using a point-based registration. The corresponding points between the input image and the atlas are marked by yellow dots. (C) Representative Iba1 staining (red) for which the cell nuclei validation was applied: Iba1 cell positions whose cell nucleus position could not be assigned (yellow dots without arrows) were deleted, whereas Iba1 positions with cell nuclei persist as valid results (white dots with arrows). (D) As a result, the cell number is expressed as cells counted per brain regions (displayed as an example as bar graph).

(PBS, Merck) followed by 20 ml of 4% phosphate buffered formaldehyde (PFA, Roti-Histofix, Carl Roth, #P087.3). After perfusion, the spinal cords and brains were isolated and transferred for 24 hours (h) to a post-fixative solution containing 4% PFA at 4°C. For longterm storage the brains were stored in 30% Sucrose solution at 4°C. After post-fixation 20 μm coronal tissue sections were prepared using a microtome with freezing stage and the slices were stored at -20°C.

Fixed tissue was immunostained following a standard protocol. The frozen sections were thawed for 10 minutes (min) at room temperature (RT) und pre-treated with sodium citrate (10 mM in PBS) for 20 min at 80°C. Afterwards the sections were washed with PBS. The sections were incubated for 1 h at RT in blocking solution (BS; PBS + 0.25% Triton X-100 + 5% goat serum), followed by incubation with one of the following antibodies: rabbit anti-DCX (1:50; Santa Cruz #sc-8066), rabbit anti-GFAP (1:500; DAKO #Z0334), rabbit anti-Iba1 (1:200; R&D #AF2535), MAP2 (1:200 Cell Signaling

#4542), and mouse anti-NeuN (1:200; Merck #MAB377) at 4°C overnight (O/N). After washing with PBS, the sections were incubated with the secondary antibodies donkey anti rabbit-Cy3 (1:500; Jackson Immuno Research #711-165-152), donkey anti rabbit-Alexa Fluor 488 (1:500; Life Technologies #A21441) or donkey anti mouse-Cy5 (1:500; Jackson Immuno Research #715-175-151) diluted in BS for 2 h at RT. Cell nuclei were stained using DAPI for 5 min after the secondary antibodies were rinsed off. The sections were washed with PBS, covered with Fluoromount G (Thermo Fisher), dried O/N at RT, and stored at 4°C. Whole brain slice stitch/merge images were acquired with the fluorescence microscopes Keyence BC8000 (10x objective) or Life Tech Evos (20x objective).

For the quantitative evaluation, the prefrontal cortex of two sets of four mice was labeled with Iba1/DAPI$_{\mathrm{I}}$ and GFAP/DAPI$_{\mathrm{G}}$, respectively. For each of the 16 datasets, we randomly sub-dived whole-brain slices into four ROIs using ImageJ (ImageJ v1.52, U. S. National Institutes of Health, Bethesda, Maryland, USA) (((30))), which led to an overall dataset of N=64 samples. The samples have different sizes in x-y-direction and varying fluorescence illuminations. The ROIs were manually analyzed by blinded and independent experts. They tagged cell nuclei and cells for both stainings, Iba1/DAPI and GFAP/DAPI, using ImageJ by visual inspection of the unmodified fluorescence microscopy images. The cell positions were saved for comparison. Only cells that were tagged by both were defined as valid (true) cells in order to increase the requirements for automatic classification ((9)).

2.2 Atlas Registration & Automatic cell detection

A detailed step-by-step procedure including recommendations to install and run AIDAhisto can be found in the manual (Supplementary Material). The code, example microscopy files and the mouse brain atlas is available on Github. In our example, the Allen Mouse Brain Atlas (ARA), 2004 Allen Institute for Brain Science, available from API was used. Any other atlas or manually drawn region-of-interest (saved as binary image) will work as well. In the following example, a simple landmark-based registration in ImageJ was used to register the microscopy image with the atlas. Notably, the registration will work only for 2D-matching slices. Any additional angle, e.g. induced by improper brain cutting, needs to be compensated during the registration procedure. For such process, we recommend QuickNII and BigStitcher which support multi-angle adjustments.

Initially, the atlas plane that relates to the microscopic image was manually extracted from the coronal ARA spaced at 10 μm (Fig.1 (A)). The subsequent transformation between the atlas as a source and the brain slice as a target image was conducted by a landmark based registration. The landmarks were manually generated with ImageJ and transferred to the plugin Landmark Correspondences (Fig.1 (B)) ((28)), which approximated a transformation between source and target image using the method of moving-least-squares. Although the resulting transformation was rendered by a 16-pixel mesh network, the transformed pixels of the transformed atlas have not been interpolated. This way, a sufficient registration was guaranteed whereas the original brain label association (in this case the grey values) remains.

Upon successful completion of the registration procedure, the application of AIDAhisto provides a freely available cell-counting tool written in Python 3.6 as well as Matlab (Matlab Version R2018a, The MathWorks Inc., Natick, USA). The software is premised on commonly used algorithm initially used for nuclei detection in digital (in-vitro) microscopic images ((4)). With extensive modification, AIDAhisto now enables the quantitative analysis of brain tissue slices for widely-used stainings e.g., DAPI, GFAP and Iba1 (Fig. 1 (A,C)) in high-resolution images within a short period of time. In order to start AIDAhisto, the microscopy file, the transformed ARA and the mean cell size σ are used as minimal input and the data is processed with the following steps.

First, two filter sets are used with the diameter proportional to σ for classifying (cellular) textures as perviously described (33). Whereas the rotationally invariant Schmid Filter Bank (S) suits perfectly for detecting round shapes, the Leung-Malik Filter Bank (LM) is not rotationally invariant and identifies more complex non-circular geometries ((29), (19)). Since, S comprises 13 isotropic filters and the LM set a total of 18 filters with 6 orientations and three scales, the convolution with the input image results in a stack of smoothed images. The average intensities of all images in the stack represent local extremas. Secondly, an iterative approach ((25)) is used to determine a pixel-wise x threshold T for the image I as follows

$$I(x) = \begin{cases} x - T & \text{if } x \geq T \\ T, & \text{otherwise.} \end{cases} \quad (1)$$

Thirdly, a grayscale dilation expands the shapes of the resulting image and applies an array with a logical image of a solid circle of size $R = \lfloor \sigma/2 \rfloor$ to identify the x- and y-positions of the local extremas. The detected extremas represent the calculated cell centroid. Finally, the detected cell position can be validated by a previously defined cell position pattern. In order to reduce the false positive rate, the identified cell nuclei (DAPI) are used as a reference for the cell detection in the immunostaining channels. The assignment is made in the radius $w \cdot R$ with scaling w using a k-Nearest-Neighbour algorithm with k equal to one. Antibody-labeled cells whose cell nucleus position can not be assigned in the direct neighborhood are automatically deleted. In the last step, every cell and nuclei count and the position are iteratively assigned with the respective atlas region or any other manually defined regions of interest (ROI) used as input. The results are stored as text files containing x,y-positions of each identified cell and the cell count per brain region (ROI), respectively. The output is also stored as an image where the identified cell positions are marked as one pixel each, which can be used for example to plot the cell number per brain region (Fig. 1 (D)).

2.3 Comparison to available cell-counting methods

We compared the performance of AIDAhisto to four freely available tools for counting cells: Incremental Cell Search (ICS) ((21)), QuickCount (QC) ((32)), QuPath (QP) ((2)) and ILASTIK (IL) ((31)). To validate the number of identified cells and related cell positions, two independent investigators M_1 and M_2 manually counted three cell lineages (DAPI, GFAP and Iba1) in $N = 64$ samples. The results of M_1 and M_2 were compared to each other and the average of both was

used as a reference to determine the consistency with the automatic cell count approaches. Statistical analysis with a log-transformed Bland & Altmann plots enabled the evaluation of individual deviations between the measurement methods taking into account the distortion and variance of the data ((3)). The logarithmic transformation was chosen to improve the comparison of the data as it enables a more uniform scattering. In a pairwise comparison of the number of identified cells, the limits of agreement (LoA) should lie in 95% of the values, where *SD* denotes the standard deviation of the differences $d \pm 2 \cdot$ *SD*. Moreover, the coefficient of variation (CV) is represented by the ratio of the *SD* to the mean.

A comparison of the absolute cell numbers, however, would be insufficient to judge the reliability of the cell counting algorithm. Therefore, the cell positions selected by the two manual raters were also used as classifiers. The classifiers´binary answer is represented by the manually marked cells, which maps to one of two states: Either both investigators tagged the same cell (true) or one investigator tagged a cell but the other not (false). The subsequent evaluation was adapted from previous studies ((20)). Briefly, an automatically marked cell is labelled as true positive (TP) or true negative (TN) if the detected result lies next to a true or a false cell within a circle radius r. Otherwise the detected results are labeled as false positive (FP) or as false negative (FN). The results have been summarized in a confusion matrix, with correct classification results on the major diagonal and faulty results in the secondary diagonal. Based on that matrix, we determined the accuracy by computing the F_1-Score.

Additionally, in order to quantify and assess whether automatic methods have sufficient matches to the manual observations, the calculation of Cohen's kappa κ ((8)) was applied.

$$\kappa = \frac{p_o - p_c}{1 - p_c} \tag{2}$$

That equation enables the examination of the agreement of two methods p_o in consideration of a randomly expected correspondence p_c. An advantage is that κ can be graded using the previously provided categories (Tab. 1) ((1)).

3 Results

The following evaluations are based on the implementation in Matlab (Matlab Version R2018a), but we also offer AIDAhisto in Python 3.6, which provides comparable results.

3.1 Cell counting performance

In a first step, we compared the agreement between both investigators to proof the reliability of the manual cell detection (Fig. 2 (A)). Despite the high correlation of $r^2 = 0.91$ and low RMSE of 0.31, the LoA of M_1 was on average 35% higher as the result of investigator M_2. The counting results of GFAP/$DAPI_G$ AIDAhisto and GFAP/$DAPI_G$ ICS also had a very low dispersion around the bisecting line in the scat-

Table 1: Categories of Altmann for the interpretation of Cohens's kappa κ are used to classify the detection accuracy ((1))

Range of κ	Agreement
0.81 - 1.00	Very good
0.61 - 0.80	Good
0.41 - 0.60	Moderate
0.21 - 0.40	Fair
0.00 - 0.20	Poor

ter plots and around zero in the Bland-Altman plots, whereas the approaches GFAP QC, GFAP QP and GFAP IL showed large deviations compared to $DAPI_G$ QC, $DAPI_G$ QP, $DAPI_G$ IL (Fig. 2 (B). Similar patterns were observed for the GFAP/$DAPI_I$ examination, though far smaller deviations were determined for the scatter of GFAP/$DAPI_I$ IL (Fig. 2 (C).
Detailed results for the individual measurements are given in table 2. Here, AIDAhisto achieved a correlation of 0.92 and ICS of 0.79 in comparison to the average of both investigators $M_{1/2}$. The other approaches presented lower correlations. As expected, further statistical analyses revealed deviations in the LoA ranging from AIDAhisto with 0.58 to QC with 1.9. The LoA of AIDAhisto is less than the LoA of 0.61 between M_1 and M_2. Furthermore, the log-retransformed data obtained by AIDAhisto were only 2% larger than the average cell counting results of both investigators.

3.2 Accuracy of cell positions

The investigators were also regarded as classifiers, with binary answers representing true or false cells. As there was no difference in the cell nuclei counting ($DAPI_G$) in the GFAP and Iba1 immunostainings (Figure 2), they were merged to a single DAPI group. Additionally, as the cell counting results of ICS were very similar to AIDAhisto in the previous examination, we next assessed whether if tagged cells were either correct or false cells with respect to the pre-set classes for all three cell lineages (DAPI, GFAP, Iba1). The confusion matrices contain the number of observations in the four possible combinations of automatically and manually marked cells (Fig. 3). True positive (TP) and true negative (TN) represent the correct classification on the main diagonal of the confusion matrix. The TP ratios of AIDAhisto were higher (74.4-88.5%) compared to ICS (20.7-75.1%), whereas ICS performed slightly better in detecting TN cells (96.4-98.6% vs. 90.8-97.4%). Furthermore, when the algorithm returned a cell position classified as incorrect, it was counted as false positive (FP), and a not automatically detected cell position was classified as false negative (FN). The FP ratio of AIDAhisto was lower for cell nuclei, as well as the GFAP and Iba1 immunostainings (11.5-25.6% vs. 24.9-79.3%), whereas ICS performed slightly better in the FN ratio (1.4-3.6% vs. 2.6-9.2%). Collectively, the F_1-Score was used to characterize the overall classification accuracy for each matrix. While the F_1-Score of AIDAhisto was not different for cell nuclei and immunostainings (82.33-88.78%), ICS reached a high F_1-Score only for cell nuclei (82.01%) and lower F_1-Score for immunostainings (45.45-72.39%).

In addition to the F_1-Score, we calculated the Cohen's Kappa κ, that measures the difference between the proportion of real agreement and the proportion of agreement that would be expected by pure coincidence (Fig. 4). The DAPI cell nuclei positions of both approaches were comparable and the median of both lies in the the second category marked as "good" (Fig. 4 (A)). For the GFAP and Iba1 immunostainings, the resulting quality category was found to be different. The median of the κ values decreased to a category marked as "Fair" for GFAP (Fig. 4 (B)) and as "Poor" (Fig. 4 (C)) for Iba1, whereas the median of AIDAhisto remained stable in the category marked as "Good". Thus, the ICS results were in less agreement with both investigators compared to the AIDAhisto results.
In all previous measures, the cell nuclei de-

Table 2: Quantitative comparison of the five tested cell counting approaches $\mathbf{A_{1-5}}$ with the average results of the two investigators $\mathbf{M_{1/2}}$ using squared Pearson correlation Value $\mathbf{r^2}$, root mean square error (RMSE), limits of agreement (LoA), difference $\mathbf{d}$ with standard deviation SD, and the coefficient of variation (CV)

Software	Acronym	RMSE	r^2	LoA	$d \pm 2 \cdot SD$	CV
Manual	-	0.31	0.91	0.61	0.30 +0.91/-0.31	6.9 %
AIDAhisto	AIDAhisto	0.30	0.92	0.58	0.03 +0.55/-0.61	6.6 %
Incremental Cell Search	ICS	0.53	0.79	1.0	0.05 +0.98/-1.1	12 %
QuickCount	QC	0.73	0.18	1.9	0.25 +2.2/-1.7	21 %
QuPath	QP	0.46	0.26	1.7	0.97 +2.6/-0.71	17 %
ILASTIK	IL	0.70	0.61	1.4	0.38 +1.0/- 1.8	16 %

tection was more accurate compared to the immunostainings. Therefore, the DAPI results were used as reference data to correct the immunostaining results. Based on the assumption that the filter detects the immunostaining in a certain range next to the cell nuclei, we implemented that feature in AIDAhisto which increased the F_1-Score by 9.88% (Fig. 4D). Here, both investigators counted N=4 DAPI/Iba1 samples again and were recruited as classified as described above. As plotted in the confusion matrix (Fig. 4D), the TP and TN were increased and FP and FN decreased by the additional cell position validation.

3.3 Qualitative examples

Next to astroglia and immune cells, we tested AIDAhisto for other experiments (Fig. 5). We achieved qualitatively good results for the detection of neuroblasts expressing doublecortin (DCX), the neuronal nuclei marker NeuN on brain and spinal cord slices as well as detection of viral tracing with tdTomato expressed in layer 5 pyramidal neurons in the mouse brain. These examples show that the AIDAhisto counting algorithm can be applied to a broad range of neuroscience applications.

4 Discussion

A detailed mapping of histopathology to brain function plays a major role in understanding brain diseases on a cellular level. A key element of that approach is a quantitative cellular analysis within a reference atlas to increase the comparability and transferability of different studies. Currently, tools for quantifying cells in histological stainings (e.g., Nissl or Hematoxilyn/Eosin) or immunohistochemical stainings were optimized for a particular feature, e.g., cell stainings or atlas-based analysis of healthy mouse brain tissue (Supplementary Material, Table 1). With AIDAhisto, we developed and validated a cell counting tool that enables a quantitative analysis of 2D whole-brain mouse preparations with respect to the provided structures of the Allen Mouse Brain Reference Atlas ((18) and (22)). Using an appropriate 3D mouse brain atlas provides in-depth characterization of brain regions across modalities and histology-to-behavior correlations. While for in-vivo imaging, such as magnetic resonance imaging (MRI),

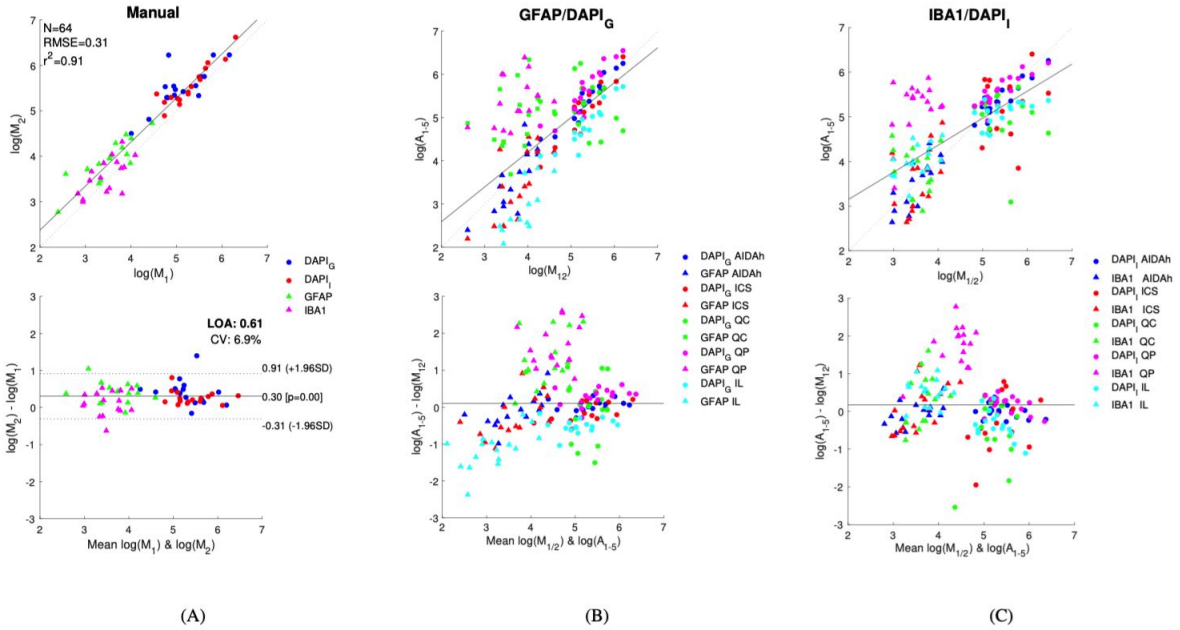


Figure 2: Comparison of the manual cell counting (manual rater $\mathbf{M_1}$ and $\mathbf{M_2}$) and the automated cell counting tools $\mathbf{A_{1-5}}$: AIDAhisto (AIDAh), Incremental Cell Search (ICS), QuickCount (QC), QuPath (QP), and ILASTIK (IL). The counting results based on N=64 samples were subdivided into the stainings Iba1/DAPI$_{\mathrm{I}}$ and GFAP/DAPI$_{\mathrm{G}}$. The results are shown as scatter plots (upper row) with bisecting (dashed line) and regression lines (solid line) as well as squared Pearson Correlation Value $\mathbf{r^2}$ and Root Mean Square Error (RMSE). The log-transformed Bland-Altman plots (bottom row) contain the limits of agreement (LoA) at $\pm\mathbf{1.96SD}$ and coefficient of variation (CV). The LoA of 0.61 is the statistical measure of uncertainty between both investigators $\mathbf{M_1}$ and $\mathbf{M_2}$, which vary with a difference of $\mathbf{d = 0.3}$ ($\mathbf{\pm 0.91 / -0.31}$). Compared to the average of both investigator $\mathbf{M_{1/2}}$ (A), the automated assessments $\mathbf{A_{1-5}}$ (B and C) show striking differences for both staining groups. In both cases, the results of each individual process should be as close as possible to the solid line. Especially the blue and red marked triangles and circles for AIDAh and ICS fulfill this criterion. For reasons of clarity, the quantitative results are listed in Table 2.

brain atlases are widely used ((23)), the histological validation is performed by delineating the specific brain regions of interest in a manual and difficult to reproduce approach ((13)). Different to the in-vivo situation ((23)), the registration of ex-vivo histology with the ARA requires a non-linear registration to match the strong tissue deformation induced by ex-vivo tissue processing. Usually, tissue cryosectioning and handling during histology induce compression and/or stretching, which prevent a direct overlay of the microscopy and the atlas. Previous approaches used principle component analysis and other algorithms to detect the tissue border zone and generate a mesh for non-linear registration ((12) and (17)) (Supplementary Material, Table 1). However, to our knowledge there is no fully automated algorithm available and we found such

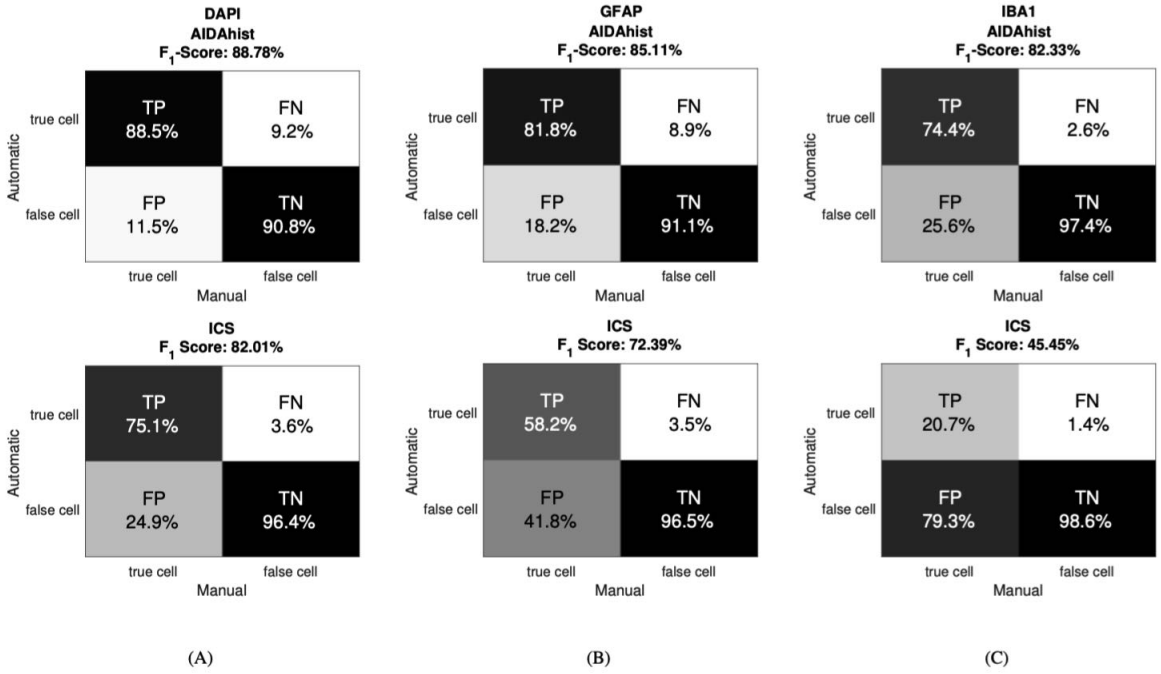
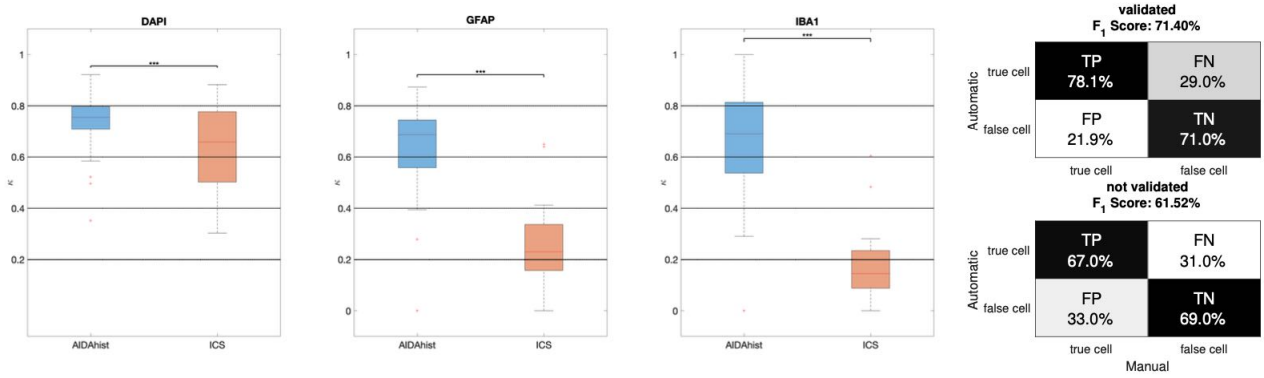


Figure 3: Confusion matrices with correct classification results in the main diagonal, incorrect results in the secondary diagonal. ICS and AIDAhisto show nearly comparable results for the DAPI samples with F_1-Scores of 82.01% and 88.78% (A). In contrast, for the other cell lineages the F_1-Score of ICS decreased because the FP hardly increased from 41.8% for GFAP (B) up to 79.3% for Iba1 (C). The F_1-Score of AIDAhisto only slightly decreased and remained stable for each cell lineage.

approaches to work well for mouse brain tissue from healthy mice but in case of brain lesions, a different approach is needed ((16)). Here, the association between imaging data and atlas regions is based on a simple and fast landmark-based registration tool that is implemented in ImageJ ((27)) and thus allows a straightforward and intuitive operation. The manual registration depends on the correct choice of landmarks and is fast and accurate when performed by an experienced user. Notably, that procedure should be applied for matching slides (histology-atlas) only. If one image is tilted, e.g. if there is an additional skewed angle from the brain sectioning, that needs to be corrected with other tools. The subsequent cell detection was fully automated and tested in Python 3.6 and Matlab (Matlab Version R2018a). To overcome the limitations of conventional rotationally invariant filters, which detect round-shaped structures but no cell structures with different shapes, we advanced the previous implementations of kernel-based cell nuclei detection detection adding S - Filter Bank (S) and LM - Filter Bank (LM) ((21),(29),(19)). Both filter banks are rotationally invariant and enabled the identification of more complex non-circular geometries. Moreover, the threshold value is not determined de-



(A) (B) (C) (D)

Figure 4: Results of the calculation of Cohen's kappa κ show large differences for AIDAhisto and ICS. The calcuation was based on all N=64 samples sub-divided in DAPI (A), GFAP (B), and Iba1 (C). Horizontal lines separate the categories of Altmann (Table 1).AIDAhisto presents a highly significant improvement to ICS with $\mathbf{p < 0.001}$ (DAPI p=5.76e-06, GFAP p= 6.86e-13, Iba1 p=3.07e-13) and the median remained steady in the category marked as "Good". The confusion matrices show the comparison of the Iba1/DAPI countings with and without additional validation based on the detected cell nuclei. The cell nuclei position was used to delete Iba1 cells with no cell nuclei in radius r. Cell nuclei validation improved Iba1 counting for all measures (TP/FP and F_1-Score increased, FN/FP rate decreased (D).

pending on the input of the user, but automatically calculated based on the input image by an established method. This allows batch processing of a large amount of data in a very short time.

In a comprehensive validation study, we could show that AIDAhisto outperforms other cell counting software using Iba1 and GFAP stainings in terms of accuracy and reliability for cell nuclei and cell structures despite different cell body geometries and pixel brightness due to the variable immunostaining and microscopy procedures. Compared to two expert raters, AIDAhisto provided the same accuracy for the detection of antibody-stained astrocytes and immune cells in the mouse brain. Furthermore, by using the highly accurate cell nuclei detection by AIDAhisto, we implemented a feature to correct the immunostaining results. This way only cells with cell nuclei will be counted in the immunostaining results.

Complex methods and architectures of pattern recognition for the automatic detection and classification of cells have become increasingly popular in recent years and show advantages over manual image processing methods ((31), (7), (26)). However, compared to sophisticated methods, like deep neuronal networks, AIDAhisto is a single system solution. Expensive hardware and large amounts of data for training procedures are not necessary. Instead, a single computer can evaluate the image data in minutes. A prerequisite for efficient and reliable cell detection is image quality and a minimal amount of image artifacts, blurring, contrast inhomogeneity, or variable intensity. AIDAhisto

was successfully tested on stitched whole mouse brain slices imaged using various microscopes and different image resolutions without differences in cell detection. AIDAhisto detects reliably various shapes of cell nuclei and cells. Here, we performed the extensive quantitative comparison only for GFAP and Iba1 as they represent a large variety of cellular shapes similar to a variety of neurons. In other examples for immunostaining of neurons on brain and spinal cord tissue as well as viral tracing, we could show that qualitatively, AIDAhisto reliably detects cell nuclei and cells. One major limitation of AIDAhisto is the cell counting in 2D only, thus the algorithm is not capable of discriminating cells in z-direction, such as in large z-stacks or 3D light sheet microscopy With the offered tool, we hope to provide a simple solution for non-expert users to achieve an accurate atlas-based cell counting and facilitate a more broader comparison of experimental approaches in preclinical studies of human brain disorders in the mouse model.

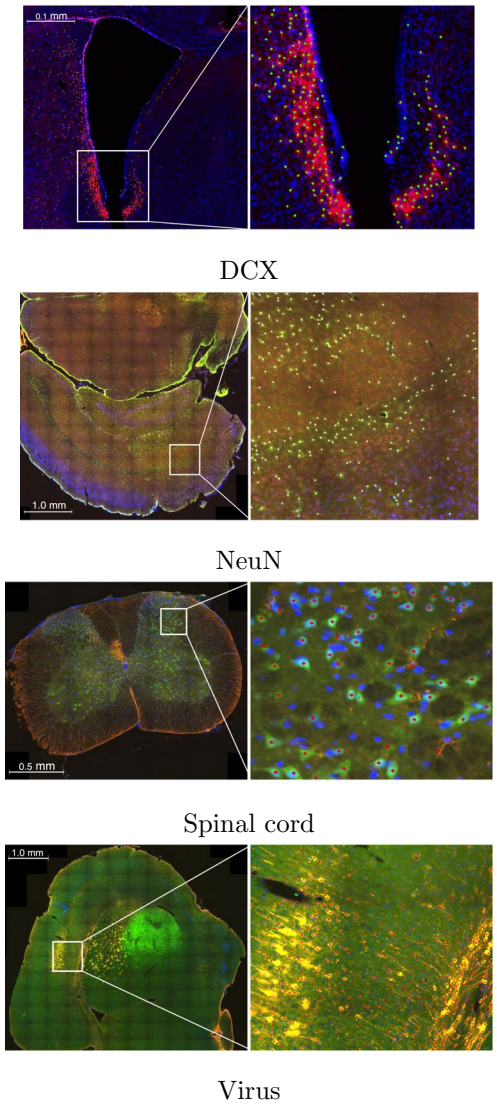


Figure 5: Qualitative results of common stainings in the mouse brain and spinal cord. Cell nuclei in all experiments counterstained with DAPI (blue). White boxes indicate zoom area. DCX (Doublecortin, red) in the subventricular zone of the lateral ventricle. Neuronal nuclei (NeuN) in the posterior part of the hippocampus and the spinal cord. Expression of td-Tomato in cortex layer 5 pyramidal neurons targeted by injection of an Adeno-associated virus.

for which other atlas tools have been proposed recently. Therefore, similar to other 2D counting tools including stereology, overcounting can appear when adjacent slices are being analyzed (a 50 μm cell might be counted twice in two adjacent 20 μm slices) ((15)). For selected brain slices, when no combined serial slicing and staining is available, AIDAhisto results can be compared to whole brain atlas-registered in-vivo imaging features, which provides a direct comparison of imaging to cellular features in the same atlas space ((23)).

5 Acknowledgements

We thank Olivia Kaesgen and Veronika Fritz for technical support. This work was supported by funding from the Friebe Foundation (T0498/28960/16).

6 Declarations of interest

None.

References

[1] D. G. Altman. *Practical statistics for medical research*. CRC press, 1990.

[2] P. Bankhead, M. B. Loughrey, J. A. Fernández, Y. Dombrowski, D. G. McArt, P. D. Dunne, S. McQuaid, R. T. Gray, L. J. Murray, H. G. Coleman, et al. Qupath: Open source software for digital pathology image analysis. *Scientific reports*, 7(1):16878, 2017.

[3] J. M. Bland and D. Altman. Statistical methods for assessing agreement between two methods of clinical measurement. *The lancet*, 327(8476):307–310, 1986.

[4] J. Byun, M. R. Verardo, B. Sumengen, G. P. Lewis, B. Manjunath, and S. K. Fisher. Automated tool for the detection of cell nuclei in digital microscopic images: application to retinal images. *Mol Vis*, 12(105-07):949–60, 2006.

[5] M. Carter and J. Shieh. *Guide to Research Techniques in Neuroscience*. Academic Press, second edition edition, 2015.

[6] J. Chalfoun, M. Majurski, T. Blattner, K. Bhadriraju, W. Keyrouz, P. Bajcsy, and M. Brady. Mist: accurate and scalable microscopy image stitching tool with stage modeling and error minimization. *Scientific reports*, 7(1):4988, 2017.

[7] D. Ciresan, A. Giusti, L. M. Gambardella, and J. Schmidhuber. Deep neural networks segment neuronal membranes in electron microscopy images. In *Advances in neural information processing systems*, pages 2843–2851, 2012.

[8] J. Cohen. A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1):37–46, 1960.

[9] K. De Vos. Cell counter. *University of Sheffield, Academic Neurology*, 2010.

[10] H.-U. Dodt, U. Leischner, A. Schierloh, N. Jährling, C. P. Mauch, K. Deininger, J. M. Deussing, M. Eder, W. Zieglgänsberger, and K. Becker. Ultramicroscopy: three-dimensional visualization of neuronal networks in the whole mouse brain. *Nature Methods*, 4(4):331–336, Mar 2007.

[11] C. Eroe, M.-O. Gewaltig, D. Keller, and H. Markram. A cell atlas for the mouse brain. *Frontiers in neuroinformatics*, 12:84, 2018.

[12] D. Fürth, T. Vaissière, O. Tzortzi, Y. Xuan, A. Märtin, I. Lazaridis, G. Spigolon, G. Fisone, R. Tomer, K. Deisseroth, et al. An interactive framework for whole-brain maps at cellular resolution. *Nature neuroscience*, 21(1):139, 2018.

[13] L. Hammelrath, S. Škokić, A. Khmelinskii, A. Hess, N. van der Knaap, M. Staring, B. P. Lelieveldt, D. Wiedermann, and M. Hoehn. Morphological maturation of the mouse brain: An in vivo mri and histology investigation. *Neuroimage*, 125:144–152, 2016.

[14] M. Hawrylycz, R. A. Baldock, A. Burger, T. Hashikawa, G. A. Johnson, M. Martone, L. Ng, C. Lau, S. D. Larsen, J. Nissanov, et al. Digital atlasing and standardization in the mouse brain. *PLoS computational biology*, 7(2):e1001065, 2011.

[15] V. Howard and M. Reed. *Unbiased stereology: three-dimensional measurement in microscopy*. Garland Science, 2004.

[16] M. Ito, M. Aswendt, A. G. Lee, S. Ishizaka, Z. Cao, E. H. Wang, S. L. Levy, D. L. Smerin, J. A. McNab, M. Zeineh, and et al. Rna-sequencing analysis revealed a distinct motor cortex transcriptome in spontaneously recovered mice after stroke. *Stroke*, 49(9):2191–2199, Sep 2018.

[17] C. D. Kopec, A. C. Bowers, S. Pai, and C. D. Brody. Semi-automated atlas-based analysis of brain histological sections. *Journal of neuroscience methods*, 196(1):12–19, 2011.

[18] E. S. Lein, M. J. Hawrylycz, N. Ao, M. Ayres, A. Bensinger, A. Bernard, A. F. Boe, M. S. Boguski, K. S. Brockway, E. J. Byrnes, and et al. Genome-wide atlas of gene expression in the adult mouse brain. *Nature*, 445(7124):168–176, Dec 2006.

[19] T. Leung and J. Malik. Representing and recognizing the visual appearance of materials using three-dimensional textons. *International journal of computer vision*, 43(1):29–44, 2001.

[20] F. Liu and L. Yang. A novel cell detection method using deep convolutional neural network and maximum-weight independent set. In *Deep Learning and Convolutional Neural Networks for Medical Image Computing*, pages 63–72. Springer, 2017.

[21] O. E. Meruvia-Pastor, J. Soh, E. J. Schmidt, J. C. Boughner, M. Xiao, H. A. Jamniczky, B. Hallgrímsson, and C. W. Sensen. Estimating cell count and distribution in labeled histological samples using incremental cell search. *Journal of Biomedical Imaging*, 2011:12, 2011.

[22] S. W. Oh, J. A. Harris, L. Ng, B. Winslow, N. Cain, S. Mihalas, Q. Wang, C. Lau, L. Kuan, A. M. Henry, et al. A mesoscale connectome of the mouse brain. *Nature*, 508(7495):207, 2014.

[23] N. Pallast, M. Diedenhofen, S. Blaschke, F. Wieters, D. Wiedermann, M. Hoehn, G. R. Fink, and M. Aswendt. Processing

pipeline for atlas-based imaging data analysis of structural and functional mouse brain mri (aidamri). *Frontiers in Neuroinformatics*, 13, Jun 2019.

[24] T. Ragan, L. R. Kadiri, K. U. Venkataraju, K. Bahlmann, J. Sutin, J. Taranda, I. Arganda-Carreras, Y. Kim, H. S. Seung, and P. Osten. Serial two-photon tomography for automated ex vivo mouse brain imaging. *Nature methods*, 9(3):255, 2012.

[25] T. Ridler, S. Calvard, et al. Picture thresholding using an iterative selection method. *IEEE trans syst Man Cybern*, 8(8):630–632, 1978.

[26] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[27] S. Saalfeld, R. Fetter, A. Cardona, and P. Tomancak. Elastic volume reconstruction from series of ultra-thin microscopy sections. *Nature methods*, 9(7):717, 2012.

[28] S. Saalfeld and P. Tomancák. Automatic landmark correspondence detection for imagej. In *Proceedings of the ImageJ User and Developer Conference*, pages 128–133, 2008.

[29] C. Schmid. Constructing models for content-based image retrieval. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 2, pages II–II. IEEE, 2001.

[30] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri. Nih image to imagej: 25 years of image analysis. *Nature methods*, 9(7):671, 2012.

[31] C. Sommer, C. Straehle, U. Koethe, and F. A. Hamprecht. Ilastik: Interactive learning and segmentation toolkit. In *2011 IEEE international symposium on biomedical imaging: From nano to macro*, pages 230–233. IEEE, 2011.

[32] K. H. Tiong, J. K. Chang, D. Pathmanathan, M. Z. Hidayatullah Fadlullah, P. S. Yee, C. S. Liew, Z. A. Abdul Rahman, K. L. Beh, and S. C. Cheong. Quickcount®: a novel automated software for rapid cell detection and quantification. *BioTechniques*, 65(6):322–330, 2018.

[33] M. Varma and A. Zisserman. A statistical approach to texture classification from single images. *International journal of computer vision*, 62(1-2):61–81, 2005.