



OPEN

Statistical physics approach to quantifying differences in myelinated nerve fibers

SUBJECT AREAS:
NEUROSCIENCE
BIOLOGICAL PHYSICS
STATISTICAL PHYSICS

César H. Comin^{1*}, João R. Santos^{2*}, Dario Corradini^{2†}, Will Morrison², Chester Curme², Douglas L. Rosene³, Andrea Gabrielli^{4,5}, Luciano da F. Costa¹ & H. Eugene Stanley²

Received
24 January 2014

Accepted
12 March 2014

Published
28 March 2014

¹Institute of Physics at São Carlos, University of São Paulo, São Carlos, SP 13560-970, Brazil, ²Center for Polymer Studies and Department of Physics, Boston University, 590 Commonwealth Avenue, Boston, Massachusetts 02215, USA, ³Boston University, School of Medicine, Department of Anatomy & Neurobiology, Boston, Massachusetts 02118, USA, ⁴Istituto dei Sistemi Complessi (ISC) - CNR, UOS "Sapienza", Dipartimento di Fisica, "Sapienza" Università di Roma, Piazzale Aldo Moro 5, 00185 Rome, Italy, ⁵IMT Alti Studi Lucca, Piazza S. Ponziano 6, 55100 Lucca, Italy.

Correspondence and requests for materials should be addressed to J.R.S. (jrsantos@bu.edu)

* These authors contributed equally to this work.

† Current address: Sorbonne Universités, UPMC Univ Paris 06, UMR 8234, PHENIX, Paris, France & CNRS, UMR 8234, PHENIX, Paris, France.

We present a new method to quantify differences in myelinated nerve fibers. These differences range from morphologic characteristics of individual fibers to differences in macroscopic properties of collections of fibers. Our method uses statistical physics tools to improve on traditional measures, such as fiber size and packing density. As a case study, we analyze cross-sectional electron micrographs from the fornix of young and old rhesus monkeys using a semi-automatic detection algorithm to identify and characterize myelinated axons. We then apply a feature selection approach to identify the features that best distinguish between the young and old age groups, achieving a maximum accuracy of 94% when assigning samples to their age groups. This analysis shows that the best discrimination is obtained using the combination of two features: the fraction of occupied axon area and the effective local density. The latter is a modified calculation of axon density, which reflects how closely axons are packed. Our feature analysis approach can be applied to characterize differences that result from biological processes such as aging, damage from trauma or disease or developmental differences, as well as differences between anatomical regions such as the fornix and the cingulum bundle or corpus callosum.

The nervous system is a complex network allowing the transmission of signals between interconnected neurons across distances which vary from fractions of millimeters to meters. Axons following similar paths are often bundled together, forming nerves in the peripheral nervous system and tracts in the central nervous system such as the corpus callosum which interconnects the two brain hemispheres. The proper functioning of such tracts depends on axon characteristics such as size, density and spatial organization. The axons populating different tracts or bundles change during development^{1–8} and aging^{9–14}, as well as a consequence of pathology^{15–26} and environmental influences^{27–29}. Alterations in specific genes might also influence the organization of bundles of axons^{2,30,31}. It is therefore important to have a means for quantifying, in an objective manner, the characteristics of these axons and their bundles and to discern which features best characterize the observed differences.

Typically, studies of differences observed in nerve fibers are limited to one or just a few geometrical properties, chosen to measure an evident and already observed difference, and then statistically evaluate that difference between groups, in an individual fashion. This methodology presents several problems, as some of the differences may be subtle and not easily identified by visual inspection and hence not chosen for quantification, restricting the identification of potential differences. As a result, there have been few systematic attempts to analyze and explore a wide range of possible features as means to identify the main effects^{32–35}.

We present here a new method to uncover which features of axons are most affected by an underlying biological process. In this systematic investigation, we consider *a priori* a large set of candidate features representative of diverse types of possible differences in axons (e.g. density, shape of axons, spatial order, etc.) and then use the feature selection technique^{36–38} to identify which combinations of such features yield the best discrimination between axons of two distinct groups. This approach enables the identification from the list of candidate features, of the *set* of features that, when taken *together*, best discriminates between groups in a general dataset.

As a case study, we apply this methodology to characterize changes observed with normal aging in the myelinated nerve fibers of the fornix of young and old rhesus monkeys. The fornix is a bundle of nerve fibers

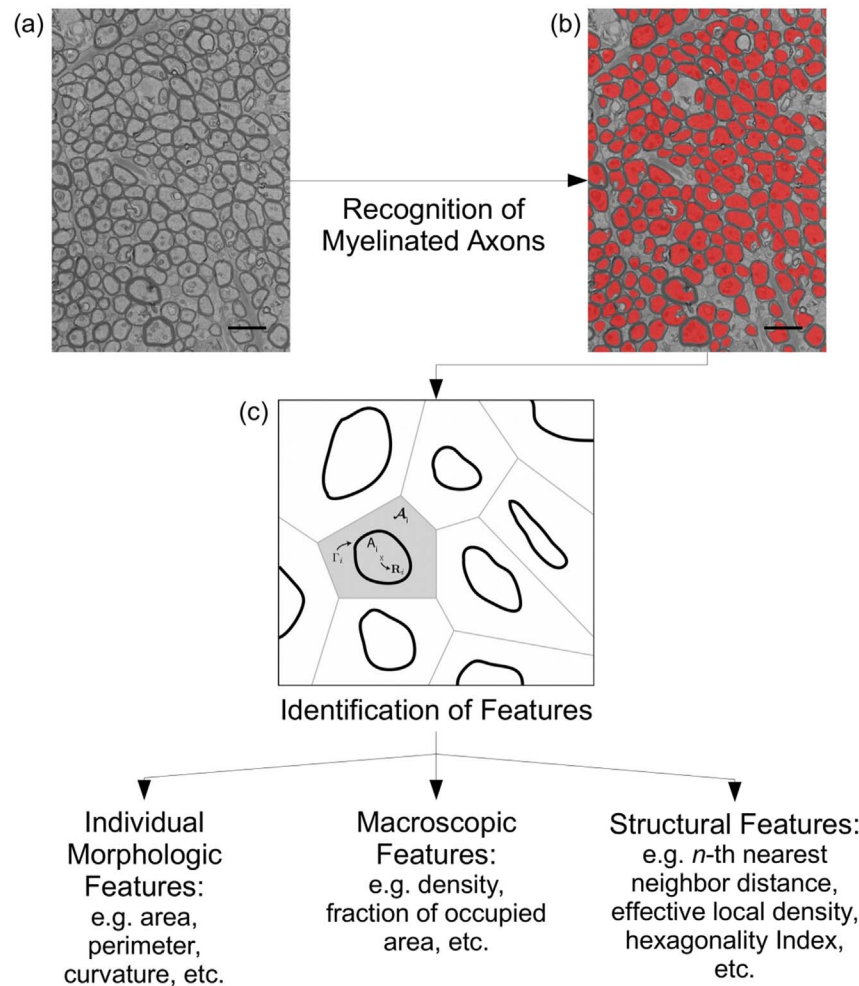


Figure 1 | Processing steps for characterizing myelinated axons and their structure. The characterization process is done on a total of 67 electron micrographs (EM images) of the fornix, belonging to 6 different subjects. (a) We show, as an example, one of the EM images of subject AM77 (6 years old). The scale bar (in black) measures $2 \mu\text{m}$. (b) Our recognition algorithm (see Methods for details) is used to segment the areas of myelinated axons (in red). (c) Schematic diagram of a small set of myelinated axon contours and the respective Voronoi tessellation of the embedding space. For one of the axons we show the relevant properties used to characterize each axon: they are the centroid position \mathbf{R}_i , axon area A_i , Voronoi cell area \mathcal{A}_i and axon shape parameters, collectively referred to as Γ_i . These properties are then used as a basis to calculate the 45 features used to describe each EM image (e.g. axon density, hexagonality index, etc.).

carrying signals from the hippocampus to other structures in the limbic system, such as the hypothalamus, and it is crucial in normal cognitive functions, especially memory formation³⁹. Previous studies have shown that the axon density declines with age⁴⁰. Our methodology shows that density related features, in particular the fraction of occupied axon area and the effective local density, which measures how closely axons are packed, are the features that provide the best discrimination between the young and old age groups. These features are shown to be better than simple density alone, which provides an accuracy of only 75% when assigning samples to their age groups. Furthermore, we show that using the combination of fraction of occupied axon area and effective local density is enough to characterize the age group separation in the samples observed, with an accuracy of 90%, and that the addition of other features adds little to the accuracy.

Results

Recognition and characterization of myelinated axons. In Fig. 1 we show a diagram illustrating the steps we take to characterize the myelinated axons from each electron microscope (EM) image. We start with an EM image of the fornix, showing a cross-sectional cut of the myelinated nerve fibers. Using the recognition algorithm

described in Methods, we obtain the outlines of the axolemma of myelinated axons. Based on this recognition protocol, we determine, for each myelinated axon, the following properties: the centroid position \mathbf{R}_i , its area A_i , the area of its Voronoi cell \mathcal{A}_i and 6 morphologic features known to provide a good description of a shape's geometrical properties, i.e. perimeter, elongation, circularity, diameter, mean curvature and bending energy⁴¹. In Fig. 1(a) we see a representative EM image of a young subject, while the recognized axons are shown overlaid on top of the original EM image in Fig. 1(b). In Fig. 1(c) we draw a schematic diagram of the system in order to clearly visualize the properties used as a basis for our analysis.

Using these 9 individual axon properties as a basis, we derive a collection of 45 different features for each sample (i.e. each EM image). These include statistical moments of the distributions (mean, standard deviation, skewness) as well as correlations between the axon properties. These features can be classified into three categories: macroscopic, morphologic and structural features. Macroscopic features, related to the characteristics of collections of axons, include quantities such as density and the fraction of occupied area (i.e. the ratio of the area occupied by the myelinated axons to the total sample area). Morphologic features refer to characteristics of individual

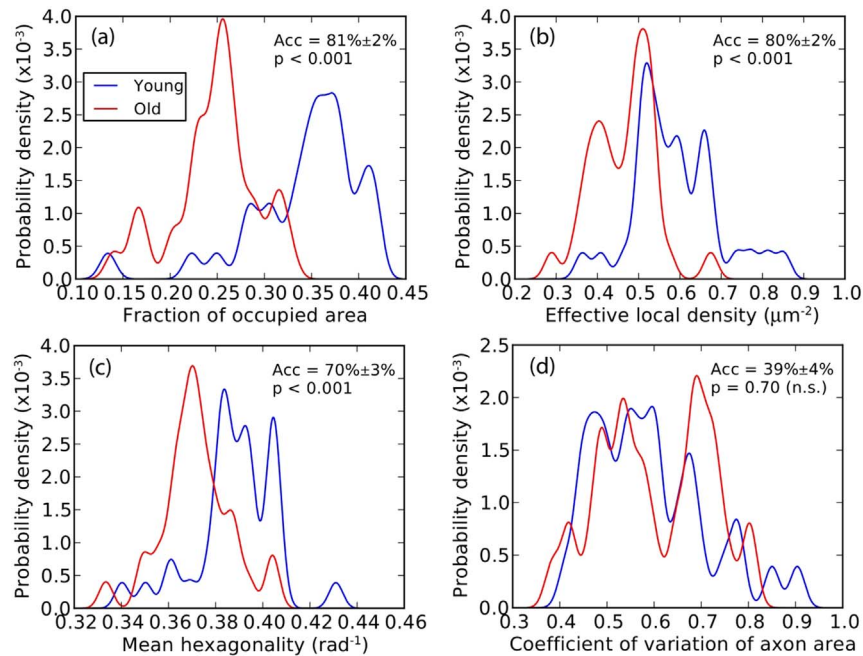


Figure 2 | Single feature analysis. Probability density functions of 4 representative features for the young group (in blue) and old group (in red). Figures (a) and (b) show the two features providing the highest accuracy in age group classification. Figure (c) shows a feature representative of the system's structural regularity. Figure (d) shows the feature providing the lowest accuracy in age group classification. The probability density functions are determined by using a gaussian kernel density estimation method. In each plot we also show the accuracy in classification using that single feature, and the Welch's t-test p-value detailing the statistical difference between the two distributions (n.s. stands for not significant, $p > 0.05$).

axons and include parameters of the area distribution, perimeter, circularity and curvature, for example. Finally, structural features, related to the relationship between different axons in each sample, include, among others, parameters of the nearest neighbors distributions and quantities derived from the Voronoi analysis such as the mean hexagonality index and the mean number of Voronoi neighbors. Supplementary Table S1 lists all the features collected together with their respective definitions.

Single feature analysis. We begin our analysis by considering each feature individually, and calculating its mean value for each EM image, or sample. We determine the accuracy of each feature when used to discriminate between young and old, defined as the cross-validated percentage of correctly classified samples⁴² and calculated using a K-nearest neighbors classifier⁴³, with $K = 3$ (see Methods for details). We note that accuracy refers to the correct classification of the individual samples, i.e. each EM image, and not of the entire set from each subject. In order to estimate the statistical difference between the two age groups, we also perform for each feature, a Welch's t-test on the mean values for each sample. The values obtained for all features can be found in the Supplementary Table S2. In Fig. 2 we show the estimated probability density functions for 4 representative features, as well as the accuracy (Acc) and the p-value of the Welch's t-test.

Fig. 2(a) and Fig. 2(b) show the two features that provide the best accuracy in classification, i.e. differentiating between membership in the young versus the old group. Since it is known that myelinated axon density declines with age⁴⁰, we expect to obtain high accuracies for density related features. The fraction of occupied area, shown in Fig. 2(a), is one such feature. It combines the macroscopic information provided by the density with the morphologic information given by the area of the axons. We find that such combination leads to an 81% accuracy for discrimination between age groups. Another interesting result comes from the effective local density, Fig. 2(b). We see that this feature has an accuracy of 80%, which is almost the same

accuracy as the fraction of occupied area, and is a better single feature discriminant of the two age groups than the actual axon density, which has an accuracy of 75%.

The effective local density introduced above is defined as the density of a random particle distribution with the same n -th nearest neighbor distance scaling (see Methods for definition). By considering only the nearest neighbor distances, the effective local density depends only on how closely packed the axons are, thus providing an estimate of the local density around the axons, ignoring the presence of any abnormally large axon-free regions occupied by excessive myelin sheath or other biological elements such as glia and blood vessels. We define an axon-free region as any region in the EM image that is not occupied by a myelinated axon. Samples with high effective local density have regions with large clustered axon-free areas with no myelinated axons, that skew the axon density measurement towards smaller values. On the other hand, samples with low effective local density have its axon-free regions spread out, i.e. equally distributed throughout the sample, with little to no clustering effect (see Supplementary Fig. S1 for details).

We observe that the axons have different structures in the two age groups, reflected in the difference of the hexagonality index, shown in Fig. 2(c). The hexagonality index measures the angular regularity of a structure, and is equal to 1 when the system is perfectly ordered as a triangular lattice, decreasing in value as a system's disorder increases⁴⁴ (see Methods for definition). From Fig. 2(c), we see that samples from the old subjects have a lower hexagonality index than the young subjects, reflecting an increase in disorder with age. Despite the observable increase in disorder with age, we find that the mean number of neighbors, measured by the number of adjacent Voronoi cells, is the same for both the young and old age groups: 5.93.

Finally, Fig. 2(d) shows the feature that gives the worst accuracy: the coefficient of variation of the axon area, a morphologic feature. In fact, considering the rest of the morphologic features, most provided low accuracies and showed no significant differences between the

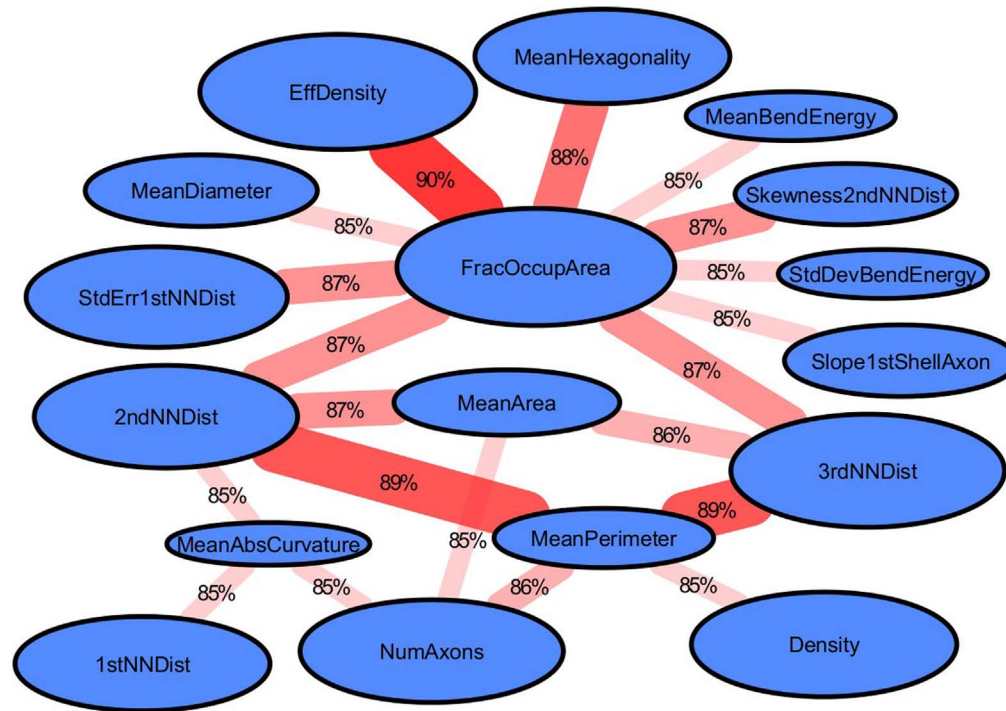


Figure 3 | Graph of the classification accuracy provided by pairs of features. The size of each ellipse is proportional to the accuracy provided by the use of that feature alone. The value on each link (illustrated by both the color and width of the link) represents the accuracy that the two features connected by the link provide when separating the samples into the two age groups. Only links having an accuracy greater or equal to 85% are shown. We note that all accuracy values shown have an associated error smaller than 2%. See Supplementary Table S1 for a list of the abbreviated names given to the features.

two age groups. In other words, there is no observable change in axon size or axon shape with aging, with the mean elongation being the only exception, with an accuracy of 61%.

Feature selection. Using single features, the highest accuracy attained distinguishing between samples from young and old is 81%. When all features are considered together the accuracy increases only to 84%, just a slight increase from the single feature case with the highest accuracy. This reflects the fact that using a high dimensional space of features comes with many issues for classification (e.g. difficulty in sampling entire space, increased execution time, additional classification noise)⁴⁵, especially when there are many irrelevant features.

To reduce the dimensionality of the fornix data and avoid problems associated with the high dimensionality one could consider using the Principal Component Analysis (PCA) procedure⁴⁶ to reduce the dimensions of the data to a few components that account for as much variability in the data as possible. However, since we are not interested in describing the variability of the age groups but in distinguishing between them and pinpointing which features are the most important for that task, we used an alternative technique called feature selection^{36,37}. This technique aims to find a subset of features that, *when combined*, gives a good separation between classes (i.e. the two age groups). In this technique, one considers a subset of features instead of each one individually, since features that alone do not give a good separation between classes can significantly improve discrimination when combined with other features^{36,37}. In a similar fashion, two highly correlated features, which in principle would hold redundant information, can also help the discrimination process in some cases.

Two feature analysis. Although the feature selection technique does not fix the size of the subset of features *a priori*, for reasons that will be clear below, we start by limiting the subset size to 2. From the total of 990 possible pairs of features, we keep only those pairs with more

than 85% of accuracy, and construct the graph shown in Fig. 3 describing the accuracy relationship between the features. An immediate characteristic we observe in the graph is that the fraction of occupied area makes connections with many other features (usually called a hub in network terminology), which confirms the importance of this feature. Considering that the fraction of occupied area alone already provides an accuracy of 81%, it is expected that a small contribution to the classification from another feature will create a link between the two features. However, this high connectivity is not replicated by other features with similarly large accuracies (e.g. effective local density, third nearest neighbor mean distance).

The highest accuracy for the age group separation is achieved for the combination of fraction of occupied area and effective local density, with a 90% accuracy. However, considering that the change of the class of a single sample roughly translates to a 1.5% difference in accuracy, other pairs of features with small differences in accuracy should also be taken into consideration. From Fig. 3, we observe that the pairs of features that result in the highest accuracies are: the combination of the fraction of occupied area (macroscopic feature) with either the effective local density or the hexagonality index (structural features); or the combination of the mean perimeter (morphologic feature) with either the second nearest neighbor mean distance or the third nearest neighbor mean distance (structural features). These 4 pairs of features result in accuracies larger than 87.5% in the age group separation. We note that the fraction of occupied area is strongly correlated with structural features such as mean hexagonality index and effective local density, which means that these feature combinations bring some redundant information to the classification (see Supplementary Fig. S2 for details). Despite this redundancy, there is still a significant improvement in the sample separation accuracy when combining the fraction of occupied area with the effective local density or the mean hexagonality index.

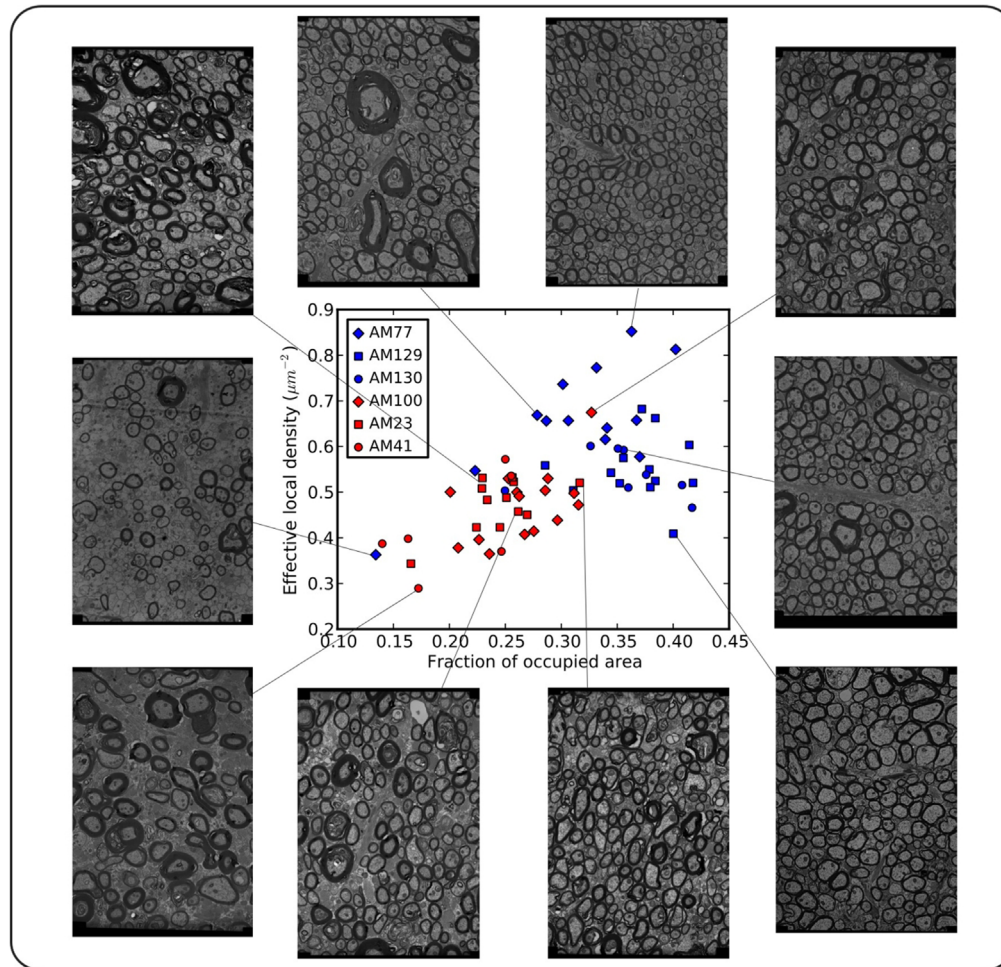


Figure 4 | Locations of the fornix samples in a two features space. At the center we show the scatter plot of all samples belonging to the 6 different subjects, depending on their fraction of occupied area and effective local density. Samples belonging to the young subjects are represented in blue while samples of the old subjects are in red. These two features not only provide a good accuracy for age group discrimination, but also have a relatively large class scatter distance. Around the scatter plot we show the EM images of selected samples having particularly interesting values for the two features. These include samples at the interface between the age groups, samples having properties of the opposite age group, as well as outliers and prototypes (i.e. samples in the core of the class) for each age group.

Multiple feature analysis. In order to determine if the age group separation can be improved by considering larger subsets of features, we perform the same calculations for subsets of three features (see Supplementary Fig. S3 for details). We note that the maximum obtained accuracy value is 91%, compared to 90% for the pairwise case, which means that little information is added when including a third feature. In fact, considering the pairs of features that result in the highest accuracies in Fig. 3, the inclusion of another degree of freedom in the classification procedure, in the form of an additional feature, would only correct, at best, one misclassified sample. Since including one degree of freedom only to correctly classify one more sample does not provide much benefit to the procedure, the addition of a third feature is unnecessary.

Removing the limit imposed on the subset size, we use feature selection to search for the highest possible classification accuracy that one can achieve for any subset of features. Since the number of possible subsets is large, we use a best-first search algorithm⁴⁷ to search for the best subset considering all possible combinations of features (see Methods for details). We find that the highest accuracy we can achieve is 94%, for a subset containing 6 features (the features in the subset are: FracOccupArea, EffDensity, PearsonR1stShellAxon, Skewness2ndNNDist, Skewness3rdNNDist and StdDevElongation; see Supplementary Table S1 for a list of the abbreviated names given

to the features). Although this subset represents a small increase from the 90% maximum accuracy we obtain for pairs of features, we note that the extra 4 features only helped correct, at most, 3 misclassified samples. Therefore, we conclude that the best approach to characterize the aging of the rhesus monkey fornix is to consider the fornix samples in a 2D space, where the samples are characterized by pairs of features, as shown in Fig. 3.

Best age discriminant. When considering the two feature analysis, several pairs of features provide a high classification accuracy. In fact, as seen in Fig. 3, four pairs of features result in accuracies larger than 87.5% in the age group separation. In order to choose the pair more pertinent to the age group separation, we measure the scatter distance between the two classes⁴⁸, revealing the difference in the mean distance between classes for pairs of features with similar resulting accuracies. Thus, we look for a pair of features that provides both a high accuracy in the age group separation as well as a large distance between the classes. Following this criterion, the most relevant pair is formed by the fraction of occupied area and the effective local density. In Fig. 4 we show a scatter plot of the fornix samples in the 2D space formed by these two features, where we see that the young and old groups are well separated, with only 6 samples falling in the wrong age class. In the same figure we also show EM



images of samples located at interesting points in this 2D space. These include samples that are on the interface between the two classes, for which the fraction of occupied area or the effective local density alone would not be enough to correctly classify the samples, but that are classified in the right age group when the combination is used.

We note that, similarly to what is revealed in the single feature analysis, the combination of effective local density and fraction of occupied area, which has an accuracy of 90%, provides a better discrimination than the pair formed by the axon density and fraction of occupied area, which has an accuracy of only 81%. In fact, by replacing axon density with effective local density, a fraction of the misclassified samples migrate to their respective age group. Although these results were obtained using the K-nearest neighbors classifier, when considering other algorithms for classification, we observe that the combination of fraction of occupied area and effective local density commonly appears as one of the most accurate pair of features, validating our results (the other classifiers used were: Naive Bayes, Bayesian Network, Logistic, C4.5, Classification And Regression Tree - CART and Multilayer Perceptron).

Finally, to validate the use of the combination of fraction of occupied area and effective local density to distinguish between young and old, we apply an unsupervised classification algorithm⁴³. This classification is done via an agglomerative hierarchical clustering method using the Ward's criterion⁴⁹, where each sample is described solely by these two features. The final result of the hierarchical clustering is presented as a dendrogram in Fig. 5, where the distance between two groups indicates how large the variance of the merged group is, compared to the variance of each group separately. We see that this unsupervised classification method, which does not know anything about the classes *a priori*, assigns most of the samples in the correct class, with only 6 samples misclassified.

Discussion

In this paper, we present a systematic procedure to quantify differences between samples of nerve fibers taken from distinct groups. This method allows us to objectively identify the most appropriate set of features that best discriminates samples from distinct groups such as young and old, but is also applicable to other conditions (development, disease) as well as for comparisons across regions of interest in the same species and across different species. By using a feature selection algorithm, this method has the advantage of considering simultaneously an extensive list of features, without the risk of ignoring features that could have been overlooked. Furthermore, this measurements are performed automatically using an algorithm, instead of relying on conventional stereological techniques. The analysis in this method is done in an incremental manner, as we perform an exhaustive comparison of single, pairs and trios of features. We also perform a search for the set of features resulting in the highest accuracy, looking for the set that provides the most accurate group separation, regardless of the number of features in that set. In this analysis, we provide a comprehensive list of features that may unveil particularly interesting characteristics in three main aspects of nerve fibers: macroscopic features (characteristics of collections of axons), morphologic features (characteristics of individual axons) and structural features (characteristics of the relationship between different axons in each sample).

In the case of the specific application to the problem of aging we find that, when considering single features individually, density related features provide the best age separation. These include fraction of occupied area, effective local density, axon density, number of axons as well as third/second/first nearest neighbor mean distance. This is expected, since it is known that one of the most pronounced effects of aging in the fornix is the loss of myelinated axons⁴⁰.

We find that the fraction of occupied area provides the best single feature accuracy. This may be an indication that the actual axon area

through which information is transmitted might be more relevant to aging than the number of elements that can transmit it, i.e. the number of axons. The mean hexagonality index also provides a good separation between the samples, with the young group having a higher mean hexagonality index, which shows that the young group has a more regular or ordered angular structure, i.e. the axons surrounding each axon are distributed more evenly. This regularity is partially lost during aging, possibly due to temporally spaced loss of axons from different regions, with temporally spaced gliosis filling in.

Features related to the area and shape of the contour of the myelinated axon, i.e. morphologic features, are not statistically different for the two age groups, as reflected by the low accuracies attained when discriminating between age groups using those features individually. This indicates that myelinated axons are lost irrespectively of their area or shape. It would be of interest to apply the same analysis to the morphologic characteristics of the myelin sheaths, since it is known that they suffer alterations upon aging⁴⁰. We reserve this task for future work.

In the analysis with two features, an exhaustive search reveals that many pairs of features can provide accuracies near 90% (these include FracOccArea + EffDensity, MeanPerimeter + 2ndNNDist, MeanPerimeter + 3rdNNDist, FracOccArea + MeanHexagonality, FracOccArea + 2ndNNDist, FracOccArea + 3rdNNDist and MeanArea + 2ndNNDist). The most relevant pair is formed by the combination of fraction of occupied area and effective local density. This pair provides the highest accuracy in the age group separation as well as a large scatter distance between the two classes. The dendrogram in Fig. 5 validates our findings, showing that even if we do not know the classes *a priori* we still find a clear separation between samples. This pair is particularly interesting, since by replacing density with effective local density, one corrects the classification of a few samples that would have been assigned to the wrong class if using the pair formed by the density with the fraction of occupied area. We find that the effective local density is, at least in the study of aging, more sensitive to axon loss than simple density. This may be the case because it is sensitive to structural changes in the spatial relationships among axons, specifically the non-random presence of axon-free regions. In turn, this may reflect the fact that axons with similar properties (size, origin, etc) may be spatially clustered. If so, then the loss of axons in these clusters might reflect their selective vulnerability.

Furthermore, we note that the pairs of features selected, where one of them is a morphologic feature, provide a good age group separation due to differences in the geometric packing of the myelinated axons. The features of axons in young samples evidently have a more restricted range of values due to a tighter geometric packing, when compared to old samples (see Supplementary Fig. S4 for details).

When considering sets of three features, the highest accuracy attained is 91%, resulting in an improvement of the classification accuracy of only 1%. That is, the inclusion of another feature only corrects, at best, the classification of one sample. We also look for the set of features (regardless of their number) that can provide the highest possible accuracy using a heuristic search, and find that the discrimination can be improved to as much as 94%, but needing a set of 6 particular features. Therefore, two features are enough to characterize the separation between the two age groups. Note that this conclusion is valid for our case study, since other cases might need more (or less) features to confidently discriminate between the two different classes.

In summary, feature selection is a novel and useful methodology to choose the best set of features to describe the samples and detect differences between classes of myelinated axons. In this specific case study of aging, we conclude that only two features are enough to discriminate the samples into their proper age groups, namely the combination of fraction of occupied area and effective local density.

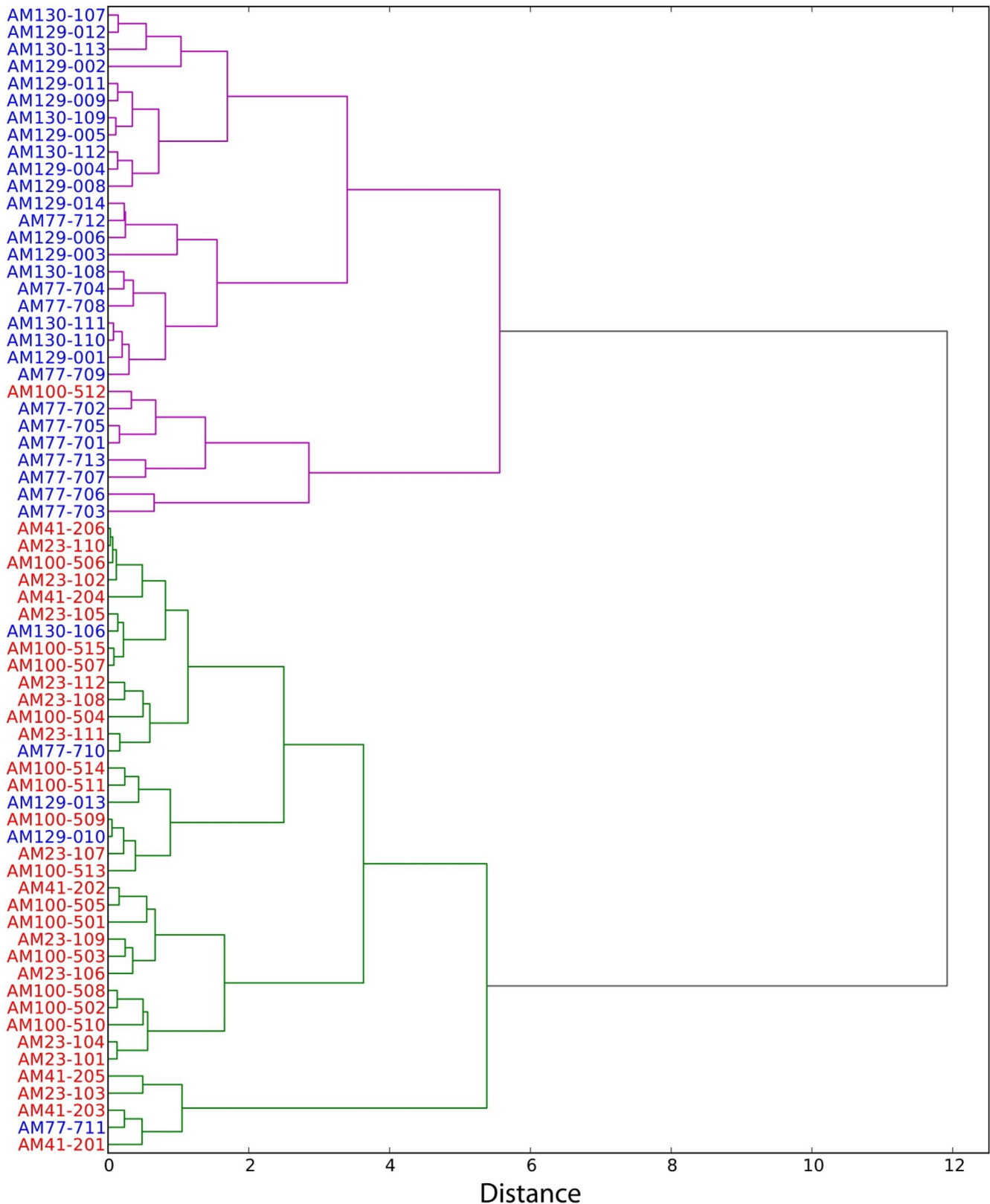


Figure 5 | Unsupervised class recognition. Dendrogram of an agglomerative hierarchical clustering algorithm⁴⁹ applied to all the fornix samples belonging to the 6 different subjects, described only by their fraction of occupied area and effective local density. Samples belonging to the young subjects are written in blue while samples of the old subjects are written in red. This algorithm is a systematic procedure to characterize the scatter plot seen in Fig. 4, without knowing the class of the samples, by grouping samples with similar characteristics together. This grouping (or clustering) is done by minimizing the variance of the cluster being merged, with that minimized variance shown as the distance of that merged group.



It is also important to note that the general methodology for feature detection presented here should be adaptable to enable the observation of changes in microscopic features caused by brain pathologies in neurodegenerative diseases like Alzheimer's Disease or by developmental disorders as well as pathological or developmental changes in other organ systems like the kidney or the pancreas. Moreover, our methodology can be potentially extended to feature detection techniques at more macroscopic scales such as MRI.

Methods

Electron micrographs. The experimental protocol for fixing the brain of the monkeys, the tissue preparation and microscopy setup are described in Ref. 40. Our measurements are performed on electron microscopic (EM) images of the fornix from six female specimens of rhesus monkey. Our cohort is divided into two groups: a young group, consisting of 3 subjects with ages ranging from 6 to 8 years; and an old group composed of the other 3 monkeys aged between 25 and 32 years old⁴⁰. In Table I we show the ages of the subjects used in this study.

A total of 67 electron micrographs (EM images) were available, of which 34 belonged to the young monkeys as follows: 13 EM images of animal AM77, 13 of animal AM129 and 6 of animal AM130. The other 33 were of the old monkeys, specifically 15 EM images of animal AM100, 12 of animal AM23 and 6 of animal AM41. These EM images were digitized with a resolution of 237 pixels per micron.

Axon recognition algorithm. The analysis of a large set of electron micrographs allowed the consideration of a large number of axons. This was required in order to perform a reliable spatial statistical analysis. Given the large number of myelinated axons, a semi-automated algorithm was developed in order to extract the coordinates and boundaries of the myelinated axons in an EM image, with the highest accuracy possible. In the EM images, the axon is very lightly stained, while the myelin sheath surrounding it is the darkest feature in the image. Our algorithm takes advantage of this contrast, and searches for continuous boundaries between a convex bright region and dark region. Note that the recognition procedure used here is original and it is based on well founded assumptions based on morphologic facts⁵⁰.

The steps involved in the recognition algorithm are:

1. Smooth the image using a Gaussian filter with a standard deviation of 6 pixels ($\sim 0.025 \mu\text{m}$).
2. Binarize the image using a threshold that varies for different samples, leaving only myelin/non-myelin objects.
3. Find the edges using a Canny edge detector⁵¹.
4. Extract contiguous regions with area larger than 5000 pixels ($\sim 0.089 \mu\text{m}^2$); this is a cut-off on the object's area implemented in order to eliminate false positives due to speckles.
5. Discard regions that have the ratio of perimeter to square root of area larger than 6.2; this is a cut-off on circularity.
6. Discard regions where the ratio between the standard deviation and the mean intensity of the pixels is larger than 0.5; this is a cut-off on uniformity.
7. Given the typical morphologic features of myelinated axons, discard regions with contours having three or more large values of convex curvature. We consider 0.03 pix^{-1} ($\sim 7.11 \mu\text{m}^{-1}$) as a large value of convex curvature; this eliminates false positives recognized in areas surrounded by the myelin sheaths of other axons.

Finally, in order to have the highest possible accuracy, a manual check is performed to eliminate remaining false positives and mark possibly missed myelinated axons.

Statistical tools. We describe in the following the quantities that we consider for our analyses. We determine the density and other quantities related to density (e.g. fraction of occupied area). This allows us to test our axon recognition procedure against the manual procedure used in Ref. 40, in which the density of myelinated axons was previously determined. We note that a reduction in the number of myelinated axons upon aging was previously observed^{40,52–54}.

Table I | List of the different monkeys used in this study with their corresponding animal number and age, see Ref. 40. All the monkeys considered were female

Animal number	Age (years)
AM 77	6
AM 129	7
AM 130	8
AM 100	25
AM 23	32
AM 41	32

Starting from the positions of the centroids R_i of the myelinated axons in our samples, we can also calculate different quantities related to their structure. The first one is the distribution of nearest neighbors distances which can be plotted, by building the histogram of the distances of the closest axon to every axon in the sample. Similarly, the n -th nearest neighbors distance distributions can be obtained considering the histogram of the distances to the n -th neighbor.

The n -th nearest distribution is also used to extract the effective local density of the sample. For a random set of points in an area (i.e. a random particle distribution described by a homogeneous Poisson point process) with an uniform mean density $\rho > 0$, the mean distance of the n -th nearest neighbor $\langle r_n \rangle$ is given by^{55–57}

$$\langle r_n \rangle = \frac{1}{\sqrt{\pi\rho}} \frac{\Gamma(n+1/2)}{\Gamma(n)}, \quad (1)$$

where $\Gamma(n)$ is the Gamma function (see Supplementary Note 1 for details). For large n , this exact formula can be approximated by

$$\langle r_n \rangle \approx \sqrt{\frac{n}{\pi\rho}}, \quad (2)$$

with a relative error smaller than 1.6% for $n \geq 8$. Although Eq. (2) was derived for a random set of points, this power law behavior for sufficiently large n is also observed in our samples.

Considering Eq. (2), we expect the n -th nearest neighbor distance to increase with the rank n , according to a power law. Therefore, we can estimate the density of the system by fitting a line for large n , on the log-log plot of the n -th nearest neighbor distance versus the rank n . We fit a line for the points $n = 8, 9, \dots, 15$ and calculate this "regression" density (named effective local density) from the intercept of the linear regression. In Fig. 6(a) and 6(b) we show illustrations of two samples with equal densities but different effective local densities, due to the presence of a large axon-free region. The calculation of the effective local density is illustrated in Fig. 6(c) for the two particular samples shown in Fig. 6(a) and 6(b). We note that the effective local density is calculated from the nearest neighbor distance behavior and, as such, it provides a measure of the local density around axons.

We also construct the Voronoi tessellation of the embedding space as determined by the spatial distribution of the centroids of the myelinated axons and study the statistical and morphologic properties of the Voronoi cells of area \mathcal{A}_i , see Fig. 1. In the Voronoi tessellation, given a discrete spatial distribution of points R_i , in our case identified with the centroids of the axons, the Voronoi cell associated to the point R_i is defined as the set of points closer to the point R_i than to any other point of the distribution. In practice, all the space is completely divided into convex non-overlapping polygons built around each centroid, where each polygon edge bisects the line segments joining the respective centroid to its neighbors (i.e. Delaunay triangulation). An interesting concept that can be investigated starting from the position of the axon R_i and the Voronoi tessellation of the embedding space is the polygonality index (PI), first defined in Ref. 44. The PI measures how close is the symmetry of the sample to a well known ordered structure. In this case we compare our samples to a regular triangular lattice, which corresponds to the 2D lattice of equal disks with the highest density. Its Voronoi tessellation is a hexagonal tiling. Following Ref. 44, for each point R_i (centroid of the axon) in the tessellation we measure the adjacent angles between line segments joining R_i with the centers of neighboring Voronoi cells, i.e. the ones with which it shares one side of the cell, which we label $1, 2, \dots, \mathcal{N}_i$. The angles are labeled $\alpha_1, \alpha_2, \dots, \alpha_{\mathcal{N}_i}$. Thus for each axon's centroid R_i we define the quantity

$$Q_i = \sum_{j=1}^{\mathcal{N}_i} |\alpha_j - \beta|. \quad (3)$$

Choosing $\beta = \pi/3$ the comparison is made with a triangular lattice. In this case the PI is called also hexagonality index (HI) and it is defined for each point R_i by

$$\Delta_i = \frac{1}{Q_i + 1}. \quad (4)$$

It follows that the mean HI is 1 for a perfect triangular lattice and it approaches 0 as the points deviate more and more from a triangular lattice (the typical mean HI for random samples is ~ 0.30).

In order to study in detail the size of the myelinated axons and their relation with the environment constituted by the other myelinated axons we consider the areas of the axons and the correlation between them, as well as the area of the respective Voronoi cells and the correlations between them. For each axon/Voronoi cell we calculate the average of the areas of the neighboring axon $\langle \mathcal{A}'_i \rangle$ /Voronoi cell $\langle \mathcal{A}'_i \rangle$. This is plotted against the area of the central axon \mathcal{A}_i /cell \mathcal{A}_i in consideration and a linear regression is performed. The procedure is then repeated for the second shell of neighbors, for which the average area is denoted $\langle \mathcal{A}''_i \rangle / \langle \mathcal{A}''_i \rangle$. We define the first shell of neighbors of an axon as the set of axons contained in neighboring Voronoi cells. Similarly, the second shell is defined by the set of axons contained in the neighboring Voronoi cells of the first neighbors themselves.

Besides the positions of the axons, the shape of their cross sections may also change with age. These morphologic changes may be due to alterations in the myelin sheath^{40,50} or even due to the compression of axons caused by the close packing of neighboring structures. Here we use six shape features to unveil if there are any differences between myelinated axons in rhesus monkeys with distinct ages. They are:

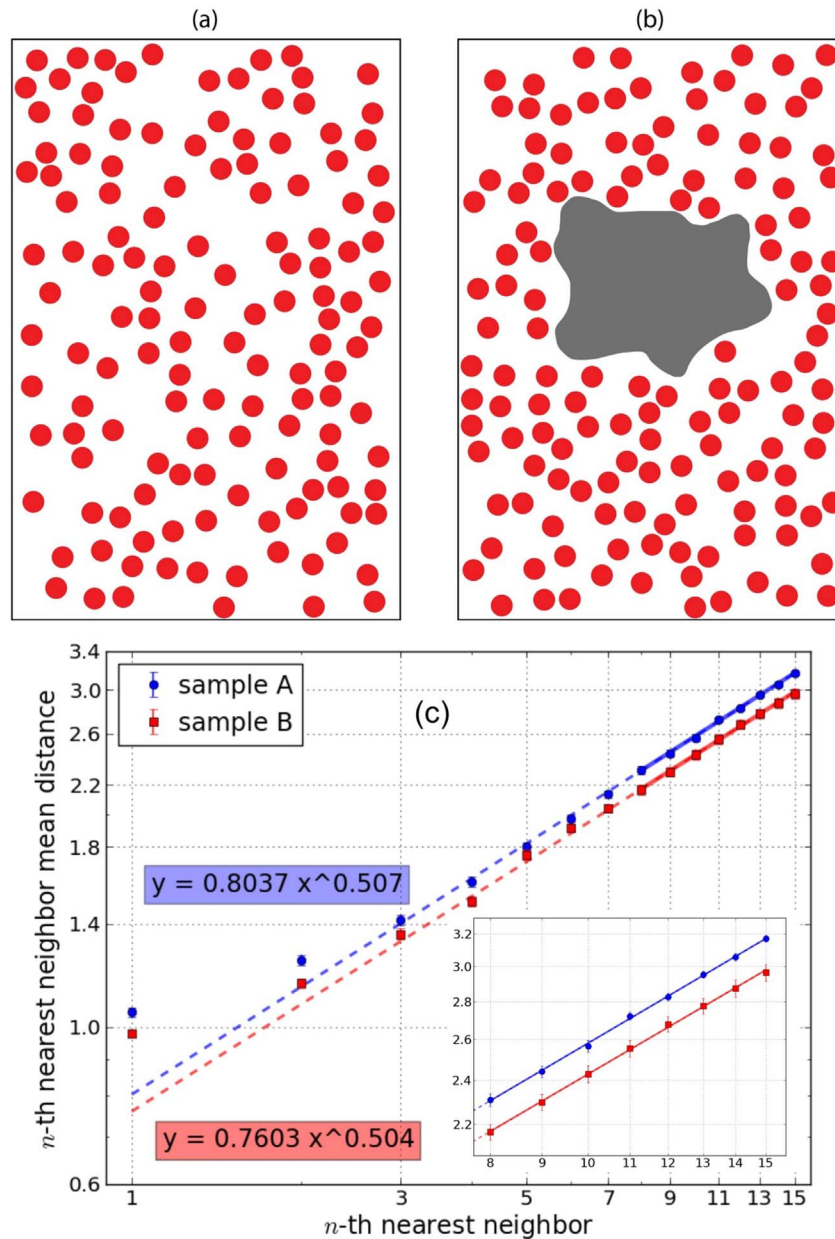


Figure 6 | Effective local density calculation. (a) and (b) Illustrations of samples with equal densities ($\rho = 0.50$) but different effective local densities. The sample in figure (a) has an effective local density value of 0.49, similar to its density value, while the sample in figure (b) has a higher effective local density value of 0.55, resulting from the axons (in red) being closer together due to the presence of a large axon-free region (in grey). (c) Plot of the average n -th nearest neighbor distance as a function of the rank n for the two samples highlighted in figures (a) and (b). The full lines are the log-log linear regressions for the points $n \geq 8$ (see inset figure for a zoomed plot), extrapolated (dashed lines) to $n = 1$, and with corresponding equations in the matching colored boxes. The effective local density is calculated from the linear regression according to Eq. 2.

perimeter, elongation, circularity, diameter, mean curvature and bending energy as defined in Ref. 41.

We present in Supplementary Table S1 all the features collected together with a short description where appropriate.

Classification of samples and feature selection. In many biological studies, it is common to quantify the relevance of a feature by using statistical tests. A typical approach is to obtain the probability distributions for the feature in different types of samples and then compare the statistical significance of the difference between their mean. Here we take a different approach, we use the K-nearest neighbors classifier with $K = 3$ to find the accuracy of the measured features when used to classify the samples in either of two classes (i.e. the young or old class). This approach has the clear advantage that the obtained value can be easily interpreted. For example, an 84% of accuracy for a feature means that in 84% of the cases the feature will correctly identify if the sample comes from a young or old subject. To perform such a task, we decided to use the K-nearest neighbors classifier mainly because of its simplicity and its straightforward interpretation in terms of class assignment (i.e. young or old).

Also, this classifier has only one parameter: the number of neighbors. We repeat the analysis for other classifiers and find little variation in the accuracy values obtained using the K-nearest neighbors classifier (the other classifiers used were: Naive Bayes, Bayesian Network, Logistic, C4.5, Classification And Regression Tree - CART and Multilayer Perceptron).

Having measured a given set of features for all the samples, the classification procedure is performed according to the following steps. We first standardize the values for each feature by subtracting their mean and dividing by their standard deviation. Then, we randomly divide the 67 samples in five different groups (henceforth called *folds*), with the only restriction that each fold has, as far as possible, the same percentage of samples from each class as the complete set. We use the samples contained in four of the folds to *train* the classifier, i.e. to define regions in the feature space that should be associated with each class. Next, we use the fold that was excluded in the training to *validate* the classifier, that is to say, we test how many samples in the excluded fold are correctly assigned to the class it originally belongs. For the 3-nearest neighbor classifier, the class assignment is done by determining, for each sample of the excluded fold, the first three neighbor samples belonging to the



training folds. We then assign the sample to the class of the majority of these three. We repeat this procedure four more times, each time excluding a different fold from the training in order to use it for validation. In the end, we have, for each fold, the number of samples correctly assigned to its class. This in turn can be represented by a percentage of correctly classified samples, which we call the accuracy for the given set of features. This process is called a *stratified cross-validation* and it is the most common approach to estimate the performance of a classifier when the number of samples is small. Finally, since there might be some variation in the accuracy value depending on how the samples were divided, we run the cross-validation procedure five times, and take the mean over the five runs. We note that increasing the number of runs does not reduce the variance of the accuracy values found, meaning that five runs is enough to obtain a reliable value for accuracy.

In order to find the set of features that can provide the best discrimination between the young and old age groups, we employ the feature selection technique^{36–38,58}. The ideal situation would be to calculate the accuracy for every single set of features, and keep only the set providing the largest accuracy. But since the number of possible sets grows exponentially with the number of features, a heuristic is needed in order to search for the optimal subset. The heuristic we use, known as best-first search⁴⁷, starts with the empty set and successively adds the features that provide the largest increase in accuracy. When the search gets stuck in a local optimum, i.e. when there are no more feature additions that can improve the accuracy, it jumps to a previously visited set and it adds the feature providing the second best accuracy found. The search continues until the number of allowed jumps is attained. In our case we allow 50000 jumps.

- Bastiani, M. J., Harrelson, A. L., Snow, P. M. & Goodman, C. S. Expression of fasciclin I and II glycoproteins on subsets of axon pathways during neuronal development in the grasshopper. *Cell* **48**, 745–755 (1987).
- Cremer, H., Chazal, G., Goridis, C. & Represa, A. NCAM Is Essential for Axonal Growth and Fasciculation in the Hippocampus. *Mol. Cell. Neurosci.* **8**, 323–335 (1997).
- Xue, Y. & Honig, M. G. Ultrastructural observations on the expression of axonin-1: Implications for the fasciculation of sensory axons during axonal outgrowth into the chick hindlimb. *J. Comp. Neurol.* **408**, 299–317 (1999).
- Wilson, S. W., Ross, L. S., Parrett, T. & Easter, S. S. Jr. The development of a simple scaffold of axon tracts in the brain of the embryonic zebrafish, *Brachydanio rerio*. *Development* **108**, 121–145 (1990).
- Scheiffele, P., Fan, J., Choih, J., Fetter, R. & Serafini, T. Neuroligin Expressed in Nonneuronal Cells Triggers Presynaptic Development in Contacting Axons. *Cell* **101**, 657–669 (2000).
- Sánchez, I., Hassinger, L., Paskevich, P. A., Shine, H. D. & Nixon, R. A. Oligodendroglia regulate the regional expansion of axon caliber and local accumulation of neurofilaments during development independently of myelin formation. *J. Neurosci.* **16**, 5095–5105 (1996).
- Gillespie, M. J. & Stein, R. B. The relationship between axon diameter, myelin thickness and conduction velocity during atrophy of mammalian peripheral nerves. *Brain Res.* **259**, 41–56 (1983).
- Chomiak, T. & Hu, B. What is the optimal value of the g-ratio for myelinated fibers in the rat CNS? A theoretical approach. *PLoS ONE* **4**, e7754 (2009).
- Ceballos, D., Cuadras, J., Verdú, E. & Navarro, X. Morphometric and ultrastructural changes with ageing in mouse peripheral nerve. *J. Anat.* **195**, 563–576 (1999).
- Elder, G. A., Friedrich, V. L. Jr., Margita, A. & Lazzarini, R. A. Age-related atrophy of motor axons in mice deficient in the mid-sized neurofilament subunit. *J. Cell Biol.* **146**, 181–192 (1999).
- Nielsen, K. & Peters, A. The effects of aging on the frequency of nerve fibers in rhesus monkey striate cortex. *Neurobiol. Aging* **21**, 621–628 (2000).
- Pfefferbaum, A. *et al.* Age-related decline in brain white matter anisotropy measured with spatially corrected echo-planar diffusion tensor imaging. *Magn. Reson. Med.* **44**, 259–268 (2000).
- Cruz, L. *et al.* Age-Related Reduction in Microcolumnar Structure Correlates with Cognitive Decline in Ventral but Not Dorsal Area 46 of the Rhesus Monkey. *Neuroscience* **158**, 1509–1520 (2009).
- Cruz, L. *et al.* Age-Related Reduction in Microcolumnar Structure in Area 46 of the Rhesus Monkey Correlates with Behavioral Decline. *Proc. Natl. Acad. Sci. USA* **101**, 15846–15851 (2004).
- Song, S.-K. *et al.* Diffusion tensor imaging detects and differentiates axon and myelin degeneration in mouse optic nerve after retinal ischemia. *NeuroImage* **20**, 1714–1722 (2003).
- Hinton, D. R., Sadun, A. A., Blanks, J. C. & Miller, C. A. Optic-nerve degeneration in Alzheimer's disease. *N. Engl. J. Med.* **315**, 485–487 (1986).
- Moreno, R. D., Inestrosa, N. C., Culwell, A. R. & Alvarez, J. Sprouting and abnormal contacts of nonmyelinated axons, and deposition of extracellular material induced by the amyloid precursor protein (APP) and other protease inhibitors. *Brain Res.* **718**, 13–24 (1996).
- Schlaepfer, W. W. & Bunge, R. P. Effects of calcium ion concentration on the degeneration of amputated axons in tissue culture. *J. Cell Biol.* **59**, 456–470 (1973).
- Blight, A. R. & Decrescito, V. Morphometric analysis of experimental spinal cord injury in the cat: the relation of injury intensity to survival of myelinated axons. *Neuroscience* **19**, 321–341 (1986).
- Xu, X. M., Guénard, V., Kleitman, N., Aebischer, P. & Bunge, M. B. A combination of BDNF and NT-3 promotes supraspinal axonal regeneration into Schwann cell grafts in adult rat thoracic spinal cord. *Exp. Neurol.* **134**, 261–272 (1995).
- Aguayo, A. J., David, S. & Bray, G. M. Influences of the glial environment on the elongation of axons after injury: transplantation studies in adult rodents. *J. Exp. Biol.* **95**, 231–240 (1981).
- Kao, C. C., Chang, L. W. & Bloodworth, J. M. B. Jr. Axonal regeneration across transected mammalian spinal cords: an electron microscopic study of delayed microsurgical nerve grafting. *Exp. Neurol.* **54**, 591–615 (1977).
- Noseworthy, J. H., Lucchinetti, C., Rodriguez, M. & Weinshenker, B. G. Multiple sclerosis. *N. Engl. J. Med.* **343**, 938–952 (2000).
- Raine, C. S., Hummelgard, A., Swanson, E. & Bornstein, M. B. Multiple sclerosis: serum-induced demyelination in vitro. A light and electron microscope study. *J. Neurol. Sci.* **20**, 127–148 (1973).
- Trapp, B. D. *et al.* Axonal transection in the lesions of multiple sclerosis. *N. Engl. J. Med.* **338**, 278–285 (1998).
- Wujek, J. R. *et al.* Axon loss in the spinal cord determines permanent neurological disability in an animal model of multiple sclerosis. *J. Neuropathol. Exp. Neurol.* **61**, 23–32 (2002).
- Stankovic, R. K., Shingde, M. & Cullen, K. M. The experimental toxicology of metallic mercury on the murine peripheral motor system: a novel method of assessing axon calibre spectra using the phrenic nerve. *J. Neurosci. Methods* **147**, 114–125 (2005).
- Stankovic, R. Atrophy of large myelinated motor axons and declining muscle grip strength following mercury vapor inhalation in mice. *Inhal. Toxicol.* **18**, 57–69 (2006).
- Ayranci, E., Altunkaynak, B. Z., Aktaş, A., Rağbetli, M.Ç. & Kaplan, S. Prenatal exposure of diclofenac sodium affects morphology but not axon number of the median nerve of rats. *Folia Neuropathol.* **51**, 76–86 (2013).
- Giese, K. P., Martini, R., Lemke, G., Soriano, P. & Schachner, M. Mouse P0 gene disruption leads to hypomyelination, abnormal expression of recognition molecules, and degeneration of myelin and axons. *Cell* **71**, 565–576 (1992).
- Readhead, C. *et al.* Expression of a myelin basic protein gene in transgenic shiverer mice: correction of the dysmyelinating phenotype. *Cell* **48**, 703–712 (1987).
- Craddock, R. C., Holtzheimer, P. E., Hu, X. P. & Mayberg, H. S. Disease state prediction from resting state functional connectivity. *Magn. Reson. Med.* **62**, 1619–1628 (2009).
- Millán, J., Franzé, M., Mouriño, J., Cincotti, F. & Babiloni, F. Relevant EEG features for the classification of spontaneous motor-related tasks. *Biol. Cybern.* **86**, 89–95 (2002).
- Cui, Y. *et al.* Identification of conversion from mild cognitive impairment to Alzheimer's disease using multivariate predictors. *PLoS ONE* **6**, e21896 (2011).
- Celebi, M. E. *et al.* A methodological approach to the classification of dermoscopy images. *Comput. Med. Imaging Graph.* **31**, 362–373 (2007).
- Guyon, I. & Elisseeff, A. An Introduction to Variable and Feature Selection. *J. Mach. Learn. Res.* **3**, 1157–1182 (2003).
- Blum, A. L. & Langley, P. Selection of relevant features and examples in machine learning. *Artif. Intell.* **97**, 245–271 (1997).
- Saeyns, Y., Inza, I. & Larrañaga, P. A review of feature selection techniques in bioinformatics. *Bioinformatics* **23**, 2507–2517 (2007).
- Gaffan, D. Recognition impaired and association intact in the memory of monkeys after transection of the fornix. *J. Comp. Physiol. Psychol.* **86**, 1100–1109 (1974).
- Peters, A., Sethares, C. & Moss, M. B. How the Primate Fornix Is Affected by Age. *J. Comp. Neurol.* **518**, 3962–3980 (2010).
- Costa, L. da F. & Cesar, R. M. Jr. *Shape Classification and Analysis: Theory and Practice*. (CRC Press, 2nd edition, Boca Raton, FL, USA, 2009).
- Witten, I. H., Frank, E. & Hall, M. A. *Data Mining: Practical Machine Learning Tools and Techniques*. (Morgan Kaufmann, 3rd edition, San Francisco, CA, USA, 2011).
- Duda, R. O., Hart, P. E. & Stork, D. G. *Pattern Classification*. (Wiley-Interscience, 2nd edition, 2000).
- Costa, L. da F., Rocha, F. & de Lima, S. M. A. Characterizing polygonality in biological structures. *Phys. Rev. E* **73**, 011913 (2006).
- Bellman, R. *Dynamic Programming*. (Dover Publications, Mineola, NY, USA, 2003).
- Jolliffe, I. T. *Principal component analysis*. (Springer, 2nd edition, New York, NY, USA, 2002).
- Russell, S. & Norvig, P. *Artificial Intelligence: A Modern Approach*. (Prentice Hall, 3rd edition, Upper Saddle River, NJ, USA, 2009).
- Fukunaga, K. *Introduction to Statistical Pattern Recognition*. (Academic Press, 2nd edition, San Diego, CA, USA, 1990).
- Everitt, B. S., Landau, S., Leese, M. & Stahl, D. *Cluster analysis*. (Wiley, 5th edition, Chichester, West Sussex, UK, 2011).
- Peters, A. & Sethares, C. F. *The fine structure of the aging brain*. URL: <http://www.bu.edu/agingbrain>, retrieved on July 8th, 2013.
- Canny, J. A Computational Approach To Edge Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **8**, 679–698 (1986).
- Pakkenberg, B. & Gundersen, H. J. G. Neocortical neuron numbers in humans: effect of sex and age. *J. Comp. Neurol.* **384**, 312–320 (1997).



53. Tang, Y., Nyengaard, J. R., Pakkenberg, B. & Gundersen, H. J. G. Age-induced white matter changes in the human brain: a stereological investigation. *Neurobiol. Aging* **18**, 609–615 (1997).
54. Marner, L., Nyengaard, J. R., Tang, Y. & Pakkenberg, B. Marked loss of myelinated nerve fibers in the human brain with age. *J. Comp. Neurol.* **462**, 144–152 (2003).
55. Kingman, J. F. C. *Poisson Processes*. (Oxford University Press, New York, USA, 1993).
56. Percus, A. G. & Martin, O. C. Scaling universalities of kth-nearest neighbor distances on closed manifolds. *Adv. Appl. Math.* **21**, 424–436 (1998).
57. Olver, F. W. J., Lozier, D. W., Boisvert, R. F. & Clark, C. W. *NIST Handbook of Mathematical Functions*. (Cambridge University Press, New York, USA, 2010).
58. Kohavi, R. & John, G. H. Wrappers for feature subset selection. *Artif. Intell.* **97**, 273–324 (1997).

Acknowledgments

We thank A. Peters and C. F. Sethares for having provided the electron micrographs used in this work and for illuminating discussions. This work was supported by FAPESP grant 2011/22639-8 (to C.H.C.), FAPESP grant 11/50761-2 and CNPQ grant 304351/2009-1 (to L. da F.C.), NIH grant 5R01AG021133-08 and NSF grant PHY-0855453 (to J.R.S., D.C., W.M., C.C., D.L.R., and H.E.S.). J.R.S. acknowledges support from Fundação para a Ciência e a Tecnologia, Portugal, through a doctoral degree fellowship (SFRH/BD/32391/2006). A.G. acknowledges support from the EU IP Project “MULTIPLEX” contract no. 317532

and from the CNR Project of National Interest “CrisisLab” financed by the Italian Government.

Author contributions

J.R.S., C.H.C., D.C. and L. da F.C. wrote the main manuscript text. J.R.S. and C.H.C. prepared Fig. 1–6. All authors discussed the results and reviewed the manuscript.

Additional information

Supplementary information accompanies this paper at <http://www.nature.com/scientificreports>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Comin, C.H. *et al.* Statistical physics approach to quantifying differences in myelinated nerve fibers. *Sci. Rep.* **4**, 4511; DOI:10.1038/srep04511 (2014).



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. The images in this article are included in the article's Creative Commons license, unless indicated otherwise in the image credit; if the image is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the image. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>