Received: 20 September 2024 | Revised: 14 April 2025 | Accepted: 16 April 2025

DOI: 10.1002/dad2.70118

Alzheimer's & Dementia
Diagnosis, Assessment & Disease Monitoring

RESEARCH ARTICLE

Fully automated MRI-based analysis of the locus coeruleus in aging and Alzheimer's disease dementia using ELSI-Net

Max Dünnwald[1,2] | Friedrich Krohn[3] | Alessandro Sciarra[1] | Mousumi Sarkar[3] | Anja Schneider[4,5] | Klaus Fliessbach[4,5] | Okka Kimmich[4] | Frank Jessen[4,6,7] | Ayda Rostamzadeh[6] | Wenzel Glanz[8] | Enise I. Incesoy[3,8,9] | Stefan Teipel[10,11] | Ingo Kilimann[10,11] | Doreen Goerss[10,11] | Annika Spottke[4,12] | Johanna Brustkern[4] | Michael T. Heneka[13] | Frederic Brosseron[4] | Falk Lüsebrink[8] | Dorothea Hämmerer[3,8,14,15,16] | Emrah Düzel[3,8,15,16] | Klaus Tönnies[2] | Steffen Oeltze-Jafra[17] | Matthew J. Betts[3,8,16]

[1]Department of Neurology, Otto von Guericke University Magdeburg (OvGU), Magdeburg, Germany

[2]Faculty of Computer Science, OvGU, Magdeburg, Germany

[3]Institute of Cognitive Neurology and Dementia Research, OvGU, Magdeburg, Germany

[4]German Center for Neurodegenerative Diseases (DZNE), Bonn, Bonn, Germany

[5]Department for Cognitive Disorders and Old Age Psychiatry, University Hospital Bonn, Bonn, Germany

[6]Department of Psychiatry, Medical Faculty, University of Cologne, Cologne, Germany

[7]Excellence Cluster on Cellular Stress Responses in Aging-Associated Diseases (CECAD), University of Cologne, Cologne, Germany

[8]German Center for Neurodegenerative Diseases (DZNE), Magdeburg, Germany

[9]Department for Psychiatry and Psychotherapy, University Clinic Magdeburg, Magdeburg, Germany

[10]German Center for Neurodegenerative Diseases (DZNE), Rostock, Germany

[11]Department of Psychosomatic Medicine, Rostock University Medical Center, Rostock, Germany

[12]Department of Neurology, University of Bonn, Bonn, Germany

[13]Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, Esch-sur-Alzette, Luxembourg

[14]Department of Psychology, University of Innsbruck, Innsbruck, Austria

[15]Institute of Cognitive Neuroscience, University College London, London, UK

[16]Center for Behavioral Brain Sciences, Universitätsplatz 2, Magdeburg, Germany

[17]Peter L. Reichertz Institute for Medical Informatics, Hannover Medical School, Hannover, Germany

Correspondence
Max Dünnwald, Universitätsklinik für Neurologie, Leipziger Strasse 44, 39120 Magdeburg, Germany.
Email: max.duennwald@med.ovgu.de

Funding information
Federal State of Saxony-Anhalt, Germany, Grant/Award Number: I88; Deutsche

Abstract

INTRODUCTION: The locus coeruleus (LC) is linked to the development and pathophysiology of neurodegenerative diseases such as Alzheimer's disease (AD). Magnetic resonance imaging–based LC features have shown potential to assess LC integrity in vivo.

Forschungsgemeinschaft, Grant/Award Number: 425899994; DFG, Sonderforschungsbereich (SFB), Grant/Award Numbers: 1436, 1436-A08, 1315, 1315-B06; Bundesministerium für Bildung und Forschung, Grant/Award Number: 01ED2102B; EU Joint Programme—Neurodegenerative Disease Research (JPND); Alzheimer's Research UK (ARUK), Grant/Award Number: SRF2018B-004; National Institutes of Health, Grant/Award Number: R01MH126971

METHODS: We present a deep learning–based LC segmentation and feature extraction method called Ensemble-based Locus Coeruleus Segmentation Network (ELSI-Net) and apply it to healthy aging and AD dementia datasets. Agreement to expert raters and previously published LC atlases were assessed. We aimed to reproduce previously reported differences in LC integrity in aging and AD dementia and correlate extracted features to cerebrospinal fluid (CSF) biomarkers of AD pathology.

RESULTS: ELSI-Net demonstrated high agreement to expert raters and published atlases. Previously reported group differences in LC integrity were detected and correlations to CSF biomarkers were found.

DISCUSSION: Although we found excellent performance, further evaluations on more diverse datasets from clinical cohorts are required for a conclusive assessment of ELSI-Net's general applicability.

KEYWORDS
biomarker, deep learning, locus coeruleus, magnetic resonance imaging, segmentation

Highlights

- We provide a thorough evaluation of a fully automatic locus coeruleus (LC) segmentation method termed Ensemble-based Locus Coeruleus Segmentation Network (ELSI-Net) in aging and Alzheimer's disease (AD) dementia.
- ELSI-Net outperforms previous work and shows high agreement with manual ratings and previously published LC atlases.
- ELSI-Net replicates previously shown LC group differences in aging and AD.
- ELSI-Net's LC mask volume correlates with cerebrospinal fluid biomarkers of AD pathology.

1 | BACKGROUND

The locus coeruleus (LC) is a small brainstem nucleus with ≈ 50.000 pigmented neurons,[1] but projects to almost all major brain regions and is the primary source of noradrenaline in the brain. It has been identified as one of the earliest brain structures to be affected in Alzheimer's disease (AD)[2] and has been linked to cognitive decline in healthy aging and progression of AD.[3] Owing to the early tau aggregation in the LC, assessing the integrity of the LC using structural magnetic resonance imaging (MRI) may be a suitable tool for obtaining pathophysiological insights in vivo.[4,5] It presents an opportunity to gain insights into cognitive and behavioral symptoms instrumental for developing effective treatments,[1] improve our understanding of AD pathogenesis, and facilitate the development of disease-modifying noradrenergic drugs.[6]

Specific MRI techniques permit the in vivo visualization of the LC by exploiting among others the magnetic properties of its neuromelanin pigmented neurons, although the exact contrast mechanisms remain unclear. We refer the reader to Betts et al.[4] and Trujillo et al.,[7] where more nuanced discussions concerning the LC MRI signal can be found. The hyperintense regions appearing in the MRI acquisitions were shown to correspond to LC properties observed in *post mortem* studies, that is, with respect to anatomical position and dimensions and LC cell density[8] and to correspond with age-related increases in neuromelanin.[9] Furthermore, associations between LC MRI contrast and AD biomarkers[10] as well as cognitive decline in health and disease have also been observed[5] suggesting these MRI techniques may be suitable for assessing LC integrity.

A reliable extraction of in vivo LC MRI biomarkers requires a robust segmentation approach. The small size and cylindrical shape of the LC (≈ 2 mm in diameter[11]) together with the comparatively coarse resolution of MRI acquisitions limits the reliability of LC integrity measures. Although reasonable compromises can be found,[4] LC MRI acquisitions are characterized by low signal-to-noise ratios and ambiguous structural boundaries, posing challenges for segmentation approaches. This is evident by particularly low inter-rater agreements between manual raters of 0.499,[12] 0.54 to 0.64,[13] and 0.67 Dice similarity coefficient (DSC)[14] as reported in the literature.

Initially, a broad majority of studies investigating the LC using specific MRI relied on manual segmentations carried out by expert raters.[15] They require considerable amounts of manual labor and the time of trained experts. In recent years, (semi-)automatic approaches that reduce the need for manual intervention have become more

common than purely manual segmentation. Many studies have used atlas- or template-based segmentation approaches.[8,12,16,17] Another type of algorithm uses the template registration as a first step to obtain a search space on which further operations, usually related to peak intensity extraction of certain rostrocaudal subparts or slices of the LC, are applied.[18,19] Although only a subset of the LC voxels are obtained, promising results using intensity-based features have been shown with respect to reproducing known cohort effects, such as structural degeneration,[19] associations to cognition, or both.[20] These methods require manual corrections in some cases as the performance relies on successful registration with a high precision. Through the remaining manual steps, rater bias may still influence the segmentation of these methods. Automatic LC segmentation algorithms, such as the approach presented here, may facilitate using LC imaging in large-scale studies by removing the need for multiple experts to invest time and manual labor on LC segmentation and introduce more objectivity by removing human bias. Our group was the first to develop a deep learning–based LC segmentation approach.[21] Its convolutional neural networks process the MRI acquisitions inherently faster than registration-based and manual methods and have shown higher objectivity by incorporating multiple experts' knowledge into the training.[14] Our pipeline comprises all steps from the LC segmentation to the reference region generation and the feature extraction. We do not manually correct the segmentations prior to our analyses and evaluations.

In the work presented here, we show an improved fully automatic LC segmentation pipeline that further increases the performance compared to our previous approach and assess its practical usability in various ways on subjects of aging and AD dementia.

RESEARCH IN CONTEXT

1. **Systematic review**: The authors reviewed the literature using traditional sources (e.g., PubMed, Google Scholar). Although there are several publications introducing semi-automatic methods for locus coeruleus (LC) segmentation, the application of deep learning methods is underexplored. To the best of our knowledge, this is the first paper using a deep learning–based approach for automated LC segmentation in Alzheimer's disease (AD) dementia.
2. **Interpretation**: Our work introduces and evaluates an improved automatic, deep learning–based LC segmentation and analysis approach. The results suggest a very high potential for practical applicability, for example, in large-scale clinical studies for neurodegenerative diseases.
3. **Future directions**: Ensemble-based Locus Coeruleus Segmentation Network (ELSI-Net) can be used to assess LC integrity on large- or small-scale studies in AD dementia. To ensure robust performance, ELSI-Net should be further evaluated in larger, more diverse datasets comprising varying LC magnetic resonance imaging protocols and clinical populations.

2 | METHODS

We apply an improved version of a recently proposed fully automatic LC analysis method[14] to two different datasets comprising MRI acquisitions from healthy aging and AD dementia. The results are compared to manual expert segmentations and their features.

2.1 | Datasets

The two datasets share the same acquisition protocol: They comprise T_1-weighted fast low angle shot (FLASH) 3 Tesla MRI scans (5.56 ms echo time, 20 ms repetition time, 23° flip angle, 130 Hz/pixel bandwidth, 7/8 partial Fourier, 13:50 minute scan time) with an isotropic resolution of 0.75 mm. The image data were upsampled using a sinc filter to achieve an isotropic voxel size of 0.375 mm and then bias field–corrected as previously described.[9]

Our Healthy Aging Dataset (HAD)[9] comprised 82 healthy subjects. There were 25 younger (22–30 years old; 13 male) and 57 older subjects (61–80 years old, 19 male). We also analyzed the T_1-weighted FLASH MRI data from the DZNE Longitudinal Cognitive Impairment and Dementia study (DELCODE).[22] This comprised 188 subjects: 68 healthy elderly adults, 22 relatives of individuals with AD, 61 subjects with subjective cognitive decline (SCD), 26 with mild cognitive impairment (MCI), and 11 with AD dementia. For our experiments, we combine the healthy elderly adults and relatives of AD subjects into one group of healthy controls.

The age of the subjects ranges from 60 to 87 years (69 on average, 102 females). The LC MRI data have been acquired at four different sites across Germany: Magdeburg, Rostock, Bonn, and Berlin. We refer the reader to Betts et al.[9] (for HAD) and Jessen et al.[22] for further details on the full DELCODE cohort.

2.2 | Segmentation methods

2.2.1 | Manual expert segmentation

Both of our trained expert raters (M. Betts, referred to as Rater 1 [R1] and M. Sarkar, referred to as Rater 2 [R2]) manually segmented all LCs in the HAD. For the DELCODE dataset, R1 manually segmented 108 subjects and R2 segmented the remaining 80 subjects.

The raters delineated the LC using ITK-SNAP[23] as previously described.[9] Briefly, the segmentation was performed on the axial slices starting at the most dorsal to ventral portion of the LC while limiting the rostrocaudal space for segmentation to slices between the inferior boundary of the interpeduncular fossa at the level of the inferior colliculus and the superior cerebellar peduncle (for reference see Betts et al.[9]).

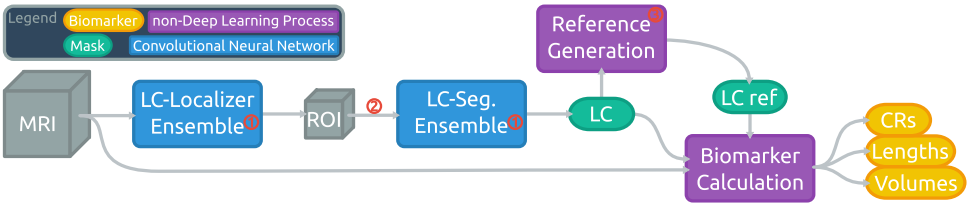

FIGURE 1 Schematic illustration of the Ensemble-based Locus Coeruleus Segmentation Network (ELSI-Net) pipeline for automated LC analysis. The encircled red numbers indicate the three main changes compared to previous work[14]: (1) ensemble networks with majority vote, (2) additional intensity normalization prior to the segmentation step, (3) new reference region generation based on LC masks without requiring a pons segmentation. CR, contrast ratio; LC, locus coeruleus; MRI, magnetic resonance imaging; ROI, region of interest.

2.2.2 | Ensemble-based Locus Coeruleus Segmentation Network

Our previously published approach[14] for fully automatic LC segmentation was further improved and used in this work. Figure 1 shows a schematic overview of the method comprising two fundamental steps that are realized using two almost identical 3D U-Net based[24] convolutional neural networks: initial LC localization followed by segmentation on an extracted patch containing only the LC and its immediate vicinity. To train and evaluate our model, we ran a 3 × 5-fold nested cross-validation which splits the subjects of the HAD dataset just as in our previous work[14] making the results directly comparable. A final training was conducted splitting the HAD in five equally sized subsets to obtain the five nets (each trained with one of these subsets as validation set and the combined rest as training set) that were used for the application to DELCODE. There was no fine-tuning or retraining with DELCODE subjects and they have not been used for either the training or validation sets (for early stopping), so that it can be seen entirely as a test set.

Three changes were introduced to the method compared to our previous work.[14] First, for the application of the models, we combined the five resulting networks in an ensemble and conducted an averaging and majority vote on the different outputs to determine the final predictions for the localization and segmentation nets, respectively. This way, the final result can profit from the information obtainable from the entire training set despite the necessity for a validation set for each individual network. Second, we normalized the intensities of the extracted image patch once again prior to passing it to the segmentation network aiming to reduce the variance of the intensity range. Third, we replaced the reference region generation relying on a sufficiently accurate pons segmentation. Instead, we determine an LC-oriented orthonormal vector base forming a coordinate system and we calculate the average offset of the semi-automatically generated reference regions on the training set in relation to the respective LCs. They are located in the pontine tegmentum—one per hemisphere. In the application case, we determine the same coordinate system and place the reference region according to the learned offset. This approach does not require a reliable pons segmentation or time-consuming registration procedures and is potentially more robust to head rotations incurred during acquisitions. The vectors for this LC-oriented coordinate system are determined as follows.

$$\vec{v}_1 = \frac{\vec{l}_1 + \vec{r}_1}{2}$$

$$\vec{v}_2 = (\vec{c}_l - \vec{c}_r) \times \vec{v}_1$$

$$\vec{v}_3 = \vec{v}_1 \times \vec{v}_2$$

with $\vec{v}_1$ the rostrocaudal LC direction derived from the two principal components obtained from two principal component analyses on the mask voxel coordinates of the LC masks of the left ($\vec{l}_1$) and right hemisphere ($\vec{r}_1$), the second base vector $\vec{v}_2$, the center of mass of the left ($\vec{c}_l$) and right LC ($\vec{c}_r$), and the third base vector $\vec{v}_3$.

2.3 | Feature extraction

The rostral LC may be particularly vulnerable in AD.[10] Hence, we extracted not only the entire LC MRI contrast ratios (CRs), but also subregional CRs and the LC mask's volume and length. Note that volume and length are estimated from the LC segmentation based solely on the in vivo MRI. Throughout this work, all reported features are bilateral, that is, they are the average of both LC hemispheres' features.

The most frequently used LC feature in the literature[15,25] are MRI intensity ratios that calculate the ratio of the maximum or median intensity value of the voxels in the LC mask (LC_{max} or LC_{median}) to the median value of a reference region (REF_{median}), positioned in the pontine tegmentum. For example, the maximum CR is defined as follows:

$$CR_{max} = \frac{LC_{max} - REF_{median}}{REF_{median}}$$

We furthermore calculate the CRs of subregions of the LC by splitting it along its axial dimension into two and three sections equal of length.

The LC signal length was measured as the number of axial slices its mask was present in, converted to millimeters.

The volume of the LC signal was determined as the number of voxels in the mask and converted into cubic millimeters.

2.4 | Experiments

We carried out the following experiments to assess the performance of Ensemble-based Locus Coeruleus Segmentation Network (ELSI-Net) in different ways. We compared its segmentations to manual expert ratings and published LC atlases, replicated subject group differences described in the literature, explored correlations of the automatically obtained LC MRI features to cerebrospinal fluid (CSF) biomarkers of AD pathology and assess the influence of acquisition-related factors. For the statistical analyses we use Jasp[26] as well as the scipy[27] library.

2.4.1 | Mask similarity

ELSI-Net's performance is evaluated by determining the agreement of its results with manual subject-wise expert segmentations. We measure the similarity of the masks with the commonly used DSC given by $DSC\,(X, Y) = \frac{2|X \cap Y|}{|X|+|Y|}$ with X and Y being sets of voxels belonging to the respective masks to be compared.

Furthermore, we compare the DSC agreement of the fully automated method, ELSI-Net, to a previously published semi-automatic approach that involves manually segmenting the LC on a study-wise template image and transforming the resulting mask into the individual subject spaces. We abbreviate this method as MT and refer the reader to Betts et al.[9] for further details.

2.4.2 | Anatomical agreement

A template-based comparison of ELSI-Net's LC masks to previously published atlases and the masks of our expert raters allows us to assess ELSI-Net's results with respect to their anatomical plausibility in terms of anatomical position and extent. To this end, we coregistered and morphed the upsampled FLASH scans of all healthy and MCI subjects together with their manual (R1) and ELSI-Net LC segmentations to FSL standard 0.5 mm asymmetric Montreal Neurological Institute (MNI) space[28] using Advanced Normalization Tools (ANTs)[29] Syn registration with bspline interpolation. We then calculated a probabilistic mask for the ELSI-Net and manual segmentations and binarized both using a 50% threshold. In total, eight subjects were excluded due to a failure of the registration process. We rendered the ELSI-Net LC template alongside the template obtained from the manual ratings as well as a recently published LC atlas by Dahl et al. (so-called meta mask[20]) that combines the information of several established LC atlases[8,9,16,18,30,31] and was brought into the same asymmetric MNI space of the other templates using Syn registration. We performed these steps analogously to Dahl et al.[20] The visualization was carried out using 3D Slicer.[32] Additionally, we determine the DSC agreement between the resulting template masks and calculate the agreement with the meta mask by Dahl et al. using the accuracy metric (mean of specificity and sensitivity) as previously described by the authors.[20] The resulting agreement is compared to a number of established and publicly available LC atlases.

2.4.3 | Feature-based comparison of subject groups

We conducted significance tests, such as *t* tests (with preceding Levene tests for equality of variances) and one-way analyses of variance (ANOVA) as well as Cohen *d* as an effect size measure.

2.4.4 | Relationship between LC signal volume and CSF measures of AD pathology

Measurements of amyloid beta ($A\beta_{42}/A\beta_{40}$) and tau proteins (total tau [t-tau], phosphorylated tau [p-tau]181) in the CSF are established biomarkers of AD pathology. From 85 of the 188 DELCODE subjects CSF measures were obtained (AD dementia: 7, MCI: 21, SCD: 22, healthy controls: 35). We correlated the automatically derived LC features to CSF biomarkers of AD pathology using Pearson *r* correlations (conditioned on variables age and sex) performed in Jasp.[26]

2.4.5 | The influence of image quality and acquisition site

We investigate the potential impact of motion artefacts on ELSI-Net. To objectively assess the MRI image quality, we make use of the convolutional neural network based approach to motion artefact quantification as recently proposed.[33] This network was trained to estimate the structural similarity (SSIM) of a single corrupted image slice to its (non-existent) ideal, uncorrupted version. The resulting predicted/estimated SSIM thus quantifies the amount of corruption by motion artefacts, 1.0 encoding perfect image quality and 0.0 the worst. We applied the network to all acquisitions from both datasets by slice-wise processing. We chose the minimum predicted SSIM out of all slices of an acquisition as its image quality score.

Apart from reporting the overall image quality of the used datasets, we investigate the relation of image quality to the DSC agreement between ELSI-Net and the expert's masks as well as to the disagreement in terms of resulting features (measured as the absolute difference between extracted features such as median and maximum CRs and LC signal volume of the two segmentation approaches). Finally, we assess the influence of the acquisition site on the agreement between ELSI-Net and the expert rating as well as on the LC CR features per se.

3 | RESULTS

ELSI-Net was applied to all subjects in the previously described fashion and produced valid output, that is, segmentation masks and the specified LC features for all subjects in both datasets.

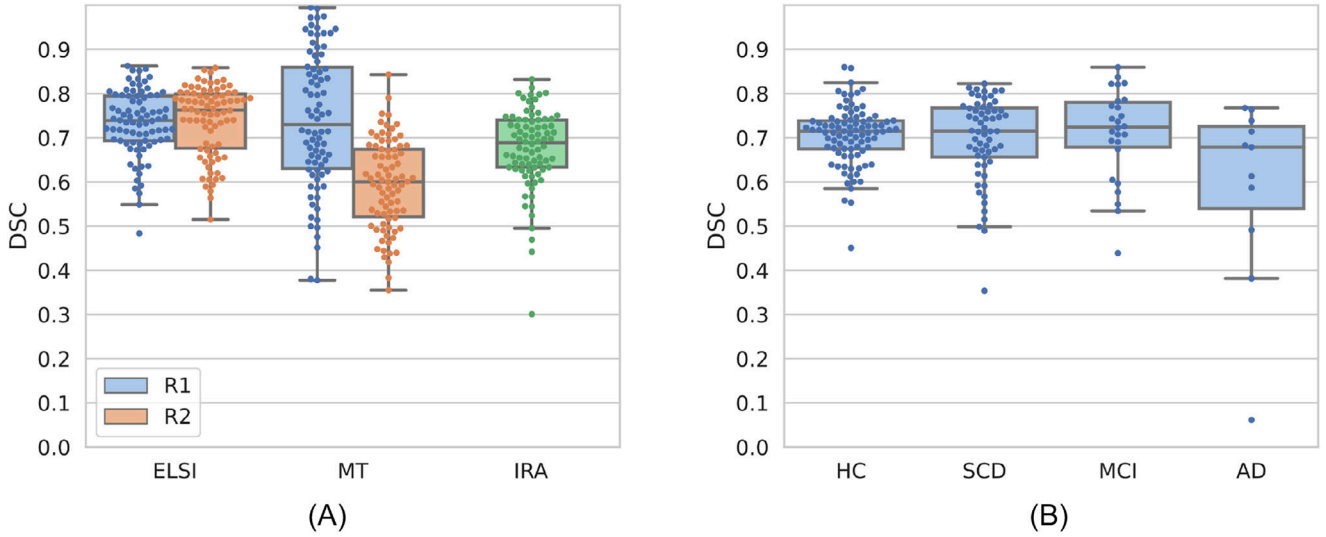

FIGURE 2 A, Healthy Aging Dataset (HAD). B, DZNE Longitudinal Cognitive Impairment and Dementia study (DELCODE) set. Box and swarm plots of Dice similarity coefficient (DSC) agreement values of ELSI-Net and the semiautomatic template registration-based method (MT)[9] on Healthy Aging Dataset (A) as well as ELSI-Net's DSC agreement on the DELCODE set (B) with respect to the manual expert segmentations by our experts Rater 1 (R1) and Rater 2 (R2). They each rated all of HAD, for which we show the agreements to each rater individually (blue and orange hue) and the inter-rater agreement (IRA, green hue). Each expert rated a complimentary subset of the DELCODE study, so that we report the agreement to the respective available rater's mask. AD, Alzheimer's disease; ELSI-Net, Ensemble-based Locus Coeruleus Segmentation Network; HC, healthy control; MCI, mild cognitive impairment; SCD, subjective cognitive decline.

3.1 | Mask similarity

3.1.1 | HAD

The HAD was rated by both expert raters with an inter-rater agreement of 67.58% ± 8.90% (mean ± standard deviation [SD]) DSC when averaging the left and right LC. The left plot in Figure 2 shows that both our automatic method (73.19% ± 7.75%) as well as the semi-automatic template registration based approach (MT;[9] 74.00% ± 15.56%) perform comparably to the inter-rater agreement in terms of mask similarity when using R1's segmentations as the reference. However, the values derived using ELSI-Net are subject to a substantially lower SD (almost half of the semi-automated method). ELSI-Net does not show a difference in its agreement to either expert rater, while with MT a strong decline of mask agreement with respect to R2's segmentations compared to those of R1 is apparent.

3.1.2 | DELCODE dataset

Across almost all subject groups from DELCODE, ELSI-Net shows relatively high agreement with a manual expert rating (see Figure 2B). The mean DSC consistently exceeds 70% and the SDs are in a range of 6.4% to 10.2%, which is comparable to the inter-rater agreement measured on HAD. The comparatively small group of AD dementia subjects however constitutes the exception with lower agreement and a larger SD in the DSC values, although the median with 67.86% is in the range of the inter-rater agreement. Figure 3 provides a qualitative visualization of two, one healthy and one AD dementia group subject.

3.2 | Anatomical agreement

Figure 4 depicts a template generated from the ELSI-Net results from DELCODE overlaid with a template from the manual ratings as well as a previously published LC atlas (meta mask) from Dahl et al.[20] ELSI-Net's agreement to the manual rating's template is very high (92.14% DSC, see Figure 4A). It confirms a good overall agreement to the manual ratings of both raters that was already suggested by the subject-level mask agreement evaluation. Only a very slight discrepancy is visible: the ELSI-Net template appears slightly shifted toward the rostral direction compared to the manual rating's template. There are, however, more deviations observable when comparing it to the meta mask (see Figure 4B), which comprises information of multiple other published atlases derived from different MRI acquisitions and modalities. The meta mask is located more rostral than the ELSI-Net template and therefore the template of the manual ratings as well. ELSI-Net's agreement with respect to the meta mask (64.33% DSC) exceeds the agreement of the manual rating's template (60.09% DSC) with the meta mask by > 4% DSC.

Other established LC atlases show agreements ranging from 63% to 47% to the meta mask as measured by the accuracy metric specified by the authors[20] (mean of sensitivity and specificity). We calculated the same metric and found ELSI-Net's template to achieve 66.62% using

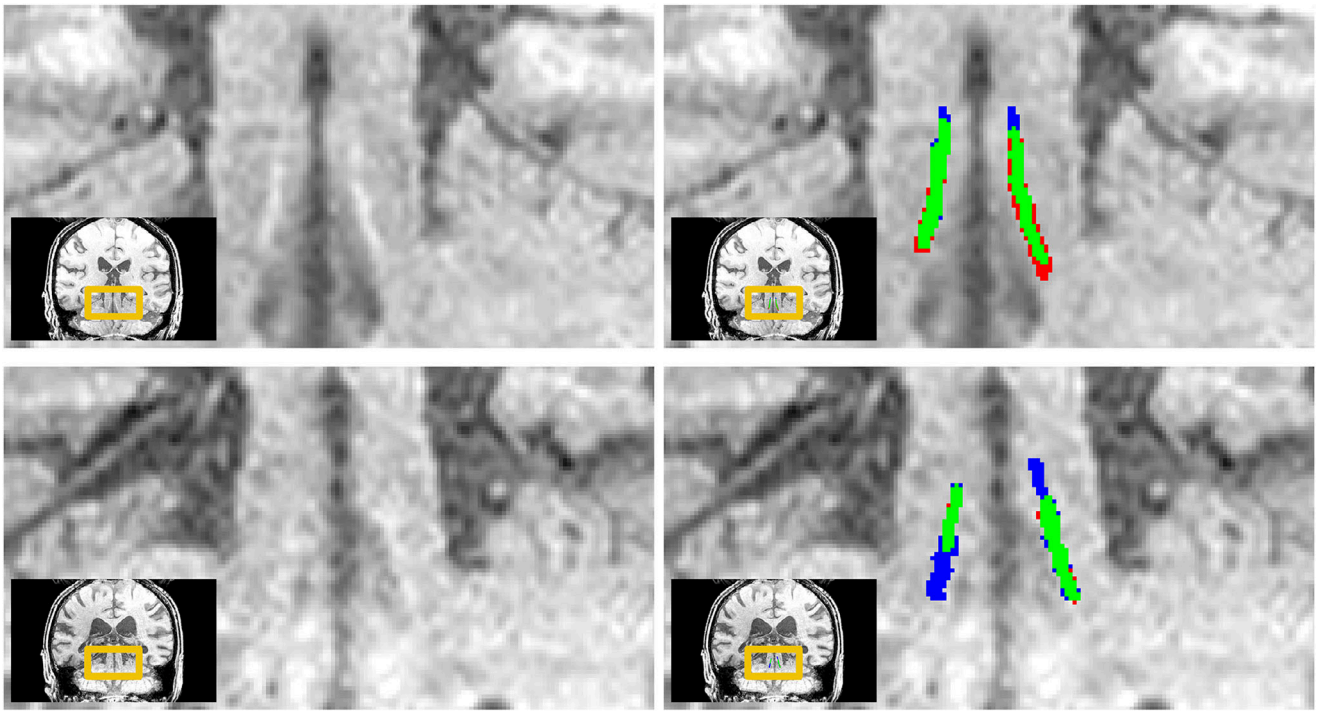

FIGURE 3 Examplary coronal slices of subjects with higher LC contrast (upper row: healthy subject, predicted SSIM (our image quality metric) is 0.8288) and low LC contrast (lower row: AD dementia subject, predicted SSIM is 0.7596). On the right, the segmentations of ELSI-Net (red) are compared to the expert rating (blue). Overlap between both is indicated in green. On the high LC contrast sample (upper row) a DSC of 74.87% was achieved, while on the low LC contrast sample 61.32% were measured. AD, Alzheimer's disease; DSC, Dice similarity coefficient; ELSI-Net, Ensemble-based Locus Coeruleus Segmentation Network; LC, locus coeruleus; SSIM, structural similarity.

this measure, which demonstrates a comparatively high agreement with the meta mask. It also exceeds the manual rating's template score, which amounts to 63.86%.

3.3 | Feature-based comparison of subject groups

We obtained the previously described LC features (see section 2.3) from HAD and DELCODE using ELSI-Net and report the resulting distributions here.

3.3.1 | HAD

We identified significant differences in LC features between young and older subject groups. The maximum CR and particularly the rostral maximum subregional CR halves and thirds show age-related increases in LC intensity on HAD. Figure 5 visualizes the resulting distributions for these features and groups. We determined Cohen *d* for estimating the effect size of the difference between young and older subjects and found that with maximum CRs R1, R2, and ELSI-Net resulted in 0.489, 0.596, and 0.662, respectively. Furthermore, all of these differences in the maximum CR were found to be statistically significant (e.g., ELSI-Net maximum CR: Student *t* test [2.759, $p = 0.007$], Levene test for equality of variances [1.857, $P = 0.177$]).

When inspecting the subregional LC CRs, it becomes evident that the age-related effect appears to be stronger in the medial and rostral LC parts. We find an increasing effect size measured by Cohen *d* from 0.374 (caudal) to 0.489 (medial) and 0.705 (rostral) for the maximum subregional CRs (splitting LC in thirds of equal length) determined by ELSI-Net.

3.3.2 | DELCODE set

Figure 6 visualizes the subject group distributions of four LC features obtained with ELSI-Net and compares them to the manual rating's distributions.

A trend of decreasing LC signal volume and length as determined by ELSI-Net appears with increasing clinical severity of the subject groups, which is not present in the expert's LC signal volume and length measurements. The visible decrease in signal volume in the AD dementia group (see Figure 6C) bears resemblance to the trend measured by the median CR with the expert ratings. A one-way ANOVA confirmed the statistical significance of this decrease (healthy controls vs. AD dementia: $F = 5.258$, $P = 0.002$ and post hoc $P_{Tukey} = 0.006$). We found Cohen *d* for the difference in signal volume between the healthy group and the AD dementia subjects to be 1.089 with ELSI-Net, which is comparable to the effect size found with the manual rating and the median CR group difference (Cohen *d*: 0.993, one-way

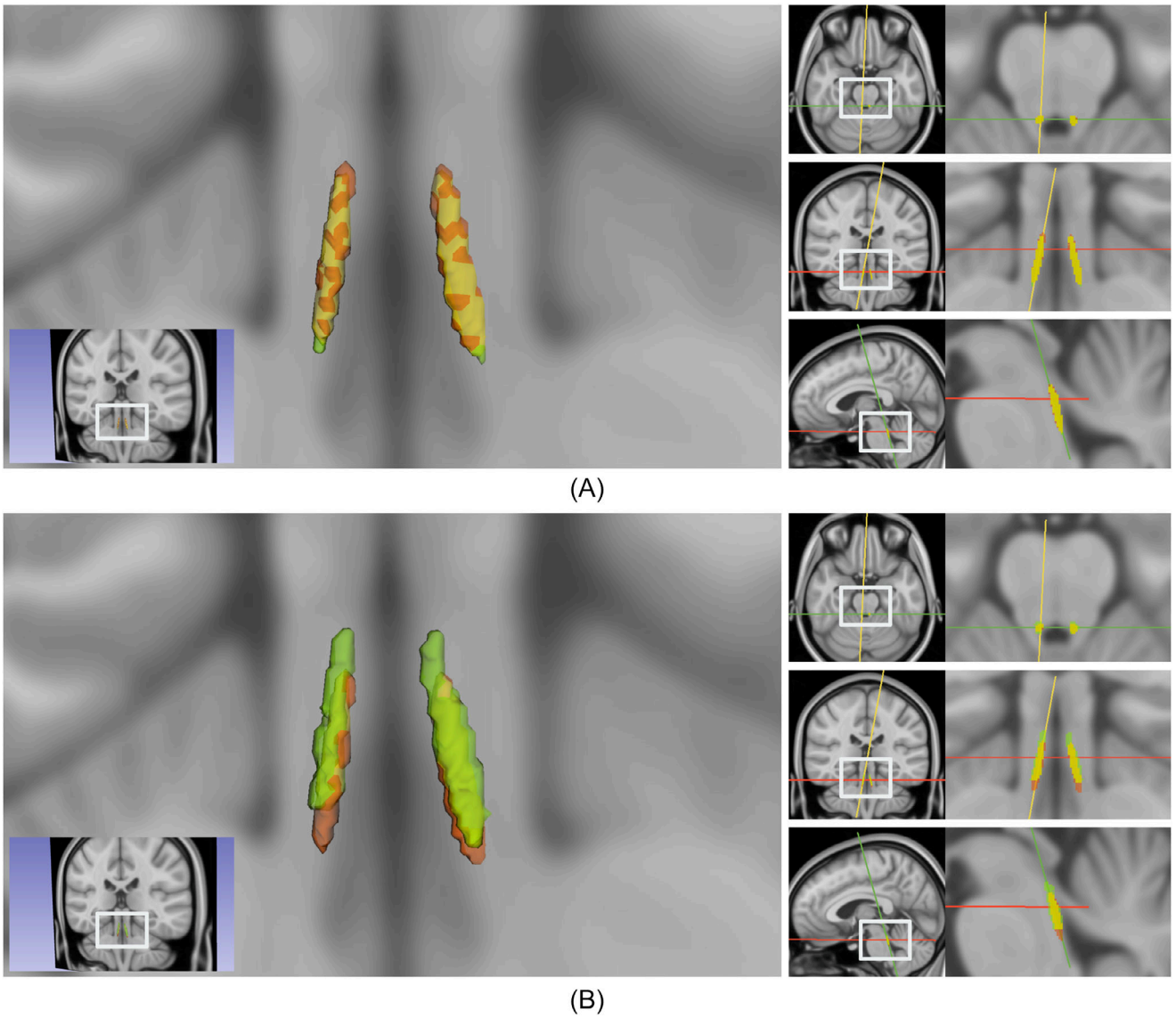

FIGURE 4 A, ELSI-Net (red) and manual ratings (green), 92.14% DSC. B, ELSI-Net (red) and the meta mask from Dahl et al.[20] (green), 64.33% DSC. Overlay of the ELSI-Net template from DELCODE with a template generated from manual ratings (A) and a published LC atlas, meta mask, that combines information from several other previously published LC atlases (B) in MNI space (0.5 mm). On the left, a 3D rendering of the masks from an example coronal slice overlaid on the MNI template is shown. On the right, the respective axial (top), coronal (middle), and sagittal (bottom) 2D slices are visualized. Red color indicates the ELSI-Net template mask, green the respective other template, and the overlapping volumes are colored in yellow. The corresponding slices are indicated by red (axial), green (coronal), and yellow (sagittal) lines on the right sides, respectively. Mask agreements are provided (DSC). Note that the agreement between the manual template (green in A) and the meta mask (green in B) is 60.09% DSC. DELCODE, DZNE Longitudinal Cognitive Impairment and Dementia study; DSC, Dice similarity coefficient; ELSI-Net, Ensemble-based Locus Coeruleus Segmentation Network; LC, locus coeruleus; MNI, Montreal Neurological Institute.

ANOVA: $F = 3.575$, $P = 0.015$ and post hoc $P_{Tukey} = 0.016$). Similar to LC signal volume, the LC signal length feature measured by ELSI-Net was also significantly decreased in AD dementia (one-way ANOVA: 4.084, $P = 0.008$ and post hoc $P_{Tukey} = 0.009$, Cohen d: 1.036; see Figure 6D).

The group differences in the CRs using ELSI-Net were not impacted by the choice of the reference region generation approach. Significant group differences were neither identified with the semi-automatically nor with the automatically generated reference regions.

3.4 | Relationship between LC signal volume and CSF measures of AD pathology

Motivated by the observed decreases in LC signal volume in MCI and AD dementia subjects measured by ELSI-Net, we correlate this feature to all available CSF measures of AD pathology. Several significant correlation results were found. They are reported in Figure 7. Although the correlations are weak, they show decreased LC signal volume is associated with higher tau and amyloid pathology. We found similar correlations between LC signal length and amyloid pathology (LC signal length

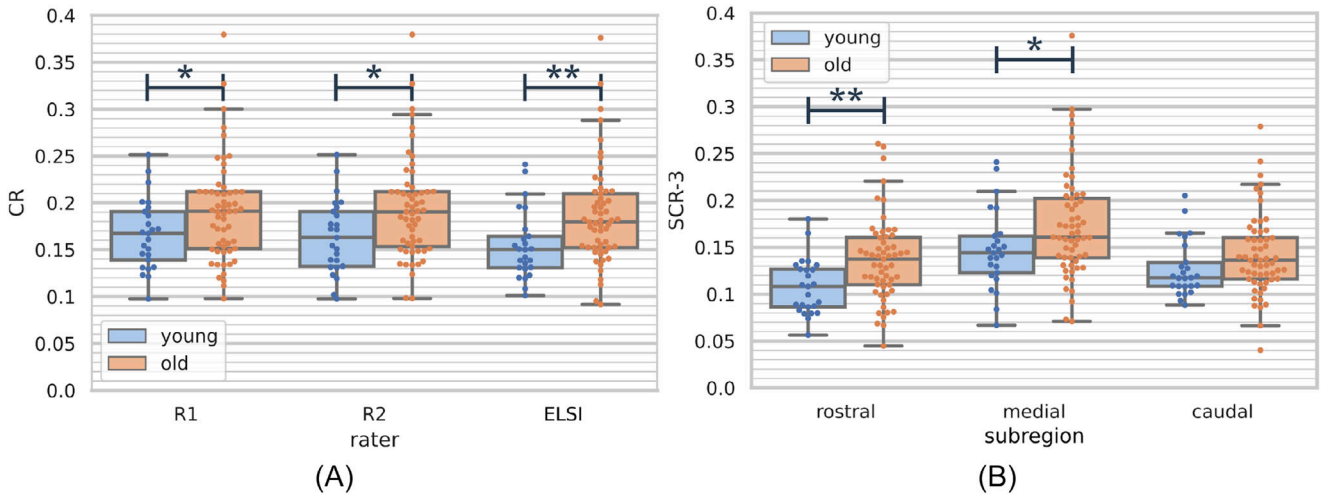

FIGURE 5 A, Maximum CRs. B, Maximum subregional CRs. Box and swarm plots of maximum LC CRs (A) and subregional maximum LC CRs (B) in young (blue) and older (orange) subject groups from Healthy Aging Dataset. For (A) the values of the expert raters R1 and R2 are reported as well as those of the fully automatic ELSI-Net. In plot (B) only the ELSI-Net results are shown. Significant differences are indicated by $^{*}p < 0.05$ and $^{**}p < 0.01$ using a two-tailed *t* test. CR, contrast ratio; ELSI-Net, Ensemble-based Locus Coeruleus Segmentation Network; LC, locus coeruleus.

and $A\beta_{42}/A\beta_{40}$: $r = 0.217$, $P = 0.049$; LC signal length and $A\beta_{42}$/p – tau181: $r = 0.296$, $P = 0.007$), but no correlation with LC CR features.

3.5 | The influence of image quality and acquisition site

Using a recently proposed method for the quantification of motion artefacts,[33] we conducted the previously described experiments to estimate the influence of acquisition artefacts on the LC metrics.

3.5.1 | Image quality of the datasets

When comparing the image quality of the two datasets, it becomes apparent that the DELCODE dataset (SSIM mean: 0.794, SD: 0.058) shows lower quality than HAD (SSIM mean: 0.827, SD: 0.035; Welch *t* test: −5.874, $P = 1.394E-8$, Levene test for equality of variances (11.505, $P = 7.987E-4$)).

A one-way ANOVA showed no significant differences between DELCODE subject groups (1.629, $P = 0.184$). Nonetheless, a coincidence of slightly decreasing image quality with increasing clinical severity is imminent in our particular dataset. The SSIM means of the healthy control, SCD, MCI, and AD dementia groups are 0.795, 0.802, 0.784, and 0.764, respectively. This motivates further investigation of a potential influence of image quality on subject group differences.

3.5.2 | Correlation of image quality with segmentation performance and feature deviation

We computed several image quality–related Pearson *r* correlations (see Table 1). One of them is between ELSI-Net's mask similarity to the manual expert rating on DELCODE quantified in terms of DSC and the measured image quality (predicted SSIM) of the samples. Although the correlation is significant it does not appear strong ($r = 0.239$, $P < 0.001$). However, a modest correlation between image quality and absolute differences in maximum CR between the two segmentation approaches was observed ($r = -0.401$, $P < 0.001$). It indicates a correlation between image quality and agreement between the CRs of ELSI-Net and manual ratings so that with increasing image quality, there is greater agreement between the two segmentation approaches. This correlation is much weaker for median LC CRs ($r = -0.255$, $P < 0.001$) and was not observed for LC signal volume feature extraction indicating LC signal volume may be influenced less by image quality.

3.5.3 | Assessment of acquisition site effects

Several one-way ANOVAs were carried out on DELCODE to gain insights into the potential influence of the acquisition sites. We found significant image quality (measured by predicted SSIM) differences between sites (one-way ANOVA: 13.274, $P = 4.121E-6$; one of the four sites was excluded since it contributed only one of the subjects).

By means of further one-way ANOVAs, we found no significant site effects on mask agreement between ELSI-Net and the expert rating (measured by DSC) as well as agreement on median and maximum CRs. This indicates that the agreement between ELSI-Net and manual expert ratings is not affected by the acquisition site. However, site-related differences could be identified in all of the CR and subregional CR features themselves (e.g., median CR, one-way ANOVA 26.809, $P = 6.057E-11$). In contrast, ELSI-Net estimates of LC signal volume and length are not subject to site effects.

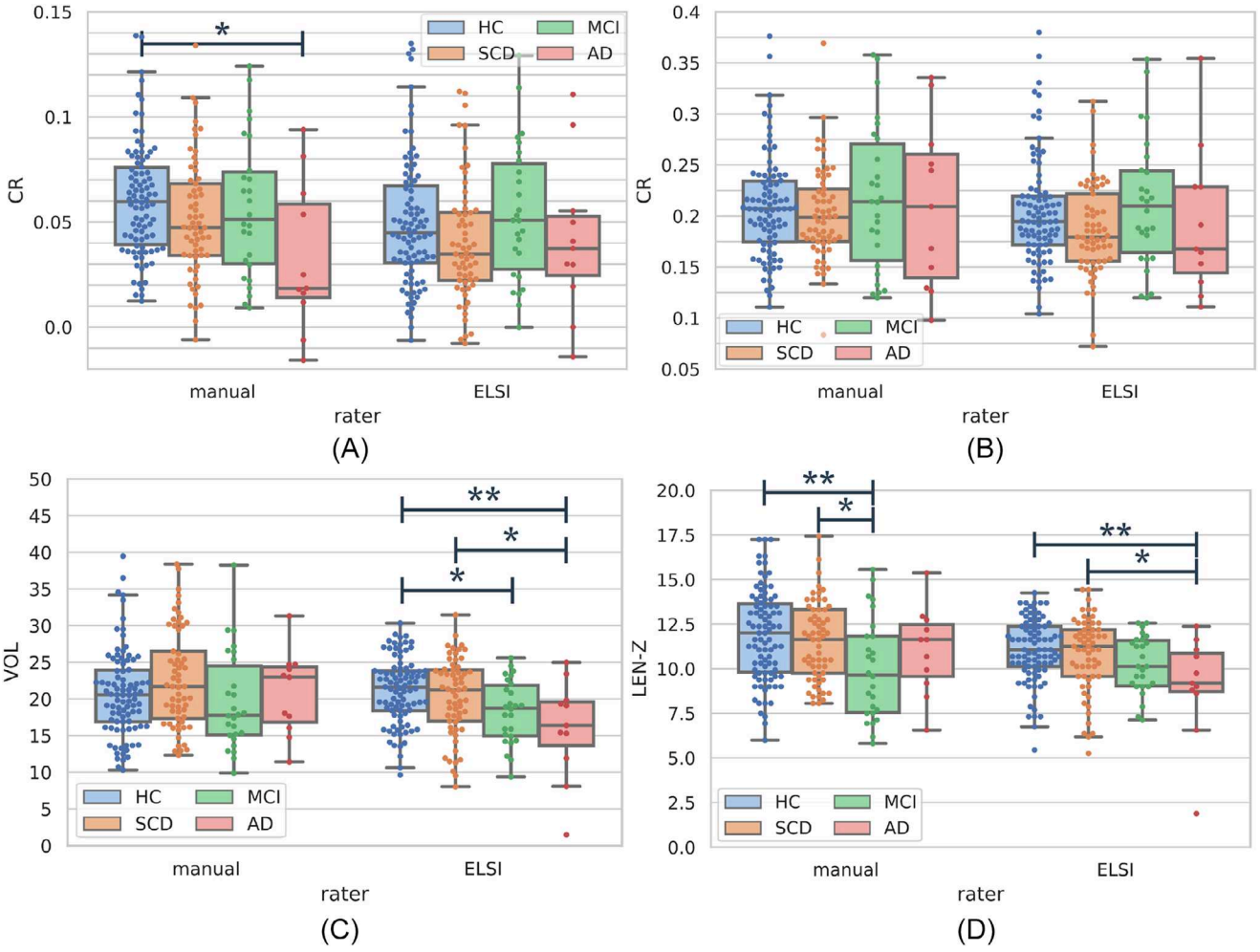

FIGURE 6 A, Median CRs. B, Maximum CRs. C, LC signal volume in mm^3. D, LC signal length in mm. Box and swarm plots of selected features of the manual rating (left) and ELSI-Net (right) on the healthy (blue), SCD (orange), MCI (green), and AD dementia (red) subject groups of the DELCODE set. Significant differences are indicated by $^*p < 0.05$ and $^{**}p < 0.01$ encoding the Tukey post hoc test result of the respective one-way ANOVA (with $p < 0.05$). AD, Alzheimer's disease; CR, contrast ratio; DELCODE, DZNE Longitudinal Cognitive Impairment and Dementia study; ELSI-Net, Ensemble-based Locus Coeruleus Segmentation Network; HC, healthy control; LC, locus coeruleus; MCI, mild cognitive impairment; SCD, subjective cognitive decline.

4 | DISCUSSION

Here, we report an improved fully automatic segmentation and feature extraction method for in vivo assessment of the LC. This method comprises an ensemble approach to applying the neural networks for LC localization and segmentation. After the initial localization, an additional intensity normalization step using a local patch surrounding the LC is applied. We propose a method for the generation of a reference region removing the need for an additional segmentation of the pons region. This approach is validated on datasets with a consistent FLASH LC MRI protocol comprising healthy young and older adults as well as in adults on the AD dementia spectrum to assess its correspondence with previously published LC atlases and clinical observations.

The proposed changes, most notably the addition of the ensemble-based inferences, led to performance improvements in terms of DSC agreement compared to our previous work,[14] in which the same nested cross-validation evaluation scheme, data splits, and dataset (HAD) were used.

In healthy aging, ELSI-Net was able to segment the LC with very high accuracy performing equal or better than an expert manual rater. In contrast to a previously published semi-automatic approach,[9] ELSI-Net exceeds the inter-rater agreement in terms of DSC with respect to both raters (compare 'ELSI' and 'MT' in Figure 2A). It is therefore arguably more objective than both the semi-automatic segmentation method and a single expert rater. A higher Cohen *d* value with respect to age-related differences between young and older adults shows that ELSI-Net could detect these differences reliably and with increased sensitivity compared to a manual segmentation approach. ELSI-Net was also able to replicate previously observed age-related increases in rostral and middle LC contrast using the same dataset.[9] ELSI-Net may potentially be deployed across sites to further reduce rater bias of LC analyses and increase comparability of studies using similar subjects and coherent FLASH MRI protocols, for example, to determine

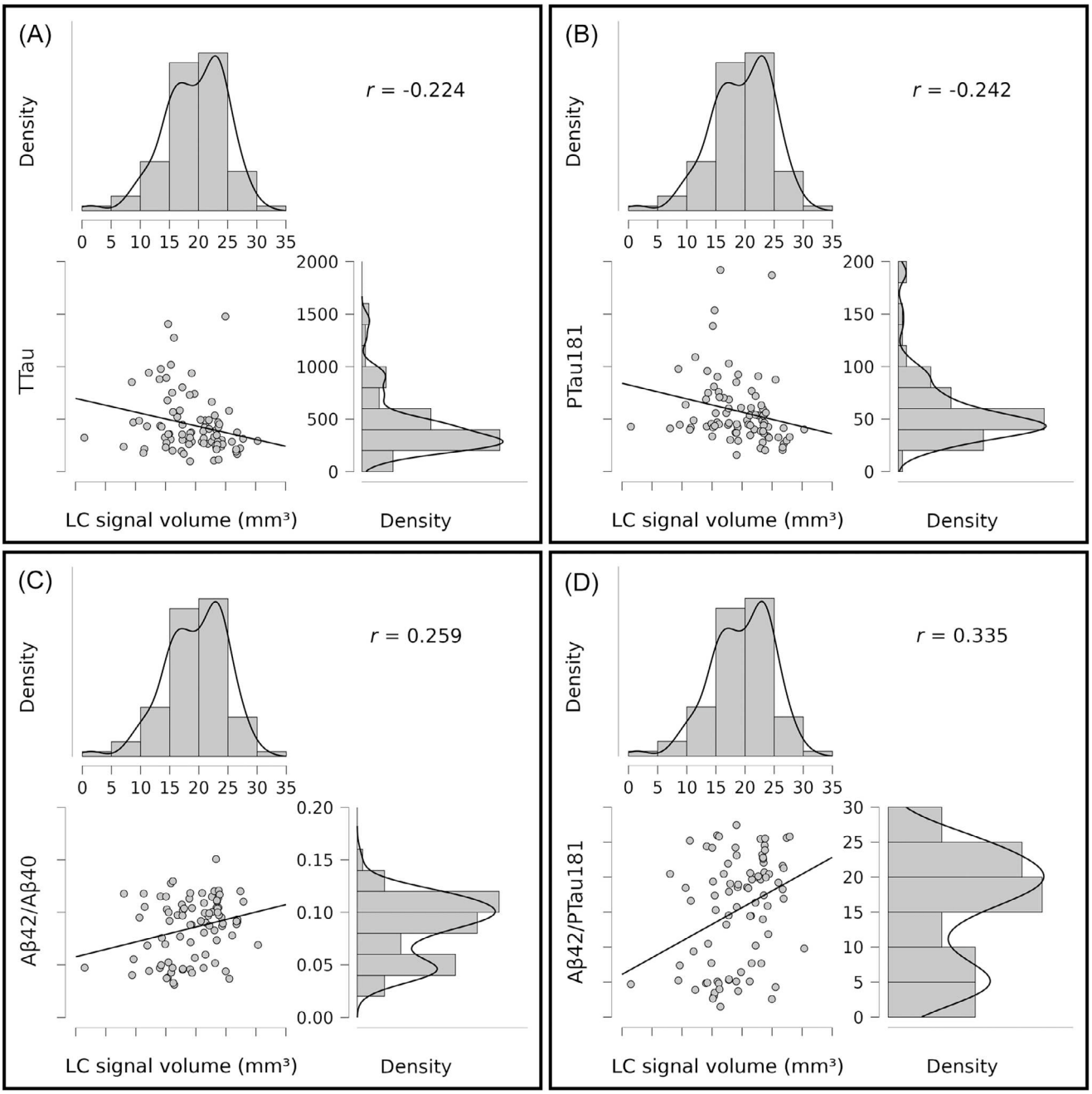

FIGURE 7 A, TTau, $p = 0.041$. B, PTau181, $p = 0.027$. C, Aβ42/Aβ40, $p = 0.018$. D, Aβ42/PTau181, $p = 0.002$. Pearson r correlations (conditioned on variables age and sex) between LC signal volume obtained with ELSI-Net and cerebrospinal fluid (CSF) measures of AD pathology. Scatterplots and value distributions are visualized. Correlation coefficient (r) and p value are reported. Aβ, amyloid beta; AD, Alzheimer's disease; ELSI-Net, Ensemble-based Locus Coeruleus Segmentation Network; LC, locus coeruleus; PTau, phosphorylated tau; TTau, total tau.

normative feature ranges of a healthy LC given a specific age. Indeed, acquisition site–related influences on the LC features were found independently of the segmentation approach and have to be considered.

In a clinical cohort of individuals with AD dementia from DELCODE, ELSI-Net could effectively segment the LC without any fine-tuning, solely being trained on the healthy aging dataset. For the most part (including the MCI subject group) a satisfactory agreement measured by DSC compared to the experts' rating was achieved. The AD dementia group was a noticeable exception, although it was the smallest group with only 11 participants and further analyses are required in larger cohorts of individuals with AD dementia to comprehensively determine its segmentation accuracy.

The overall anatomical plausibility of the automatically obtained LC masks and a performance comparable to those of experts is indicated by the very high agreement between the LC template of the ELSI-Net DELCODE masks and the template generated from the manual ratings, but also a meta mask comprising a number of previously published atlases.[20] ELSI-Net's template shows the highest agreement with the

TABLE 1 Pearson *r* correlation of the image quality measured as predicted structural similarity (SSIM) of DELCODE and the specified other values: The DSC agreement of the ELSI-Net and expert's masks and the absolute differences of several extracted LC signal features.

Value name	*r*	*p*
DSC(ELSI-Net, experts)	0.239***	9.546E-4
$\lvert(\text{medianCR}_{\text{ELSI-Net}} - \text{medianCR}_{\text{experts}})\rvert$	−0.255***	4.098E-4
$\lvert(\text{maximumCR}_{\text{ELSI-Net}} - \text{maximumCR}_{\text{experts}})\rvert$	−0.401***	1.132E-8
$\lvert(\text{VOL}_{\text{ELSI-Net}} - \text{VOL}_{\text{experts}})\rvert$	−0.084	0.249
$\lvert(\text{LEN}_{\text{ELSI-Net}} - \text{LEN}_{\text{experts}})\rvert$	−0.008	0.908

Abbreviations: CR, contrast ratio; DELCODE, DZNE Longitudinal Cognitive Impairment and Dementia study; DSC, Dice similarity coefficient; ELSI-Net, Ensemble-based Locus Coeruleus Segmentation Network; LC, locus coeruleus; LEN, length; VOL, volume.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

meta mask among all published atlases that were used for the meta mask creation. This indicates that ELSI-Net can generate anatomically precise LC segmentations in the context of the described FLASH MRI protocol removing the need for semi-automatic or manual segmentation. With ELSI-Net we found significant differences between healthy controls and subjects with AD dementia with respect to LC signal volume and length but not median LC contrast, as observed with the expert ratings. This could indicate a deviation of segmentation style between ELSI-Net and manual LC segmentation. It is conceivable that ELSI-Net is differentially influenced by a reduction in LC MRI contrast present in AD dementia subjects compared to expert raters that rely more on anatomical prior knowledge. ELSI-Net was only trained on LC segmentations from healthy young and older adults. Because it has never seen the variance introduced by AD dementia during training, the lower LC contrast or additional data characteristics unknown to us may yield smaller LCs in ELSI-Net's assessment in AD. This would explain the lower DSC agreement between ELSI-Net and manual raters observed in the AD dementia group as well as the deviations seen in LC CRs between ELSI-Net and manual rating on DELCODE. Therefore LC signal volume and length estimates using ELSI-Net might be more accurate in clinical cohorts due to reduced human bias that manual raters may be prone to. The examples shown in Figure 3 indicate the shape and size variance in ELSI-Net's segmentation, which appears to be more directly influenced by local image properties. Hence we presume ELSI-Net is unlikely to merely produce a learned average LC mask. It should be noted that ELSI-Net may not only rely on intensity, but on additional characteristics such as the surrounding anatomy and shape of the LC during the segmentation procedure.

We further investigated the influence of image quality on our (automatic) LC analysis in multiple ways. We found significant correlations between image quality and agreement on CR-based features (as quantified by absolute difference between ELSI-Net and the experts' rating in particular for maximum intensity CRs). However, no such correlation was observed with respect to ELSI-Net's LC signal volume measure, where we observed most pronounced differences between healthy controls and AD dementia.

We found significant acquisition site–related effects on the image quality and CR features in general. No influence of site could be found on the agreement between ELSI-Net and the experts' rating with respect to mask and CR feature agreement as well as ELSI-Net's measurements of LC signal volume and length indicating a robust performance across multiple sites.

As an additional validation step, we further assessed how LC segmentations generated by ELSI-Net are related to previously reported associations with AD pathology. In a subset of subjects with known CSF status from DELCODE, we found reduced LC signal volume was significantly associated with increased tau and amyloid pathology in agreement with previous findings.[10,34,35,36]

4.1 | Limitations

An important caveat in our analyses is the rather small number of AD dementia subjects ($n = 11$) in our clinical cohort of 188 participants. Further evaluations on more datasets with larger groups of AD dementia subjects, ideally with amyloid and tau pathology biomarkers, are required to ascertain the performance of ELSI-Net more conclusively in this population. Another interesting aspect that was left unexplored is the robustness of ELSI-Net with respect to acquisition parameters and differing MRI protocols. To this end, more extensive evaluation work on a variety of datasets comprising diverse anisotropic resolutions, differing LC MRI sequences such as turbo spin echo and magnetization transfer–based acquisitions need to be carried out, and ELSI-Net's performance on slab acquisitions needs to be investigated. We aim to conduct these experiments, publish them, and release ELSI-Net as an easy-to-use Docker in the near future. Of course, a necessary requirement for the application of ELSI-Net is the acquisition of specific sequences with LC contrast in general. Recently, automatic methods were applied to LC segmentation in acquisitions without LC contrast as an alternative to atlas-based approaches.[37] While inherently lacking precision and the possibility to extract LC contrast, signal volume, or length features, they may allow functional MRI or diffusion MRI analyses in datasets in which LC MRI was not acquired.

Finally, the image quality assessed here was based on the quality of the whole-brain acquisition and was not LC specific, which may not reflect the quality of the visualization of the LC and its immediate vicinity.

5 | CONCLUSION

In this work, we evaluate an improved version of a previously proposed, fully automatic approach to LC segmentation and feature extraction termed ELSI-Net. Evaluation on LC imaging data acquired from young and older adults but also from subjects across the AD dementia continuum show that ELSI-Net reliably generates anatomically plausible results with excellent agreement to established LC atlases given a consistent FLASH LC MRI protocol. We found increased objectivity with ELSI-Net compared to single expert raters and a semi-automatic

LC segmentation method. LC features by ELSI-Net demonstrate high sensitivity replicating previously shown subject group differences in healthy aging and AD dementia. We saw correlations of LC signal volume measured by ELSI-Net to tau and amyloid pathology and robust performance with respect to data acquired across multiple sites.

ELSI-Net provides a means to automatically segment the LC with high accuracy particularly in aging cohorts in the context of the FLASH LC MRI protocol. Further analyses are required to determine its effectiveness in segmenting the LC in different LC MRI contrasts; longitudinal datasets; and additional clinical cohorts, for example including subjects with Parkinson's disease, depression, and further neurological disorders.

ACKNOWLEDGMENTS/CONSENT STATEMENT

This research was supported by the federal state of Saxony-Anhalt, Germany (Project I 88). M. J. B. is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project-ID 425899994 - Sonderforschungsbereiche 1436 (SFB 1436) and by the German Federal Ministry of Education and Research (BMBF, funding code 01ED2102B) under the aegis of the EU Joint Programme—Neurodegenerative Disease Research (JPND). D. H. is supported by Sonderforschungsbereich 1315, Project B06, Sonderforschungsbereich 1436, Project A08, ARUK SRF2018B-004, NIH R01MH126971. The DELCODE study was approved by the ethical committees of the medical faculties of all participating sites: the ethical committees of Berlin (Charité, University Medicine), Bonn, Cologne, Göttingen, Magdeburg, Munich (Ludwig-Maximilians-University), Rostock, and Tübingen. The process was led and coordinated by the ethical committee of the medical faculty of the University of Bonn. The registration number of the trial at the ethical committee in Bonn is 117/13. The Healthy Aging study received ethical approval from the ethics commission of the Faculty of Medicine at the University of Magdeburg. Written informed consent was obtained from all subjects in accordance with the Declaration of Helsinki.

Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare no potential conflicts of interests. Author disclosures are available in the supporting information.

ORCID

Max Dünnwald https://orcid.org/0000-0003-3838-3345

REFERENCES

1. Theofilas P, Ehrenberg AJ, Dunlop S, et al. Locus coeruleus volume and cell population changes during Alzheimer's disease progression: a stereological study in human postmortem brains with potential implication for early-stage biomarker discovery. *Alzheimer's & Dementia*. 2017;13(3):236-246. http://www.sciencedirect.com/science/article/pii/S1552526016326796. doi:10.1016/j.jalz.2016.06.2362
2. Braak H, Del Tredici K. Where, when, and in what form does sporadic Alzheimer's disease begin?. *Curr Opin Neurol*. 2012;25(6):708-714. doi:10.1097/WCO.0b013e32835a3432
3. Mather M, Harley CW. The Locus Coeruleus: essential for Maintaining Cognitive Function and the Aging Brain. *Trends Cogn Sci*. 2016;20(3):214-226. https://www.cell.com/trends/cognitive-sciences/abstract/S1364-6613(16)00018-8. doi:10.1016/j.tics.2016.01.001
4. Betts MJ, Kirilina E, Otaduy MCG, et al. Locus coeruleus imaging as a biomarker for noradrenergic dysfunction in neurodegenerative diseases. *Brain*. 2019;142(9):2558-2571. doi:10.1093/brain/awz193
5. Krohn F, Lancini E, Ludwig M, et al. Noradrenergic neuromodulation in ageing and disease. *Neurosci Biobehav Rev*. 2023;152:105311. https://www.sciencedirect.com/science/article/pii/S0149763423002804. doi:10.1016/j.neubiorev.2023.105311
6. Galgani A, Giorgi FS. Exploring the Role of Locus Coeruleus in Alzheimer's Disease: a Comprehensive Update on MRI Studies and Implications. *Curr Neurol Neurosci Rep*. 2023;23(12):925-936. doi:10.1007/s11910-023-01324-9
7. Trujillo P, Aumann MA, Claassen DO. Neuromelanin-sensitive MRI as a promising biomarker of catecholamine function. *Brain*. 2023:awad300. doi:10.1093/brain/awad300
8. Keren NI, Lozar CT, Harris KC, Morgan PS, Eckert MA. In vivo mapping of the human locus coeruleus. *Neuroimage*. 2009;47(4):1261-1267. https://linkinghub.elsevier.com/retrieve/pii/S1053811909006235. doi:10.1016/j.neuroimage.2009.06.012
9. Betts MJ, Cardenas-Blanco A, Kanowski M, Jessen F, Düzel E. In vivo MRI assessment of the human locus coeruleus along its rostrocaudal extent in young and older adults. *Neuroimage*. 2017;163:150-159. doi:10.1016/j.neuroimage.2017.09.042
10. Betts MJ, Cardenas-Blanco A, Kanowski M, et al. Locus coeruleus MRI contrast is reduced in Alzheimer's disease dementia and correlates with CSF Aβ levels. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*. 2019;11:281-285. http://www.sciencedirect.com/science/article/pii/S2352872919300144. doi:10.1016/j.dadm.2019.02.001
11. Fernandes P, Regala J, Correia F, Gonçalves-Ferreira AJ. The human locus coeruleus 3-D stereotactic anatomy. *Surg Radiol Anat*. 2012;34(10):879-885. doi:10.1007/s00276-012-0979-y
12. Ariz M, Abad RC, Castellanos G, et al. Dynamic Atlas-Based Segmentation and Quantification of Neuromelanin-Rich Brainstem Structures in Parkinson Disease. *IEEE Trans Med Imaging*. 2019;38(3):813-823. doi:10.1109/TMI.2018.2872852
13. Tona K-D, Keuken MC, de Rover M, et al. In vivo visualization of the locus coeruleus in humans: quantifying the test–retest reliability. *Brain Structure and Function*. 2017;222(9):4203-4217. doi:10.1007/s00429-017-1464-5
14. Dünnwald M, Ernst P, Düzel E, Tönnies K, Betts MJ, Oeltze-Jafra S. Fully automated deep learning-based localization and segmentation of the locus coeruleus in aging and Parkinson's disease using neuromelanin-sensitive MRI. *Int J Compr Assisted Radiology and Surgery*. 2021. doi:10.1007/s11548-021-02528-5
15. Liu KY, Marijatta F, Hämmerer D, Acosta-Cabronero J, Düzel E, Howard RJ. Magnetic resonance imaging of the human locus coeruleus: a systematic review. *Neurosci Biobehav Rev*. 2017;83:325-355. http://www.sciencedirect.com/science/article/pii/S0149763417304219. doi:10.1016/j.neubiorev.2017.10.023
16. Ye R, Rua C, O'Callaghan C, et al. An in vivo Probabilistic Atlas of the Human Locus Coeruleus at Ultra-high Field. *Biorxiv*. 2020:2020. https://www.biorxiv.org/content/10.1101/2020.02.03.932087v1
17. Puigvert M, Gabriel C, Javier U, et al. Atlas-based segmentation of brainstem regions in neuromelanin-sensitive magnetic resonance images. *Medical Imaging 2015: Computer-Aided Diagnosis*. International Society for Optics and Photonics; 2015:94140Y. doi:10.1117/12.2080779

18. Dahl MJ, Mather M, Düzel S, et al. Higher rostral locus coeruleus integrity is associated with better memory performance in older adults. *Biorxiv*. 2019:332098. doi:10.1101/332098
19. Sibahi A, Gandhi R, Al-Haddad R, et al. Characterization of an automated method to segment the human locus coeruleus. *Hum Brain Mapp*. 2023;44(9):3913-3925. doi:10.1002/hbm.26324
20. Dahl MJ, Mather M, Werkle-Bergner M, et al. Locus coeruleus integrity is related to tau burden and memory loss in autosomaldominant Alzheimer's disease. *Neurobiol Aging*. 2022;112:39-54. doi:10.1016/j.neurobiolaging.2021.11.006
21. Dünnwald M, Betts MJ, Sciarra A, Düzel E, Oeltze-Jafra S. Automated Segmentation of the Locus Coeruleus from Neuromelanin-Sensitive 3T MRI Using Deep Convolutional Neural Networks. 2020:61-66. doi:10.1007/978-3-658-29267-6_13
22. Jessen F, Spottke A, Boecker H, et al. Design and first baseline data of the DZNE multicenter observational study on predementia Alzheimer's disease (DELCODE). *Alzheimer's Research & Therapy*. 2018;10:15. doi:10.1186/s13195-017-0314-2
23. Yushkevich PA, Piven J, Hazlett HC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage*. 2006;31(3):1116-1128. doi:10.1016/j.neuroimage.2006.01.015
24. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: learning Dense Volumetric Segmentation from Sparse Annotation. In: Ourselin S, ed. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*. Springer International Publishing; 2016:424-432. doi:10.1007/978-3-319-46723-8_49
25. Galgani A, Lombardo F, Della Latta D, et al. Locus Coeruleus Magnetic Resonance Imaging in Neurological Diseases. *Curr Neurol Neurosci Rep*. 2020;21(1):2. doi:10.1007/s11910-020-01087-7
26. JASP Team. JASP (Version 0.18.0)[Computer software]. 2023. https://jasp-stats.org/
27. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental Algorithms for Scientific Computing in Python. *Nat Methods*. 2020;17:261-272. doi:10.1038/s41592-019-0686-2
28. Fonov VS, Evans AC, Botteron K, et al. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. In: *Neuroimage*. Organization for Human Brain Mapping 2009 Annual Meeting 47 2009:S102. issn: 1053-8119. doi:10.1016/S1053-8119(09)70884-5. https://www.sciencedirect.com/science/article/pii/S1053811909708845
29. Advanced Normalization Tools. original-date: 2013-01-23T15:43:41Z. Oct. 23, 2023. https://github.com/ANTsX/ANTs
30. Liu KY, Acosta-Cabronero J, Cardenas-Blanco A, et al. In vivo visualization of age-related differences in the locus coeruleus. *Neurobiol Aging*. 2019;74:101-111. doi:10.1016/j.neurobiolaging.2018.10.014
31. Tona K-D, van Osch MJP, Nieuwenhuis S, Keuken MC. Quantifying the contrast of the human locus coeruleus in vivo at 7 Tesla MRI. *PLoS One*. 2019;14(2):e0209842. doi:10.1371/journal.pone.0209842
32. Kikinis R, Pieper SD, Vosburgh KG. 3D Slicer: a Platform for Subject-Specific Image Analysis, Visualization, and Clinical Support. In: Jolesz FA, ed. *Intraoperative Imaging and Image-Guided Therapy*. Springer; 2014:277-289. doi:10.1007/978-1-4614-7657-3_19
33. Sciarra A, Chatterjee S, Dünnwald M, et al. Reference-less SSIM Regression for Detection and Quantification of Motion Artefacts in Brain MRIs. In: 2022. https://openreview.net/forum?id=24cqMfboXhH
34. Jacobs HIL, Becker JA, Kwong K, et al. In vivo and neuropathology data support locus coeruleus integrity as indicator of Alzheimer's disease pathology and cognitive decline. *Sci Transl Med*. 2021;13(612):eabj2511. doi:10.1126/scitranslmed.abj2511
35. Cassidy CM, Therriault J, Pascoal TA, et al. Association of locus coeruleus integrity with Braak stage and neuropsychiatric symptom severity in Alzheimer's disease. *Neuropsychopharmacology*. 2022;47(5):1128-1136. doi:10.1038/s41386-022-01293-6
36. Van Egroo M, Riphagen JM, Ashton NJ, et al. Ultra-high field imaging, plasma markers and autopsy data uncover a specific rostral locus coeruleus vulnerability to hyperphosphorylated tau. *Mol Psychiatry*. 2023;28(6):2412-2422. doi:10.1038/s41380-023-02041-y
37. Aganj I, Mora J, Fischl B, Augustinack JC. Automatic geometry-based estimation of the locus coeruleus region on T1-weighted magnetic resonance images. *Front Neurosci*. 2024;18. doi:10.3389/fnins.2024.1375530

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Dünnwald M, Krohn F, Sciarra A, et al. Fully automated MRI-based analysis of the locus coeruleus in aging and Alzheimer's disease dementia using ELSI-Net. *Alzheimer's Dement*. 2025;17:e70118. https://doi.org/10.1002/dad2.70118