

Review article

Advances in MRI optic nerve segmentation

Carla Xena-Bosch^{a,*}, Sriki Kirti Kodali^b, Nitin Sahi^b, Declan Chard^{b,c}, Sara Llufríu^d,
Ahmed T Toosy^b, Eloy Martínez-Heras^d, Ferran Prados^{a,b,e}

^a e-Health Center, Universitat Oberta de Catalunya, Barcelona, Spain

^b Queen Square MS Centre, Department of Neuroinflammation, UCL Institute of Neurology, Faculty of Brain Sciences, University College London, London, United Kingdom

^c National Institute for Health Research (NIHR) University College London Hospitals (UCLH) Biomedical Research Centre, United Kingdom

^d Neuroimmunology and Multiple Sclerosis Unit, Hospital Clínic de Barcelona, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS) and Universitat de Barcelona, Barcelona, Spain

^e Centre for Medical Image Computing, Department of Medical Physics and Biomedical Engineering, University College London, London, United Kingdom

ARTICLE INFO

Keywords:

MRI
Optic nerve
Segmentation
Neurodegenerative disease
Deep learning

ABSTRACT

Understanding optic nerve structure and monitoring changes within it can provide insights into neurodegenerative diseases like multiple sclerosis, in which optic nerves are often damaged by inflammatory episodes of optic neuritis. Over the past decades, interest in the optic nerve has increased, particularly with advances in magnetic resonance technology and the advent of deep learning solutions. These advances have significantly improved the visualisation and analysis of optic nerves, making it possible to detect subtle changes that aid the early diagnosis and treatment of optic nerve-related diseases, and for planning radiotherapy interventions. Effective segmentation techniques, therefore, are crucial for enhancing the accuracy of predictive models, planning interventions and treatment strategies. This comprehensive review, which includes 27 peer-reviewed articles published between 2007 and 2024, examines and highlights the evolution of optic nerve magnetic resonance imaging segmentation over the past decade, tracing the development from intensity-based methods to the latest deep learning algorithms, including multi-atlas solutions using single or multiple image modalities.

1. Introduction

The optic nerve (ON) connects the eye to the brain (see Fig. 1). Morphologically, ONs are thin, tortuous structures that extend from the globe of the eye to the optic chiasm, exhibiting significant anatomical variability in size, shape and curvature, as well as variable amounts of surrounding cerebrospinal fluid (CSF) along their length (van Elst et al., 2023). The ONs transmit all visual information from the retina through to the hemidecussation at the optic chiasm. Thereafter, visual information travels along the optic tracts to synapse in the lateral geniculate nucleus (see Fig. 1). From there, visual information is conveyed in the optic radiation which synapses in the primary visual cortex. The ON is a bundle of axons (the healthy ON has between 693,000 to 1685,000 fibres), the number of which declines substantially as we age (with an estimated annual loss of 4000 to 5426 nerve fibres), and there is no significant difference between males and females, or between the left and right optic nerves (Chow and Paley, 2021).

ON structural features can be very useful to understand disease *in-*

vivo. Associations of structural changes in the ON with objective measures of function such as vision and visual evoked potential (VEP), can be studied. The ON is often the earliest site affected in multiple sclerosis (MS), with optic neuritis presenting as visual impairment associated with focal damage of the nerve. In addition, the ON is a critical structure in radiotherapy planning and neuro-oncology, as it is one of the organs at risk because of its sensitivity to radiation. Excessive radiation exposure can lead to optic neuropathy, resulting in vision impairment or even blindness. Therefore, precise delineation of the ON is essential for effectively treating tumors located near the optic pathways.

The growing interest in the ON is linked to technological improvements in magnetic resonance imaging (MRI). In 2007, (Hickman, 2007) reviewed major advances in ON segmentation in MS over the previous 10 years. These advances started with the demonstration of ON atrophy after optic neuritis particularly in MS, thanks to the development of fat- and CSF-suppressed imaging sequences (Yiannakas et al., 2010). Another milestone in ON research, beyond segmentation, was the ability to measure magnetization transfer ratio (MTR) over the entire ON,

* Corresponding author at: NeuroADaS Lab - e-Health Center, Universitat Oberta de Catalunya, Rambla del Poblenou, 154, 08018 Barcelona, Spain.

E-mail address: carlaxena@uoc.edu (C. Xena-Bosch).

<https://doi.org/10.1016/j.msard.2025.106437>

Received 11 November 2024; Received in revised form 5 March 2025; Accepted 8 April 2025

Available online 8 April 2025

2211-0348/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

providing an indication of the degree of demyelination/remyelination within the lesion (Hickman, 2007). Additionally, the ability to measure microstructural properties in the orbital ON was made possible by the emergence of quantitative diffusion-weighted imaging (DWI) techniques using diffusion tensor imaging (DTI) measurements (Wheeler-Kingshott et al., 2002; Wheeler-Kingshott et al., 2006). These acquisitions introduced the possibility to perform nerve-tracking analysis using zonal oblique multislice echo planar imaging (ZOOM-EPI), which has a shortened echo train length that increases resolution and decreases distortions, thereby visualising axonal integrity in structures as small as the ON (Dowell et al., 2009). In this manner, MRI advances have helped to increase resolution, sensitivity and contrast to noise ratio, thereby improving visualisation and easy detection of small structures like the ON (Chow and Paley, 2021). MRI ON segmentation allows direct visualisation of pathology in the ON for a wide range of neurodegenerative diseases. It also has an important role in planning radiotherapy interventions. Although there are published reviews on segmentation of “organs-at-risk” using computed tomography (CT) mainly for planning radiotherapy treatments (Cardenas et al., 2019; Vrtovec et al., 2020; Liu et al., 2023), none are dedicated specifically to MRI ON segmentation.

In this review, we will discuss the evolution of MRI ON segmentation over the past decade, from intensity based methods or multi-atlas solutions to the latest deep learning algorithms using single or multiple image modalities. In this era of the artificial intelligence (AI) solutions for precision medicine, effective segmentation techniques that allow precise measurement of nerve atrophy and lesion load in neurodegenerative diseases are key indicators of progression and severity. These techniques could also improve the accuracy of predictive models, leading to better patient management and personalised treatment strategies.

2. Material and methods

Our aim is to bridge the 15-year gap from 2007 to the present and so we only included works published from 2007 (when Hickman et al. published their comprehensive review (Hickman, 2007)) up to the 30th of August 2024. This review used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology (Page et al., 2021). It includes peer-reviewed journal articles ON automated segmentation using MRI techniques. We conducted a PubMed search to

identify pertinent articles. PubMed was chosen as the primary source of this review due to its comprehensive coverage of peer-reviewed biomedical literature, ensuring the inclusion of high-quality and relevant studies on the topic. The search criteria included the following search terms: “optic nerve”, “magnetic resonance” and “automatic” or “automated”, alongside “segmentation” or “detection” keywords. Additionally, the bibliographic references cited in the identified papers were scrutinised to find potential missing articles. Non-peer-reviewed articles, such as those found on preprint online repositories and conference proceedings were not included.

When available in their respective articles, we have also provided metrics to quantify the algorithms’ performance in segmenting the ON. As an evaluation score, the reviewed articles have included the Dice Similarity Coefficient (DSC), which is a measure of the spatial overlap between two masks. Higher values are better, and its range goes from 0 to 1 (Dice, 1945). The DSC is calculated as follows:

$$DSC(GT, PM) = \frac{2 \times |GT \cap PM|}{|GT| + |PM|}$$

DSC is a function of the predicted mask (PM) and the manual binary segmentation masks, which are considered the ground truth (GT) (Prados et al., 2017).

The Hausdorff Distance (HD) has been used in order to evaluate the accuracy of shape and boundary predictions since it is an effective indicator to assess contour similarity, and is occasionally employed as Hausdorff distance 95th percentile (HD95) to eliminate outliers.

$$D(X \rightarrow Y) = \max(d_i^{X \rightarrow Y}), i = 1 \dots N_x$$

$$HD(GT, PM) = \max(D(GT \rightarrow PM), D(PM \rightarrow GT))$$

where d is the Euclidean distance between voxel x and y (van Elst et al., 2023; Prados et al., 2017).

The Hausdorff distance 95th percentile (HD95) is defined as:

$$HD95(GT, PM) = \max_{95\%}(HD(GT, PM))$$

Another common measure to assess the performance is using the Symmetric Mean Absolute Surface Distance (MSD), which captures the average Euclidean distance between the predicted and ground truth values (Prados et al., 2017).

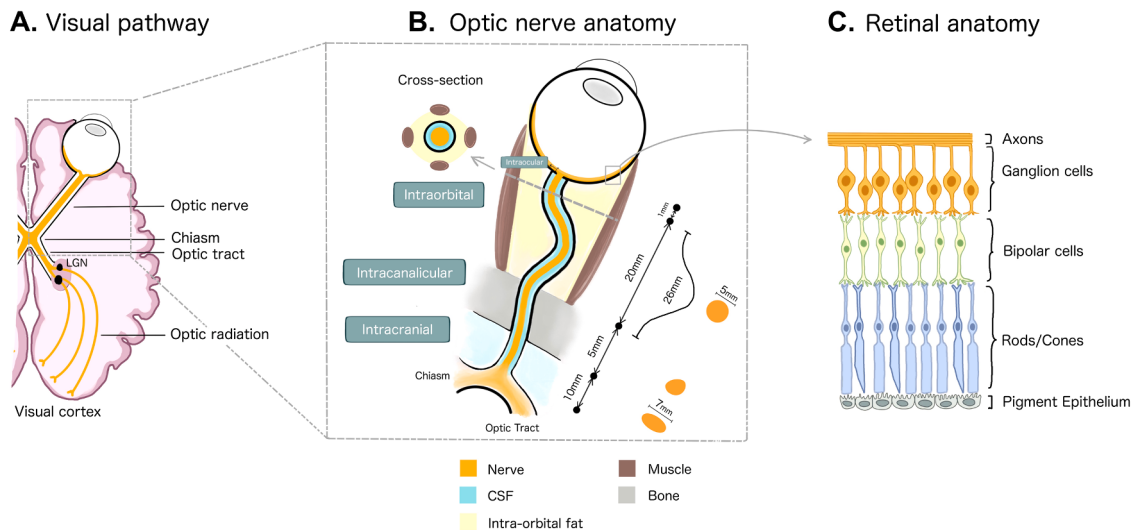


Fig. 1. A. Schematic overview of the visual pathway; B. Optic nerve anatomy; and C. Retinal anatomy. The optic nerve connects the retina to the brain and is surrounded by cerebrospinal fluid from the globe of the eye to the optic chiasm. It can be divided into intraorbital, intracanalicular and intracranial segments. The retina is composed of optic nerve axons, ganglion cells (cell bodies of axons that form the optic nerve), bipolar cells, rods, cones, and pigment epithelium, amongst others. Abbreviations: lateral geniculate nucleus (LGN).

$$MSD(GT, PM) = \frac{1}{N_{GT} + N_{PM}} \left(\sum_{i=1}^{N_{GT}} |d_i^{GT \rightarrow PM}| + \sum_{i=1}^{N_{PM}} |d_i^{PM \rightarrow GT}| \right)$$

where N_{GT} and N_{PM} are the total number of voxels in the contour for GT and PM respectively.

3. Results

Fig. 2 summarises the identification of publications included in this review. From the initial 44 articles identified in the previously defined PubMed searches, in a first screening one was excluded as it had been published before 2007, another was Hickman's review itself and another was a recent review about MRI radiotherapy methods (Liu et al., 2023). Afterwards, the remaining articles were assessed and 22 of those were discarded: 12 because the ON was not segmented, 5 did not use MRI, 2 were clinical applications of other methods, one was a conference paper, another featured manual segmentation and the last one was a review of ensemble methods. Additionally, 20 more articles were identified from the bibliographic references of the previously selected papers. Of those, 6 were discarded for being conference papers, 6 for not segmenting the ON, 2 for not using MRI and 1 for performing a manual segmentation of the ON. Therefore, 19 papers found via the PubMed search and 8 via citation searching were included in this review, resulting in a final set of 27 papers.

We categorised the 27 articles from this review into two main groups based on the methodology that they employed. The first group comprises classic or traditional techniques, most of which predate the appearance of deep learning. This group, which includes 15 papers (see Table 1), encompasses intensity based methods, single and multi-atlas template techniques, and other non deep learning based approaches. The second group comprises 12 papers (see Table 2) that use deep learning segmentation techniques to delineate the ON. Each of these two main groups have been further divided into two subgroups based on the number of images involved. Within these two categories, we distinguish

between single image modality, which benefits from a single MRI acquisition protocol for delineating the ON, and multi-modality approach, which includes more than one image type, sometimes from different sources (i.e. CT scans), to leverage different contrast to optimally delineate the ON boundaries.

3.1. Classical approaches

ON segmentation, similar to other tissue or organ segmentation, initially used techniques that were intensity based, single or multi atlas/template propagation or shape models which are recognized in this paper as classical approaches. Within this group, we categorise the techniques into single modality or multi-modality, depending on the number of input images they used.

3.1.1. Single image modality

The included publications in this section use only MRI as input which typically is T1-weighted (T1w), T2-weighted (T2w), T1 with Gadolinium (Gad), proton density (PD) or even DWI.

3.1.1.1. Single atlas or intensity based methods. In 2008, (Isambert et al., 2008), presented a single atlas-based segmentation approach on T1-weighted MRI to segment the ONs. This method was named atlas-based automatic segmentation (ABAS). The ABAS is composed of an synthetic MRI or atlas of the brain on which each cerebral structure of interest was manually delineated by an expert using a pair of rigid registered CT/MR images. Using ABAS, the labels from the atlas were automatically propagated to the input image.

Yiannakas et al. (2010, 2013), and then later (Nguyen et al., 2018), both used an Active Shape Model (ASM) segmentation to semi-automatically delineate the ON. This model captures both the shape of deformation of structures and the intensity variations. Firstly, the ASM is constructed. Manual segmentations of the lens and ON were performed and then used to define a volume of interest (VOI) in the eye

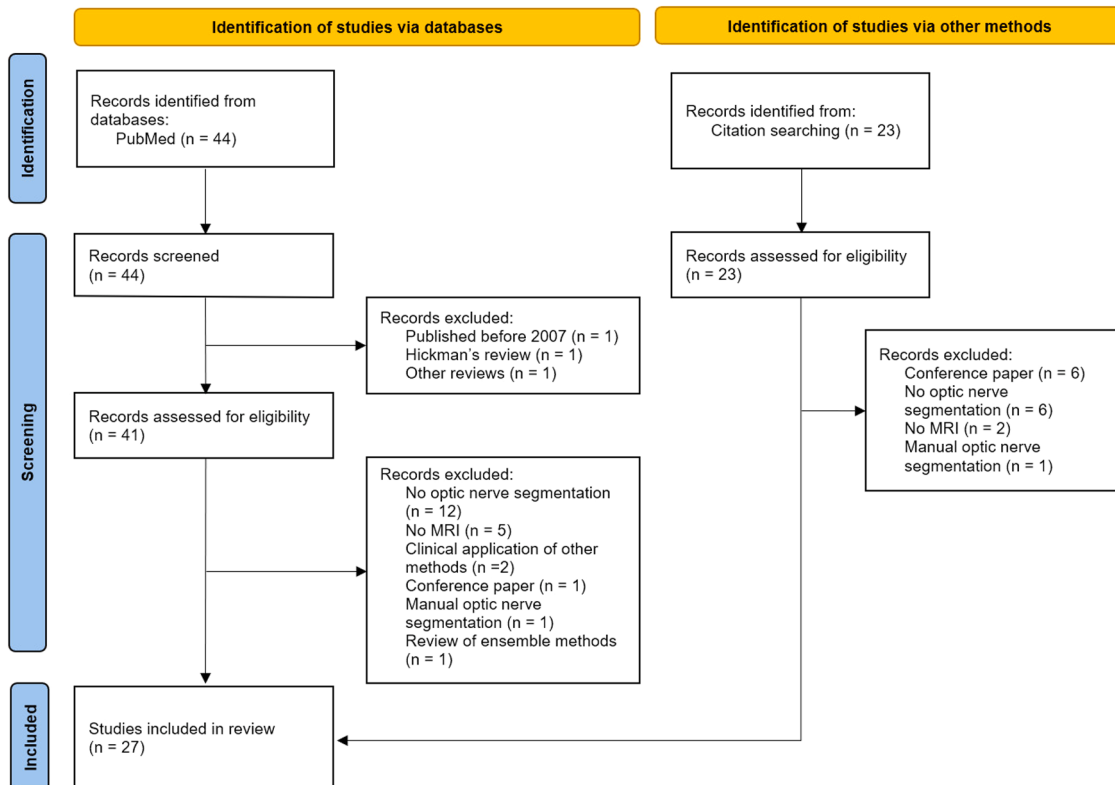


Fig. 2. PRISMA 2020 flow diagram for new systematic reviews. The process of identification of eligible articles is illustrated.

Table 1
Summary of published articles that have used classic methods for optic nerve segmentation. From left to right; number of imaging modalities used, the year of publication, reference, short description of each study, imaging modalities used, summary of results, and specific disease or application. The highlighted results only apply to the optic nerve, which if not available, the whole visual pathway was considered. Results are derived from different datasets as reported by the respective studies.

Modality	Year	Author	Short description	Imaging modalities	Results	Disease / Application
Single	2008	Isambert et al. (2008)	Atlas-based automatic segmentation software (ABAS)	T1w	DSC = 0.38 (range, 0.40–0.53)	Radiotherapy
	2010	Yiannakas et al. (2010)	Active Shape Model (ASM)	FSE	–	Atrophy
	2011	Hernowo et al. (2011)	Segmentation based on voxel-based morphometry (VBM) using a study-specific tissue probability map (TPM).	T1w	–	Glaucoma
	2013	Yiannakas et al. (2013)	Active Shape Model (ASM)	T2w	–	Atrophy
	2014	Ramli et al. (2014)	ON segmentation using the NeuRoi Software followed by manual outline refinement.	T1w	–	Glaucoma
	2014	Panda et al. (2014)	Several multi-atlas label fusion algorithms are assessed.	T2w	DSC = 0.81, HD = 2.18 mm and MSD = 0.41 mm	MS and mimics
	2016	Harrigan et al. (2016)	Multi-atlas segmentation	T2w	–	MS and mimics
	2018	Nguyen et al. (2018)	Active Shape Model (ASM)	T1w contrast-enhanced and non-contrast-enhanced	DSC = 0.82 and HD = 1.86 mm	Uveal melanoma
	2021	He et al. (2021)	Four different tractography methods	dMRI	–	Optic nerve related diseases
	2022	Crouzen et al. (2022)	Automated atlas-based segmentation	T1w	Internal cohort DSC _L = 0.83 ± 0.04 DSC _R = 0.84 ± 0.05 External cohort DSC _L = 0.84 ± 0.10 DSC _R = 0.85 ± 0.10	Radiotherapy
	2022	Tan et al. (2022)	Three segmentation methods: Adaptively Regularized Kernel-based Fuzzy-C Mean (AFCM), Level Set Method (LSM) and Multiplicative Intrinsic Component Optimization (MICO); and three interpolation methods: Reverse Diffusion (RD), Iterative Curvature Based Interpolation (ICBI) and Contrast-Guided (CG)	T1w or T2w	–	Optic neuritis
	2023	Feng et al. (2023)	Gradient-based edge detection with skeletonization (GES) with bicubic interpolation	T1w	DSC = 0.81 ± 0.04	Optic nerve related diseases
	2024	Chow et al. (2024)	Three segmentation methods: spatial-FCM (sFCM); modified-FCM (mFCM); and LSM; and three interpolation methods: Lanczos; iterative curvature-based interpolation (ICBI); and CG	PD or T2w	DSC _{PD} = 0.92 ± 0.01 DSC _{T2w} = 0.88 ± 0.01	Optic nerve related diseases
Multi	2011	Noble and Dawant (2011)	Tubular structure localization algorithm	T1w contrast-enhanced and non-contrast-enhanced and CT	DSC ≈ 0.80	Radiotherapy
	2019	Miller et al. (2019)	Semi-automated diffusion magnetic resonance imaging (dMRI) tractography method	dMRI and T1w	–	Glaucoma

from specific MRI acquisitions or from a structural T1-weighted image. Nguyen et al. (2018). used T1-weighted images with and without contrast and obtained a DSC value of 0.82 and a HD of 1.86 mm for the ON segmentation.

In 2011, (Hernowo et al., 2011). introduced a voxel-based morphometry (VBM) approach over T1-weighted images to determine the volume of the ONs. To do so, a study-specific tissue probability map (TPM) is first computed by registering a set of brains to a common space, followed by skull-stripping using the Brain Extraction Tool (BET). The brains images are then divided into six tissue classes by applying the FMRIB Automated Segmentation Tool (FAST) and an average tissue class volume is created from all the segmentations in common space and the average image from all the processed images, this average image is going to be used as template reference (i.e. atlas) for the posterior segmentations. Finally, any individual ON segmentation can be computed. The source input T1-weighted image to process is segmented following the VBM approach that is part of the Statistical Parametric Mapping (SPM) software. In short, using the previously created TPM, the input source image is registered and modulated in order to detect the different tissue classes, including the ON.

Subsequently in 2014 and also using single modality, (Ramli et al., 2014). measured the ON’s volume on T1-weighted MRI using the NeuRoi Software. The ON was identified on an axial MRI slice as an

isointense signal in the optic tract similar to brain white matter and separated from the hypointense signal of the dural sheath. Finally, the wall of the ON was manually outlined by drawing the region of interest (ROI). The ON was outlined from the most anterior part of the visible ON (posterior to the eye globe) to the anterior optic chiasm.

In 2022, (Tan et al., 2022). suggested a novel image processing mode to address the limitations in MRI by studying three segmentation methods and three interpolation methods on T1 and T2-weighted images. The methods were: Adaptively Regularized Kernel-based Fuzzy-C Mean (AFCM), Level Set Method (LSM) and Multiplicative Intrinsic Component Optimization (MICO); and the interpolation methods: Reverse Diffusion (RD), Iterative Curvature Based Interpolation (ICBI) and Contrast-Guided (CG). The MRIs were processed in two stages: segmentation and interpolation to obtain an improved spatial resolution to integrate the image processing of ONs as the diagnosis of optic neuritis. The measurements of six datasets were compared with the mean cross-sectional area of the normal ONs. The combination LSM-ICBI over T1 images obtained the closest value (26.52 mm²) of mean cross-sectional area to the reference value (27.51 ± 0.83) mm². In T2-weighted images, the pair RD-LSM obtained the closest value (22.43 mm²) of area measurement to the reference value of (22.26 ± 1.29) mm².

Table 2

Summary of published articles that have used deep learning methods for optic nerve segmentation. From left to right; number of imaging modalities used, the year of publication, reference, short description of each study, imaging modalities used, summary of results, and specific disease or application. The highlighted results only apply to the optic nerve, which if not available, the whole visual pathway was considered. Results are derived from different datasets as reported by the respective studies.

Modality	Year	Author	Short description	Imaging modalities	Results	Disease / Application
s	2019	Chen et al. (2019)	Recursive ensemble organ segmentation (REOS) framework	T1w	DSC _L = 0.78 ± 0.11 HD _L = 3.20 ± 2.20 mm DSC _R = 0.82 ± 0.59 HD _R = 1.80 ± 0.70 mm	Brain cancer
	2020	Mlynarski et al. (2020)	Modified version of 2D U-Net CNN	T1w Gad	DSC _{Raw} = 0.67 DSC _{Tot} = 0.87 HD = 6.30 mm	Brain cancer
	2020	Ai et al. (2020)	Spatial probabilistic distribution map (SPDM)-based two channel 3D U-Net	T1w or CT	DSC = 0.86 ± 0.01 HD = 3.56 ± 1.89 mm	Preoperative planning
	2022	Dai et al. (2022)	Mask scoring R-CNN	T1w	DSC _L = 0.67 ± 0.11 HD95 _L = 3.25 ± 2.11 mm MSD _L = 0.91 ± 0.38 mm DSC _R = 0.68 ± 0.11 HD95 _R = 2.96 ± 1.32 mm MSD _R = 0.87 ± 0.29 mm	Head and neck cancer
	2022	Martí-Juan et al. (2022)	Specific image processing followed by an automated pipeline to extract the optic nerve and a 3D CNN to detect ON lesions	T2w	–	MS and mimics
	2023	van Elst et al. (2023)	3D U-Net	T2w	DSC = 0.84 ± 0.03 DSC _{Tot} = 0.91 ± 0.04 HD95 = 0.60 [range 0.42 - 1.02] mm	Retinoblastoma
Multi	2016	Mansoor et al. (2016)	PAScAL (PArTitioned Shape and Appearance Learning)	T1w Gad, T2w and FLAIR	DSC = 0.79 ± 0.07	Several, tumors
	2019	Tong et al. (2019)	Generative adversarial network (GAN) with shape constraint (SC-GAN)	Low-field MRI and CT	DSC _L = 0.72 ± 0.05 DSC _R = 0.69 ± 0.07	Radiotherapy
	2020	Liu et al. (2020)	Dual pyramid networks (DPN)	T1w and CT	DSC _L = 0.72 ± 0.12 HD95 _L = 3.15 ± 5.14 mm MSD _L = 0.94 ± 1.32 mm DSC _R = 0.72 ± 0.20 HD95 _R = 3.14 ± 5.11 mm MSD _R = 1.10 ± 2.01 mm	Head and neck cancer
	2021	Dai et al. (2021)	Dual pyramid networks (DPN)	T1w and CBCTs	DSC _L = 0.78 ± 0.05 HD95 _L = 1.86 ± 1.73 mm MSD _L = 0.55 ± 0.18 mm DSC _R = 0.77 ± 0.04 HD95 _R = 2.06 ± 2.69 mm MSD _R = 0.58 ± 0.32 mm	Head and neck cancer
	2023	Xie et al. (2023)	CNTSeg	T1w and dMRI	DSC = 0.82	None in particular
	2023	Alzahrani et al. (2023)	Combination of atlas and 3D U-Net (Çiçek et al., 2016)	T1w Gad and CT	DSC _L = 0.65 ± 0.09 MSD _L = 0.09 ± 0.03 mm DSC _R = 0.68 ± 0.08 MSD _R = 0.09 ± 0.03 mm	Radiotherapy

3.1.1.2. Label fusion or multi-atlas based methods. Rather than relying on a single atlas label template or TPM, the label fusion or multi-atlas based strategy is based on registering separately each single template image from a database with multiple templates to the input subject image. These pairwise warps are then used to transfer the template labels into the input subject space. Then, a label fusion algorithm combines the anatomical variability from all the registered templates to generate the final segmentation, this process is accounting for individual

differences and improving the reliability of segmentation results.

Label fusion methods offer two main advantages: first across-subject anatomical variability is better captured than in a single atlas, which can be viewed as a parametric model that typically uses single mode distributions (e.g., Gaussian) to encode anatomical appearance, and, second, the fusion of results coming from multiple registrations improve robustness against occasional registration failures. The main drawback is the computational burden introduced by the multiple registrations

and information fusion from the entire training data (Sabuncu et al., 2010). Empirical results in these studies suggest that errors in the manual labelling and registration procedures are reduced during label fusion, resulting in more accurate segmentation.

On the ON, in 2014 (Panda et al., 2014), used a multi-atlas label fusion algorithm on T2-weighted images. In this work, the template database manual segmentation was built using the Medical Image Analysis Processing and Visualization software package v7 (MIPAD) for the full length of the left and the right ONs and the optic chiasm on all the subjects. Once the template database was built, seven statistical and multi-atlas label fusion algorithms to segment the ONs were assessed. This included majority vote, simultaneous truth and performance level estimation (STAPLE), spatial STAPLE, local weighted vote, non-local STAPLE, non-local spatial STAPLE and joint label fusion. From these, the most consistent segmentations were obtained by non-local spatial STAPLE, which achieved a median DSC of 0.81, MSD 0.41 mm, and HD 2.18 mm for the ONs. Joint label fusion achieved a slightly superior median performance for the ONs (DSC = 0.82, MSD = 0.39 mm, and HD = 2.15 mm), but ultimately non-local spatial STAPLE was selected given its slight advantage in the outlier reduction and lower surface distance measures (Panda et al., 2014).

Some time later, in 2016 (Harrigan et al., 2016), also used a multi-atlas segmentation approach to localise the ON and sheath on T2-weighted MRI followed by a slice-wise Gaussian mixture approach for fine tuning the obtained segmentations and constructing a model of two concentric tubes. The combined approach from Harrigan et al. was characterised by noisy results due to the slice-wise approach and with difficulties to manage bended ONs.

Crouzen et al. (2022), developed in 2022 a multi-atlas-based segmentation method using T1-weighted MR scans and rigid image registration to propagate the labels. The method performance is compared to manual delineation by two specialists. The DSC of the left ON respectively in the internal and external evaluation cohort was of 0.83 ± 0.04 and of 0.84 ± 0.10 ; and for the right ON, it was of 0.84 ± 0.05 in the internal evaluation cohort and of 0.85 ± 0.10 in the external one.

3.1.1.3. Tractography-based and other methods. DWI has allowed the opportunity to track white matter projections within the brain. These tractography methods can delineate anterior and posterior visual pathways. They are characterised by the detection of the brain regions like the ON using an automatic selection of strategic ROIs for seeding and afterwards clustering a high definition of fibres that identifies the structure.

He et al. (2021), investigated in 2021 the performance of multiple tractography methods for reconstruction of the complete retinogeniculate visual pathway (RGVP) including the four anatomical subdivisions, using diffusion MRI (dMRI). The four different methods included two methods based on the constrained spherical deconvolution (CSD) model, the deterministic (SD-Stream) and the probabilistic (iFOD1); and two that used the unscented Kalman filter (UKF) tractography framework, the one-tensor (UKF-1T) and the two-tensor (UKF-2T). For each RGVP subdivision, three ROIs were used, including one of the ON ROIs (left or right), one of the optic tract ROIs (left or right), and the optic chiasm ROI. Finally, the performance of each tractography method across subjects was evaluated using the normalised overlap score (NOS) method and a comparison to anatomical T1-weighted-based RGVP segmentation. The UKF-2T method obtained the highest score (NOS = 0.718), indicating the highest overlap of tractography across subjects. The next highest scores were obtained by iFOD1 (NOS = 0.605) and UKF-1T (NOS = 0.508). SD-Stream obtained the lowest score (NOS = 0.398).

Two years after, in 2023, Feng et al. (Feng et al., 2023), proposed a segmentation method called gradient-based edge detection with skeletonization (GES) for the cross-sectional on magnetic resonance (MR) images acquired with T1-weighted fast spoiled gradient-echo without

fat saturation. Firstly, the images were pre-processed with bicubic interpolation to improve the spatial resolution. Secondly, the proposed GES segmentation was applied to produce a distinct ON image. It was a semi-automated method using prior knowledge of the location of the ON. The edges of the ON were identified by finding the largest gradient changes in signal intensity between the ON region and its surrounding CSF. Particle swarm optimization (PSO) and level set method (LSM) segmentations were applied for comparison. Manual segmentation performed by a certified radiologist was used as the ground truth. The bicubic-GES processed ON images were used for the quantitative measurement on ten datasets. The DSC indexes were calculated for every slice in each portion of the ON. The proposed GES segmentation demonstrated superior results compared to PSO and LSM, obtaining a mean and standard deviation DSC of 0.81 ± 0.04 .

Most recently, in 2024, (Chow et al., 2024), strived to find an optimum and automated interpolation and segmentation method, in order to improve the spatial resolution of the ONs on fat-saturated MRI, which produced two types of images: PD and T2-weighted. They compared three interpolation methods: Lanczos; iterative curvature-based interpolation (ICBI); and CG along with three segmentation methods: spatial-FCM (sFCM); modified-FCM (mFCM); and LSM. Nine methods were investigated by combining different interpolation and segmentation methods in a different order. Lanczos-mFCM was identified as the best model in this study to process the ON images based on four factors: image quality; DSC value; the percentage difference in the area; the signal-to-noise ratio; and contrast-to-noise ratio. The best-identified interpolation and segmentation combination was then applied to measure the ON mean area from 10 datasets. The Lanczos-mFCM method produced ONs with DSC values of 0.92 ± 0.01 from the PD images and 0.88 ± 0.01 from the T2-weighted images measured on 10 datasets and all slices. In addition, the Student's *t*-test at 99 % confidence level showed that there is no significant difference in the ON areas measured using manually segmented and Lanczos-mFCM processed images. Therefore, the manual and Lanczos-mFCM methods are equivalent.

3.1.2. Multi-modality images

Multimodal methods typically combine MRI and CT. A necessary preprocessing step for these methods is a body registration between the input MR and CT images. This group of approaches starts with (Noble and Dawant, 2011), in 2011, who used both contrast-enhanced and non-contrast enhanced T1-weighted MR, and CT images to segment the ONs and chiasm. In this case, Noble et al. used a tubular structure localization algorithm in which a statistical model and image registration were used to incorporate a priori local intensity and shape information from an atlas. The structures of interest were segmented manually in each of these MR/CT pairs, and the centerline of these structures was extracted using a thinning method. Finally, a correspondence between points along the centerlines was established. It is relevant to mention that Noble et al. extended the ONs past the chiasm and considered the ON and the contra-lateral optic tract as a single structure. In addition, the optic chiasm was not explicitly segmented, instead the chiasm is found as the intersection of the two ONs, which is as its anatomical definition. The method resulted in a mean DSC of approximately 0.80 for both ONs.

Subsequently, in 2019, (Miller et al., 2019), developed an advanced, semi-automated dMRI tractography based method to identify and analyse the ONs and using the T1-weighted MRI to place the ROIs. Using a pair of diffusion scans, a low-noise field-corrected volume could be created, allowing the ONs to be isolated using probabilistic tractography. To improve the quality of the tractography, the two reverse-encoded diffusion scans were combined into a single corrected volume using the FMRIB Software Library (FSL) software (University of Oxford, Oxford, England). Three ROIs along the brain's visual pathway were manually identified. The T1-weighted image was used to place the left and right ONs and the optic chiasm by gross anatomy. Visual pathways were derived through probabilistic diffusion-weighted

tractography using MRtrix2 (Brain Research Institute, Melbourne, Australia) and fiber groups were cleaned using the Automated Fiber Quantification toolkit (Stanford University). Fibers were overlaid on the anatomical T1-weighted volume and any fibers that were found to be anatomically implausible were manually removed.

3.2. Deep learning approaches

An explosive growth in image processing power has helped the development of AI solutions in many fields including ON segmentation. This second group comprises articles which use deep learning solutions to automatically delineate the ON; they are also split in two subgroups depending on the number of input modalities.

3.2.1. Single image modality

In 2019, (Chen et al., 2019) used a recursive ensemble organ segmentation (REOS) framework on T1-weighted MRI to segment six brain regions, including the ON. Eighty images were retrospectively collected with the gold-standard manual contours. Among these, sixty were assigned for model training and five-fold cross-validation, and the other twenty for testing. In these recursive frameworks, each brain region is grouped into different levels according to key factors that determine auto-segmentation accuracy. Large-volume and high-contrast brain regions are assigned to low-level groups and small-volume and low-contrast brain regions are assigned to high-level groups. Low-level brain regions are segmented first and the segmented low-level organs are used as constraints to guide the high-level brain regions segmentation. In each level, an ensemble of two 3D U-Nets, namely an EnUNet, is present. The EnUNet architecture contains three modules each one of them based on a 3D convolutional neural network (CNN) that localise the brain region, then detect the contour and finally ensemble the results obtained from the previous steps. The three modules are trained sequentially. The proposed REOS method achieved a segmentation with mean DSCs of 0.78 ± 0.105 , $0.822 \pm 0.59\%$ for the left and right ON, respectively. The HD was 3.20 ± 2.20 mm for the left ON and 1.80 ± 0.70 mm for the right one.

At the start of the decade, in 2020, (Mlynarski et al., 2020). instead used a modified version of the 2D U-Net CNN, on Gad-enhanced T1-weighted images as well, addressing problems related to computational costs and missing ground truth segmentations for a subset of classes. Feature maps of the encoding part as concatenated in the decoding part in order to combine low-level and high-level features and to ease the flow of gradients during the optimization process. The final convolutional layer (the segmentation layer) of the standard U-Net had two feature maps, representing pixel-wise classification scores of the class 0 ("background") and the class 1 ("segmented area"). The mean raw DSC (fivefold cross-validation) of the ONs, obtained on a set of 44 MRIs after majority voting and postprocessing, was of 0.67, and the HD was of 6.30 mm. To take into account the uncertain borders of the ground truth, DSC ignoring mismatches on the border of the ground truth, were also reported. As most mismatches between the outputs and the ground truth were on noisy borders of organs, there was a considerable difference between the two values. The DSC (five-fold cross-validation) with tolerance to one voxel of the ONs, after majority voting and postprocessing, was of 0.87.

In the same year, 2020, (Ai et al., 2020). proposed an automated segmentation method for T1-weighted and CT scans which used a spatial probabilistic distribution map (SPDM)-based two channel 3D U-Net to make shape and position prior information available for deep learning. First, an atlas was calculated by group-wise registration, and then it was used to non-rigidly register each training volume image getting a deformation field. Second, the deformation field was used to transform the label of the corresponding training image to the template space, and then all the warped labels were summed up to create an SPDM. Third, the region of interest of the image and SPDM were sent to the network as two channels, namely, SPDMfuse, to predict the final segmentation. The

proposed method was evaluated and compared against a conventional 3D U-Net on two datasets, T1-weighted MRI and CT. The proposed two-channel 3D U-Net (SPDMfuse) achieved a DSC of 0.86 ± 0.01 and a HD of 3.56 ± 1.89 mm for the segmentation of the visual pathway as a whole.

Dai et al. (2022). implemented later in 2022 a regional CNN (R-CNN) on T1-weighted MRI for multi-organ auto-delineation, which was a variation of mask R-CNN. This architecture predicts organ positions via bounding boxes, outputting vectors for both the bounding box index and the organ class. First, a backbone network extracted a coarse feature map from the MRI patch, followed by a regional proposal network to compute ROI candidates. These were used to crop the coarse feature map (ROI alignment), and an attention gate highlighted informative areas. Next, R-CNN refined the feature map and extracted the ROI for each organ. The mask network then applied initial segmentations to the refined feature map cropped by the ROI. Finally, the segmentation was achieved through weighted averaging, fusing the initial segmentations based on mask scores from the mask scoring network. Trained with paired MRI and ground truth contours using four types of loss functions the model could derive contours of the ON and other brain parts from an input MRI. Five-fold cross-validation was performed for the assessment of the method. The DSC of the left and right ON were 0.67 ± 0.11 and 0.68 ± 0.11 respectively. The HD95 of the left ON was 3.25 ± 2.11 mm and 2.96 ± 1.32 mm for the right one. The MSD was 0.91 ± 0.38 mm and 0.87 ± 0.29 mm for the left and right ONs, respectively.

Martí-Juan et al. (2022) presented also in 2022 an automated pipeline to extract the ON from T2-weighted fat-saturated scans and developed a 3D CNN model that learned to detect ON lesions in them. Prior to training, scans were processed to remove irrelevant parts of the image and reduce its dimensionality, leaving only the ON and surrounding area. Afterwards, a 3D CNN was implemented for the classification task. It received as input the previously mentioned 3D crops of the ON and was composed of two 3D convolutional layers, each one having a rectifying linear unit non-linearity layer, a max pooling layer to reduce dimensionality, and a dropout layer. A final dense, fully connected layer led to a softmax layer that produced probabilities for the two possible outputs: presence of lesion or not. Additionally, two simpler classification models were implemented to compare the performance of the CNN model and assess its robustness. The first model used a Support Vector Machine (SVM) and the second one, a Random Forest (RF). Both models had as input all the voxels of the 3D crop of the ON. The three models were then evaluated. The only output of the network was a probability value informing if the model detected a lesion in an image or not. The results showed balanced accuracies of around 68 % for validation, with similar sensitivity and specificity, so classification results were not affected by the uneven proportion of positive and negative labels. Results of the model were consistently better than the two other classification methods, obtaining results around mid 50 %, meaning that those simpler models were not able to distinguish the presence/absence of ON lesions.

van Elst et al. (2023). proposed in the following year, 2023, a 3D pipeline for automatic segmentation and quantification of the ON 3D T2-weighted MRI, while accurately differentiating it from CSF along the entire length of the nerve. Multicenter data were obtained and performance was assessed in a tenfold cross-validation ($n = 32$) and on a separate test-set ($n = 8$). The segmentation network architecture was based on the U-Net and manual ON segmentations were performed by an experienced reader using 3D Slicer for generating the training dataset. The pipeline consisted of two main steps. Firstly, a 3D U-Net was employed to automatically segment the ON from high-resolution T2-weighted MRI. The dataset of each site was randomly split into two subsets for training and testing with a partition rate of 0.8 and 0.2, respectively. Secondly, they implemented a quantification method that used the resulting 3D segmentations to extract ON diameter and cross-sectional area along the centerline of the nerve. This automatic approach addressed the limitations of other quantification methods by

enabling quantitative measurements independent of image intensity values or ON orientation. The segmentation network achieved a mean DSC score of 0.84 and a median HD of 0.64 mm. The average spatial agreement achieved by the model was a DSC 0.84 ± 0.03 on the test-set. By allowing a margin of one voxel tolerance to account for uncertainty at the borders of the manual segmentation, the DSC was increased to 0.91 ± 0.04 . The HD95 was 0.60 [0.42 to 1.02] mm on the test set.

3.2.2. Multi-modality images

The first deep learning based method that we found that uses multimodal data was (Mansoor et al., 2016), which developed in 2016 the PASCAL (PARTitioned Shape and Appearance Learning) framework, which automatically segmented the anterior visual pathway. To that end, it used joint partitioned shape models steered by an appearance model that was created using a combination of Gad-enhanced T1, T2-weighted, and Fluid-attenuated inversion recovery (FLAIR) MRI scans along with deep-learning features. PASCAL consisted in a shape localization method using conditional space deep learning, a volumetric multiscale curvelet transform-based intensity normalisation method for robust statistical model, and optimally partitioned statistical shape and appearance models based on regional shape variations for greater local flexibility. A mean DSC of 0.78 ± 0.12 was obtained for the segmentation of the entire AVP, and a DSC of 0.79 ± 0.07 for the ON only using a leave-one-out validation strategy.

Tong et al. (2019), developed later in 2019 a method based on a generative adversarial network (GAN) with shape constraint (SC-GAN) for fully automated head and neck (H&N) organs segmentation on CT and low-field 0.35 T MRI. A deep supervised fully convolutional DenseNet was employed as the segmentation network for voxel-wise prediction. Afterwards, a CNN-based discriminator network was utilised to correct predicted errors and image-level inconsistencies between the prediction and ground truth. The proposed segmentation method was first benchmarked on a public CT dataset, and then on MR images. The performance of the proposed SC-GAN was compared with GAN alone and GAN with the shape constraint (SC) but without the DenseNet (SC-GAN-ResNet). The DSC for the segmentation of the left ON using SC-GAN was 0.72 ± 0.05 and 0.69 ± 0.07 for the right ON.

In 2020, (Liu et al., 2020), developed a dual pyramid networks (DPN) method combining both CT and synthetic MR (sMR), since the first provides bony structure information and the latter superior soft tissue information. Firstly, the sMR images were obtained by feeding CT images into the cycle-consistent adversarial networks (CycleGAN) model. This model was trained using pre-aligned CT and MRI pairs to generate sMR images based on CT input images. Then, independent features were then extracted from CT and sMR separately: the first pyramid network was used to extract semantic features from CT that represent bony structures, and the second pyramid network was used to explore semantic features from sMR that represent the soft tissues. These independent features were then combined and refined via the DPN network to segment several organs including the ON. Both the first and second pyramid networks had a U-Net like architecture. Deep-supervision was used to force the intermediate feature maps to be semantically discriminative at each image scale. Fivefold cross-validation was used to train and validate the proposed segmentation algorithm data. The obtained mean DSC for the segmentation of the left ON was 0.72 ± 0.12 and 0.72 ± 0.20 for the right ON. The mean HD95 of the left ON was 3.15 ± 5.14 mm and 3.14 ± 5.11 mm for the right one; while the mean MSD was 0.94 ± 1.32 mm and 1.10 ± 2.01 mm for the left and right ON, respectively.

One year later, in 2021, (Dai et al., 2021), adopted a similar sMR-aided strategy from their previous work in (Liu et al., 2020), and further improved the network architecture using DPNs. The obtained DSC for the segmentation of the left ON using cone-beam CT (CBCT)+sMRI was 0.78 ± 0.05 and 0.77 ± 0.04 for the right ON. The HD95 and MSD for the left ON were 1.86 ± 1.73 mm and 0.55 ± 0.18 mm respectively; and 2.06 ± 2.69 mm and 0.58 ± 0.32 mm for the right one.

Xie et al. (2023), proposed in 2023 a novel multimodal deep-learning based multi-class network for automated cranial nerves (CNs) tract segmentation without using tractography, ROI placement or clustering, called CNTSeg. To that end, they introduced T1-weighted MRIs, fractional anisotropy (FA) images, and fiber orientation distribution function (fODF) peaks into the training data set. TCNTSeg consisted of two parts: the segmentation network, to generate a cranial nerves tract segmentation binary mask by entering data from different modalities, and the data fusion module, to fuse the CNs features of different modal data. The segmentation network was composed of a 2D encoder-decoder structure built upon the U-Net architecture. The CNs segmentation was responsible for the prediction of the CNs structure on T1-weighted, FA, and fODF images. The multimodal data fusion module used a back-end fusion method to extract the CNs features of different modal data (i.e., T1-weighted, FA, and fODF). The three modalities had three identical but independent U-Net networks, whose outputs were then used to obtain the final CNs prediction. CNTSeg achieved an average DSC for the ON of 0.82.

Finally, later in 2023, Alzahrani et al. (2023), trained and evaluated separate CT and MRI deep learning segmentation models in RayStation (RaySearch AB, Stockholm). Sixty brain CT scans and T1-weighted Gad-enhanced brain MRI available were used. Firstly a brain with OAR atlas was developed as a gold standard, with the OAR being manually delineated using CT and MRI scans in combination. Afterwards, a commercially available 3D U-Net (Çiçek et al., 2016) was used to train all the autosegmentation models (RayStation 11 A, RaySearch Laboratories AB, Stockholm, Sweden). Three MRI models were trained (i) using the original clinical contours based on planning CT and rigidly registered T1-weighted Gad-enhanced MRI (MRIu), (ii) as (i), further edited based on CT anatomy, to meet international consensus guidelines (MRIeCT), and (iii) as (i), further edited based on MRI anatomy (MRIeMRI). Also, two additional CT models were trained using: (iv) original clinical contours (CTu) and (v) clinical contours edited based on CT anatomy (CTeCT). After training, all the models were used to generate automatic contours on the independent validation dataset. The evaluation was done by comparing the generated contour to the gold standard contours in each modality. The obtained DSC values of the segmentation using the MRIeMRI model were of 0.65 ± 0.09 for the left ON, and of 0.68 ± 0.08 for the right one. The achieved MSD values were 0.09 ± 0.03 mm for both the left and right ONs.

4. Discussion

We found that ON segmentation has evolved from classical (including pre-deep learning methods and semi-automated strategies) to deep learning approaches, and concurrently both improved DSC and HD scores, with more automated processing.

Fifteen papers employing traditional methods were reviewed: the methods included intensity-based techniques, single and multi-atlas methods. These methods, from the pre-deep learning era, played a crucial role in the early stages of ON segmentation. They relied heavily on techniques such as atlas-based segmentation, ASM, or VBM. For example, (Isambert et al., 2008; Nguyen et al., 2018), utilised an atlas-based segmentation approach, while (Yiannakas et al., 2010; Yiannakas et al., 2013) and (Nguyen et al., 2018), employed ASM to capture the ON's shape and intensity variations. These early methods, despite being effective and accurate, often required manual intervention to capture the complex anatomy of the ON. It is also important to highlight that single-modality methods, (e.g., T1-weighted or T2-weighted images), were prevalent in early studies. However, these methods had potential limitations in fully depicting ON structure, particularly when image contrast was suboptimal.

The emergence of deep learning has significantly transformed ON segmentation, reflected in the 12 studies reviewed above. These methods, particularly those using CNNs and U-Net architectures, demonstrated good performance in fully automatic ON delineation. For

instance, (Chen et al., 2019). used an REOS framework that incorporated multiple levels of segmentation based ON size and contrast with surrounding tissues. Additionally, studies by (Ai et al., 2020) and (Dai et al., 2022; Dai et al., 2021) developed advanced 3D CNN-based methods that combined intensity data with spatial probabilistic distribution maps, enhancing segmentation accuracy by incorporating anatomical priors. These approaches not only reduced the need for manual intervention, becoming fully automated, but also improved the robustness and reliability of ON segmentation across different imaging datasets. Overall deep learning approaches, particularly those leveraging multi-modality data, offer promising improvements in accuracy and automation, making them more suitable for clinical applications.

Due to its complex morphology and proximity to surrounding structures (see Fig. 1), there are many challenges in the image acquisition of the ON (Chow and Paley, 2021) (see Fig. 3). Its mobility, small size, and the signal interaction between CSF and the orbital soft tissue can affect MR image quality. Accordingly, across the included papers we did not appreciate consensus for the most optimal MRI protocol for ON segmentation. We included multi-modal sequences ranging from highly dedicated and optimised ones for ON (Yiannakas et al., 2010) to more clinical scans such as T1-weighted, T2-weighted and FLAIR, or others that include the use of contrast agents like Gadolinium, and even DWI for delineating the ON through tractography. Faster sequences are being developed, to reduce motion artefacts, without compromising signal-to-noise ratio and/or spatial resolution. With a standard voxel size of $1 \times 1 \times 1$ mm at 3T MRI, there are still significant partial volume effects when delineating ON cross-sectional areas, which range from about 10 to 20 mm². The current voxel size makes cross-sectional area measurement difficult due to partial volume effects. In addition, longitudinal atrophy measurement as indirect change (Prados and Barkhof, 2018), can introduce a systematic error through independent delineation of two separate acquisitions. As for the spinal cord, ON longitudinal atrophy measurements could benefit from registration-based algorithms for changes in its boundary (Prados et al., 2020; Valsasina et al., 2022; Luchetti et al., 2024).

A further consideration is that multimodal MR acquisition is advantageous for capturing the entire ON length, which traverses distinct environments, such as orbital fat, bone from the canalicular segment, and brain tissue in the intracranial region (see Fig. 1). These varying environments pose unique challenges for segmentation. Multimodal imaging allows for the visualisation of different parts of the ON at distinct contrasts, with each modality highlighting specific tissues along its path, providing a more comprehensive view and potentially improving the accuracy of segmentation. However, the challenge of eye movement between acquisitions can introduce misalignments, complicating the integration of data from multiple modalities. Although common space registration techniques can mitigate some of these issues by aligning sequences accurately, these solutions remain complex since they are a combination of rigid registrations for the canicular and cranial area, and non-rigid registrations for the orbital area that might affect the segmentation results. Also, the clinical translation of such multimodal approaches remains difficult due to some singularities of ON MRI

acquisition and the benefits of multimodality, such as enhanced tissue differentiation, must be balanced against practical issues of implementing these techniques into routine clinical practice.

Over time, we have observed a transition from single-atlas methods to multi-atlas methods, accompanied by ongoing debates regarding the advantages of single versus multiple modality approaches. More recently, DL techniques have demonstrated significant improvements in segmentation accuracy, underscoring the impact of these methods in overcoming the limitations associated with traditional approaches. However, progress in ON segmentation development is hindered by the absence of a public MRI dataset with their corresponding labels that can serve as a benchmark for comparing various algorithms. Consequently, while reported Dice coefficients provide useful insights, they are insufficient to determine the best methodology. This challenge may be partly due to a lack of consensus in the research context on the optimal MRI sequences or segmentation methods. Moreover, the lack of comparative studies of different methodologies, especially in MS patients, remains a significant gap.

Nonetheless, there is a pressing need for greater efforts to address the challenge in ON segmentation or MS lesion segmentation within the ON. This would involve curating and making available a comprehensive and diverse dataset with manual annotations to train and benchmark newly developed algorithms effectively. But this public dataset needs to come with an agreement in the measures for reporting and benchmarking the output quality of the ON segmentations. Currently, a wide variety of approaches are used; in the papers we reviewed the three most commonly used measures were (in order) DSC, HD and MSD. These three measures have been proven useful for understanding the goodness of an ON segmentation showing the amount of overlapping (DSC), the size of the miss-segmentations or absolute error (HD) or the precision of the overall shape or mean error (MSD). However, due to the characteristics of the ON shape, the field will benefit from using the skeletonized version of HD and MSD. The skeletonized HD will help us to understand if a method has local errors in a specific area of the ON (i.e. orbital area), values close to 0 will mean that overall shape follows the ground truth or expected result. Whereas that global errors will be assessed by the skeletonized version of MSD, which will show up misalignments or important deviations between the obtained and the expected results and values close to 0 will mean follow the same path.

Although optimized MR protocols may enhance image quality for clinical evaluation, a modality-agnostic approach remains more generalizable for segmentation algorithms, due to the variability in MRI protocols across centers. In the upcoming years, we anticipate a substantial growth in research on the ON, driven by a recognition of its clinical relevance and the availability of increasingly robust automated segmentation methods. The inclusion of the ON in the diagnostic criteria for MS (Foster et al., 2024) will create a demand for improved methods to visualise this region of the central nervous system, accurately delineate it, and detect MS lesions.



Fig. 3. From left to right, 3D T1, 3D FLAIR and 3D T2 axial MRIs of the brain showing healthy optic nerves and chiasm. These conventional MRI sequences are routinely used for assessing the optic nerve.

5. Conclusions

In conclusion, the segmentation of the ON has evolved significantly over the past decades, driven by advances in MRI techniques and the emergence of deep learning algorithms. This review highlights the transition from traditional methods to modern deep learning approaches. The integration of single and multi-modality imaging has further enhanced a precise and reliable ON segmentation. As MRI technology continues to progress, and the inclusion of the ON in the diagnostic criteria of some neurodegenerative diseases, the field of ON segmentation is poised to make substantial contributions in medical diagnostics and treatment planning, offering deeper insights into the mechanisms underlying ON conditions. Additionally, improving visualisation of the ON will help develop markers specific to myelination and neurodegeneration which will help future trials of emerging neuroprotective, remyelinating therapies.

Code availability

This publication does not have any code related to its development and no code has been published.

Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study just reviewing previous literature research.

Ethics approval

This study has been approved by the Ethics Committee of the University Oberta de Catalunya (UOC) stating that this research does not include human subjects participation or any processing of personal data and the research fulfils current legislation on data protection.

Statements and declaration of conflicts of interest

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. C.X.-B., S.K., and E.M.-H. have nothing to disclose. N.S. is a clinical research fellow in a post previously supported by Merck and subsequently by MRC (MR/W019906/1), and has received speaker honoraria from Merck. D.C. is a consultant for Hoffmann-La Roche. In the last three years he has received research funding from Hoffmann-La Roche, the International Progressive MS Alliance, the MS Society, the Medical Research Council, and the National Institute for Health Research (NIHR) University College London Hospitals (UCLH) Biomedical Research Centre. He co-supervises a clinical fellowship at the National Hospital for Neurology and Neurosurgery, London, which is supported by Merck. S.L. received compensation for consulting services and speaker honoraria from Biogen Idec, Novartis, TEVA, Genzyme, Sanofi, Bristol Myers Squibb and Merck, is holder of grants from the Spanish Ministry of Science and Innovation (PI21/0118) and of the Agencia de Gestión de Ayudas Universitarias y de Investigación from the Generalitat de Catalunya (AGAUR SGR 01325). A.T.T. is supported by grants from the MRC (MR/S026088/1), NIHR BRC (541/CAP/OC/818,837), and RoseTrees Trust (PGL21/10,079). FP is partially funded by the National Institute for Health Research Biomedical Research Centre Initiative at UCL and UCLH and received a Guarantors of Brain fellowship 2017–2020.

Consent to participate

This research does not include human subjects participation or any processing of personal data.

Consent for publication

All authors gave consent for the publication of this paper.

CRedit authorship contribution statement

Carla Xena-Bosch: Writing – review & editing, Writing – original draft, Investigation, Formal analysis, Data curation, Conceptualization. **Srikiirti Kodali:** Writing – review & editing, Writing – original draft, Conceptualization. **Nitin Sahi:** Writing – review & editing. **Declan Chard:** Writing – review & editing. **Sara Llufriu:** Writing – review & editing. **Ahmed T Toosy:** Writing – review & editing. **Eloy Martinez-Heras:** Writing – review & editing, Writing – original draft, Conceptualization. **Ferran Prados:** Writing – review & editing, Writing – original draft, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- Ai, D., Zhao, Z., Fan, J., Song, H., Qu, X., Xian, J., et al., 2020. Spatial probabilistic distribution map-based two-channel 3D U-net for visual pathway segmentation. *Pattern. Recognit. Lett.* 138, 601–607. <https://doi.org/10.1016/j.patrec.2020.09.003>.
- Alzahrani, N., Henry, A., Clark, A., Murray, L., Nix, M., Al-Qaisieh, B., 2023. Geometric evaluations of CT and MRI based deep learning segmentation for brain OARs in radiotherapy. *Phys. Med. Biol.* 68. <https://doi.org/10.1088/1361-6560/acf023>.
- Cardenas, C.E., Yang, J., Anderson, B.M., Court, L.E., Brock, K.B., 2019. Advances in auto-segmentation. *Semin. Radiat. Oncol.* 29, 185–197. <https://doi.org/10.1016/j.semradi.2019.02.001>.
- Chen, H., Lu, W., Chen, M., Zhou, L., Timmerman, R., Tu, D., et al., 2019. A recursive ensemble organ segmentation (REOS) framework: application in brain radiotherapy. *Phys. Med. Biol.* 64, 025015. <https://doi.org/10.1088/1361-6560/aaf83c>.
- Chow, L.S., Paley, M.N.J., 2021. Recent advances on optic nerve magnetic resonance imaging and post-processing. *Magn. Reson. Imaging* 79, 76–84. <https://doi.org/10.1016/j.mri.2021.03.014>.
- Chow, L.S., Paley, M.N.J., Hickman, S.J., 2024. Evaluation of optimal interpolation and segmentation of the optic nerves on magnetic resonance images for cross-sectional area measurement. *Int. J. Imaging Syst. Technol.* 34. <https://doi.org/10.1002/ima.23030>.
- Çiçek Ö., Abdulkadir A., Lienkamp S.S., Brox T., Ronneberger O. 3D U-net: learning dense volumetric segmentation from sparse annotation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*. Cham: Springer International Publishing; 2016. pp. 424–432. https://doi.org/10.1007/978-3-319-46723-8_49.
- Crouzen, J.A., Petoukhova, A.L., Wiggeraad, R.G.J., Hutschemaekers, S., Gadella-van Hooijdonk, C.G.M., van der Voort van Zyp, N.C.M.G., et al., 2022. Development and evaluation of an automated EPTN-consensus based organ at risk atlas in the brain on MRI. *RadiOther Oncol.* 173, 262–268. <https://doi.org/10.1016/j.radonc.2022.06.004>.
- Dai, X., Lei, Y., Wang, T., Dhakaan, A.H., McDonald, M., Beitler, J.J., et al., 2021. Head-and-neck organs-at-risk auto-delineation using dual pyramid networks for CBCT-guided adaptive radiotherapy. *Phys. Med. Biol.* 66, 045021. <https://doi.org/10.1088/1361-6560/abd953>.
- Dai, X., Lei, Y., Wang, T., Zhou, J., Rudra, S., McDonald, M., et al., 2022. Multi-organ auto-delineation in head-and-neck MRI for radiation therapy using regional convolutional neural network. *Phys. Med. Biol.* 67. <https://doi.org/10.1088/1361-6560/ac3b34>.
- Dice, L.R., 1945. Measures of the amount of ecologic association between species. *Ecology* 26, 297–302. <https://doi.org/10.2307/1932409>.
- Dowell, N.G., Jenkins, T.M., Ciccarelli, O., Miller, D.H., Wheeler-Kingshott, C.A.M., 2009. Contiguous-slice zonally oblique multislice (CO-ZOOM) diffusion tensor imaging: examples of in vivo spinal cord and optic nerve applications. *J. Magn. Reson. Imaging* 29, 454–460. <https://doi.org/10.1002/jmri.21656>.
- Feng, Y., Chow, L.S., Gowdh, N.M., Ramli, N., Tan, L.K., Abdullah, S., et al., 2023. Gradient-based edge detection with skeletonization (GES) segmentation for magnetic resonance optic nerve images. *Biomed. Signal. Process. Control* 80, 104342. <https://doi.org/10.1016/j.bspc.2022.104342>.
- Foster, M.A., Pontillo, G., Davagnanam, I., Collorone, S., Prados, F., Kanber, B., et al., 2024. Improving criteria for dissemination in space in multiple sclerosis by including additional regions. *Ann. Clin. Transl. Neurol.* <https://doi.org/10.1002/actn.3.52170>.
- Harrigan, R.L., Plassard, A.J., Bryan, F.W., Caires, G., Mawn, L.A., Dethrage, L.M., et al., 2016. Disambiguating the optic nerve from the surrounding cerebrospinal fluid:

- application to MS-related atrophy. *Magn. Reson. Med.* 75, 414–422. <https://doi.org/10.1002/mrm.25613>.
- He, J., Zhang, F., Xie, G., Yao, S., Feng, Y., Bastos, D.C.A., et al., 2021. Comparison of multiple tractography methods for reconstruction of the retinogeniculate visual pathway using diffusion MRI. *Hum. Brain Mapp.* 42, 3887–3904. <https://doi.org/10.1002/hbm.25472>.
- Hernowo, A.T., Boucard, C.C., Jansonius, N.M., Hooymans, J.M.M., Cornelissen, F.W., 2011. Automated morphometry of the visual pathway in primary open-angle glaucoma. *Invest. Ophthalmol. Vis. Sci.* 52, 2758–2766. <https://doi.org/10.1167/iovs.10-5682>.
- Hickman, S.J., 2007. Optic nerve imaging in multiple sclerosis. *J. Neuroimaging* 17, 42S–45S. <https://doi.org/10.1111/j.1552-6569.2007.00136.x>. Suppl 1.
- Isambert, A., Dhermain, F., Bidault, F., Commowick, O., Bondiau, P.-Y., Malandain, G., et al., 2008. Evaluation of an atlas-based automatic segmentation software for the delineation of brain organs at risk in a radiation therapy clinical context. *Radiother Oncol.* 87, 93–99. <https://doi.org/10.1016/j.radonc.2007.11.030>.
- Liu, P., Sun, Y., Zhao, X., Yan, Y., 2023. Deep learning algorithm performance in contouring head and neck organs at risk: a systematic review and single-arm meta-analysis. *Biomed. Eng. Online* 22, 104. <https://doi.org/10.1186/s12938-023-01159-y>.
- Liu, Y., Lei, Y., Fu, Y., Wang, T., Zhou, J., Jiang, X., et al., 2020. Head and neck multi-organ auto-segmentation on CT images aided by synthetic MRI. *Med. Phys.* 47, 4294–4302. <https://doi.org/10.1002/mp.14378>.
- Luchetti, L., Prados, F., Cortese, R., Gentile, G., Calabrese, M., Mortilla, M., et al., 2024. Evaluation of cervical spinal cord atrophy using a modified SIENA approach. *Neuroimage*, 120775. <https://doi.org/10.1016/j.neuroimage.2024.120775>.
- Mansoor, A., Cerrolaza, J.J., Idrees, R., Biggs, E., Alsharid, M.A., Avery, R.A., et al., 2016. Deep learning guided partitioned shape model for anterior visual pathway segmentation. *IEEE Trans. Med. Imaging* 35, 1856–1865. <https://doi.org/10.1109/TMI.2016.2535222>.
- Martí-Juan, G., Frías, M., García-Vidal, A., Vidal-Jordana, A., Alberich, M., Calderon, W., et al., 2022. Detection of lesions in the optic nerve with magnetic resonance imaging using a 3D convolutional neural network. *Neuroimage Clin.* 36, 103187. <https://doi.org/10.1016/j.nicl.2022.103187>.
- Miller, N., Liu, Y., Krivochenitser, R., Rokers, B., 2019. Linking neural and clinical measures of glaucoma with diffusion magnetic resonance imaging (dMRI). *PLoS One* 14, e0217011. <https://doi.org/10.1371/journal.pone.0217011>.
- Mlynarski, P., Delingette, H., Alghamdi, H., Bondiau, P.-Y., Ayache, N., 2020. Anatomically consistent CNN-based segmentation of organs-at-risk in cranial radiotherapy. *J. Med. Imaging (Bellingham)* 7, 014502. <https://doi.org/10.1117/1.JMI.7.1.014502>.
- Nguyen, H.-G., Sznitman, R., Maeder, P., Schalenbourg, A., Peroni, M., Hrbacek, J., et al., 2018. Personalized anatomic eye model from T1-weighted volume interpolated gradient echo magnetic resonance imaging of patients with uveal melanoma. *Int. J. Radiat. Oncol. Biol. Phys.* 102, 813–820. <https://doi.org/10.1016/j.ijrobp.2018.05.004>.
- Noble, J.H., Dawant, B.M., 2011. An atlas-navigated optimal medial axis and deformable model algorithm (NOMAD) for the segmentation of the optic nerves and chiasm in MR and CT images. *Med. Image Anal.* 15, 877–884. <https://doi.org/10.1016/j.media.2011.05.001>.
- Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., et al., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 372, n71. <https://doi.org/10.1136/bmj.n71>.
- Panda, S., Asman, A.J., Khare, S.P., Thompson, L., Mawn, L.A., Smith, S.A., et al., 2014. Evaluation of multi-atlas label fusion for in vivo MRI orbital segmentation. *J. Med. Imaging (Bellingham)* 1. <https://doi.org/10.1117/1.JMI.1.2.024002>.
- Prados, F., Ashburner, J., Blaiotta, C., Brosch, T., Carballido-Gamio, J., Cardoso, M.J., et al., 2017. Spinal cord grey matter segmentation challenge. *Neuroimage* 152, 312–329. <https://doi.org/10.1016/j.neuroimage.2017.03.010>.
- Prados, F., Barkhof, F., 2018. Spinal cord atrophy rates: ready for prime time in multiple sclerosis clinical trials? *Neurology* 91, 157–158. <https://doi.org/10.1212/WNL.0000000000005873>.
- Prados, F., Moccia, M., Johnson, A., Yiannakas, M., Grussu, F., Cardoso, M.J., et al., 2020. Generalised boundary shift integral for longitudinal assessment of spinal cord atrophy. *Neuroimage* 209, 116489. <https://doi.org/10.1016/j.neuroimage.2019.116489>.
- Ramli, N.M., Sidek, S., Rahman, F.A., Peyman, M., Zahari, M., Rahmat, K., et al., 2014. Novel use of 3T MRI in assessment of optic nerve volume in glaucoma. *Graefes. Arch. Clin. Exp. Ophthalmol.* 252, 995–1000. <https://doi.org/10.1007/s00417-014-2622-6>.
- Sabuncu, M.R., Yeo, B.T.T., Van Leemput, K., Fischl, B., Golland, P., 2010. A generative model for image segmentation based on label fusion. *IEEE Trans. Med. Imaging* 29, 1714–1729. <https://doi.org/10.1109/TMI.2010.2050897>.
- Tan, Y.H., Chow, L.S., Chuah, J.H., Lai, K.W., 2022. Diagnosis of optic neuritis using magnetic resonance images. *Multimed. Tools. Appl.* 81, 41979–41993. <https://doi.org/10.1007/s11042-022-13520-9>.
- Tong, N., Gou, S., Yang, S., Cao, M., Sheng, K., 2019. Shape constrained fully convolutional DenseNet with adversarial training for multiorgan segmentation on head and neck CT and low-field MR images. *Med. Phys.* 46, 2669–2682. <https://doi.org/10.1002/mp.13553>.
- Valsasina, P., Horsfield, M.A., Meani, A., Gobbi, C., Gallo, A., Rocca, M.A., et al., 2022. Improved assessment of longitudinal spinal cord atrophy in multiple sclerosis using a registration-based approach: relevance for clinical studies. *J. Magn. Reson. Imaging* 55, 1559–1568. <https://doi.org/10.1002/jmri.27937>.
- van Elst, S., de Bloeme, C.M., Noteboom, S., de Jong, M.C., Moll, A.C., Göricke, S., et al., 2023. Automatic segmentation and quantification of the optic nerve on MRI using a 3D U-net. *J. Med. Imaging (Bellingham)* 10, 034501. <https://doi.org/10.1117/1.JMI.10.3.034501>.
- Vrtovec, T., Močnik, D., Strojani, P., Pernuš, F., Ibragimov, B., 2020. Auto-segmentation of organs at risk for head and neck radiotherapy planning: from atlas-based to deep learning methods. *Med. Phys.* 47. <https://doi.org/10.1002/mp.14320> e929–e950.
- Wheeler-Kingshott, C.A.M., Parker, G.J.M., Symms, M.R., Hickman, S.J., Tofts, P.S., Miller, D.H., et al., 2002. ADC mapping of the human optic nerve: increased resolution, coverage, and reliability with CSF-suppressed ZOOM-EPI. *Magn. Reson. Med.* 47, 24–31. <https://doi.org/10.1002/mrm.10016>.
- Wheeler-Kingshott, C.A.M., Trip, S.A., Symms, M.R., Parker, G.J.M., Barker, G.J., Miller, D.H., 2006. In vivo diffusion tensor imaging of the human optic nerve: pilot study in normal controls. *Magn. Reson. Med.* 56, 446–451. <https://doi.org/10.1002/mrm.20964>.
- Xie, L., Huang, J., Yu, J., Zeng, Q., Hu, Q., Chen, Z., et al., 2023. CNTSeg: a multimodal deep-learning-based network for cranial nerves tract segmentation. *Med. Image Anal.* 86, 102766. <https://doi.org/10.1016/j.media.2023.102766>.
- Yiannakas, M.C., Toosy, A.T., Raftopoulos, R.E., Kapoor, R., Miller, D.H., Wheeler-Kingshott, C.A.M., 2013. MRI acquisition and analysis protocol for in vivo intraorbital optic nerve segmentation at 3T. *Invest. Ophthalmol. Vis. Sci.* 54, 4235–4240. <https://doi.org/10.1167/iovs.13-12357>.
- Yiannakas, M.C., Wheeler-Kingshott, C.A.M., Berry, A.M., Chappell, K., Henderson, A., Kolappan, M., et al., 2010. A method for measuring the cross sectional area of the anterior portion of the optic nerve in vivo using a fast 3D MRI sequence. *J. Magn. Reson. Imaging* 31, 1486–1491. <https://doi.org/10.1002/jmri.22202>.