



## OPEN ACCESS

## EDITED BY

Xing Lu,  
University of California, San Diego,  
United States

## REVIEWED BY

Yuening Zhang,  
University of Oklahoma University College,  
United States  
Yixiong Zhou,  
Shanghai Jiao Tong University, China

## \*CORRESPONDENCE

Daniel T. Ginat,  
✉ dtg1@uchicago.edu,  
✉ ginatd01@gmail.com

RECEIVED 02 September 2025

REVISED 01 December 2025

ACCEPTED 15 December 2025

PUBLISHED 08 January 2026

## CITATION

Shetty A, Babu Jai Shanker RR, Illimoottil M,  
Chohdry Q, Kadkol S, Illimoottil S, Holliman M  
and Ginat DT (2026) Automated segmentation  
of the lacrimal gland on non-contrast versus  
post-contrast T1-weighted MRI sequences.  
*Front. Phys.* 13:1697903.  
doi: 10.3389/fphy.2025.1697903

## COPYRIGHT

© 2026 Shetty, Babu Jai Shanker, Illimoottil,  
Chohdry, Kadkol, Illimoottil, Holliman and  
Ginat. This is an open-access article  
distributed under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#). The  
use, distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Automated segmentation of the lacrimal gland on non-contrast versus post-contrast T1-weighted MRI sequences

Amogh Shetty<sup>1</sup>, Ramkumar Rajabathar Babu Jai Shanker<sup>2</sup>,  
Mathew Illimoottil<sup>3</sup>, Qasim Chohdry<sup>3</sup>, Shrinidhi Kadkol<sup>4</sup>,  
Sarah Illimoottil<sup>3</sup>, Matthew Holliman<sup>2</sup> and Daniel T. Ginat<sup>5\*</sup>

<sup>1</sup>Department of Biology, Rensselaer Polytechnic Institute, Troy, NY, United States, <sup>2</sup>Biological Sciences Division, University of Chicago, Chicago, IL, United States, <sup>3</sup>University of Missouri-Kansas City School of Medicine, Kansas City, MO, United States, <sup>4</sup>University of Illinois at Chicago College of Medicine, Chicago, IL, United States, <sup>5</sup>Department of Radiology, Section of Neuroradiology, University of Chicago, Chicago, IL, United States

**Purpose:** The lacrimal glands are small orbital exocrine structures responsible for tear production. Segmentation on MRI is challenging due to their small size, low contrast with adjacent tissues, and partial representation across slices. This study evaluates U-Net based models for automated lacrimal gland segmentation on non-contrast T1-weighted (AX-T1) and contrast-enhanced fat-suppressed (POST-AX-T1-FS) MRI.

**Methods:** Eighty-six patients with high-resolution orbital MRI were retrospectively analyzed. Manual gland annotations were created in 3D Slicer. A U-Net architecture was trained with 4-fold cross-validation on an 80:20 train-test split. Performance was assessed on a hold-out set using Dice Similarity Coefficient (DSC), Intersection over Union (IoU), and Hausdorff Distance.

**Results:** POST-AX-T1-FS achieved the highest performance (mean DSC  $0.79 \pm 0.19$ , IoU  $0.68 \pm 0.19$ ), outperforming AX-T1. Volume correlation with ground truth was 0.81 for POST-AX-T1-FS and 0.71 for AX-T1. Most errors were false negatives in abnormal gland morphology. Qualitative review showed anatomically consistent segmentations, especially with region-prioritized sampling.

**Conclusion:** CNN-based models show ability to segment lacrimal glands from orbital MRI, though performance is moderate with Dice scores around 0.79. Non-contrast sequences may provide reasonably accurate segmentations, but further refinement and broader validation are required. With continued optimization and larger, more diverse datasets, these models may eventually support more consistent gland delineation in research and early exploratory clinical use.

## KEYWORDS

artificial intelligence, contrast enhanced MRI, fat suppression, lacrimal gland, magnetic resonance imaging

# 1 Introduction

The lacrimal gland is an exocrine gland located in the upper outer portion of each eye socket, within the lacrimal fossa of the frontal bone. Its main role is to produce the aqueous layer of the tear film, which is essential for lubricating the eye, maintaining clear vision, and protecting the ocular surface [1, 2]. The gland receives sensory innervation from the lacrimal nerve (CN V1) and parasympathetic input from the facial nerve (CN VII), which regulates tear production in response to stimuli like irritation or emotion [3]. Accurate segmentation of the lacrimal gland in medical imaging holds diagnostic and research value. In cases of autoimmune diseases, such as Sjogren syndrome, or orbital tumors, identifying subtle morphological changes in the gland can aid early detection and monitoring [4–6]. In particular, measuring the volume of the lacrimal gland provides a valuable biomarker for detecting inflammation, atrophy, or abnormal growth. Gland enlargement may indicate infection, sarcoidosis, or neoplasia, while atrophy is often seen in chronic autoimmune conditions or after radiation exposure [7]. Automated segmentation enables volumetric assessment and structural analysis, which can be used to track disease progression or response to therapy over time.

Gadolinium-based contrast agents (GBCAs) are commonly used in MRI to enhance the visibility of certain tissues by shortening T1 relaxation times in tissues that are in the vicinity of GBCAs, thereby increasing signal intensity on T1-weighted images [8]. While effective, their use is associated with potential risks, including nephrogenic systemic fibrosis, allergic reactions, and accumulation of gadolinium in neural tissues [9–11]. As a result, there is growing interest in non-contrast imaging techniques. However, the accuracy of these non-contrast methods for lacrimal gland segmentation remains unclear.

Fat suppression is a category of magnetic resonance imaging (MRI) techniques used to reduce or eliminate the signal from adipose tissue, allowing for clearer visualization of surrounding structures [12]. This is particularly important in regions like the orbit, where the lacrimal gland is surrounded by a significant amount of orbital fat. Without fat suppression, the high signal intensity of fat on T1-weighted images can obscure or mask subtle pathological changes in the lacrimal gland, such as inflammation, atrophy, or neoplastic lesions [13]. Fat suppression is especially useful in post-contrast imaging, where it helps highlight areas of pathological enhancement that might otherwise blend in with the bright signal of surrounding fat [14]. However, the extent to which fat suppression helps with lacrimal gland segmentation remains unclear.

Artificial intelligence, particularly Convolutional Neural Networks (CNNs), has proven effective in medical imaging tasks such as tumor detection, segmentation, and characterization [15]. The U-Net architecture, a CNN designed for semantic segmentation, is widely used in medical imaging [16]. It features an encoder-decoder structure with skip connections that preserve spatial detail by linking corresponding layers. U-Net performs well even with limited labeled data and has inspired variants like 3D U-Net and Attention U-Net [17, 18]. For instance, it has been used to segment extraocular muscles on CT, demonstrating its utility in delineating fine anatomical structures in orbital imaging [19].

Several studies have attempted automated segmentation of the lacrimal gland from MRI or CT scans. One study trained models ranging from standard U-Nets to nnU-Net across a variety of head and neck organs, including the lacrimal gland [20]. Using the Dice Similarity Coefficient (DSC), it reported mean DSC values between 0.396 and 0.663 for lacrimal gland segmentation in T1-weighted MR images, and identified the lacrimal gland as one of the most difficult structures to segment because of its small size, variable shape, and limited contrast. Another multi-institutional study further advanced head and neck segmentation by training an nnU-Net pipeline on paired CT and T1-weighted MRI from 296 patients, combining data from the HaN-Seg Challenge and TCIA datasets. MRI was rigidly registered to CT, and both modalities were stacked during training, with modality dropout used to enable both single- and dual-modality input. The pipeline achieved state-of-the-art performance on 30 organs-at-risk with a mean DSC of 78.12% and a mean Hausdorff distance of 3.42 mm, suggesting that similar multi-modal strategies could help with lacrimal gland segmentation even though lacrimal performance was not separately reported [21].

Beyond healthy lacrimal glands, closely related peri-orbital tasks have demonstrated that deep learning can perform segmentation and volumetric measurement of ocular adnexal lymphoma (OAL), using T1-weighted, T2-weighted, and contrast-enhanced sequences with and without fat suppression [22]. The network achieved excellent agreement with expert annotations and particularly strong performance on fat-suppressed T2-weighted images, enabling reliable volumetric tumor burden assessment in a multi-center setting. This work shows that nnU-Net can handle subtle, heterogeneous soft-tissue lesions in the crowded orbital region, which is encouraging for lacrimal and peri-lacrimal target structures.

Other studies have focused on anatomically adjacent components of the lacrimal drainage system. One study proposed a fully automated pipeline for 3D reconstruction of the bony nasolacrimal canal (NLC) from CT, using intensity-based preprocessing and rule-based region growing, and demonstrated accurate canal extraction relative to expert annotations [23]. More recently, Haylaz et al. applied nnU-Net v2 to segment the NLC on cone-beam CT (CBCT) images from 100 patients, reporting strong performance metrics (Dice coefficient ~0.8465) [24]. These results indicate that self-configuring encoder-decoder architectures can reliably segment thin, tubular structures within the orbit, despite variable canal morphology and limited contrast, and reinforce the suitability of nnU-Net-style pipelines for lacrimal-system tasks.

Functional imaging work in prostate-specific membrane antigen (PSMA) PET/CT has also shown that glands with high physiologic tracer uptake, including the lacrimal glands, can be incorporated into multi-organ deep-learning segmentation frameworks. Notably, the use of PSMA results in high contrast within lacrimal gland tissues, making them especially suitable for deep-learning segmentation. One group used a combination of self-supervised pre-training on 526 unlabeled scans and supervised fine-tuning on 100 labelled cases, achieving high DSC and sensitivity across organs with intense tracer uptake, including lacrimal glands [25]. More recently, Yazdani et al. proposed a Swin UNETR-based model for lesion and organ-at-risk segmentation on [<sup>68</sup>Ga]Ga-PSMA-11 PET/CT images using self-supervised pre-training followed by fine-tuning on 100 annotated patients [26]. Their framework segmented

10 organs-at-risk, explicitly including bilateral lacrimal glands as a small-volume class that was delineated on PET-only slices because of their small size and intense PSMA avidity. These PSMA PET/CT studies demonstrate that transformer-based architectures such as Swin-UNETR can successfully learn small gland classes embedded in a whole-body context, and they provide further evidence that lacrimal gland segmentation can be integrated into multi-organ pipelines.

Lacrimal gland segmentation from CT scans has also been explored, especially in patients with Graves' orbitopathy. Using orbital CT from 701 patients, they trained a specialised encoder-decoder architecture and compared it to several conventional networks (Attention U-Net, DeepLabV3+, SegNet, HarDNet-MSEG) for segmenting the eyeball, extra-ocular muscles, optic nerve, and lacrimal gland [27]. On selected axial and coronal slices, their proposed network achieved high Dice coefficients ( $>0.9$  for several orbital tissues) and substantially improved qualitative boundary delineation compared to baseline models. For the lacrimal gland, they reported DSC values on the order of 0.87 and 0.79 in axial and coronal views, respectively, highlighting both the feasibility and the residual difficulty of lacrimal segmentation in routine CT. Taken together, these MRI, CT, CBCT, meibography, and PSMA PET/CT studies indicate that small, anatomically variable peri-orbital structures, tumor and non-tumor, can be segmented reliably using modern CNN and transformer-based architectures, and they motivate the development of specialized lacrimal gland segmentation models that leverage both structural and functional imaging.

To our knowledge, no prior study has demonstrated automated segmentation of the healthy lacrimal gland from contrast-enhanced and fat suppressed MRI. This study aims to compare U-Net performance on both contrast-enhanced and non-contrast MRI sequences for lacrimal gland segmentation.

## 2 Materials and methods

### 2.1 Dataset

We analyzed 86 sets of pretreatment baseline MRI scans of the head and neck region collected between January and September of 2018. The MRI scans were selected based on whether they had non-contrast axial T1-weighted (AX-T1) and contrast-enhanced T1-weighted fat-suppressed (POST-AX-T1-FS). Scans with artifacts were excluded. We ultimately had 74 POST-AX-T1-FS scans and 80 AX-T1 scans from 81 patients' cases, comprising 55 females (68%), 26 males (32%), and five patients of unreported gender (6%). The median age of the patients was 54.5 years (17–90 years).

### 2.2 Manual segmentation of lacrimal glands

All images were conducted with 1.5 and 3.0 T with a slice thickness of 3–4 mm. Scans had axial resolutions of  $512 \times 512$  or  $256 \times 256$  voxels, with voxel sizes from 0.3 to 0.7 mm. The imaging protocol included 2D T1-weighted axial fast spin-echo sequence (AX-T1) and 2D T1-weighted axial sequence after contrast injection

with fat suppression (POST-AX-T1-FS). The gadolinium-based contrast agent DOTAREM was used to enhance tissue contrast.

The segmentation of the lacrimal glands was performed on POST-AX-T1-FS and AX-T1 images by a group of students and confirmed by a senior radiologist. Manual segmentations were performed using 3D Slicer (version 5.6.2, <https://www.slicer.org/>) (Figure 1). Each MRI scan included the region of the head extending from the upper chin to the mid-scalp. Segmentations on POST-AX-T1-FS and AX-T1 images were independently performed. All original segmentations were manually corrected according to suitability and lacrimal gland margins.

### 2.3 Data preprocessing and augmentation

Both contrast-enhanced and non-contrast sequences were subjected to pre-processing transformations like intensity normalization (maximum-minimum rescaling) and isometric resampling (ensuring the pixels of different scan matrices have standardized scales) to obtain 512 by 512 pixel width scan slices. Data preprocessing helps the segmentation algorithm during the learning process by standardizing a certain set of features without removing scan specific information.

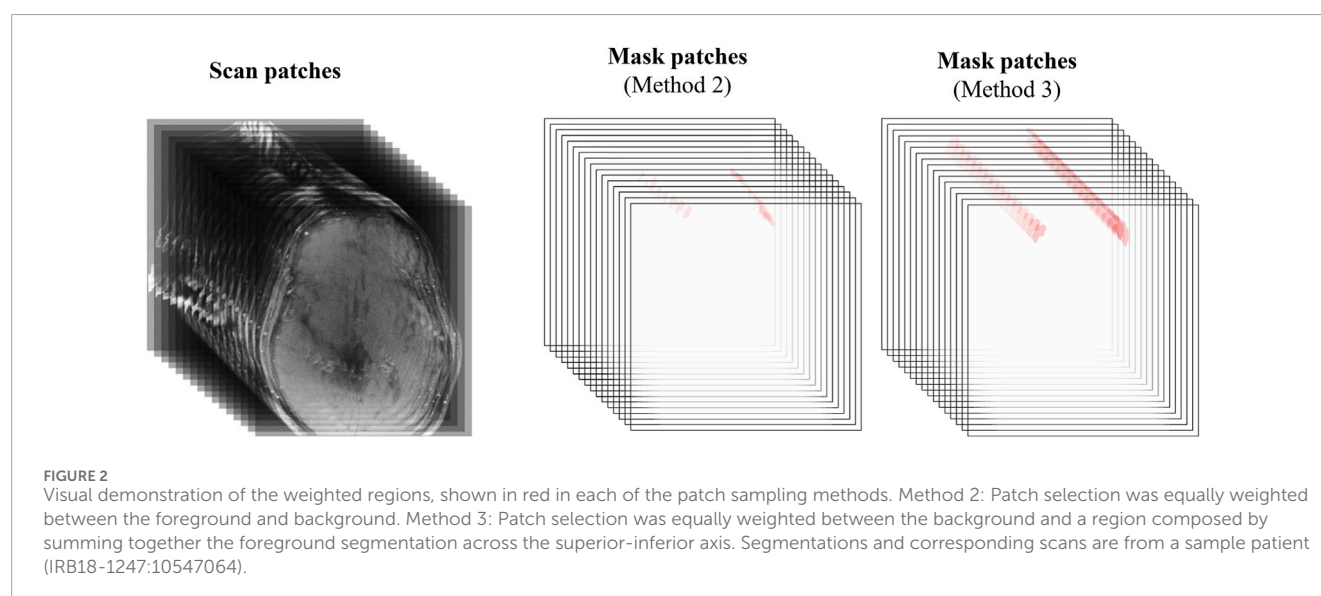
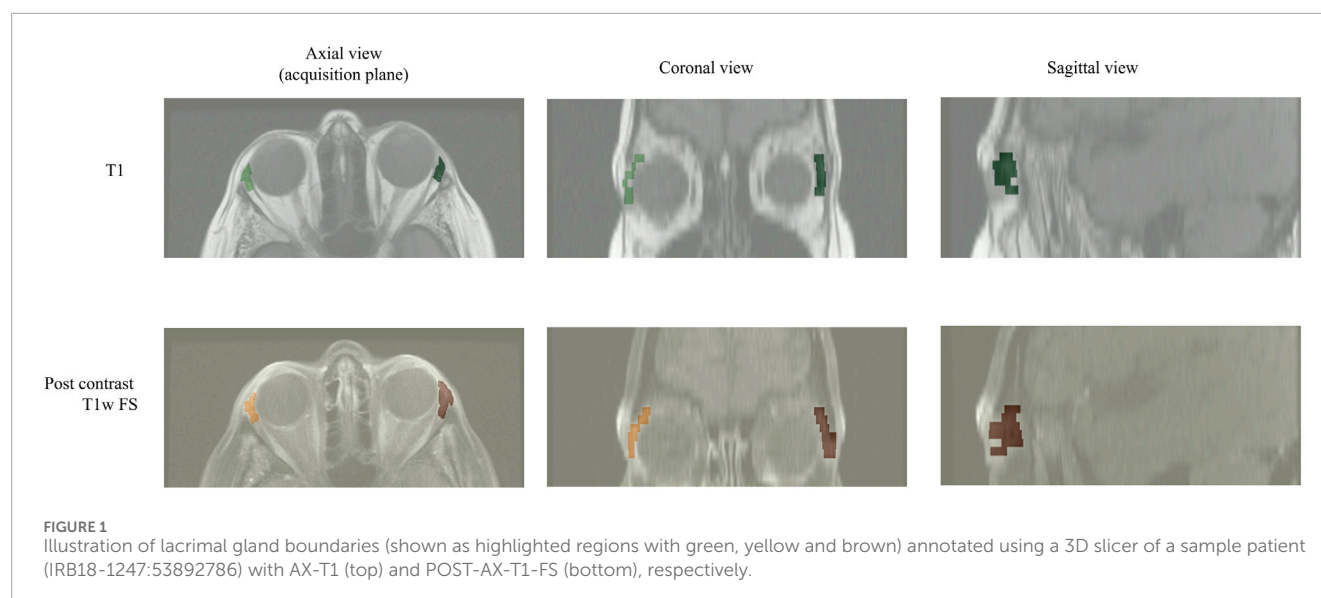
Data augmentation and pre-processing were used to improve generalizability. Both contrast-enhanced and non-contrast sequences underwent spatial transformations such as flipping, scaling, Gaussian noise addition, and limited-angle rotations. Uniform 2D patches were extracted from each scan and used for training to increase sample size and reduce computational load. Since the lacrimal glands occupy only ~2% of the scan and are absent from many slices, three oversampling methods were used to increase foreground patch frequency: (i) random sampling (control), (ii) sampling weighted toward foreground segmentations, and (iii) sampling weighted toward the expected lacrimal region regardless of gland presence (Figure 2). Method (iii) was ultimately used to reduce false positives by exposing the model to more samples from the orbital region.

### 2.4 Architecture

A standard 2D U-Net was used for lacrimal gland segmentation due to its strong performance on small datasets and low computational demands. The 2D design enables efficient, slice-wise segmentation of axial MRI scans, balancing spatial feature capture with processing efficiency. To reduce false positives, post-processing retained only the largest contiguous segmented region per prediction.

### 2.5 Loss functions

Loss functions guided the model by quantifying differences between predicted and true segmentations. Region-based metrics, including Dice and Jaccard (IoU) losses, were evaluated for overlap quality. Weighted cross-entropy was also considered to address class imbalance. Dice loss was ultimately chosen for its effectiveness in reducing false positives by emphasizing overlap with ground truth.



## 2.6 Training and experimental design

The dataset included eighty-one patients: seventy-four contrast-enhanced POST-AX-T1-FS and eighty non-contrast AX-T1 scans, with seventy-three patients having both modalities. Fourteen patients with both scan types were randomly selected as a hold-out test set, following an 80:20 train-test split. From the remaining patients, individual axial slices were resampled and randomly assigned to the training and validation sets, yielding 1,307 AX-T1 and 1,139 POST-AX-T1-FS slices for training (Figure 3). A 2D U-Net was trained on individual slices using four-fold cross-validation over 200 epochs, chosen empirically. In each fold, one of four subsets served as validation while the rest were used for training; final predictions were averaged across folds (Figure 4). To maintain consistency, models were trained and evaluated separately on each modality, enabling a direct comparison of segmentation performance between contrast-enhanced and non-contrast scans.

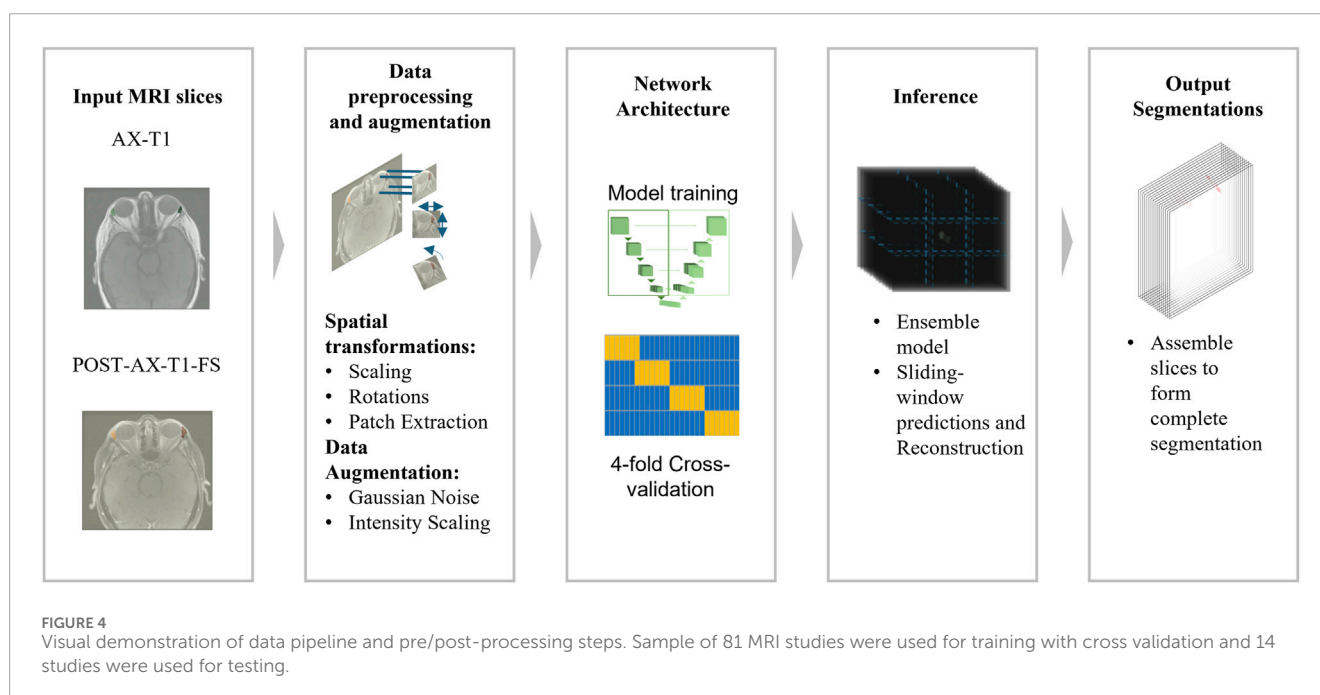
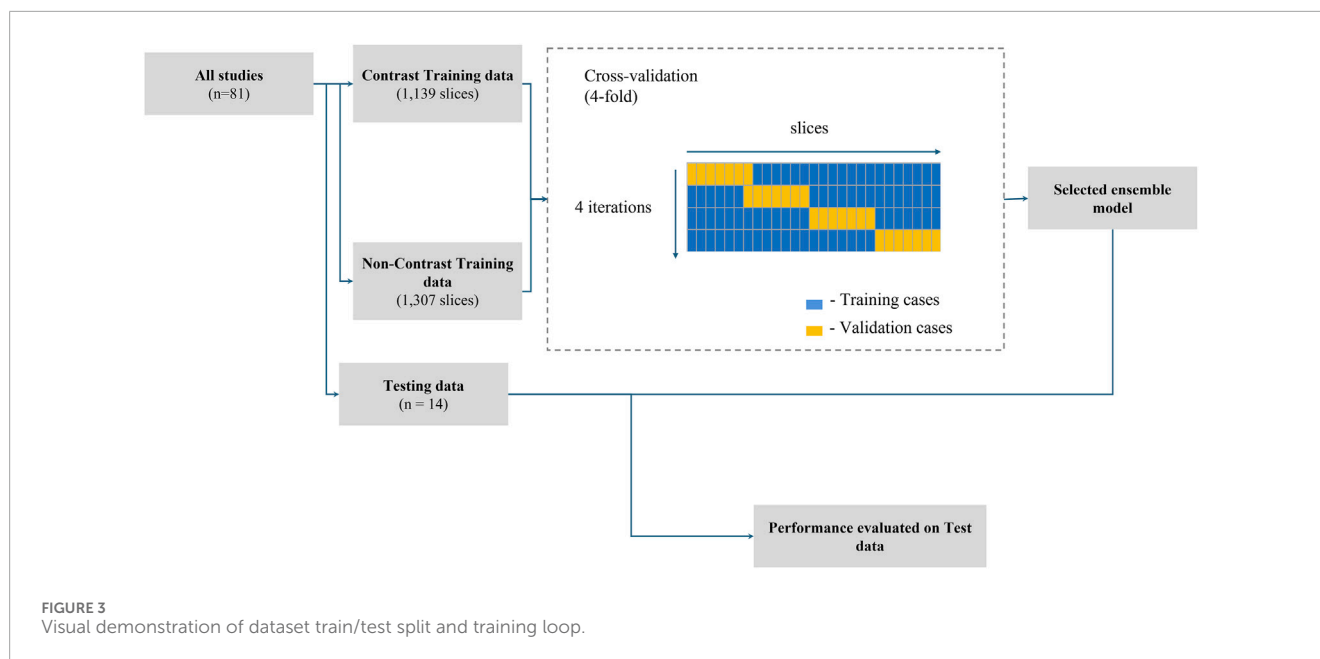
## 2.7 Evaluation methodology

Segmentation performance was quantitatively evaluated using region-based metrics, primarily the Dice score and Intersection over Union (IoU), with Dice serving as the main measure for volumetric accuracy. Hausdorff Distance (HD) was also evaluated to capture the maximum contour deviation.

## 2.8 Saliency maps

Saliency maps help provide insight into an AI algorithm's decision-making process by visually representing parts of the input that are most important in a model's prediction. These visual explanations are generated from the gradients, which help train the model by adjusting the model parameters. In the context of multi-layer CNNs, activations are the output of a particular layer after the





convolution/up-sampling operation is carried out, and gradients are the derivatives of the loss function that are propagated back to the layer under consideration. So, while activations can be thought of as the “state” of a layer, gradients are the “direction and magnitude” of updates that need to be made to the convolutional layer’s weights.

Gradient-weighted Class Activation Mapping (Grad-CAM), as the name indicates, uses the gradients to weight the activations within a particular layer to create a visual heat map that highlights the regions of the image that are important to that layer [28]. By strategically selecting one representative layer at each level of

depth of the UNet model, valuable insights can be derived on how the model processes information and makes segmentation decisions at different stages of feature extraction and reconstruction. In this study, the Gradient-weighted Class Activation Mapping (Grad-CAM) methodology proposed by Selvaraju et al. is used to generate saliency maps. The visualizations are derived using the PyTorch 2.6.0 implementation of the GradCAM (version 1.5.5) methodology. The codes used for this section are publicly available and can be found in the supplemental data section.

## 3 Results

### 3.1 Patient characteristics

The baseline characteristics of patients in training and test set are included in [Tables 1, 2](#). Statistical tests were carried out using Python Scipy library (version 1.15.2). No significant differences were found between the training and testing set for age and sex.

### 3.2 Quantitative evaluation and model performance

The model was trained using 4-folds of the training set to predict segmentations on the hold-out test set. Full-image predictions were reconstructed by stitching together individual patches using a sliding window approach with 25% overlap between adjacent patches. The performances of segmentations are assessed using several different metrics with the mean and interquartile ranges of the metrics demonstrated in [Table 3](#). Shapiro test confirmed non-normal distribution ( $p < 0.001$  for all). The scatter plots of LGV from ground truth and predicted segmentations are presented in [Figure 5](#), where a clearer correlation is seen between predicted and ground truth volumes for POST-AX-T1-FS ( $p < 0.001$ ) compared to the AX-T1 ( $p = 0.004$ ).

**TABLE 1** Baseline characteristics of patients having AX-T1 scans; p-values calculated using two-tailed, unpaired t-test (age), and chi-squared test (sex).

Variables	Train	Test	p-value
Age in years (mean $\pm$ sd)	54.9 $\pm$ 20.7	48.4 $\pm$ 17.6	0.29
Sex			0.63
M	19 (29%)	4 (29%)	
F	43 (65%)	10 (71%)	
Unreported	4 (6%)	0 (0%)	

**TABLE 2** Baseline characteristics of patients having POST-AX-T1-FS scans; p-values calculated using two-tailed, unpaired t-test (age), and chi-squared test (sex).

Variables	Train	Test	p-value
Age in years (mean $\pm$ sd)	56.6 $\pm$ 19.8	48.4 $\pm$ 17.6	0.17
Sex			0.48
M	17 (28%)	4 (29%)	
F	38 (63%)	10 (71%)	
Unreported	5 (8%)	0 (0%)	

### 3.3 Saliency maps

The Grad-CAM visualizations are illustrated in [Figure 7](#). In layers (1) and (2) of the encoder pathway, the gradient-weighted activations are diffuse and widespread. This shows that the earlier layers of the trained U-net capture general contextual information and basic features from the input image. In layers (3) and (4), the activations get more concentrated around certain specific regions. The brighter spots indicate regions of higher gradient-weighted activations, which coincide with the location of lacrimal glands, orbital area, and skull boundary. The bottleneck layer (5) represents the deepest part of the network and the highest level of abstraction extracted by the network. In this layer, the bilateral activation pattern suggests that the model identifies the paired nature of the lacrimal glands, even though the model is trained and evaluated on individual patches that may not contain both lacrimal glands in the same patch. The activations also seem more focused on the precise location of the lacrimal glands than those from the previous layers. Still, they also have not precisely delineated the boundaries in this stage. In layers (6) and (7) of the decoder pathway, the activations become more anatomically precise as the network starts the image reconstruction process. In layers (8) and (9), the spatial features from the previous layers in the encoder pathway are used to further refine the lacrimal gland boundaries. The saliency maps demonstrate that the model focuses on anatomically correct regions relevant to the precise segmentation of the lacrimal gland.

## 4 Discussion

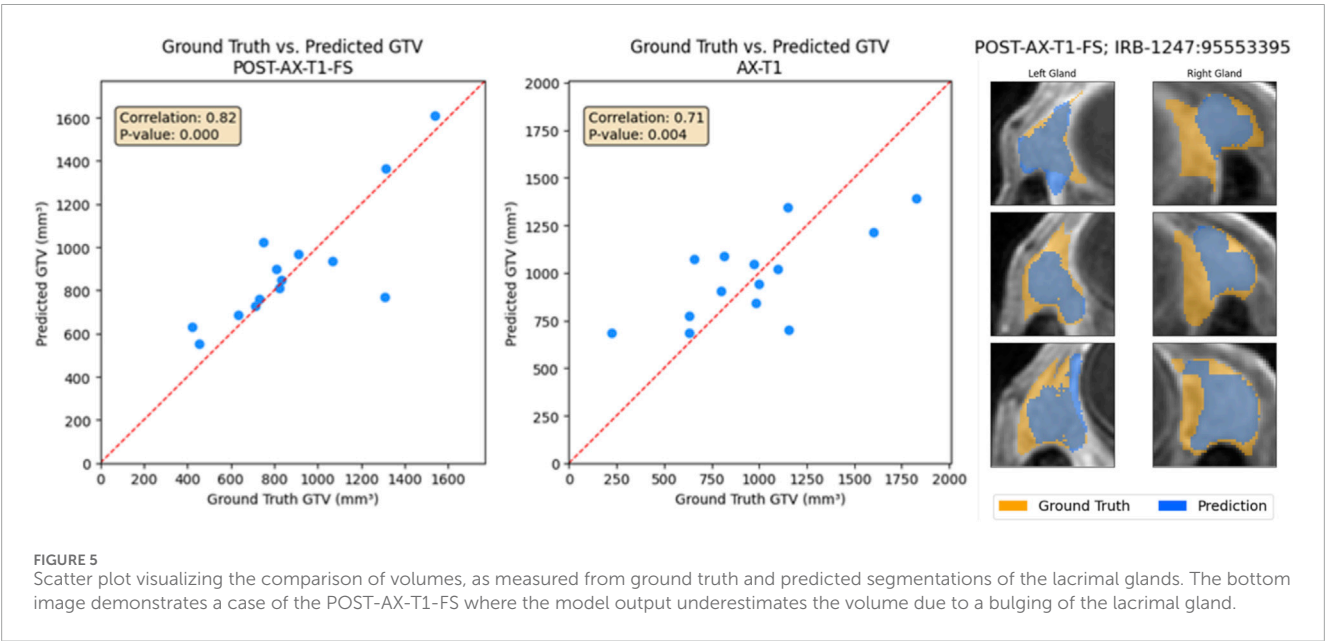
### 4.1 Comparison between POST-AX-T1-FS and AX-T1 MRI performance

Our study evaluated the performance of deep learning models in segmenting the lacrimal gland using both contrast-enhanced-fat-suppressed and non-contrast T1-weighted MRI sequences. A 2D UNet model was trained and evaluated on POST-AX-T1-FS sequences, demonstrating superior segmentation performance, achieving a mean Dice Similarity Coefficient (DSC) of  $[0.79 \pm 0.19; 0.77 \pm 0.23]$  and a mean Intersection over Union (IoU) of  $[0.68 \pm 0.19; 0.67 \pm 0.23]$ . In comparison, models trained on AX-T1 sequences achieved lower performance, with a mean Dice of  $[0.67 \pm 0.17; 0.61 \pm 0.23]$  and IoU of  $[0.52 \pm 0.18; 0.47 \pm 0.21]$ . This disparity can be attributed to the enhanced contrast between the lacrimal gland and surrounding orbital fat in POST-AX-T1-FS, which aids in more accurate boundary delineation. AX-T1 scans, lacking both contrast agent and fat suppression, frequently produced segmentations with boundary leakage or underrepresentation of gland volume ([Figure 6](#)).

Additionally, models trained on POST-AX-T1-FS demonstrated greater robustness in volume estimation, with a correlation coefficient of 0.82 compared to 0.71 for AX-T1. These findings suggest that fat-suppressed, contrast-enhanced imaging not only improves average segmentation accuracy but also yields more reliable predictions across patients. While AX-T1 sequences offer a gadolinium-free alternative, their reduced performance indicates that they may not yet serve as a full substitute for contrast-enhanced imaging in tasks requiring high segmentation precision.

TABLE 3 Evaluation metrics.

Organ Scan type	Dice score (mean ± sd) (25%–75% IQR)	IOU score (mean ± sd) (IQR)	HD <sub>95</sub> (in mm) (mean ± sd) (IQR)
Right lacrimal gland POST-AX-T1-FS	0.79 ± 0.19 0.78 to 0.89	0.68 ± 0.19 0.63 to 0.80	2.21 ± 1.80 1.05 to 2.84
Left lacrimal gland POST-AX-T1-FS	0.77 ± 0.23 0.79 to 0.90	0.67 ± 0.23 0.65 to 0.82	4.31 ± 13.14 0.84 to 3.11
Right lacrimal gland AX-T1	0.67 ± 0.17 0.58 to 0.79	0.52 ± 0.18 0.41 to 0.66	3.73 ± 2.43 1.79 to 4.68
Left lacrimal gland AX-T1	0.61 ± 0.23 0.47 to 0.76	0.47 ± 0.21 0.31 to 0.62	4.84 ± 7.50 2.17 to 4.73



These results highlight a trade-off between diagnostic clarity and material cost.

## 4.2 Clinical significance

Our findings highlight the potential value of accurate lacrimal gland segmentation in clinical imaging, particularly for tracking gland volume in diagnostic or treatment contexts. Automated segmentation provides a reliable means for volumetric analysis, which could help identify gland enlargement or atrophy, key features in conditions like Sjögren’s Syndrome, sarcoidosis, and orbital tumors. While DSC scores indicate room for improvement, particularly for non-contrast scans, the ability to quantitatively track volumetric changes over time could still facilitate treatment monitoring, such as assessing response to radiotherapy or detecting disease recurrence.

While U-Net has shown some promise in lacrimal gland segmentation, further exploration of alternative deep learning

models is crucial for improving accuracy, particularly for challenging cases or different imaging modalities. Models like DeepLabV3+ and Attention U-Net, with multi-scale context and attention mechanisms, have been successfully used for segmenting complex structures in similar contexts [24]. Additionally, architectures such as FCN and hybrid models combining convolutional and transformer-based approaches, as seen in studies like Yazdani et al.’s work on lesion and organ-at-risk segmentation using a SwinUNETR-based model [26], could improve segmentation, especially for non-contrast MRI scans. Generative adversarial networks (GANs), which have been applied in other segmentation tasks, may also help refine the delineation of poorly defined gland borders. Furthermore, nnU-Net has been effectively applied in other segmentation problems, such as segmenting the NLC on cone-beam CT images [24]. Additionally, a study using orbital CT from 701 patients trained a specialized encoder-decoder architecture and compared it to several conventional networks (Attention U-Net, DeepLabV3+, SegNet, HardNet-MSEG) for segmenting the eyeball, extraocular

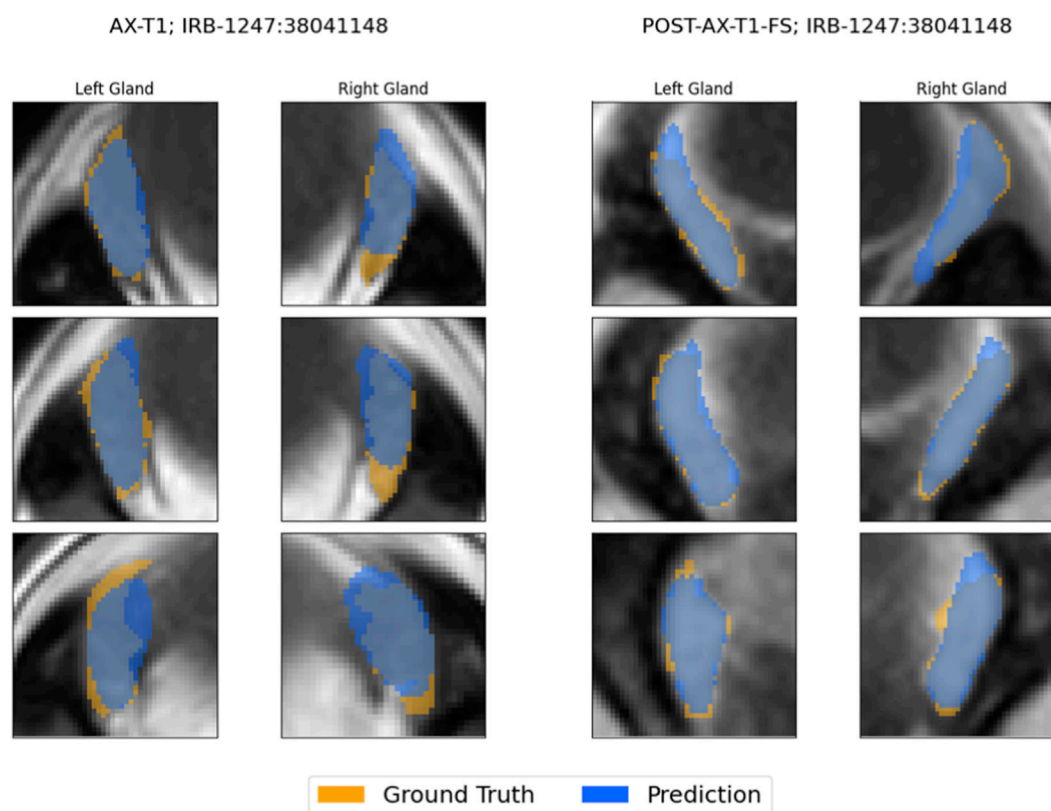


FIGURE 6

Visualization showing three representative slices containing the ground truth segmentations (top) and predicted segmentations (bottom) of the lacrimal glands for patient IRB-1247:38041148.

muscles, optic nerve, and lacrimal gland [28]. Integrating these models into the pipeline could enhance segmentation accuracy and adaptability, supporting broader clinical applications across diverse patient populations and imaging conditions.

From a workflow perspective, integrating automated lacrimal gland segmentation into radiology platforms can significantly streamline clinical evaluations, reducing the time needed for manual contouring and allowing radiologists to focus on higher-level decision-making. Importantly, the use of deep learning models also contributes to reduced interobserver variability, a common challenge in segmenting small or poorly defined structures such as the lacrimal gland, especially in non-contrast scans.

### 4.3 Limitations and future directions

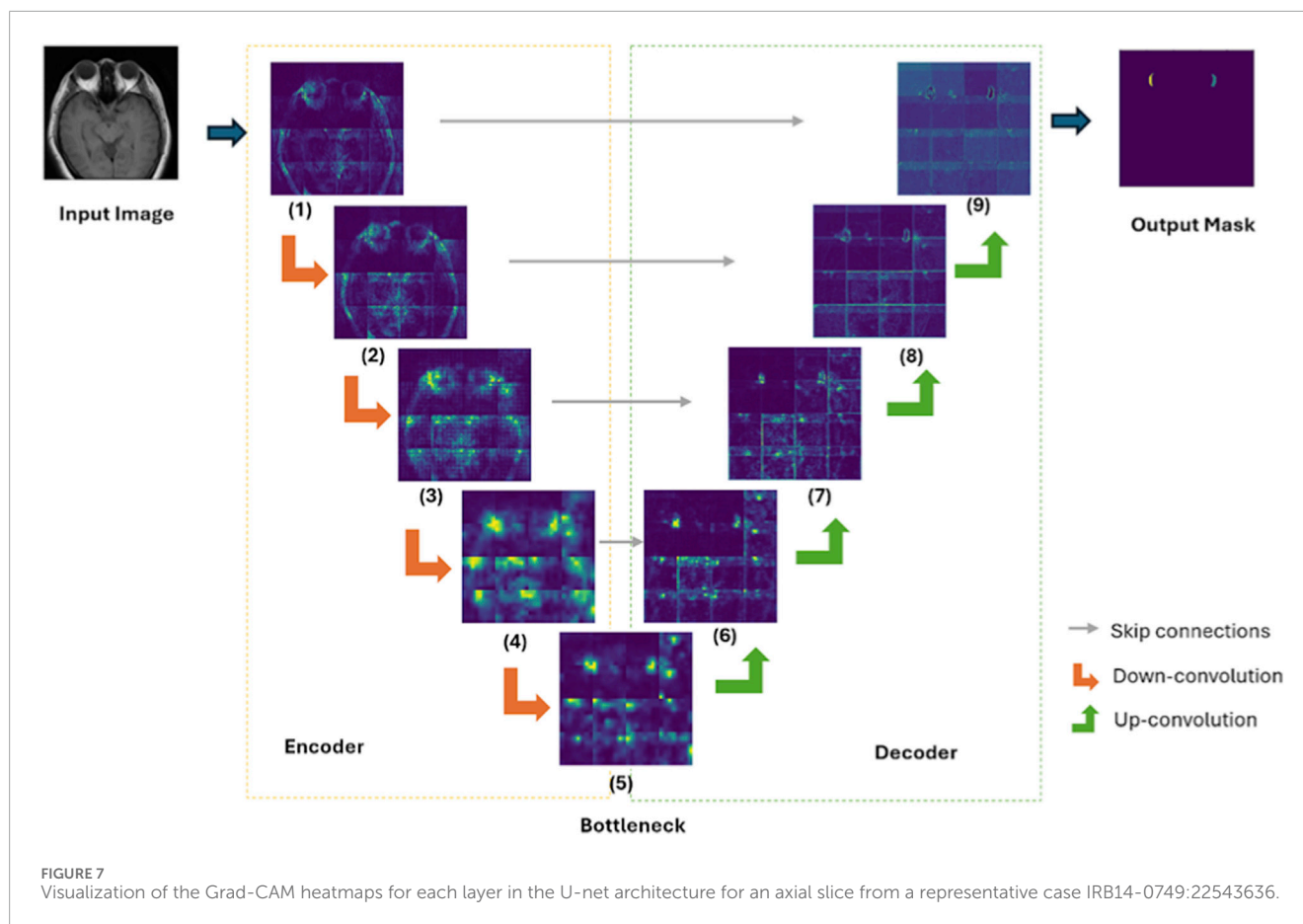
The small dataset size presents a significant constraint; the limited number of annotated MRIs may hinder the model's ability to generalize across diverse patient populations, scanner types, and imaging protocols. In addition, all MRIs were acquired from a single institution using a consistent scanner and protocol, which introduces the possibility of site-specific bias and limits external validity. To address these limitations, future work should prioritize the expansion of the dataset to include multi-institutional MRIs with varied demographics and lacrimal gland morphologies. This would

enable more robust model training and evaluation across diverse clinical settings.

Another question left unanswered is whether the improved segmentation performance observed in fat-suppressed contrast-enhanced scans is driven primarily by fat suppression or by the presence of the gadolinium-based contrast agent itself. Contrast-enhanced, fat-suppressed sequences simultaneously reduce background adipose signal and increase the relative conspicuity of the lacrimal gland, making it unclear which factor contributes more strongly to the model's success. To disentangle these effects, future studies should include a dedicated analysis of fat-suppressed, non-contrast MRI sequences. Such an investigation would clarify whether fat suppression alone provides sufficient gland-to-background contrast to support accurate segmentation, or whether contrast uptake is necessary to achieve the observed performance gains. This distinction will be essential for determining the optimal imaging protocol for both model training and eventual clinical deployment, particularly in cases where contrast administration may be contraindicated.

Additional model refinement is also warranted. Incorporating attention mechanisms, boundary-aware loss functions, or even transformer-based segmentation networks could improve the delineation of gland margins. Finally, future studies should include prospective clinical trials to assess the integration of these segmentation tools into the diagnostic workflow. Key





outcomes would include not only segmentation accuracy but also diagnostic impact, workflow efficiency, and physician satisfaction, providing a comprehensive evaluation of real-world clinical utility.

## 5 Conclusion

This study demonstrates that auto-segmentation algorithms perform better on contrast-enhanced (POST-AX-T1-FS) MR sequences compared to non-contrast (AX-T1) scans, achieving lower Hausdorff distances and higher Dice and IoU scores. However, despite this performance gap favoring contrast-enhanced imaging, neither model reached a level of accuracy that would be considered clinically optimal, especially for non-contrast scans. This result suggests that the current segmentation methods are not yet robust enough for reliable clinical use, particularly in cases where contrast agents are unavailable or contraindicated. This retrospective single-institution study carries a Level 4 grade of evidence [29].

Due to the modest performance observed in both settings, there is a clear need to explore alternative deep learning architectures that may better capture the lacrimal gland's small size and variable appearance. Models like Attention U-Net, DeepLabV3+, SegNet, and HarDNet-MSEG have been successfully applied to similar segmentation challenges and may offer improvements over the baseline models.

Accurate lacrimal gland segmentation has potential value for volumetric analysis in clinical imaging, especially for tracking gland changes associated with conditions such as Sjögren's syndrome, sarcoidosis, and orbital tumors. Although combining contrast and non-contrast sequences led to slight performance improvements, significant advancement is still necessary before automated segmentation can be fully integrated into routine clinical workflows. Future efforts should focus on expanding model comparisons, validating findings in multi-center datasets, and assessing how segmentation tools could help support clinical decision-making and reduce interobserver variability.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation. The computer code for training and validation can be found in the links below. Model training: [https://github.com/amoghs11/lacrimal\\_gland\\_segmentation/tree/main](https://github.com/amoghs11/lacrimal_gland_segmentation/tree/main) Explainable AI code: [https://github.com/rbramkumar/lg\\_segmentation\\_gradcam](https://github.com/rbramkumar/lg_segmentation_gradcam).

## Ethics statement

The studies involving humans were approved by University of Chicago institutional review board. The studies were conducted in

accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participant's legal guardians/next of kin in accordance with the national legislation and institutional requirements.

## Author contributions

AS: Data curation, Investigation, Software, Methodology, Writing – original draft, Writing – review and editing, Formal Analysis. RB: Software, Investigation, Methodology, Writing – original draft, Writing – review and editing, Formal Analysis. MI: Data curation, Investigation, Writing – review and editing. QC: Data curation, Writing – review and editing. SK: Data curation, Writing – review and editing. SI: Data curation, Writing – review and editing, Validation. MH: Data curation, Writing – review and editing. DG: Conceptualization, Writing – review and editing, Methodology, Investigation.

## Funding

The author(s) declared that financial support was not received for this work and/or its publication.

## Acknowledgements

This work was completed in part with resources provided by the University of Chicago's Research Computing Center. The study

was approved by the Institutional Review Board of the University of Chicago IRB18-1247.

## Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Machiele R, Lopez MJ, Czyz CN. Anatomy, head and neck: eye lacrimal gland. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing (2025). Available online at: <https://www.ncbi.nlm.nih.gov/books/NBK532914/> (Accessed June 9, 2025).
- Seifert P, Spitznas M, Koch F, Cusumano A. The architecture of human accessory lacrimal glands. *Ger J Ophthalmol* (1993) 2(6):444–54.
- Hacking C, Hapugoda S, Knipe H. Lacrimal nerve. *Radiopaedia.org*. doi:10.5334/rID-52112
- Carsons SE, Patel BC. *Sjogren syndrome*. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing (2023). Available online at: <https://www.ncbi.nlm.nih.gov/books/NBK431049/> (Accessed June 9, 2025).
- Brito-Zerón P, Baldini C, Bootsma H, Bowman SJ, Jonsson R, Mariette X, et al. Sjögren syndrome. *Nat Rev Dis Primers* (2016) 2:16047. doi:10.1038/nrdp.2016.47
- Negrini S, Emmi G, Greco M, Borro M, Sardaneli F, Murdaca G, et al. Sjögren's syndrome: a systemic autoimmune disease. *Clin Exp Med* (2022) 22(1):9–25. doi:10.1007/s10238-021-00728-6
- Agarwal A, Chandak S. Sarcoidosis presenting as lacrimal gland enlargement: eyes speak the truth. *Taiwan J Ophthalmol* (2019) 10(3):227–30. doi:10.4103/tjo.tjo\_125\_18
- Blumfield E, Swenson DW, Iyer RS, Stanescu AL. Gadolinium-based contrast agents - review of recent literature on magnetic resonance imaging signal intensity changes and tissue deposits, with emphasis on pediatric patients. *Pediatr Radiol* (2019) 49(4):448–57. doi:10.1007/s00247-018-4304-8
- Idée JM, Port M, Medina C, Lancelot E, Fayoux E, Ballet S, et al. Possible involvement of gadolinium chelates in the pathophysiology of nephrogenic systemic fibrosis: a critical review. *Toxicology* (2008) 248(2-3):77–88. doi:10.1016/j.tox.2008.03.012
- Guo BJ, Yang ZL, Zhang LJ. Gadolinium deposition in brain: current scientific evidence and future perspectives. *Front Mol Neurosci* (2018) 11:335. doi:10.3389/fnmol.2018.00335
- McDonald RJ, McDonald JS, Kallmes DF, Jentoft ME, Murray DL, Thielen KR, et al. Intracranial gadolinium deposition after contrast-enhanced MR imaging. *Radiology* (2015) 275(3):772–82. doi:10.1148/radiol.15150025
- Delfaut EM, Beltran J, Johnson G, Rousseau J, Marchandise X, Cotten A. Fat suppression in MR imaging: techniques and pitfalls [published correction appears in *Radiographics* 1999 Jul-Aug;19(4):1092]. *Radiographics* (1999) 19(2):373–82. doi:10.1148/radiographics.19.2.g99mr03373
- Rana K, Juniat V, Patel S, Selva D. Normative lacrimal gland dimensions by magnetic resonance imaging in an Australian cohort. *Orbit* (2023) 42(2):157–60. doi:10.1080/01676830.2022.2055085
- Simon J, Szumowski J, Totterman S, Kido D, Ekholm S, Wicks A, et al. Fat-suppression MR imaging of the orbit. *AJNR Am J Neuroradiol* (1988) 9(5):961–8.
- Ilesanmi AE, Ilesanmi TO, Ajayi BO. Reviewing 3D convolutional neural network approaches for medical image segmentation. *Heliyon* (2024) 10:e27398. doi:10.1016/j.heliyon.2024.e27398
- Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. *arXiv [Preprint]*. Available online at: <https://arxiv.org/abs/1505.04597>.
- Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Olaf R. 3D U-Net: learning dense volumetric segmentation from sparse annotation. 424–32. (2016). doi:10.48550/arxiv.1606.06650
- Oktay O, Schlemper J, Folgoc LL. Attention U-Net: learning where to look for the pancreas. *arXiv.org*. doi:10.48550/arXiv.1804.03999
- Shanker RRB, Zhang MH, Ginat DT. Semantic segmentation of extraocular muscles on computed tomography images using convolutional neural networks. *Diagnostics* (2022) 12(7):1553. doi:10.3390/diagnostics12071553
- Podobnik G, Ibragimov B, Tappeiner E, Lee C, Kim JS, Mesbah Z, et al. HaN-Seg: the head and neck organ-at-risk CT and MR segmentation challenge. *Radiother Oncol* (2024) 198:110410. doi:10.1016/j.radonc.2024.110410

21. Quetin S, Heschl A, Murillo M, Murali R, Pater P, Shenouda G, et al. Automatic segmentation of organs at risk in head and neck cancer patients from CT and MR scans. *arXiv [Preprint]* (2024). doi:10.48550/arXiv.2405.10833
22. Wang G. Fully automated segmentation and volumetric measurement of ocular adnexal lymphoma (OAL) using nnU-net (2024). Available online at: [www.ncbi.nlm.nih.gov/articles/PMC11424727/](https://www.ncbi.nlm.nih.gov/articles/PMC11424727/) (Accessed June 9, 2025).
23. Jañez-García L, Saenz-Frances F, Ramirez-Sebastian JM, Toledano-Fernandez N, Urbasos-Pascual M, Jañez-Escalada L. Three-dimensional reconstruction of the bony nasolacrimal canal by automated segmentation of computed tomography images. *PLoS ONE* (2016) 11(5):e0155436. doi:10.1371/journal.pone.0155436
24. Haylaz E, Gumusoy I, Duman SB, Kalabalik F, Eren MC, Demirsoy MS, et al. Automatic segmentation of the nasolacrimal canal: application of the nnU-Net v2 model in CBCT imaging. *J Clin Med* (2025) 14(3):778. doi:10.3390/jcm14030778
25. Klyuzhin IS, Chaussé G, Bloise I, Lavista Ferres JM, Uribe C, Rahmim A, et al. *PSMA-Hornet: fully-automated, multi-target segmentation of healthy organs in PSMA PET/CT images* (2022) medRxiv. doi:10.1101/2022.02.02.22270344
26. Yazdani E, Karamzadeh-Ziarati N, Cheshmi SS, Sadeghi M, Geramifar P, Vosoughi H, et al. Automated segmentation of lesions and organs at risk on [<sup>68</sup>Ga]Ga-PSMA-11 PET/CT images using self-supervised learning with swin UNETR. *Cancer Imaging* (2024) 24(1):30. doi:10.1186/s40644-024-00675-x
27. Lee SH, Lee S, Lee J, Lee JK, Moon NJ. Effective encoder-decoder neural network for segmentation of orbital tissue in computed tomography images of graves' orbitopathy patients. *PLoS ONE* (2023) 18(5):e0285488. doi:10.1371/journal.pone.0285488
28. Selvaraju RR. Grad-CAM: visual explanations from deep networks via gradient-based localization. In: *ICCV* (2017).
29. Pham N, Hill V, Rauschecker A, Lui Y, Niogi S, Fillipi CG, et al. Critical appraisal of artificial intelligence-enabled imaging tools using the levels of evidence system. *AJNR Am J Neuroradiol* (2023) 44(5):E21–E28. doi:10.3174/ajnr.A7850