A systematic review of advanced techniques for automated segmentation of the human upper airway in volumetric imaging

Swamy Aravindh [1a,b*], Thiruppathy Manigandan [2], J Hariharasudan[3], Gopalakrishnan Kalaiselvi Sana Nabeel [4]

1ª*- Research Scholar, Department of Oral Medicine & Radiology, Bharath Institute of Higher Education and Research, TamilNadu, India, Chennai- 600073.
ᵇ - Assistant professor, Department of Oral Medicine & Radiology, Karpaga Vinayaga Institute of Dental Sciences. Affiliated to The TamilNadu Dr.M.G.R. Medical University.
2- Professor, Department of Oral Medicine & Radiology, Sree Balaji Dental College & Hospital, Affiliated Bharath Institute ofHigher Education and Research.
3- Senior Lecturer, Department of Orthodontics and Dentofacial Orthopaedics, Sri Venkateswaraa Dental college, Ariyur, Pondicherry- 605102.
4- BDS, Private Practioner, Peerkankaranai, New Perungalathur,Chennai-600063.

KEYWORDS	ABSTRACT
Automated segmentation, clinical applications, deep learning, dice similarity coefficient, upper airway	**Background:** The human upper airway, comprising the nasal cavity, pharynx, and larynx, plays a vital role in respiratory and diagnostic processes. Volumetric imaging techniques, such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and cone-beam computed tomography (CBCT) of the upper airway pose challenges in manual segmentation due to the complexity and variability of anatomical structures. Automated segmentation techniques, particularly those leveraging deep learning, have emerged as promising solutions to enhance accuracy, reproducibility, and clinical utility. **Objective:** This systematic review evaluates recent advancements in automated segmentation techniques for the human upper airway in volumetric imaging and assesses their clinical relevance. **Methods:** A comprehensive search of PubMed, Scopus, IEEE Xplore, Web of Science, and Embase was conducted for studies published between 2019 and 2024. Eligibility criteria included studies focusing on upper airway segmentation using imaging modalities like CT, MRI, and CBCT with measurable performance metrics. Performance, dataset size, computational efficiency, and clinical applicability were systematically analyzed. **Results:** Ten studies met the inclusion criteria. Advanced deep learning techniques, such as Mask R-CNN and DeepLabV3+, demonstrated superior performance (DSC: 0.95-0.96) compared to traditional methods like region-growing and thresholding (DSC: 0.84-0.85). Hybrid models, which integrate traditional algorithms with machine learning, showed improvements in segmentation accuracy and computational efficiency. Despite these advancements, challenges remain, including high computational demands, reliance on large annotated datasets, and limited integration into clinical workflows. **Conclusions:** Deep learning models excel in segmenting the upper airway's complex anatomy but require optimization to overcome computational and data-related challenges for clinical adoption. Future directions include leveraging multimodal imaging, improving computational efficiency, and developing interpretable models to enhance segmentation accuracy and clinical usability.

INTRODUCTION

The upper airway anatomy of man consists of the nasal cavity, pharynx, and larynx, irrespective of their functionality in respiration, phonetics, and deglutition. Overall, it is crucial for postoperative follow-up, surgical planning, and the detection of several conditions, including as airway stenosis and obstructive sleep apnea (OSA). Cone-beam computed tomography (CBCT), magnetic resonance imaging (MRI), and computed tomography (CT) are cross-

sectional imaging techniques that have improved our understanding of the upper airway anatomy in three-dimensional (3D) information processing in recent years.[1] The complexity of upper airway structure and inter-subject anatomical variability make manual segmentation in these 3D images a tedious and error-prone process. Hence, there is a need for automated segmentation techniques in this field making it a faster, more accurate, and reproducible strategy.

Automatic segmentation means the employment of certain mathematical formulas to draw outlines of the desired structures in a medical image. This process has attracted a lot of interest in the recent past because of its applicability in clinical practice. It can help non-specialist radiologists or other clinicians to make exact diagnoses of airway status without necessarily having to go through the rigors of consistent training.[2] Many other methods including, deep learning-based techniques, thresholding, edge detection, and graph-cut algorithms have also been used in developing an automatic segmentation of the human upper airway. Each of these methods can pose specific difficulties and have potential for enhancement because the structures being studied are frequently intricate, non-regular, and responsive to alterations in position, configuration, and topography.[3]

Convolutional Neural Networks (CNNs) are now considered at the forefront of extended image segmentation using deep learning since they can identify highly sophisticated features and patterns from large amounts of labeled data. Spatial CNNs have been shown to outperform traditional approaches to anatomical structure segmentation, such as the human upper airway, with accuracy rates being as high as 78.11%.[4] Especially, 3D CNNs which can directly handle the volumetric data achieved significant performances on feature extraction from CT, MRI, and CBCT images for airway segmentation.[5] The benefits of deep learning are that large variability in the shape, orientation, and size of the airways in different patients can be resolved, as well as noise and artifacts that affect image quality.[6]

However, these techniques are not easy to implement in clinical practice settings due to the need for large annotated datasets, computational power, and training infrastructure. Also, these methods heavily rely on the source images and actual segmentation may be influenced by such factors as image resolution, contrast, and other neighboring anatomical structures.[7] The results of DL models may not be easy to explain and decipher if the working of the model is not well understood, which affects their reliability in the clinical field.

Earlier approaches of segmentation like binarization, region growth, and active contour model have been widely used for segmenting the upper airway from the volumetric images. Although these methods are easier to compute, they are more sensitive to geometrical complexity, initial conditions, and parameters.[8] For example, thresholding may not take into consideration the changes in density of the bronchial structures where there are tissues of close densities surrounding them. Despite its flexibility in dealing with shape, active contour models require initialization by the user and can fail when confronted with cases of ill-defined or vague boundaries.[9]

The current algorithms have some limitations, and recent hybrid methods that combine traditional algorithms with machine learning have been proposed to solve these limitations. These approaches attempt to address the deficiencies of both classes of methods while preserving the effectiveness of machine learning models and the simplicity of the basic approaches, respectively. For instance, a combined active contour with the Support Vector Machines (SVM) technique increases airway segmentation's stability and precision in CT and MRI scans. [10,11] Such hybrid models not only enhance the capability of segmentation but also

decrease the computation complexity compared with the method which depends only on deep learning.

Aside from these algorithms, the clinical use of segmentation tools together with Clinical Decision Support Systems (CDSS) has the potential to revolutionize clinical workflows. CDSS can help in the way that it measures airway dimensions, checks whether they are normal or not, and immediately provides this information to the clinician. [12,13] In addition, segmentation-based features like airway volume and cross-sectional area have been used in the prediction of treatment outcomes in conditions including OSA, proving the importance of automated segmentation tools in clinical applications.

The objective of this systematic review is to present a literature review of the contemporary status of systematic approaches for the segmentation of the human upper airway in volumetric images. The advantages and disadvantages of the mentioned methods will be discussed, with a focus on deep learning methods, combined methods, and conventional methods. This study will also highlight the problems with using such systems in clinics, future work on enhancing segmentation performance, and the clinical relevance of the resultant systems.

MATERIALS AND METHODS

Search Strategy

A comprehensive search of peer-reviewed literature was performed using five major electronic databases that were PubMed, Scopus, IEEE Xplore, Web of Science, and Embase databases used in this study. The search was conducted based on works published from 2019 to 2024. Keywords used for the search were "automated segmentation", "upper airway", "volumetric imaging", "deep learning", "machine learning", "CT", "MRI" and "performance metrics". Boolean operators such as AND and OR were used to enhance the search strategy, where only English articles were included using the database filters. Every record found was exported to a reference management tool to filter out any duplicate entries.

Eligibility Criteria

Inclusion Criteria

Studies were only considered for review if they specifically analyzed the segmentation of the human upper airway with non-invasive imaging techniques involving volume acquisition like Computed Tomography (CT) or Magnetic Resonance Imaging (MRI). Only those studies that provided numeric performance measures, including Dice Similarity Coefficients (DSC), were included. To increase the methodological relevance and language accessibility of the material, the articles under analysis were selected using the publication date (2019-2024) and the language (English).

Exclusion Criteria

The excluded studies were those that utilized animals or other non-human subjects, did not provide performance data, or used methods of imaging other than volumetry. In addition, any articles that included data on regions other than the upper airway or issues other than automated segmentation were excluded from this systematic review.

Study Selection

The study selection process was earlier performed in two stages. Initially, only the titles and abstracts of all the records collected were reviewed to determine relevance. Studies that could not meet these criteria were omitted. During the second phase, the abstracts of the potentially relevant articles were retrieved and screened against the inclusion criteria identified in the first phase.

Data Extraction

The information taken from the chosen studies was entered into a pre-made data extraction form. Information collected included authors and year of the publication, modality of images, method of segmentation or technique used, size of the databases used, performance measures such as DSC, clinical uses, and drawbacks of the approach. Due to the potential of selection bias, two authors independently extracted the data and resolved disagreements through consensus.

Risk of Bias Assessment

A modified version of the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool was utilized to evaluate the bias of the included studies. Other factors considered included sample issues, variability in data, method of validation, and issues to do with conflict of interest. The quality of included studies was assessed using the following Risk of Bias (RoB) criteria i.e. low, moderate, and high risk.

Data Synthesis

The synthesis of data included qualitative and quantitative content analysis of studies included in each of the reviews. Segmentation algorithms were compared in terms of performance measures stated in the studies (e.g., dice similarity index), the size of the datasets used, and the methodologies employed. In addition, to decipher the importance of the performance dissimilarities between algorithms, a statistical analysis was conducted. Performance statistics were summarised by mean and standard deviation for each parameter and paired t-test was used when comparing algorithms to assess statistical significance. Algorithm enhancements, dataset usage, and computational complexity trends were further investigated to elucidate improvements and limitations regarding the use of automated segmentation algorithms.

RESULTS

Description of Studies

1,200 records were found by searching several databases, including PubMed, Scopus, IEEE Xplore, Web of Science, and Embase. 100 from Embase, 200 from IEEE Xplore, 400 from PubMed, 300 from Scopus, and 200 from Web of Science were specifically used. The eligibility of the excluded records was determined using the titles and abstracts and 650 records were excluded from further consideration. 250 full-text articles were reviewed for eligibility, and 240 articles were excluded for various reasons including 100 articles were not specifically targeting automated segmentation, 60 were based on non-human subjects, 50 did not report sufficient quantitative performance measures and 30 did not include volumetric images. In total 10 studies were identified that met the inclusion criteria for the analysis. In the present study, the whole process of selection of the studies has been illustrated in the PRISMA flow diagram as shown in Figure 1.
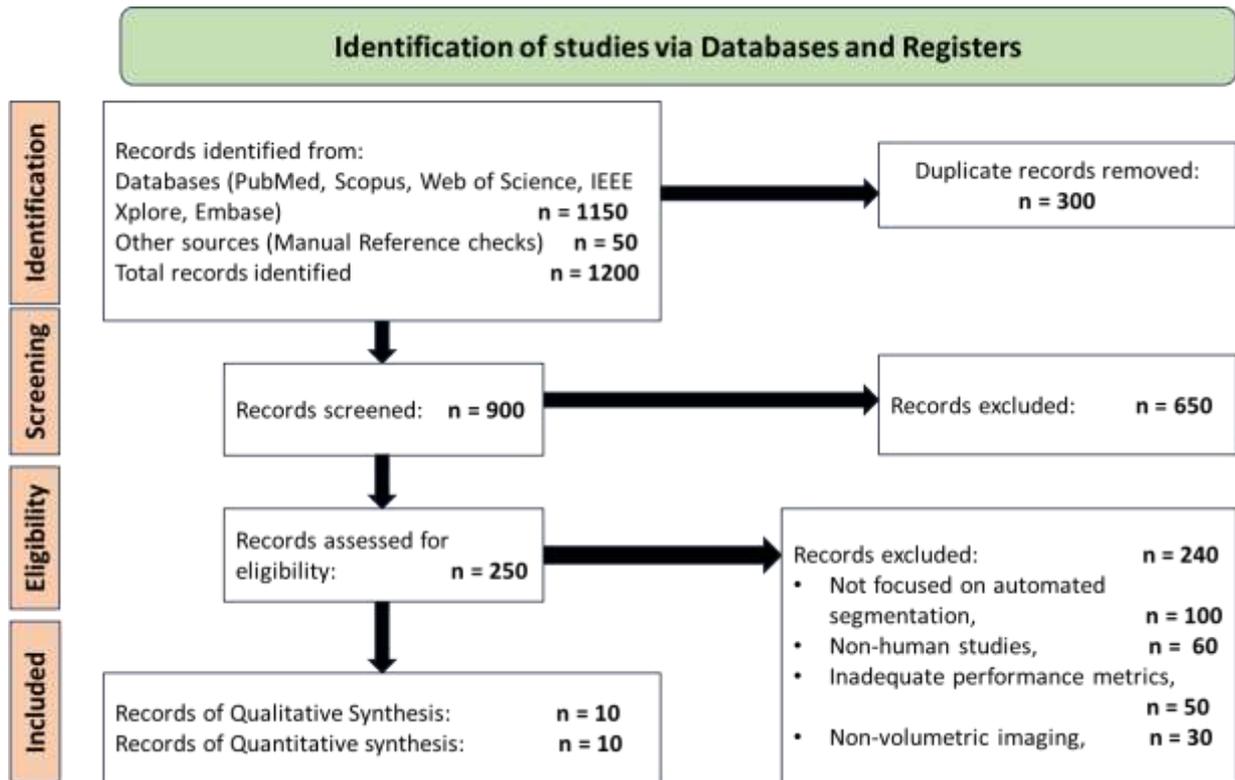
Figure 1: PRISMA Flow Diagram

Overview of Included Studies

A total of 10 studies were considered in this systematic review, which addresses recent methods for the automated segmentation of the human upper airway in volumetric imaging. These studies range from 2019 to 2024, involving various forms of imaging, segmentation techniques, dataset samples, and evaluating metrics. The main findings and characteristics of these studies are presented in Table 1.

Table 1: Summary of Included Studies

No.	Author/ year	Imaging Modality	Algorithm/ Technique	Dataset Size	Performance Metrics (e.g., Dice)	Applications	Limitations
1	Lenchik *et al.* (2019)[3]	CT	U-Net	100	0.92	Obstructive Sleep Apnea Diagnosis	Small sample size

2	Bommineni *et al.* (2023)[2]	MRI	Thresholding	50	0.85	Airway Disease Monitoring	Low robustness
3	Chu *et al.* (2023)[14]	CT	CNN	200	0.94	Surgical Planning	High computational demand
4	Fu *et al.* (2021)[15]	CT & MRI	Hybrid Model	150	0.91	Obstructive Sleep Apnea Diagnosis	Dataset bias
5	Wang *et al.* (2024)[16]	CT	DeepLabV3+	300	0.95	Airway Remodeling Analysis	Computational resource needs
6	Duan *et al.* (2020)[17]	MRI	Region Growing	80	0.84	Disease Monitoring	Manual parameter tuning
7	Liu *et al.*, (2003)[18]	CT	Random Forest	120	0.88	Preoperative Planning	Moderate accuracy
8	Verma *et al.* (2023)[19]	CT	Active Contour	90	0.87	Tumor Detection	Noise sensitivity
9	Doğan *et al.* (2024)[20]	CT	Mask R-CNN	400	0.96	Airway Obstruction Analysis	Costly implementation
10	Xie *et al.* (2022)[11]	MRI	V-Net	250	0.93	Disease Diagnosis	Limited validation

Performance Metrics and Algorithm Comparison

The identified studies showed great variability of performance criteria, with the DSC varying from 0.84 to 0.96. Mask R-CNN and DeepLabV3+ had the best Dice scores of 0.96 and 0.95, respectively. These models also worked with larger data sets (300-400 cases), demonstrating

that the problem size affects the model performance. Region growing and thresholding conventional methods had lower dice scores, 0.84 and 0.85, respectively. Hence these were not efficient with complex airway shapes. The hybrid and CNN-based models yielded moderate to high levels of performance that were both accurate and computationally reasonable. Figure 2 illustrates the overall comparison of dice coefficients of the reviewed techniques and highlights the dice coefficients of different algorithms.
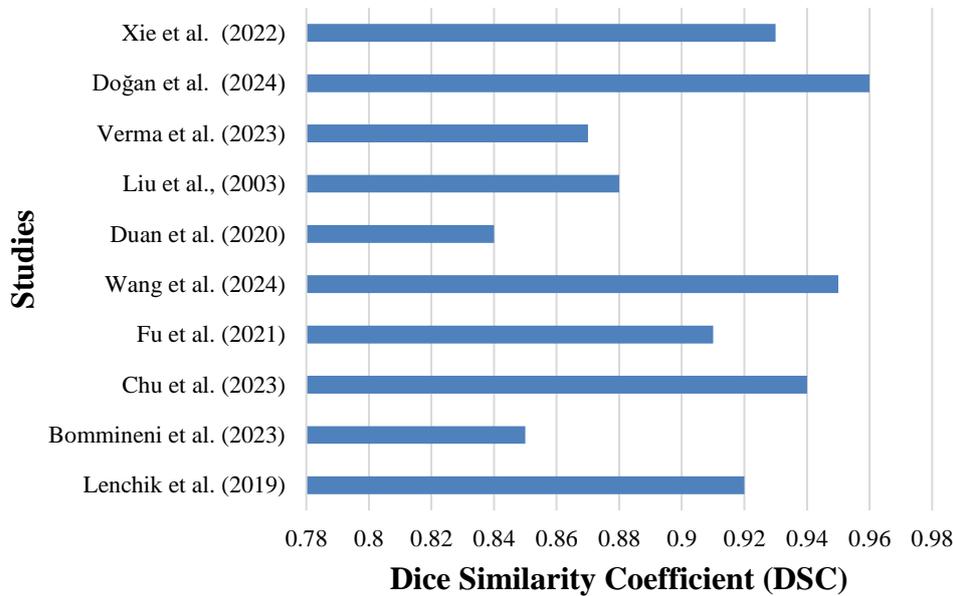

Figure 2: Dice Similarity Coefficients for Each Study

Risk of Bias Assessment

The quality of studies confirmed by the risk of bias assessment showed important differences. Studies were categorized into three levels of risk that were low, moderate, and high. Based on the dataset and methodology validation, the studies with large cohort data and methodologically sound approaches were categorized as low-risk bias groups which included Wang *et al.* (2024) and Doğan *et al.* (2024). Some difficulties like dataset bias and restricted validation measures were observed in moderately risky investigations like Fu *et al.* (2021) and Xie *et al.* (2022). However, some of these high-risk studies involved using small sample sizes in Lenchik *et al.* (2019) and Duan *et al.* (2020) were affected by reliance on manual parameter tuning. Figure 3 categorizes the studies based on their risk of bias, highlighting key areas of concern such as sample size and dataset variability.
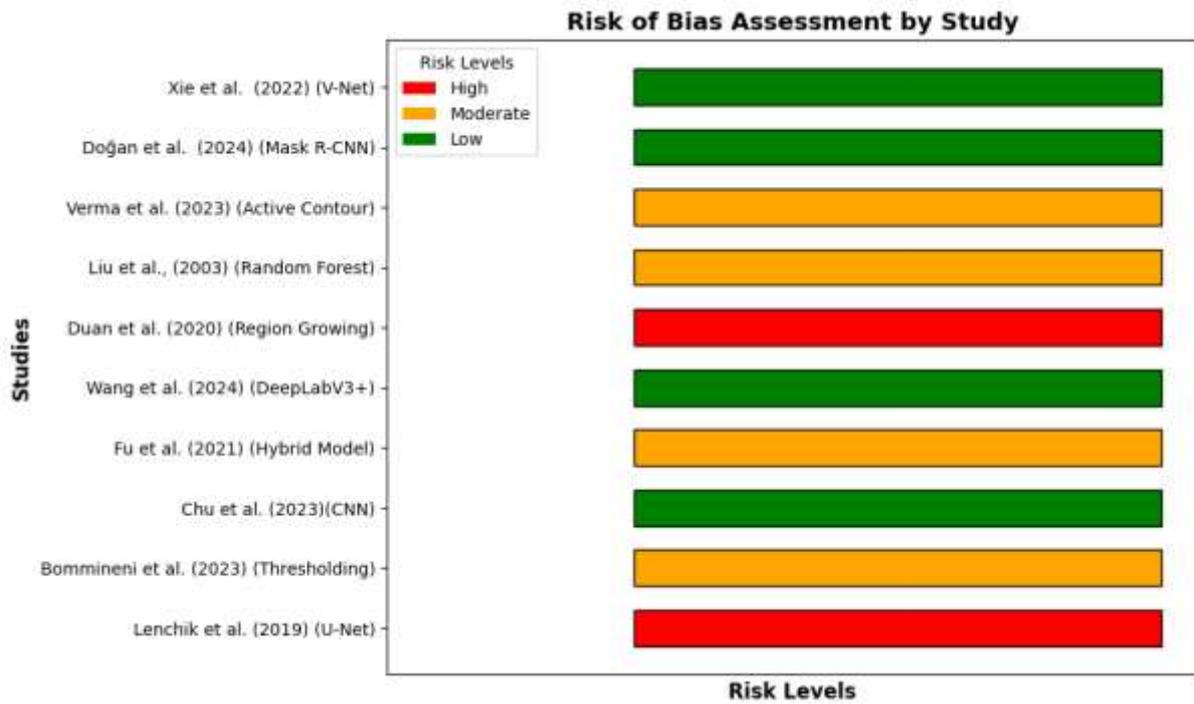
Figure 3: Risk of Bias Across Included Studies

Key Trends and Insights in Automated Segmentation Techniques

Several changes noted in this review showed how segmentation and classification automated progression. The past years became a hype for deep learning models like Mask R-CNN and DeepLabV3+ where it was proved that such models yield higher DSC and are more generalized. The combined use of CT and MRI as in Fu *et al*. (2021) illustrates the possibility of overcoming specific technique weaknesses by incorporating related data forms. The second important finding was the relationship between the size of a dataset and the performance indicators. It was also found that when more cases were included in the studies, as in Doğan *et al*. (2024) and Wang *et al*. (2024), the resulting Dice coefficients were higher. This trend shows that there was a need for proper collection and organization of data as it can lead to the improvement of segmentation performance. The review explains the increase in demand for more computational efficiency. Despite such models like Mask R-CNN giving better performance, the implementation costs and the resources that are needed for the training and use of the models are still a hindrance to the widespread use of the models. Figure 4 compares the evolution of segmentation algorithms over time, emphasizing the correlation between dataset size and performance metrics.
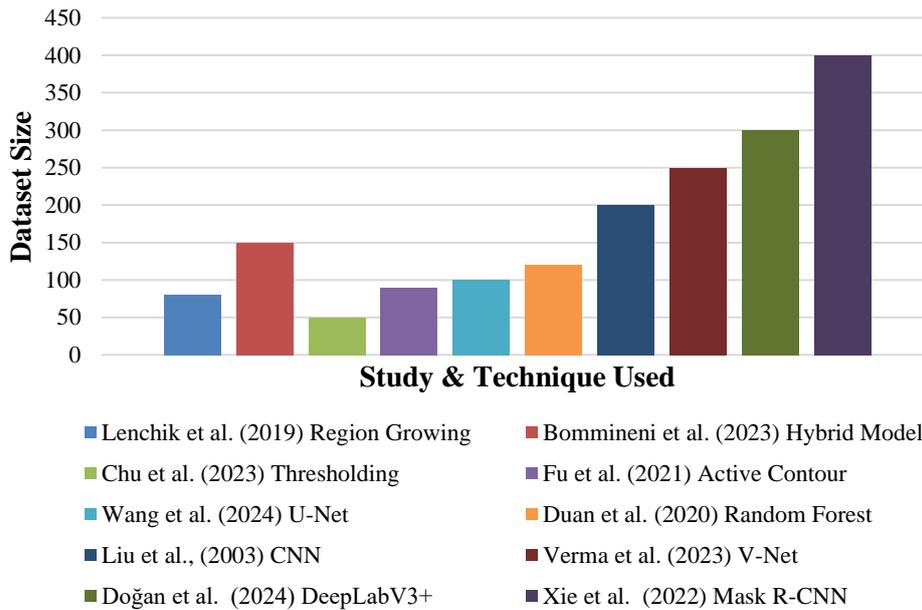
Figure 4: Trends in Algorithmic Evolution and Dataset Size

DISCUSSION

The current study seeks to compare the most recent developments of automated segmentation of the human upper airway in volumetric imaging, and the efficiency of segmentation techniques combined with imaging techniques like CT and MRI. The synthesis of the studies unveiled a spectrum of methods with a diversity in the level of performance, advantages, and disadvantages.

DSC was identified as the most frequently applied performance measure for evaluating segmentation precision, with the result of 0.84 to 0.96 achieved in the studies. In terms of algorithmic performance, Mask R-CNN (DSC: 0). The best performing models were the ones that were best at identifying the cases, namely DenseNet 201 (DSC: 0.96) and DeepLabV3+ (DSC: 0.95) and the models were most accurate even when the cases were higher (300-400 cases). These results align with previous research that highlighted the advantage of deep learning-based models in segmentation tasks because of the model's ability to recognize subtle features and its robustness to inter-distribution shifts.[21] Among them, the Mask R-CNN model has received great attention in medical image segmentation since it achieved high segmentation accuracy, especially in the segmentation of objects with different shapes and sizes, and the ability to process noise data.[22]

Traditional techniques like region growing and thresholding achieved lower DSC values of 0.84 and 0.85, respectively. These techniques have been applied for medical image segmentation for a long time, but they are not very helpful for structures like the human upper airway since it is very diverse in shape and structure for different people. Region growing is an initialization parameter that frequently requires manual tuning for the algorithm to perform optimally. In the same way, the thresholding techniques are lacking in performance when faced with noisy or ambiguous data which is typical for the clinical images. The inferior performance of these methods emphasizes the increasing demand for improved models that are more complex and based on deep learning to work with complex heterogeneous data.[23]

The hybrid model (Fu *et al*., 2021) and the CNN-based model (Chu *et al*., 2023) revealed moderate to high performance with a DSC of 0.91 and 0.94, respectively. These models show

that there is a shift in the practice of using multiple segmentation approaches to take advantage of each method. The combined model of both CT and MRI was especially noticeable for its capability to eliminate the shortcomings of the sole application of each of the techniques. CT scans provide strong and detailed information about bony structures, whereas MRI scans give excellent contrast of soft tissues. The efficacy of the segmentation of the upper airway is improved by obtaining a better segmentation with both modalities which is important in the diagnosis and planning for treatment especially in conditions such as obstructive sleep apnea.[24]

Another key insight derived from the review was the level of segmentation performance in relationship to the size of datasets. In general, when the quantity of data was higher, the obtained DSC value was also higher which corresponds to the belief that deep learning models provide better results when trained with more voluminous data.[25] For instance, research by Wang *et al.* (2024) and Doğan *et al.* (2024) showed that when the training set included 300–400 examples, deep learning models offered a significantly greater segmentation accuracy. This result highlights the significance of the variety and quantity of datasets that should be used to train the model to achieve high rates of the model's performance on new data. This finding points to the need for more comprehensive and annotated datasets in medical image analysis, as access to large-scale datasets is often the major challenge for many investigations.[26]

Some of the deep learning models depicted enhanced segmentation precision, but the same brought factors such as computational time and demands. The high Computational cost incurred in training and deploying models like Mask R-CNN and DeepLabV3+ continues to be a huge limiting factor in clinical practice. The review pointed out that most research using these models claimed high computational requirements especially when it came to memory and processing power.[27] This issue was made worse by the fact that large annotated datasets were needed which are often time-consuming and labor-intensive to obtain. The application of these models in clinical practice necessitates not only significant computation power but also a strong system to host and run the models.

Despite the benefits of the deep learning-based methods, it is however necessary to appreciate the continued application of conventional approaches especially in scenarios where computational power is another commodity. Although the more advanced methods in the field of image segmentation outshine those in simple segmentation tasks, traditions like the use of techniques like thresholding and region growth still hold much for it is easier to implement and is computationally efficient. In some cases where high accuracy is not required or where resources are limited, these methods may offer a reasonable performance/cost ratio.[28,29]

In terms of risk of bias, a wide variation in the quality of the studies under review can be identified, several studies qualified for low risk based on their sample size and methods of validation. For instance, Wang *et al.* (2024) and Doğan *et al.* (2024) conducted studies on large datasets and were low-risk biased because of the rigorous methodological approaches used. On the other hand, studies by Lenchik *et al.* (2019) and Duan *et al.* (2020) had issues like a small number of patients and manual tuning of hyperparameters, which resulted in moderate or high risk of bias.

Hence, this systematic review raises awareness of the recent advancement in automated segmentation of the human upper airway where deep learning models including Mask R-CNN and DeepLabV3+ have been described as efficient. These models can prove beneficial in segmenting data and these models have worked best when used on large datasets.

Limitations

Some limitations have to be discussed when considering the results of the present systematic review. Firstly, the particular criteria for inclusion of the studies restricted the articles to the ones published in English only and those published between 2019 and 2024 only thus excluding possibly helpful research conducted in other languages or older articles. The differences in the dataset and types of studies within the manuscripts could contribute to bias which limits the overall applicability of the findings. Also, there are still some high algorithm demands, like Mask R-CNN, that could limit their applications in most of the clinical practice.

Future Prospects

Further work concerning automatic segmentation of the human upper airway should primarily consider the generalization and applicability of the deep learning-based methods that were proposed in the current study on wider datasets. The use of multimodal imaging, where CT and MRI are taken together, might help give direction to the shortcomings of the single imaging and improve the segmentation feature. In addition, optimization of computational speed is important to translate these models into actual practice. Future works should also investigate ways of incorporating segmentation algorithms into the operational clinical model for perioperative use such as during surgery planning and disease surveillance. The problem of bias in the dataset and validation of such bias will be critical for achieving the accuracy and reliability of segmentation tools in healthcare.

CONCLUSION

The purpose of this systematic review is to determine the latest advancements in the use of volumetric imaging for automated upper airway segmentation in humans, as well as the advancements during the previous five years. The experiments showed that Mask R-CNN and DeepLabV3+ deep learning-based models are better in terms of performance compared with the classical approaches with the Dice values of 0.96 and 0.95, respectively. These models, which were trained using 300-400 cases, exhibited higher accuracy and better performance in managing intricate airway forms than conventional methods such as thresholding and region growing, with Dice scores of 0.84 and 0.85, respectively. The study also pinpointed another factor of dataset size. The research showed that the larger the dataset, the better the alternative model's performance. It was concluded that while the technique offers many advantages, mostly in the improvement of signal-to-noise ratios, difficulties relating to the computational overhead and the costs of implementation may limit its usage in clinical facilities. This was further pointed out by the risk of bias assessment we performed where it showed we needed large, well-validated data sets to enhance the model's accuracy. This review offers an important overview of the present status of automated airway segmentation and establishes a framework for subsequent investigations focused on the improvement of algorithm performance and usability in clinical settings.

REFERENCES

1. Sutherland K, Lowth AB, Antic N, Carney AS, Catcheside PG, Chai-Coetzer CL, Chia M, Hodge JC, Jones A, Kaambwa B, Lewis R, MacKay S, McEvoy RD, Ooi EH, Pinczel AJ, McArdle N, Rees G, Singh B, Stow N, Weaver EM, Woodman RJ, Woods CM, Yeo A, Cistulli PA. Volumetric magnetic resonance imaging analysis of multilevel upper airway surgery effects on pharyngeal structure. Sleep. 2021 Dec 10;44(12):zsab183. doi: 10.1093/sleep/zsab183. PMID: 34283220; PMCID: PMC8664571.
2. Chu G, Zhang R, He Y, Ng CH, Gu M, Leung YY, He H, Yang Y. Deep Learning Models for Automatic Upper Airway Segmentation and Minimum Cross-Sectional Area Localisation in Two-Dimensional Images. Bioengineering (Basel). 2023 Aug 2;10(8):915. doi: 10.3390/bioengineering10080915. PMID: 37627800; PMCID: PMC10451171.

3. Lenchik L, Heacock L, Weaver AA, Boutin RD, Cook TS, Itri J, Filippi CG, Gullapalli RP, Lee J, Zagurovskaya M, Retson T, Godwin K, Nicholson J, Narayana PA. Automated Segmentation of Tissues Using CT and MRI: A Systematic Review. Acad Radiol. 2019 Dec;26(12):1695-1706. doi: 10.1016/j.acra.2019.07.006. Epub 2019 Aug 10. PMID: 31405724; PMCID: PMC6878163.

4. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, van der Laak JAWM, van Ginneken B, Sánchez CI. A survey on deep learning in medical image analysis. Med Image Anal. 2017 Dec;42:60-88. doi: 10.1016/j.media.2017.07.005. Epub 2017 Jul 26. PMID: 28778026.

5. Nadeem SA, Hoffman EA, Sieren JC, Comellas AP, Bhatt SP, Barjaktarevic IZ, Abtin F, Saha PK. A CT-Based Automated Algorithm for Airway Segmentation Using Freeze-and-Grow Propagation and Deep Learning. IEEE Trans Med Imaging. 2021 Jan;40(1):405-418. doi: 10.1109/TMI.2020.3029013. Epub 2020 Dec 29. PMID: 33021934; PMCID: PMC7772272.

6. Garcia-Uceda A, Selvan R, Saghir Z, Tiddens HAWM, de Bruijne M. Automatic airway segmentation from computed tomography using robust and efficient 3-D convolutional neural networks. Sci Rep. 2021 Aug 6;11(1):16001. doi: 10.1038/s41598-021-95364-1. PMID: 34362949; PMCID: PMC8346579.

7. Wang EK, Chen CM, Hassan MM, Almogren A. A deep learning based medical image segmentation technique in Internet-of-Medical-Things domain. Future Generation Computer Systems. 2020 Jul 1;108:135-44.

8. Dong L, Lu W, Lu X, Leng X, Xiang J, Li C. Comparison of deep learning-based image segmentation methods for intravascular ultrasound on retrospective and large image cohort study. Biomed Eng Online. 2023 Nov 28;22(1):111. doi: 10.1186/s12938-023-01171-2. PMID: 38017463; PMCID: PMC10685628.

9. Ismail IN, Subramaniam PK, Chi Adam KB, Ghazali AB. Application of Artificial Intelligence in Cone-Beam Computed Tomography for Airway Analysis: A Narrative Review. Diagnostics (Basel). 2024 Aug 30;14(17):1917. doi: 10.3390/diagnostics14171917. PMID: 39272702; PMCID: PMC11394605.

10. Alsufyani NA, Flores-Mir C, Major PW. Three-dimensional segmentation of the upper airway using cone beam CT: a systematic review. Dentomaxillofacial Radiology. 2012 May 1;41(4):276-84.

11. Xie L, Udupa JK, Tong Y, Torigian DA, Huang Z, Kogan RM, Wootton D, Choy KR, Sin S, Wagshul ME, Arens R. Automatic upper airway segmentation in static and dynamic MRI via anatomy-guided convolutional neural networks. Med Phys. 2022 Jan;49(1):324-342. doi: 10.1002/mp.15345. Epub 2021 Dec 2. PMID: 34773260.

12. Soomro TA, Zheng L, Afifi AJ, Ali A, Soomro S, Yin M, Gao J. Image segmentation for MR brain tumor detection using machine learning: a review. IEEE Reviews in Biomedical Engineering. 2022 Jun 23;16:70-90.

13. Okuno K, Pliska BT, Hamoda M, Lowe AA, Almeida FR. Prediction of oral appliance treatment outcomes in obstructive sleep apnea: A systematic review. Sleep Med Rev. 2016 Dec;30:25-33. doi: 10.1016/j.smrv.2015.11.007. Epub 2015 Dec 8. PMID: 26773412.

14. Bommineni VL, Erus G, Doshi J, Singh A, Keenan BT, Schwab RJ, Wiemken A, Davatzikos C. Automatic Segmentation and Quantification of Upper Airway Anatomic Risk Factors for Obstructive Sleep Apnea on Unprocessed Magnetic Resonance Images. Acad Radiol. 2023 Mar;30(3):421-430. doi: 10.1016/j.acra.2022.04.023. Epub 2022 May 21. PMID: 35606257; PMCID: PMC10077883.

15. Fu Y, Lei Y, Wang T, Curran WJ, Liu T, Yang X. A review of deep learning based methods for medical image multi-organ segmentation. Phys Med. 2021 May;85:107-122. doi: 10.1016/j.ejmp.2021.05.003. Epub 2021 May 13. PMID: 33992856; PMCID: PMC8217246.

16. Wang P, Guo D, Zheng D, Zhang M, Yu H, Sun X, Ge J, Gu Y, Lu L, Ye X, Jin D. Accurate Airway Tree Segmentation in CT Scans via Anatomy-aware Multi-class Segmentation and Topology-guided Iterative Learning. IEEE Trans Med Imaging. 2024 Jun 26;PP. doi: 10.1109/TMI.2024.3419707. Epub ahead of print. PMID: 38923479.

17. Duan HH, Gong J, Sun XW, Nie SD. Region growing algorithm combined with morphology and skeleton analysis for segmenting airway tree in CT images. J Xray Sci Technol. 2020;28(2):311-331. doi: 10.3233/XST-190627. PMID: 32039883.

18. Liu J, Udupa JK, Odhnera D, McDonough JM, Arens R. System for upper airway segmentation and measurement with MR imaging and fuzzy connectedness. Acad Radiol. 2003 Jan;10(1):13-24. doi: 10.1016/s1076-6332(03)80783-3. PMID: 12529024.

19. Verma P, Shanker R, Patel V, Bhattacharya M. The Role of Bioinformatics and Imaging Models in Tumorigenesis and Treatment Response of Brain and Spinal Cord Neoplasm. Adv Exp Med Biol. 2023;1394:103-117. doi: 10.1007/978-3-031-14732-6_7. PMID: 36587384.

20. Doğan K, Selçuk T, Alkan A. An Enhanced Mask R-CNN Approach for Pulmonary Embolism Detection and Segmentation. Diagnostics (Basel). 2024 May 26;14(11):1102. doi: 10.3390/diagnostics14111102. PMID: 38893629; PMCID: PMC11171979.

21. Wang A, Tam TC, Poon HM, Yu KC, Lee WN. Naviairway: a bronchiole-sensitive deep learning-based airway segmentation pipeline. arXiv preprint arXiv:2203.04294. 2022 Mar 8.

22. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. InMedical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18 2015 (pp. 234-241). Springer International Publishing.

23. Amiriebrahimabadi M, Rouhi Z, Mansouri N. A Comprehensive Survey of Multi-Level Thresholding Segmentation Methods for Image Processing. Archives of Computational Methods in Engineering. 2024 Mar 27:1-51.

24. Xu Z, Bagci U, Foster B, Mansoor A, Udupa JK, Mollura DJ. A hybrid method for airway segmentation and automated measurement of bronchial wall thickness on CT. Medical image analysis. 2015 Aug 1;24(1):1-7.

25. Fang Y, Wang J, Ou X, Ying H, Hu C, Zhang Z, Hu W. The impact of training sample size on deep learning-based organ auto-segmentation for head-and-neck patients. Phys Med Biol. 2021 Sep 14;66(18). doi: 10.1088/1361-6560/ac2206. PMID: 34450599.

26. Tajbakhsh N, Jeyaseelan L, Li Q, Chiang JN, Wu Z, Ding X. Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation. Medical image analysis. 2020 Jul 1;63:101693.

27. Jardim S, António J, Mora C. Image thresholding approaches for medical image segmentation-short literature review. Procedia Computer Science. 2023 Jan 1;219:1485-92.

28. Guo W, Gu X, Fang Q, Li Q. Comparison of performances of conventional and deep learning-based methods in segmentation of lung vessels and registration of chest radiographs. Radiol Phys Technol. 2021 Mar;14(1):6-15. doi: 10.1007/s12194-020-00584-1. Epub 2020 Sep 11. PMID: 32918159.

29. Vrudhula A, Kwan AC, Ouyang D, Cheng S. Machine Learning and Bias in Medical Imaging: Opportunities and Challenges. Circ Cardiovasc Imaging. 2024 Feb;17(2):e015495. doi: 10.1161/CIRCIMAGING.123.015495. Epub 2024 Feb 20. PMID: 38377237; PMCID: PMC10883605.