# Automated vision-based assistance tools in bronchoscopy: stenosis severity estimation

Clara Tomasini[1] · Javier Rodriguez-Puigvert[1] · Dinora Polanco[2] · Manuel Viñuales[2] · Luis Riazuelo[1] · Ana C. Murillo[1]

## Abstract

**Purpose** Subglottic stenosis refers to the narrowing of the subglottis, the airway between the vocal cords and the trachea. Its severity is typically evaluated by estimating the percentage of obstructed airway. This estimation can be obtained from CT data or through visual inspection by experts exploring the region. However, visual inspections are inherently subjective, leading to less consistent and robust diagnoses. No public methods or datasets are currently available for automated evaluation of this condition from bronchoscopy video.

**Methods** We propose a pipeline for automated subglottic stenosis severity estimation during the bronchoscopy exploration, without requiring the physician to traverse the stenosed region. Our approach exploits the physical effect of illumination decline in endoscopy to segment and track the lumen and obtain a 3D model of the airway. This 3D model is obtained from a single frame and is used to measure the airway narrowing.

**Results** Our pipeline is the first to enable automated and robust subglottic stenosis severity measurement using bronchoscopy images. The results show consistency with ground-truth estimations from CT scans and expert estimations and reliable repeatability across multiple estimations on the same patient. Our evaluation is performed on our new Subglottic Stenosis Dataset of real bronchoscopy procedures data.

**Conclusion** We demonstrate how to automate evaluation of subglottic stenosis severity using only bronchoscopy. Our approach can assist with and shorten diagnosis and monitoring procedures, with automated and repeatable estimations and less exploration time, and save radiation exposure to patients as no CT is required. Additionally, we release the first public benchmark for subglottic stenosis severity assessment.

## Introduction

Subglottic stenosis (SGS) is a condition characterized by the narrowing of the airway in the subglottis, just below the vocal cords and before the trachea, which results in airway obstruction airflow restriction. The percentage of narrowed airway or stenosis index (SI) is obtained by comparing the obstructed airway area to that of a posterior part of the airway without obstruction. In clinical practice, physicians estimate the SI using either a CT scan, which involves exposure to radiation, or during a bronchoscopy procedure. In bronchoscopy, the estimation is performed visually or by fitting endotracheal tubes of known dimensions through the stenosis. Bronchoscopy-based estimates are inherently subjective, making diagnosis less robust and potentially leading to non-optimal treatment strategies [1].

The subjectivity of bronchoscopy-based assessments has been previously shown [1, 2]. When asked to assess stenosis grade in 1447 patients [2], physicians tended to wrongly classify it (53% of cases assessed the wrong severity level), mostly underestimating it (91% of wrongly classified cases were underestimations). While CT offers a more objective alternative, its utility for continuous monitoring is limited due to the patient cumulative radiation exposure. To reduce assessment subjectivity, there is a need for a standardized measuring tool [1], using bronchoscopy images to avoid repeated patient radiation exposure from CT scans. This work

✉ Clara Tomasini
ctomasini@unizar.es

1 DIIS, i3A. Universidad de Zaragoza, Zaragoza, Spain

2 Hospital Universitario Miguel Servet, Zaragoza, Spain

introduces the first pipeline to enable automated and robust estimation of subglottic stenosis severity using bronchoscopy images only.

However, public real-world bronchoscopy image data is notably scarce, in particular including subglottic stenosis recordings. This motivates the self-supervised nature of the main step in our approach, as discussed later, and also the collection and release of our Subglottic Stenosis Dataset from real procedures. Figure 1 summarizes the proposed pipeline, that relies on illumination decline in endoscopy to segment and track the airway lumen, and to obtain a 3D model of the explored airway applying LigthDepth [3]. The 3D model is obtained from a single frame automatically selected right after the vocal cords and is used to estimate airway narrowing.

Overall, the contributions of this work are twofold:

- A novel approach to enable automated Subglottic Stenosis measurement based only on bronchoscopy video, acquired during regular physicians exploration but without requiring the endoscope to traverse the stenosis. Our work is the first to demonstrate the use of *illumination decline* as a supervisory signal for the acquisition of 3D reconstructions in bronchoscopy procedures. Our promising results indicate that our approach can help reduce the subjectivity involved in diagnosing of SGS. Besides, it can reduce exploration time and minimize the use of CT scans, benefiting patients and assisting physicians in diagnosis and monitoring procedures.
- A novel benchmark, Subglottic Stenosis Dataset, for SGS severity estimation[1], with data obtained from real bronchoscopy procedures along with ground-truth from CT scans and expert estimations. Public benchmarks are needed to encourage further research on automated stenosis evaluation.

## Related work

### Stenosis estimation

Physicians generally agree [1] that a standardized stenosis measuring tool is needed to reduce uncertainty and variability observed in traditional clinical methods. Fully automatic approaches for SGS severity estimation have been proposed, though most use other imaging modalities such as CT, MRI or Optical Coherence Tomography [4]. Bronchoscopy images have been used in [5] to estimate tracheal stenosis. The narrowing is measured by considering two frames in inspiration and expiration phases and using depth estimation to evalu-

ate the area of the opening without and with obstruction [5]. Alternatively, airway areas can be estimated through segmentation of the obstructed lumen and reference healthy tracheal ring [6]. Segmentation of the lumen can be obtained from depth estimation [7] or image pixels intensity [8], since the lumen usually corresponds to the deepest or darkest area of the image. However, these methods cannot be directly applied to SGS, as they rely on tracheal anatomical features. The subglottis, being the lower portion of the larynx cavity, does not exhibit the cartilage rings and the breathing deformation seen in the trachea and used in [5, 6]. As very little data is available for stenosis estimation, training-free methods such as SLIC superpixels [9], and foundation models for image segmentation like SAM [10] can also prove relevant. For specific SGS estimation using bronchoscopy images, [11, 12] propose a semi-automatic method that requires physician input to measure the stenosis using image processing software.

### SGS estimation data

No publicly available datasets currently exist for stenosis estimation, limiting the development of a fully automatic tool. A few datasets exist, however, for the task of lumen segmentation, such as the CVC-LumenDB dataset [8], which has very low resolution, or the Phantom dataset [13], consisting of images obtained in a phantom airway instead of real bronchoscopy images. Endoscopy datasets with high-resolution real images are more widely available for gastroscopy and colonoscopy, like EndoMapper [14] or Kvasir [15].

## Methodology

### Overview

During a bronchoscopy, physicians place the bronchoscope in front of the vocal cords to ensure a clear view of the airway cavity to navigate towards the trachea and avoid visual occlusions that could affect accurate stenosis assessment. They proceed through the airway towards the lumen, which usually appears as the darkest area in the image, and visually estimate stenosis by evaluating the lumen visible in the image, using their prior knowledge and the bronchoscope diameter.

Building on this foundation, our approach, presented in Fig. 2, first detects and tracks the airway lumen to select the optimal frame for stenosis evaluation. In that selected frame, we identify the stenosed region and estimate its severity by using a detailed 3D model generated based on the illumination decline.

We use illumination decline for the two key steps of our proposed approach: (1) to track the deepest area in the image, corresponding to the airway lumen, to detect when the camera passes the vocal cords and to discern between stenosed and

---

[1] Data available at https://sites.google.com/unizar.es/subglottic-stenosis-estimation/home.
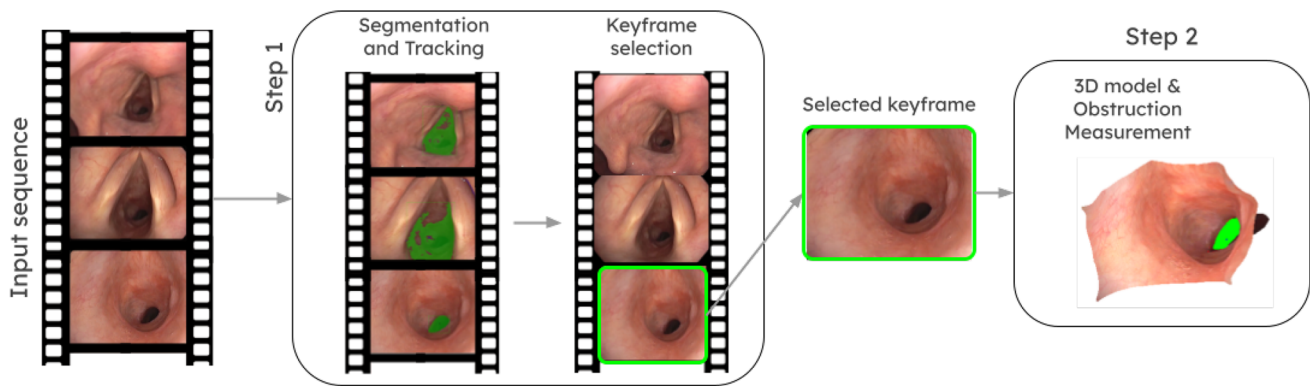
**Fig. 1** Overview of the proposed automatic stenosis estimation approach. Our two-step pipeline takes as input the bronchoscopy video, starting in front of the vocal cords. In each frame, the lumen is segmented and tracked until the camera reaches the subglottis and starts to see the stenosis. At that selected keyframe, we compute a 3D reconstruction of the explored airway, in which the stenosis is measured

non-stenosed areas (see "Segmentation, tracking and frame selection" Section), and (2) to build a 3D model where we estimate the stenosis (see "3D reconstruction and stenosis estimation" Section).

In the unique environment of endoscopic imaging, the co-location of camera and illumination source enables the recovery of geometric information through photometric analysis [3, 16, 17]. We observe that the colour palette in bronchoscopic images is significantly influenced by the scene illumination. Therefore, we employ the spotlight light model of [18], later used in [3] to reconstruct the scene geometry by employing the illumination decline information. Moreover, in endoscopic imagery, the intensity of the image $\mathcal{I}$ decreases with the distance $d$, following the inverse square law and the gamma correction $\gamma$, $\mathcal{I} \approx (1/d^2)^{1/\gamma}$. When the intensity falls below a certain intensity level, the available photometric information might no longer be sufficient for reliable 3D reconstruction. However, this information is useful for navigation during the procedure following the darkest area and identifying the stenosed tissue.

### Segmentation, tracking and frame selection

This step segments and tracks the darkest area along the sequence until it finds the best keyframe for subglottic stenosis estimation, just below the vocal cords. The darkest regions are segmented with a binary intensity threshold over the grayscale image. Then, we track the resulting segments with an IoU (Intersection Over Union) tracker [19]. It is necessary to allow a few consecutive frames to fail in this tracking to make the approach robust to sudden illumination changes.

When the tracker stops, it means that the darkest area segment has significantly changed, i.e. its IoU with the previous frame tracked segment is below the minimum IoU established. At this point, we consider the camera has passed the vocal cords and reached the subglottis, and the darkest part of

the image has shifted from corresponding to the vocal cords to the stenosis. This frame is selected as keyframe and used to run the second step of our pipeline.

### 3D reconstruction and stenosis estimation

Once the keyframe is selected, we compute the 3D reconstruction from this single view, employing the LightDepth model, as introduced by [3]. This model processes an input image $I$ and outputs an up-to-scale depth map $\widehat{d}$, an albedo map $\widehat{\rho}$, and normals $\widehat{n}$ derived from $\widehat{d}$. LightDepth presents state-of-the-art results in other endoscopic settings, ideal for the illumination decline hypothesis, while experiments with DepthAnything [20] and COLMAP [21, 22] produced less detailed, noisier reconstructions.

Our 3D model is defined by a set of 3D points $X_{3D} = \pi^{-1}(\widehat{d}, K)$, where $\pi^{-1}$ denotes the projection function, and $K$ represents the camera intrinsics. As described in "Segmentation, tracking and frame selection" Section, we segment the stenosed area using illumination decline, i.e. we detect the darkest region in the image. By segmenting the stenosed region in the image and finding its correspondence in the obtained 3D model, we can precisely delineate the 3D contour of the stenotic region. This set of 3D points, $X_{3Dstenosis}$, outline the region and define a common plane, $\mathcal{P}$. We intersect $\mathcal{P}$ with the 3D model to obtain the area corresponding to the stenosis, $A_{stenosis}$ (see green region in the example from Fig. 2 step 2):

$$X_{3Dstenosis} = \{X \in X_{3D} | X \in \mathcal{P}\}, \tag{1}$$

In addition to delineating the stenosed area, we generate a series of planes $\mathcal{Q}_k$ oriented perpendicular to the camera to intersect with the 3D reconstruction $\mathcal{X}_{3D}$. The plane that results in the maximum intersection area $A_k$ is selected as
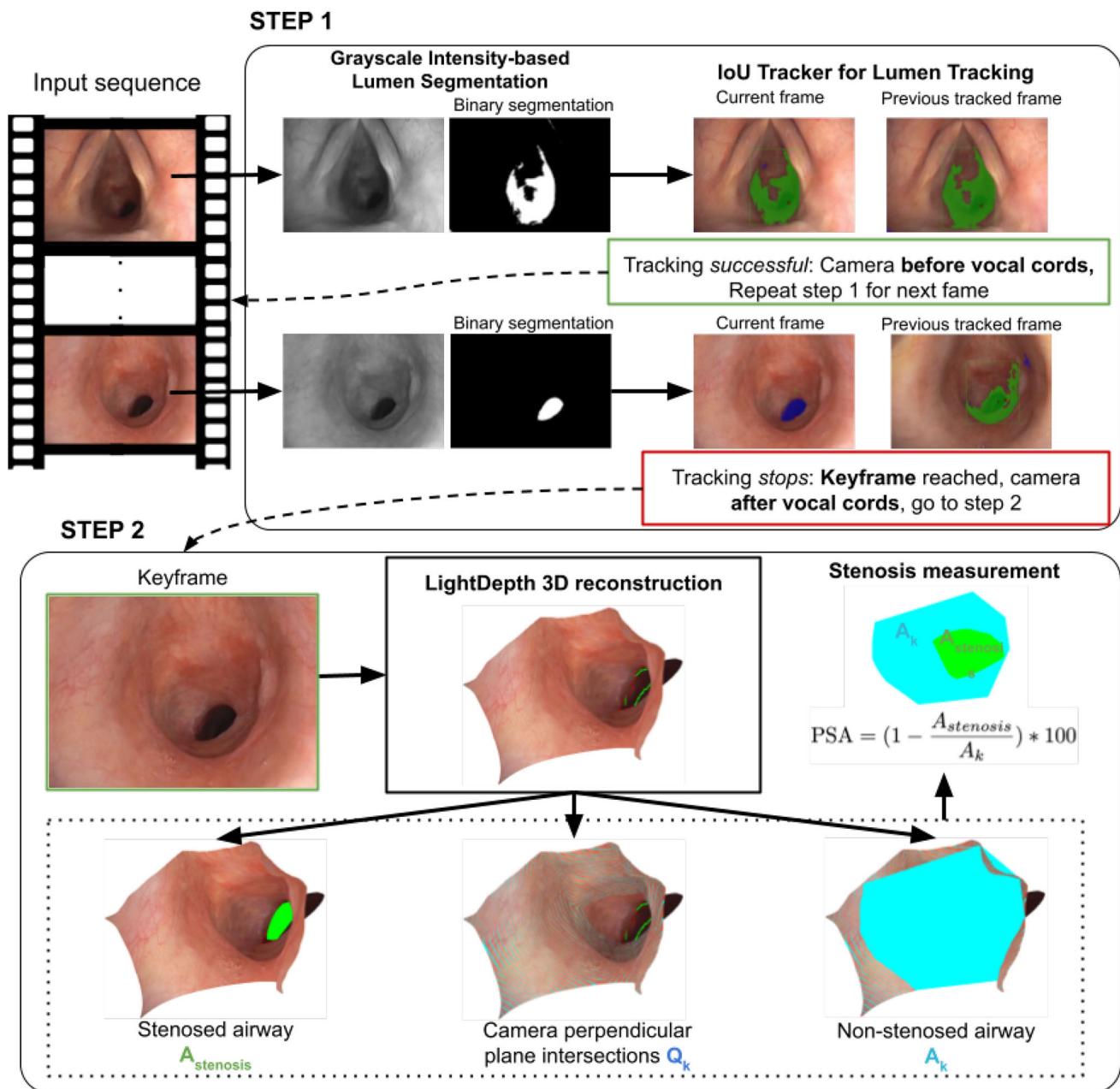
**Fig. 2** Overview of our proposed approach. The first step detects the darkest area through grayscale intensity segmentation and tracks it along the sequence using IoU tracker. Once the camera reaches the adequate location to measure the stenosis, the darkest segment changes shape significantly. The tracking is lost and the keyframe is reached. That keyframe is then passed to our second step, which produces a 3D reconstruction of the airway using LightDepth model. The stenosed and non-stenosed airway areas are measured in that model

the optimal reference area just beyond the vocal cords (blue region in the example from Fig. 2 step 2).

$$k = \underset{k}{\operatorname{argmax}} A_k \qquad (2)$$

We quantify the degree of airway stenosis observed, $PSA$, by computing the ratio of the stenosed area ($A_{stenosis}$) to the maximum area measured in the airway between the vocal cords and the stenosis ($A_k$). Additionally, to approximate the obstruction in terms of airway diameter, similar to the method proposed [23], we fit circles to the delineated stenosed and non-stenosed areas in the 3D reconstruction, measure their diameters and compute the ratios ($PSD$).

$$PSA = \left(1 - \frac{A_{stenosis}}{A_k}\right) * 100,$$

$$PSD = \left(1 - \frac{diameter_{stenosis}}{diameter_k}\right) * 100 \qquad (3)$$

### Subglottic stenosis dataset

Our subglottic stenosis (SGS) dataset encompasses 16 bronchoscopy videos from 11 patients (A through J), recorded during real medical procedures using three bronchoscopes: Olympus BF-H1100, Olympus BRF-180 and Olympus BF-1TH1100. Table 1 summarizes the dataset contents, with examples in Fig. 3. For Patients A, C and D, we include CT scans with ground-truth stenosis measurements in terms of areas, obtained from intersections between a predefined plane—at the trachea's widest point and stenosis location—and the 3D CT reconstruction. For patients A, B and D to J, obstruction percentage visually estimated by experts—by comparing estimated stenosis diameter to average airway diameter without stenosis as done in [23]—is available as reference. Patients G, H and I are healthy without stenosis. Patients B, D and E have multiple bronchoscopy videos under identical pathological conditions by different physicians.

## Experiments

### Experimental setup

*Segmentation and tracking parameters* Segmentation involves detecting the pixels in the grayscale image with intensity below 50, chosen following the illumination decline and the inverse square law, described in "Overview" Section, and tested though visual assessment on real colonoscopy recordings from EndoMapper dataset [14]. This threshold remains effective despite small illumination variations within each bronchoscopy sequence. Tracking relies on a minimum IoU of 50% for detection and 25 frames before a tracked segment is lost.

*3D reconstruction model training* The 3D reconstruction at the keyframe is obtained using LightDepth [3], in its LightDepth [3] U-Net [24] version for faster inference time. We train LightDepth on 2800 frames from sequence 1 of the Subglottic Stenosis dataset, covering various anatomical regions, including the subglottic and tracheal areas, with the corresponding geometric and photometric calibration of the bronchoscope. This sequence is used for LightDepth training, but is kept to evaluate the quality of stenosis estimation, as it includes scarce CT ground-truth data and that errors in this sequence are similar than in the rest. LightDepth is trained for 5 epochs using a learning rate of $lr = 10^{-4}$, a depth smoothness regularization weight of $\lambda_s = 0.1$, and a

specularity-aware loss component with $\lambda_{sp} = 1$. The encoder is initialized from ImageNet pre-trained weights. The albedo and depth decoders are trained from scratch.

*Metrics* The main metrics computed to evaluate the accuracy and consistency of our approach are:

- Percentage of correctly selected keyframes (**Correct Keyframes**), which measures if the keyframe selected, using our first step ("Segmentation, tracking and frame selection" Section), falls within a manually defined optimal frame interval, spanning from when the bronchoscope is just below the vocal cords to just before reaching the stenosis.
- Value range and absolute difference (**Diff.**) between estimations (PSD or PSA) computed on sequences of the same patient and condition.
- Absolute error (AE) for each sequence and mean absolute error (**MAE**) to quantify the average error in estimated PSA and PSD for all sequences with PSA or PSD reference. Given a sequence $i$ and its corresponding ground-truth $GT_i$ and computed estimation $ce_i$, the MAE over N sequences is computed as follows:

$$MAE = \frac{\sum_{i=1}^{N} |GT_i - ce_i|}{N} = \frac{\sum_{i=1}^{N} AE_i}{N}. \qquad (4)$$

## Results

This section presents the quantitative evaluation, regarding accuracy and consistency, of our full proposed pipeline. The supplementary video in Online Resource 2 shows how our pipeline runs on sequences 1 and 2 of the SGS dataset.

*Comparison to CT scan and expert estimations* First, we compare the stenosis estimations obtained using our pipeline, on the 16 sequences of SGS dataset, to the CT scan PSA ground-truth and/or expert PSD estimations.

As no baseline is available for the task of SGS estimation, up to our knowledge, we implement several baselines to better illustrate the performance of our approach. In particular, we consider several alternatives for the critical initial stage of keyframe selection. We replace our intensity-based segmentation (*Ours*, "Segmentation, tracking and frame selection" Section) with well-known segmentation alternatives to achieve the target (lumen) segment, maintaining the rest of the keyframe selection process ("Segmentation, tracking and frame selection" Section, tracking and end-point selection) and 3D reconstruction and estimation step ("3D reconstruction and stenosis estimation" Section) as state-of-the-art methods. The alternative baselines considered are:

**Table 1** Overview of the released dataset. 16 bronchoscopy videos from 11 patients. 12 sequences with reference obstruction percentage estimation from experts (**Expert estimation**), 5 with corresponding CT scan. Repeated procedures for patients B, D and E from different physicians

| Sequence | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Patient ID | A | B | B | B | C | D | D | D | E | E | F | G | H | I | J | K |
| CT scan | ✓ | ✗ | ✗ | ✗ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Expert estimation | ✓ | ✓ | ✓ | ✗ | ✗ | ✓ | ✓ | ✗ | ✓ | ✗ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Scope ID | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 3 | 3 | 1 | 3 | 1 | 3 |
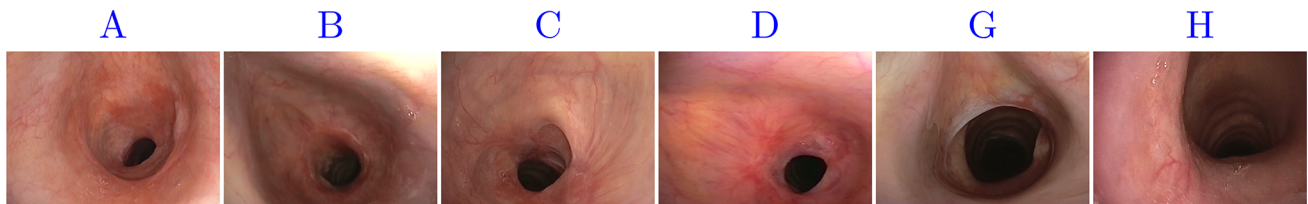


**Fig. 3** Examples of subglottic area appearance for patients A, B, C, D, G and H from our proposed dataset. Patients A, B, C and D present varying degrees of stenosis. Patients G and H are healthy, and the trachea with its characteristic cartilaginous rings can be seen in the back

- *SAM-based*, uses SAM [10] segmentation prompted with the darkest pixel.
- *SLIC-based*, uses SLIC [9] superpixel segmentation, selecting the segment that contains the darkest pixel.
- *Manual*, manually selecting the first frame right after the bronchoscope passes the vocal cords, as a reference lower bound for the error of our second stage ("3D reconstruction and stenosis estimation" Section).

Table 2 presents the average results obtained over all sequences. Detailed results per sequence are shown in Tables 1 and 2 of Online Resource 1. Our pipeline, using the proposed intensity-based segmentation, consistently selects keyframes within the ground-truth interval for all sequences, while the other alternatives present several failure cases. Moreover, our proposed approach has significantly lower PSA and PSD error and achieves the closest error to that obtained when manually given the best keyframe possible. Additionally, paired *t*-tests between ground-truth PSA (reference expert PSD respectively) and corresponding estimates obtained using our full method, with *p*-value 0.6039 (0.2766 respectively), greater than 0.05, shows no statistically significant difference between means. Overall, this experiment highlights the ability of our approach to accurately estimate the stenosis.

*Same patient and interexpert consistency* This second experiment analyses the consistency of our estimations in the SGS Dataset. This analysis can only use sequences corresponding to patients (B, D and E) with several procedures without patient condition changes between the sequences. The PSA and PSD for both sequences of the same patient should be similar or close in range. For patients B and D, two expert PSD estimations are also available, and their variations are

shown as reference. Table 3 presents estimation range and difference between minimum and maximum estimations. Using our approach, estimations differ on average 1.79% in PSA and 5.07% in PSD between the two sequences, while expert PSD estimates vary on average 7.5%. Overall, our pipeline shows consistency over repeated estimations for the same patient on comparable ranges than expert estimations.

## Conclusion and discussion

In this work, we have presented a pipeline to automatically evaluate subglottic stenosis using only bronchoscopy videos, without traversing the stenosed region. Our two-step approach first selects the optimal frame for measuring the stenosis through segmentation and tracking and then uses a dynamically computed 3D airway reconstruction to measure the stenosis obstruction at that keyframe.

We evaluate our approach using our novel benchmark for subglottic stenosis estimation, the first of its kind being publicly released to facilitate reproducibility and encourages further research. Our results show that our pipeline provides an assistance tool in real procedures for first robust stenosis severity estimation and to monitor the condition during follow-up procedures, to assess the stenosis evolution. Our method standardizes stenosis measurement, reducing the subjectivity traditionally observed, and easily integrates into existing bronchoscopy workflows without needing extra hardware or doctor training. Moreover, measuring at the keyframe takes 7 s on average on a GPU GeForce RTX3090 and could run in parallel to doctor assessment.

However, there are few limitations to address in future work. Our method slightly underestimates the stenosis index (SI), and our healthy reference measurement location dif-

**Table 2** Proposed stenosis estimation (Ours) compared to other baselines with different keyframe selection approaches. Mean absolute error (MAE) is computed for **all tests** but also considering only tests where the selected keyframes are correct (**Correct KF only**)

| Keyframe Selection | Correct Keyframes (%) | PSA MAE (%) | | PSD MAE (%) | |
|---|---|---|---|---|---|
| | | Correct KF only | All tests | Correct KF only | All tests |
| SAM-based | 37.5 | 6.43 | 39.84 | 15.33 | 34.72 |
| SLIC-based | 25 | 16.07 | 52.40 | 5.34 | 28.44 |
| Ours | **100** | 4.84 | **4.84** | 6.70 | **6.70** |
| Manual | 100 | 1.21 | 1.21 | 5.31 | 5.31 |

**Table 3** Stenosis estimation consistency analysis. PSA and PSD estimation range (**Range**) and difference (**Diff.**) between estimations on two sequences corresponding to the same patient (**S.-P.**). Estimates obtained using two methods: our approach for PSA and PSD, and from expert estimations for PSD (patients B and D)

| S.-P. | Method | PSA range (Diff.) (%) | PSD range (Diff.) (%) |
|---|---|---|---|
| 2,3-B | Ours | 58.51-62.16 (3.65) | 62.94-55.70 (7.24) |
| 2,3-B | Experts | / | 55-60 (5) |
| 6,7-D | Ours | 61.90-62.35 (0.45) | 57.89-59.09 (1.20) |
| 6,7-D | Experts | / | 60-70 (10) |
| 9,10-E | Ours | 64.58-65.85 (1.27) | 58.95-65.71 (6.76) |
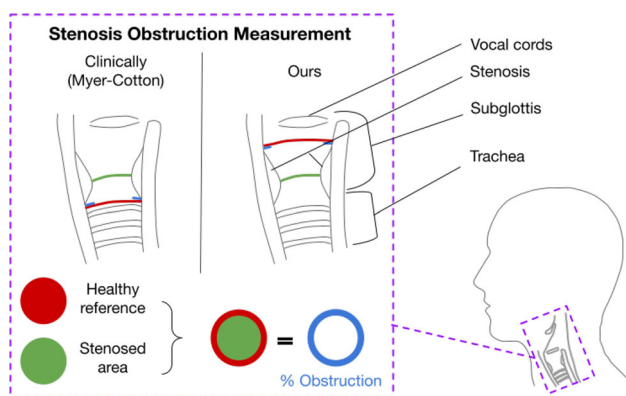


**Fig. 4** Obstruction measurement as done clinically (Myer–Cotton classification) compared to our approach. The green area represents the stenosed area, for which the measurement is the same clinically and in our case. The red area corresponds to the healthy reference airway: clinically measured in the trachea below the stenosis, in our approach measured between the vocal cords and the stenosis. The obstruction (in blue) is the difference between the healthy and stenosed areas

fers from the clinically used Myer–Cotton (MC) classification [23], as seen in Fig. 4. We measure the largest airway area right after the vocal cords, before traversing the stenosis, while the MC classification typically considers the area of the trachea after the stenosis. Future work thus includes collecting and releasing more expert-supervised data to calibrate our approach and validate a direct comparison with the MC classification and possibly correct the slight underestimation.

Our work paves the way for future research using 3D geometric information to reduce subjectivity and procedure time, benefiting patients and physicians in diagnosis and monitoring.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

# References

1. Begnaud A, Connett JE, Harwood EM, Jantz MA, Mehta HJ (2015) Measuring central airway obstruction. What do bronchoscopists do? Ann Am Thorac Soc 12(1):85–90
2. Murgu S, Colt H (2013) Subjective assessment using still bronchoscopic images misclassifies airway narrowing in laryngotracheal stenosis. Interact Cardiovasc Thorac Surg 16(5):655–660
3. Rodríguez-Puigvert J, Batlle VM, Montiel JMM, Martinez-Cantin R, Fua P, Tardós JD, Civera J (2023) Lightdepth: Single-view depth self-supervision from illumination decline. In: IEEE/CVF Int. Conf. on Computer Vision (ICCV), pp. 21273–21283
4. Sharma GK, Chin Loy A, Su E, Jing J, Chen Z, Wong BJ, Verma S (2016) Quantitative evaluation of adult subglottic stenosis using intraoperative long-range optical coherence tomography. Annal Otol Rhinol Laryngol 125(10):815–822
5. Banach A, Naito M, King F, Masaki F, Tsukada H, Hata N (2023) Computer-based airway stenosis quantification from bronchoscopic images: preliminary results from a feasibility trial. Int J Comput Assist Radiol Surg 18(4):707–713
6. Sánchez C, Bernal J, Sánchez FJ, Diez M, Rosell A, Gil D (2015) Toward online quantification of tracheal stenosis from videobronchoscopy. Int J Comput Assist Radiol Surg 10:935–945
7. Keuth R, Heinrich M, Eichenlaub M, Himstedt M (2023) Weakly supervised airway orifice segmentation in video bronchoscopy. In: Medical Imaging 2023: Image Processing, vol. 12464, pp. 58–65. SPIE
8. Sánchez C, Bernal J, Gil D, Sánchez FJ (2014) On-line lumen centre detection in gastrointestinal and respiratory endoscopy. In: CLIP 2013, Held in Conjunction with MICCAI 2013, pp. 31–38. Springer
9. Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Süsstrunk S (June 2010) Slic superpixels. EPFL Technical Report 149300
10. Kirillov A, Mintun E, Ravi N, Mao H, Rolland C, Gustafson L, Xiao T, Whitehead S, Berg AC, Lo W-Y, Dollár P, Girshick R (2023) Segment anything. IEEE/CVF Int. Conf. on Computer Vision, 4015–4026
11. Murgu S, Colt HG (2009) Morphometric bronchoscopy in adults with central airway obstruction: case illustrations and review of the literature. Laryngoscope 119(7):1318–1324
12. Francom CR, Best CA, Eaton RG, Pepper V, Onwuka AJ, Breuer CK, Lind MNM, Grischkan JM, Chiang T (2019) Clinical validation and reproducibility of endoscopic airway measurement in pediatric aerodigestive evaluation. Int J Pediatr Otorhinolaryngol 116:65–69
13. Visentini-Scarzanella M, Sugiura T, Kaneko T, Koto S (2017) Deep monocular 3d reconstruction for assisted navigation in bronchoscopy. Int J Comput Assist Radiol Surg 12:1089–1099
14. Azagra P, Sostres C, Ferrández Á, Riazuelo L, Tomasini C, Barbed OL, Morlana J, Recasens D, Batlle VM, Gómez-Rodríguez JJ, Elvira R, López J, Oriol C, Civera J, Tardós J, Murillo AC, Lanas Á, Montiel JMM (2023) Endomapper dataset of complete calibrated endoscopy procedures. Sci. Data 10(1):671
15. Pogorelov K, Randel KR, Griwodz C, Eskeland SL, Lange T, Johansen D, Spampinato C, Dang-Nguyen D-T, Lux M, Schmidt PT, Riegler M, Halvorsen P (2017) Kvasir: A multi-class image dataset for computer aided gastrointestinal disease detection. In: 8th ACM on Multimedia Systems Conference, pp. 164–169
16. Batlle VM, Montiel JMM, Fua P, Tardós JD (2023) Lightneus: Neural surface reconstruction in endoscopy using illumination decline. In: MICCAI, pp. 502–512. Springer, Cham
17. Batlle VM, Montiel JMM, Tardós JD (2022) Photometric single-view dense 3D reconstruction in endoscopy. In: IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS), pp. 4904–4910. https://doi.org/10.1109/IROS47612.2022.9981742
18. Modrzejewski R, Collins T, Hostettler A, Marescaux J, Bartoli A (2020) Light modelling and calibration in laparoscopy. Int. J. Comput Assist Radiol Surg 15(5):859–866
19. Deshpande AM (2020) Multi-object trackers in Python. GitHub
20. Yang L, Kang B, Huang Z, Xu X, Feng J, Zhao H (2024) Depth anything: Unleashing the power of large-scale unlabeled data. In: CVPR
21. Schönberger JL, Zheng E, Pollefeys M, Frahm J-M (2016) Pixel-wise view selection for unstructured multi-view stereo. In: ECCV
22. Schönberger JL, Frahm J-M (2016) Structure-from-motion revisited. In: CVPR
23. Myer CM III, O'Connor DM, Cotton RT (1994) Proposed grading system for subglottic stenosis based on endotracheal tube sizes. Annal Otol Rhinol Laryngol 103(4):319–323
24. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. In: MICCAI, pp. 234–241. Springer