BioMedInformatics

MDPI

Article

AdaDenseNet-LUC: Adaptive Attention DenseNet for Laryngeal Ultrasound Image Classification

Cunyuan Luan [1] and Huabo Liu [1,2,*]

[1] School of Automation, Qingdao University, Qingdao 266071, China; luancunyuan@qdu.edu.cn
[2] Shandong Key Laboratory of Industrial Control Technology, Qingdao 266071, China
* Correspondence: hbliu@qdu.edu.cn

Abstract

Evaluating the difficulty of endotracheal intubation during pre-anesthesia assessment has consistently posed a challenge for clinicians. Accurate prediction of intubation difficulty is crucial for subsequent treatment planning. However, existing diagnostic methods often suffer from low accuracy. To tackle this issue, this study presented an automated airway classification method utilizing Convolutional Neural Networks (CNNs). We proposed Adaptive Attention DenseNet for Laryngeal Ultrasound Classification (AdaDenseNet-LUC), a network architecture that enhances classification performance by integrating an adaptive attention mechanism into DenseNet (Dense Convolutional Network), enabling the extraction of deep features that aid in difficult airway classification. This model associates laryngeal ultrasound images with actual intubation difficulty, providing healthcare professionals with scientific evidence to help improve the accuracy of clinical decision-making. Experiments were performed on a dataset of 1391 ultrasound images, utilizing 5-fold cross-validation to assess the model's performance. The experimental results show that the proposed method achieves a classification accuracy of 87.41%, sensitivity of 86.05%, specificity of 88.59%, F_1 score of 0.8638, and AUC of 0.94. Grad-CAM visualization techniques indicate that the model's attention is attention to the tracheal region. The results demonstrate that the proposed method outperforms current approaches, delivering objective and accurate airway classification outcomes, which serve as a valuable reference for evaluating the difficulty of endotracheal intubation and providing guidance for clinicians.

Keywords: difficult airway; intubation difficulty; artificial intelligence; attention mechanism

Check for updates

Academic Editors: James C. L. Chow and Alexandre G. De Brevern

Received: 2 December 2025
Revised: 12 January 2026
Accepted: 15 January 2026
Published: 16 January 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

A difficult airway denotes the clinical challenges faced by anesthesiologists during mask ventilation or endotracheal intubation [1]. The reported incidence of difficult mask ventilation ranges from 1.4% to 5.0%, whereas difficult endotracheal intubation varies between 1.9% and 10% [2–4]. Such unpredictable events increase the risks of brain injury and death [5,6] and require specialized skills and complex procedures. Consequently, preoperative identification of patients at risk of a difficult airway is essential to reduce complications and anesthesia-related mortality [7]. Despite decades of study, predicting a difficult airway remains challenging in routine practice. Bedside screening tools such as the Mallampati score, thyromental distance and related tests are constrained by subjectivity and substantial interobserver variability, which can lead to misclassification. Moreover, these tools capture only coarse surrogates of airway anatomy and often fail to model the complex contextual and sequential dependencies that determine intubation difficulty. AI

systems based on facial photographs or other external markers also face practical barriers: they are susceptible to imaging artifacts, generalize poorly across acquisition protocols, and therefore struggle to translate across clinical settings. These limitations motivate more objective imaging based assessment. Ultrasound offers a noninvasive modality available at the bedside that directly visualizes internal airway structures, for example the trachea, the epiglottis and the tongue base, and it provides more reproducible anatomic information than external inspection. Laryngeal ultrasound enables dynamic evaluation during quiet breathing or phonation and offers repeatability without ionizing radiation. However, it also presents distinct challenges, including acoustic shadowing caused by air and cartilage, dependence on the operator and on probe orientation, and heterogeneity across scanners and protocols, compounded by typically small or weakly labeled datasets. Accordingly, there is a strong need for ultrasound based deep learning frameworks that extract discriminative features directly from internal airway images and that are designed for robustness and clinical utility in difficult airway assessment. Moreover, due to subjectivity and contextual variability, the predictive performance of the modified LEMON criteria and the Simplified Airway Risk Index SARI remains unstable, with limited consistency across protocols and populations [8,9].

With the global rapid expansion of Artificial Intelligence (AI) applications, its auxiliary role in medical practice has increasingly become a focus of academic and clinical attention. Deep learning, a core technology in AI [10], offers substantial advantages by effortlessly identifying trends in data that are challenging for experts to detect [11]. Significant progress has been made by deep learning in recent years, especially in the detection and diagnosis of clinical conditions including lung cancer [12], breast cancer [13], diabetic retinopathy [14], stroke sequelae [15], and early Alzheimer's disease. For patients with high risk of difficult intubation, anatomical airway abnormalities are usually present, and anesthesiologists rely on visual observation to identify these abnormalities. However, AI-based technologies can objectively, accurately, and reliably identify visual clues and handle subtle differences. To date, several studies have attempted to develop AI-based image systems for the identification and management of difficult airways, covering methods such as attention mechanisms combined with manual Mallampati scoring [16], deep learning models based on facial images [17], and fully automated semi-supervised deep learning methods [18]. Difficult airway management continues to encounter multiple obstacles in current AI methodologies, including algorithmic obsolescence, variability in imaging protocol adherence, and suboptimal performance in clinical prediction tasks.

This study sought to apply advanced deep learning technology through the development of AdaDenseNet-LUC, a specialized AI model designed to correlate ultrasound images of surgical patients with the real-world challenges of intubation. The proposed AdaDenseNet-LUC framework aims to provide doctors with a more precise and reliable preoperative assessment of intubation difficulty by leveraging its adaptive attention mechanism and deep feature extraction capabilities. This innovative approach is expected to optimize the surgical process, reduce intubation-related risks, and ultimately improve patient outcomes through enhanced prediction accuracy. By implementing AdaDenseNet-LUC in clinical practice, we anticipate significant improvements in patient rescue success rates and overall surgical safety.

2. Methods

2.1. Study Design and Participants

This study was conducted on 300 patients planned for surgery under total anesthesia at the Affiliated Hospital of Qingdao University in China. Each participant provided written informed consent. All patients participating in this study consented to the inclusion

of their images, clinical data, and other related information for publication. This study adheres to all applicable regulations, and patients were included according to the following four criteria: (1) the patient is over 18 years old; (2) patients undergoing planned surgery under general anesthesia; (3) patients categorized as having an ASA (American Society of Anesthesiologists) clinical status of I–III; (4) patients who voluntarily signed the informed consent. Acquisition of Patients' Laryngeal Ultrasound images (Figure 1). To maintain the integrity of anatomical structures, the original laryngeal ultrasound images were preserved in their native aspect ratio during preprocessing. The raw output of the ultrasound device generated images with a resolution proportional to 549 × 364 pixels. Therefore, we adopted this non-standard size to avoid distortion that would occur if the images were resized to common square dimensions (e.g., 224 × 224). This choice ensured that spatial relationships among airway structures, such as the trachea and epiglottis, were retained without stretching or compression, while also balancing feature preservation with computational efficiency. In clinical ultrasound practice, the imaging depth (echo depth) may vary across subjects because of differences in neck circumference and soft-tissue thickness; consequently, minor depth adjustments are sometimes required to ensure complete visualization of the laryngeal region. In this study, all images were retained at their native resolution and aspect ratio to avoid geometric distortion. To alleviate residual scale variability associated with depth changes, we employed a data augmentation pipeline that included mild scaling transformations. In this study, the ground-truth labels for airway difficulty were obtained from intraoperative video laryngoscopic assessments, following the Cormack–Lehane (C-L) classification: Grade I, full view of the glottis; Grade IIa, posterior glottis visible; Grade IIb, no glottic opening with only arytenoids/posterior epiglottis visible; Grade III, epiglottis only; Grade IV, neither epiglottis nor glottic structures visible. The grading was independently performed by two anesthesiologists with $\geq$5 years of clinical experience. In cases of disagreement, a consensus was reached through discussion. To assess inter-observer agreement prior to consensus, we calculated Cohen's Kappa, which showed a substantial level of agreement ($\kappa = 0.81$). To ensure consistency, all evaluations were conducted at the best obtainable view in the sniffing position, permitting external laryngeal pressure (BURP) if needed. According to clinical convention, we defined C-L Grades I–II as the non-difficult (simple) group and Grades III–IV as the difficult group. This labeling strategy ensured that the dataset reflects intraoperative ground-truth outcomes rather than purely visual interpretation of ultrasound images. Among the 300 enrolled subjects, ultrasound image acquisition was performed using a standardized protocol, where one echo image was captured per subject at a fixed view/angle. Three images with extremely poor quality (e.g., severe artifacts or inadequate visualization of laryngeal structures) were excluded during quality control, resulting in a final dataset of 297 images. The class-wise distribution was as follows: Class I ($n = 121$), Class II ($n = 68$), Class III ($n = 55$), and Class IV ($n = 53$). Accordingly, the simple-airway group (Classes 1–2) contained 189 images, whereas the difficult-airway group (Classes 3–4) contained 108 images.

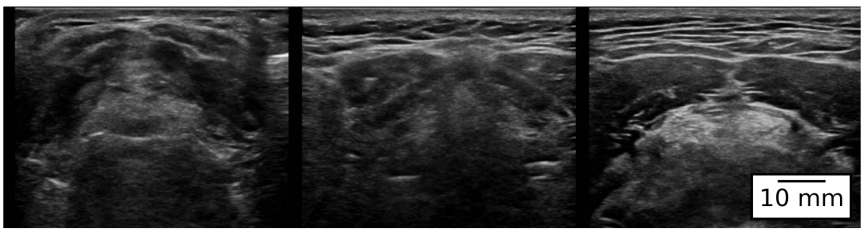


Figure 1. Laryngeal ultrasound image of surgical patients.

2.2. Data Augmentation

In the obtained images, 80% of the data was allocated for training, while the rest was used for inference and evaluation. Additionally, we observed that the original training set had an imbalanced number of samples between difficult and simple airways. The number of simple airway samples was approximately 1.75 times higher than that of difficult airway samples. To tackle this issue, we employed data augmentation methods, including random horizontal flipping ($p = 0.5$), random rotation within ±15°, scaling within the range of 0.9–1.1, brightness adjustment by ±20%, and contrast adjustment by ±20%. Across Datasets 1–5, we started from the *original* ultrasound dataset and applied *class-specific* augmentation to expand the training data: the *simple-airway* class was augmented by $\times 7$ and the *difficult-airway* class by $\times 3$ using the same transformation pipeline across all datasets. This procedure yielded a total of 1391 images per extended training set (class-wise counts are reported in Table 1). To clarify the acquisition and augmentation statistics, the raw dataset contained 189 images from the simple-airway category and 108 images from the difficult-airway category. To alleviate class imbalance, we applied a class-dependent augmentation strategy, in which difficult-airway images were augmented by a factor of 7, while simple-airway images were augmented by a factor of 3. Consequently, the effective numbers of samples after augmentation were 756 (108×7) for difficult airways and 567 (189×3) for simple airways. Table 1 reports the dataset partition based on the augmented dataset (i.e., the effective sample sizes used in model development), rather than the raw acquisition counts. We enforced patient-level 5-fold cross-validation (i.e., all images from the same patient were assigned to the same fold), and the test folds contained only original (non-augmented) images. The advantages of data augmentation are as follows: (1) It expanded the dataset, mitigating the issue of limited samples. (2) It enhanced sample diversity, which contributed to improving the model's generalization capability. (3) It balanced the sample distribution between the two classes, addressing the issue of sample imbalance.

Table 1. Training data augmentation and five-fold cross-validation for both test data and training data.

	Test Data		**Training Data**	
	Simple	**Difficult**	**Simple**	**Difficult**
Dataset1	150	128	599	514
Dataset2	150	128	599	514
Dataset3	149	129	600	513
Dataset4	150	128	599	514
Dataset5	149	129	600	513

To reduce the potential data bias introduced by random splitting of image data for training and inference evaluation, we employed 5-fold cross-validation [19], as shown in Figure 2. Specifically, during the construction of these five datasets, we used a stratified k-fold cross-validation strategy to ensure that the distribution of simple and difficult intubation groups in each dataset remained consistent, thereby avoiding bias caused by class imbalance. Importantly, we confirm that the 5-fold cross-validation was strictly performed at the patient level, and no augmented versions of the images from the test fold were included in the training fold. In each fold, we trained the model on the training set and evaluated it on the corresponding validation set, ensuring the stability and generalization ability of the model's performance. Finally, we calculated the AUC (Area Under the Curve) for each dataset as a metric for evaluating the model's classification performance. This approach allowed us to effectively assess the model's performance across different datasets, ensuring the reliability and accuracy of the results.

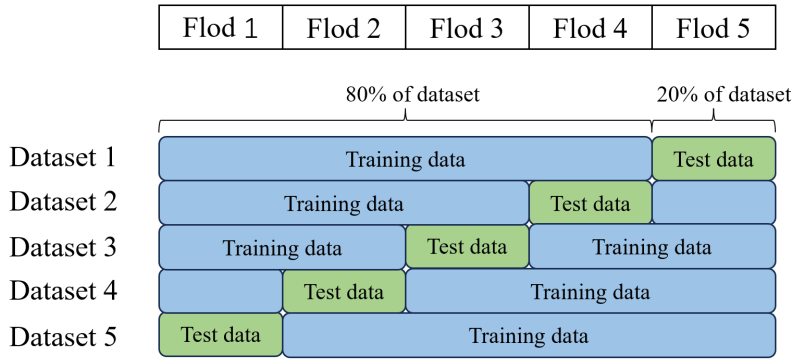


Figure 2. 5-fold cross validation.

2.3. Network Architecture

The AdaDenseNet-LUC network architecture diagram, illustrated in Figure 3, utilizes two deep learning techniques: transfer learning and fine-tuning [20]. Transfer learning improves the accuracy of models by applying pre trained models on large datasets to new tasks. Through this approach, we are able to achieve higher classification accuracy with less image data, as the transferred model has already learned effective features. In this study, we used DenseNet as the base model [21]. In addition, fine-tuning techniques were used to adjust the model based on the Ultrasonic scan images of the surgical patients in this study, to classify the difficulty of intubation. After training, we performed inference evaluation using a pre-segmented image dataset (test set), thereby assessing the model's accuracy in predicting intubation difficulty.

In AdaDenseNet-LUC, two complementary attention modules are integrated to improve feature learning. The SE block recalibrates channel-wise features, highlighting informative channels and reducing redundancy. These refined features are then processed by the LSTM-based adaptive attention module, which assigns dynamic weights to different spatial regions to capture long-range dependencies and global context. The fused representation is finally propagated through the DenseNet backbone for classification. This design allows the model to combine local channel refinement with global contextual modeling, thereby enhancing the accuracy of difficult airway prediction from laryngeal ultrasound images.

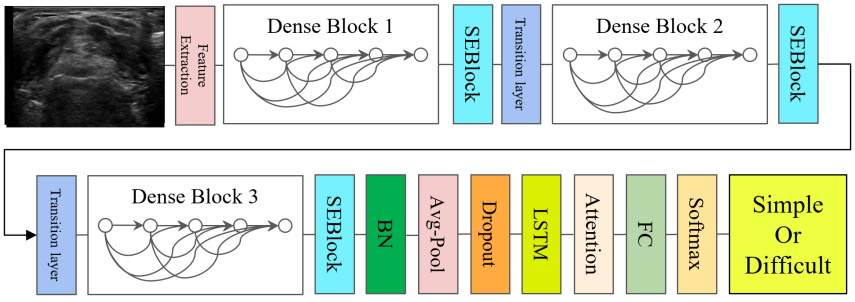


Figure 3. AdaDenseNet-LUC network architecture diagram.

2.4. Experimental Setup and Implementation Details

All experiments were conducted in Python 3.9.19 with PyTorch 2.3.1 under the Anaconda 2.5.2 environment on Windows 11. The hardware platform consisted of an Intel Core i5-12600KF CPU, 32 GB of DDR4 RAM, and an NVIDIA GeForce RTX 4070 Super GPU with 12 GB GDDR6 memory. The model was trained for 70 epochs with a mini-batch size

of 64 using mixed-precision computation to improve training efficiency and reduce GPU memory consumption.

Optimization was performed with Stochastic Gradient Descent (SGD) using a momentum of 0.9 and a weight decay coefficient of 1×10^{-4}. A phased learning rate schedule was applied: the learning rate was initially set to 0.01 for the first 20 epochs, reduced to 0.001 between epochs 20 and 50, and further decreased to 0.0001 after epoch 50 until the completion of training. No early stopping strategy was applied, and the training proceeded for a fixed number of 70 epochs.

3. The Basic Principle of DenseNet

The architectural configuration of DenseNet—a convolutional neural network featuring dense interlayer connectivity—was first introduced by Huang et al. in 2017 [21], with its topological representation detailed in Figure 4.

DenseNet has significant advantages through feature sharing and dense inter layer connections. Its main advantage lies in the ability to reuse feature maps from different layers, thereby reducing inter layer dependencies, and providing diverse input features through quick connections, addressing the issue of gradient vanishing in deep networks. The final prediction results are based on the features from all layers, allowing the model to achieve improved performance and robustness while maintaining a smaller size and lower computational complexity. The architectural design of DenseNet necessitates cross-layer concatenation of feature maps, inducing persistent feature reuse throughout network propagation. This structural characteristic precipitates a linear escalation in parameter volume proportional to network depth progression, ultimately manifesting as exponential amplification of computational complexity and memory consumption during training iterations [22].

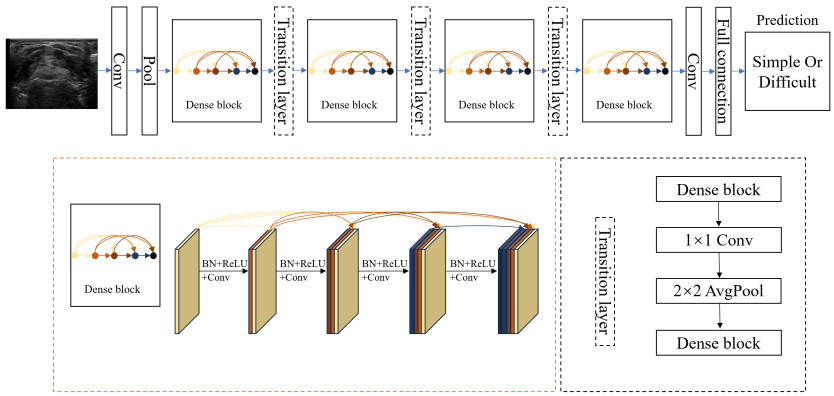


Figure 4. The structure of DenseNet.

The foundational structure of DenseNet includes dense blocks, transition layers, convolutional layers, and fully connected layers. A dense block is composed of multiple dense units, each containing batch normalization (BN), ReLU activation, and convolution operations. These units are tightly connected and employ a pre activation strategy to improve training efficiency and generalization ability. The feature propagation mechanism within dense blocks involves concatenating each unit's input tensor with the complete set of outputs from preceding computational units, thereby generating composite feature maps that undergo iterative transmission through the network's hierarchical architecture. This feature reuse effectively alleviates the gradient vanishing problem and allows for the generation of more features with fewer convolution kernels. The transition layer, situated between adjacent dense blocks, generally includes a 1×1 convolutional layer and a 2×2 average pooling layer. It serves to compress input features, reducing the size and

dimensionality of feature maps, which in turn decreases the number of parameters and helps prevent overfitting. Ultimately, the fully connected layer handles the classification task by combining category feature information from the network and applying weights to make classification predictions [23].

4. Model Design and Implementation

4.1. Squeeze-and-Excitation (SE) Block

The SE Block is a lightweight attention module designed to enhance the effectiveness of convolutional neural networks by adaptively recalibrating the importance of feature channels [24], as shown in Figure 5. It operates through two main steps: Squeeze, which aggregates global spatial information into a channel descriptor using global average pooling, and Excitation, which employs fully connected layers with non-linear activations to generate channel-wise weights. Finally, the recalibration step multiplies the original feature maps by these learned weights, emphasizing informative channels while suppressing less relevant ones.

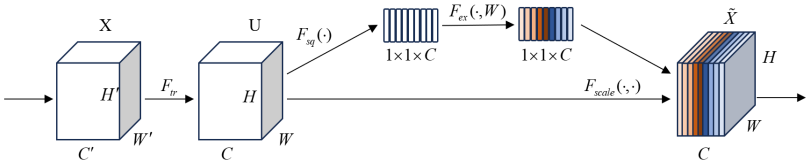


Figure 5. Combine squeeze and excitation steps together.

4.1.1. Squeeze

At this stage, the SE Block reduces the feature maps of each channel into a scalar through Global Average Pooling, as shown in Formula (1): This is achieved by calculating the average of all spatial positions for each channel. Specifically, if the input feature map is X (The size is $H \times W \times C$,Where H is the height, W is the width, and C is the number of channels),Then the Squeeze operation generates a vector z of shape $1 \times 1 \times C$.

$$z_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_{ijc} \tag{1}$$

4.1.2. Excitation

At this stage, the SE Block determines the importance of each channel using a small fully connected neural network, typically composed of two fully connected layers and ReLU activation, as shown in Formula (2). Firstly, the dimension is reduced to C/r through a ReLU activated fully connected layer (where r is the scaling factor),Then, a fully connected layer activated by Sigmoid is used to restore the dimension to C, ultimately generating a channel weight vector s:

$$s = \sigma(W_2 \cdot \mathrm{Re}\,LU(W_1 \cdot z)) \tag{2}$$

4.1.3. Recalibration

Finally, the SE Block recalibrates the feature map by multiplying the feature values of each channel in the initial map X with the corresponding channel weights. As shown in Formula (3):

$$\tilde{X} = X \cdot s \tag{3}$$

In this study, we specifically chose the SE Block for several important reasons. It provides a computationally efficient solution that is particularly suitable for relatively small-scale medical datasets such as ours. Moreover, it has been widely validated in nu-

merous medical imaging tasks, consistently demonstrating robust and stable performance. In addition, its channel recalibration function complements the LSTM-based adaptive attention mechanism integrated in our framework, enabling the model to achieve a balance between local feature refinement and global contextual modeling. Taken together, these advantages allow the SE Block to improve feature expression, robustness, and noise resistance, while ensuring an effective balance between efficiency, accuracy, and interpretability in clinical applications.

4.2. Long Short-Term Memory Adaptive Attention Module

Long Short-Term Memory (LSTM) is a type of recurrent neural network that successfully captures contextual information in long sequences, overcoming the gradient vanishing and explosion problems commonly faced in traditional RNNs [25], as demonstrated in Figure 6. In traditional RNNs, parameters are updated through time backpropagation (BPTT), but as the time step increases, the gradient gradually decreases, making it difficult to effectively update the preceding nodes, ultimately leading to optimization failure. This makes it difficult for subsequent nodes to obtain effective information from previous nodes during the testing phase when the sequence is long, resulting in a "long-term dependency" problem that affects prediction accuracy. To address this issue, LSTM introduces memory units and gating mechanisms. Memory units serve to retain state information, while gating mechanisms control how these states are updated.

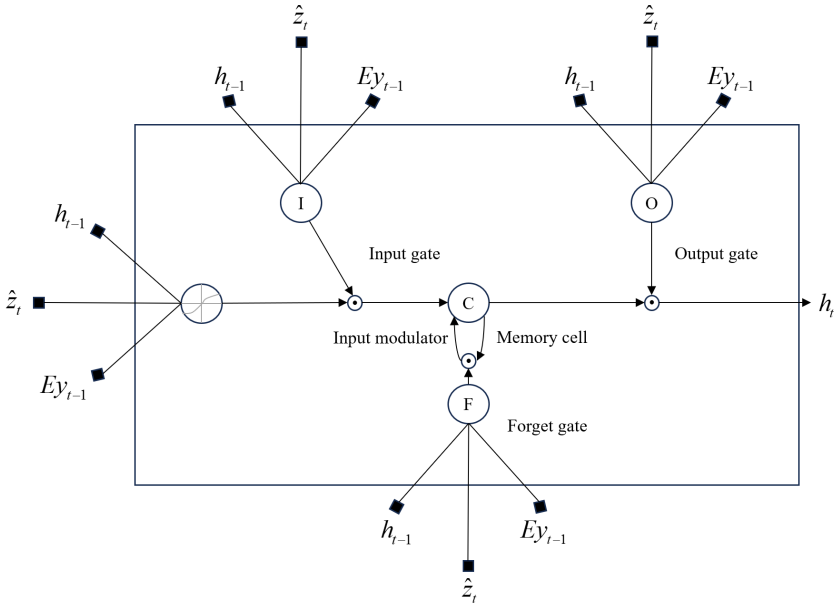


Figure 6. In an LSTM unit, there are three gates: input gate, forget gate, and output gate.

The structure of the LSTM model and the corresponding mathematical formulas for the input gate, output gate, and forget gate are shown in Formulas (4)–(6):

$$\begin{pmatrix} i_t \\ f_t \\ o_t \\ g_t \end{pmatrix} = \begin{pmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{pmatrix} \mathbf{T}_{D+m+n,n} \begin{pmatrix} Ey_{t-1} \\ h_{t-1} \\ \hat{z}_t \end{pmatrix} \tag{4}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \tag{5}$$

$$h_t = o_t \odot \tanh(c_t) \tag{6}$$

Among them, i_t, f_t, c_t, o_t and h_t are, respectively, the input state, forget state, memory state, output state, and hidden state of LSTM. Ey_{t-1} is a lookup table, y_{t-1} is used to obtain m-dimensional word vectors of words, h_{t-1} is the hidden state of the previous moment, and $\tilde{z}_t \in R^D$ represents the actual meaning of the LSTM "input", which represents a context vector for a specific region and can capture visual information. Because of its dependence on time t, it is considered a dynamic variable that captures the relevant image regions corresponding to that specific time at various time points. The initial values for the hidden state and cell state are calculated using two separate multi-layer perceptrons (MLPs), with the input to the perceptrons being the average value of the features from each image region, as shown in Formulas (7) and (8):

$$c_0 = f_{init,c}(\frac{1}{L}\sum_{i=1}^{L} a_i) \tag{7}$$

$$h_0 = f_{init,h}(\frac{1}{L}\sum_{i=1}^{L} a_i) \tag{8}$$

4.3. LSTM Adaptive Attention Mechanism

In our framework, the LSTM does not process temporal sequences but instead models spatial dependencies within the 2D feature maps. Each spatial slice of the feature map (such as a row or a column) is treated as a sequential step, allowing the LSTM to propagate contextual information across anatomical regions. This sequential treatment enables the LSTM to capture dependencies between different spatial regions. The adaptive attention layer then computes dynamic weights based on the hidden states of the LSTM, enabling the network to selectively emphasize diagnostically relevant regions such as the trachea and epiglottis.

In LSTM, the output and state update of each time step mainly depend on the current input and the previous hidden state. Although it can capture some spatial dependencies, it may ignore earlier key information when processing long sequences. To address this limitation, we introduce the adaptive attention mechanism, which assigns attention weights to allow LSTM to focus on different input parts at each time step according to task requirements. This makes more effective use of global information [26]. As shown in Figure 7.

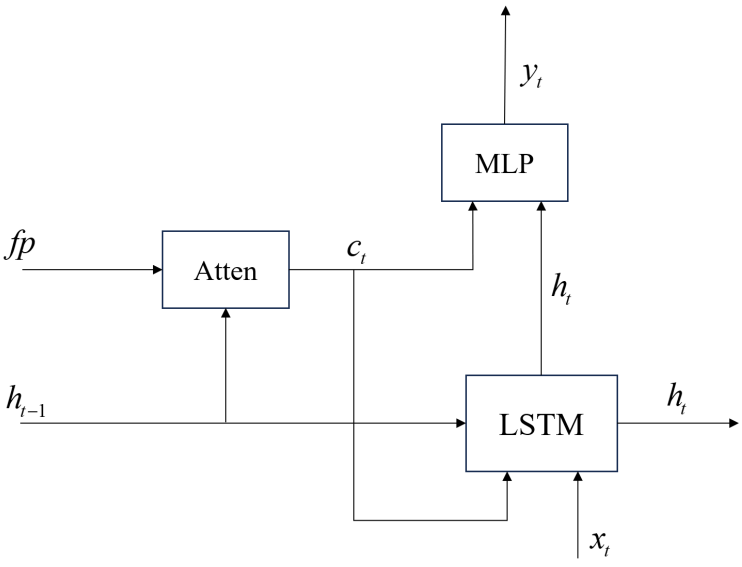


Figure 7. Demonstration of LSTM adaptive attention mechanism.

At step t, we compute attention scores using Formula (9)

$$e_t = w^\top \tanh(W_h h_t + W_a \tilde{a}_t), \quad \alpha_t = \text{softmax}(e_t), \quad z = \sum_{t=1}^{L} \alpha_t \, \tilde{a}_t. \tag{9}$$

The context z summarizes diagnostically salient regions (e.g., trachea/epiglottis). We fuse z with the SE-recalibrated global descriptor by concatenation followed by a small MLP before classification, combining channel-wise refinement (SE) and global spatial context (LSTM-attention).

In a single channel image, the information of each pixel is relatively limited. The LSTM-based adaptive attention mechanism captures long-range spatial dependencies and highlights key diagnostic regions by combining hierarchical features with global context. The adaptive attention mechanism can focus on different regions of input features at different time steps according to the current task requirements, because single channel medical image features are sparse and difficult to extract. The attention mechanism enables the model to concentrate on key regions.

5. Evaluation

In this article, we adopt accuracy, sensitivity, specificity, F_1 score, and Matthews Correlation Coefficient (MCC) as the evaluation criteria. Notably, MCC provides a more balanced assessment of classification performance, especially under class-imbalanced settings. These indicators are calculated by substituting true positive (TP), true negative (TN), false positive (FP), and false negative (FN) into the corresponding equations, as shown in Formulas (10)–(15).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{10}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{11}$$

$$Specificity = \frac{TN}{FP + TN} \tag{12}$$

$$Precision = \frac{TP}{FP + TP} \tag{13}$$

$$F_1 - score = \frac{2 \times Precision \times Sensitivity}{Precision + Sensitivity} \tag{14}$$

$$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}. \tag{15}$$

In addition to the previously mentioned evaluation metrics, we also use the receiver operating characteristic (ROC) curve and the area under the curve (AUC) to evaluate the model's performance. The ROC curve shows the relationship between sensitivity and specificity, while the AUC value represents the area under the curve, typically ranging from 0.5 to 1. Generally, the closer the ROC curve is to the upper left corner of the coordinate system, and the closer the AUC is to 1, the better the model's performance.

6. Experimental Results

Hyperparameters such as learning rate, weight decay, and batch size were determined based on prior studies and empirical exploration. To ensure robustness and reduce variance caused by random data splitting, all results were reported as the mean performance over 5-fold cross-validation. Visualization outputs, including ROC curves and accuracy progression plots for both training and testing sets, were generated as illustrated in Figures 8 and 9.

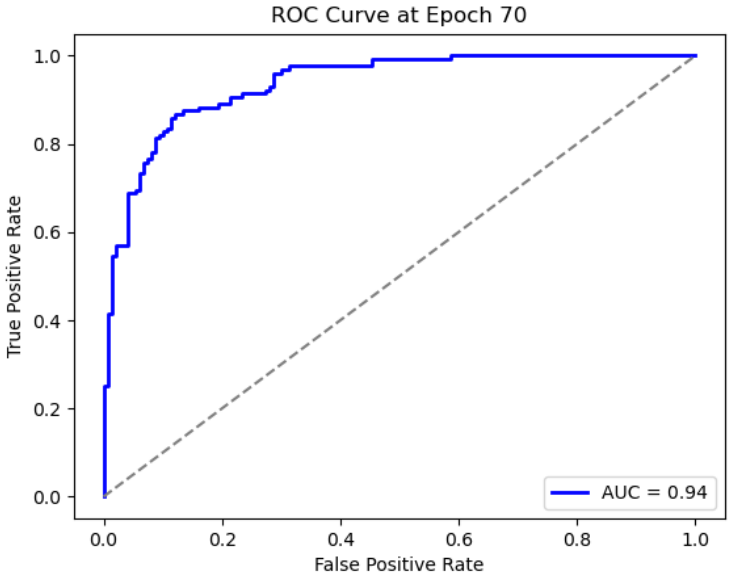


Figure 8. Receiver Operating Characteristic (ROC) Curve of the proposed model at 70 epochs.

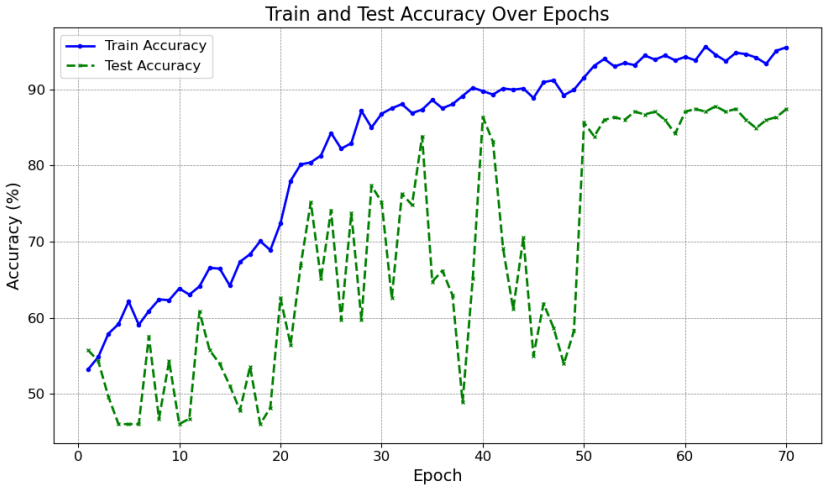


Figure 9. Training and testing accuracy progression across 70 epochs.

In the transfer learning stage, the convolutional layers of the pretrained DenseNet backbone up to the third dense block were frozen to retain generic feature representations, while the subsequent dense block and the LSTM-based attention module were fine-tuned on the target dataset. A lower learning rate of 0.0001 was applied to the unfrozen layers to ensure stable adaptation, whereas the newly added fully connected classification layer was trained with a higher learning rate of 0.001 to accelerate convergence. The LSTM hidden layer was set to a size of 256, allowing the model to capture rich temporal dependencies in the sequential features. The final fully connected classification layer (MLP) had an input size of 256 (matching the LSTM hidden size) and an output size of 2, representing the two classification categories. The activation function used in the fully connected layer was softmax, ensuring probabilistic classification across the two classes. This fine-tuning scheme allowed us to effectively balance the preservation of pretrained knowledge with the adaptation to task-specific features.

The ablation study in Table 2 evaluates the impact of different model components, with all numbers representing the average results from 5-fold cross-validation. The DenseNet-only model achieved an AUC of 0.866 with modest performance. Adding

the SE Block improved the AUC to 0.886, highlighting its contribution to feature refinement. The LSTM attention-only model performed poorly with an AUC of 0.686, showing the importance of combining components. The full AdaDenseNet-LUC model, incorporating DenseNet, SE Block, and LSTM attention, achieved the best performance with an AUC of 0.914, demonstrating the effectiveness of combining all components.

Table 2. Ablation study results.

Model	AUC	Accuracy	Sensitivity	Specificity	F1 Score
DenseNet Only	0.866	0.792	0.782	0.796	0.764
SE Block Only	0.886	0.800	0.788	0.808	0.776
LSTM Attention Only	0.686	0.658	0.638	0.676	0.618
AdaDenseNet-LUC	**0.914**	**0.822**	**0.812**	**0.830**	**0.800**

The ablation study results presented in Table 2 provide valuable insights into the model's performance. The "Attention Only LSTM" model performs significantly worse (ASC 0.686). This performance degradation can be attributed to the fact that the LSTM module, when used in isolation, may not be able to fully capture the spatial dependencies or contextual information necessary for effective classification. In contrast, when integrated with the attention mechanism, the LSTM benefits from dynamic weight allocation, allowing it to focus on the most diagnostically relevant features across different spatial regions. This fusion enhances the model's ability to capture long-range dependencies and emphasize key areas such as the trachea and epiglottis, leading to improved performance.

In our experiments, we evaluated the model at the patient level, extracting one ultrasound frame per specific angle for each patient, ensuring that the data used for predictions represents the most relevant anatomical information. Therefore, the results are reported at the patient level, with each patient having a single prediction. In this study, DenseNet is used as the core architecture for deep feature extraction. To highlight the advantages of using the improved DenseNet for deep feature extraction in this task, we compare it with Vgg16 [27], ResNet [28], AlexNet [29], MobileNet [30], EfficientNet [31], EfficientNetV2 [32], DenseNet [21],Transformer [33] and Mamba [34]. Specifically, transfer learning methods are applied to transfer each network pre-trained on ImageNet to our dataset for retraining, with all networks trained using the same parameter configuration. The results from the 5-fold cross-validation experiment are presented in Tables 3 and 4. From Table 4, it is evident that the AUC of our model surpasses that of other models, reaching 0.888.

Therefore, the experimental results show that the improved DenseNet exhibits better deep feature extraction ability compared to other common network architectures in this task [35–37], especially in terms of AUC value. This demonstrates the effectiveness of DenseNet's efficient feature transfer and deeper feature learning in solving this problem. In addition, the improved model in this article not only has high accuracy, but also demonstrates better stability and convergence during the training process, further verifying its potential application in the field of medical imaging. Therefore, we believe that the model based on improved DenseNet has strong competitiveness and practicality in tracheal Ultrasound image classification tasks. Future work will aim to further optimize the network architecture and expand the datasets to enhance the performance and generalization capability of models in real-world applications.

Table 3. The model prediction performance of five-fold cross-validation.

Dataset	AUC	Accuracy	Sensitivity	Specificity	F_1 Score	MCC
Dataset1	0.91	0.80	0.80	0.81	0.78	0.60
Dataset2	0.90	0.82	0.83	0.82	0.80	0.65
Dataset3	0.94	0.87	0.86	0.89	0.86	0.75
Dataset4	0.93	0.83	0.80	0.86	0.82	0.66
Dataset5	0.89	0.79	0.77	0.77	0.74	0.58

Table 4. Performance comparison of different models on the dataset, reporting the average results from five-fold cross-validation.

Model	Dataset1	Dataset2	Dataset3	Dataset4	Dataset5	Average
Vgg16	0.75	0.90	0.57	0.94	0.94	0.825
ResNet	0.95	0.90	0.78	0.95	0.92	0.880
AlexNet	0.70	0.85	0.68	0.78	0.92	0.793
MobileNet	0.90	0.87	0.84	0.85	0.93	0.885
EfficientNet	0.92	0.69	0.57	0.39	0.82	0.678
EfficientNetV2	0.89	0.50	0.46	0.47	0.68	0.604
DenseNet	0.88	0.85	0.76	0.86	0.76	0.822
Transformer	0.88	0.85	0.93	0.86	0.88	0.886
Mamba	0.90	0.89	0.78	0.92	0.77	0.857
AdaDenseNet-LUC	**0.91**	**0.90**	**0.94**	**0.93**	**0.89**	**0.914**

Compare the accuracy between existing prediction indicators and the best artificial intelligence model in the field of difficult airway prediction. The data in Tables 5 and 6 clearly shows that the method proposed in this study outperforms in multiple indicators. In terms of sensitivity, the approach introduced in this article successfully attained 86.0%, which has significant advantages compared to other indicators such as incisor spacing (44.4%) and protruding teeth (24.1%). This suggests that the approach introduced in this article demonstrates greater sensitivity in detecting potential difficult intubation scenarios. In terms of specificity, this article method performed exceptionally well, reaching 88.6%, far exceeding other indicators such as Mallampati grading (52.7%), though slightly lower than the Teeth Condition method (89.9%), indicating that the laryngeal ultrasound image classification can more accurately exclude non difficult intubation situations and reduce misjudgments. In terms of AUC, a comprehensive measure of accuracy and discrimination, the method proposed in this paper achieved a high value of 0.940, slightly better than facial image classification (0.864), and significantly higher than indicators such as sternotomy distance (0.587). This clearly illustrates that the method presented in this paper offers higher accuracy and reliability in predicting difficult intubation, thereby providing a more valuable predictive tool for clinical practice. We assess statistical significance using paired two-sided t-tests on fold-wise AUROCs computed on identical splits. For each model, we report the mean AUROC with a 95% t-interval across 5 folds, and we test the per-fold AUC difference (Model − AdaDenseNet-LUC) to obtain the t statistic and p-value. As summarized in Table 5, AdaDenseNet-LUC shows significant improvements over DenseNet ($t = 3.328$, $p = 0.029$), EfficientNetV2 ($t = 3.580$, $p = 0.023$), and Transformer ($t = 2.915$, $p = 0.043$), while differences against other baselines are not statistically significant ($p \geq 0.05$). Given $n = 5$, we also report Cohen's d and 95% CIs to quantify effect sizes and uncertainty.

Table 5. Model performance comparison with AUROC, sensitivity, and specificity across different architectures.

Model	Avg AUC	AUC CI (95%)	Avg Sen	Sen CI (95%)	Avg Spe	Spe CI (95%)
DenseNet	0.822	(0.751, 0.893)	0.782	(0.752, 0.811)	0.788	(0.749, 0.826)
EfficientNet	0.678	(0.419, 0.936)	0.636	(0.448, 0.823)	0.664	(0.410, 0.917)
EfficientNetV2	0.600	(0.370, 0.829)	0.544	(0.332, 0.755)	0.624	(0.439, 0.808)
MobileNet	0.878	(0.832, 0.920)	0.792	(0.752, 0.831)	0.804	(0.770, 0.837)
AlexNet	0.786	(0.661, 0.913)	0.716	(0.556, 0.875)	0.748	(0.603, 0.892)
ResNet	0.900	(0.812, 0.987)	0.800	(0.753, 0.847)	0.814	(0.763, 0.864)
Vgg16	0.820	(0.621, 1.019)	0.778	(0.623, 0.933)	0.816	(0.673, 0.958)
Transformer	0.886	(0.841, 0.918)	0.782	(0.746, 0.817)	0.805	(0.767, 0.840)
Mamba	0.857	(0.763, 0.940)	0.786	(0.757, 0.816)	0.801	(0.751, 0.848)
AdaDenseNet-LUC	**0.914**	**(0.888, 0.939)**	**0.812**	**(0.769, 0.854)**	**0.830**	**(0.772, 0.887)**

Table 6. Statistical comparison of AUC values across models using five-fold cross-validation.

Model	T	P
DenseNet	3.328	0.029
EfficientNet	2.360	0.077
EfficientNetV2	3.580	0.023
MobileNet	1.438	0.223
AlexNet	2.425	0.072
ResNet	0.377	0.725
Vgg16	1.208	0.293
Transformer	2.915	0.043
Mamba	1.909	0.128
AdaDenseNet-LUC	**–**	**–**

Figure 10 presents the classification outcomes for two representative examples from the test set, with the proposed model predicting the class for each instance. Grad-CAM visualization techniques indicate that the model's attention is mainly focused on the tracheal region. However, we acknowledge that structures such as the epiglottis, tongue, and hyoid bone play significant roles in airway assessment, particularly when evaluating the potential for obstruction or difficulty in intubation. While the trachea's visibility is essential for identifying difficult airways, these additional anatomical features must also be considered in clinical evaluations. Grad-CAM works by visualizing which regions of the image contribute most to the model's predictions. In the case of simple airways, Grad-CAM typically highlights clearer, more defined anatomical structures, such as the central airway passages, which are easier for the model to recognize. In contrast, for difficult airways, Grad-CAM focuses more on complex features, such as irregularities, narrow passages, or areas with occlusions, which are more challenging for the model to differentiate. These differences in the Grad-CAM visualizations reflect the model's sensitivity to variations in airway structures, helping us understand how the model distinguishes between simple and difficult airways based on anatomical features. Future work could expand the model to incorporate these additional structures more effectively, thus improving the accuracy and clinical utility of the approach in predicting difficult intubation scenarios.

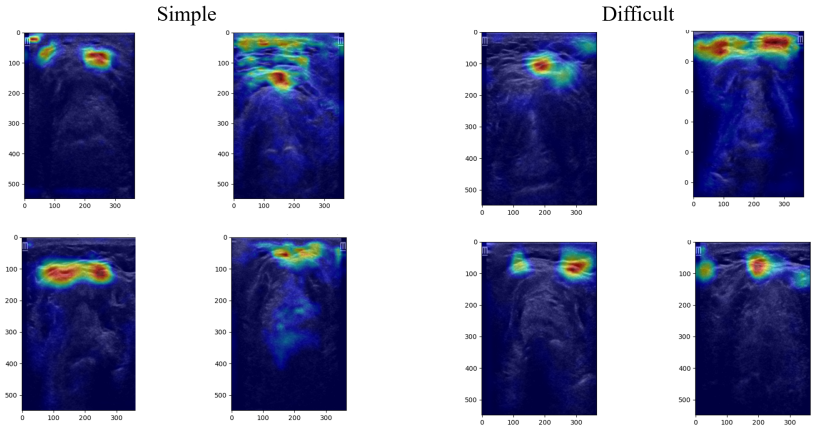


Figure 10. Grad CAM visualization of two types of samples.

7. Discussion

In this study, we propose AdaDenseNet-LUC for difficult airway classification based on laryngeal ultrasound images. Experiments on 1391 images with 5-fold cross-validation show that the proposed method achieves 87.41% accuracy, 86.05% sensitivity, 88.59% specificity, and an AUC of 0.94, indicating strong discrimination while keeping a favorable balance between missed difficult cases and false alarms.

Consistently, fold-wise comparisons show AdaDenseNet-LUC achieves the highest mean AUROC (0.914) and yields statistically significant improvements over several strong baselines (e.g., DenseNet and Transformer) under identical splits.

In our experiments, all models were trained and evaluated on the same dataset under identical preprocessing and parameter settings to ensure fairness. Table 7 presents the best-performing results, which is consistent with previous studies in the literature where the best values are typically highlighted. To ensure comparability with prior work, our study also reports the best results. This reporting strategy helps to align our findings with those of existing research.

Table 7. Performance of the proposed method compared with baseline approaches, reporting the best-performing results for consistency with prior studies.

Indicator	Sensitivity (%)	Specificity (%)	AUC
Mallampati Classification (MPC) (1/2/3/4)	79.6	52.7	0.673
Inter-Incisor Gap (IIG) (cm)	44.4	75.0	0.633
Head and Neck Mobility (HNM)	53.7	77.7	0.670
Thyromental Distance (TMD) (cm)	53.7	58.1	0.587
Horizontal Length of Mandible (HLM) (cm)	48.1	64.9	0.558
Teeth Condition (BT) (Normal/Mild/Severe)	24.1	89.9	0.572
Upper Lip Bite Test (ULBT) (1/2/3)	48.1	70.3	0.607
Facial Image Classification	81.8	83.3	0.864
Laryngeal Ultrasound Image Classification	**86.0**	**88.6**	**0.940**

The clinical predictor values shown in Table 7 are derived from the literature and are presented for contextual reference. They do not represent a head-to-head comparison on the present cohort. It is important to interpret the comparison in Table 7 with caution, as the literature-derived clinical predictor metrics reflect populations and protocols that differ from those of the present study. Such heterogeneity may bias the absolute differences in performance.

Interpretability results from Grad-CAM indicate that the model mainly attends to the tracheal region, which aligns with clinical intuition for airway assessment. Nevertheless, this study is limited by its single-center cohort (300 patients) and the lack of explicit non-imaging clinical variables (e.g., obesity, snoring, smoking status). Future work will focus on multicenter external validation and multimodal integration to better reflect real-world decision-making.

8. Limitations and Future Work

This study has several limitations. First, it is a single-center study based on 300 patients, which may introduce sampling bias and limit the generalizability of the findings. In addition, the incidence of difficult airway cases may vary regionally. For example, in parts of southern China, the habit of chewing betel nuts is associated with a higher prevalence of oral cancer, which could increase the likelihood of difficult airway cases; meanwhile, smoking—an important risk factor for airway difficulties [38]—may further exacerbate airway problems in some patients. Second, compared with certain AI studies based on magnetic resonance imaging (MRI), the number of cases and images in the present work is relatively small. Finally, in real-world clinical practice, experienced anesthesiologists consider not only visual indicators but also clinical expertise and patient history when assessing airway difficulty. Our current model does not explicitly incorporate such non-imaging clinical variables, which represents a limitation.

To address these limitations, we plan to conduct multicenter collaborations across different regions to build a larger, more diverse database and collect additional clinical images, thereby improving external validity. We have also begun integrating the proposed model into a computer-aided application to support efficient airway assessment by anesthesiologists. In parallel, we will incorporate key clinical variables (e.g., obesity, snoring, smoking status, and relevant comorbidities) into the system to enhance diagnostic accuracy and better reflect real-world decision-making.

9. Conclusions

This article proposes AdaDenseNet-LUC, a novel deep feature extraction framework that combines transfer learning and adaptive attention mechanisms for automatic classification of difficult airways. We collected and constructed a dataset containing 1391 throat Ultrasonic scan images for training and evaluating the model. The integration of an LSTM-based adaptive attention mechanism into DenseNet effectively enhances the extraction of deep features tailored for throat ultrasound image classification. At the same time, using transfer learning to accelerate model training and alleviate the problem of small datasets has improved classification performance. The experimental results showed that after 5-fold cross validation, the classification accuracy of this method on the dataset was 87.41%, sensitivity was 86.05%, specificity was 88.59%, F_1 score was 0.8638, and AUC value reached 0.94. In the future, combined with more data from throat Ultrasonic images, artificial intelligence models are expected to help predict difficulties in intubation, thereby timely requesting experienced doctors' assistance, reducing visual field loss caused by improper intubation, and saving patients' lives.

Author Contributions: Conceptualization, C.L. and H.L.; methodology, C.L.; software, C.L.; validation, C.L. and H.L.; formal analysis, C.L.; investigation, C.L.; resources, H.L.; data curation, C.L.; writing—original draft preparation, C.L.; writing—review and editing, C.L. and H.L.; visualization, C.L.; supervision, H.L.; project administration, H.L.; funding acquisition, H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, Grant Number 62273189.

Institutional Review Board Statement: This study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of The Affiliated Hospital of Qingdao University.

Informed Consent Statement: Patient consent was waived by the Institutional Review Board due to the retrospective nature of this study and anonymization of the data.

Data Availability Statement: The clinical data used in this study are not publicly available due to patient privacy and institutional ethics regulations. De-identified data may be available from the corresponding author on reasonable request and with approval from the institutional review board (IRB).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Apfelbaum, J.L.; Hagberg, C.A.; Caplan, R.A.; Blitt, C.D.; Connis, R.T.; Nickinovich, D.G.; Benumof, J.L.; Berry, F.A.; Bode, R.H.; Cheney, F.W.; et al. Practice guidelines for management of the difficult airway: An updated report by the American Society of Anesthesiologists Task Force on Management of the Difficult Airway. *Anesthesiology* **2013**, *118*, 251–270. [PubMed]
2. Nørskov, A.K.; Rosenstock, C.V.; Wetterslev, J.; Astrup, G.; Afshari, A.; Lundstrøm, L.H. Diagnostic accuracy of anaesthesiologists' prediction of difficult airway management in daily clinical practice: A cohort study of 188 064 patients registered in the Danish Anaesthesia Database. *Anaesthesia* **2015**, *70*, 272–281. [CrossRef]
3. Langeron, O.; Masso, E.; Huraux, C.; Guggiari, M.; Bianchi, A.; Coriat, P.; Riou, B. Prediction of difficult mask ventilation. *Anesthesiology* **2000**, *92*, 1229–1236. [CrossRef]
4. Levitan, R.M.; Heitz, J.W.; Sweeney, M.; Cooper, R.M. The complexities of tracheal intubation with direct laryngoscopy and alternative intubation devices. *Ann. Emerg. Med.* **2011**, *57*, 240–247. [CrossRef] [PubMed]
5. Cook, T.M. Major complications of airway management in the UK: Results of the Fourth National Audit Project of the Royal College of Anaesthetists and the Difficult Airway Society. Part 1: Anaesthesia. *Br. J. Anaesth.* **2011**, *106*, 617–631. [CrossRef]
6. Cook, T.M.; MacDougall-Davis, S.R. Complications and failure of airway management. *Br. J. Anaesth.* **2012**, *109* , i68–i85. [CrossRef] [PubMed]
7. Heidegger, T. Management of the difficult airway. *N. Engl. J. Med.* **2021**, *384*, 1836–1847. [CrossRef]
8. Hagiwara, Y.; Watase, H.; Okamoto, H.; Goto, T.; Hasegawa, K. Japanese Emergency Medicine Network Investigators. Prospective validation of the modified LEMON criteria to predict difficult intubation in the ED. *Am. J. Emerg. Med.* **2015**, *33*, 1492–1496. [CrossRef]
9. Nørskov, A.K.; Wetterslev, J.; Rosenstock, C.V.; Afshari, A.; Astrup, G.; Jakobsen, J.C.; Thomsen, J.L.; Bøttger, M.; Ellekvist, M.; Schousboe, B.M.B.; et al. Effects of using the simplified airway risk index vs usual airway assessment on unanticipated difficult tracheal intubation: A cluster randomized trial with 64,273 participants. *BJA Br. J. Anaesth.* **2016**, *116*, 680–689. [CrossRef]
10. Vinisha, A.; Boda, R. DeepBrainTumorNet: An effective framework of heuristic-aided brain tumour detection and classification system using residual Attention-Multiscale Dilated inception network. *Biomed. Signal Process. Control* **2025**, *100*, 107180. [CrossRef]
11. Lu, M.Y.; Chen, T.Y.; Williamson, D.F.K.; Zhao, M.; Shady, M.; Lipkova, J.; Mahmood, F. AI-based pathology predicts origins for cancers of unknown primary. *Nature* **2021**, *594*, 106–110. [CrossRef] [PubMed]
12. Murugesan, M.; Kaliannan, K.; Balraj, S.; Singaram, K.; Kaliannan, T.; Albert, J.R. A hybrid deep learning model for effective segmentation and classification of lung nodules from CT images. *J. Intell. Fuzzy Syst.* **2022**, *42*, 2667–2679. [CrossRef]
13. Han, Y.; Chen, W.; Heidari, A.A.; Chen, H.; Zhang, X. A solution to the stagnation of multi-verse optimization: An efficient method for breast cancer pathologic images segmentation. *Biomed. Signal Process. Control* **2023**, *86*, 105208. [CrossRef]
14. Huang, S.; Li, J.; Xiao, Y.; Shen, N.; Xu, T. RTNet: Relation transformer network for diabetic retinopathy multi-lesion segmentation. *IEEE Trans. Med. Imaging* **2022**, *41*, 1596–1607. [CrossRef] [PubMed]
15. Murray, N.M.; Unberath, M.; Hager, G.D.; Hui, F.K. Artificial intelligence to diagnose ischemic stroke and identify large vessel occlusions: A systematic review. *J. Neurointerv. Surg.* **2020**, *12*, 156–164. [CrossRef]
16. Zhang, F.; Xu, Y.; Zhou, Z.; Zhang, H.; Yang, K. Critical element prediction of tracheal intubation difficulty: Automatic Mallampati classification by jointly using handcrafted and attention-based deep features. *Comput. Biol. Med.* **2022**, *150*, 106182. [CrossRef]
17. Hayasaka, T.; Kawano, K.; Kurihara, K.; Suzuki, H.; Nakane, M.; Kawamae, K. Creation of an artificial intelligence model for intubation difficulty classification by deep learning (convolutional neural network) using face images: An observational study. *J. Intensive Care* **2021**, *9*, 38 . [CrossRef]

18. Wang, G.; Li, C.; Tang, F.; Wang, Y.; Wu, S.; Zhi, H.; Zhang, F.; Wang, M.; Zhang, J. A fully-automatic semi-supervised deep learning model for difficult airway assessment. *Heliyon* **2023**, *9*, e15629.
19. Wong, T.-T.; Yeh, P.-Y. Reliable accuracy estimates from k-fold cross validation. *IEEE Trans. Knowl. Data Eng.* **2019**, *32*, 1586–1594. [CrossRef]
20. Lu, J.; Behbood, V.; Hao, P.; Zuo, H.; Xue, S.; Zhang, G. Transfer learning using computational intelligence: A survey. *Knowl.-Based Syst.* **2015**, *80*, 14–23. [CrossRef]
21. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
22. Zhou, T.; Ye, X.; Lu, H.; Zheng, X.; Qiu, S.; Liu, Y. Dense convolutional network and its application in medical image analysis. *BioMed Res. Int.* **2022**, *2022*, 2384830. [CrossRef]
23. Li, Z.; Sun, N.; Gao, H.; Qin, N.; Li, Z. Adaptive subtraction based on U-Net for removing seismic multiples. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 9796–9812. [CrossRef]
24. Li, X.; Wu, J.; Lin, Z.; Liu, H.; Zha, H. Recurrent squeeze-and-excitation context aggregation net for single image deraining. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 254–269.
25. Zou, L.; Xia, L.; Ding, Z.; Song, J.; Liu, W.; Yin, D. Reinforcement learning to optimize long-term user engagement in recommender systems. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, 4–8 August 2019; pp. 2810–2818.
26. Deng, Z.; Jiang, Z.; Lan, R.; Huang, W.; Luo, X. Image captioning using DenseNet network and adaptive attention. *Signal Process. Image Commun.* **2020**, *85*, 115836. [CrossRef]
27. Simonyan, K. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
28. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
29. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1097–1105. [CrossRef]
30. Howard, A.G. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861. [CrossRef]
31. Tan, M.; Le, Q. EfficientNet: Rethinking model scaling for convolutional neural networks. In Proceedings of the International Conference on Machine Learning, Long Beach, CA, USA, 9–15 June 2019; pp. 6105–6114.
32. Tan, M.; Le, Q. EfficientNetV2: Smaller models and faster training. In Proceedings of the International Conference on Machine Learning, Virtual, 18–24 July 2021; pp. 10096–10106.
33. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 5998–6008.
34. Gu, A.; Dao, T.; Ermon, S.; Rudra, A.; Ré, C. Mamba: Linear-Time Sequence Modeling with Selective State Spaces. *arXiv* **2023**, arXiv:2312.00752. [CrossRef]
35. Seo, S.-H.; Lee, J.-G.; Yu, S.-B.; Kim, D.-S.; Ryu, S.-J.; Kim, K.-H. Predictors of difficult intubation defined by the intubation difficulty scale (IDS): Predictive value of 7 airway assessment factors. *Korean J. Anesthesiol.* **2012**, *63*, 491. [CrossRef]
36. Eberhart, L.H.J.; Arndt, C.; Cierpka, T.; Schwanekamp, J.; Wulf, H.; Putzke, C. The reliability and validity of the upper lip bite test compared with the Mallampati classification to predict difficult laryngoscopy: An external prospective evaluation. *Anesth. Analg.* **2005**, *101*, 284–289. [CrossRef]
37. Safavi, M.; Honarmand, A.; Zare, N. A comparison of the ratio of patient's height to thyromental distance with the modified Mallampati and the upper lip bite test in predicting difficult laryngoscopy. *Saudi J. Anaesth.* **2011**, *5*, 258–263. [CrossRef] [PubMed]
38. Warnakulasuriya, S.; Chen, T.H.H. Areca nut and oral cancer: Evidence from studies conducted in humans. *J. Dent. Res.* **2022**, *101*, 1139–1146. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.