International Journal of Computer Information Systems and Industrial Management Applications.
ISSN 2150-7988 Volume 15 (2023) pp. 257-265

Received: 13 Nov. 2022; Accepted: 7 May, 2023; Published: 9 June, 2023

Automatic Diagnosis Chest-X-Ray-Based-Framework for Semantic Segmentation fnd Placement Errors Detection of Catheters and Tubes

Abdelfettah Elaanba[1], Mohammed Ridouani[2] and Larbi Hassouni[3]

[1] Hassan II University, RITM Laboratory, CED Engineering Sciences,
Casablanca, Morocco
Abdelfettah.elaanba@gmail.com

[2] Hassan II University, RITM Laboratory, CED Engineering Sciences,
Casablanca, Morocco
mohammed.ridouani@gmail.com

[3] Hassan II University, RITM Laboratory, CED Engineering Sciences,
Casablanca, Morocco
lhassouni@hotmail.com

***Abstract*: In daily healthcare work routines, automatic diagnosis systems are essential. Human errors are very likely when working in those dangerous environments with a heavy workload and stress. One of the medical procedures where mistakes are risky and can result in severe complications if not caught in time is the task of positioning tubes and catheters. A type of tube is inserted for a patient as part of the tube placement procedure. The position of the installed tube is then determined by screening the patient. Waiting for a radiologist to confirm the diagnosis will delay the tube adjustment. Indeed, more complications may result from the tube adjustment delay or any potential diagnostic mistakes. Through this work, we propose a framework for diagnosis and validation for in-time tube placement error detection. The framework analyzes the chest X-ray right after the tube is inserted and generates a segmentation mask along with classification values for possible errors. Our proposed framework is founded on a customized U-net model that provides competitive segmentation results (dice coefficient of 94,5%) compared to the original U-Net model version. Moreover, the proposed framework is optimized to support deployment on production-edge mobile devices with 75% fewer training parameters.**

***Keywords*:** Semantic segmentation, U-Net, Endotracheal tube, Central venous catheter, Nasogastric tube, Chest X-ray image, Convolutional networks, Deep learning.

I. Introduction

Diagnosis errors are frequent in healthcare (Diagnosis is wrong 10-15% of the time) [1]. The tubes and catheters positioning task are not an exception. Research revealed that 10-20 % of inserted airway tubes (ETT and ET) are mispositioned [2]. Furthermore, 3 % of NG tubes are positioned abnormally within 40% of those errors causing grave complications [3-5]. There are multiple causes for this high number of complications that happened while or after installing a tube for the patient. First, patients needing these devices are in critical health situations or unconscious. So, the doctor gets no feedback from the patient helping him to adjust the device during the tube placement operation. Therefore, if an error occurs is not detected during the process. The second reason is that complications can be developed if a mispositioned tube is not adjusted immediately after the placement operation. This problem is due to delay caused by retarding the post-operative diagnosis. The post-placement diagnosis occurred by performing a Chest-X-Ray screening after the tube placement. Then the X-Ray image must be examined immediately by a radiologist who decides if the tube needs an adjustment. Delays or errors during this operation can cause serious complications. Another reason for this high number of complications is the low number of radiologists examining the post-operative X-Rays for a fast tube adjustment in case of errors. The diagnosis is suspended until a doctor is available, causing severe complications. An example of a serious problem is lung perforation [6]. The lung perforation is due to placing the ETT (Endotracheal tube) tube in one lung instead of the correct position between the two lungs (7cm above the carina [7]). Therefore, the other lung is still not ventilated, which causes lung perforation. Apart from these three reasons, there is always room for errors caused by human nature. A radiologist can misdiagnose a Chest-X-Ray for an inserted tube. So, an abnormally placed tube can be classified as normal, and the patient's case will get worst. Those types of errors can be avoided by using an automatic diagnosis system to help the experts to validate their diagnosis. We propose through this work a framework for automatic diagnosis verification. The framework can be used by a radiologist to get a prediction for possible errors within the mask images for those tubes. The final diagnosis is the

radiologist opinion based on the framework results. The proposed framework is designed to use in work environments and can be deployed on mobile devices with minimum performance. We demonstrated that we could reduce the number of parameters by 50%. We still get highly accurate results (approximately the same result as the full parameter model -1%) with an inference time of 30% less. The proposed framework could be embedded into hospital portable devices to help doctors' diagnosing accurately the patients' X-Ray or to notify the doctors to adjust the tube for a suspect error. This paper is arranged as follows: Section II presents some background in Deep learning segmentation methods. In Section III, we introduce the Tubes and Catheters abnormal positioning segmentation-related works. In Section IV, we introduce the used Dataset. Section V introduces the methodology and presents the proposed Tube diagnosis framework. Section VI presents the results and the discussion. Finally, in section VII, we introduce the conclusion.

II. Background

In this section, we present the existing image segmentation approaches and their applications on medical images. Then we examine and compare the use of the traditional segmentation methods with deep learning-oriented segmentation methods. At the end of this section, we give an overview of the U-net model.

A. Medical images segmentation

Before diving into the medical image segmentation subject, we begin by giving a general definition of image segmentation. Image segmentation is the operation of partitioning an image into non-intersecting regions. Each region is homogenous. The goal is to detect an object in the image or simply remove the complexity of the image and make it easy to analyze. This process applies to different digitalized image types; light-intensity images, magnetic images (MRI), or X-ray images which is the subject of this study. Now move on to be specific and develop the medical image segmentation. Medical image segmentation consists of the extraction or isolation of interest regions from medical images such as MRI, CT scans, or X-ray images. The isolated area can contain bones, and medical devices, on any desired image part to focus. The segmentation helps to eliminate unnecessary details such as soft tissue, air, or boundaries from the image. There are two families of segmentation: traditional methods and deep learning-based methods. The first method uses techniques like edge detection and mathematical filters [8]. The second method uses the learning from annotated dataset Image-mask to train convolutional neural architectures like U-net, Mask R-CNN, or Feature Pyramid Network (FPN) [9]. The trained networks are saved and then used for inference on new patients' Chest-X-Ray images.

B. No deep learning segmentation methods

Before the introduction of deep learning in medical image segmentation, classic segmentation technics were used. Classic segmentation focuses on the processed image and doesn't require previous training on annotated images. Therefore, the approach is limited when images contain complex patterns, like in the case of medical images. An example of known classic segmentation technics is edge-based, region-based, and thresholding-based segmentations. One old, simple, and popular segmentation method is gray-level thresholding. The gray level segmentation consists of subdividing the image based on the variation of the gray level intensity (gray histogram). There are two known gray threshold families: global and local threshold segmentation. The global threshold is done by setting only one threshold value for the whole image [10]. Otherwise, the local threshold segmentation is performed by splitting the image into regions and setting a specific threshold for each image region [11-13]. Another segmentation method to mention is the edge-based approach. This technique aims to look at the pixel level point in which the gray level is burnt by examining neighbor points [14-16]. We exemplify only these two methods. However, there are plenty of ways for no deep learning image segmentation available in the literature. The choice of one technic or another depends on the presence or absence of edges, noise, and the nature of the image histogram. An example of the application of the classic segmentation method is threshold-based leg segmentation in chest x-ray images [17] and a CAD system for Lung Nodule Detection proposed by [18].

C. Deep learning medical image segmentation methods

One of the most known deep learning segmentation methods for medical images is U-Net. U-Net is a convolutional neural network introduced recently by Ronneberger [19]. Later several modified versions were developed to be accurate or to reduce train and inference time. A reduced U-Net version was proposed by Matuszewski [20]. This reduced U-Net architecture is designed to recognize viruses from electron microscopy images. The deep learning methods consist of training a deep convolutional network like U-Net and then using this model for inference to segment new images related to the images into the initial training database. The advantage of this method is that it permits the extraction of complex shapes from images independently to their position or dimension into the inference image.

D. U-Net Architecture

The U-Net network is a convolutional neural network first designed for semantic segmentation (known as pixel-level segmentation) on biomedical images. The U-Net is a CNN encoder-decoder model with skipped and concatenated connections between equivalent levels from the encoder to the decoder. This symmetric connection between the same level layers from the encoder to the decoder gives the model the shape of U(U-Net). The convolutional layers are used to search features, and the skipped connections are used to localize them and put them in the right position in the segmented image. The architecture of U-Net is a succession of repeated blocs (convolution or deconvolution bloc). In the encoder: each bloc is formed by a convolution operation followed by the Relu activation function and Maxpooling for downsampling the input. For the decoder blocs, each bloc is formed by a deconvolution operation followed by the Relu activation function and Maxpooling for upsampling (see Fig.1). Sometimes during training to help the model better

generalize, a dropout layer is placed after the Relu layer. This helps the model avoid overfitting. In the original U-net version, the output layer is (1x1) convolution with a sigmoid activation function. The sigmoid function generates the final mask using a pixel-level classification. The output mask (segmentation result) has the exact dimensions as the input image.

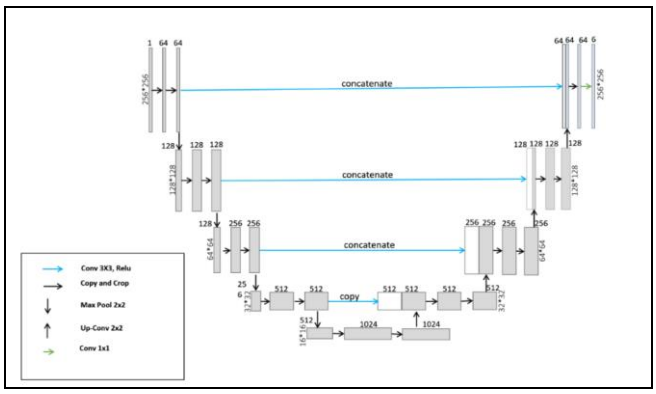

Figure 1. U-NET-64 Architecture.

The power of U-Net is the ability to learn from a few samples which are suitable for the medical field. Indeed, the cost of labeling medical images is very height. The U-net is considered an improvement of its precursor Network FCN for semantic image segmentation developed by Jonathan Long et al. (2015) [21].

III. Related works

In this section, we give an overview of related works to tube segmentation based on chest x-ray images for adult and infant patients. Moreover, we scroll the related work to the models' performance optimization and the number of parameter reductions.

A. Tubes segmentation based on adult CXR images

Medical segmentation is a hot research subject. Many recent works targets CXR-images segmentation; In one related work, Xiaoyan Wang [22] proposes a method for peripherally inserted central catheter (PICC) extraction and tips positioning predictions using Chest-X-ray images. The proposed method (MAG-Net) (Automatic multi-stage attention-guided framework) consists of two stages. The first stage is the coarse stage which the training is performed on the low-resolution image to identify the region that contains the PICC tube. Then, the desired area is extracted and replaced by a high-resolution part from the original image in the training dataset. The new image is used as an input to train the next stage model (fine tuning) to extract the tube. This approach permits the reduction of the training computing costs for the second stage model. Instead of training on the entire high-resolution image, a small area of the original image is used (which contains the tube). The MAG-Net framework is adapted to extract very thin tubes such as PICC from high-resolution CXR images.

B. Deep learning methods for segmentation of lines in pediatric chest radiographs

Another study by Ryan Sullivan [23] presents a two-stage segmentation model en-coder-decoder that focuses on the segmentation of tubes from pediatric CXR images. The encoder is ResNet50, and the decoder is ResNet101. The objective is to check the existence of tubes in the image and then highlight them in the CXR image if found. The importance of this method is saving performance by selecting the only image with lines for training and inference of the segmentation model. So, the first stage operates like a filter selecting only images which contain tubes. The second stage process only the filtered data with images containing tubes. Therefore, the processing time is reduced for training and inference.

C. Methods used to reduce u-net parameter number and increase Model performance

1) Truncated Architectures

Recently CNN models have been implemented in real-world medical devices and equipment. Those devices perform inference in real-time and require small models' architectures with few parameters to function. So, the research focuses more on the performance optimization when developing models for medical devices use. The objective is to get more accurate results with less computing time and using less resources. An example of research proposed by [24] tests and compares several truncated medical-oriented architectures with fewer parameters like ResNet18 and Effi-cientNetB0 for diseases detection tasks on the CheXpert dataset. The tested models give equivalent results to the big architectures. Models' truncation is performed by dropping some dense layers of the original models' structure and keeping the low-level layers unchanged.

2) knowledge Distillation

knowledge distillation technic used to tackle the large models' inference latency on edge mobile devices. It permits building fast inference models with equivalent results as large models or an ensemble of models. The knowledge distillation method can reduce the model parameters by 50x [25] or high. The principle of the distillation is to train a large model known as the teacher network. The teacher is used to train a smaller target model known as the student model. This approach is first announced by [26]. There are three ways to apply knowledge distillation. First, use the teacher probabilities predictions to train the student model [26]. The second consists of using a noise-based regularizer according to the teacher while the student training [27]. The last distillation technique uses in-between replicas learned by the teacher combined with the final predicted probabilities distributions to train the student model [28].

3) Deep convolutional neural networks techniques for abnormal tubes detection task

The Use of DCNNs to handle the task of abnormal tubes tends to replace methods based on feature extraction and SVM (Support Vector Machine) [29]. DCNNs benefit from a large amount of labeled available data that may be employed to identify tube mispositioning issues with excellent outcomes. Tubes, lines, and catheters are visible on 33% of chest X-ray

images [30]. Even yet, not much research focuses on automated multi-tubes position detection. Most of the work is focused on a specific type of tube, namely the ET tube. Otherwise, there are various efforts to prevent overfitting [31] by adding Gaussian noise, random brightness, and random gamma. The horizontal flip augmentation is utilized to simulate the AP/PA techniques used to capture X-ray images.

IV. Dataset

The following section introduces the dataset used for classification and segmentation tasks. We give the structure, labels, and technical parameters related to this dataset with brief data analysis for this dataset.

A. Classification data

The dataset we work with in this paper is the Royal Australian and New Zealand College of Radiologists (RANZRC) dataset [32]. The dataset was published publicly in 2021 by the Australian college to accurate the tubes and catheter detection task based on chest X-ray images. Twenty nine radiologists participate in the annotation of 40,000 samples with 11 tubes of positioning labels: Three classes for ETT (Ab-normal, Borderline, Normal), four classes for NGT (Abnormal, Borderline, Incomplete, Normal), three classes for CVC (Abnormal, Borderline, Normal), and finally one class for Swan Ganz tube (present)(see Fig.2).The training dataset is divided to train and validation parts; we resize the images to the classification model corresponding input. All the training images are augmented using random brightness.

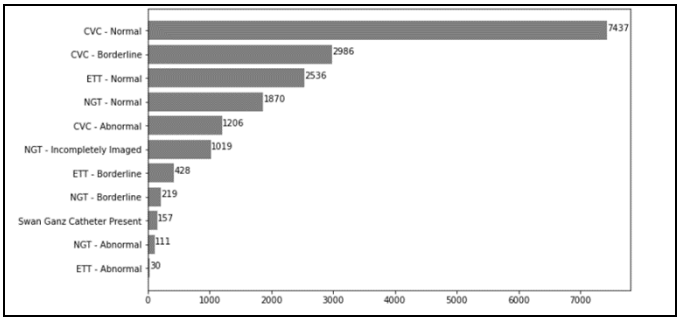

Figure 2. Dataset classes distribution.

B. Segmentation data

The dataset also contains 18000 images with their corresponding masks for the tube segmentation task. For the data repartition, 40% of the sample corresponds to unique patients (9095 patients). Otherwise, 20% correspond to patients having more than two x-ray images. In this study we use the segmentation part of the dataset to train our segmentation models, and we use the classification part to train classification models for the task of abnormal tube positioning detection. To prepare the data for the segmentation task. We start by resizing the images and masks to the target model input size. Then we normalize the images to get all pixel values between 0 and 1. After we split the data to train, validate, and test data. Then we convert the dataset into a TensorFlow tensor. Given the specificity of the tube segmentation task and that the data is enough to perform well for this task, we don't introduce any augmentation during the training process.

V. Methodology

This work aims at extracting tubes and lines from a patient CXR image using a segmentation model (Customized U-net). The extracted tubes can be associated with their prediction labels using a Deep convolutional neural network. The segmentation results (Image mask) associated with the classification results (prediction about the position of these tubes) can be presented visually to the medical staff to make a final diagnosis. To achieve this purpose, we will test some U-net architectures for medical segmentation to choose the more accurate and less compute cost architectures. We trained several U-Net variants with different number of parameters. All the training processes and parameters are detailed below.

A. Customized U-Net architectures

We take the original U-Net architecture as an initial point to develop a compressed customized version of the U-net. For the rest of this paper, we call the developed version Custom-U-Net-F. F in Fig.3 refers to the number of filters. Custom-U-net model is formed by varying the number of filters for bought convolutional and deconvolutional U-Net blocs. Then we replace the output with a multiclass layer. The number of classes is 6, which is the possible types and placements of tubes we can find in a training mask image. Next, we adapt the network input to our desired training input; (256,256,3) for the input images and (256,256,1) for the mask's images. The figure below shows the Custom-U-Net proposed architecture. We trained different Custom-U-Net variants by changing the number of filters. For evaluating and selecting the best Custom-U-Net variants, we compared the variants with the original trained U-Net network.

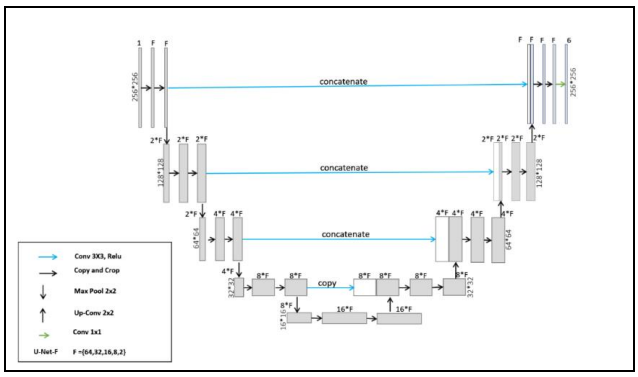

Figure 3. Customized U-Net Architecture.

B. Segmentation training process

The training process described below is applied to each model independently. We use this process to optimize and acquire the highest outcome for each model so that we can compare them later using the inference dataset. The process starts by training the customs U-Net segmentation models with different numbers of parameters (see table .2) on the segmentation dataset. Each image in the training dataset corresponds to a mask. The training is performed on the train dataset splitted into train (60%) and validation (20%) (see Table 1.). We keep 20% of the data for tests. We use the

parameters in the table below (see Table 4.) as a starting point for all the models' training. The loss function is categorical cross-entropy, and we use dice-coefficient as a metric. We run all the pieces of training on a local CentOS machine using Keras framework for 50 epochs. Then, we save all the models weights for future inferences on the test dataset.

Table 1. Dataset subsets for training, validation, and test.

Number of Samples par dataset	**Splits (60% for Training, 20% for validation and 20% for test)**
Training	10799 Sample
Validation	3600 Sample
Test	3600 Sample

The table below chows all trained models in this work.

Table 2. U-net Trained models list

Custom U-Net Variant	**Number of trainable parameters**
U-Net-64	34 550 598
U-Net-32	8 639 910
U-Net-16	2 161 110
U-Net-08	540 846
U-Net-04	135 498
U-Net-02	34 020

C. Classification models training process

In addition, the classification task is presented as a complement to the segmentation task, which is the main value of this study. Classification prediction associated with the segmentation image helps the medical staff to quickly validate the diagnosis and see if there is any issue. For the classification task, we select to test five models that fit well with the lines and catheters classification (ResNet50V2, DenseNet121, EfficientNetB4, DenseNet169, and Xception). We modified the original output to 11 classes new classification layer. We use transfer learning from ImageNet as initial weights initialization. To avoid training overfitting, we apply image augmentations such as rotation, horizontal flip, and random brightness. Moreover, we use early stopping and Learning rate reducer functions to stop training when the result doesn't relieve. Please find below the list of classification models (see Table 3.). Notice that we perform all the training and inferences tasks using the Keras framework.

Table 3. List of the classification trained models.

Model Name	**Number of trainable parameters**
ResNet50V2	23,541,899
DenseNet121	6,965,131
EfficientNetB4	17,568,339
DenseNet169	12,502,795
Xception	20,829,491

D. Classification models evaluation

During training, performances are evaluated by AUC (area under the curve). We choose AUC because it's an excellent metric for multiclass classification [33] problems. AUC estimates how the model can distinguish or separate classes. The higher the AUC, the better the model can classify 0s as 0s and 1s as 1s. Finding the ideal threshold values allows for the best possible separation between classes. AUC is calculated by measuring the percentage of TN (True negatives) and TP (True positives) compared to the sum of all True and False.

Accuracy: The total classes predicted correctly by the model.

$$Accuracy = \frac{TP + TN}{TP+TN+FP+FN} \quad (1)$$

E. Inference methodology

In practice, the inference is the function of exposing a trained model to a new image never seen before by this model, which is a real-world use of a trained model. In this paper, we use the test data splits for both segmentation and classification data to test the inference. For the segmentation task we fed the model an image without a mask from the test data. The model provides the mask image as output. Otherwise, for the classification model, we fed the model by the same chest X-Ray image, the one used to feed the segmentation model, and we get as an output the probabilities predictions for each tube position. Then we combine them(predictions) with their corresponding segmented tubes.

F. Framework of diagnosis validation

The diagnosis validation framework proposed by this work will help the medical staff, to prevent complications caused by the tubes positioning errors, by providing highly accurate results. The framework uses the trained models to output a mask image with the present tubes and the associated positioning labels (classification prediction) for each tube. In practice, we give an input image to the framework, and the framework generates the tubes masks using the segmentation model and predicts their positions by the classification model. After that, the radiologist can decide if the tube is normally positioned or suggest an adjustment. The figures (see Fig.4 and Fig.5) below shows how the framework works.

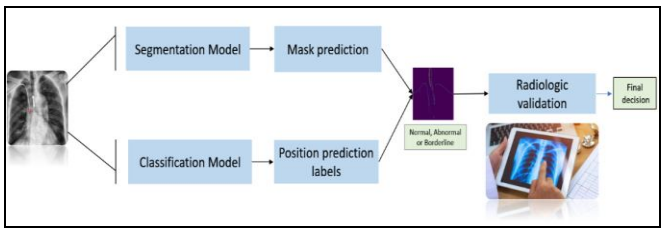

Figure 4. Diagnosis validation and abnormal tubes detection Framework.

To build an instance of this framework, we begin by choosing separately the best segmentation and classification models based on the test data. Both models have to accept the same input format, or we insert an adaptation transformer if the inputs are different. The segmentation and classification models are placed after the framework input layer. The second layer is used to combine the mask (segmentation output) with their corresponding prediction (classification output). Finally,

the last layer is a graphical interface to visualize and present the results to the radiologist who gives the diagnosis validation. The validation represents the output of our proposed framework. The image below shows the complete framework pipeline.

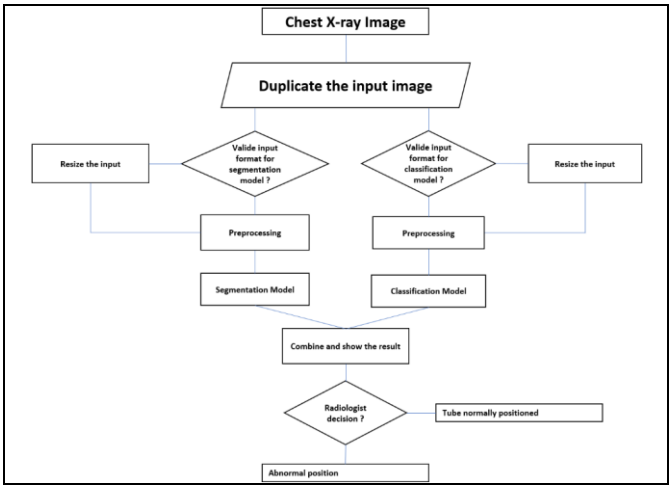

Figure 5. Framework pipeline

VI. EXPERIMENTS AND RESULTS

A. Segmentation with customized U-Net architectures

We trained six U-Net versions with different parameter numbers. All the models are trained using the same training pipeline. For the training pipeline, all U-Net models use the hyperparameters below (see Table 4.). For the optimizer, we use Adam [34] with a learning rate of 10e-3. Finally, we choose sparsecategorialCrossentropy as a loss function adapted to multiclass segmentation tasks. Indeed, we want that the mask distinct between different tube types highlighting them using different contrasts see Fig.6.

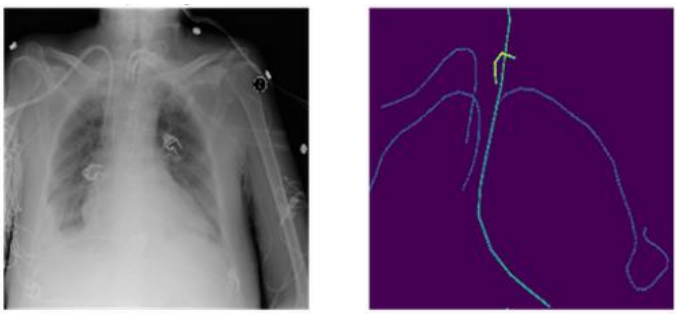

Figure 6. Mask with multi-tubes

Table 4. Segmentation training Hyperparameters.

Models' Training Hyperparameters	
Initial Learning rate	1e-3
Gradient descent Optimizer	Adam
Loss Function	SparseCategoricalCrossentropy
Input Image Format	(256,256,3)
Mask Image Format	(256,256,1)
EPOCHS Number	25

The table below shows the results of the six U-Net versions for the tubes and catheters segmentation task. The table reports the dice coefficients for each model inference on the 3600 samples (test dataset). The result also shows the number of trained parameters for all U-Net variants. Moreover, we associate the result with the training execution time to evaluate the training speed of the models.

Table 5. Tubes and catheters segmentation results.

Custom U-Net Variant	**Number of trainable parameters**	**Dice Coefficient**	**Training time (Min) for 15 epochs**	**Inference time For 3600 test Image (Second)**
U-Net-64	34 550 598	96.36%	4793.95	2010s
U-Net-32	8 639 910	94.50%	2296.52	1498s
U-Net-16	2 161 110	79.57%	1276.50	1299s
U-Net-08	540 846	41.06%	702.13	1182s
U-Net-04	135 498	5.35%	451.75	1167s
U-Net-02	34 020	0,13%	279.71	1153s

The objective is to select models accurate enough for the tube segmentation task and slight enough to run on medical devices. The more the model is parameters-less, the more is suitable for real-time inference in the work environments. The table below shows the dice coefficient degradation associated with the performance gain for different models' lengths (see Table 6.).

We can see from the table below that the dice coefficient falls with the reduction of the number of model parameters. The DC values for the models U-Net-02, U-net-04, and U-Net-08 are extremely low and not accurate enough to work with. However, for the other models U-Net16 and U-Net-32 the decrease of the DC coefficients 2% for U-Net-16 and 17% for U-Net32 is still acceptable for the segmentation task (see Fig.7). The best results are registered for the model U-Net-32 with 75% parameter less compared to U-Net-64 while the dice coefficient is still approximately the same as U-Net-64 (94.5%).

Table 6. Parameters reduction percentage for customized U-net Models.

Custom U-Net Variant	**Parameter reduction (U-Net-64 as a reference)**	**Dice Coefficient**	**Custom U-Net Variant**	**Parameter reduction (U-Net-64 as a reference)**
U-Net-64	0%	0%	0%	0%
U-Net-32	75%	1,93%	52,10%	25,47%
U-Net-16	93,75%	17,42%	73,37%	35,37%
U-Net-08	98,43%	57,39%	85,35%	41,19%
U-Net-04	99,61%	94,45%	90,58%	41,94%
U-Net-02	99,90%	99,87%	94,17%	42,64%

The image below shows the mask output for the tested models. We noticed that the results are acceptable for the models; U-Net (64,32, and 16). However, for the other models (U-net 8,4, and 2), the results are not intriguing. So, even though there is a huge drop in parameters for those models, they are useless (U-Net-2,4 and 8). For the selected models we can see

clearly that the results are exploitable. So, we can tolerate some loss of dice coefficient without a clear impact on the visual result (mask). Those models with fewer parameters can be deployed on production devices with low computing resources. Another advantage is the inference time for U-Net16 and U-Net32. We register 25% and 35% less inference time compared to U-Net64. Which is perfect in the production environment. For the 3600 tested images; U-Net-32 inference time is 1299s (360ms per image), and U-Net16 inference time is 1498 (416ms per image). Both models are considered acceptable compared to the inference time of the reference model U-Net-64 (0.55s per image). The large results gap is noticed for training time. For the reference model U-net64 the training takes 4793s (>17 hours) for 15 epochs compared to the selected models U-net16 and U-net32 which take successively 2296s and 1276s (< 12 hours). The reduced train time will help developers update and accurate quickly the models while collecting more data in the production environment.

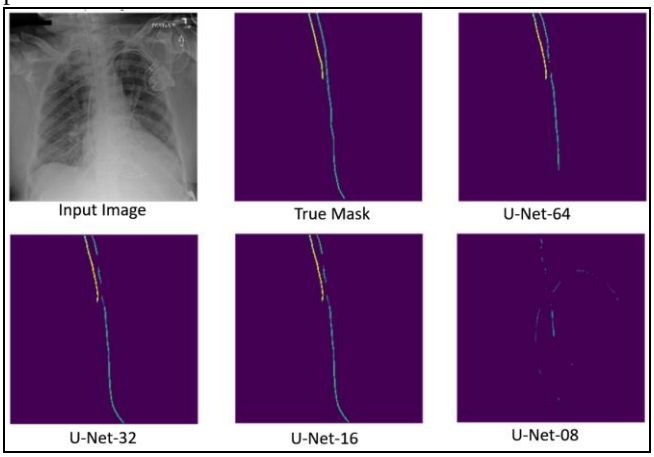

Figure 7. Segmentation predicted Masks for the trained U-net Models.

B. Classification models results.

The classification task is a supervised learning task performed using Deep Convolutional Neural Networks. Other technics are used to increase the accuracy of the classification tasks; for instance, stacked generalization [35-38]. We included a classification model optimization in this work as a complement to testing the proposed validation framework. The validation framework required two models: the segmentation and the classification model. For the segmentation, we proposed a detailed framework for selection and optimization. Otherwise, for the classification task we used an extension of one of our previous works [39-40] as a starting point. Then we adapted the model to the validation framework. The table below shows the training and inference results on the classification test dataset (see Table 7.).

Table 7. Inference results on the classification test dataset.

Model Name	Number of trainable parameters	AUC (Test dataset)
ResNet50V2	23,541,899	89.89%
DenseNet121	6,965,131	92.02%
EfficientNetB4	17,568,339	87.23%
DenseNet169	12,502,795	90.28%
Xception	20,829,491	88.88%

All the models are trained on the segmentation task. The task consists of predicting 11 possible placements for the inserted tube and catheters. The output for a tested Chest-X-Ray image is (11,1) vector with the AUCs for each class. The columns AUC in Table 7 show the AUC average for the inference on the test dataset. DenseNet121 is our selected model for the validation framework. The DenseNet121 has less trained parameter compared to the other models in the list, with the highest AUC (92.02%). All the tests carried out in the next section are performed using the DenseNet121 model.

C. Validation framework results

We test our framework using the test dataset. For the choice of the appropriate models, we pick the best classification and segmentation models based on the results in sections (6.2 and 6.3). The schema below shows the diagnosis of a patient's CXR image processed by the proposed framework (see Fig.8). We can see from the classification prediction that the CXR contains an ETT tube correctly positioned with a percentage of 98%. Otherwise, the segmentation model outputs a mask that visualizes the ETT tube in the exact position. With this information, the radiologist can decide if the ETT tube was positioned normally in the patient's X-ray image.

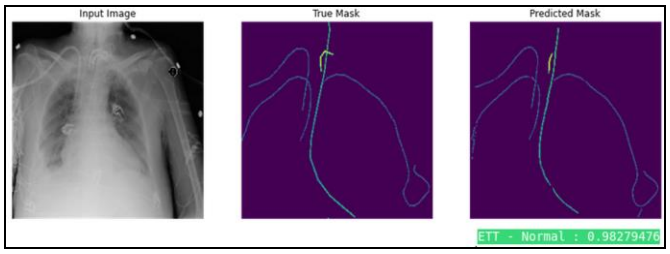

Figure 8. Validation Framework Output Example 1.

In the second example (see Fig.9). There is a mismatch between the segmentation result and the classification prediction. The segmentation mask shows a CVC tube incompletely segmented; otherwise, the classification prediction showed that the CVC tube is borderline. In such a case, we move to the next layer of the validation framework (Radiologist). The radiologist can see the contradiction and decide according to what he sees on the mask and his experience.

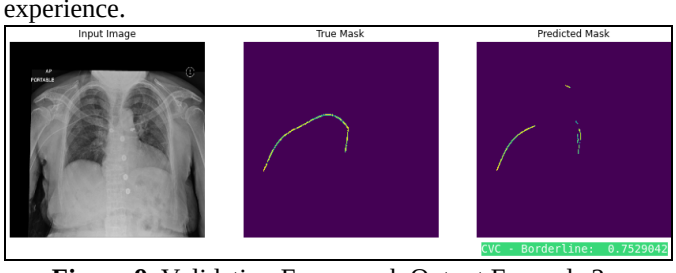

Figure 9. Validation Framework Output Example 2.

D. Discussion

Our proposed segmentation model performs as well as the available related work to the tube segmentation task. Wang et al.[22] proposed a segmentation method for the PICC tubes with a dice coefficient of 44%. Considering that our work targets the segmentation of tubes other than PICC, it is still an equivalent task of general tube segmentation. So, we can compare our results with the Wang proposed model. Our model outperforms this result for all the segmented tubes in our study. We registered a dice coefficient of 94% for the U-Net-32 model. Moreover, comparing our result to Ryan Sullivan et al. The work of Ryan Sullivan addresses the tube segmentation of lines in a pediatric chest radiograph; the study reported a dice coefficient of 73% less than our dice coefficient of 94% (see Table. 8).

Table 8. Related works results comparison.

Model	Dice Coefficient
U-Net-32(Our)	94.5%
TwoStage-UNet-EfficientNetB3[22]	73.7%
Mag-Net [23]	44.11%

All the tests presented in this work were carried out independently using the models forming the framework. We test the classification model on the test dataset. Then the segmentation model. The test of the complete framework in not been performed yet. Indeed, the test includes the view of a radiologist as a final step to decide if the tube is normal or abnormally positioned which is not available in this work project. We planned the includes this test in a future article.

VII. CONCLUSION

Tubes and catheters are saving lives devices; these devices are used in intensive care units to support ill patients in breeding, administrating drugs, or feeding unconscious patients [41][42]. Some errors in the placement process by doctors can cause crucial health complications. Moreover, correcting these errors as quickly as possible is mandatory[43]. Through this work, we propose a framework for accelerating tube readjustment procedures. The framework diagnoses the post-operation chest image and output immediately visual and numerical results for each present tube. The framework results are easily understandable by the medical staff. Then, a radiologist is needed for an urgent diagnosis validation of the suspect's possible error in priority to avoid the risk caused by the retard. Finally, the proposed tubes and catheter diagnosis validation framework gives a relevant diagnosis methodology and results. However, the framework still needs to be tested in real situations by medical experts in intensive care units.

Endnote

Part of this work was presented at the 13th International Conference on Innovations in Bio-Inspired Computing and Applications(IBICA'22), December 15-17, 2022 [41].

References

[1] "Graber MLThe incidence of diagnostic error in medicineBMJ Quality & Safety 2013;22: ii21-ii27."

[2] "Tracheal intubation in the ICU: Life saving or life threatening? Indian J Anaesth. 2011 Sep-Oct; 55(5): 470–475."

[3] "Koopmann MC, Kudsk KA, Szotkowski MJ, Rees SM." A Team-Based Protocol and Electromagnetic Technology Eliminate Feeding Tube Placement Complications [Internet]". Vol. 253, Annals of Surgery. 2011. p. 297–302."

[4] " Sorokin R, Gottlieb JE. "Enhancing patient safety during feeding-tube insertion: a review of more than 2,000 insertions". JPEN J Parenter Enteral Nutr. 2006 Sep;30(5):440–5."

[5] "Marderstein EL, Simmons RL, Ochoa JB. Patient safety: effect of institutional protocols on adverse events related to feeding tube placement in the critically ill. J Am Coll Surg. 2004 Jul;199(1):39–47; discussion 47–50."

[6] "Fraser RS. Lung perforation complicating tube thoracostomy: pathologic description of three cases. Hum Pathol 1988;19:518-23"

[7] "Pollard, Richard J. MD; Lobato, Emilio B. MD Endotracheal Tube Location Verified Reliably by Cuff Palpation, Anesthesia & Analgesia: July 1995 - Volume 81 - Issue 1 - p 135-138"

[8] "Ronneberger, O. et al., U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science(), vol 9351."

[9] "Damian J. Matuszewski et al., Reducing the U-Net size for practical scenarios: Virus recognition in electron microscopy images, Computer Methods and Programs in Biomedicine, Volume 178, 2019, Pages 31-39, ISSN 0169-2607."

[10] "T. Taxt, P.J. Flynn and A. K. Jain, Segmentation of document images, IEEE Trans. Pattern Analysis Mach. lntell. 11(12), 1322-1329 (1989)."

[11] "C. K. Chow and T. Kaneko, Automatic boundary detection of the left-ventricle from cineangiograms, Comput. Biomed. Res. 5, 388-410 (1972). "

[12] "Y. Nakagawa and A. Rosenfeld, Some experiments on variable thresholding, Pattern Recognition ll, 191-204 (1979). "

[13] " S. D. Yanowitz and A. M. Bruckstein, A new method for image segmentation, Comput. Vision Graphics Image Process. 416, 82-95 (1989)."

[14] "A. Rosenfeld and A. C. Kak, Digital Picture Processing. Academic Press, New York (1982). "

[15] "E. L Hall, Computer Image Processing and Recognition. Academic Press, New York (1979). "

[16] "R. C. Gonzalez and P. Wintz, Digital Image Processing. Addison-Wesley, Reading, Massachusetts (1987)."

[17] "M. N. Saad et al., Image segmentation for lung region in chest X-ray images using edge detection and morphology, 2014 IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2014), 2014, pp. 46-51."

[18] " Goo JM et al., Automated lung nodule detection at low-dose CT: preliminary experience. Korean J Radiol. 2003 Oct-Dec;4(4):211-6. doi: 10.3348/kjr.2003.4.4.211. PMID: 14726637; PMCID: PMC2698098. "

[19] "Ronneberger, O. et al., U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science(), vol 9351."

[20] "Damian J. Matuszewski et al., Reducing the U-Net size for practical scenarios: Virus recognition in electron microscopy images, Computer Methods and Programs in Biomedicine, Volume 178, 2019, Pages 31-39, ISSN 0169-2607."

[21] Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.

[22] "Wang, Xiaoyan, et al. Automatic and accurate segmentation of peripherally inserted central catheter (PICC) from chest X-rays using multi-stage attention-guided learning. Neurocomputing 482 (2022): 82-97."

[23] "Sullivan, Ryan, et al. "Deep learning methods for segmentation of lines in pediatric chest radiographs." Medical Imaging 2020: Computer-Aided Diagnosis. Vol. 11314. SPIE, 2020."

[24] "Ke, Alexander, et al. "CheXtransfer: performance and parameter efficiency of ImageNet models for chest X-Ray interpretation." Proceedings of the Conference on Health, Inference, and Learning. 2021."

[25] "Liu, Yifan, et al. "Structured knowledge distillation for semantic segmentation." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019."

[26] "G. Hinton, O. Vinyals, and J. Dean. Distilling the knowledge in a neural network. arXiv preprint arXiv:1503.02531, 2015."

[27] "Meng, Zhong, et al. "Adversarial teacher-student learning for unsupervised domain adaptation." 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2018."

[28] "Romero, A., Ballas, N., Kahou, S. E., Chassang, A., Gatta, C., & Bengio, Y. (2014). Fitnets: Hints for thin deep nets. arXiv preprint arXiv:1412.6550."

[29] "Widodo, A., & Yang, B. S. (2007). Support vector machine in machine condition monitoring and fault diagnosis. Mechanical systems and signal processing, 21(6), 2560-2574."

[30] "Van Ginneken, B., Hogeweg, L., Prokop, M.: Computer-aided diagnosis in chest radiography: beyond nodules ‹Eur. J. Radiol., 2009, 72, (2), pp.226–230"

[31] "Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting.. Journal of Machine Learning Research, 15, 1929-1958."

[32] "Tang, J.S.N., Seah, J.C.Y., Zia, A. et al. CLiP, catheter and line position dataset. Sci Data 8, 285 (2021). https://doi.org/10.1038/s41597-021-01066-8"

[33] "Techa, C., Ridouani, M., Hassouni, L., & Anoun, H. (2022, November). Alzheimer's Disease Multi-class Classification Model Based on CNN and StackNet Using Brain MRI Data. In Proceedings of the 8th International Conference on Advanced Intelligent Systems and Informatics 2022 (pp. 248-259). Cham: Springer International Publishing."

[34] "Jais, Imran Khan Mohd, Amelia Ritahani Ismail, and Syed Qamrun Nisa. "Adam optimization algorithm for wide and deep neural network." Knowledge Engineering and Data Science 2.1 (2019): 41-46."

[35] "Rabbah, J., Ridouani, M., Hassouni, L.: A New Classification Model Based on Stacknet and Deep Learning for Fast Detection of COVID 19 Through X Rays Images. In: Fourth International Conference on Intelligent Computing in Data Sciences (ICDS), pp. 1-8. (2020).

[36] "Rabbah, J., Ridouani, M., Hassouni, L.: A New Churn Prediction Model Based on Deep Insight Features Transformation for Convolution Neural Network Architecture and Stacknet. (IJWLTT), 17(1), 1-18. (2022)."

[37] "Benazzouza, S., Ridouani, M.,Salahdine, F., Hayar, A.:A Novel Prediction Model for Malicious Users Detection and Spectrum Sensing Based on Stacking and Deep Learning. Sensors 22 (17), 6477 (2022)."

[38] "Techa, C., Ridouani, M., Hassouni, L., & Anoun, H. (2023, March). Automated Alzheimer's Disease Classification from Brain MRI Scans Using ConvNeXt and Ensemble of Machine Learning Classifiers. *In Proceedings of the 14th International Conference on Soft Computing and Pattern Recognition (SoCPaR 2022) (pp. 382-391). Cham: Springer Nature Switzerland."*

[39] "Elaanba, Abdelfettah, Mohammed Ridouani, and Larbi Hassouni. "A Stacked Generalization Chest-X-Ray-Based Framework for Mispositioned Medical Tubes and Catheters Detection." Biomedical Signal Processing and Control 79 (2023): 104111."

[40] "Elaanba A, Ridouani M, Hassouni L. Automatic detection using deep convolutional neural networks for 11 abnormal positioning of tubes and catheters in chest x-ray images. In *2021 IEEE World AI IoT Congress (AIIoT) 2021 May 10 (pp. 0007-0012). IEEE*."

[41] "Elaanba, A., Ridouani, M., & Hassouni, L. (2023, March). Automatic Diagnosis Framework for Catheters and Tubes Semantic Segmentation and Placement Errors Detection. In *Innovations in Bio-Inspired Computing and Applications: Proceedings of the 13th International Conference on Innovations in Bio-Inspired Computing and Applications (IBICA 2022) Held During December 15-17, 2022 (pp. 176-188). Cham: Springer Nature Switzerland."*

[42] L Dora, S Agrawal, R Panda, A Abraham, Optimal breast cancer classification using Gauss–Newton representation based algorithm, Expert Systems with Applications, 97: 134-145, 2017.

[43] A Abraham, Intelligent systems: Architectures and perspectives, Recent advances in intelligent paradigms and applications, Springer, 1-35, 2003.