

A DATA-CONSTRAINED APPROACH FOR OCCUPATIONAL SILICOSIS DETECTION ON CHEST X-RAYS WITH FEW-SHOT LEARNING

NGUYEN THI TAN TIEN¹, BUI QUOC BAO², PHAM VAN CUONG^{3,*}

¹*Thai Nguyen University of Medicine and Pharmacy, 284 Luong Ngoc Quyen Street,
Thai Nguyen City, Thai Nguyen Province, Viet Nam*

²*VinUniversity, Vinhomes Ocean Park, Gia Lam District, Ha Noi, Viet Nam*

³*Posts and Telecommunications Institute of Technology, Nguyen Trai Street, Mo Lao Ward,
Ha Dong District, Ha Noi, Viet Nam*



Abstract. Occupational silicosis is a serious lung disease caused by long-term exposure to silica dust, mainly affecting workers in industries such as mining and construction. Diagnosis of silicosis is challenging due to subtle disease manifestations on chest X-rays (CXRs) and limited labeled medical data. Traditional deep learning models, such as Convolutional Neural Networks (CNNs), often require large datasets, which are often heavily expensive and time-consuming for collection and annotation, yet useful for specialized medical applications. To address these challenges, we present the use of Few-Shot Learning (FSL) to enable accurate detection of occupational silicosis with a minimal number of labeled examples. Our experimental results demonstrate that the FSL-based model achieves 84.4% accuracy and 46.0% mIoU in the 1-shot setting and 89.52% accuracy with 47.89% mIoU in the 4-shot setting. These findings highlight the potential of FSL to improve diagnostic accuracy in data-limited environments, making it a viable solution for improving medical image analysis in resource-constrained settings.

Keywords. Few-shot learning, few-shot segmentation, image segmentation, silicosis detection.

1. INTRODUCTION

Occupational silicosis is a chronic and potentially fatal lung disease caused by prolonged inhalation of silica dust, primarily affecting workers in industries such as mining, construction, and manufacturing [1]. The early and accurate diagnosis of silicosis is crucial for preventing disease progression and improving patient outcomes. However, the diagnosis of silicosis is challenging due to the subtle and varied presentation of the disease on chest X-rays (CXR), and the scarcity of labeled medical images available for training deep-learning models [2].

In recent years, deep learning, particularly Convolutional Neural Networks (CNNs), has shown great promise in medical image analysis, significantly advancing the field of automated

*Corresponding author.

E-mail addresses: nguyenthitantien@tnmc.edu.vn (N.T.T. Tien); bao.BQ222002M@sis.hust.edu.vn (B.Q. Bao); cuongpv@ptit.edu.vn (P.V. Cuong).

disease detection. However, traditional deep learning models typically require large amounts of annotated data to achieve high accuracy, which is often not feasible in specialized medical fields like silicosis diagnosis. This limitation presents a significant challenge in developing effective diagnostic models for silicosis [3].

To address this challenge, Few-Shot Learning (FSL) offers a promising approach by enabling models to generalize well with only a few labeled examples. FSL is particularly well-suited for scenarios where obtaining large, annotated datasets is time-consuming and expensive. This paper explores the application of FSL in diagnosing occupational silicosis from chest X-ray images, aiming to develop a model that can perform accurately even with a limited number of training samples. By leveraging FSL techniques, we aim to improve the diagnostic process, making it more accessible and reliable in settings with constrained data availability [4].

Chest X-ray (CXR) imaging is a frequently employed method for diagnosing patients exposed to risk factors [5, 6]. Moreover, integrating X-ray analysis into decision support systems can utilize deep learning models to enhance diagnostic accuracy and support healthcare professionals in their decision-making processes. As a result, it's vital to build a CXR image dataset that includes silicosis cases and to develop deep learning models specifically designed for detecting this disease.

2. RELATED WORK

2.1. Radiologist-centric methods

For over a century, chest X-rays (CXR) have been a fundamental tool in diagnosing lung diseases [7]. Traditionally, radiologists analyze these images visually to detect patterns and abnormalities indicative of various pulmonary conditions. While this method is invaluable, it is also time-consuming and demands substantial training and expertise. Moreover, the subjective nature of visual interpretation can result in variability between radiologists, where different professionals may draw different conclusions from the same image [8].

As lung disease incidence continues to rise, coupled with the growing need for rapid and accurate diagnostics, there has been a push towards exploring alternative approaches that can enhance and potentially replace traditional methods. In response, research has increasingly focused on developing computer-aided diagnostic (CAD) systems [9] designed to support radiologists in interpreting CXR images. These systems often utilize advanced image processing and analysis techniques to extract key features from the images, offering additional insights and decision support to improve diagnostic accuracy and efficiency.

Beyond CAD systems, research has also explored other innovative methods for diagnosing lung diseases using CXR. Some studies have investigated the potential of clinical decision support systems (CDSS), which combine patient history and other clinical data with CXR findings to provide comprehensive recommendations for diagnosis and treatment [10]. Additionally, efforts have been made to develop standardized reporting systems for CXR, aimed at enhancing consistency and accuracy in interpretations across different healthcare providers [11]. These initiatives are critical in reducing variability between radiologists and improving communication between radiologists and other clinicians.

2.2. Traditional machine learning methods

Detecting lung diseases, particularly pneumonia, from CXR images presents a significant challenge, mainly due to the scarcity of annotated datasets [5, 12]. Various traditional machine-learning techniques have been employed to tackle this issue. For instance, Chandra et al. [13] focused on segmenting lung regions and extracting eight statistical features, using classifiers such as multilayer perceptron (MLP), random forest, sequential minimal optimization (SMO), regression-based classification, and logistic regression. Their work demonstrated strong potential, achieving a notable 95.39% accuracy with MLP on a dataset containing 412 images. Similarly, Kuo et al. [14] utilized 11 features to identify pneumonia in patients with schizophrenia, achieving a remarkable 94.5% accuracy using a decision tree classifier. Yue et al. [15] explored six features on chest CT scans, resulting in a peak AUC of 97%. However, these conventional approaches often struggle with the complexity of multi-dimensional X-ray images and sophisticated medical data, which can lead to limitations in their adaptability and flexibility.

2.3. Deep-learning based methods

The field of medical image analysis has seen a profound shift, gradually moving away from conventional machine learning methods towards the prevalent use of deep learning techniques. Unlike traditional algorithms that depend on manually crafted features for classification or segmentation [16, 17], deep learning particularly Convolutional Neural Networks (CNNs) enables end-to-end learning, where relevant features are automatically derived from raw input data [18, 19]. CNNs have become fundamental in image classification due to their capacity to extract translation-invariant features through convolutional layers. For instance, Sharma et al. [20] and Stephen et al. [21] illustrate this transition by designing simple CNN architectures for classifying pneumonic chest X-rays (CXR). Despite data limitations, they applied data augmentation to improve model accuracy. However, this technique also exposed the limitations of data augmentation in offering truly new information that could significantly enhance CNN performance.

Rajpukar et al. [22] demonstrated a pioneering application of CNNs with CheXNet, a 121-layer model built for lung disease analysis, which achieved a 76.8% F1-score while highlighting the challenges posed by the absence of patient history in affecting model performance. CheXNet not only forecasts lung diseases but also generates heatmaps indicating areas of interest on X-rays. Further research [23] delved into using cutting-edge CNN architectures for detecting pneumothorax, achieving an AUC of 0.75. Ozturk et al. [24] introduced Darknet, a new method for automatically detecting COVID-19 from raw chest X-ray images, with high accuracy in both binary and multi-class classifications. AlMamlook et al. [25] proposed an integrated model that combines seven machine-learning models with renowned CNN architectures, achieving a remarkable overall accuracy of 98.46%. These studies highlight the crucial role of CNNs in advancing medical image analysis, particularly for critical respiratory conditions.

In summary, the adoption of deep learning, especially Convolutional Neural Networks (CNNs), has fundamentally transformed the field of medical image analysis, overcoming the constraints of traditional machine learning approaches.

2.4. Few-shot learning method

For medical imaging data related to occupational silicosis, the use of Few-Shot Learning (FSL) is particularly important for several reasons. First, the amount of labeled data in this field is often very limited due to the challenges in collecting and annotating medical images. FSL allows models to learn from a small number of samples, reducing the need for large datasets while still maintaining high accuracy in diagnosis. Second, FSL can help in the early and accurate detection of signs of occupational silicosis, thereby supporting more effective treatment and prevention efforts. Finally, applying FSL in the diagnosis of silicosis also helps minimize variability in assessments by doctors, thanks to the model's strong generalization capability across classes that it has not been trained on.

Few-shot learning (FSL) is a specialized branch of meta-learning [26], focused on training models that can perform well on new data by using only a few samples from related tasks. The main idea of FSL is to identify a hypothesis space and determine the optimal hypothesis. FSL includes interesting variants such as one-shot learning [27–29], where each label is classified using only one sample per class, and zero-shot learning [30–33], which handles unseen data by relying on data descriptions without any labeled samples.

FSL has demonstrated significant potential across various applications in medical imaging. A systematic review by Pachetti and Colantonio (2024) identified that FSL techniques, including Prototypical Networks and Model-Agnostic Meta-Learning (MAML), are particularly useful for tasks such as segmentation and classification, where annotated data are limited [34]. This review underscored FSL's role in improving the adaptability of diagnostic models to new imaging modalities and disease types.

In tuberculosis (TB) diagnosis, Pramana et al. applied Prototypical Networks with ResNet architectures on chest X-rays, achieving high classification accuracy despite class imbalances in the dataset. This approach demonstrated FSL's capacity to deliver reliable classification performance with minimal labeled samples, which is particularly valuable in resource-limited settings [35]. Likewise, Galán-Cuenca et al. used a Siamese network-based FSL model to classify COVID-19 cases in chest X-rays, addressing the challenges of intra- and inter-domain imbalances, which are common in pandemic-related data [36]. For segmentation tasks, Wang et al. introduced SSA-Net, a model combining spatial self-attention and semi-supervised FSL, to segment COVID-19 pneumonia in CT scans. This model tackled data imbalance and effectively segmented complex lesion patterns, demonstrating FSL's ability to derive meaningful insights even from limited labeled data [37]. Additionally, in skeletal imaging, Xie et al. developed an FSL framework using ResNet, VGG, and Xception models to diagnose osteopenia and osteoporosis from knee X-rays. The FSL model surpassed radiologists in specific diagnostic scenarios, underscoring FSL's capability in cases with restricted labeled data [38].

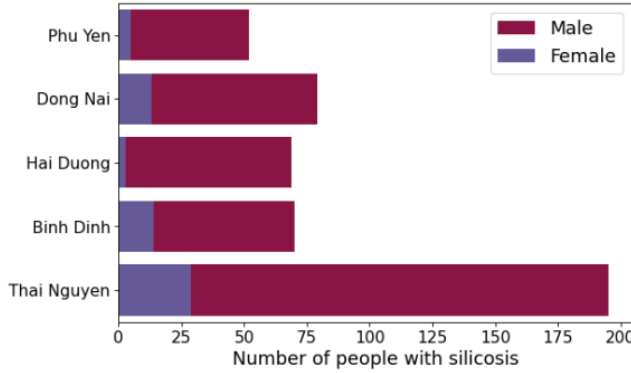
For occupational silicosis, a rare and challenging disease to diagnose via chest X-rays, FSL could be an ideal solution for early and precise detection of lung lesions, even with limited data availability.

3. DATASET

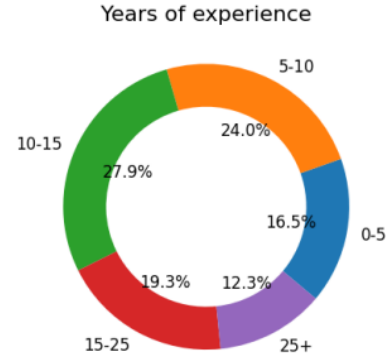
3.1. Data collection

The research team selected a sample of 2,089 workers directly engaged in production activities, frequently exposed to silica dust at various sites across Vietnam, including Thai Nguyen (metalworking), Hai Duong (cement production), Binh Dinh, Phu Yen (granite mining), and Dong Nai (brick manufacturing). These workers were part of the national project titled: *“Research on the Molecular Epidemiology, Risk Factors, and Application of Advanced Techniques for Early Diagnosis of Silicosis in Vietnam”*, led by Hanoi Medical University.

To qualify for the study, participants needed a minimum of one year of work experience by the time of data collection. Each patient had chest X-rays and lung dust films measuring 35×43 cm, adhering to ILO 2000 standards. All participants were thoroughly informed of the study’s purpose and voluntarily consented to participate in interviews. Investigators ensured full respect for each participant’s decisions and responses, without criticism or judgment during data collection. Furthermore, participants were provided clear explanations of the medical examination procedures, including chest X-rays and respiratory function tests, and gave informed consent for these tests. All personal information collected from participants was kept strictly confidential.



(a) Statistics by gender and location



(b) Statistics by years of experience

Figure 1: Statistics of individuals diagnosed with silicosis pneumonia

Of the data collected, 535 cases were confirmed as silicosis. Figure 1(a) shows the distribution of patients by gender and location, revealing a stark imbalance, with the majority of cases being male. This reflects the predominance of men in high-risk industries. Thai Nguyen, in particular, had the highest number of silicosis cases, underscoring the area’s elevated risk for this occupational disease.

Figure 1(b) breaks down the data by years of work experience, demonstrating a clear link between prolonged exposure to silica dust and the likelihood of developing the disease. Workers with 10-15 years of experience comprised the largest share of cases, followed by those with 5-10 years. Interestingly, the incidence of silicosis was lower among those with over 25 years of experience, possibly due to a smaller workforce in that age group or workers leaving high-risk jobs before the disease could fully manifest.

This analysis highlights a strong correlation between work experience, geographic location, gender, and the risk of silicosis. It also emphasizes the critical need for early detection and intervention, especially for workers with prolonged exposure to silica dust in high-risk environments.

3.2. Data processing

The dataset consists of DICOM (.dcm) files, a common format used for storing medical imaging data due to its capacity to retain important clinical information. While these files can be visualized effectively using specialized software such as MicroDicom or RadiAnt, their large size and complex structure make them impractical for direct use in machine-learning algorithms. Additionally, the process of converting and viewing DICOM images in formats like JPG or PNG is both time-consuming and inefficient.

To address these limitations, the X-ray images were converted to more widely used formats such as JPEG and PNG. This conversion facilitates easier processing and analysis, ensuring compatibility with popular machine learning frameworks like PyTorch. The conversion process was automated using Python scripts to consistently handle the transformation of DICOM files to jpg format, ensuring uniformity across the dataset.

The final dataset comprises 2005 images, of which 445 represent cases of silicosis, while the remaining 1560 depict non-disease cases.

To ensure consistency during model training, all images were resized to 256×256 pixels. This specific resolution was chosen because it is well-suited to most convolutional neural network (CNN) architectures, which often require input images to be of a fixed size. Resizing the images helps mitigate issues that could arise from varying image dimensions, thereby improving the efficiency of the model training process.

By standardizing both the format and size of the images, we were able to reduce the complexity of data processing, optimize storage, and ensure that the dataset is well-prepared for use in machine learning applications. This preprocessing step plays a crucial role in improving the performance and scalability of the machine learning models.

Table 1: Our collected dataset

Class	Id	Number of images
Background	0	2005
Bone	1	2005
Normal	2	509
Silicosis	3	445
Viral	4	520
Bacterial	5	531

4. FEW-SHOT LEARNING FOR OCCUPATIONAL SILICOSIS DETECTION

4.1. Problem definition

Few-shot learning is a technique designed to address the challenge of performing tasks with a very limited amount of labeled data, which is particularly useful in scenarios where

acquiring large, annotated datasets is impractical or costly. In our study, we work with two distinct image datasets: (1) a base set, denoted as $\mathcal{D}^T$, comprising classes that are well-represented and extensively labeled, and (2) a novel set, referred to as $\mathcal{D}^E$, which contains classes that are underrepresented or entirely new. The novel set specifically consists of chest X-ray images, with the lung disease classes in this set being different from those present in the base set. This difference in class distribution poses a unique challenge, as the model must generalize to the novel classes with minimal exposure.

In our methodology, the LungDiseasePTIT dataset is strategically divided into two subsets: the training set ($\mathcal{D}^T$) and the test set ($\mathcal{D}^E$). The training set includes a range of base classes, while the test set introduces novel classes that the model has not encountered during training. We denote the sets of classes corresponding to $\mathcal{D}^T$ and $\mathcal{D}^E$ as $\mathcal{C}^T$ and $\mathcal{C}^E$, respectively, with each subset having a distinct set of lung disease categories. Notably, in the context of the LungDiseasePTIT dataset, the combined number of classes across both subsets is $|\mathcal{C}^T| + |\mathcal{C}^E| = 6$, reflecting a small yet diverse set of conditions that the model must learn to identify.

To detect lung diseases within a few-shot learning framework, we implement episodic learning, as described in [39], a method renowned for its effectiveness in few-shot scenarios. This approach involves generating multiple few-shot tasks during training, where each task selects a small subset of data from the training set $\mathcal{D}^T$ and divides it into a support set $\mathcal{S}^T$ and a query set $\mathcal{Q}^T$.

The support set $\mathcal{S}^T$ contains labeled images with information about different types of lung diseases, providing the model with the necessary context for learning. The query set $\mathcal{Q}^T$, composed of unlabeled images, challenges the model to both classify and segment lung disease regions based on the knowledge acquired from the support set. This structured approach simulates the few-shot learning condition, allowing the model to adapt to new, unseen scenarios with minimal data.

Following the training phase, the model is evaluated using the novel test set $\mathcal{D}^E$, which is similarly divided into a support set $\mathcal{S}^E$ and a query set $\mathcal{Q}^E$. The evaluation involves using the model to predict labels and segment lung disease regions in the query set, which contains chest X-ray images of novel classes. The support set $\mathcal{S}^E$ includes a few examples of a new type of lung disease, and the model must generalize its learning to accurately predict and segment these new instances.

In this episodic framework, each episode is defined by two primary sets.

$$\mathcal{S}^* = \{(x_i^s, a_i^s) \mid x_i^s \in \mathbb{R}^{I \times W \times C}, a_i^s \in \mathbb{R}^{I \times W}\}_{i=1}^{NK}, \quad (1)$$

where, x_i^s and a_i^s represent a support image and its corresponding label, respectively. The superscript * indicates the training (T) or test (E) phase, N is the number of classes, and K is the number of labeled instances per class, embodying the N -way K -shot problem. H , W , and C denote the height, width, and channels of the image.

$$\mathcal{Q}^* = \{x_j^q \mid x_j^q \in \mathbb{R}^{I \times W \times C}\}_{j=1}^M, \quad (2)$$

where, x_j^q is a query image, and M is the number of images in the query set. The superscript * denotes the phase (training or test), and j indexes the samples in the query set.

This method ensures the model is well-equipped to handle the complexities of occupational silicosis detection, even when only a few labeled examples are available for training.

In the classification task, our goal is to determine the multi-hot class occurrence vector $\mathbf{y}_C \in \mathbb{R}^N$ using a classification function f_C . Simultaneously, for the segmentation task, we aim to generate a segmentation mask $\mathbf{Y}_S \in \mathbb{R}^{I \times W}$ associated with these classes via a segmentation function f_S . These objectives are mathematically formulated as.

$$\begin{aligned}\mathbf{y}_C &= f_C(x_j^q, \mathcal{S}^*; \theta_C), \\ \mathbf{Y}_S &= f_S(x_j^q, \mathcal{S}^*; \theta_S),\end{aligned}\tag{3}$$

where, θ_C and θ_S represent the learnable parameters of the classification and segmentation models, respectively.

Rather than optimizing these functions f_C and f_S separately, this study focuses on a unified approach, aiming to jointly optimize a single function f_{CS} that integrates both few-shot classification and segmentation tasks (FS-CS). This unified model f_{CS} , parameterized by θ_{CS} , takes a query image x_j^q and a support set $\mathcal{S}^*$ as inputs, and outputs both a multilabel class occurrence vector $\mathbf{y}_C$ and a corresponding segmentation map $\mathbf{Y}_S$.

$$\{\mathbf{y}_C, \mathbf{Y}_S\} = f_{CS}(x_j^q, \mathcal{S}^*; \theta_{CS}),\tag{4}$$

here, $\mathbf{y}_C \in \mathbb{R}^N$ denotes the predicted multi-hot class occurrence vector, and $\mathbf{Y}_S \in \mathbb{R}^{I \times W}$ represents the predicted class-wise segmentation mask.

The FS-CS approach not only generalizes the few-shot classification (FS-C) but also offers significant advantages over both FS-C and few-shot segmentation (FS-S):

- FS-CS can classify query images that either belong to none or multiple target classes. For instance, if the disease is not present, the system accurately identifies this and avoids issuing false warnings by classifying the image as background.

- Unlike conventional FS-S methods [40–42], which assume that the query class set exactly matches the support class set, FS-CS relaxes this constraint, allowing the query class set to be a subset of the support class set.

To solve Eq. (4), the model extracts N probability maps corresponding to each class in the support set, forming the class-wise foreground map set $\mathcal{Y}$. Each probability map $\mathbf{Y}^{(n)} \in \mathbb{R}^{I \times W}$ represents the likelihood of a pixel belonging to the foreground of a particular class, maintaining the same resolution as the input image. The overall process is expressed as.

$$\mathcal{Y} = f(x_j^q, \mathcal{S}^*; \theta) = \left\{ \mathbf{Y}^{(n)} \right\}_{n=1}^N,\tag{5}$$

where, f is the model before post-processing and θ are the model's learnable parameters. The output $\mathcal{Y}$ is then post-processed to extract both $\mathbf{y}_C$ and $\mathbf{Y}_S$.

5. EXPERIMENTS

5.1. Experimental setup

In this study, we evaluate our proposed method by comparing it to four prominent few-shot learning algorithms that have set benchmarks in instance segmentation and object

detection. The Path Aggregation Network (PANet) [43] has excelled in the COCO 2017 Challenge, achieving first place in Instance Segmentation and second in Object Detection through effective feature integration across multiple scales. The Prior Guided Feature Enrichment Network (PFENet) [44], recognized for its strength in Few-Shot Learning (FSL), enhances generalization from limited data by leveraging prior knowledge. Meanwhile, the Hypercorrelation Squeeze Networks (HSNet) [45] introduce a novel framework for Few-Shot Classification and Segmentation (FS-CS) with lightweight 4D convolutions, balancing efficiency and accuracy. Lastly, the Attentive Squeeze Network (ASNet) [46] demonstrates state-of-the-art performance in both FS-CS and FS-S tasks by employing an attention mechanism that prioritizes relevant features for improved class discrimination. This comparative analysis aims to underscore the strengths of our method within the current landscape of advanced techniques.

For the evaluation metrics in multi-label classification, we utilize the exact ratio of 0/1 : ER = 1 $[\mathbf{y}_C = \mathbf{y}_{gt}]$, where $\mathbf{y}_{gt}$ represents the ground truth multi-hot class occurrence vector, and $\mathbf{y}_C$ denotes the predicted vector generated by our model. This metric provides a clear measure of the model’s accuracy in classifying multiple labels, which is increasingly relevant in real-world applications where multiple objects may need to be detected simultaneously. In the case of segmentation, we adopt the mean Intersection over Union (mIoU) metric: $mIoU = \frac{1}{N} \sum_n IoU_n$, with IoU_n indicating the Intersection over Union value for the n^{th} class. This comprehensive assessment ensures that our method is rigorously evaluated, particularly in distinguishing between similar classes, thus providing a more nuanced understanding of its performance.

5.2. Training configurations

We selected ResNet50, pre-trained on ImageNet, as the backbone network for our method, ensuring a robust and fair comparison with other state-of-the-art techniques. The IGFL model was optimized using the Adam optimizer with a learning rate of 0.001, specifically tuned for the label segmentation task to enhance convergence stability. The training process was conducted over 100 epochs on an NVIDIA Tesla T4 GPU, equipped with 16GB of memory, which provided the necessary computational resources for efficient processing of the datasets. These fundamental hyperparameters were carefully chosen through preliminary experimentation to ensure optimal performance across our experiments.

5.3. Experimental results

5.3.1. Comparison results with baseline methods

Overall, the results in Table 2 indicate that all models perform adequately in both classification and segmentation, which aligns with the general understanding of few-shot learning. Notably, in the 1-way 1-shot setting, the classification performance of IG-FSL (84.40% ER) and ASNet (83.13% ER) stands out, surpassing the other methods, while PFENet and PANet show relatively lower classification performance (77.57% and 74.70%, respectively). Similarly, segmentation (mIoU) results follow the same trend, with IG-FSL and ASNet achieving the highest scores at 46.00% and 45.16%, respectively.

In the 2-way 1-shot case, all methods exhibit a decline in both classification and segmentation performance, which is expected, as the models are required to make predictions

Table 2: The performance results of various methods in two experimental settings: 1-way 1-shot and 2-way 1-shot, were evaluated on classification (0/1 ER) and segmentation (mIoU) metrics.

Performance results				
methods	1-way 1-shot		2-way 1-shot	
metric	cls. 0/1 ER	seg. mIoU	cls. 0/1 ER	seg. mIoU
PANet	74.70	40.10	62.62	39.78
PFENet	77.57	40.82	62.40	40.22
HSNet	82.11	43.78	63.79	41.52
ASNet	83.13	44.16	63.40	42.35
IG-FSL	84.40	46.00	65.00	43.05

with more options, making the task more challenging. However, IG-FSL still leads with a classification ER of 65.00% and a segmentation mIoU of 43.42%. ASNet follows closely with 64.40% ER and 43.05% mIoU. In contrast, PANet and PFENet show the largest performance drops, particularly in segmentation, with PFENet’s mIoU decreasing to 38.22%.

Additionally, we observe a substantial performance drop when comparing the 1-way and 2-way cases. In the 1-way classification setting, the model only needs to predict whether the query instance belongs to the single support class or not, effectively a binary decision. In contrast, the 2-way classification task introduces an additional option, requiring the model to distinguish between two support classes, which inherently increases complexity and reduces accuracy. Similarly, in segmentation tasks, the model must differentiate between two distinct objects in the 2-way case, as opposed to a single object in the 1-way setting. This added complexity provides a clear rationale for the observed performance decline across both classification and segmentation metrics.

Overall, IG-FSL and ASNet demonstrate superior performance across both experimental settings and metrics, while PFENet and PANet trail behind, particularly in the more complex 2-way 1-shot setting.

5.3.2. A comparative analysis based on the number of ways

In this experiment, we systematically varied the number of ways to assess its impact on model performance. For the 1-way scenario, we selected silicosis lung disease as the target for the test set. When the number of ways exceeds one ($N > 1$), specifically for $N \in \{2, 3, 4\}$, we partition the training set into $N - 1$ lung disease classes, reserving the remaining classes for testing.

As illustrated in Figure 2(a), the superiority of IG-FSL in segmentation tasks becomes increasingly apparent as the number of ways rises. This finding underscores the robustness of IG-FSL across varying levels of complexity, although some fluctuations in performance are anticipated due to changes in the number of classes within the test set.

Figure 2(b) further demonstrates that under the 1-way 1-shot configuration, the IG-FSL method achieves the highest performance, boasting an accuracy of 84.4%. However, as we increase the number of ways from 1 to 4, a noticeable decline in both classification (ER) and segmentation (mIoU) metrics is observed. This trend highlights the challenges associated with multi-class prediction, where increased complexity can hinder performance. Nevertheless, it is noteworthy that IG-FSL consistently outperforms all other methods throughout these variations, reaffirming its effectiveness and reliability in diverse scenarios.

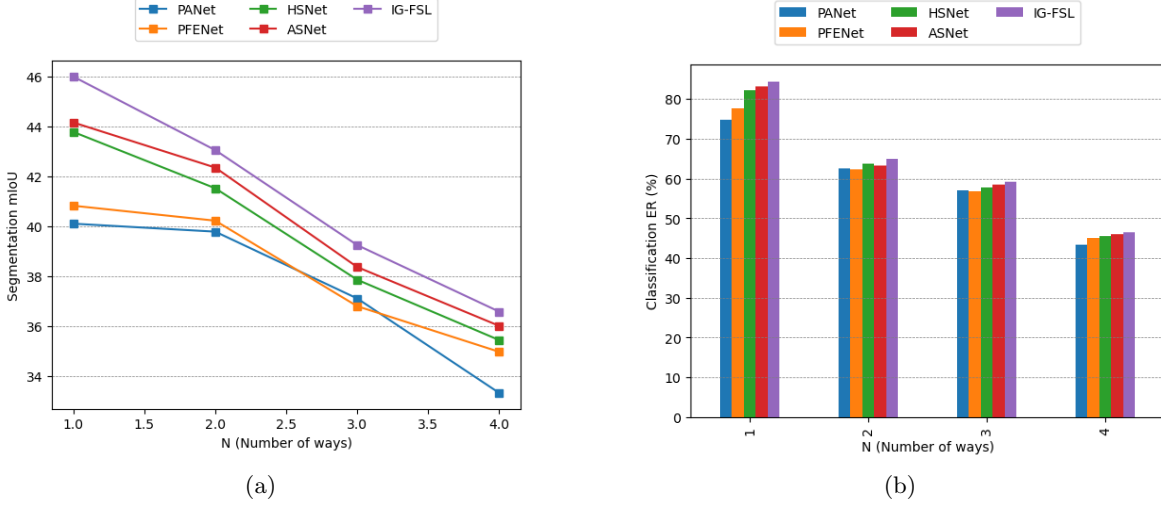


Figure 2: Few-shot learning results in varying number of ways (a) and comparisons (b)

5.3.3. Comparison between all competitive schemes with respect to the number of shots

Table 3: Comparisons on various methods based on classification (0/1 ER) and segmentation (mIoU) metrics across different shot configurations (1-shot to 4-shot).

metric	cls. 0/1 ER				seg. mIoU			
methods	1-shot	2-shot	3-shot	4-shot	1-shot	2-shot	3-shot	4-shot
PANet	74.70	74.91	76.42	78.53	40.10	40.98	42.72	44.84
PFENet	77.57	78.42	79.21	79.97	40.82	41.17	43.34	45.11
HSNet	82.11	83.78	85.34	86.74	43.78	44.44	45.02	46.29
ASNet	83.13	84.22	86.37	88.36	44.16	45.74	46.33	47.55
IG-FSL	84.46	86.09	87.66	89.52	46.00	46.27	47.32	47.89

Overall, performance improves with an increasing number of shots for all methods, highlighting the benefits of additional training examples. PANet shows moderate gains, with classification error decreasing from 74.70% to 78.53% and segmentation mIoU rising from 40.10% to 44.84%. PFENet follows closely, improving from 77.57% to 79.97% in classification and from 40.82% to 45.11% in segmentation.

HSNet and ASNet exhibit more substantial improvements, particularly in the 3-shot and 4-shot scenarios, with ASNet reaching 88.36% accuracy and 47.55% mIoU in the 4-shot case. IG-FSL stands out as the top performer, with classification error dropping to 89.52% and segmentation mIoU increasing to 47.89% in the 4-shot setting.

With an increase in the number of shots, meaning a higher number of support examples per class, each model is provided with a more extensive set of examples for each class. This facilitates learning of essential features, enabling the models to better recognize variations within each class. Additional examples help the models identify shared characteristics while filtering out irrelevant noise, thereby enhancing generalization capability. This trend is consistently observed across all five models in both classification and segmentation evaluations.

However, increasing the shot count also raises computational demands, as the model must perform additional comparisons between input images and the support examples, which, in turn, increases processing time.

In summary, all methods benefit from increased shots, but IG-FSL, along with ASNet and HSNet, demonstrates the most significant performance gains, reaffirming their competitiveness in few-shot learning.

5.3.4. Qualitative study

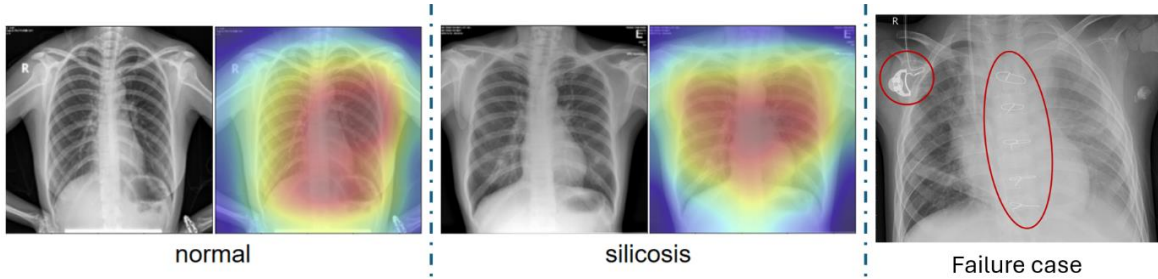


Figure 3: GradCam visualization of a normal case and a silicosis case; and an example image of a failure case.

GradCAM analysis highlights the model’s focus on lung regions, particularly small white nodules in silicosis cases, aiding in disease detection. However, misclassified cases reveal limitations, as non-diagnostic artifacts like medical devices can introduce noise, affecting predictions. This underscores the model’s sensitivity to irrelevant features, suggesting the need for further refinement or preprocessing. Overall, GradCAM is a valuable tool for improving model interpretability and diagnostic reliability.

6. CONCLUSIONS

This paper explored the application of Few-Shot Learning (FSL) to address the data scarcity issue in diagnosing occupational silicosis using chest X-ray (CXR) images. Our approach successfully demonstrated that FSL can yield accurate results even with a limited number of labeled samples, making it a promising technique for data-constrained environments.

The experimental outcomes illustrate the strength of the FSL model in handling both classification and segmentation tasks. In the 1-shot setting, the model reached an accuracy of 84.4% and a mean Intersection over Union (mIoU) of 46.0%, while in the 4-shot scenario, it achieved 89.52% accuracy and 47.89% mIoU. These results confirm the viability of FSL in improving diagnostic accuracy in contexts where acquiring large datasets is challenging.

While promising, our study has limitations that warrant further exploration. One notable limitation is the dependency on the quality and representativeness of the few labeled samples, which, if unrepresentative, could hinder the model’s generalizability. Additionally, FSL’s effectiveness across a broader range of medical conditions or image modalities remains to be thoroughly evaluated. Future research could focus on incorporating diverse data sources, such as combining CXR images with clinical data, or refining model architectures to

further enhance performance. Investigating these avenues could improve FSL’s adaptability to various diagnostic challenges, thus strengthening its potential as an effective, scalable solution in medical diagnostics, especially for data-limited scenarios.

REFERENCES

- [1] F. Brims, D. Chambers, D. Deller, G. Edwards, R. Hoy, C. Jones, R. Knight, D. Yates, C. Kela-her, J. Shepherd *et al.*, “National Guidance for doctors assessing workers exposed to respirable crystalline silica dust with specific reference to the occupational respiratory diseases associated with engineered stone,” Tech. Rep., 2022.
- [2] E. K. Austin, C. James, and J. Tessier, “Early detection methods for silicosis in Australia and internationally: A review of the literature,” *International Journal of Environmental Research and Public Health*, vol. 18, no. 15, p. 8123, Jul. 2021.
- [3] A. U. Ibrahim, M. Ozsoz, S. Serte, F. Al-Turjman, and P. S. Yakoi, “Pneumonia classification using deep learning from chest X-ray images during COVID-19,” *Cognitive Computation*, vol. 16, no. 4, pp. 1589–1601, Jan. 2021.
- [4] Y. Wang, J. Wu, and D. Wu, “A survey on the application of convolutional neural networks in the diagnosis of occupational pneumoconiosis,” *Journal of Biomedical Engineering*, vol. 41, no. 2, pp. 413–420, 2024.
- [5] R. Kundu, R. Das, Z. W. Geem, G.-T. Han, and R. Sarkar, “Pneumonia detection in chest X-ray images using an ensemble of deep learning models,” *PLOS ONE*, vol. 16, no. 9, p. e0256630, Sep. 2021.
- [6] H.-Y. Chiu, R. H.-T. Peng, Y.-C. Lin, T.-W. Wang, Y.-X. Yang, Y.-Y. Chen, M.-H. Wu, T.-H. Shiao, H.-S. Chao, Y.-M. Chen, and Y.-T. Wu, “Artificial intelligence for early detection of chest nodules in X-ray images,” *Biomedicines*, vol. 10, no. 11, p. 2839, Nov. 2022.
- [7] S. R. Teixeira and A. Naves, “Chest X-ray: an examination that has been in use for centuries but is still essential, especially in the clinical management of newborns in the neonatal intensive care unit,” *Radiologia Brasileira*, vol. 51, no. 1, pp. VII–VIII, Feb. 2018.
- [8] G. Tourassi, S. Voisin, V. Paquit, and E. Krupinski, “Investigating the link between radiologists’ gaze, diagnostic decision, and image content,” *Journal of the American Medical Informatics Association*, vol. 20, no. 6, pp. 1067–1075, Nov. 2013.
- [9] M. L. Giger and K. Suzuki, “Computer-Aided Diagnosis,” in *Biomedical Information Tech-nology*. Elsevier, 2008, pp. 359–XXII.
- [10] Y. Lee, Y. Chae, and S. Jeon, “Integration and evaluation of clinical decision support systems for diagnosis idopathics pulmonary fibrosis (IPF),” *Healthcare Informatics Research*, vol. 16, no. 4, p. 260, 2010.
- [11] R. Najjar, “Redefining radiology: A review of artificial intelligence integration in medical imag-ing,” *Diagnostics*, vol. 13, no. 17, p. 2760, Aug. 2023.

- [12] S. Albahli, H. T. Rauf, A. Algosaiibi, and V. E. Balas, "AI-driven deep CNN approach for multi-label pathology classification using chest X-Rays," *PeerJ Computer Science*, vol. 7, p. e495, Apr. 2021.
- [13] T. B. Chandra and K. Verma, "Pneumonia detection on chest X-Ray using machine learning paradigm," in *Proceedings of 3rd International Conference on Computer Vision and Image Processing*. Springer Singapore, Nov. 2019, pp. 21–33.
- [14] K. M. Kuo, P. C. Talley, C. H. Huang, and L. C. Cheng, "Predicting hospital-acquired pneumonia among schizophrenic patients: a machine learning approach," *BMC Medical Informatics and Decision Making*, vol. 19, no. 1, Mar. 2019.
- [15] H. Yue, Q. Yu, C. Liu, Y. Huang, Z. Jiang, C. Shao, H. Zhang, B. Ma, Y. Wang, G. Xie, H. Zhang, X. Li, N. Kang, X. Meng, S. Huang, D. Xu, J. Lei, H. Huang, J. Yang, J. Ji, H. Pan, S. Zou, S. Ju, and X. Qi, "Machine learning-based CT radiomics method for predicting hospital stay in patients with pneumonia associated with SARS-CoV-2 infection: a multicenter study," *Annals of Translational Medicine*, vol. 8, no. 14, pp. 859–859, Jul. 2020.
- [16] T. Meraj, H. T. Rauf, S. Zahoor, A. Hassan, M. I. Lali, L. Ali, S. A. C. Bukhari, and U. Shoaib, "Lung nodules detection using semantic segmentation and classification with optimal features," *Neural Computing and Applications*, vol. 33, no. 17, pp. 10 737–10 750, May 2020.
- [17] V. Rajinikanth, S. Kadry, R. Damaševičius, D. Taniar, and H. T. Rauf, "Machine learning scheme to detect choroidal-neovascularization in retinal OCT image," in *2021 Seventh International Conference on Bio Signals, Images, and Instrumentation (ICBSII)*. IEEE, Mar. 2021, pp. 1–5.
- [18] S. Kadry, Y. Nam, H. T. Rauf, V. Rajinikanth, and I. A. Lawal, "Automated detection of brain abnormality using deep learning scheme: A study," in *2021 Seventh International Conference on Bio Signals, Images, and Instrumentation (ICBSII)*. IEEE, Mar. 2021, pp. 1–5.
- [19] V. Rajinikanth, S. Kadry, D. Taniar, R. Damaševičius, and H. T. Rauf, "Breast cancer detection using thermal images with Marine-Predators-Algorithm selected features," in *2021 Seventh International Conference on Bio Signals, Images, and Instrumentation (ICBSII)*. IEEE, Mar. 2021, pp. 1–6.
- [20] H. Sharma, J. S. Jain, P. Bansal, and S. Gupta, "Feature extraction and classification of chest X-Ray images using CNN to detect pneumonia," in *2020 10th International Conference on Cloud Computing, Data Science and Engineering (Confluence)*. IEEE, Jan. 2020, pp. 227–231.
- [21] O. Stephen, M. Sain, U. J. Maduh, and D.-U. Jeong, "An efficient deep learning approach to pneumonia classification in healthcare," *Journal of Healthcare Engineering*, vol. 2019, pp. 1–7, Mar. 2019.
- [22] P. Rajpurkar, "CheXNet: Radiologist-level pneumonia detection on chest X-Rays with deep learning," *ArXiv abs/1711*, vol. 5225, 2017.

- [23] A. G. Taylor, C. Mielke, and J. Mongan, “Automated detection of moderate and large pneumothorax on frontal chest X-rays using deep convolutional neural networks: A retrospective study,” *PLOS Medicine*, vol. 15, no. 11, p. e1002697, Nov. 2018.
- [24] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. Rajendra Acharya, “Automated detection of COVID-19 cases using deep neural networks with X-ray images,” *Computers in Biology and Medicine*, vol. 121, p. 103792, Jun. 2020.
- [25] R. E. Al Mamlook, S. Chen, and H. F. Bzizi, “Investigation of the performance of machine learning classifiers for pneumonia detection in chest X-ray images,” in *2020 IEEE International Conference on Electro Information Technology (EIT)*. IEEE, Jul. 2020, pp. 098–104.
- [26] J. Vanschoren, “Meta-learning,” in *Automated Machine Learning*. Springer International Publishing, 2019, pp. 35–61.
- [27] M. Woodward and C. Finn, “Active one-shot learning,” *arXiv preprint arXiv:1702.06559*, 2017.
- [28] B. Lake, R. Salakhutdinov, J. Gross, and J. Tenenbaum, “One shot learning of simple visual concepts,” in *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 33, no. 33, 2011.
- [29] L. Fei-Fei, R. Fergus, and P. Perona, “One-shot learning of object categories,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 4, pp. 594–611, Apr. 2006.
- [30] C. H. Lampert, H. Nickisch, and S. Harmeling, “Attribute-based classification for zero-shot visual object categorization,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 3, pp. 453–465, Mar. 2014.
- [31] H. Larochelle, D. Erhan, and Y. Bengio, “Zero-data learning of new tasks,” in *AAAI*, vol. 1, no. 2, p. 3, 2008.
- [32] M. Rohrbach, M. Stark, and B. Schiele, “Evaluating knowledge transfer and zero-shot learning in a large-scale setting,” in *CVPR 2011*. IEEE, Jun. 2011.
- [33] Z. Ding, M. Shao, and Y. Fu, “Low-rank embedded ensemble semantic dictionary for zero-shot learning,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Jul. 2017, pp. 6005–6013.
- [34] E. Pachetti and S. Colantonio, “A systematic review of few-shot learning in medical imaging,” *Artificial Intelligence in Medicine*, vol. 156, p. 102949, Oct. 2024.
- [35] A. A. G. Pramana, F. I. Permana, M. F. Maulana, and D. R. Fudholi, “Few-shot learning approach on tuberculosis classification based on chest X-Ray images,” *arXiv preprint arXiv:2409.11644*, 2024.
- [36] A. Galán-Cuenca, A. J. Gallego, M. Saval-Calvo, and A. Pertusa, “Few-shot learning for COVID-19 chest X-ray classification with imbalanced data: an inter vs. intra domain study,” *Pattern Analysis and Applications*, vol. 27, no. 3, Jun. 2024.

- [37] X. Wang, Y. Yuan, D. Guo, X. Huang, Y. Cui, M. Xia, Z. Wang, C. Bai, and S. Chen, “SSA-Net: Spatial self-attention network for COVID-19 pneumonia infection segmentation with semi-supervised few-shot learning,” *Medical Image Analysis*, vol. 79, p. 102459, Jul. 2022.
- [38] H. Xie, C. Gu, W. Zhang, J. Zhu, J. He, Z. Huang, J. Zhu, and Z. Xu, “A few-shot learning framework for the diagnosis of osteopenia and osteoporosis using knee X-ray images,” *Journal of International Medical Research*, vol. 52, no. 9, Sep. 2024.
- [39] A. Li, T. Luo, T. Xiang, W. Huang, and L. Wang, “Few-shot learning with global class representations,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 9715–9724.
- [40] G. S. Dhillon, P. Chaudhari, A. Ravichandran, and S. Soatto, “A baseline for few-shot image classification,” *arXiv preprint arXiv:1909.02729*, 2019.
- [41] D. Chen, Y. Chen, Y. Li, F. Mao, Y. He, and H. Xue, “Self-supervised learning for few-shot image classification,” in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, Jun. 2021, pp. 1745–1749.
- [42] A. Majee, K. Agrawal, and A. Subramanian, “Few-shot learning for road object detection,” in *AAAI Workshop on Meta-Learning and Metadl Challenge*. PMLR, 2021, pp. 115–126.
- [43] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, “Path aggregation network for instance segmentation,” in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, Jun. 2018.
- [44] Z. Tian, H. Zhao, M. Shu, Z. Yang, R. Li, and J. Jia, “Prior guided feature enrichment network for few-shot segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 2, pp. 1050–1065, Feb. 2022.
- [45] J. Min, D. Kang, and M. Cho, “Hypercorrelation squeeze for few-shot segmenation,” in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE, Oct. 2021.
- [46] D. Kang and M. Cho, “Integrative few-shot learning for classification and segmentation,” in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Jun. 2022.

Received on October 05, 2024

Accepted on November 28, 2024