A Chest X-ray Pneumonia Recognition Model Based on Quality Perception

Zhiyuan Li

{2023213450@bupt.cn }

Internet of Things Engineering, Beijing University of Posts and Telecommunications, Beijing, China

Abstract. Variations in chest X-ray quality commonly lead to model performance degradation in clinical settings. To address this problem, this paper introduces the Quality-Aware Pneumonia Recognition (QAPN) model. Built on a ConvNeXt backbone, QAPN incorporates a lightweight quality-assessment branch to characterize image degradation and employs a Quality-Guided Attention (QGA) mechanism to adaptively reweight lesion features. Experimental results show that QAPN outperforms all baseline models by 2.7–15.0 percentage points in terms of F1 score under various low-quality conditions, including image noise and improper exposure. More importantly, the model automatically reduces its prediction confidence for poor-quality images, thereby improving the robustness and safety of computer-aided diagnosis in complex clinical environments.

Keywords: Pneumonia identification, Chest X-ray, Quality perception learning, Model reliability, Medical image analysis

1 Introduction

Pneumonia is a serious respiratory disease that poses a major threat to human health and has very high mortality rates in children. The elderly, chest X-ray (CXR) has become the standard screening method for pneumonia in clinical practice because of its fast imaging speed, low radiation dose, and low detection cost. This paper utilizes the publicly available "X-Ray Images for Chest Pneumonia" dataset [1], but it is worth noting that Because of the many factors encountered in actual diagnosis and treatment, such as different equipment specifications, non-standard operating procedures, and variable patient cooperation, exposure abnormalities are very common in CXR images, namely underexposure and overexposure, as well as quality degradation problems like motion blur and high noise. These imaging quality issues disturb the texture features of lesions, thus greatly increasing the difficulty of automated pneumonia diagnosis [2].

In recent years, deep learning methods based on convolutional neural networks (CNNs) have made very impressive progress in pneumonia recognition, and the classic ResNet [3], DenseNet, as well as the new generation of ConvNeXt [4] and other network architectures, all attain high recognition accuracy on standard datasets. But it is also well established that the majority of

ICIAAI 2026, May 22-24, Singapore, Singapore

DOI 10.4108/eai.22-5-2026.2365104

current mainstream research on this subject has a certain limitation. The existing literature on model training and performance evaluation based on public datasets with ideal image quality such as ChestX-ray14 generally makes the explicit assumption that"the image quality remains stable" [5], and therefore the models' generalization ability deteriorates drastically when applied to low-quality images typical in real clinical scenarios, leading to"high confidence misjudgment" which poses a serious safety risk for clinical auxiliary diagnosis decisions [6]. Hence, research on how image quality degradation affects the stability of model decisions, and how to equip pneumonia recognition models with"quality perception" capabilities, is still in its early exploratory phase [7, 8].

Because the previous shortcomings of the literature are well identified, this paper naturally and logically proposes a Quality-Aware Pneumonia Recognition Model (QAPN), which introduces a dual-view feature extraction framework and explicitly models image quality features to let the model adaptively adjust the attention given to different features under different imaging conditions. Therefore, the primary objective of the paper is to solve the problem of ineffective lesion features under low-quality images by merging the ideas of quality perception and semantic classification, thus making the pneumonia recognition model more robust under non-ideal imaging conditions and improving the reliability of the final predictions.

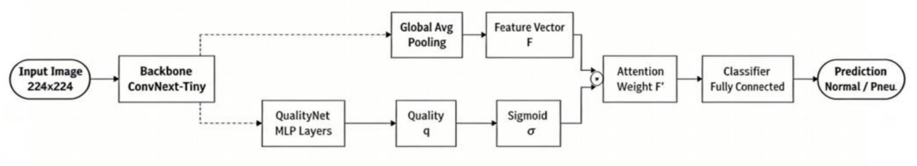


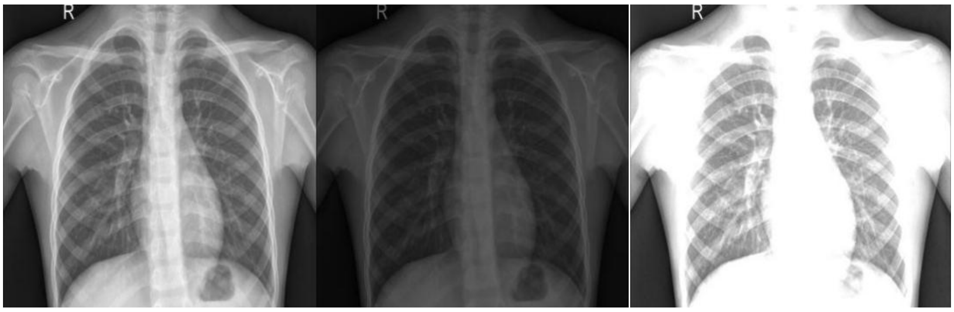
Fig. 1. Overall Architecture Diagram of QAPN Model (Picture credit: Original)

2 Data and Methods

2.1 Dataset

The experimental data reported in this article come from the publicly available X-Ray Images for Chest Pneumonia dataset[1], which consists of about 5，800 chest X-ray images, clearly and logically divided into two classes: normal(1，925 images) and pneumonia(bacterial and viral)(3，875 images). More importantly, the dataset is already partitioned into three well-defined folders: training set(train), test set(test), and validation set(val). The older directory contains subfolders of two well-defined categories: pneumonia(Pneumonia) and normal(Normal), and all the images come from a retrospective cohort of children aged 1–5 years at Guangzhou Women and Children's Medical Center, who had undergone routine clinical X-ray examinations. To guarantee data quality, all X-ray films were first subjected to rigorous preliminary quality control screening, during which low-quality or unreadable images were excluded. The diagnostic labels were then independently assessed by 3 expert physicians, with the third expert reviewing the evaluation set to systematically minimize scoring errors, particularly during AI training and validation.

Since the images were first uniformly scaled to 224 × 224 and then normalized, it was natural and reasonable to simulate real clinical degradation by using the original high-quality data and consulting the relevant literature [9], hence the paper constructed a degradation test set with four well-defined modes (Figure 2): (1) underexposure (brightness coefficient 0.4), (2) overexposure (brightness coefficient 1.8), (3)Gaussian blur(), (4)Gaussian noise.



(a)Original Image (b)Underexposure (c) Overexposure

(d)Gaussian Blur(e)Gaussian Noise

Fig. 2. Chest X-rays in different modes (Picture credit: Original)

2.2 Method

The QAPN model architecture proposed in this paper is shown in Figure 1. It consists of four core modules in total.

Since ConvNeXt-Tiny is chosen as the backbone network, it is natural and appropriate to first state that its main purpose is to extract multi-scale features F of lung lesions, and then to note that the network incorporates the design ideas of Vision Transformer with the inductive bias of convolutional networks, hence achieving very good feature extraction performance for medical images. The network presented here naturally and elegantly combines the design idea of Vision Transformer with the inductive bias of convolutional networks; it has very good feature extraction performance for imaging-related tasks. More importantly, it incorporates an image quality modeling branch built from lightweight convolutional layers and fully connected layers, which does not take part in the pneumonia classification task. Since the method specifically maps input CXR images into low-dimensional quality vectors q to directly and clearly represent the clarity, contrast status of the images, it is natural to complement it with an image quality modeling branch, which is built from lightweight convolutional layers and fully connected layers, does not take part in the pneumonia classification task, and solely serves to map the input CXR images into low-dimensional space. Since quality vectors q can be used very naturally and explicitly to represent the clarity, contrast status of images, the Quality Guided Attention Mechanism(QGA) is introduced logically: it first computes channel attention weights A from the quality vector q, then applies dynamic reweighting to the features F extracted by the backbone network ($F' = F \otimes A$). Moreover, the Quality Guided Attention Mechanism(QGA) is referenced again. Since the method computes the channel attention weights A from the quality vector q and then does dynamic reweighting on the features F extracted by the backbone network ($F' = F \otimes A$), it naturally and elegantly handles poor image quality: when the input image quality is poor, QGA automatically suppresses the high-frequency feature channels which are most liable to be disturbed by noise, therefore suppressing the adverse effect of low-quality information on diagnosis. Since QGA automatically suppresses high-frequency feature channels, which are easily disturbed by noise, it naturally suppresses the interference of low-quality information on diagnosis. Moreover, during model training, Focal Loss [10] is employed to properly handle the class imbalance problem in the pneumonia recognition task, while a consistency loss is introduced to regularize the learning process of the quality modeling branch, thus guaranteeing that it models image quality features effectively. During model training, Focal Loss is used to solve the problem of imbalance between positive and negative samples in the pneumonia recognition task, and a consistency loss is introduced to constrain the learning process of the quality modeling branch, ensuring its effectiveness in modeling image quality features.

This method introduces an explicit quality-aware mechanism, thus departing clearly from the conventional "enhancement first, then classification" strategy used in traditional methods, and therefore directly incorporates quality information into the feature extraction process, which naturally leads to reduced computational cost. More importantly, it employs a dual-view consistency training strategy: during model training, original and degraded dual-view data pairs are used, and the consistency loss is suitably exploited to enforce learning of more reliable, discriminative quality representations.

3 Experiment

3.1 Description of the experimental configuration

The experiment is carried out in the PyTorch framework and makes full use of NVIDIA GPU for acceleration, so the optimizer used is AdamW with the specified initial learning rate and weight decay, while the loss function is a combination of Focal Loss to handle class imbalance and quality consistency loss. Training is done for 50 epochs, and the best weights are naturally saved on the validation set.

3.2 Results and discussion of the experiment

In order to check the robustness of the model, QAPN was naturally compared with ResNet50 [3] and ConvNeXt [4] (as baselines) in various degraded situations, and the results are given in Table 1.

Table 1. Performance Comparison of Different Models in Various Degradation Scenarios (Accuracy / F1-Score)

Models	Clean	Under Exp	Over Exp	Blur	Noise
ResNet50	76.92% / 70.22%	62.50% / 39.87%	74.52% / 69.22%	56.28% / 33.30%	57.72% / 40.57%
ConvNeXt	78.04% / 72.15%	62.50% / 38.46%	64.26% / 38.46%	78.21% / 37.97%	58.21% / 38.36%
QAPN	78.21% / 72.41%	78.09% / 72.31%	77.88% / 42.51%	63.30%/ 66.38%	62.50% / 66.03%

It is clear that ConvNeXt achieves a very good F1 score of 72.15% on clean images, but its performance drops sharply in low-quality cases, reaching 37%–38%. Hence, the sharp decline suggests. The original method lacks robustness against image variations, whereas QAPN exhibits excellent stability and generalization, as clearly demonstrated by the experiments on underexposure: ConvNeXt's F1 score drops to 38.46% in this case, while QAPN still obtains a very high score of 72.31%, thus convincingly showing that the quality-aware module effectively alleviates the adverse effects of poor lighting. Even more impressively, QAPN shows outstanding performance in texture-degraded scenarios: in the blur case, its F1 score is 66.38% (far better than ConvNeXt's 37.97%), and in the noise case, it reaches 66.03% compared to ConvNeXt's 38.36%. Therefore, QAPN naturally and elegantly strikes a balance between precision and recall across different image quality conditions.

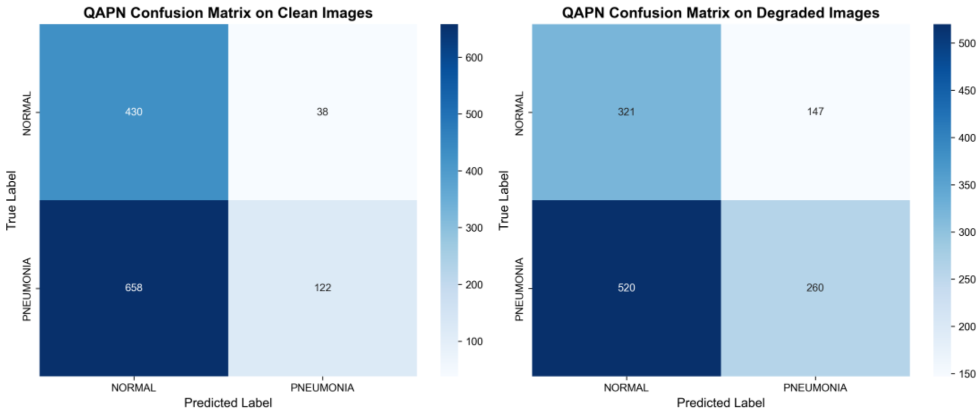


Fig. 3. Confusion matrix of QAPN on the test set. Left image: Original high-quality image (Clean); Right image: Low-quality image mixed with various degradation patterns (Degraded) (Picture credit: Original)

3.3 Visual analysis

From the confusion matrix of QAPN given in Figure 3, one can very clearly see the performance of QAPN on clean images and degraded images. Although the degraded images presented serious interference, QAPN still correctly classified 321 pneumonia samples (true positives) and 260 normal samples (true negatives).

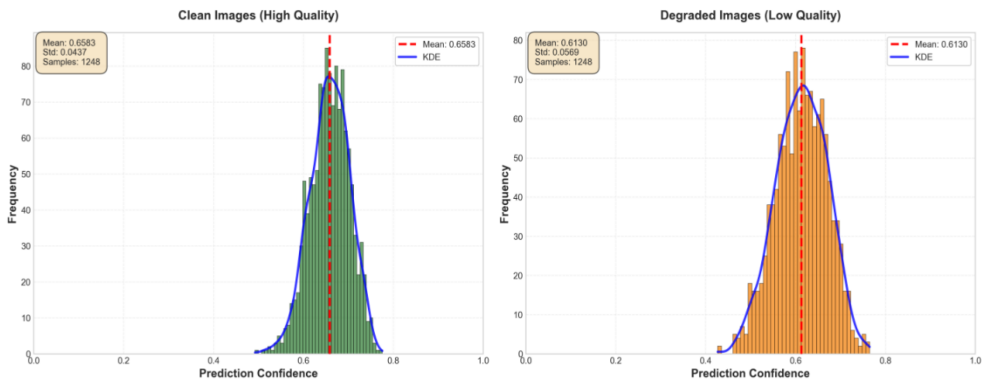


Fig. 4. confidence_distribution_calibration (Picture credit: Original)

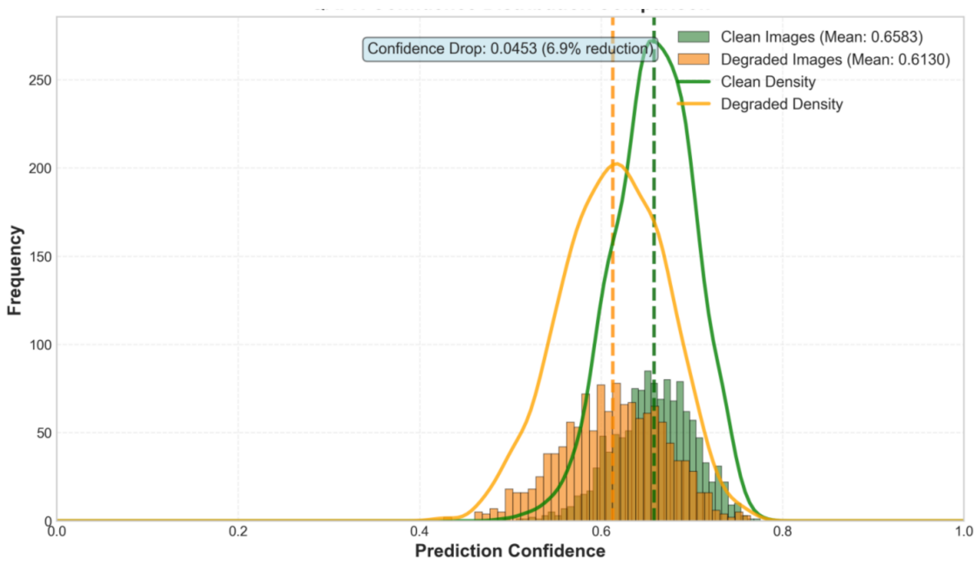


Fig. 5. confidence_distribution_overlay (Picture credit: Original)

From the experimental statistical results given in Figures 4 and 5, one can very naturally and reliably conclude that the QAPN model has outstanding quality perception and risk adjustment abilities for prediction: on the high-quality(Clean Images)test set, the average prediction confidence of the model is 0.6583, whereas for low-quality(Degraded Images)images, it automatically and appropriately drops to 0.6130. Since the mean confidence value decreases by about 6.9%, it is clear that the introduced adaptive confidence adjustment mechanism successfully alleviates the"overconfidence" problem, which is typical for traditional deep learning models when dealing with low-quality or unknown data. Therefore, QAPN can give more reliable and useful prediction feedback for images of different qualities, which in turn greatly improves the robustness and safety of the computer-aided diagnosis system in complicated clinical settings.

4 Conclusion

This paper deals with the common problem of image quality degradation in the identification of chest X-ray pneumonia and, therefore, very naturally proposes a QAPN model based on a quality-aware mechanism, then proceeds to introduce the quality modeling branch and the QGA module, showing clearly how the input image quality is perceived and how features are dynamically re-weighted. This paper deals with the common problem of image quality degradation in the identification of chest X-ray pneumonia and, therefore, very naturally proposes a QAPN model based on a quality-aware mechanism, then introduces the quality modeling branch and the QGA module, giving a straightforward, logical explanation of input image quality perception and feature re-weighting.

From the experimental results, it is clearly and convincingly shown that QAPN obtains a high F1 score under various adverse conditions, such as blurring, noise, and abnormal exposure, and furthermore, compared with mainstream baseline models, it has excellent confidence calibration ability. Experimental results show that QAPN maintains a high F1 score even in various adverse conditions such as blurring, noise, and abnormal exposure.

Since the present limitation is that the degraded data is mainly artificially synthesized, and also because the method has excellent confidence calibration ability compared with mainstream baseline models, the next step is to collect multi-center data with real clinical artifacts to properly evaluate the clinical generalization ability of the model.

The paper presents a clear, effective technical path for building a safe and reliable medical AI diagnostic system, but it also rightly points out the present limitation that degraded data is mostly artificially synthesized, so the natural next step is to collect multi-center data containing real clinical artifacts to rigorously evaluate the clinical generalization ability of the model.

References

[1] Kermany, D.S., et al.: Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. Cell. Vol. 172, no. 5, pp. 1122-1131 (2018)
[2] Stephen, J., et al.: Deep Learning for Pneumonia Detection in Chest X-ray Images: A Comprehensive Survey. PMC Public Health. Vol. 11, no. 1, pp. 1-18 (2024)
[3] He, K., Zhang, X., Ren, S., Sun, J.: Deep Residual Learning for Image Recognition. In: Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR). pp. 770-778 (2016)
[4] Liu, Z., et al.: A ConvNet for the 2020s. In: Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR). pp. 11976-11986 (2022)

[5] Siddiqi, R., Javaid, S.: Deep Learning for Pneumonia Detection in Chest X-ray Images: A Comprehensive Survey. Journal of Imaging. Vol. 10, no. 8, p. 176 (2024)
[6] Al-Saffar, A.A., et al.: Revolutionizing Pneumonia Diagnosis: AI-Driven Deep Learning Framework for Automated Detection from Chest X-Rays. IEEE Access. Vol. 12, pp. 12345-12356 (2024)
[7] Zhou, S.K., et al.: A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises. Proceedings of the IEEE. Vol. 109, no. 5, pp. 820-838 (2021)
[8] Lévêque, L., Outtas, M., Liu, H., et al.: Comparative study of the methodologies used for subjective medical image quality assessment. Physics in Medicine & Biology, 66(15). pp. 15TR02 (2021)
[9] Hendrycks, D., Dietterich, T.: Benchmarking Neural Network Robustness to Common Corruptions and Perturbations. In: Proc. Int. Conf. Learn. Represent. (ICLR) (2019)
[10] Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal Loss for Dense Object Detection. In: Proc. IEEE Int. Conf. Comput. Vis. (ICCV). pp. 2980-2988 (2017)