Automated Quality Evaluation of Panoramic Dental Radiographs Using a Domain-Adapted Transfer Learning

Nur Nafiiyah[1]*, Rifky Aisyatul Faroh[2], Eha Renwi Astuti[3], Rini Widyaningrum[4],
Agus Harjoko[5], Kang-Hyun Jo[6], Alhidayati Asymal[7], Youan Nhareswary Dwike Prasetya[8]
Department of Informatics, Universitas Islam Lamongan, Indonesia[1, 8]
Electrical Engineering, Universitas Islam Lamongan, Indonesia[2]
Faculty of Dental Medicine-Department of Dentomaxillofacial Radiology, Universitas Airlangga, Indonesia[3, 7]
Faculty of Dentistry-Department of Dentomaxillofacial Radiology, Universitas Gadjah Mada, Indonesia[4]
Department of Computer Science and Electronics, Universitas Gadjah Mada, Yogyakarta, Indonesia[5]
Department of Electrical-Electronic and Computer Engineering, University of Ulsan, Ulsan, South Korea[6]

***Abstract*—Assessing the quality of panoramic dental radiographs is essential to ensure diagnostic accuracy and patient safety. However, existing CNN-based approaches for radiograph quality assessment often emphasize architectural comparisons, while providing limited discussion on training stability and generalization, particularly when applied to relatively small and heterogeneous datasets. To address this gap, this study proposes a transfer learning-based framework that integrates Global Average Pooling (GAP) and Batch Normalization (BN) to enhance feature robustness and reduce overfitting in panoramic dental radiograph quality classification. Three pretrained CNN architectures: ResNet50, VGG16, and VGG19 were evaluated using panoramic radiographs collected from two tertiary hospitals in Indonesia. Experimental results using *k*-fold cross-validation indicate that the proposed GAP+BN refinement improves classification consistency across models, with VGG16 demonstrating the most stable and reliable performance. These findings suggest that domain-adapted transfer learning with appropriate feature aggregation and normalization can support the development of automated and clinically reliable quality assurance systems for panoramic dental imaging.**

***Keywords*—Batch Normalization; image quality; panoramic radiograph; transfer learning**

I. Introduction

Panoramic radiography is one of the most widely used imaging modalities in dentistry because it provides comprehensive visualization of maxillary and mandibular structures with relatively low radiation exposure. However, the diagnostic reliability of panoramic radiographs is strongly dependent on image quality, which can be adversely affected by patient positioning errors, improper exposure parameters, and technical disturbances. Poor-quality radiographs reduce diagnostic accuracy and increase the likelihood of retakes, thereby exposing patients to unnecessary additional radiation. Consequently, systematic quality assessment of panoramic radiographs is a critical aspect of dental radiology practice.

Previous studies have evaluated panoramic radiograph quality from both technical and diagnostic perspectives. Comparative analysis of different imaging devices has shown relatively small variations in image quality, indicating that multiple systems may be clinically acceptable [1],[2]. Other investigations reported that positioning errors, including head tilt and patient movement, occur in up to 95% of panoramic radiographs and represent the primary contributors to reduced diagnostic quality [3]. Although these studies provide valuable insights into common sources of image degradation, quality assessment in routine clinical practice still relies heavily on manual evaluation by operators and radiologists, which is inherently subjective and time-consuming.

Several studies have emphasized the role of operator training and quality control in improving panoramic radiograph quality. Investigations conducted in South Wales and other dental institutions revealed that a large proportion of panoramic radiographs failed to meet national quality standards, primarily due to patient positioning errors and insufficient operator instructions [4],[5],[6]. In addition, improper tongue positioning against the palate has been identified as one of the most frequent errors, leading to radiolucent artifacts in the maxillary region and compromising diagnostic interpretation [7], [8]. These findings consistently highlight that human-related factors remain the dominant cause of poor panoramic radiograph quality.

With the advancement of artificial intelligence, deep learning, particularly Convolutional Neural Networks (CNNs), has shown promising performance in automating image quality assessment across various radiographic modalities. CNN-based approaches have achieved high accuracy in evaluating diagnostic quality in ankle radiographs [9], and have been successfully applied to automated quality control in chest and elbow radiography [10], [11]. In the context of dental imaging, recent studies utilizing YOLOv8 and Generative Adversarial Networks have primarily focused on image enhancement or anatomical anomaly detection rather than explicitly addressing acquisition-related errors as indicators of panoramic radiograph quality [12], [13]. Despite these advances, several challenges remain unresolved. Reviews have reported inconsistencies in labeling strategies and dataset utilization for machine learning-based panoramic

*Corresponding author.

radiograph quality assessment, which limits model generalization and reproducibility [14]. Moreover, most deep learning applications in panoramic radiography continue to prioritize anomaly detection or disease diagnosis, such as osteoporosis assessment, rather than systematic quality evaluation based on acquisition technique and patient positioning errors [15], [16], [17]. This indicates that research specifically targeting automated panoramic radiograph quality evaluation remains limited, particularly in terms of methodological robustness and clinical applicability.

To address these limitations, this study proposes a transfer learning-based CNN framework for panoramic radiograph quality classification by systematically comparing three widely used architectures: ResNet50, VGG16, and VGG19. These models were selected due to their proven effectiveness in medical image classification tasks and their distinct feature representation capabilities. Previous studies have demonstrated the strong performance of VGG16 in binary medical image classification tasks, achieving high accuracy and F1-scores [18], while ResNet50 is known for its ability to mitigate the vanishing gradient problem through residual connections. Building upon these findings, the present study integrates Global Average Pooling (GAP) and Batch Normalization (BN) to enhance training stability and generalization, particularly when dealing with limited datasets.

Unlike previous research that focused on dental anomaly detection or disease diagnosis using CNN-based models with Batch Normalization [16], This study emphasizes the evaluation of panoramic radiograph quality based on acquisition-related errors. The main contributions of this research are threefold: 1) providing a comparative analysis of popular pretrained CNN architectures for panoramic radiograph quality assessment, 2) demonstrating the effectiveness of GAP and BN in improving training stability and classification performance on limited datasets, and 3) presenting empirical evidence from multi-institutional panoramic radiographs collected from Universitas Gadjah Mada and Universitas Airlangga hospitals in Indonesia to support the development of AI-based automatic quality assurance systems in dental radiology.

This study is organized as follows: Section II presents a review of related literature. Section III describes the dataset used in this research and explains the proposed method in detail. Section IV discusses the experimental results and findings. Finally, Section V concludes the study and outlines directions for future research.

II. Literature Review

The quality of panoramic radiographs determines the diagnostic value and reliability of image interpretation. Several research have evaluated the consistency and quality of images obtained from different radiographic machines and the factors influencing them. An early research [1] compared two panoramic units and found that both produced equivalent diagnostic results for anatomical evaluation, with negligible differences in image quality, indicating that both devices were clinically acceptable. This suggests that technical acquisition parameters and patient positioning have a greater impact on diagnostic quality than differences between imaging devices.

Automated approaches have also been introduced to assess radiographic acceptability. In [2], the authors developed an automated CNN-based framework to determine whether panoramic radiographs were suitable for further analysis. The system demonstrated high accuracy and showed potential to support diagnostic workflows, particularly in regions with limited access to orthodontic radiologists.

Further applications of deep learning have extended to non-panoramic radiographs. For example, research [9] evaluated the diagnostic quality of ankle radiographs based on anatomical features, achieving 94.1% accuracy, surpassing average radiologist performance. Similarly, research [19] proposed a CNN with multi-scale feature extraction to automatically assess panoramic radiograph quality, outperforming conventional methods. Several studies have also focused on improving image quality. Research [20] introduced a Multi-Scale Top-Hat with Geodesic Reconstruction (MSTHGR) method to enhance image contrast and detail, while research [13] used a Generative Adversarial Network (Pix2Pix) to correct blurred and noisy images, significantly improving PSNR and SSIM metrics. Another research [21] employed a deep learning-based super-resolution technique (LTE) to enhance image resolution with visually and quantitatively superior results.

Panoramic radiography plays an essential role in dental diagnosis; however, patient positioning errors frequently occur due to equipment complexity and anatomical variations, leading to reduced image quality and diagnostic accuracy. Research developed a deep learning model to detect six common positioning errors: slumped posture, low chin, open lips, head rotation, head tilt, and tongue not against the palate, using 552 panoramic radiographs. Six CNN architectures were employed for feature extraction, and the extracted features were fused and classified using a Support Vector Machine (SVM). The model achieved an accuracy of 0.832 with high precision and recall, proving effective for automatic identification of patient positioning errors and improving diagnostic reliability.

Positioning errors are among the main causes of reduced diagnostic quality in panoramic imaging. Research [3] reported that 95% of 500 radiographs contained one or more positioning errors, particularly head tilt. Similarly, research [4] found that 54.6% of radiographs in South Wales exhibited positioning errors and 47.9% were affected by poor patient preparation or instruction. Research [5], which analyzed 1,000 panoramic radiographs, found that the most common error was the tongue not being pressed against the palate (69.5%), followed by patient motion (0.4%). Research [6] involving 2,629 patients revealed that only 32.8% of radiographs were error-free, with tongue positioning errors being the most prevalent. Research [7] compared image quality across three radiographic technicians and found significant inter-operator variation, with the “gazebo effect” and condyle positioning errors being the most frequent findings.

In Indonesia, research [8] reported that 98.78% of panoramic radiographs exhibited patient positioning errors, most commonly due to the tongue not adhering to the palate (49.68%), although the overall diagnostic quality remained

acceptable. Similar errors have been reported in other research [22], [23], [24], [25], where head alignment, lateral rotation, and palatoglossal air space were identified as the primary factors reducing image quality [26].

Several studies have highlighted the impact of imaging technique and operator skill on radiographic quality. In [27], the authors emphasized the importance of head positioning, occlusal plane orientation, mandibular condyle alignment, and metal artifact control in producing symmetric and diagnostically acceptable panoramic images. Research [28] added that horizontal spatial distortion up to ≤20% remains clinically acceptable, while a minimum spatial resolution of 1.88 lp/mm is necessary to maintain anatomical readability.

In [9], the authors showed that while radiographic quality did not vary significantly among technicians, the types of errors did, indicating that operator training and experience play a major role. This finding was supported by research [23], which concluded that radiographic errors are not directly correlated with professional experience but are influenced by patient cooperation and attentiveness during image acquisition. In addition to operator-related issues, technical parameters such as exposure settings and image processing also affect diagnostic quality. In [24], [25], the authors found that post-trauma radiographs frequently suffered from exposure errors, out-of-focus positioning, and metal artifacts, which limited diagnostic interpretation.

Efforts to improve radiographic quality control have increasingly employed deep learning methods. In [12], the authors developed a YOLOv8-based model to assess panoramic radiograph quality using five clinical visual criteria, achieving an overall accuracy of 81.4%. Similar approaches have been applied to chest [10] and elbow [11] radiographs, demonstrating strong potential for automated medical image quality assessment.

Moreover, research [14] conducted a systematic review of 41 research on panoramic radiograph image quality for machine learning model training. The review found that most research lacked standardized image quality criteria, and defective images were often excluded from datasets without detailed analysis. This practice introduces bias and reduces model generalizability, emphasizing the need for automated methods to assess and manage radiograph quality consistently.

The quality of panoramic radiographs is predominantly affected by acquisition technique and patient positioning, with the latter being the leading cause of diagnostically unacceptable images. Although deep learning systems have been widely used for image enhancement and quality evaluation, research that integrates expert clinical annotations with CNN models for objective assessment of panoramic radiograph quality remains limited.

This highlights an existing research gap that needs to be addressed through the development of CNN-based automated systems capable of accurately identifying technical and positioning errors while providing feedback to improve image acquisition quality. Table I summarizes the key differences between the present research and the work of [12].

TABLE I. COMPARISON BETWEEN THE PROPOSED MODEL AND PREVIOUS RESEARCHES

Aspect	Research [12]	Proposed	Contribution/Difference
Objective	Automated assessment of panoramic radiograph quality based on four criteria (artifacts, coverage area, patient positioning, and contrast/density) and clinical acceptability (acceptable vs. not acceptable).	Development of a CNN-based panoramic radiograph quality evaluation model using a domain-adapted ResNet50 to classify image quality into good and poor categories.	Both assess image quality; however, the proposed research focuses on optimizing a domain-adapted CNN and evaluating inter-architecture performance rather than single-model classification using YOLO.
Model	YOLOv8 (single-stage real-time classifier with 141 layers, Adam-W optimizer, learning rate=0.0001).	Domain-adapted ResNet50 architecture with additional Global Average Pooling (GAP) and Batch Normalization (BN) layers for improved stability and generalization.	The proposed research presents a transfer learning-based modification of a classical CNN architecture, which is referred to as a domain-adapted transfer learning to emphasize its adaptation to the specific characteristics of panoramic dental radiographs. This adaptation was achieved through architectural modification (GAP and BN layers) and fine-tuning on domain-specific data.
Evaluation Criteria	Four aspects (artifacts, coverage, positioning, contrast/density) plus overall image quality.	Focuses on overall radiograph quality (good or poor) based on coverage, vertical and horizontal distortion, detail and sharpness, artifacts and ghost images, and radiolucent areas in the anterior maxillary apical region.	The proposed method performs holistic radiograph quality assessment rather than per-aspect evaluation, providing faster and more efficient results for clinical application.
Accuracy	74.1-97.9% per criterion, with an average of 81.4% for clinical classification.	Domain-adapted ResNet50 achieved 82-85% accuracy and the highest F1-score compared to baseline CNNs (VGG architectures).	The proposed model delivers comparable or slightly higher performance with lower model complexity than YOLOv8.
Optimization	Strategy Data augmentation, Adam-W optimizer, and automatic mixed precision.	Transfer learning, Batch Normalization, and adaptive learning rate scheduler.	The proposed approach emphasizes training efficiency and domain adaptability without requiring high-end GPU resources.
Clinical Contribution	Provides a real-time radiograph quality control system and an educational tool for dental students.	Offers a lightweight CNN model for automated radiograph quality evaluation in resource-limited institutions.	The proposed method is easier to implement in local dental clinics without the need for high computational infrastructure.

A. Conventional Approaches and Their Limitations

Early research relied on manual visual inspection by radiologists, which is subjective and time-consuming. Later, rule-based metrics such as contrast, entropy, or edge sharpness were used to quantify quality, but these fail to capture anatomical distortion or exposure errors. Classical machine-learning models (e.g., SVM, *k*-NN) using handcrafted texture or wavelet features improved objectivity but depended heavily on feature design and were sensitive to acquisition variability. Recent deep-learning approaches applied CNNs for anomaly detection or anatomy segmentation rather than direct good/poor quality classification, and often used single-center data without cross-institution validation. This research differs by applying a transfer-learning CNN (ResNet50, VGG16, VGG19) directly to the binary quality classification task of panoramic radiographs, reducing the need for handcrafted features and improving robustness for small, domain-specific datasets.

III. Methods

A. Dataset

The dataset used in this research was obtained from the Dental and Oral Hospital (RSGM) of Universitas Airlangga, Surabaya, and the Dental and Oral Hospital of Universitas Gadjah Mada, Yogyakarta. Ethical approval for the use of radiographic data was granted by both institutions under approval numbers 0910/HRECC.FODM/VIII/2025 and 158/UNI/KEP/FKG-RSGM/EC/2025, respectively. In total, 1,285 panoramic radiographic images were collected, consisting of 820 images from Universitas Airlangga and 465 images from Universitas Gadjah Mada.

Each radiograph was independently evaluated by four dental radiology experts based on five key diagnostic quality aspects (see Table II), with a maximum total score of five. Radiographs with a total score of five were categorized as good quality, indicating no retake was required, while radiographs with a score below five were categorized as poor quality, indicating a retake was recommended. To ensure annotation reliability and reduce label ambiguity, only radiographs with full consensus among all four experts were included in the final dataset used for model training and evaluation.

Based on this strict selection criterion, 442 panoramic radiographs were retained, including 221 good-quality and 221 poor-quality images (see Table III). Specifically, 105 good-quality and 105 poor-quality radiographs were obtained from Universitas Airlangga, while 116 good-quality and 116 poor-quality radiographs were obtained from Universitas Gadjah Mada. This balanced class distribution was intentionally maintained to prevent model bias towards a dominant class.

Examples of panoramic radiographs used in this study are shown in Fig. 1. Good-quality radiographs exhibit complete anatomical coverage, clear visualization of the maxillomandibular structures, sharp anatomical details, and the absence of distortion or artifacts that may interfere with diagnosis. In contrast, poor-quality radiographs are characterized by positioning-related distortions, asymmetry, reduced sharpness due to patient movement, and the presence of artifacts such as ghost images or extensive radiolucent regions that compromise diagnostic reliability.

TABLE II. Image Quality Assessment Criteria

No.	Aspect	1	0
1	Coverage	The image includes the lower border of the mandible, both right and left condyles, and the lower border of the orbits.	One or more of these anatomical areas are not included in the image.
2	Vertical and Horizontal Distortion	No vertical or horizontal distortion is observed.	Distortion interferes with interpretation (e.g., inverted or V-shaped dental arch, asymmetrical mandibular shape).
3	Detail and Sharpness	The image shows clear anatomical detail and sharpness.	Image detail is unclear or overlapping due to patient head movement during exposure.
4	Artifacts and Ghost Images	No artifacts or ghost images interfere with diagnostic interpretation.	Ghost images from foreign objects or large radiolucent areas (air spaces) obscure dental or maxillofacial structures.
5	Radiolucent Area in the Anterior Maxillary Apical Region	No radiolucent area interferes with interpretation.	A radiolucent area is present in the anterior maxillary apical region, reducing diagnostic clarity.

TABLE III. Dataset Distribution

Class	Training	Testing	Total
Good	170	51	221
Poor	170	51	221
Total	340	102	442

Although the final dataset used in this study consists of 442 curated panoramic radiographs, which may be considered relatively small for deep learning applications, this reflects realistic clinical constraints where high-quality, expert-labeled radiographic data are limited. The reduced dataset size increases the risk of overfitting; therefore, several mitigation strategies were employed. First, transfer learning from pretrained CNN models was used to leverage previously learned generic visual features. Second, Global Average Pooling (GAP) was applied to reduce the number of trainable parameters, while Batch Normalization (BN) was incorporated to stabilize training and improve generalization. In addition, *k*-fold cross-validation was adopted to maximize data utilization and provide a more robust estimate of model performance across different data partitions.

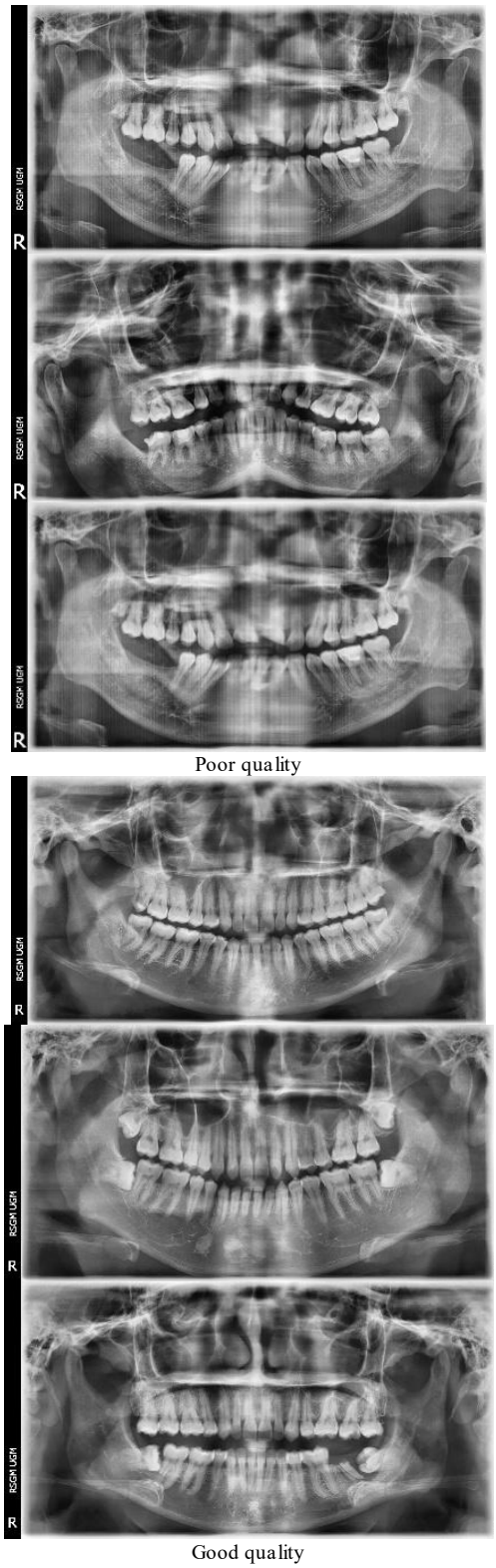


Fig. 1. Examples of panoramic radiographs.

Although the final dataset size is relatively limited, this scenario reflects realistic clinical conditions where curated, high-quality labeled radiographic data are often scarce. To mitigate the risk of overfitting associated with limited data, this study employed a transfer learning strategy using pretrained CNN models, combined with Global Average Pooling and Batch Normalization to enhance feature robustness and training stability. Furthermore, *k*-fold cross-validation was applied to maximize data utilization and provide a more reliable estimation of model generalization performance across different data splits.

B. Proposed Method

This research proposes a transfer learning-based approach using three CNN architectures: ResNet50, VGG16, and VGG19 to classify the quality of panoramic dental radiographs into two classes: good and poor (Fig. 2). All pretrained models were initialized with ImageNet weights and modified by adding Global Average Pooling (GAP) and Batch Normalization (BN) layers to improve adaptability to the specific characteristics of dental radiographs. In this context, the term domain-adapted refers to the adaptation of a pretrained network through transfer learning and fine-tuning using dental imaging data, rather than an explicit domain adaptation mechanism such as adversarial alignment or feature distribution matching. The proposed method aims to automate the panoramic radiograph quality assessment process and empirically verify the reliability of ResNet50 compared to the other architectures. Model performance was evaluated using *k*-fold cross-validation to ensure robustness and reduce sampling bias. The main stages of the proposed method include the following steps (as illustrated in Fig. 2):

a) Preprocessing data: All panoramic radiograph images were converted into RGB format with a resolution of 512×512 pixels and normalized to a range of [0,1].

b) Feature extraction (transfer learning): Three pretrained CNN models were employed as feature extractors: ResNet50, which utilizes residual connections to preserve gradient flow and is well-suited for detecting complex radiographic textures; VGG16 and VGG19, which employ simple stacked convolutional layers effective for capturing edge and shape representations.

c) Architecture modification: To enhance model performance, several layers were added: Global Average Pooling (GAP) to reduce feature dimensions while retaining essential spatial information; Batch Normalization (BN) to stabilize inter-layer activation distributions and accelerate convergence; and Dense Layers (128-256 units) with ReLU activation as the classifier head. The models were optimized using the Adam or AdamW optimizer with a learning rate of 0.0001, batch size of 2, and 30 training epochs.

d) Training and evaluation: The models were trained on a labeled dataset of 442 panoramic radiographs (340 for training and 102 for testing). Evaluation metrics included accuracy [Eq. (1)], precision [Eq. (2)], recall (sensitivity) [Eq. (3)], specificity [Eq. (4)], and F1-Score [Eq. (5)] analysis [18]. These metrics were selected to provide a comprehensive understanding of diagnostic reliability, particularly regarding false negatives, which may correspond to clinically unacceptable images being classified as acceptable. This evaluation framework supports the model's applicability for clinical quality assurance and radiation dose reduction in dental radiology workflows.

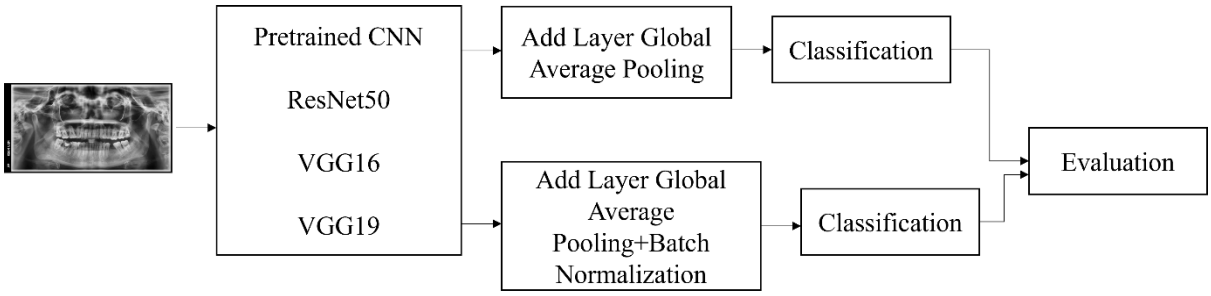


Fig. 2. Proposed method framework.

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

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP+TN}{TP+FN} \quad (3)$$

$$Specificity = \frac{TN}{TP+FP} \quad (4)$$

$$F1-Score = \frac{2.TP}{2.TP+FP+FN} \quad (5)$$

IV. Results and Discussion

Tables IV to XII summarize the class-wise evaluation metrics obtained from *k*-fold cross-validation for all evaluated models. Rather than focusing solely on absolute metric values, the results reveal clear performance trends across architectures and configurations. Overall, VGG16-based models consistently demonstrated more stable and balanced performance compared to ResNet50 and VGG19, particularly when combined with Global Average Pooling (GAP) and Batch Normalization (BN).

ResNet50 achieved relatively high AUC values in several folds, indicating strong discriminative potential. However, its performance exhibited substantial variability in recall and specificity across folds (Tables IV to VI and Tables XIII to XV). This instability suggests sensitivity to data partitioning and potential overfitting, which can be attributed to the model's higher architectural complexity relative to the limited size of the curated panoramic radiograph dataset. Although residual connections are designed to facilitate deeper feature learning, they may lead to unstable optimization when the available training data are insufficient to support robust parameter estimation.

In contrast, VGG16 demonstrated consistently lower variance across folds and more balanced class-wise performance (Tables VII to IX and Tables XVI to XVIII). The integration of GAP reduced the number of trainable parameters, while BN improved gradient flow and training stability. These architectural refinements enabled VGG16 to achieve the most reliable overall performance, with an average accuracy of 85.00% and an AUC of 94.72%. The relatively shallow and uniform convolutional structure of VGG16 appears to be better suited for capturing global structural distortions and acquisition-related artifacts commonly observed in panoramic radiographs.

TABLE IV. ResNet50 (GAP) Evaluation Results for Each Class

k-fold=1	Recall	Precision	F1-score	Specificity
Good	55.56	80.00	65.57	84.38
Poor	84.38	62.79	72.00	55.56
mean	69.97	71.40	68.79	69.97
std	20.38	12.17	4.54	20.38
k-fold=2	Recall	Precision	F1-score	Specificity
Good	28.57	58.82	38.46	78.79
Poor	78.79	50.98	50.98	28.57
mean	53.68	54.90	44.72	53.68
std	35.51	5.55	8.85	35.51
k-fold=3	Recall	Precision	F1-score	Specificity
Good	71.88	62.16	66.67	61.11
Poor	61.11	70.97	65.67	71.88
mean	66.49	66.56	66.17	66.49
std	7.61	6.23	0.70	7.61
k-fold=4	Recall	Precision	F1-score	Specificity
Good	100.00	52.94	69.23	0.00
Poor	0.00	0.00	0.00	100.00
mean	50.00	26.47	34.62	50.00
std	70.71	37.44	48.95	70.71
k-fold=5	Recall	Precision	F1-score	Specificity
Good	100.00	60.78	75.61	45.95
Poor	45.95	100.00	62.96	100.00
mean	72.97	80.39	69.29	72.97
std	38.22	27.73	8.94	38.22

TABLE V. ResNet50 (GAP+BN) Evaluation Results for Each Class

k-fold=1	Recall	Precision	F1-score	Specificity
Good	66.67	88.89	76.19	90.63
Poor	90.63	70.73	79.45	66.67
mean	78.65	79.81	77.82	78.65
std	16.94	12.84	2.31	16.94
k-fold=2	Recall	Precision	F1-score	Specificity
Good	22.86	88.89	36.36	96.97

Poor	96.97	54.24	54.24	22.86
mean	59.91	71.56	45.30	59.91
std	52.41	24.50	12.64	52.41
k-fold=3	Recall	Precision	F1-score	Specificity
Good	0.00	0.00	0.00	100.00
Poor	100.00	52.94	69.23	0.00
mean	50.00	26.47	34.62	50.00
std	70.71	37.44	48.95	70.71
k-fold=4	Recall	Precision	F1-score	Specificity
Good	100.00	52.94	69.23	0.00
Poor	0.00	0.00	0.00	100.00
mean	50.00	26.47	34.62	50.00
std	70.71	37.44	48.95	70.71
k-fold=5	Recall	Precision	F1-score	Specificity
Good	100.00	45.59	62.63	0.00
Poor	0.00	0.00	0.00	100.00
mean	50.00	22.79	31.31	50.00
std	70.71	32.24	44.28	70.71

TABLE VI. RESNET50 (GAP+BN) OPTIMIZER ADAMW EVALUATION RESULTS FOR EACH CLASS

k-fold=1	Recall	Precision	F1-score	Specificity
Good	100.00	61.02	75.79	28.13
Poor	28.13	100.00	43.90	100.00
mean	64.06	80.51	59.85	64.06
std	50.82	27.57	22.55	50.82
k-fold=2	Recall	Precision	F1-score	Specificity
Good	100.00	51.47	67.96	0.00
Poor	0.00	0.00	0.00	100.00
mean	50.00	25.74	33.98	50.00
std	70.71	36.40	48.06	70.71
k-fold=3	Recall	Precision	F1-score	Specificity
Good	0.00	0.00	0.00	100.00
Poor	100.00	52.94	69.23	0.00
mean	50.00	26.47	34.62	50.00
std	70.71	37.44	48.95	70.71
k-fold=4	Recall	Precision	F1-score	Specificity
Good	97.22	62.50	76.09	34.38
Poor	34.38	91.67	50.00	97.22
mean	65.80	77.08	63.04	65.80
std	44.44	20.62	18.45	44.44
k-fold=5	Recall	Precision	F1-score	Specificity
Good	100.00	45.59	62.63	0.00
Poor	0.00	0.00	0.00	100.00
mean	50.00	22.79	31.31	50.00
std	70.71	32.24	44.28	70.71

TABLE VII. VGG16 (GAP) EVALUATION RESULTS FOR EACH CLASS

k-fold=1	Recall	Precision	F1-score	Specificity
Good	86.11	72.09	78.48	62.50
Poor	62.50	80.00	70.18	86.11
mean	74.31	76.05	74.33	74.31
std	16.70	5.59	5.87	16.70
k-fold=2	Recall	Precision	F1-score	Specificity
Good	60.00	87.50	71.19	90.91
Poor	90.91	68.18	77.92	60.00
mean	75.45	77.84	74.55	75.45
std	21.86	13.66	4.76	21.86
k-fold=3	Recall	Precision	F1-score	Specificity
Good	84.38	64.29	72.97	58.33
Poor	58.33	80.77	67.74	84.38
mean	71.35	72.53	70.36	71.35
std	18.41	11.66	3.70	18.41
k-fold=4	Recall	Precision	F1-score	Specificity
Good	97.22	67.31	79.55	46.88
Poor	46.88	93.75	62.50	97.22
mean	72.05	80.53	71.02	72.05
std	35.60	18.70	12.05	35.60
k-fold=5	Recall	Precision	F1-score	Specificity
Good	93.55	69.05	79.45	64.86
Poor	64.86	92.31	76.19	93.55
mean	79.21	80.68	77.82	79.21
std	20.28	16.45	2.31	20.28

TABLE VIII. VGG16 (GAP+BN) EVALUATION RESULTS FOR EACH CLASS

k-fold=1	Recall	Precision	F1-score	Specificity
Good	69.44	100.00	81.97	100.00
Poor	100.00	74.42	85.33	69.44
mean	84.72	87.21	83.65	84.72
std	21.61	18.09	2.38	21.61
k-fold=2	Recall	Precision	F1-score	Specificity
Good	62.86	95.65	75.86	96.97
Poor	96.97	71.11	82.05	62.86
mean	79.91	83.38	78.96	79.91
std	24.12	17.35	4.38	24.12
k-fold=3	Recall	Precision	F1-score	Specificity
Good	96.88	81.58	88.57	80.56
Poor	80.56	96.67	87.88	96.88
mean	88.72	89.12	88.23	88.72
std	11.54	10.67	0.49	11.54
k-fold=4	Recall	Precision	F1-score	Specificity
Good	100.00	83.72	91.14	78.13
Poor	78.13	100.00	87.72	100.00
mean	89.06	91.86	89.43	89.06
std	15.47	11.51	2.42	15.47
k-fold=5	Recall	Precision	F1-score	Specificity
Good	93.55	76.32	84.06	75.68
Poor	75.68	93.33	83.58	93.55
mean	84.61	84.82	83.82	84.61
std	12.64	12.03	0.34	12.64

TABLE IX. VGG16 (GAP+BN) OPTIMIZER ADAMW EVALUATION RESULTS FOR EACH CLASS

k-fold=1	Recall	Precision	F1-score	Specificity
Good	63.89	95.83	76.67	96.88
Poor	96.88	70.45	81.58	63.89
mean	80.38	83.14	79.12	80.38
std	23.32	17.95	3.47	23.32
k-fold=2	Recall	Precision	F1-score	Specificity
Good	97.14	73.91	83.95	63.64
Poor	63.64	95.45	76.36	97.14
mean	80.39	84.68	80.16	80.39
std	23.69	15.23	5.36	23.69
k-fold=3	Recall	Precision	F1-score	Specificity
Good	90.63	74.36	81.69	72.22
Poor	72.22	89.66	80.00	90.63
mean	81.42	82.01	80.85	81.42
std	13.01	10.82	1.20	13.01
k-fold=4	Recall	Precision	F1-score	Specificity
Good	100.00	80.00	88.89	71.88
Poor	71.88	100.00	83.64	100.00
mean	85.94	90.00	86.26	85.94
std	19.89	14.14	3.71	19.89
k-fold=5	Recall	Precision	F1-score	Specificity
Good	16.13	100.00	27.78	100.00
Poor	100.00	58.73	74.00	16.13
mean	58.06	79.37	50.89	58.06
std	59.31	29.18	32.68	59.31

TABLE X. VGG19 (GAP) EVALUATION RESULTS FOR EACH CLASS

k-fold=1	Recall	Precision	F1-score	Specificity
Good	72.22	74.29	73.24	71.88
Poor	71.88	69.70	70.77	72.22
mean	72.05	71.99	72.00	72.05
std	0.25	3.24	1.75	0.25
k-fold=2	Recall	Precision	F1-score	Specificity
Good	80.00	70.00	74.67	63.64
Poor	63.64	75.00	68.85	80.00
mean	71.82	72.50	71.76	71.82
std	11.57	3.54	4.11	11.57
k-fold=3	Recall	Precision	F1-score	Specificity
Good	90.63	65.91	76.32	58.33
Poor	58.33	87.50	70.00	90.63
mean	74.48	76.70	73.16	74.48
std	22.83	15.27	4.47	22.83
k-fold=4	Recall	Precision	F1-score	Specificity
Good	88.89	68.09	77.11	53.13
Poor	53.13	80.95	64.15	88.89
mean	71.01	74.52	70.63	71.01
std	25.29	9.10	9.16	25.29
k-fold=5	Recall	Precision	F1-score	Specificity
Good	77.42	77.42	77.42	81.08
Poor	81.08	81.08	81.08	77.42
mean	79.25	79.25	79.25	79.25
std	2.59	2.59	2.59	2.59

TABLE XI. VGG19 (GAP+BN) EVALUATION RESULTS FOR EACH CLASS

k-fold=1	Recall	Precision	F1-score	Specificity
Good	86.11	79.49	82.67	75.00
Poor	75.00	82.76	78.69	86.11
mean	80.56	81.12	80.68	80.56
std	7.86	2.31	2.81	7.86
k-fold=2	Recall	Precision	F1-score	Specificity
Good	100.00	59.32	74.47	27.27
Poor	27.27	100.00	42.86	100.00
mean	63.64	79.66	58.66	63.64
std	51.43	28.76	22.35	51.43
k-fold=3	Recall	Precision	F1-score	Specificity
Good	96.88	50.00	65.96	13.89
Poor	13.89	83.33	23.81	96.88
mean	55.38	66.67	44.88	55.38
std	58.68	23.57	29.80	58.68
k-fold=4	Recall	Precision	F1-score	Specificity
Good	72.22	86.67	78.79	87.50
Poor	87.50	73.68	80.00	72.22
mean	79.86	80.18	79.39	79.86
std	10.80	9.18	0.86	10.80
k-fold=5	Recall	Precision	F1-score	Specificity
Good	87.10	90.00	88.52	91.89
Poor	91.89	89.47	90.67	87.10
mean	89.49	89.74	89.60	89.49
std	3.39	0.37	1.51	3.39

TABLE XII. VGG19 (GAP+BN) OPTIMIZER ADAMW EVALUATION RESULTS FOR EACH CLASS

k-fold=1	Recall	Precision	F1-score	Specificity
Good	72.22	86.67	78.79	87.50
Poor	87.50	73.68	80.00	72.22
mean	79.86	80.18	79.39	79.86
std	10.80	9.18	0.86	10.80
k-fold=2	Recall	Precision	F1-score	Specificity
Good	100.00	70.00	82.35	54.55

Poor	54.55	100.00	70.59	100.00
mean	77.27	85.00	76.47	77.27
std	32.14	21.21	8.32	32.14
k-fold=3	Recall	Precision	F1-score	Specificity
Good	90.63	80.56	85.29	80.56
Poor	80.56	90.63	85.29	90.63
mean	85.59	85.59	85.29	85.59
std	7.12	7.12	0.00	7.12
k-fold=4	Recall	Precision	F1-score	Specificity
Good	86.11	83.78	84.93	81.25
Poor	81.25	83.87	82.54	86.11
mean	83.68	83.83	83.74	83.68
std	3.44	0.06	1.69	3.44
k-fold=5	Recall	Precision	F1-score	Specificity
Good	19.35	100.00	32.43	100.00
Poor	100.00	59.68	74.75	19.35
mean	59.68	79.84	53.59	59.68
std	57.02	28.51	29.92	57.02

TABLE XIII. RESNET50 (GAP) *K*-FOLD CROSS-VALIDATION RESULTS

k-fold	Accuracy	Precision	Recall	F1-score	AUC	Specificity
1	69.12	62.79	84.38	72.00	78.21	55.56
2	52.94	50.98	78.79	61.90	69.52	28.57
3	66.18	70.97	61.11	65.67	72.92	71.88
4	52.94	0.00	0.00	0.00	76.39	100.00
5	70.59	100.00	45.95	62.96	92.07	100.00
max	70.59	100.00	84.38	72.00	92.07	100.00
mean	62.35	56.95	54.05	52.51	77.82	71.20
std	8.74	36.62	33.80	29.61	8.63	30.50

TABLE XIV. RESNET50 (GAP+BN) *K*-FOLD CROSS-VALIDATION RESULTS

k-fold	Accuracy	Precision	Recall	F1-score	AUC	Specificity
1	77.94	70.73	90.63	79.45	91.58	66.67
2	58.82	54.24	96.97	69.57	91.26	22.86
3	52.94	52.94	100.00	69.23	90.23	0.00
4	52.94	0.00	0.00	0.00	87.85	100.00
5	45.59	0.00	0.00	0.00	93.98	100.00
max	77.94	70.73	100.00	79.45	93.98	100.00
mean	57.65	35.58	57.52	43.65	90.98	57.91
std	12.28	33.23	52.62	40.06	2.22	45.28

TABLE XV. RESNET50 (GAP+BN) OPTIMIZER ADAMW *K*-FOLD CROSS-VALIDATION RESULTS

k-fold	Accuracy	Precision	Recall	F1-score	AUC	Specificity
1	66.18	100.00	28.13	43.90	86.20	100.00
2	51.47	0.00	0.00	0.00	89.61	100.00
3	52.94	52.94	100.00	69.23	93.06	0.00
4	67.65	91.67	34.38	50.00	89.06	97.22
5	45.59	0.00	0.00	0.00	95.90	100.00
max	67.65	100.00	100.00	69.23	95.90	100.00
mean	56.77	48.92	32.50	32.63	90.77	79.44
std	9.68	48.06	40.90	31.22	3.77	44.43

TABLE XVI. VGG16 (GAP) *K*-FOLD CROSS-VALIDATION RESULTS

k-fold	Accuracy	Precision	Recall	F1-score	AUC	Specificity
1	75.00	80.00	62.50	70.18	78.21	86.11
2	75.00	68.18	90.91	77.92	85.45	60.00
3	70.59	80.77	58.33	67.74	74.48	84.38
4	73.53	93.75	46.88	62.50	86.89	97.22
5	77.94	92.31	64.86	76.19	85.00	93.55
max	77.94	93.75	90.91	77.92	86.89	97.22
mean	74.41	83.00	64.70	70.91	82.01	84.25
std	2.67	10.44	16.20	6.29	5.38	14.55

TABLE XVII. VGG16 (GAP+BN) *K*-FOLD CROSS-VALIDATION RESULTS

k-fold	Accuracy	Precision	Recall	F1-score	AUC	Specificity
1	83.82	74.42	100.00	85.33	95.75	69.44
2	79.41	71.11	96.97	82.05	96.19	62.86
3	88.24	96.67	80.56	87.88	94.88	96.88
4	89.71	100.00	78.13	87.72	94.79	100.00
5	83.82	93.33	75.68	83.58	91.98	93.55
max	89.71	100.00	100.00	87.88	96.19	100.00
mean	85.00	87.11	86.27	85.31	94.72	84.55
std	4.08	13.35	11.34	2.55	1.64	17.11

TABLE XVIII.VGG16 (GAP+BN) OPTIMIZER ADAMW *K*-FOLD CROSS-VALIDATION RESULTS

k-fold	Accuracy	Precision	Recall	F1-score	AUC	Specificity
1	79.41	70.45	96.88	81.58	95.05	63.89
2	80.88	95.45	63.64	76.36	95.93	97.14
3	80.88	89.66	72.22	80.00	93.84	90.63
4	86.76	100.00	71.88	83.64	90.00	100.00
5	61.76	58.73	100.00	74.00	91.63	16.13
max	86.76	100.00	100.00	83.64	95.93	100.00
mean	77.94	82.86	80.92	79.12	93.29	73.56
std	9.47	17.57	16.39	3.91	2.45	35.14

TABLE XIX. VGG19 (GAP) *K*-FOLD CROSS-VALIDATION RESULTS

k-fold	Accuracy	Precision	Recall	F1-score	AUC	Specificity
1	72.06	69.70	71.88	70.77	84.38	72.22
2	72.06	75.00	63.64	68.85	80.26	80.00
3	73.53	87.50	58.30	70.00	72.60	90.63
4	72.06	80.95	53.13	64.15	77.08	88.89
5	79.41	81.08	81.08	81.08	88.14	77.42
max	79.41	87.50	81.08	81.08	88.14	90.63
mean	73.82	78.85	65.61	70.97	80.49	81.83
std	3.19	6.76	11.09	6.21	6.07	7.79

TABLE XX. VGG19 (GAP+BN) *K*-FOLD CROSS-VALIDATION RESULTS

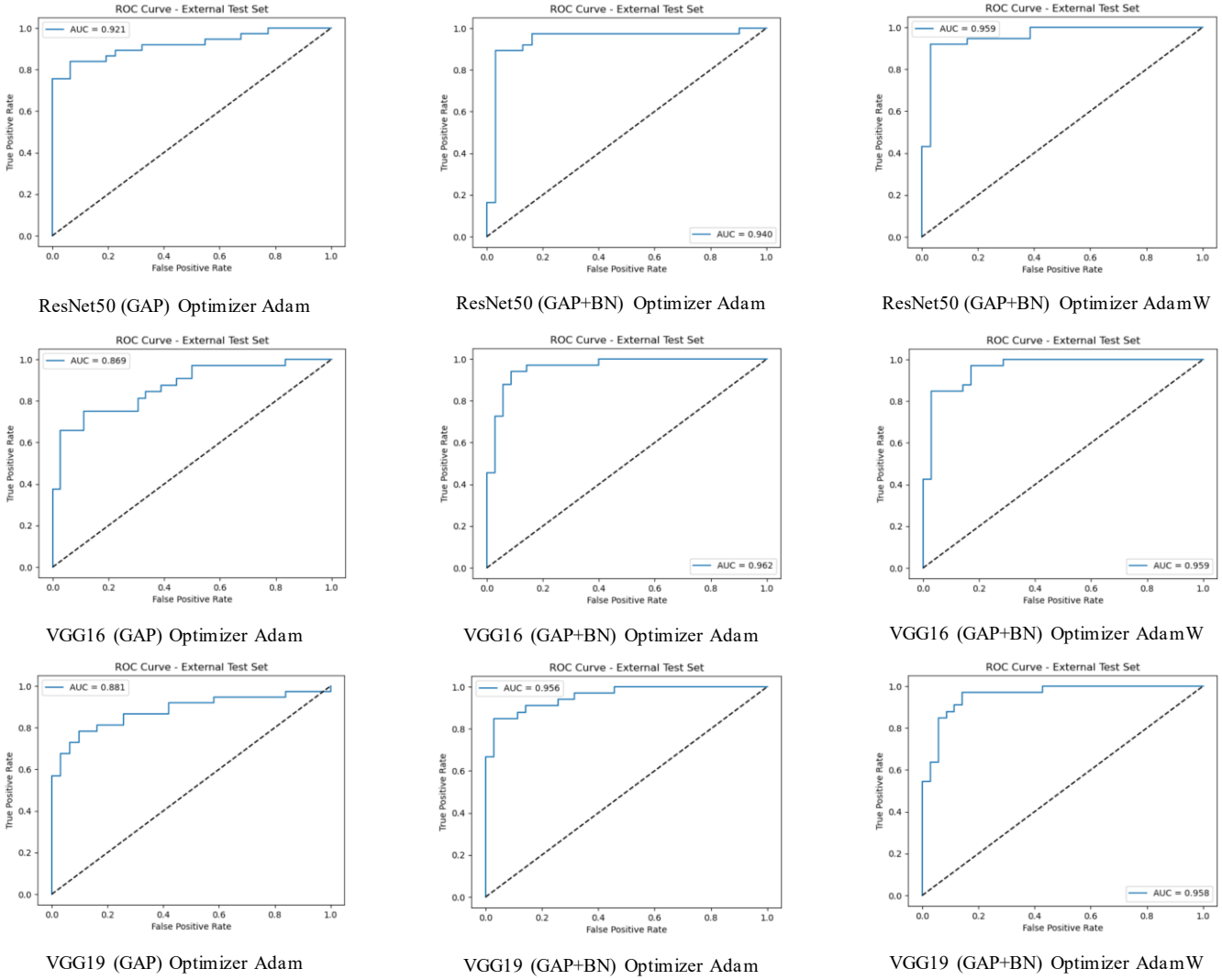
k-fold	Accuracy	Precision	Recall	F1-score	AUC	Specificity
1	80.88	82.76	75.00	78.69	86.72	86.11
2	64.71	100.00	27.27	42.86	95.58	100.00
3	52.94	83.33	13.89	23.81	92.27	96.88
4	79.41	73.68	87.50	80.00	92.71	72.22
5	89.71	89.47	91.89	90.67	93.46	87.10
max	89.71	100.00	91.89	90.67	95.58	100.00
mean	73.53	85.85	59.11	63.21	92.15	88.46
std	14.60	9.71	36.03	28.47	3.29	10.90

TABLE XXI. VGG19 (GAP+BN) OPTIMIZER ADAMW *K*-FOLD CROSS-VALIDATION RESULTS

k-fold	Accuracy	Precision	Recall	F1-score	AUC	Specificity
1	79.41	73.68	87.50	80.00	89.15	72.22
2	77.94	100.00	54.55	70.59	95.76	100.00
3	85.29	90.63	80.56	85.29	91.93	90.63
4	83.82	83.87	81.25	82.54	92.88	86.11
5	63.24	59.68	100.00	74.75	93.55	19.35
max	85.29	100.00	100.00	85.29	95.76	100.00
mean	77.94	81.57	80.77	78.63	92.65	73.66
std	8.76	15.56	16.60	5.94	2.41	31.97

VGG19 exhibited competitive discriminative performance, with mean AUC values around 92% (Table X to Table XII and Table XIX to Table XXI). However, its deeper architecture resulted in slightly higher performance variability compared to VGG16, suggesting increased sensitivity to limited data and fold composition. Among VGG19 variants, the use of adaptive optimization (AdamW) improved convergence consistency, reinforcing the importance of optimizer selection in small-scale medical imaging datasets.

Fig. 3 illustrates the ROC curves for all evaluated configurations, further confirming that VGG16 with GAP and BN achieved superior and more consistent discrimination between good- and poor-quality radiographs. The proximity of the ROC curve to the upper-left corner indicates a favorable balance between sensitivity and specificity, which is critical in quality assessment tasks where both false positives and false negatives carry clinical consequences.



ResNet50 (GAP) Optimizer Adam | ResNet50 (GAP+BN) Optimizer Adam | ResNet50 (GAP+BN) Optimizer AdamW

VGG16 (GAP) Optimizer Adam | VGG16 (GAP+BN) Optimizer Adam | VGG16 (GAP+BN) Optimizer AdamW

VGG19 (GAP) Optimizer Adam | VGG19 (GAP+BN) Optimizer Adam | VGG19 (GAP+BN) Optimizer AdamW

Fig. 3. ROC-AUC curves of transfer learning models.

From an error analysis perspective, false negatives where poor-quality radiographs are misclassified as good quality pose a greater clinical risk than false positives. Such errors may result in diagnostically inadequate images being accepted, potentially leading to misdiagnosis or delayed treatment. The VGG16 (GAP+BN, Adam) model exhibited a

lower false-negative rate compared to other configurations, suggesting improved reliability for clinical deployment. False positives, while less critical from a safety standpoint, may still contribute to unnecessary retakes and workflow inefficiencies; however, these errors are generally preferable to false negatives in radiographic quality assurance.

A comparative evaluation with previous studies (Table XXII) demonstrates that the proposed VGG16-based approach achieves competitive performance relative to more complex frameworks such as CNN-SVM hybrids or object detection-based models (e.g., YOLOv8). Importantly, the proposed method maintains a lightweight architecture and efficient training process, making it more practical for real-world clinical implementation.

Overall, these findings indicate that domain-adapted transfer learning, when combined with architectural simplification and normalization strategies, can effectively enhance model generalization for automatic panoramic radiograph quality assessment. By providing reliable differentiation between diagnostically acceptable and poor-quality images, the proposed approach has the potential to reduce unnecessary retakes, improve patient safety by minimizing radiation exposure, and support standardized quality control in dental radiology practice.

Despite the promising results, this study has several limitations. The relatively small dataset, although carefully curated with expert consensus, may limit the generalizability of the proposed models to broader clinical settings. While the use of transfer learning, GAP, BN, and *k*-fold cross-validation helps mitigate overfitting, external validation using independent datasets from additional institutions would provide stronger evidence of generalizability. Furthermore, ablation studies investigating the effects of dataset size and class balance were not conducted and will be considered in future research.

TABLE XXII. CITATION-ONLY CONTEXTUAL COMPARISON

No	Model	Research Focus	Evaluation Results
1	Model CNN feature fusion+SVM [28]	Patient positioning errors (multi-class)	Accuracy=83.2%, AUC up to 0.998
2	YOLOv8 Classification Model [12]	Radiograph quality	Artifacts (Accuracy=87.2%, Precision=88.9%, F1-Score=0.864-0.943); Coverage area (Accuracy=74.1%, Precision =83.3-91.3%, F1-Score=0.769-0.941); Patient positioning (Accuracy=77.3%, Precision=76.9%, F1-Score=0.72-0.67); Contrast/Density (Accuracy=97.9%, Precision=83.3-84.6%, F1-Score=0.782-0.765); Overall image quality (Accuracy=79.3%, Precision=81.4%, F1-Score=0.809-0.820)
3	Proposed Method	Radiograph quality (good and poor)	Accuracy=85.00%, Precision=87.11%, Recall=86.27%, F1-score=85.31%, AUC=94.72%, Specificity=84.55%

V. CONCLUSION

Based on the experimental results and analysis, the proposed transfer learning-based CNN models demonstrated promising performance in the automatic quality classification of panoramic dental radiographs. Among the evaluated architectures, the VGG16 model enhanced with Global Average Pooling (GAP) and Batch Normalization (BN) and optimized using the Adam optimizer achieved the best overall performance under *k*-fold cross-validation, with an average accuracy of 85.00%, precision of 87.11%, recall of 86.27%, F1-score of 85.31%, AUC of 94.72%, and specificity of 84.55%. The integration of GAP and BN effectively improved training stability and reduced overfitting, allowing the model to generalize well across folds. These findings confirm that the domain-adapted transfer learning approach can provide a stable and reliable framework for automated panoramic radiograph quality assessment. In the future, expanding the dataset and incorporating external validation across multiple institutions will further strengthen model generalizability and clinical applicability. This study contributes to advancing AI-based quality assurance systems in dental radiology, enabling more consistent diagnostic outcomes and reducing retake rates in clinical practice.

Based on these findings, several directions for future research are recommended: Increasing data quantity and diversity to improve model robustness and generalization. Further fine-tuning of pretrained models to better adapt to radiographic features. System implementation and integration into clinical workflows for real-time image quality assessment.

ACKNOWLEDGMENT

Thanks to the Deputy for Strengthening Research and Development, Ministry of Research and Technology/National Research and Innovation Agency, Indonesia, for providing research funding through the Domestic Cooperation Research scheme with research contract number 001/SP2H/PL/LITBANG PEMAS/2025. Thanks to the Dental and Oral Hospital of Universitas Airlangga, Surabaya and Universitas Gadjah Mada, Yogyakarta, for providing dental radiographic panoramic data.

REFERENCES

[1] G. Kaeppler, D. Axmann-Krcmar, I. Reuter, J. Meyle, and G. Gómez-Román, "A clinical evaluation of some factors affecting image quality in panoramic radiography," Dentomaxillofacial Radiology, vol. 29, no. 2, 2000, doi: 10.1038/sj.dmfr.4600505.

[2] J. Faure and A. Engelbrecht, "A Convolutional Neural Network for Dental Panoramic Radiograph Classification," in ACM International

Conference Proceeding Series, Association for Computing Machinery, Apr. 2021, pp. 54–59. doi: 10.1145/3461598.3461607.

[3] A. Khator, M. Motwani, and A. Choudhary, "A study for determination of various positioning errors in digital panoramic radiography for evaluation of diagnostic image quality," Indian Journal of Dental Research, vol. 28, no. 6, 2017, doi: 10.4103/ijdr.IJDR_781_16.

[4] A. Loughlin, N. Drage, C. Greenall, and D. J. J. Farnell, "An investigation in to the impact of acquisition location on error type and rate when undertaking panoramic radiography," Radiography, vol. 23, no. 4, 2017, doi: 10.1016/j.radi.2017.07.004.

[5] N. Kumar, "Assessment of common errors and subjective quality of digital panoramic radiographs in a dental institution," Dentistry and Medical Research, vol. 8, no. 1, 2020, doi: 10.4103/dmr.dmr_22_19.

[6] A. S. Lingam, P. Koppolu, R. Abdulsalam, R. L. Reddy, A. Anwarullah, and D. Koppolu, "Assessment of common errors and subjective quality of digital panoramic radiographs in dental institution, Riyadh," Ann Afr Med, vol. 22, no. 1, 2023, doi: 10.4103/aam.aam_213_21.

[7] H. Abdul-Wahab, D. J. Ferguson, and N. Abou-Kheir, "Assessment of panoral radiograph quality in a dental treatment center," APOS Trends in Orthodontics, vol. 6, 2016, doi: 10.4103/2321-1407.177960.

[8] N. R. Suparno, A. Faizah, and A. N. Nafisah, "Assessment of Panoramic Radiograph Errors: An Evaluation of Patient Preparation and Positioning Quality at Soelastri Dental and Oral Hospital," Open Dent J, vol. 17, no. 1, 2023, doi: 10.2174/0118742106261974230925073155.

[9] H. Sun et al., "An AI-Based Image Quality Control Framework for Knee Radiographs," J Digit Imaging, vol. 36, no. 5, 2023, doi: 10.1007/s10278-023-00853-6.

[10] K. Nousiainen, T. Mäkelä, A. Piilonen, and J. I. Peltonen, "Automating chest radiograph imaging quality control," Physica Medica, vol. 83, 2021, doi: 10.1016/j.ejmp.2021.03.014.

[11] Q. Lai et al., "Quality control of elbow joint radiography using a YOLOv8-based artificial intelligence technology," Eur Radiol Exp, vol. 8, no. 1, Dec. 2024, doi: 10.1186/s41747-024-00504-7.

[12] N. Ameli, M. Miri Moghaddam, H. Lai, and C. Pacheco-Pereira, "Automated quality evaluation of dental panoramic radiographs using deep learning," Imaging Sci Dent, vol. 55, no. 2, p. 175, 2025, doi: 10.5624/isd.20240232.

[13] H. S. Kim et al., "Refinement of image quality in panoramic radiography using a generative adversarial network," Dentomaxillofacial Radiology, vol. 52, no. 5, 2023, doi: 10.1259/dmfr.20230007.

[14] E. Delamare, X. Fu, Z. Huang, and J. Kim, "Panoramic imaging errors in machine learning model development: a systematic review," Dentomaxillofac Radiol, vol. 53, no. 3, 2024, doi: 10.1093/dmfr/twae002.

[15] M. Y. Ichahane and N. Assad, "Multi-Model Attention-Enhanced CNN Ensemble for Automated Osteoporosis Detection in Radiographic Knee Images," International Journal of Intelligent Engineering and Systems, vol. 18, no. 3, pp. 177–194, 2025, doi: 10.22266/ijies2025.0430.13.

[16] A. Fariza, R. Asmara, E. R. Astuti, and R. H. Putra, "Tooth and Supporting Tissue Anomalies Detection from Panoramic Radiography Using Integrating Convolution Neural Network with Batch Normalization," International Journal of Intelligent Engineering and Systems, vol. 17, no. 2, 2024, doi: 10.22266/ijies2024.0430.19.

[17] N. Nafi'iyah, C. Fatichah, D. Herumurti, E. R. Astuti, R. H. Putra, and A. S. Akbar, "Automatic for Generating Landmark Mandibular Panoramic Radiography Image," International Journal of Intelligent Engineering and Systems, vol. 17, no. 1, pp. 584–595, 2024, doi: 10.22266/ijies2024.0229.49.

[18] Kamal and H. Ez-zahraouy, "A comparison between the VGG16, VGG19 and ResNet50 architecture frameworks for classification of normal and CLAHE processed medical images," Res Sq, 2023.

[19] J. Lin, Y. Ma, W. Lu, Z. Qu, Z. Jin, and J. Zhou, "Panoramic radiograph quality assessment: Database and algorithm," Displays, vol. 82, 2024, doi: 10.1016/j.displa.2023.102625.

[20] J. C. Mello Román et al., "Panoramic dental radiography image enhancement using multiscale mathematical morphology," Sensors, vol. 21, no. 9, 2021, doi: 10.3390/s21093110.

[21] H. Mohammad-Rahimi et al., "Super-Resolution of Dental Panoramic Radiographs Using Deep Learning: A Pilot Study," Diagnostics, vol. 13, no. 5, 2023, doi: 10.3390/diagnostics13050996.

[22] A. C. Subbulakshmi, N. Mohan, R. Thiruneervannan, S. Naveen, and S. Gokulraj, "Positioning errors in digital panoramic radiographs: A study," J Orofac Sci, vol. 8, no. 1, 2016, doi: 10.4103/0975-8844.181922.

[23] H. Maghbuli, T. Razi, E. Banakar, P. Emamverdizadeh, and S. Razi, "Positioning Errors in Panoramic Images Based on the Dentition Type of Patients Referring to the Oral and Maxillofacial Radiology Department of Tabriz Dental School During 2017-2018," Avicenna Journal of Dental Research, vol. 15, no. 2, 2023, doi: 10.34172/ajdr.2023.535.

[24] O. Almășan et al., "Post-Traumatic-Related Technical Errors in Orthopantomographic Imaging," Medicines, vol. 9, no. 12, 2022, doi: 10.3390/medicines9120063.

[25] E. D. Costa, W. G. Cral, F. P. Murad, M. L. Oliveira, G. M. B. Ambrosano, and D. Q. Freitas, "Prevalence of Errors and Number of Retakes in Panoramic Radiography: Influence of Professional Training and Patient Characteristics," International journal of odontostomatology, vol. 15, no. 3, 2021, doi: 10.4067/s0718-381x2021000300719.

[26] R. Izzetti, M. Nisi, G. Aringhieri, L. Crocetti, F. Graziani, and C. Nardi, "Basic knowledge and new advances in panoramic radiography imaging techniques: A narrative review on what dentists and radiologists should know," Applied Sciences (Switzerland), vol. 11, no. 17, 2021, doi: 10.3390/app11177858.

[27] H. G. Yeom et al., "Correlation between spatial resolution and ball distortion rate of panoramic radiography," BMC Med Imaging, vol. 20, no. 1, 2020, doi: 10.1186/s12880-020-00472-5.

[28] H. Y. Su et al., "Fusion extracted features from deep learning for identification of multiple positioning errors in dental panoramic imaging," J Xray Sci Technol, vol. 31, no. 6, 2023, doi: 10.3233/XST-230171.