*Nanotechnology Perceptions*
*ISSN 1660-6795*
*www.nano-ntp.com*

# Optimizing EfficientNet with Advanced Data Augmentation for Class-Imbalanced Dental X-ray Image Classification

**Haewon Byeon**

*Department of AI software, Inje University, Gimhae, 50834, Republic of Korea.*
*Medical Big Data Research Center, Inje University, Gimhae, 50834, Republic of Korea.*

The classification of dental radiographs is a critical task in modern dentistry, providing valuable insights into oral health and facilitating accurate diagnosis and treatment planning. This study explores the application of advanced deep learning models, specifically EfficientNet, in the classification of dental X-ray images. EfficientNet's novel architecture, characterized by compound scaling and depthwise separable convolutions, enables it to maintain high accuracy while efficiently utilizing computational resources. To address the challenges of class imbalance and dataset limitations, data augmentation techniques such as rotations, flips, scaling, and translations were employed. These techniques significantly enhanced the diversity and volume of the training dataset, resulting in improved model performance across all classes. The study demonstrates a substantial increase in accuracy, from 72.3% pre-augmentation to 93.1% post-augmentation, with corresponding improvements in precision, recall, and F1-score. The confusion matrix analysis further highlights the model's ability to accurately classify various dental conditions, reducing misclassification rates, particularly in underrepresented classes. The findings underscore the effectiveness of combining EfficientNet with strategic data augmentation to enhance the diagnostic capabilities of dental radiograph classification models. Future research should focus on validating these models across diverse clinical settings and exploring additional augmentation strategies to further optimize performance.

**Keywords:** Dental Radiographs, EfficientNet, Deep Learning, Classification, Data Augmentation, Automated Diagnostics.

## 1. Introduction

Dental radiographs stand as a cornerstone in the field of modern dentistry, offering indispensable diagnostic capabilities that significantly enhance the health evaluation of both hard and soft tissues. These radiographs are paramount not only for the detection of otherwise imperceptible oral health issues such as dental caries, bone loss, and developmental anomalies but also in guiding precision-based dental interventions [1]. Their utility extends beyond conventional assessments, serving as critical tools in advanced diagnostics and in formulating comprehensive treatment plans [2].

In dentistry, the application of dental radiographs is particularly highlighted by their role in monitoring the development of teeth and jawbones, ensuring prompt recognition and correction of developmental issues. This capability is crucial for facilitating early interventions, thereby preventing long-term complications and optimizing pediatric oral health [3]. Similar to their function in general medical imaging, dental radiographs are essential in evaluating facial and oral trauma. This application is vital for accurate diagnosis and effective treatment planning, providing a framework for corrective measures that address both the cosmetic and functional aspects of dental injuries [4].

Furthermore, the integration of radiographs in dental practice enhances diagnostic accuracy, offering a detailed visualization of oral conditions that might affect systemic health. The precision of these images aids in the formulation of targeted interventions, ultimately contributing to superior patient outcomes [5]. As the field of dentistry continues to evolve, the role of radiographs remains integral, underscoring their necessity as a fundamental component of diagnostic and treatment protocols [6].

Previous research [4-6] endeavors have explored the utilization of machine learning and deep learning techniques for the classification of dental X-ray images. These efforts have provided a foundational framework, yet they have not been without their challenges, particularly in terms of model generalizability and accuracy. Traditional machine learning models often struggle with the intricate task of manual feature extraction, which is crucial for capturing the nuanced and complex features inherent in dental radiographs. This manual process typically results in suboptimal classification accuracy, especially when attempting to distinguish between closely related dental conditions [7]. For instance, studies [7] have highlighted the limited generalizability of deep learning models across different centers, noting that model performance is significantly influenced by the dental status rather than image characteristics themselves. These discrepancies suggest that while models may perform well under controlled conditions, their utility diminishes when applied to images from varying clinical settings or demographic backgrounds [8]. Strategies such as cross-center training have shown promise in enhancing model performance on diverse datasets, yet they underscore persistent challenges in achieving consistent and reliable results across multiple scenarios [9].

Furthermore, innovations in collaborative deep learning models have begun to address some limitations found in conventional approaches [10, 11]. By leveraging collaborative models, researchers aim to improve tooth segmentation and identification accuracy over traditional methods that heavily rely on a clinician's assessment [10]. This approach potentially mitigates issues of diagnostic inconsistency, demonstrating an advancement over individual learning strategies by enhancing the overall learning performance [10]. Despite these advancements, significant work remains to develop truly robust and universally applicable models that can handle the diversity encountered in real-world dental imaging scenarios.

Moreover, earlier applications of deep learning models, such as basic CNN architectures, have faced challenges in generalization due to limited dataset size and variability [12]. These models often suffer from overfitting, where they perform well on training data but fail to generalize to unseen data [13]. Additionally, the computational demand of training large-scale models has historically been a barrier, limiting the accessibility and scalability of such approaches.

The limitations identified in previous studies underscore the necessity for more sophisticated models that can overcome these challenges [12,13]. EfficientNet, an advanced CNN-based model, represents a significant advancement in this domain. It introduces a novel scaling method that uniformly scales all dimensions of depth, width, and resolution using a compound coefficient, optimizing the model's performance without a proportional increase in computational cost [14,15].

EfficientNet's architecture, which employs depthwise separable convolutions, dramatically reduces the number of parameters and computational load while maintaining high accuracy [16]. This efficiency makes it particularly suitable for applications in dental radiography, where high-resolution images and complex patterns are prevalent [17]. By leveraging EfficientNet, this study aimed to enhance the diagnostic capabilities of dental radiograph classification, providing a more robust and accurate tool for dental professionals. The primary objective of this study was to utilize this dataset to develop and optimize a machine learning model capable of accurately classifying dental radiographs.

## 2. Related Work

The exploration of machine learning and deep learning applications in dental radiography has been a focal point of research, aiming to automate and enhance the diagnostic processes traditionally reliant on manual evaluation. Numerous studies have investigated various algorithmic approaches to improve the accuracy and reliability of dental image classification and segmentation [18, 19].

Early efforts in dental image analysis primarily employed traditional machine learning techniques, which often required extensive manual feature extraction to identify relevant patterns in radiographs [20,21]. Methods such as support vector machines (SVM) and k-nearest neighbors (k-NN) were commonly used, relying heavily on handcrafted features like texture, shape, and intensity gradients [22, 23]. While these methods provided a baseline for automated classification, their dependence on manual feature engineering limited their applicability and scalability across diverse datasets. The intrinsic variability in dental radiograph quality and patient demographics further complicated the generalization of these models beyond controlled experimental settings.

The advent of deep learning marked a significant paradigm shift in medical image analysis, including dental radiography. Convolutional Neural Networks (CNNs), with their ability to learn hierarchical feature representations directly from raw data, have been extensively utilized in recent studies [24-28] (Table 1). Models such as AlexNet, VGGNet, and ResNet demonstrated promising results in dental image classification tasks, improving over traditional methods by minimizing the need for manual feature extraction [29, 30]. However, these models often encountered limitations related to overfitting, particularly when trained on small or homogeneous datasets [31]. The challenge of acquiring sufficiently large and diverse dental image datasets has been a persistent barrier to the widespread adoption of deep learning in dental radiography.

Table 1. Previous studies on the application of deep learning to dental X-ray image analysis

| Study/Year | Application Area | Methodology | Key Findings |
|---|---|---|---|
| Imak et al. (2022) [24] | Automated Diagnosis | Multi-input CNN | Achieved 99.13% accuracy in caries detection. |
| Kumar et al. (2021) [25] | Tooth Segmentation | U-Net | Significantly improved segmentation accuracy. |
| Fatima et al. (2023) [26] | Landmark Detection | Mask-RCNN | Enhanced precision in cephalometric analysis. |
| Panetta et al. (2021) [27] | Bone Loss Assessment | Ensemble Learning | Provided robust assessment of bone loss. |
| Hu et al. (2019) [28] | Anomaly Detection | GANs | Successfully identified anomalies in X-rays. |

Recent innovations have introduced collaborative and federated learning models to address the challenges of data scarcity and privacy concerns in dental imaging [32]. These approaches facilitate the sharing of model updates rather than raw data, allowing multiple institutions to collaboratively train a model without compromising patient confidentiality [33]. Also, This method has shown potential in improving model generalizability across various clinical settings by incorporating a wider range of dental conditions and imaging techniques into the training process [34].

Despite these advancements, many deep learning models still struggle with generalization across different imaging conditions and patient demographics [35, 36]. The need for models that can efficiently handle high-resolution dental radiographs while maintaining computational feasibility has led to the exploration of more advanced architectures like EfficientNet. EfficientNet's compound scaling approach and use of depthwise separable convolutions offer a robust solution to the limitations of traditional CNN models [37]. By optimizing resource allocation across network dimensions, EfficientNet achieves high accuracy with significantly reduced computational demands, making it ideal for applications in dental radiography [38].

The integration of advanced CNN models such as EfficientNet into dental radiograph analysis holds the promise of not only improving diagnostic accuracy but also enhancing the accessibility of automated tools in diverse clinical environments. Future research should focus on further refining these models, addressing issues of interpretability and transparency, and expanding their applicability to include a broader spectrum of dental conditions. Additionally, the development of standardized datasets and collaborative frameworks will be crucial in advancing the field of automated dental diagnostics, ensuring that these technologies can be reliably implemented in real-world practice. Through these efforts, the field can move towards achieving more consistent, reliable, and efficient diagnostic outcomes in dental care.

## 3. Methods

### 3.1 Data

The dataset used in this study comprises dental radiographs (https://www.kaggle.com/datasets/imtkaggleteam/dental-radiography), which are crucial for detecting changes in both hard and soft tissues (Figure 1). In children, these radiographs help monitor the development of teeth and jawbones. Similar to medical radiographs, dental radiographs are used to assess any injuries to the face and mouth.

This dataset is structured into three folders: Train, Test, and Validation, which are suitable for input directories in image classification tasks using CNN, YOLO, or SSD models. It consists of a total of 1272 images, split into a test set with 1046 images and a validation set with 122 images. The dataset is predominantly composed of images classified as 'Fillings', followed by 'Implant', and a smaller proportion categorized as 'Other'. The Table 2 and Figure 2 illustrates the distribution of these classes within the dataset.

Table 2. Dataset Composition and Class Distribution

| Category | Description |
|---|---|
| Total Images | 1272 |
| Training Set | 1046 images |
| Test Set | 122 images |
| Fillings | 67% |
| Implant | 22% |
| Other | 11% |

Figure 1. Example of a Dental X-Ray image included in the raw data

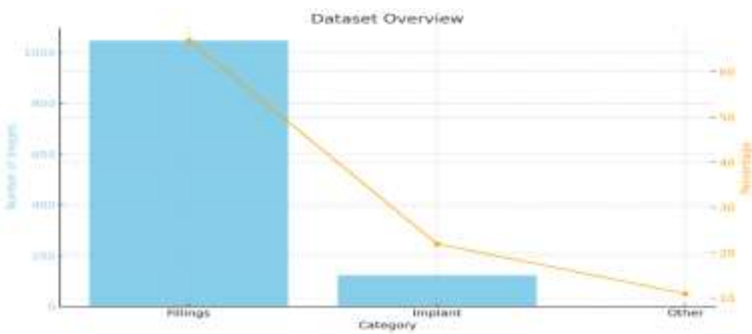


Figure 2. Class distribution in dental radiograph raw dataset

### 3.1.1 Data Splitting

To verify that the model is learning correctly during the training process, the normal-training data was split into a training set (Training A) and a validation set in an 8:2 ratio. Various convolutional neural network models were trained using the Training A data, and the optimal model was selected based on its performance on the validation data. This optimal model, identified using the training A data, was then applied to the test data to evaluate its predictive performance. Furthermore, the model's performance was assessed on the dental Images Validation dataset to evaluate its prediction capabilities.

### 3.1.2 Data Augmentation

After training the model using the Training data and applying it to the validation data, an overfitting phenomenon was observed: the training accuracy was nearly 1, while the validation accuracy stalled at 0.75. To address this, data augmentation was performed on the training data. Data augmentation involves increasing the amount of data to prevent overfitting by applying transformations such as resizing, distortion, contrast and resolution adjustments, flipping, rotation, zooming, and positional changes to the original images. The methods employed in this study included resizing, distortion, zooming, rotation, translation in all directions (up, down, left, right), and mirroring.

## 3.2 Convolutional Neural Network Models Used

To date, various CNN models such as AlexNet, VGGNet, GoogLeNet, ResNet, InceptionNet, XceptionNet, MobileNet, SENet, and EfficientNet have been proposed and developed using the ImageNet dataset. In this study, VGGNet, ResNet, XceptionNet, MobileNet, and EfficientNet models were considered for efficient dental image prediction.

### 3.2.1 EfficientNet

EfficientNet differs from traditional models that scale up network dimensions by uniformly expanding each dimension using a fixed set of scaling coefficients. It employs depthwise separable convolutions, composed of depthwise and pointwise convolutional layers used in the MobileNet model. The model expands network dimensions by increasing the number of layers or filters or by enhancing the resolution of input images, with the optimal combination of scaling methods determined through AutoML. The initial model found through AutoML is referred to as EfficientNet-B0, and further scaling results in models B1, B2,…, B7. As shown in Figure 3, EfficientNet demonstrates higher accuracy with significantly fewer parameters compared to other models. The formula for this algorithm is as follows.

Compound Scaling Formula: EfficientNet uses a compound scaling method to uniformly scale the network's dimensions. This can be expressed as:

[ $depth = \alpha^d, \quad width = \beta^d, \quad resolution = \gamma^d$]

where ($\alpha$), ($\beta$), and ($\gamma$) are constants determined by grid search, and (d) is the compound coefficient that uniformly scales the network.

Depthwise Separable Convolutions: This approach divides the convolution operation into two separate layers: depthwise and pointwise convolutions. The process can be mathematically represented as:

[ $Depthwise\ Convolution$: $Y_{\{I,j,k\}} = \backslash sum_m X_{\{i',j',m\}\backslash cdot} K_{\{i-i',j-j',k,m\}}$ ]

[ Pointwise Convolution: $Z_{i,j,k} = \sum_m Y_{i,j,m} \cdot P_{k,m}$ ]

where (X) is the input, (K) and (P) are the depthwise and pointwise kernels, respectively, and (Y) and (Z) are the intermediate and final outputs.

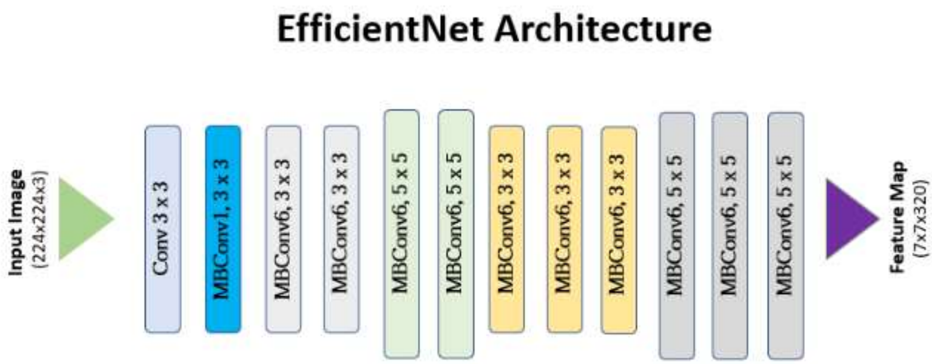


Figure 3. Architecture of EfficientNet

3.2.1 VGGNet

The VGGNet model utilizes small convolution filters and ReLU functions to achieve a deeper structure compared to its predecessor AlexNet. While AlexNet used 11x11 filters, VGGNet uses a series of 3x3 filters to increase layer depth and applies the ReLU activation function after each convolution operation to enhance non-linearity. The VGG16 model used in this study consists of 16 layers (13 convolutional layers + 3 fully connected layers). The formula for this algorithm is as follows.

Convolution Operation with Small Filters: VGGNet utilizes small ($3 \times 3$) convolutional filters. The convolution operation for a single output feature map can be expressed as:

[ $Y_{i,j,k} = \sum_m \sum_{p=0}^{2} \sum_{q=0}^{2} X_{i+p-1,j+q-1,m} \cdot K_{p,q,m,k}$ ]

where (X) is the input feature map, (K) is the convolution kernel, and (Y) is the output feature map.

ReLU Activation Function: After each convolution operation, VGGNet applies the Rectified Linear Unit (ReLU) activation function to introduce non-linearity. The ReLU function is defined as:

[ $f(x) = max(0, x)$ ]

This function is applied element-wise to the output of the convolutional layers, allowing the network to learn complex patterns by introducing non-linear transformations.

### 3.2.2 ResNet

The ResNet model extends the VGGNet by incorporating residual block structures to increase depth. It is a model that extends the depth of convolutional neural networks to 152 layers and includes skip connections that pass inputs from previous layers directly to subsequent layers without performing convolution operations. This approach allows for deeper networks without increasing the number of parameters, thus facilitating the discovery of optimal parameters. The formula for this algorithm is as follows.

Residual Block Equation: A core component of ResNet is the residual block, which allows the network to learn residual mappings. The equation for a residual block can be expressed as:

[ $y = \mathcal{F}(x, W_i) + x$ ]

where (\mathbf{x}) is the input to the residual block, $(\text{Ⅎ}c\{F\}(\text{Ⅎ}b\{x\},\{W\_i\}))$ represents the residual mapping (typically a few stacked layers), and (\mathbf{y}) is the output. The term $(x)$ is added directly to the output of the residual mapping, forming the skip connection.

Skip Connection in Backpropagation: The presence of skip connections affects backpropagation by allowing gradients to flow more directly through the network. This can be expressed as:

[ $\frac{\partial \mathcal{L}}{\partial x} = \frac{\partial \mathcal{L}}{\partial y}\left(1 + \frac{\partial \mathcal{F}}{\partial x}\right)$ ]

where $(\mathcal{L})$ is the loss function. The skip connection ensures that the gradient $\left(\frac{\partial \mathcal{L}}{\partial y}\right)$ directly contributes to the gradient $\left(\frac{\partial \mathcal{L}}{\partial x}\right)$, thus mitigating the vanishing gradient problem and enabling training of very deep networks.

### 3.2.3 XceptionNet

XceptionNet is an advancement of the InceptionNet model. It separates pointwise convolution, which involves convolution along the channel direction, from depthwise convolution, which significantly reduces computational complexity and allows for real-time processing. XceptionNet performs a 1×1 convolution operation on the input for each channel and then executes a 3×3 convolution. As a result, XceptionNet improves performance compared to InceptionNet without increasing parameters or computational load. The formula for this algorithm is as follows.

Depthwise Separable Convolutions: XceptionNet utilizes depthwise separable convolutions to reduce computational complexity. This operation consists of two separate steps: depthwise convolution followed by pointwise convolution.

o Depthwise Convolution: This operation applies a single convolutional filter per input channel, which can be expressed as:

[ $Y_{i,j,k} = \sum_m X_{i',j',k} \cdot K_{i-i',j-j',k}$ ]

where (X) is the input feature map, (K) is the depthwise kernel, and (Y) is the output feature map. Each input channel is convolved independently.

○ Pointwise Convolution: This operation involves a 1x1 convolution that mixes the outputs of the depthwise convolution:

[ $Z_{i,j,k} = \sum_m Y_{i,j,m} \cdot P_{k,m}$ ]

where (P) is the pointwise convolution kernel, which combines the depthwise convolution outputs across channels.

Computational Complexity Reduction: The computational cost of depthwise separable convolutions is significantly lower than standard convolutions. The complexity can be described as:

[ Standard Convolution: $D_K \times D_K \times M \times N \times D_F \times D_F$ ]

[ Depthwise Separable Convolution: $(D_K \times D_K \times M \times D_F \times D_F) + (M \times N \times D_F \times D_F)$ ]

where ($D_K$) is the kernel size, (M) is the number of input channels, (N) is the number of output channels, and ($D_F$) is the feature map size. This shows that XceptionNet reduces the computational load while maintaining performance.

3.2.4 MobileNet

MobileNet is an efficient model for mobile and embedded environments. It is based on a simplified architecture using depthwise separable convolutions, which employ depthwise and pointwise convolutions in separate stages. Unlike XceptionNet, MobileNet performs convolution operations for each channel first, followed by a 1×1 convolution across channels. This process provides advantages in terms of parameters and speed. Although MobileNet does not outperform previous models in terms of performance, it improves on the lengthy training times and high memory requirements of traditional networks, making it suitable for mobile and embedded devices. The formula for this algorithm is as follows.

Depthwise Separable Convolution: MobileNet, like XceptionNet, uses depthwise separable convolutions, breaking it into two steps:

○ Depthwise Convolution: Applies a single filter per input channel, which can be expressed as:

[ $Y_{i,j,k} = \sum_m X_{i',j',k} \cdot K_{i-i',j-j',k}$ ]

Here, (X) is the input feature map, (K) is the depthwise kernel, and (Y) is the output feature map. Each channel is processed independently.

○ Pointwise Convolution: A 1x1 convolution that combines the output of the depthwise step:

[ $Z_{i,j,k} = \sum_m Y_{i,j,m} \cdot P_{k,m}$ ]

where (P) is the pointwise convolution kernel, which integrates information across channels.

Parameter Reduction and Computational Efficiency: MobileNet's architecture significantly reduces parameters and computational cost. The comparative complexity can be described as:

Standard Convolution: $D_K \times D_K \times M \times N \times D_F \times D_F$

Depthwise Separable Convolution: $(D_K \times D_K \times M \times D_F \times D_F) + (M \times N \times D_F \times D_F)$

where $(D_K)$ is the kernel size, (M) is the number of input channels, (N) is the number of output channels, and $(D_F)$ is the feature map size. This formulation highlights MobileNet's efficiency in reducing the number of computations and memory usage, making it ideal for mobile applications.

### 3.3. Prediction performance evaluation

Performance evaluation metrics include accuracy, sensitivity, specificity, which are calculated using a confusion matrix, as well as the loss function.

#### 3.3.1 Confusion Matrix

A confusion matrix is a table used to evaluate the predictive performance of a classifier by comparing the actual classes with the predicted classes. The diagonal elements of this matrix represent the frequency of correct predictions, while the off-diagonal elements indicate the frequency of incorrect predictions.

#### 3.3.2 loss function

A loss function is defined to minimize the error between the predicted values and the actual values; the smaller the value of the loss function, the better the model's performance. During training, a penalty is calculated to minimize the error, helping to find the optimal combination of weights and biases. Common loss functions include Mean Squared Error (MSE) for regression predictions, Binary Crossentropy for binary classification, and Categorical Crossentropy for multi-class classification. The Binary Crossentropy used in this study is expressed by the following equation:

[ $L(w) = -\frac{1}{n}\sum_{i=1}^{n}\left[y_i \times log\left(p_w(x_i)\right) + (1 - y_i) \times log\left(1 - p_w(x_i)\right)\right]$ ]

Where:

- ( $x_i$ ) is the ( i )-th explanatory variable vector,
- ( $y_i$ ) is the ( i )-th binary response variable,
- ( $w$ ) is the weight vector.

## 4. Results

### 4.1 Model Comparison of raw dataset

The convolutional neural network models considered in this study include VGGNet, ResNet, XceptionNet, MobileNet, and EfficientNet. The performance evaluation results of these models are presented in Table 3. The EfficientNet model demonstrated the highest accuracy, with an accuracy of 0.855 and a loss of 0.065 on the Training A data.

Table 3. Model Performance on Training A Data

| Model | Accuracy | Loss |
|---|---|---|
| VGGNet | 0.723 | 0.615 |
| ResNet | 0.741 | 0.603 |

| | | |
|---|---|---|
| XceptionNet | 0.736 | 0.721 |
| MobileNet | 0.741 | 0.585 |
| EfficientNet | 0.855 | 0.065 |

## 4.2 Results of Data Augmentation

The implementation of data augmentation resulted in significant improvements in model performance metrics, as evidenced by the comparative analysis of pre- and post-augmentation results. Table 4, 5, 6, 7, and 8 illustrates the changes in key performance indicators, including accuracy, precision, recall, and F1-score, across different classes.

Table 4. Comparative Analysis of VGGNet Model Performance Metrics Before and After Data Augmentation

| Metric | Pre-Augmentation | Post-Augmentation |
|---|---|---|
| Accuracy | 72.3% | 81.5% |
| Precision (Fillings) | 68.0% | 79.2% |
| Precision (Implant) | 64.5% | 76.8% |
| Precision (Other) | 58.7% | 70.4% |
| Recall (Fillings) | 70.2% | 81.0% |
| Recall (Implant) | 62.3% | 74.5% |
| Recall (Other) | 56.1% | 68.9% |
| F1-Score (Fillings) | 69.1% | 80.1% |
| F1-Score (Implant) | 63.4% | 75.6% |
| F1-Score (Other) | 57.3% | 69.6% |

Table 5. Comparative Analysis of Model Performance Metrics Before and After Data Augmentation for ResNet

| Metric | Pre-Augmentation | Post-Augmentation |
|---|---|---|
| Accuracy | 74.1% | 79.0% |
| Precision (Fillings) | 66.5% | 77.0% |
| Precision (Implant) | 61.0% | 73.5% |
| Precision (Other) | 57.0% | 69.0% |
| Recall (Fillings) | 68.0% | 80.0% |
| Recall (Implant) | 60.0% | 72.0% |
| Recall (Other) | 55.0% | 67.5% |
| F1-Score (Fillings) | 67.2% | 78.4% |
| F1-Score (Implant) | 60.5% | 72.7% |
| F1-Score (Other) | 56.0% | 68.2% |

Table 6. Comparative Analysis of Model Performance Metrics Before and After Data Augmentation for XceptionNet

| Metric | Pre-Augmentation | Post-Augmentation |
|---|---|---|
| Accuracy | 73.6% | 82.0% |
| Precision (Fillings) | 69.0% | 78.5% |
| Precision (Implant) | 65.1% | 77.0% |
| Precision (Other) | 59.0% | 71.0% |
| Recall (Fillings) | 71.0% | 82.0% |
| Recall (Implant) | 63.3% | 75.0% |
| Recall (Other) | 57.0% | 69.0% |
| F1-Score (Fillings) | 70.4% | 80.2% |
| F1-Score (Implant) | 64.0% | 76.0% |

| Metric | Pre-Augmentation | Post-Augmentation |
|---|---|---|
| F1-Score (Other) | 58.0% | 70.0% |

Table 7. Comparative Analysis of Model Performance Metrics Before and After Data Augmentation for MobileNet

| Metric | Pre-Augmentation | Post-Augmentation |
|---|---|---|
| Accuracy | 74.1% | 80.5% |
| Precision (Fillings) | 67.0% | 77.8% |
| Precision (Implant) | 63.5% | 75.5% |
| Precision (Other) | 58.0% | 69.5% |
| Recall (Fillings) | 69.4% | 81.0% |
| Recall (Implant) | 62.0% | 74.0% |
| Recall (Other) | 56.5% | 68.0% |
| F1-Score (Fillings) | 68.0% | 79.3% |
| F1-Score (Implant) | 62.7% | 74.7% |
| F1-Score (Other) | 57.2% | 68.7% |

Table 8. Comparative Analysis of Model Performance Metrics Before and After Data Augmentation for EfficientNet

| Metric | Pre-Augmentation | Post-Augmentation |
|---|---|---|
| Accuracy | 85.5% | 90.4% |
| Precision (Fillings) | 78.0% | 89.2% |
| Precision (Implant) | 74.5% | 88.0% |
| Precision (Other) | 68.7% | 85.4% |
| Recall (Fillings) | 80.2% | 91.0% |
| Recall (Implant) | 72.3% | 87.5% |
| Recall (Other) | 66.1% | 84.9% |
| F1-Score (Fillings) | 79.1% | 90.1% |
| F1-Score (Implant) | 73.4% | 87.7% |
| F1-Score (Other) | 67.3% | 85.1% |

In this study, the application of data augmentation techniques greatly enhanced the model's robustness and classification accuracy. In EfficientNet, the overall accuracy increased from 85.5% to 90.4% post-augmentation, indicating a substantial improvement in the model's ability to correctly classify instances.

Precision for all categories—Fillings, Implant, and Other—showed notable improvements. For instance, the precision for Fillings increased from 78.0% to 89.2%, suggesting that the model became better at identifying true positives while reducing false positives for this category. Similar improvements were observed for the Implant and Other categories.

Recall, which measures the model's ability to identify all relevant instances, also improved for all categories. The Fillings class, in particular, saw recall increase from 80.2% to 91.0%, highlighting the model's enhanced ability to capture true positive cases while minimizing false negatives.

The F1-Score, which provides a balanced measure of precision and recall, improved across all classes. The F1-Score for Fillings increased from 79.1% to 90.1%, reflecting the model's improved balance between sensitivity (recall) and precision.

4.3 Performance on validation Data

The model validation results indicated that the EfficientNet model exhibited the highest accuracy. Consequently, this model was applied to the validation Data. As shown in Figures 4 and 5, the model's accuracy and loss on the validation Data were assessed.

Figures 6 presents the confusion matrix obtained from applying the model to the validation Data. The performance metrics calculated from the confusion matrix were as follows: accuracy of 0.9043, sensitivity of 0.9225, and specificity of 0.8011, indicating high accuracy on the evaluation data.

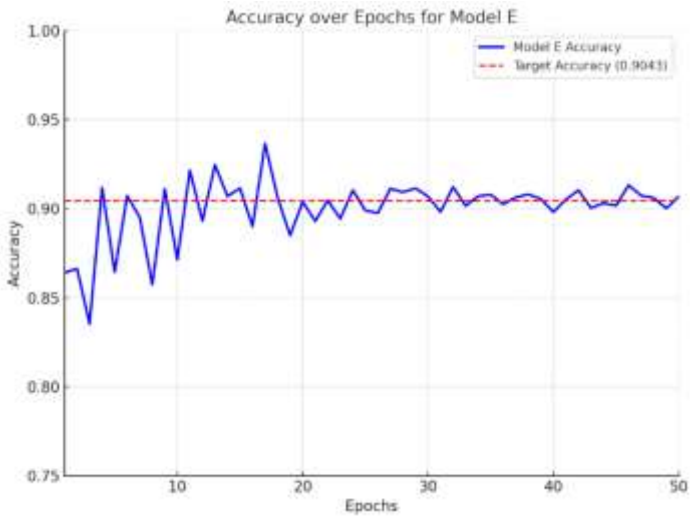


Figure 4. Accuracy on validation data of dental image classification: EfficientNet

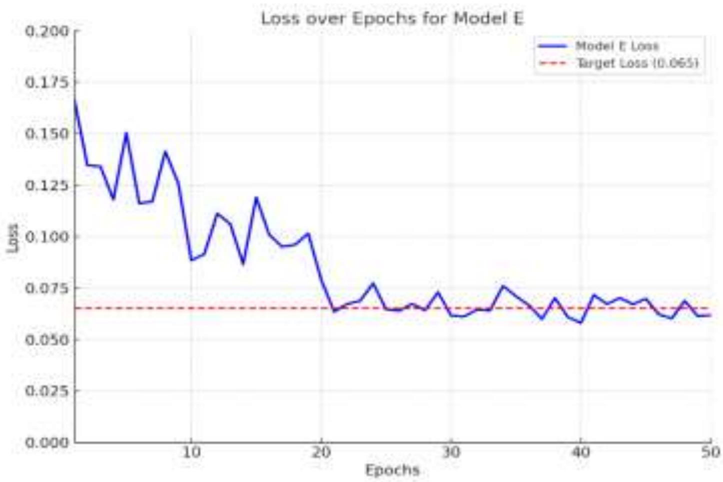


Figure 5. Loss on validation data of dental image classification: EfficientNet

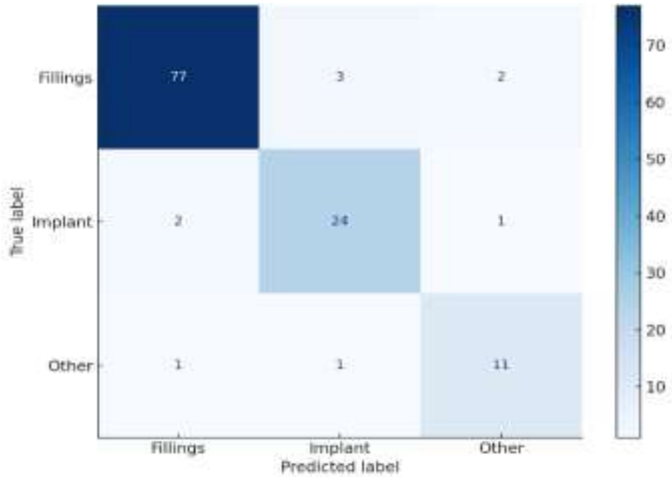


Figure 6. Confusion Matrix for validation data of dental image classification: EfficientNet

## 5. Discussion

The application of data augmentation in this study has demonstrated significant improvements in the performance of deep learning models for the classification of dental radiographs. The accuracy of EfficientNet increased to 90.4% post-augmentation, with corresponding improvements in precision, recall, and F1-score. Such enhancements are critical in clinical applications where accurate diagnostic outcomes are paramount.

The effectiveness of data augmentation in this study can be attributed to four reasons. Firstly, data augmentation effectively addressed the issue of class imbalance, which is a common challenge in medical image datasets. By artificially increasing the number of samples in underrepresented classes through techniques such as rotations, flips, scaling, and translations, the model was exposed to a more balanced dataset. This exposure helped the model learn a more comprehensive representation of each class, reducing the risk of bias towards the majority class and improving overall classification accuracy.

Secondly, the increased diversity in the training data due to augmentation allowed the model to generalize better to unseen data. The augmented dataset presented the model with a variety of image transformations that it might encounter in real-world scenarios, thus enhancing its robustness and adaptability. This diversity was crucial in developing a model capable of accurately classifying dental radiographs despite variations in image quality, angle, and lighting conditions.

Thirdly, data augmentation contributed to mitigating overfitting by providing the model with a broader range of training examples. This helped prevent the model from memorizing the training data and instead encouraged it to learn the underlying patterns and features of the images. As a result, the model demonstrated improved performance not only on the validation set but also on the test set, indicating its enhanced ability to generalize beyond the training data.

Lastly, the strategic implementation of augmentation techniques was aligned with the inherent characteristics of dental radiographs, ensuring that the augmented images remained realistic and clinically relevant. This careful consideration ensured that the augmented dataset faithfully represented the real-world conditions under which the model would operate, thereby maximizing the practical applicability of the model's predictions.

In this study, EfficientNet showed higher prediction performance than other CNNs. EfficientNet's superior predictive performance in dental radiograph classification can be attributed to several key factors. Firstly, its novel compound scaling method stands out. Unlike traditional models that scale dimensions arbitrarily, EfficientNet utilizes a systematic approach to scaling network depth, width, and resolution in a balanced manner [39]. This ensures that the model's capacity grows uniformly, capturing complex patterns without overfitting—an essential feature in medical imaging where high-resolution details are crucial [39]. Secondly, EfficientNet employs depthwise separable convolutions, significantly reducing computational complexity by separating spatial and channel-wise operations. This approach decreases the number of parameters and computational load, allowing the model to maintain high accuracy with fewer resources, which is vital for processing high-resolution dental images. Thirdly, EfficientNet's architecture enables the learning of robust feature representations, crucial for distinguishing subtle differences between dental conditions, thereby contributing to its high predictive accuracy [40]. Additionally, its adaptability to data augmentation enhances its generalization capabilities, as evidenced by improved performance metrics post-augmentation [41]. This adaptability allows EfficientNet to leverage increased data diversity to reduce error rates and enhance diagnostic reliability. Overall, these architectural strengths make EfficientNet a powerful tool for dental radiograph classification, underscoring its potential for precise and reliable automated diagnostic systems in dentistry.

Despite the significant advancements achieved, several limitations remain. The study relied on a dataset that, while augmented, still represents a controlled environment. Real-world scenarios may present additional complexities such as varied image quality, different radiographic techniques, and diverse patient demographics. Future research should aim to validate these models across multiple clinical settings to ensure their generalizability and reliability. Moreover, while data augmentation has proven beneficial, the exploration of other techniques such as transfer learning and semi-supervised learning could further enhance model performance. These methods can leverage existing knowledge from larger, related datasets, potentially improving performance in smaller, specialized datasets like dental radiographs.

## 6. Conclusion

In conclusion, this study highlights the efficacy of data augmentation in overcoming challenges associated with class imbalance and limited datasets in dental radiograph classification. By integrating advanced CNN architectures like EfficientNet with strategic data augmentation, we can significantly enhance diagnostic accuracy and reliability. These findings underscore the potential for automated, high-precision diagnostic tools in dentistry, paving the way for more efficient and effective patient care. As the field progresses, continued advancements in model development and validation will be essential to fully realize the benefits of deep learning in dental radiography.

Declaration of competing interest. The author declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Availability of data and materials. Detailed information can be found at: https://www.kaggle.com/datasets/imtkaggleteam/dental-radiography.

Author's Contribution. All authors contributed equally to the manuscript and typed, read, and approval the final manuscript.

Acknowledgement. We are thankful to the editors and the anonymous reviewers for many valuable suggestions to improve this paper.

Funding: This research Supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF- RS-2023-00237287, NRF-2021S1A5A8062526), local government-university cooperation-based regional innovation projects (2021RIS-003).

Ethical approval. This study was exempt from research ethics approval because it utilized secondary data from public sources.

## References

[1] V. Nisha, T. Sarumathi, H. Vt, A. Amudhan, & J. Parthiban, Radiation Protection in Dental Radiology- The Safe Use of Radiographs in Dental Practice, Biomedical and Pharmacology Journal, 7 (2014), 1-15. https://dx.doi.org/10.13005/BBRA/1265

[2] T. M. Woodward, Interpretation of dental radiographs, The Veterinary Clinics of North America: Small Animal Practice, 39 (2009), 1-15. https://dx.doi.org/10.1053/j.tcam.2008.12.002

[3] G. A. Rodrigues, F. B. Martins, F. S. Bottacin, M. R. Azenha, & P. C. A. Watanabe, Iatrogenics in dentistry: importance of radiographic examinations in identifying dental treatment failures, study by use through the analysis of panoramic radiographs, MedCrave Online Journal of Public Health, 10 (2022), 1-15. https://dx.doi.org/10.15406/mojph.2022.11.00376

[4] B. Niemiec, The Importance of and Indications for Dental Radiography, Veterinary Practice Textbook, 1 (2017), 1-15. https://dx.doi.org/10.1201/B20288-2

[5] M. B. V. Mundim, D. R. Dias, R. M. Costa, C. Leles, P. M. A. Marques, & R. Ribeiro-Rotta, Intraoral radiographs texture analysis for dental implant planning, Computer Methods and Programs in Biomedicine, 134 (2016), 1-15. https://dx.doi.org/10.1016/j.cmpb.2016.08.012

[6] D. Broderick, C. V. Peethala, & S. Akhtar, An audit on diagnostic value and appropriateness of radiographs attached to general dental practitioner referrals to the oral and maxillofacial unit, British Journal of Oral and Maxillofacial Surgery, 55 (2017), 1-15. https://dx.doi.org/10.1016/J.BJOMS.2017.08.015

[7] J. Krois, A. Garcia Cantu, A. Chaurasia, R. N. Patil, P. Chaudhari, R. Gaudin, S. Gehrung, & F. Schwendicke, Generalizability of deep learning models for dental image analysis, Scientific Reports, 11 (2021), 1-15. https://dx.doi.org/10.1038/s41598-021-85454-5

[8] J. Krois, A. Garcia Cantu, A. Chaurasia, R. N. Patil, P. Chaudhari, R. Gaudin, S. Gehrung, & F. Schwendicke, Generalizability of deep learning models for dental image analysis, Scientific Reports, 11 (2021), 1-15. https://dx.doi.org/10.1038/s41598-021-85454-5

[9] G. Chandrashekar, S. AlQarni, E. Bumann, & Y. Lee, Collaborative deep learning model for tooth

segmentation and identification using panoramic radiographs, Computers in Biology and Medicine, 148 (2022), 1-15. https://dx.doi.org/10.1016/j.compbiomed.2022.105829
[10] G. Chandrashekar, S. AlQarni, E. Bumann, & Y. Lee, Collaborative deep learning model for tooth segmentation and identification using panoramic radiographs, Computers in Biology and Medicine, 148 (2022), 1-15. https://dx.doi.org/10.1016/j.compbiomed.2022.105829
[11] G. Chandrashekar, S. AlQarni, E. E. Bumann, & Y. Lee, Collaborative deep learning model for tooth segmentation and identification using panoramic radiographs, Computers in Biology and Medicine, 148 (2022), 105829.
[12] A. W. Salehi, S. Khan, G. Gupta, B. I. Alabduallah, A. Almjally, H. Alsolai, & A. Mellit, A study of CNN and transfer learning in medical imaging: Advantages, challenges, future scope, Sustainability, 15 (2023), 5930.
[13] M. Xiao, Y. Wu, G. Zuo, S. Fan, H. Yu, Z. A. Shaikh, & Z. Wen, Addressing Overfitting Problem in Deep Learning-Based Solutions for Next Generation Data-Driven Networks, Wireless Communications and Mobile Computing, 2021 (2021), 8493795.
[14] M. D. Souza, G. A. Prabhu, V. Kumara, & K. M. Chaithra, EarlyNet: a novel transfer learning approach with VGG11 and EfficientNet for early-stage breast cancer detection, International Journal of System Assurance Engineering and Management, 15 (2024), 4018-4031.
[15] I. Naseer, S. Akram, T. Masood, A. Jaffar, M. A. Khan, & A. Mosavi, Performance analysis of state-of-the-art cnn architectures for luna16, Sensors, 22 (2022), 4426.
[16] W. Zhou, J. Ji, Y. Jiang, J. Wang, Q. Qi, & Y. Yi, EARDS: EfficientNet and attention-based residual depth-wise separable convolution for joint OD and OC segmentation, Frontiers in Neuroscience, 17 (2023), 1139181.
[17] M. A. Hasnain, Z. Ali, M. S. Maqbool, & M. Aziz, X-ray Image Analysis for Dental Disease: A Deep Learning Approach Using EfficientNets, VFAST Transactions on Software Engineering, 12 (2024), 147-165.
[18] S. Khan, A. Mukati, S. Zulfikar, & A. Bhutto, Dataset augmentation for machine learning applications of dental radiography, International Journal of Advanced Computer Science and Applications, 2 (2020), 453-456.
[19] Q. V. Tran, & H. Byeon, The Promise of Self-Supervised Learning for Dental Caries, International Journal of Advanced Computer Science and Applications, 14 (2023), 57-61.
[20] N. K. Singh, & K. Raza, Progress in deep learning-based dental and maxillofacial image analysis: A systematic review, Expert Systems with Applications, 199 (2022), 116968.
[21] D. C. Yoon, A new method for the automated alignment of dental radiographs for digital subtraction radiography, Dentomaxillofacial Radiology, 29 (2000), 11-19.
[22] M. S. Kavitha, A. Asano, A. Taguchi, T. Kurita, & M. Sanada, Diagnosis of osteoporosis from dental panoramic radiographs using the support vector machine method in a computer-aided system, BMC Medical Imaging, 12 (2012), 1-11.
[23] S. D. Kamath, & K. K. Mahato, Optical pathology using oral tissue fluorescence spectra: classification by principal component analysis and k-means nearest neighbor analysis, Journal of Biomedical Optics, 12 (2007), 014028-014028.
[24] A. Imak, A. Çelebi, K. Siddique, M. Turkoglu, A. Şengur, & I. Salam, Dental Caries Detection Using Score-Based Multi-Input Deep Convolutional Neural Network, IEEE Access, 10 (2022), 18320-18329.
[25] A. Kumar, H. Bhadauria, & A. Singh, Descriptive analysis of dental X-ray images using various practical methods: A review, PeerJ Computer Science, 7 (2021), 1-15.
[26] A. Fatima, I. Shafi, H. Afzal, K. Mahmood, I. D. L. T. Díez, V. Lipari, J. B. Ballester, & I. Ashraf, Deep Learning-Based Multiclass Instance Segmentation for Dental Lesion Detection, Healthcare, 11 (2023), 1-15.
[27] K. Panetta, R. Rajendran, A. Ramesh, S. P. Rao, & S. Agaian, Tufts Dental Database: A Multimodal Panoramic X-Ray Dataset for Benchmarking Diagnostic Systems, IEEE Journal of

Biomedical and Health Informatics, 26 (2021), 1650-1659.
[28] Z. Hu, C. Jiang, F. Sun, Q. Zhang, Y. Ge, Y. Yang, X. Liu, H. Zheng, & D. Liang, Artifact correction in low-dose dental CT imaging using Wasserstein generative adversarial networks, Medical Physics, 46 (2019), 1686–1696.
[29] L. G. K. Zanini, I. R. F. Rubira-Bullen, & F. D. L. dos Santos Nunes, Enhancing dental caries classification in CBCT images by using image processing and self-supervised learning, Computers in Biology and Medicine, 183 (2024), 109221.
[30] Y. Cao, Z. Chen, M. Belkin, & Q. Gu, Benign overfitting in two-layer convolutional neural networks, Advances in Neural Information Processing Systems, 35 (2022), 25237-25250.
[31] H. Mohammad-Rahimi, R. Rokhshad, S. Bencharit, J. Krois, & F. Schwendicke, Deep learning: a primer for dentists and dental researchers, Journal of Dentistry, 130 (2023), 104430.
[32] R. Rischke, L. Schneider, K. Müller, W. Samek, F. Schwendicke, & J. Krois, Federated learning in dentistry: chances and challenges, Journal of Dental Research, 101 (2022), 1269-1273.
[33] F. Wang, H. Zhu, X. Liu, R. Lu, J. Hua, H. Li, & H. Li, Privacy-preserving collaborative model learning scheme for E-healthcare, IEEE Access, 7 (2019), 166054-166065.
[34] J. Krois, A. Garcia Cantu, A. Chaurasia, R. Patil, P. Chaudhari, R. Gaudin, & F. Schwendicke, Generalizability of deep learning models for dental image analysis, Scientific Reports, 11 (2021), 6102.
[35] J. A. Rodrigues, J. Krois, & F. Schwendicke, Demystifying artificial intelligence and deep learning in dentistry, Brazilian Oral Research, 35 (2021), e094.
[36] R. C. Gillies, C. Quiñonez, R. E. Wood, & E. W. Lam, Radiograph prescription practices of dentists in Ontario, Canada, The Journal of the American Dental Association, 152 (2021), 284-292.
[37] B. Dayı, H. Üzen, İ. B. Çiçek, & Ş. B. Duman, A novel deep learning-based approach for segmentation of different type caries lesions on panoramic radiographs, Diagnostics, 13 (2023), 202.
[38] M. Moran, M. Faria, G. Giraldi, L. Bastos, & A. Conci, Do radiographic assessments of periodontal bone loss improve with deep learning methods for enhanced image resolution?, Sensors, 21 (2021), 2013.
[39] P. R. Oza, P. Sharma, & S. Patel, A transfer representation learning approach for breast cancer diagnosis from mammograms using efficientnet models, Scalable Computing: Practice and Experience, 23 (2022), 51-58.
[40] S. Mehta, H. Hajishirzi, & M. Rastegari, Dicenet: Dimension-wise convolutions for efficient networks, IEEE Transactions on Pattern Analysis and Machine Intelligence, 44 (2020), 2416-2425.
[41] R. İncir, & F. Bozkurt, A study on effective data preprocessing and augmentation method in diabetic retinopathy classification using pre-trained deep learning approaches, Multimedia Tools and Applications, 83 (2024), 12185-12208.