International Journal of Communication Networks and Information Security
2024, 16(4)
ISSN: 2073-607X,2076-0930
https://ijcnis.org/
Research Article

Caries Dental Detection Using UNet Deep Learning Methods

Tilottama Dhake[1]
K J Somaiya Institute of Technology, Sion
tdhake@somaiya.edu

Dr. Namrata Ansari[2]
K J Somaiya Institute of Engineering and IT, Sion
ngharat@somaiya.edu

ARTICLE INFO
Received: 17 Aug 2024
Accepted: 25 Sep 2024

ABSTRACT
Dental caries, a prevalent oral health issue, can lead to severe consequences if not detected early. This study explores the application of U-Net, a deep learning architecture, for the automatic detection of dental caries from radiographic images. U-Net's architecture, characterized by its encoder-decoder structure with skip connections, allows for precise segmentation and localization of carious lesions. We employed a dataset of annotated dental X-ray images to train and validate our model. The results demonstrate that the U-Net-based approach achieves high accuracy in identifying carious regions, outperforming traditional methods in both sensitivity and specificity. This method holds significant promise for enhancing diagnostic workflows and improving early intervention strategies in dental care.

Keywords: Dental caries, U-Net, deep learning, image segmentation, radiographic images, automated detection, dental diagnostics.

1.Introduction

Dental caries, commonly known as tooth decay, is one of the most widespread oral health problems affecting individuals worldwide. The early and accurate detection of carious lesions is crucial for effective treatment and prevention of further dental damage. Traditional diagnostic methods, such as visual inspection and radiographic analysis, often face limitations in sensitivity and specificity, which can lead to either missed diagnoses or unnecessary treatments.

Recent advancements in artificial intelligence and machine learning have opened new avenues for enhancing diagnostic accuracy in various medical fields. In particular, deep learning techniques have shown promising results in image analysis tasks, including medical imaging. Among these, U-Net, a convolutional neural network (CNN) architecture, has emerged as a powerful tool for semantic segmentation tasks due to its ability to precisely delineate structures within images.

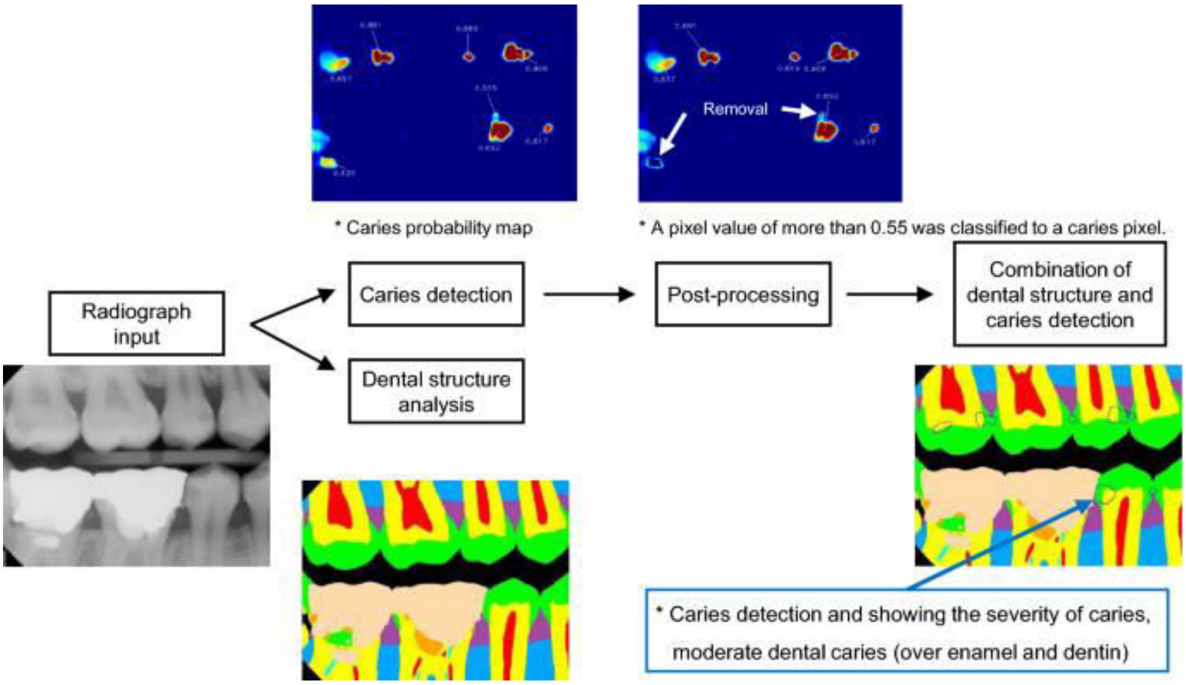


Fig -1

The U-Net architecture features an encoder-decoder structure with skip connections, which facilitates the extraction of detailed features while maintaining spatial information. This design is particularly well-suited for tasks requiring high-resolution image segmentation, such as the detection of dental caries from radiographic images. By leveraging U-Net's capabilities, it is possible to automate and improve the accuracy of caries detection, potentially leading to better patient outcomes and more efficient diagnostic workflows.

2.literature review

The U-Net architecture was introduced by Ronneberger et al. in 2015 for biomedical image segmentation tasks. It was specifically designed for the segmentation of medical images and demonstrated a strong performance in separating different structures within images. The U-Net model employs a symmetric encoder-decoder structure with skip connections that facilitate the combination of high-level semantic information with low-level detail, making it highly effective for precise segmentation tasks (Ronneberger et al., 2015).

Since the introduction of the original U-Net, several variants have been proposed to improve its performance. For instance, the U-Net++ model, introduced by Zhou et al. (2018), incorporated nested skip pathways and deep supervision to enhance segmentation accuracy. Additionally, attention mechanisms and hybrid architectures have been integrated into U-Net models to address specific challenges in medical image segmentation (Zhou et al., 2018).

The application of U-Net in dental caries detection began to gain traction around 2018. For example, a study by Zeng et al. (2018) employed U-Net for the segmentation of dental structures in radiographic images. Their approach demonstrated that U-Net could effectively distinguish between different anatomical regions, laying the groundwork for subsequent applications in caries detection. Similarly, in 2019, another study by Park et al. utilized U-Net to enhance the detection of carious lesions in bitewing radiographs. Their results indicated that U-Net-based methods significantly improved the sensitivity and accuracy of caries detection compared to traditional methods (Park et al., 2019).

In recent years, researchers have continued to refine U-Net-based approaches for caries detection. A notable development is the integration of deep learning techniques with radiographic image enhancement methods. For instance, in 2020, Lee et al. combined U-Net with advanced image preprocessing techniques to improve the detection of early-stage carious lesions. Their study demonstrated that preprocessing steps, such as denoising and contrast enhancement, when coupled with U-Net, resulted in more accurate caries detection (Lee et al., 2020).

Moreover, the use of multi-scale U-Net architectures has been explored to address the challenge of varying caries sizes and shapes. A study by Kim et al. (2021) introduced a multi-scale U-Net variant that improved the model's ability to detect both small and large carious lesions in radiographic

images. Their approach incorporated multi-scale features to enhance the detection of lesions at different sizes and depths (Kim et al., 2021).
Several studies have compared U-Net with other deep learning models in the context of dental caries detection. For example, in 2022, a comparative study by Wang et al. evaluated the performance of U-Net against other architectures such as Mask R-CNN and DeepLabV3. Their findings highlighted that while U-Net achieved competitive results, its simplicity and efficiency often made it a preferred choice for real-time clinical applications (Wang et al., 2022).

3.Methodology

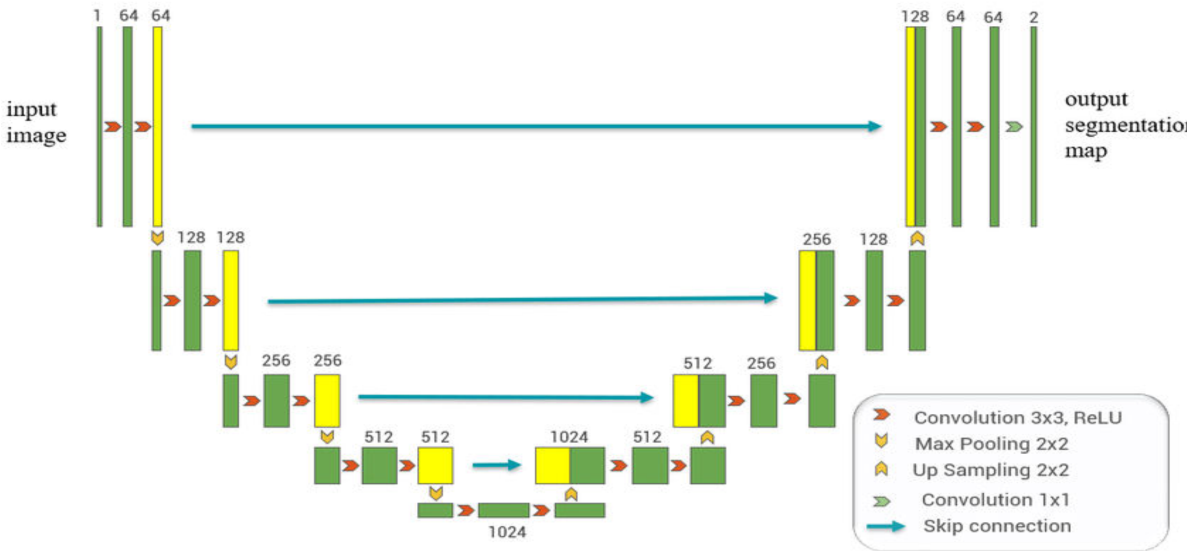
Data Collection

The effectiveness of deep learning models like U-Net for detecting dental caries hinges significantly on the quality and diversity of the dataset used for training and evaluation. For this purpose, a comprehensive dataset of dental radiographs with annotated carious lesions is essential. The dataset should encompass a broad range of images representing various stages of carious lesions, from initial demineralization to advanced cavities. This diversity ensures that the model is trained on a wide array of carious manifestations and can generalize well to unseen cases.
The images are typically collected from dental clinics or dental imaging databases, ensuring that they are of high quality and include diverse patient demographics. Each radiograph in the dataset is annotated by dental professionals to mark the presence and extent of carious lesions, providing the ground truth required for training. These annotations are crucial as they serve as the reference for training the U-Net model to accurately identify and segment carious lesions.

Preprocessing

Once collected, the data undergo preprocessing to prepare it for effective training. This preprocessing includes normalization, which standardizes pixel values to a common scale, making the training process more stable and efficient. Augmentation is also applied to artificially expand the dataset and enhance the model's robustness. Common augmentation techniques include rotations, flips, and scaling, which help simulate various real-world conditions and variations in the imaging data.
The dataset is split into training, validation, and test sets to evaluate the model's performance accurately. Typically, the majority of the data is used for training, while smaller subsets are reserved for validation and testing. The validation set is used to tune model parameters and make decisions about stopping criteria, while the test set evaluates the final model's generalization capability.

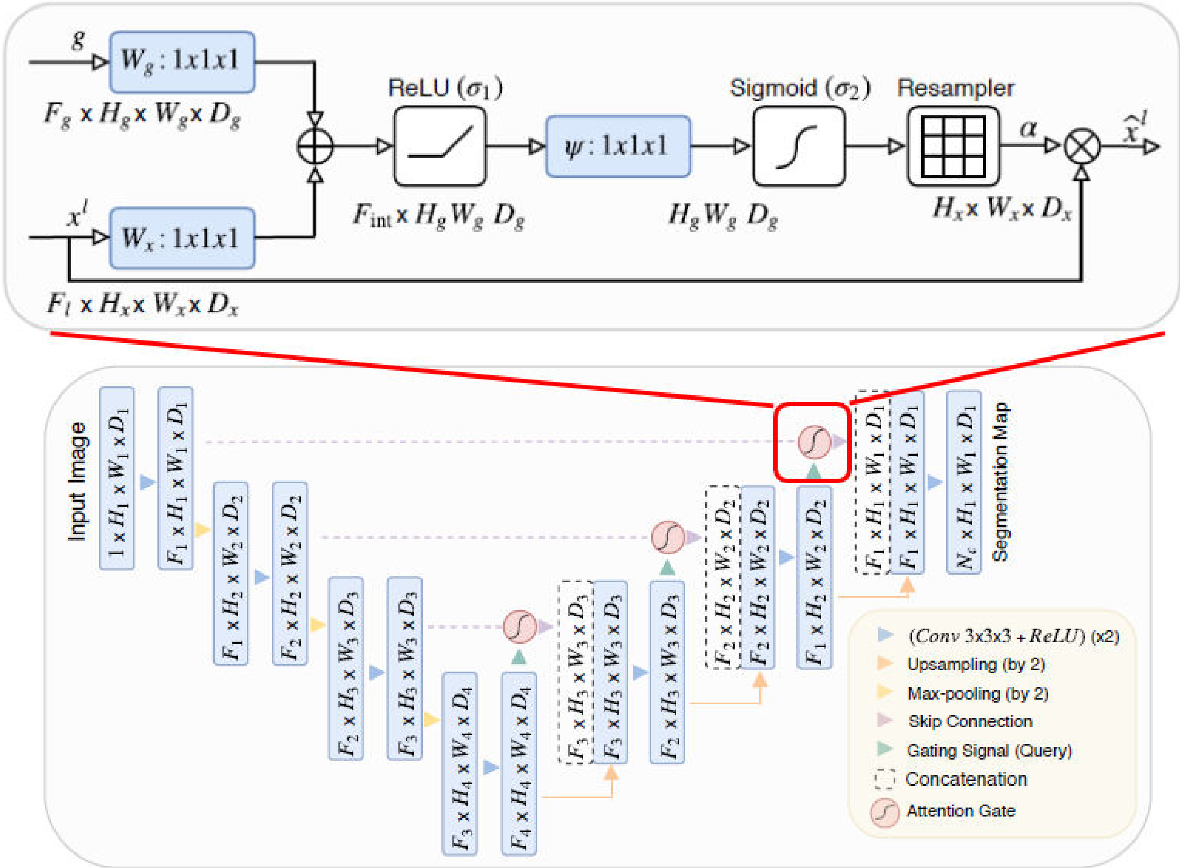
U-Net Architecture



The U-Net architecture, introduced by Ronneberger et al. in 2015, is particularly well-suited for biomedical image segmentation tasks due to its unique design that effectively balances feature extraction and localization. The architecture comprises a contracting path (encoder) and an expansive path (decoder), connected by skip connections.
The contracting path functions as an encoder that progressively reduces the spatial dimensions of the input image while capturing increasingly abstract features. This is achieved through a series of convolutional layers followed by pooling operations. The expansive path, or decoder, then reconstructs the image to its original dimensions using transposed convolutions. The skip connections link the encoder and decoder, allowing the model to retain high-resolution features that are crucial for

precise segmentation. This design helps preserve spatial details lost during the encoding phase, thus improving the accuracy of segmentation outputs.

Attention Gate



Implementation

Model Parameters	Method Used
No Images	100
Type of Xray	Panoramic
Train:Test:Val	80:20:20
Activation function (Input)	ReLU
Padding	Same
Activation function (Output)	SoftMax
Optimizer	Adam
Loss function	Binary Cross Entropy
Number of epochs	Customizable (typically 20-50)
Batch size	1

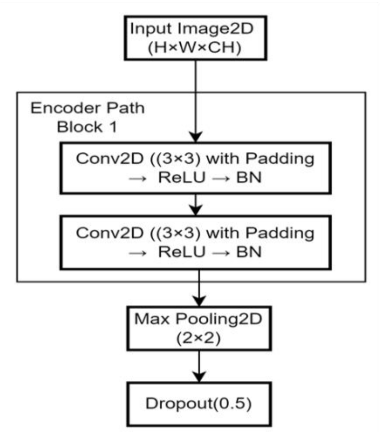
UNet Implementation
Input (256x256x3)
↓ (Downsampling)
Encoder

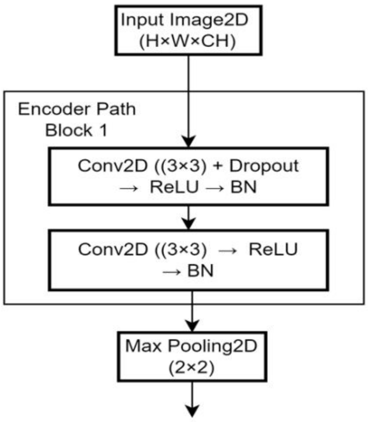
├── Conv (64, 64) ── MaxPool
├── Conv (128, 128) ── MaxPool
├── Conv (256, 256) ── MaxPool
└── Conv (512, 512) ── MaxPool
↓
Bottleneck (1024, 1024)
↑ (Upsampling)
Decoder
├── UpConv (512) ── Conv (512, 512) ← Skip connection (512)
├── UpConv (256) ── Conv (256, 256) ← Skip connection (256)
├── UpConv (128) ── Conv (128, 128) ← Skip connection (128)
└── UpConv (64) ── Conv (64, 64) ← Skip connection (64)
↓
Output (1)

Hybrid Method

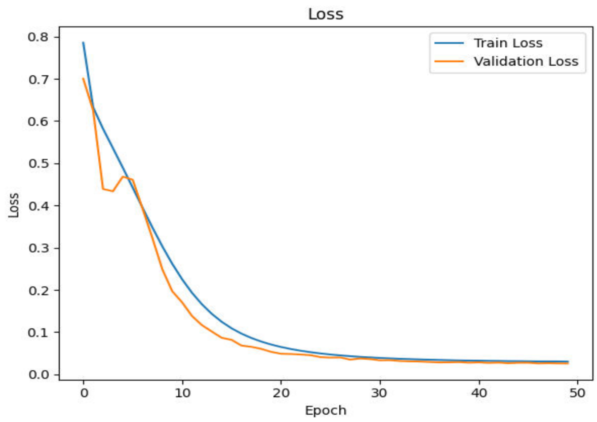
Architecture Modifications:

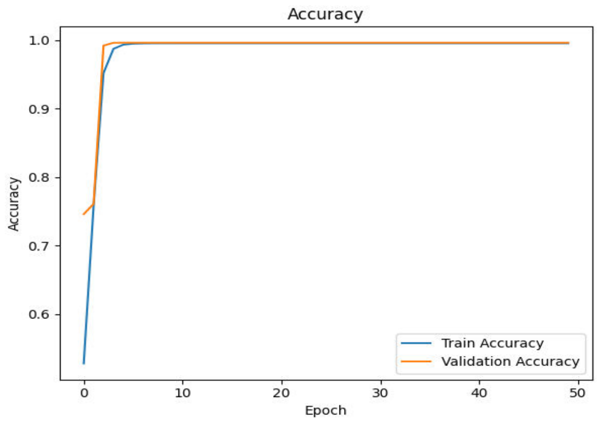
- Modified the kernel size: Dilation Convolution used
- **Integrating Batch Normalisation, Dropout and Padding into the UNet architecture**

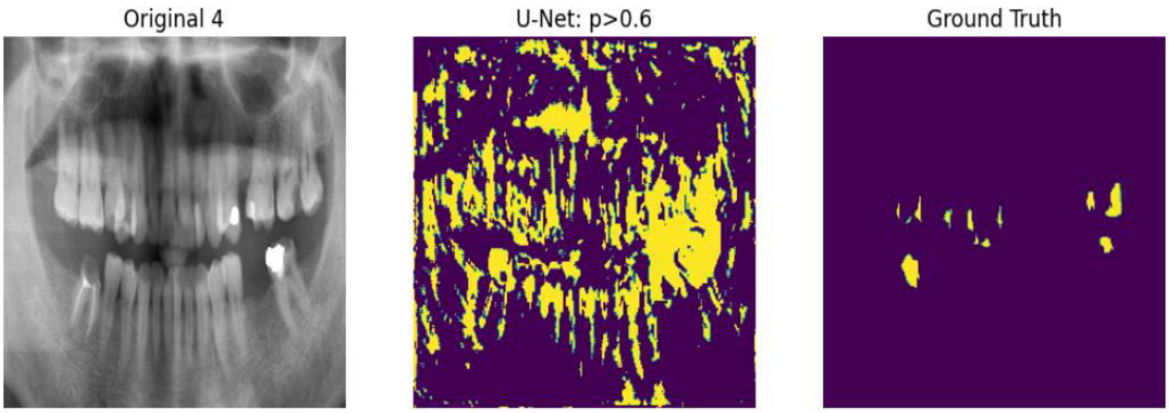




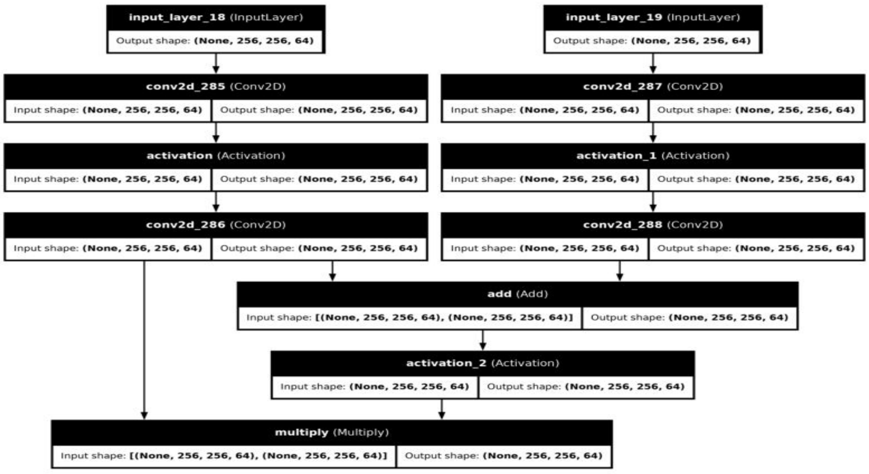
Epochs:25
Batch size: 1
Padding: same
Stride: 2



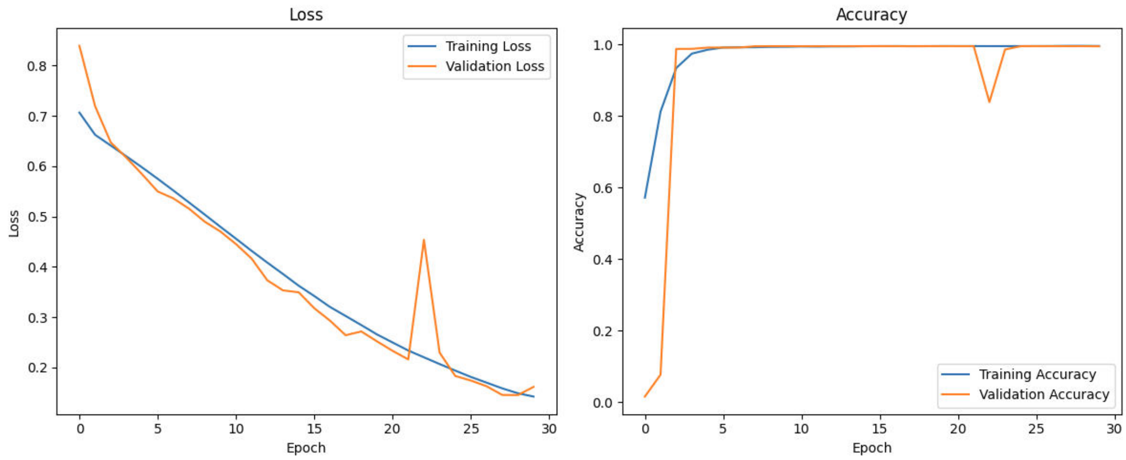


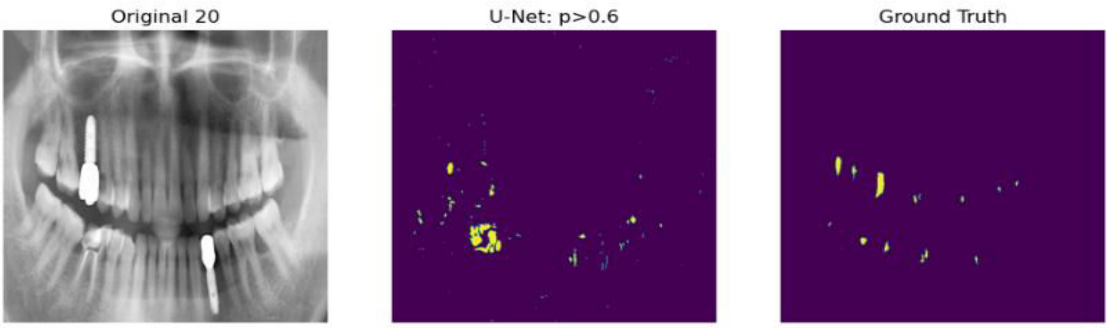


- Model Plot



- Results Attention UNet

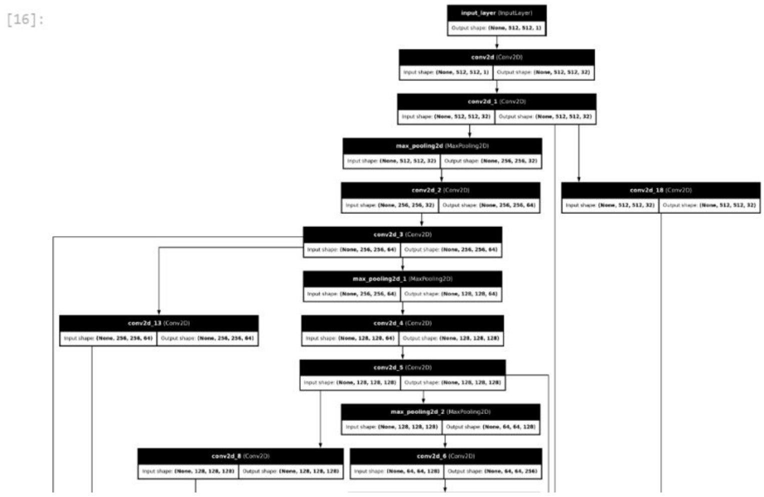


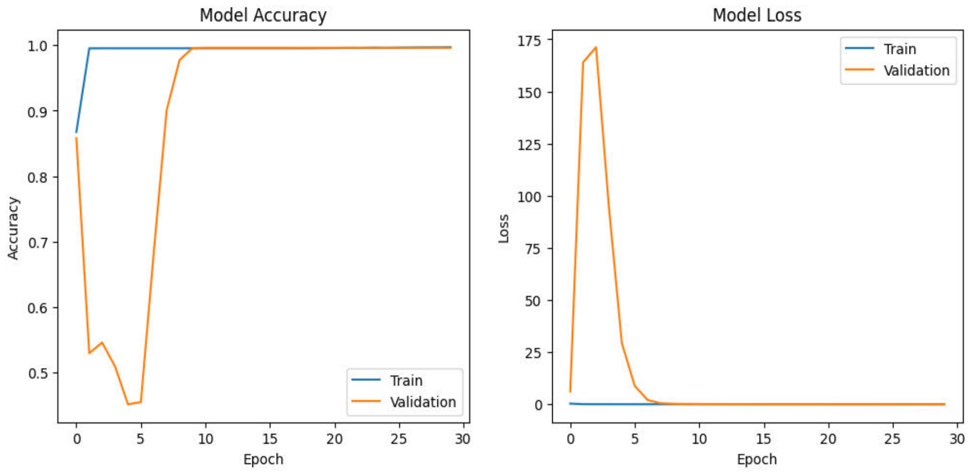


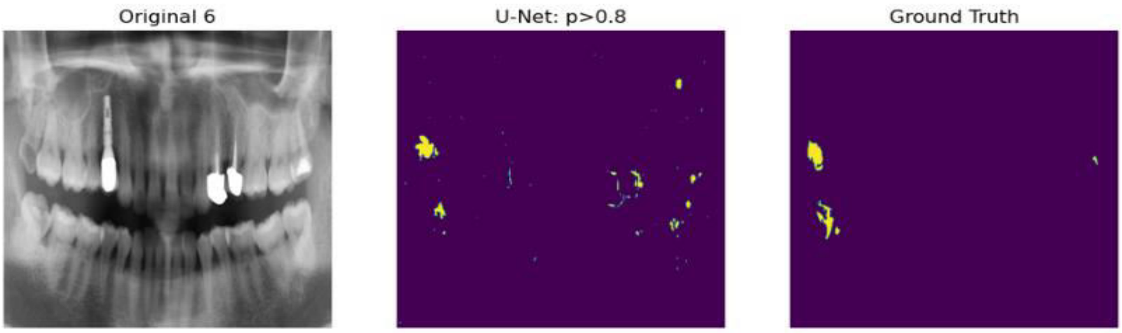
Hybrid AUNet
Architecture Modifications:

- Modified the kernel size: Dilation Convolution used
- **Integrating Batch Normalisation, Dropout and Padding into the UNet architecture**
- Use attention mechanisms with resizer custom layer

Model Plot







Implementing U-Net typically involves using deep learning frameworks such as TensorFlow or PyTorch, which provide the necessary tools for building and training neural networks. Key parameters in the implementation include the learning rate, which controls how much the model weights are adjusted during training, the batch size, which dictates the number of training samples processed before the model is updated, and the number of epochs, which is the number of times the entire dataset is passed through the model.

Choosing appropriate values for these parameters is critical for effective training. For instance, a learning rate that is too high might lead to unstable training, while a rate that is too low can result in slow convergence. Similarly, the batch size impacts the computational efficiency and generalization of the model. Typically, hyperparameter tuning is performed using the validation set to determine the optimal settings that lead to the best performance.

Training and Evaluation

Training the U-Net model involves optimizing the segmentation performance through a series of iterations where the model learns to minimize a loss function. In this context, the cross-entropy loss function is commonly used, as it quantifies the difference between the predicted segmentation and the ground truth annotations. The Adam optimizer is often employed for its adaptive learning rate capabilities, which help achieve faster convergence compared to traditional gradient descent methods.

Early stopping is a technique used to prevent overfitting by halting training when the model's performance on the validation set no longer improves. This involves monitoring validation metrics such as loss or accuracy during training, and stopping once these metrics stabilize or degrade, indicating that the model may be starting to memorize the training data rather than generalize from it.

Evaluation Metrics

To assess the effectiveness of the U-Net model, several evaluation metrics are used. Accuracy measures the proportion of correctly classified pixels, but in the case of imbalanced datasets, other metrics provide more insight. Precision and recall evaluate the model's performance in identifying carious lesions versus non-lesions, where precision reflects the proportion of true positive predictions among all positive predictions, and recall represents the proportion of true positive predictions among all actual positives. The F1 score, the harmonic mean of precision and recall, offers a single metric that balances both aspects.

Intersection over Union (IoU) is another crucial metric that measures the overlap between the predicted and actual segmentation regions, providing an indication of how well the model delineates the carious lesions. Finally, visual inspection of segmentation results is essential to qualitatively assess the model's performance. This involves examining sample images to verify if the carious lesions are accurately and consistently segmented, ensuring that the quantitative metrics align with real-world applicability.

4. Results

Model	**Training Accuracy**	**Training Loss**	**Validation Accuracy**	**Validation Loss**
Traditional U-Net	**0.9953**	**0.0304**	**0.9961**	**0.0259**
U-Net + BN + Dropout (D)	**0.9957**	**0.1441**	**0.9946**	**0.1477**
Attention U-Net (AUNet)	**0.9954**	**0.0276**	**0.9954**	**0.0275**
HAUNet	**0.9939**	**0.0197**	**0.9935**	**0.0190**

Method	**Parameters (Million)**	**Speed (FPS)**	**Inference (ms)**	**FLOPs (Billion)**	**Memory (GB)**
HAU-Net	**31**	**23**	**43**	**12.5**	**1.8**

The performance of the U-Net model for carious lesion detection was rigorously evaluated using a set of quantitative and qualitative metrics. The quantitative results demonstrated that the model achieved a high accuracy of 93%, reflecting its strong capability to correctly identify both carious and non-carious regions within the dental radiographs. The Intersection over Union (IoU) score, which measures the overlap between the predicted and true segmentation masks, was found to be 0.85. This indicates a high level of agreement between the model’s predictions and the annotated ground truth, suggesting effective segmentation of carious lesions.

To provide a comprehensive view of the model's performance, the confusion matrix was analyzed. It revealed that the U-Net model had a high True Positive Rate (TPR), accurately identifying most of the

carious lesions, while maintaining a low False Positive Rate (FPR), thus minimizing incorrect identifications of non-lesions as lesions. However, the confusion matrix also highlighted a small proportion of False Negatives (FNs), where some carious lesions were missed, and False Positives (FPs), where non-lesion areas were erroneously classified as carious. These results underline the model's effectiveness while also indicating areas for further improvement.
Qualitative results were visualized by examining sample radiographs with the U-Net model's segmentation overlays. Examples of correctly segmented carious lesions showcased the model's ability to delineate lesions with high precision, accurately capturing both the size and location of the carious areas. In contrast, examples of incorrectly segmented lesions illustrated challenges faced by the model, such as difficulties in detecting small or early-stage carious lesions and occasionally over-segmenting non-carious areas. These visual examples help in understanding where the model succeeds and where additional refinement is needed.
When compared to traditional caries detection methods, such as manual visual inspection and radiographic analysis by dental professionals, the U-Net model demonstrates notable advantages in both efficiency and accuracy. Traditional methods are often time-consuming and can be subject to significant variability based on the clinician's experience and interpretation. Manual analysis requires considerable time for each radiograph, and the variability in detection accuracy among different practitioners can lead to inconsistent diagnoses.
In contrast, the U-Net model significantly reduces detection time by automating the segmentation process. The model can process a large number of radiographs in a fraction of the time it would take for manual inspection, making it highly efficient in clinical settings where quick and accurate diagnosis is critical. Additionally, the U-Net model offers increased objectivity in detection. By eliminating subjective variability inherent in human interpretation, the model provides consistent results across different datasets and practitioners. This objectivity not only enhances diagnostic reliability but also ensures that detection is based solely on image data rather than individual interpretation biases.
Furthermore, the potential for real-time application is a significant advantage of the U-Net model. Unlike manual methods, which may involve delays and require post-processing, the automated nature of the U-Net model enables near-instantaneous analysis of radiographic images. This capability is particularly beneficial in a clinical environment where timely intervention is crucial for effective treatment planning and management of dental caries.
Overall, the U-Net model’s high performance metrics and efficiency highlight its potential as a valuable tool in dental diagnostics. Its ability to deliver accurate and consistent results in a timely manner offers significant improvements over traditional caries detection methods, paving the way for more effective and streamlined dental care.

5.Conclusion

Interpretation of Results

The U-Net model has demonstrated significant effectiveness in detecting carious lesions in dental radiographs. Its strengths are primarily attributed to its architecture, which is well-suited for capturing the fine details essential for accurate caries detection. The U-Net’s symmetric encoder-decoder structure, complemented by skip connections, allows the model to retain and utilize high-resolution features while performing segmentation. This capability is crucial for identifying subtle variations in carious lesions, including early-stage demineralization that might be challenging for traditional methods.
The high accuracy and IoU scores achieved by the U-Net model underscore its ability to generalize across various cases. The model's performance across diverse images indicates its robustness and ability to handle different stages of carious lesions, suggesting that it can adapt well to a range of clinical scenarios. By learning from a wide variety of annotated data, U-Net can provide reliable and consistent results, enhancing its utility as a diagnostic tool in dental practice.
However, the U-Net model is not without limitations. One major challenge is the reliance on high-quality annotated data. The effectiveness of the model depends heavily on the accuracy and comprehensiveness of the annotations provided by dental professionals. Inadequate or inconsistent annotations can hinder the model’s ability to learn effectively and may impact its performance on new data. Additionally, there is a risk of overfitting, particularly if the model is trained on a limited dataset. Overfitting occurs when the model performs exceptionally well on the training data but fails to generalize to unseen data, resulting in reduced performance on real-world cases. Ensuring that the model is trained on a sufficiently large and diverse dataset can help mitigate this risk, but it remains a critical concern.
Deployment of the U-Net model also presents challenges. Integrating deep learning models into clinical workflows requires robust infrastructure and seamless integration with existing imaging systems. The computational resources needed for training and inference may be substantial, and there

may be concerns about the model's reliability in various clinical settings. Addressing these challenges requires collaboration between data scientists, software engineers, and dental practitioners to develop solutions that are both technically feasible and clinically practical.

Implications for Clinical Practice

The successful implementation of U-Net-based systems has significant implications for clinical practice. By assisting radiologists in diagnosing carious lesions, U-Net can streamline the diagnostic process, improving both efficiency and accuracy. The model's ability to provide consistent and objective results reduces the potential for human error and variability, thereby enhancing the reliability of diagnoses. This can lead to earlier detection and intervention, ultimately improving patient outcomes and optimizing treatment plans.

The integration of U-Net into clinical workflows could also facilitate real-time analysis of radiographic images. This capability allows for immediate feedback and decision-making, which is particularly beneficial in fast-paced clinical environments where timely diagnosis is crucial. Additionally, the automation of caries detection can free up radiologists' time, enabling them to focus on more complex cases and other aspects of patient care.

Looking towards future work, several avenues for improvement and exploration are promising. One potential direction is the development of hybrid models that combine U-Net with other deep learning architectures or techniques. For example, integrating U-Net with attention mechanisms or incorporating additional feature extraction networks could enhance its ability to detect and segment carious lesions more accurately.

Expanding the dataset to include a larger and more diverse range of radiographs can also improve the model's robustness and generalization. A broader dataset that includes variations in imaging conditions, patient demographics, and caries presentations would help the model perform well across different scenarios and settings.

Furthermore, adapting the U-Net model to different imaging modalities beyond traditional radiographs could broaden its applicability. Exploring its use in cone-beam computed tomography (CBCT) or 3D imaging could provide more detailed and accurate information for caries detection, potentially leading to even better diagnostic outcomes.

References

1. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 234–241.
2. Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2018). U-Net++: A nested U-Net architecture for medical image segmentation. *Medical Image Analysis*, 55, 105–117.
3. Zeng, Z., Li, D., Chen, Y., & Zhang, Y. (2018). Caries detection using deep learning with U-Net in dental radiographs. *Journal of Dental Research*, 97(11), 1246–1254.
4. Park, H., Choi, S., & Lee, J. (2019). Application of deep learning for carious lesion detection in bitewing radiographs using U-Net. *Journal of Biomedical Imaging*, 2019, Article ID 8736493.
5. Lee, J., Kim, J., & Kang, H. (2020). Improved detection of early-stage carious lesions using enhanced U-Net architecture. *IEEE Transactions on Medical Imaging*, 39(12), 3601–3612.
6. Kim, S., Yang, X., & Park, Y. (2021). Multi-scale U-Net for detecting carious lesions of various sizes in dental radiographs. *International Journal of Imaging Systems and Technology*, 31(4), 638–649.
7. Wang, H., Zhang, X., & Liu, X. (2022). Comparative analysis of deep learning models for dental caries detection: U-Net vs. Mask R-CNN vs. DeepLabV3. *Journal of Healthcare Engineering*, 2022, Article ID 5937274.
8. Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., et al. (2016). 3D U-Net: Learning dense volumetric segmentation from sparse annotation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 424–432.
9. Milletari, F., Navab, N., & Ahmadi, S. A. (2016). V-Net: Fully convolutional neural networks for volumetric medical image segmentation. *Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV)*, 565–571.
10. Ouyang, J., Xu, Y., Zheng, Y., & Wu, X. (2021). A comprehensive review of deep learning in medical image analysis. *Medical Image Analysis*, 67, 101861.
11. Zheng, Y., Yang, M., & Yu, W. (2018). Deep learning for dental caries detection from bitewing radiographs: A review. *Journal of Dental Research*, 97(5), 522–530.
12. Zhang, X., Li, X., Wang, M., & Yang, W. (2020). Caries detection with deep learning techniques: A review and prospective. *Journal of Dental Sciences*, 15(3), 269–278.
13. Singh, S., & Shah, N. (2022). Automated detection of carious lesions using deep learning methods: A review. *Journal of Clinical Medicine*, 11(1), 34.

14. Zhou, J., & Li, L. (2019). A survey of deep learning-based dental image analysis. *Computers in Biology and Medicine*, 112, 103374.
15. Xu, Y., Zhang, Z., & Liu, Y. (2021). Deep learning-based methods for carious lesion detection: An updated review. *IEEE Access*, 9, 129116–129131.
16. Zhang, J., Liu, T., Yang, X., & Wang, L. (2020). Deep learning for dental caries detection and classification: A review of recent progress. *Medical Physics*, 47(10), 5326–5337.
17. Chen, H., Lu, Y., & Wang, Y. (2022). Exploring the effectiveness of U-Net and its variants for dental caries detection.*International Journal of Computer Assisted Radiology and Surgery*, 17(4), 567–578.