DEVELOPMENT OF AN AI TOOL TO IDENTIFY TEETH FROM 3D IMAGING

Final Report

Data Science Capstone Project

DATA5709

1. Ka Wing Cheng (520409224)

ABSTRACT

To develop an e-learning platform to enhance students' self-directed learning on the anatomical and morphological features of teeth, this report started the investigation on AI models to classify different types of teeth and segment features from 3D images. There have been many convolutional neural network (CNN) applications on the classifications of 2D and 3D teeth computed tomography (CT) images and the segmentation of 2D teeth images. However, the CNN applications on segmentation of tooth features from 3D teeth CT images are rare. Therefore, the development roadmap of the segmentation of tooth features is proposed. The methodology of 3D features segmentation and annotation was developed. With reference to recent 3D medical images segmentation, nnU-Net was identified as a potential AI model. A classification model was developed as a starter model to assist the segmentation task. Among several commonly used CNN architectures and some modified 3D architectures were tried, video ResNet achieved a micro F1-score of 0.793478 on test data and was adopted for the model of the classification task.

TABLE OF CONTENTS

1. INTRODUCTION

In dentomaxillofacial radiology, computed tomography (CT) or cone beam computed tomography (CBCT) is used to capture the three-dimensional structures of teeth, jaw, and surrounding structures. In a dental CT scan, many different anatomical and morphological features can be observed on a tooth. Such features are different for different tooth types. Certain types of teeth can only have some kinds of features but not the others. Meanwhile, even for some common features, they can have different specific variations on different teeth. Therefore, tooth morphology plays an important role in dental science.

However, tooth morphology is a difficult topic for dentistry students. A common traditional way of teaching and learning is to use plastic teeth. Plastic teeth can be fabricated quantitatively mimicking real human teeth. The plastic models have advantages over real human teeth in hygienic and ethical aspects but lack the variations observed with extracted human teeth. (Lone, McKenna, Cryan, Downer, & Toulouse, 2018) With the technology advancement, there are quests for 3D model applications on learning tools in tooth morphology.

Artificial intelligence (AI) can be useful for developing such 3D model applications. The development of AI has entered an explosive regime globally. With inspiration from human neuron operation, deep learning makes use of artificial neural networks with one or more than one hidden layer of artificial neural networks between the input and output layers, have done excellent work on computer vision among different kinds of machine learning algorithms. (Shen, Wu, & Suk, 2017) Different architectures of convolutional neural networks (CNNs) being applied in medical imaging with record-breaking performance (Singh et al., 2020).

For the AI applications in dentomaxillofacial radiology, there have been applications on 2D and 3D teeth classification using CNNs. (Schwendicke, Golla, Dreher, & Krois, 2019) There are also CNN applications on 2D tooth structure segmentation. (Schneider et al., 2022) However, there are very few applications of AI on 3D tooth CT scans. This knowledge gap hampered the development of 3D models in tooth morphology and has hindered on the teaching and learning process in dental education. Without 3D tooth morphological models, it is difficult to create any online learning tools to facilitate the self-directed learning of dentistry students. Therefore, in this report, deep learning models are developed for classification of teeth, and methods are investigated for segmentation of tooth structures. The resultant classification model produced will assist the segmentation of tooth structures in the future and will be further implemented on a e-learning platform for Sydney Dental School to enhance dentistry student's learning experience.

2. RELATED LITERATURE

2.1 Literature Review

For the general review on CNNs for dental images, work by Schwendicke et al., 2019 summarised 36 research publications published globally between 2015 – 2019. The reviewed studies covered different aspects of dentistry and types of tasked involved with CNNs. Different types of dental images, including CT, CBCT, panoramic, periapical radiograph, bitewing radiograph, and others are also mentioned for the studies. From the review, for classification tasks, AlexNet (Krizhevsky, Sutskever, & Hinton, 2012), VGGNet (Simonyan & Zisserman, 2014), ResNet (He, Zhang, Ren, & Sun, 2016), DenseNet (G. Huang, Liu, Van Der Maaten, & Weinberger, 2017) and GoogleNet (Szegedy et al., 2015) were the most used CNN architectures. Following the review, it was found that AlexNet has been used for CBCT teeth classification with a 7-classes accuracy of 91.0% (Yuma et al., 2017).

For the segmentation task, there were some studies for 2D teeth segmentation. Panoramic X-ray images were segmented with modified U-Net (Ronneberger, Fischer, & Brox, 2015) with an mean DICE score of 0.744 (Wirtz, Mirashi, & Wesarg, 2018), and ResNet with a F1-score of 0.88 (Jader et al., 2018). There was a segmentation on periapical images for oriented tooth proposals with a mean average best overlap of 0.7138 (Eun & Kim, 2016). For segmentation of tooth structures, i.e., caries, enamel, dentin, pulp, crown, restoration, and root canal treatment, U-Net has been used for bitewing images with a F1-score of 0.567 (Ching-Wei et al., 2016). There was also a benchmarking (Schneider et al., 2022) of various CNN architectures for bitewing images for enamel, dentin, pulp, fillings, and crowns segmentation with a median F1-score 0.86 with U-Net++ (Zhou, Rahman Siddiquee, Tajbakhsh, & Liang, 2018), U-Net, and LinkNet (Chaurasia & Culurciello, 2017).

There has been plenty of segmentation studies for 3D medical images but little for dentistry. A key review (Singh et al., 2020) summarised the state-of-the-art CNN architectures for 3D medical imaging segmentation. There were different CNN architectures in the brain tumour segmentation (BraTS) challenges and most winners were 3D U-Nets based. There were MRI data of glioma patients in the BraTS dataset (Menze et al., 2015). 3D Slicer (Fedorov et al., 2012), an open-source software, was used for the annotation of the 3D tumour structures. An example of brain tumour annotation of the BraTS dataset is shown in Figure 1. It can be concluded that 3D U-Nets were the best performers of CNN architectures.

More advanced deep CNNs can be found with more recent BraTS challenges. nnU-Net (Isensee, Jäger, Full, Vollmuth, & Maier-Hein, 2021) was the champion BraTS 2020. In the latest BraTS 2021 challenge, the winner was still using nnU-Net (Luu & Park, 2022).

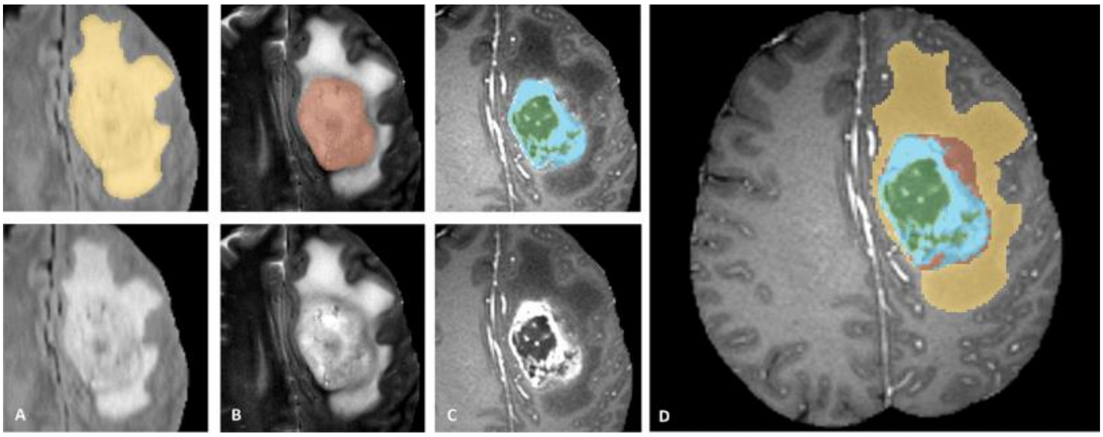

Figure 1. Manual annotation of brain tumour structures in BraTS using 3D Slicer. (Menze et al., 2015)

3. RESEARCH/PROJECT PROBLEMS

3.1 Research/Project Aims & Objectives

The aim this project was to develop an AI model for identification of the anatomical and morphological features from the 3D X-ray CT images of human primary and permanent teeth. The AI model will be useable by Sydney Dental School's front-end developer for building a e-learning platform, which will serve the learning of teeth structures from CT scans for dentistry students.

3.2 Research/Project Questions

Which AI strategies should be used for the AI tool? For the 3D images of X-ray CT scans, what pre-processing steps must be taken by the expert of Sydney Dental School to provide data annotations of the tool's development?

3.3 Research/Project Scope

To develop an AI model for the identification of the anatomical and morphological features from the 3D images of human primary and permanent in X-ray CT scans using deep learning algorithms in Python. For the data needed for the development, provide data annotation instructions to Sydney Dental School's experts.

4. METHODOLOGIES

4.1 Methods

In view of the success of nnU-Net in brain tumour segmentation (Luu & Park, 2022), it will be used for differentiating different anatomical and morphological structures of human primary and permanent teeth from 3D X-ray images. However, since tooth morphology features are far more numerous than brain tumour structures, it is a much harder task than the segmentation in BraTS challenges. In BraTS challenges, there are 4 sub-structures of the whole tumour. However, in teeth morphology, there are at least 46 structures to be identified. Therefore, it will need a large sample size, and a huge labour input for 3D annotations. Hence the development time will be long. Here, a two-stage strategy is used to achieve the objective. Since some tooth morphological features are only present on certain types of teeth, the classification of teeth type can precede the segmentation task. A deep CNN model is first developed to classify different types of teeth from X-ray CT images. The developed model can be integrated into the future segmentation model. Since the development of models are done on Pytorch, the resultant model can be saved into a model file with .pth extension using the torch.save() function. The model can be reloaded with Pytorch on Python, C++, or Java platforms.

For the classification model, the three most frequently used CNN architectures from the review (Schwendicke et al., 2019), namely AlexNet, VGGNet, and ResNet have been adopted as the candidate models. Since these models were originally designed for 2D colour image classification, a trick has been used to adapt the 3D single channel data for the 2D 3-channels models. A 2D convolution filter nn.Conv2d, with a 1 × 1 kernel size and stride of (1, 1) without padding, has been applied on the 3D CT data to extract the features from the height and reduce the height dimension into 3 to match the 3-channels requirements of the 2D models. In addition to adaptation approach, the AlexNet model was also modified into another 3D AlexNet by replacing all the 2D convolution filters, maximum and adaptive average pooling layers by their 3D versions. Moreover, because of the similarity of data structures between video and volumetric data, one existing 3D variation of ResNet used in video application (Tran et al., 2018) has also been used. The 3D ResNet model with 18 layers was originally designed for action recognition. There is only one adaptation of CT images, which is to change the number of channels from 1 to 3 for colourful videos nowadays. The adaption used the same trick of convolution filter but now a 3D 1 × 1 × 1 convolution filter nn.Conv3d is used to manipulate the number of channels.

4.2 Data Collection

Around 300 extracted teeth have been listed in an array and scanned by CT by the Sydney Dental School. The CT images were stored in DICOM (Digital Imaging and Communications in Medicine) format. More extracted teeth samples are to be scanned in the future.

For the classification task, each tooth was isolated from the array as shown in Figure 2. Here a simple thresholding method was used to separate the teeth from the background. By setting the threshold of voxel intensity 500, the teeth can be differentiated well from the low-density supporting materials. Individual tooth isolation was done by the library connected-components-3d (cc3d) for volumes with at least 1,000 voxels. Isolated tooth images were saved in .nrrd (Nearly Raw Raster Data) format, which can be easily opened by 3D Slicer. The cross-section plot of an example of an isolated tooth is shown in Figure 3. The labelling of tooth type was then done by the Sydney Dental School. The labels and the corresponding encoding in the programming are shown in Table 1. The data were then split into training, validation, and test sets using sklearn stratified train_test_split to preserve label's ratio as much as possible in each data set. After data splitting, the data were resized such that each volume can be bounded by [128, 64, 64]. The volumes were then padded to the size of [128, 91, 91]. All data sets were fed into the model in a batch of 4 volumes and the voxel values of each batch was normalised by the batch's mean and standard deviation as following:

$$\text{Normalised Image} = \frac{\text{Image} \ - \text{mean}}{\text{standard deviation}}.$$

For the training set, additional data augmentation was performed. Each 3D images were rotated between [-90, 90] degrees randomly. This is desirable as a tooth viewed from different angle should have the same conclusion on the same tooth type. However, since a human has symmetrical left and right sides, no flipping or mirroring was done to prevent flipping of left and right. Figure 4 shows the cross-section of a padded and rotated tooth image.

The whole dataset was split into the training, validation, and test sets. The training set was for the training and building of model, and the validation set was used for hyper-parameters tuning. The test set was used for evaluation of the final performance, which was not involved in any training and hyper-parameters tuning processes.

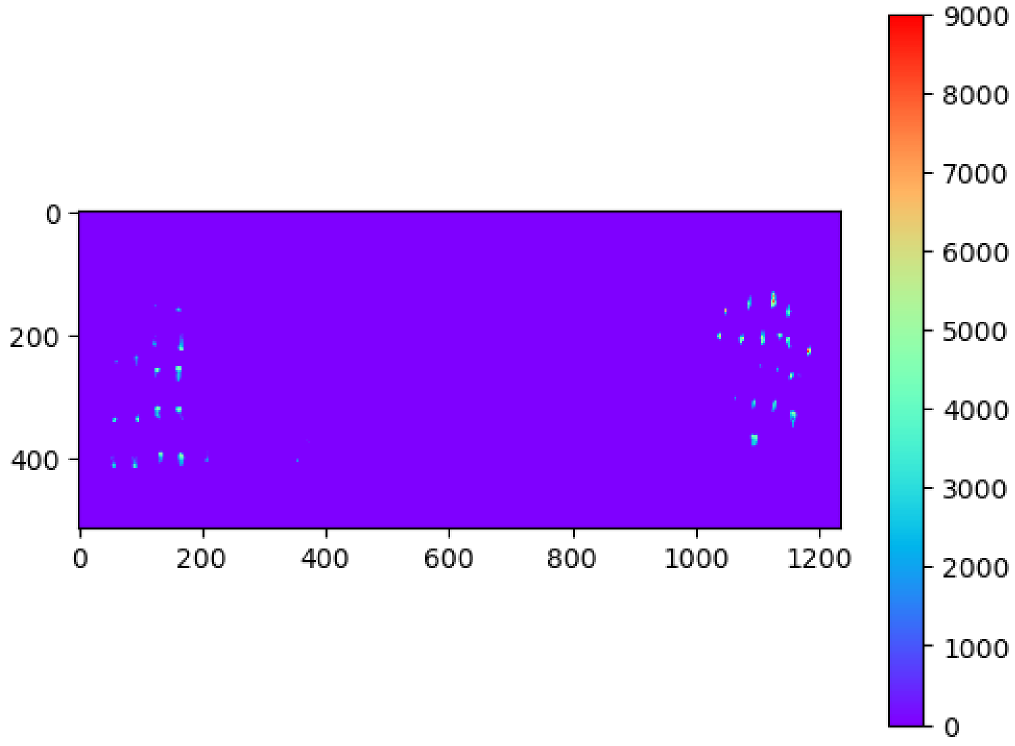

Figure 2. A scanned array of extracted human teeth.

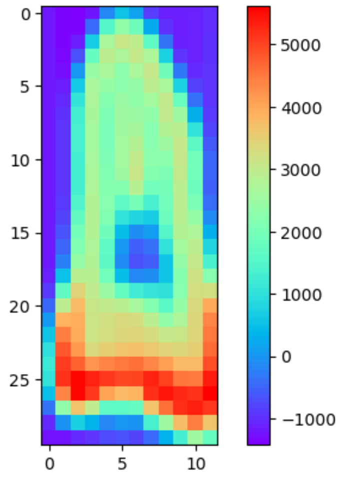

Figure 3. A vertical cross-section plot of an isolated tooth.

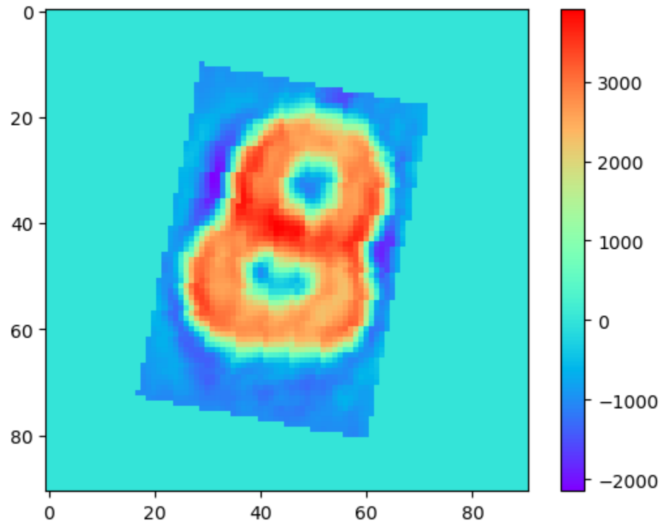

Figure 4. A horizontal cross-section plot if a padded and rotated tooth image.

Table 1

The labels and the correspond encoding.

Labels	**Class Numbers**
mandibular canine	0
mandibular incisor	1
mandibular molar	2
mandibular premolar	3
maxillary canine	4
maxillary central incisor	5
maxillary incisor	6
maxillary lateral incisor	7
maxillary molar	8
maxillary premolar	9

4.3 Data Analysis

For the classification task, the commonly used F1-score is used. F1-score has the advantage of balancing the precision and recall:

$$F_1 = \frac{2}{recall^{-1} + precision^{-1}} = \frac{2tp}{2tp + fp + fn}$$

$$recall = \frac{tp}{tp + fn}$$

$$precision = \frac{tp}{tp + fp}$$

where *tp* is the true positives, *fp* is the false positives, and *fn* is the false negatives. Since there are different type of teeth, it is a multi-classes classification problem, which has different considerations of F1-score calculation among different classes (scikit-learn, 2023). The micro F1-score calculate score using the global true positives, false positives, and false negatives. This calculation is the most common method used but it ignores the label the label imbalance. The macro F1-score, on the contrary, calculate the scores for each class first and then average the scores of all labels. This calculation can reflect the performance of each class, no matter how large or small the class sample sizes are. In our analysis, both micro and macro F1-scores are calculated.

5. RESOURCES

5.1 Hardware & Software

For the classification task, the hardware is as follow:

- CPU: Intel(R) Core (TM) i7-10750H CPU 2.60 GHz
- RAM: 32 GB
- GPU: NVIDIA GeFore RTX 2070 Super with Max-Q Design
- GPU Memory: 8 GB

The software is as follows:

- Operating system: Microsoft Windows 11 Home 64bit
- Annotation software: 3D Slicer
- IDE: Jupyter Notebook

- Python version: 3.9
- Libraries used: os, re, math, random, numpy, PIL, pandas, sklearn, nrrd, matplotlib, pytorch, torchvision, cv2, cc3d

For the segmentation task, the high-performance computing cluster Gadi will be used to run the program. On Gadi, there are GPU resources, and the operating system is Linux. The Python version is 3.9 with pytorch.

5.2 Roles & Responsibilities

All the literature review, model development, coding, data processing, segmentation method provision, data analysis, and report writing was done by Ka Wing Cheng.

The labelling and segmentation of X-ray CT teeth data was done by experts in the Sydney Dental School.

The development of the student learning website will be done by the front-end developer appointed by the Sydney Dental School.

6. MILESTONES / SCHEDULE

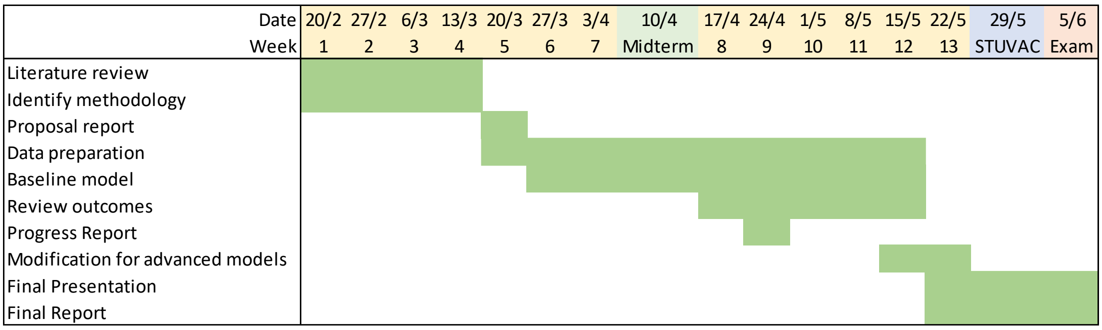

Figure 5. The Gantt chart for the project milestones.

Table 2

Details to measure progress and completion.

Task	Progress and completion measure	Status
Literature review	Identified relevant literatures	Completed in week 4
Identify methodology	Suggestion of model and data annotation strategy	Completed in week 4
Proposal report	Written report	Completed in week 5
Data preparation	Successful annotation of tooth structures	Completed in week 12

Baseline model	Successful building of model	Completed in week 12
Review outcomes	Calculation of performance scores, identify possible area of improvement	Completed in week 12
Progress Report	Written report	Completed in week 9
Modification for advanced models	Implementation of model modification, calculation of performance changes	Completed in week 13
Final Presentation	PowerPoint presentation	Completed on 1st June
Final Report	Written report	Completed on 11th June

7. RESULTS

After isolation of individual tooth from the CT scan, there were 360 useable teeth 3D images. After labelling, the data set was split into training, validation, and test sets and the distribution is shown in Table 3. The distribution of the data is plotted in Figure 6.

Table 3

Data distribution of the training, validation, and test set.

Label	Training	Validation	Test	Total
maxillary molar	62	32	32	126
mandibular molar	47	25	24	96
maxillary premolar	22	12	11	45
mandibular premolar	15	7	8	30
maxillary central incisor	8	4	4	16
mandibular canine	7	4	4	15
mandibular incisor	6	3	4	13
maxillary canine	3	2	2	7
maxillary lateral incisor	3	2	2	7
maxillary incisor	3	1	1	5

Table 4

Micro F1-scores of the trial run of different models.

	Training Data	Validation Data	Test Data
ResNet	0.863636	0.858696	0.684783
VGG16	0.710227	0.771739	0.684783
AlexNet	0.795455	0.782609	0.673913
3D AlexNet	0.693182	0.836957	0.619565
3D ResNet	0.943182	**0.902174**	0.728261

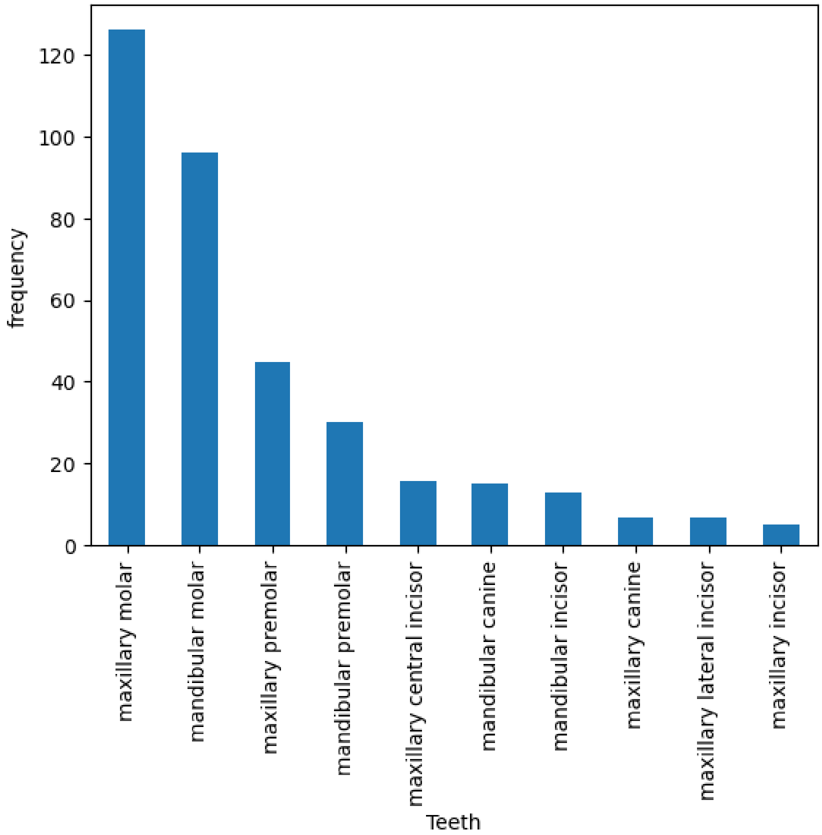

Figure 6. The distributed of all usable labelled data.

The training data set was input for the trial training of ResNet, VGG16, AlexNet, 3D Alex Net and 3D ResNet and validated with the validation set. Adam optimiser with a learning rate 10^{-4} as well as cross-entropy loss was used, and epochs were 30. The models were recorded with the best validation F1-scores. The micro and macro F1-scores are tabulated in Table 4 and Table 5.

Table 5

Macro F1-scores of the trial run of different models

	Training Data	**Validation Data**	**Test Data**
ResNet	0.799432	0.684686	0.574779
VGG16	0.630240	0.506027	0.486936
AlexNet	0.687771	0.645796	0.450561
3D AlexNet	0.545427	0.482550	0.383500
3D ResNet	0.896104	**0.766879**	0.575931

From the trial runs, 3D ResNet had the best validation F1-scores, so it was chosen for hyperparameters tuning with 30 epochs. The tuning of the learning rate is shown in Figure 7. From Figure 7, 10^{-4} was already the best values of the learning rate.

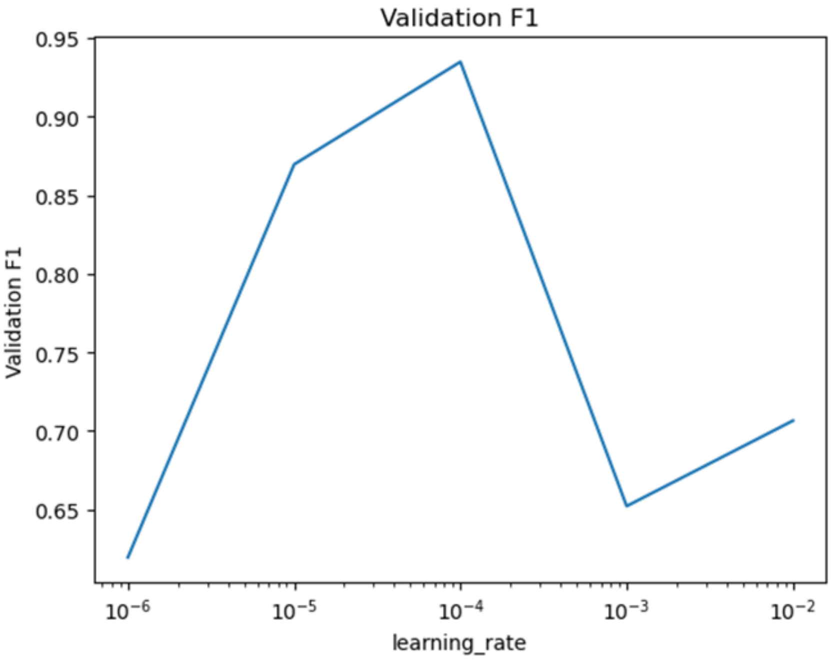

Figure 7. The hyper-parameter tuning of the learning rate of 3D ResNet.

After hyper-parameter tuning, the 3D ResNet was re-trained with both the training and validation data. Since there were no validation data in re-training, the stopping criterion was set as the same micro F1-score of the previous highest validation micro F1-score in the hyper-parameter tuning. The final model was tested with the test data set and the results are shown in Table 6 and the classification report is shown in Table 7. The confusion matrix on test data is shown in Figure 8.

Table 6

The results of the final 3D ResNet model on the test data set.

Micro F1-score	0.793478
Macro F1-score	0.602551
Time (per sample)	0.22 s

Table 7

Classification report of the final 3D ResNet model on the test data.

	precision	recall	f1-score	support
mandibular canine	0.5	0.75	0.6	4
mandibular incisor	0.6	0.75	0.67	4
mandibular molar	0.9	0.79	0.84	24
mandibular premolar	0.8	0.5	0.62	8
maxillary canine	0	0	0	2
maxillary central incisor	1	1	1	4
maxillary incisor	0.5	1	0.67	1
maxillary lateral incisor	0	0	0	2
maxillary molar	0.83	0.94	0.88	32
maxillary premolar	0.69	0.82	0.75	11
accuracy			0.79	92
macro avg	0.58	0.65	0.6	92
weighted avg	0.77	0.79	0.78	92

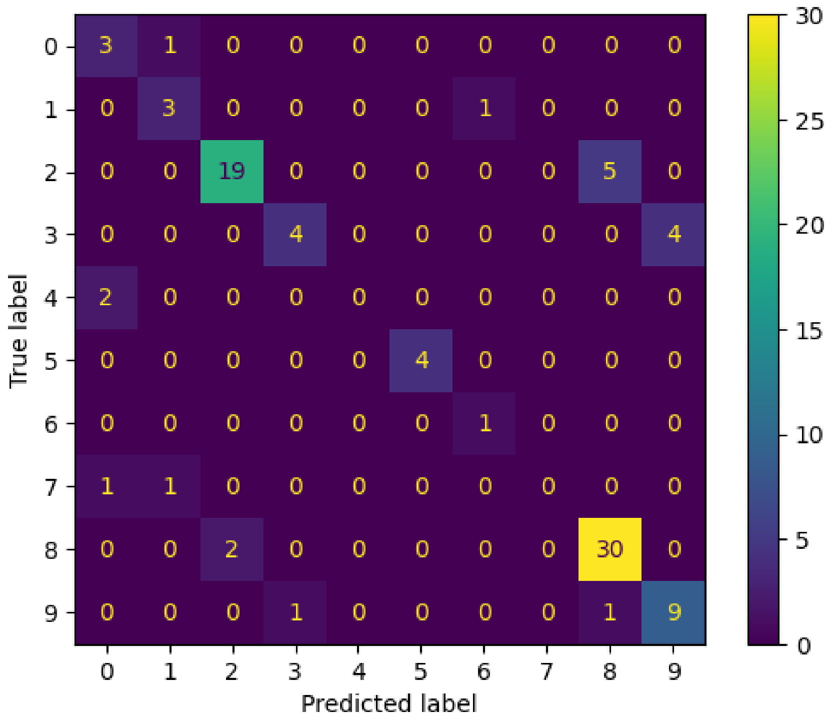

Figure 8. Confusion matrix of the final 3D ResNet model on the test data.

8. DISCUSSION

Comparing the results of the final model, 10-classes accuracy of 79% in Table 7 with the 7-classes accuracy of 91.0% of AlexNet (Yuma et al., 2017), there should still be room for improvement. From Table 7 and Figure 8, the inference of maxillary canine and maxillary lateral incisor were all zero. From Figure 6, one can observe that these two classes had the second and third least counts among all classes. On the contrary, the two classes having the best performance was the top two most populated classes. Therefore, one reason of the lower performance should be the low sample sizes of some classes. This caused a class imbalance problem in the training process, in which the training was dominated by the more populated classes. More data collection on the less populated classes is needed to overcome the problem. Before obtaining a balanced data set, strategies like weighted loss, focal loss (Lin, Goyal, Girshick, He, & Dollar, 2017) or weighted sampling with more data augmentation can be used.

Regarding the models, the performance of 3D ResNet was better than the 2D ResNet. Although the 3D ResNet was originally designed for video recognition, it seems that it can handle volumetric data well. On the contrary, using 2D 1×1 convolutional filter to adapt the 3D data for 2D model and extract the height information may compress the height information, although the method is fast and simple. For 3D AlexNet, its performance is less than its 2D version. However, from the F1-scores of the training and validation data, the lower validation performance suggests an under-training. A larger number of epochs than 30 is needed but it also means a longer training time is needed.

During the tooth isolation and labelling process, some images were found to have low resolution and were hard to separate and label. One example of low-resolution image is shown in Figure 9. Since artificial neural networks mimics the human way of capturing the features of an image or data, if the resolution of an image is too low for human recognition, it is also difficult for AI models to extract the features. Therefore, it is important to ensure the resolution of CT scans in the future data collection process. Nevertheless, if the high-resolution data collection is enough to get a high-performance model in the future, lower resolution data or even artificially burred images can be incorporated into the data with a view to extend the model's ability towards lower resolution images.

Figure 9. An example of low-resolution tooth CT image.

9. LIMITATIONS AND FUTURE WORKS

The major limitation of the project was the complexity of data collection stage. Since there are many features in the CT scans of teeth, the annotation process is labour intensive and time consuming. Therefore, in this report, the focus is put on the classification task to allow more time for the annotation for the segmentation task. Moreover, because of time limitation, only limited data augmentation and hyper-parameter tuning were done. As mentioned in the discussion, class imbalance and low-resolution data are also limitations of the works.

In the future, more high-resolution CT scans of teeth to achieve a more balance data set will be done. More data augmentation, hyper-parameters tuning can be done for a more sophisticated model. If the class imbalance of data cannot be solved in the short future, loss function for imbalance data, and weight sampling should be tried. For both classification and segmentation tasks, since the labelling or annotation works need a lot of labour and time, semi-supervised learning strategies, like cross-pseudo supervision (Y. Huang, Zhang, Yan, & Hassan, 2022), can be carried out.

REFERENCES

Chaurasia, A., & Culurciello, E. (2017). *LinkNet: Exploiting encoder representations for efficient semantic segmentation*.

Ching-Wei, W., Cheng-Ta, H., Jia-Hong, L., Chung-Hsing, L., Sheng-Wei, C., Ming-Jhih, S., . . . Claudia, L. (2016). A benchmark for comparison of dental radiography analysis algorithms. *Medical Image Analysis, 31*, 63-76. doi:https://doi.org/10.1016/j.media.2016.02.004

Eun, H., & Kim, C. (2016). *Oriented tooth localization for periapical dental X-ray images via convolutional neural network*.

Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.-C., Pujol, S., . . . Kikinis, R. (2012). 3D Slicer as an Image Computing Platform for the Quantitative Imaging Network. *Magnetic Resonance Imaging, 30*(9), 1323-1341. doi:10.1016/j.mri.2012.05.001

He, K., Zhang, X., Ren, S., & Sun, J. (2016). *Deep residual learning for image recognition.* Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). *Densely connected convolutional networks.* Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.

Huang, Y., Zhang, H., Yan, Y., & Hassan, H. (2022). *3D Cross-Pseudo Supervision (3D-CPS): A Semi-supervised nnU-Net Architecture for Abdominal Organ Segmentation*, Cham.

Isensee, F., Jäger, P. F., Full, P. M., Vollmuth, P., & Maier-Hein, K. H. (2021). *nnU-Net for Brain Tumor Segmentation*, Cham.

Jader, G., Fontineli, J., Ruiz, M., Abdalla, K., Pithon, M., & Oliveira, L. (2018). *Deep Instance Segmentation of Teeth in Panoramic X-Ray Images*.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). *ImageNet Classification with Deep Convolutional Neural Networks*. https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollar, P. (2017). *Focal Loss for Dense Object Detection*. Paper presented at the Proceedings of the IEEE International Conference on Computer Vision (ICCV).

Lone, M., McKenna, J. P., Cryan, J. F., Downer, E. J., & Toulouse, A. (2018). A Survey of tooth morphology teaching methods employed in the United Kingdom and Ireland. *Eur J Dent Educ, 22*(3), e438-e443. doi:10.1111/eje.12322

Luu, H. M., & Park, S.-H. (2022). *Extending nn-UNet for Brain Tumor Segmentation*, Cham.

Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., . . . Van Leemput, K. (2015). The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Trans Med Imaging, 34*(10), 1993-2024. doi:10.1109/TMI.2014.2377694

Ronneberger, O., Fischer, P., & Brox, T. (2015). *U-Net: Convolutional Networks for Biomedical Image Segmentation*, Cham.

Schneider, L., Arsiwala-Scheppach, L., Krois, J., Meyer-Lueckel, H., Bressem, K. K., Niehues, S. M., & Schwendicke, F. (2022). Benchmarking Deep Learning

Models for Tooth Structure Segmentation. *J Dent Res, 101*(11), 1343-1349. doi:10.1177/00220345221100169

Schwendicke, F., Golla, T., Dreher, M., & Krois, J. (2019). Convolutional neural networks for dental image diagnostics: A scoping review. *J Dent, 91*, 103226. doi:10.1016/j.jdent.2019.103226

scikit-learn, d. (2023). sklearn.metrics.f1_score — scikit-learn 1.2.2 documentation. Retrieved from https://scikit-learn.org/stable/modules/generated/sklearn.metrics.f1_score.html

Shen, D., Wu, G., & Suk, H. I. (2017). Deep Learning in Medical Image Analysis. *Annu Rev Biomed Eng, 19*, 221-248. doi:10.1146/annurev-bioeng-071516-044442

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Singh, S. P., Wang, L., Gupta, S., Goli, H., Padmanabhan, P., & Gulyas, B. (2020). 3D Deep Learning on Medical Images: A Review. *Sensors (Basel), 20*(18). doi:10.3390/s20185097

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., . . . Rabinovich, A. (2015). *Going Deeper With Convolutions*. Paper presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Tran, D., Wang, H., Torresani, L., Ray, J., LeCun, Y., & Paluri, M. (2018). *A Closer Look at Spatiotemporal Convolutions for Action Recognition*.

Wirtz, A., Mirashi, S. G., & Wesarg, S. (2018). *Automatic Teeth Segmentation in Panoramic X-Ray Images Using a Coupled Shape Model in Combination with a Neural Network*, Cham.

Yuma, M., Chisako, M., Tatsuro, H., Xiangrong, Z., Takeshi, H., Akitoshi, K., & Hiroshi, F. (2017). Classification of teeth in cone-beam CT using deep convolutional neural network. *Computers in Biology and Medicine, 80*, 24-29. doi:https://doi.org/10.1016/j.compbiomed.2016.11.003

Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., & Liang, J. (2018). *UNet++: A Nested U-Net Architecture for Medical Image Segmentation*, Cham.

Declaration

I, *(your name)*……Ka Wing CHENG…………………………………………………,

agree ☑

disagree ☐

(please tick one of the boxes)

that my thesis is made available as a sample thesis on the School of CS's eLearning platform for incoming students to view.

Signature: K.W. Cheng

Date: 10 June 2023