Progress in Microbes and Molecular Biology

HH PUBLISHER

Original Research Article

Incisor Malocclusion Using Cut-out Method and Convolutional Neural Network

Muhamad Farhin Harun[1], Azurah A Samah[1*], Muhammad Imran Ahmad Shabuli[1], Hairudin Abdul Majid[1], Haslina Hashim[1], Nor Azman Ismail[1], Syiral Mastura Abdullah[2], Aspalilah Alias[3,4,5]

Article History

Received: 17 May 2022;

Received in Revised Form: 25 September 2022;

Accepted: 4 October 2022;

Available Online: 6 October 2022

[1]Faculty of Computing, Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor, Malaysia; muhamadfarhin@gmail.com (MFH); imranaini96@gmail.com (MIAS); hairudin@utm.my (HAM); haslinah@utm.my (HH); azman@utm.my (NAI)

[2]Faculty of Dentistry, Universiti Sains Islam Malaysia, Jalan Pandan Utama, 55100 Kuala Lumpur, Malaysia; syiral@usim.edu.my (SMA)

[3]Department of Basic Sciences and Oral Biology, Faculty of Dentistry, Universiti Sains Islam Malaysia, Kuala Lumpur, Malaysia; aspalilah@usim.edu.my (AA)

[4]Forensic Odontology Unit, Department of Imaging & Pathology, KU Leuven, Belgium; aspalilah@usim.edu.my (AA)

[5]Department of Forensic Odontology, Faculty of Dental Medicine, Universitas Airlangga, Indonesia; aspalilah@usim.edu.my (AA)

*Corresponding author: Azurah A Samah; Faculty of Computing, Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor, Malaysia; azurah@utm.my (AAS)

Abstract: Malocclusion is a condition of misaligned teeth or irregular occlusion of the upper and lower jaws. This condition leads to poor performance of vital functions such as chewing. A common procedure in orthodontic treatment for malocclusion is a conventional diagnostic procedure where a dental health professional takes dental x-rays to examine the teeth to diagnose malocclusion. However, the manual orthodontic diagnostic procedure by dental experts to identify malocclusion is time-consuming and vulnerable to expert bias that results in delayed treatment completion time. Recently, artificial intelligence technology in image processing has gained attention in orthodontics treatment, accelerating the diagnosis and treatment process. However, several issues concerning the dental images as input of the classification model may affect the accuracy of the classification. In addition, unstructured images with varying sizes and the problem of a machine learning algorithm that does not focus on the region of interest (ROI) for incisor features bring challenges in delivering the treatment. This study has developed a malocclusion classification model using the cut-out method and Convolutional Neural Network (CNN). The cut-out method restructures the input images by standardising the sizes and highlighting the incisor sections of the images which assisted the CNN in accurately classifying the malocclusion. From the results, the implementation of the cut-out method generates higher accuracy across all classes of malocclusion compared to the non-implementation of the cut-out method.

Keywords: Malocclusion; orthodontics; incisor; classification; convolutional neural network; class activation mapping, cut-out method; image processing

1. Introduction

Malocclusion is a common dental problem. It is a dentoskeletal condition that can impact patients' quality of life and social interactions by affecting both function and aesthetics [1]. Malocclusion may aggravate or even create health problems such as eating disorders, migraines, speech difficulties and sleep disorders [2]. The genesis of malocclusion is complicated by several genetic, environmental and local factors and it can be caused by a skeletal or dental-related discrepancy [3]. Various types and symptoms give difficulties to dental solutions to overcome the problems. The conventional manual classification procedure by the dentist consumes time and requires more evaluation processes. At the same time, most studies have not confirmed the anticipation of which orthodontic treatments that applies modern technology will minimise treatment time [4–8]. There are three main classes of malocclusion, which are Class I, Class II and Class III malocclusion [9]. The classes are used to classify the condition of teeth and to give a clear view of the malocclusion problem to the dentist. Manual analysis by the dentist and orthodontic experts is currently the most common technique for the majority of orthodontic problems.

The manual malocclusion classification procedure is time-consuming due to the need for a thorough manual evaluation by an expert. Some dental images are challenging to classify, costly to collect and revolve around various legal problems for patient privacy [10]. The manual classification procedure may consume weeks to months. This lengthy and laborious diagnostic process contributed to a long waiting time for patients waiting for malocclusion treatment. The computational deep learning-based model proposed in this study is able to speed up the classification process with better decision consistency at a reduced cost. This will lead to improved quality of services in terms of reduced cost of treatment and waiting time for patients waiting for malocclusion treatment at various dental clinics.

This paper is organised as follows. Section II presents the literature review on categories of malocclusion, Convolutional Neural Network (CNN), previous studies on the classification of malocclusion and Class Activation Mapping (CAM). This is followed by Section III on the methodology of the study. Briefly, Section III highlights the main activities in conducting the study, including CNN development, dataset, experimental design, performance measure and data analysis. In Section IV, the experimental results are presented and discussed. This paper continues in Section V on results and discussion. The paper ends with a conclusion that is included in Section VI.

2. Literature Review

2.1. Categories of Malocclusion

As described by World Health Organization (WHO), malocclusion is a handicapping dentofacial anomaly, which refers to irregular occlusion or abnormal craniofacial relationships which can affect aesthetic appearance, function, facial harmony and psychosocial well-being [6]. Malocclusion is the third most popular oral pathologic disorder after tooth decay and periodontal disease that requires special treatment. Malocclusions such as deep overbite, midline deviation, extreme overjet, anterior cross bite, mal-alignment, and open bite are common in clinics. The Angle Malocclusion classification based on the relative position of the permanent maxillary first molar is commonly used for assessing malocclusion [11]. The four front teeth in the upper and lower jaws are called incisors. The central incisors are the two incisors on either side of the midline, while the lateral incisors are the teeth next to the central incisors [12]. The condition of the problem based on the British Standard Institute (BSI) also categorizes the classes into Class I, Class II/1, Class II/2 and Class III, as shown in Table 1 and Figure 1.

Table 1. Incisor Classification Description [13].

Class	Description
Class I	Lower incisors' incisal edge bites on or below upper incisors' cingulum plateau
Class II/1	Upper incisors are upright or proclined and lower incisors bite behind the upper incisors' cingulum plateau
Class II/2	Lower incisors bite behind the cingulum plateau and the higher incisors are retroclined
Class III	Maxillary incisors' cingulum plateau is anterior to mandibular incisors' cingulum plateau

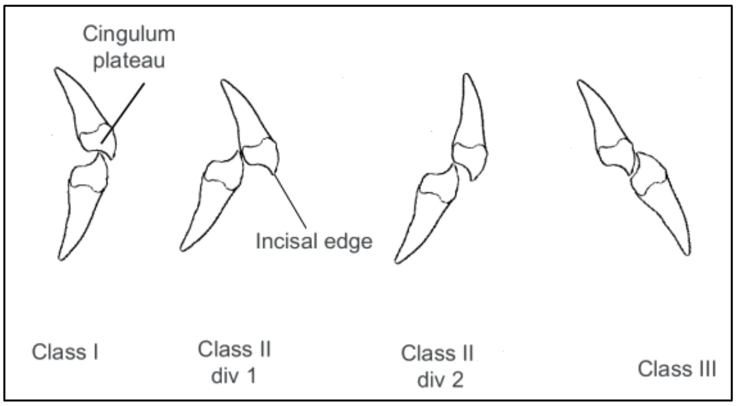

Figure 1. Incisor Classification.

Other types of teeth, such as canines, molars and premolars, have their own malocclusion classification. All incisor classification patterns will be the same for three malocclusion classes, and Class II/1 and II/2 will be combined to form Class II malocclusion. It should be emphasised that the incisor section will decide the outcome of the execution processes, which will be classified into one of three malocclusion classifications based on the image input.

2.2. Convolutional Neural Network (CNN)

CNNs have recently been used in medical trials and the findings showed that this approach is one of the potential computer-aided methods for medical practitioners [14]. This medical imaging study works to design orthodontic procedures based on CNN all imaging modalities used in orthodontics. As shown in Figure 2, the CNN cycle will conclude the cephalometric x-ray implemented together to be evaluated using CNN. Along with the growing demands for dental healthcare, a recent advance in the image recognition technology using the CNN model brings drastic improvements in diagnostic imaging and is eagerly awaited in the field of orthodontics as well [15].

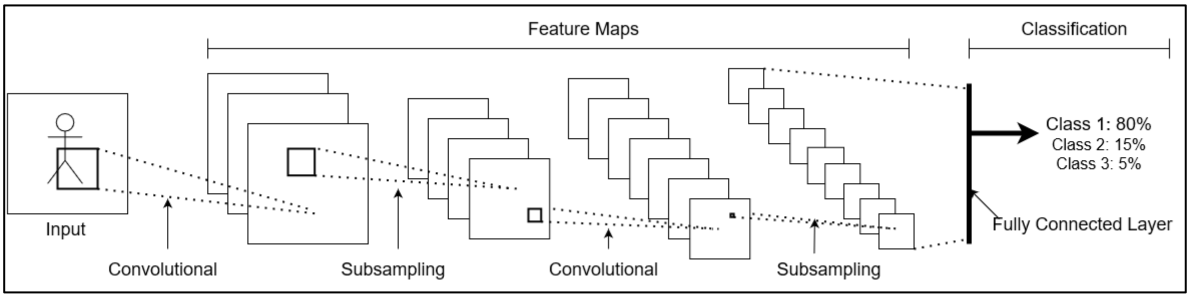

Figure 2. Convolutional Neural Network (CNN).

When there is a paucity of training data for image classification, the learnt model may be fooled by irrelevant local information, which might be considered noise. Furthermore, the overall structure of a CNN model, rather than the particular local pixels, has a significant impact on its performance in most classification tasks and if pixels within local structures change in a way that does not impair the overall perspective of a picture, the model should be able to accurately recognise the input image [16].

Overfitting is a serious difficulty in training robust CNN models and regularisation is a method to prevent CNN from overfitting [17]. Dropout is a regularisation method where square areas of an image is obscured. As visualised in Figure 3, the technique is important to help present a method for efficiently mixing a large number of distinct neural network topologies.

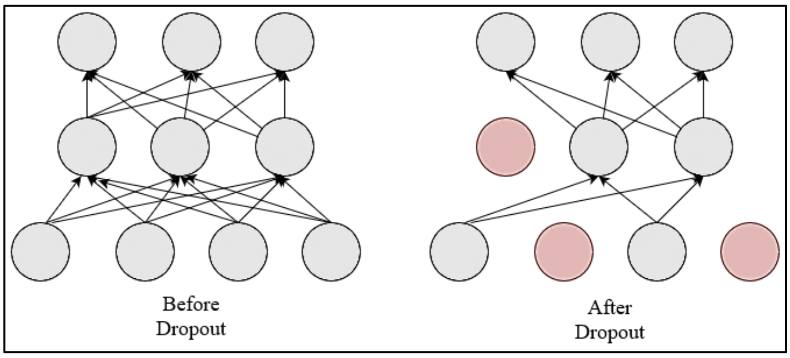

Figure 3. Dropout Method Implementation.

During training, hidden unit activations are set to zero with a pre-set probability, resulting in dropout and when evaluating the network, all activations are maintained. However, the output is scaled according to the dropout probability [18]. This strategy approximates averaging over an exponential number of smaller sub-networks and it performs well as a robust type of bagging that prevents feature detectors from co-adapting within the network [19].

2.3. Previous Studies on Classification of Malocclusion

Despite a few attempts to develop more understanding of malocclusions, the orthodontic discipline is still mostly confined to Edward H. Angle's malocclusion classification, which was developed nearly 30 years before cephalometry was introduced [13,20].The studying of the lateral cephalogram image features is a conventional method for malocclusion class detection that is still used today. As time passes, researchers have employed a variety of techniques to improve the malocclusion classification procedure [21]. The complete structure of the teeth can be examined through the cephalogram images as mentioned earlier. For the development of the study, results of classification based on incisor focus are extremely limited.

2.4. Class Activation Mapping (CAM)

To sort different classes, the CAM method has been used to highlight the class-specific areas that the network learned from the data [22]. The CAM method can also be combined with fine-grained visualisation techniques to produce high-resolution class-discriminative representations. This applies to a wide range of CNN model families without any architectural modifications or re-training. However, their effects, are often accompanied by random noises that are unrelated to the image's target item and the weight does not adequately capture the value of each activation map [23]. The class-specific area highlighted using the CAM method can assist the understanding of how the CNN operates on classification and what features the CNN had extracted.

3. Materials and Methods

3.1. Materials

This retrospective study uses existing data recorded by the Faculty of Dentistry, Universiti Sains Islam Malaysia (USIM), which does not require any written informed consent from the patient. The patient id was replaced with a unique id and no personal details were exposed during the execution of this study. This study has been approved by the Research and Ethics committee of USIM. The Ethics approval code for this study is USIM/JKEP/2020-88

3.2. Methodology

There are four main activities in conducting the study, as shown in Figure 4. The four main activities are Problem Articulation, Data Collection, Experimental Design, Performance Measure and Analysis.

Figure 4. Four Phases of Research Methodology.

The problem articulation phase starts with understanding malocclusion, CNN modelling solution and techniques used in the malocclusion classification study. In total, there were 454 images, including 166 incisor images of Class I, 254 incisor images of Class II and 125 incisor images of Class III used in this study. The next phase is data collection, where lateral cephalogram images of the entire facial skeleton used in this study were gathered as input to measure the classification performance of CNN for analysis.

There were several tasks in the experimental design phase, which were data pre-processing (Table 2), development of the CNN model, hyperparameter tuning, implementation of the cut-out method and CAM (Refer to Figure 6). Each image input is fixed at a size of 128x128 pixels. The objective of this phase is to design the CNN model to identify patterns by analysing the structure of each image in each class. The pattern of the incisor is identified as a significant feature based on the single input image. To produce good ROI, the cut-out method is implemented to improve the performance of CNN. Thus, the size of the image data was adjusted to enhance the malocclusion classification. Then, CAM highlights the extracted features from CNN in terms of visualisation.

Table 2. Data Augmentation Parameters.

Parameter	Value/Type
Zoom Range	0.5
Shear Range	0.2
Height Shift Range	0.2
Width Shift Range	0.2
Fill Mode	Nearest

Lastly, the performance measure and analysis were conducted. A confusion matrix which consisted of four elements, as shown in Figure 5 was generated to aid in the analysis of the results. There are four elements in the confusion matrix which are True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN).

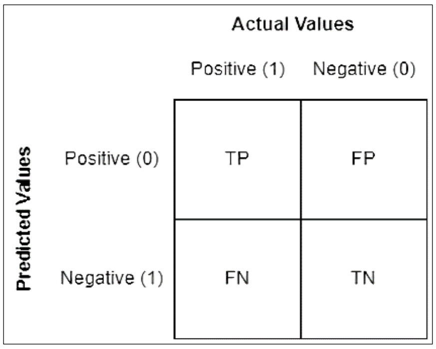

Figure 5. Confusion Matrix.

The accuracy, precision, recall and F1 score are measures derived from the confusion matrix. Accuracy is a measure of how many correct predictions the proposed model made for the complete test dataset. Precision is a measure of the correctness of a positive prediction. On the other hand, recall is a measure to determine how many true positives get predicted out of all positives in the dataset. Finally, the F1 score is a statistic that takes into account both precision and recall. Equation 1 to 4 shows the formulation of performance measures used in this study.

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

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

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

$$F_1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} = \frac{2TP}{2TP+FP+FN} \quad (4)$$

4. Experimental Results

4.1. Experimental Setup

The flow of the experimental setup for this study is as shown in Figure 6. The layers were implemented in a sequence to establish a series of evaluations toward input during training. The CNN layers of the study are as specified in Table 3.

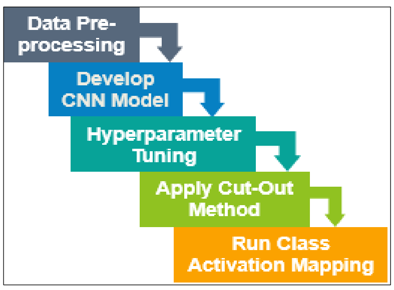

Figure 6. The flow of experimental setups.

Table 3. CNN layers architecture.

Layer	Specification
Conv2D	Input Size = 128×128×3
3×3 convolution, 16 filters	
Conv2D	3×3 convolution, 16 filters
Batch Normalization	-
Activation	ReLu
MaxPooling2D	2×2 max-pooling
Conv2D	3×3 convolution, 32 filters
Conv2D	3×3 convolution, 32 filters
Batch Normalization	-
Activation	ReLu
MaxPooling2D	2×2 max-pooling
Conv2D	3×3 convolution, 64 filters
Conv2D	3×3 convolution, 64 filters
Batch Normalization	-
Activation	ReLu
MaxPooling2D	2×2 max-pooling
Conv2D	3×3 convolution, 128 filters
Conv2D	3×3 convolution, 128 filters
Batch Normalization	-
Activation	ReLu
MaxPooling2D	2×2 max-pooling
Flatten	-
Dropout	50% dropout
Activation	ReLu
Dense	3 filters
Softmax	softmax

4.2. Hyperparameters

The hyperparameter settings for the CNN model were configured to deliver better outcomes and maximise the capabilities of the training process. The parameters are as shown in Table 4.

Table 4. Hyperparameters and training options for CNN.

Hyperparameters and Training options	Assigned Values
Training optimisation function	Adam
Initial Learn Rate	1e-07
Maximum epoch	200
Training:Validation:Test ratio	7:2:1
Learning Rate	0.001
Number of Iterations	200

4.3. Cut-out Method

The input data was in the form of single images. Most of the images were unstructured images where the size of images varied from one image to another. Additionally, the incisor section for each image was relatively small and confined. The cut-out method was used to restructure the image such that it focuses on the incisor section and provided a clear view for model development. The size of the image input was initially different for all images included in the study. Therefore, size adjustments were required. The adjustments involve changing the pixel size of an image manually for horizontal and vertical scaling using a resize function attained from the Opencv library to scale the image and create a cut-out effect. The newly restructured image which includes the incisor part is displayed and scaled as shown in Figure 7.

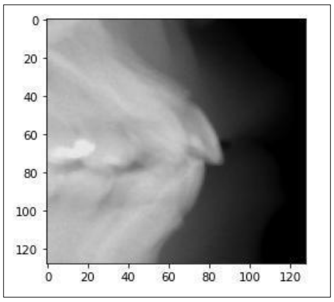

Figure 7. Cut-out technique from a selected input image.

The cut-out method was further validated during the post-processing phase using CAM heatmap which is further discussed in the next section (B. CAM Heatmap Validation and Analysis).

5. Results and Discussion

In this section, a confusion matrix for classification, accuracy and loss from the model is presented. Additionally, a heatmap from CAM is included and the validation of the cut-out method is discussed to give additional perspective to the evaluation. There are others malocclusion classification studies using CNN on different malocclusion datasets. As our cut-out method is a very niche improvement over the private dataset, it was very difficult to conduct a one-to-one comparison with a published dataset. Instead, we measured the effectiveness of our experimental framework with other malocclusion studies.

5.1. Confusion Matrix

The performance measures of incisor classification were analysed using the confusion matrix. The result of the analysis is shown in Figure 8.

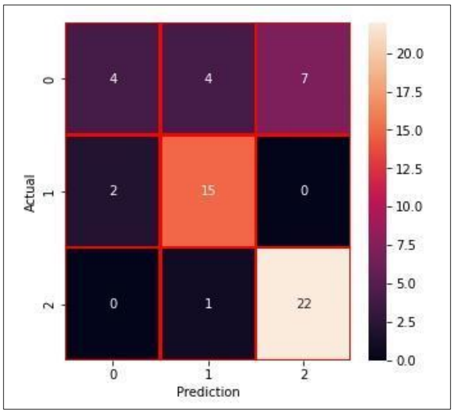

Figure 8. Confusion matrix of classification processes.

For the true and false values, nine-sectioned blocks were set up to represent each class. Details of the performance matrices at a 75% accuracy and 0.64 loss rates are shown in Table 5.

Table 5. Accuracy, Precision, Recall and F1-Score.

	Precision	Recall	F1-Value	Support
Class I	0.67	0.27	0.38	15
Class II	0.75	0.88	0.81	17
Class III	0.76	0.96	0.85	23

The recall value of 0.96 for Class III was the highest among the other classes. A high recall value indicates low misclassification of the predicted class. Features for Class I and Class II were quite similar to each other, which was why their recall values were lower than Class III. The recall value for Class I was the lowest amongst the classes presented, which indicates that the model cannot correctly identify the features for Class I. From the results obtained, it was seen that many of the Class I samples were misclassified as either Class II or Class III.

Class III achieved the highest F1value based on its corresponding precision and recall values. The performance of incisor classification for malocclusion was highly satisfied for Class III. In contrast, Class I achieved the lowest F1 value as both of its corresponding precision and recall values were low.

5.2. CAM Heatmap Validation and Analysis

The CAM heatmap was used to validate the CNN results. The validation was done using single image classification. In particular, CAM heatmap was used to detect the focal area of the incisor part for a single image. A sample of a single image was chosen to demonstrate the validation process. Figure 9 and Figure 10 show the heatmap output of the same image. In Figure 9, the heatmap output was generated without implementing the cut-out method. On the other hand, Figure 10 shows a heatmap output that was generated after a cut-out method has been implemented.

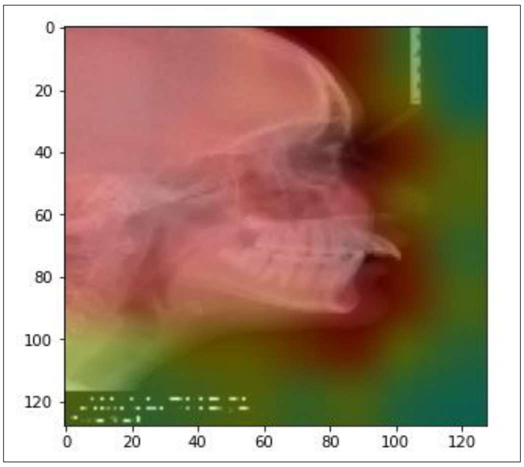

Figure 9. Heatmap output without cut-out method.

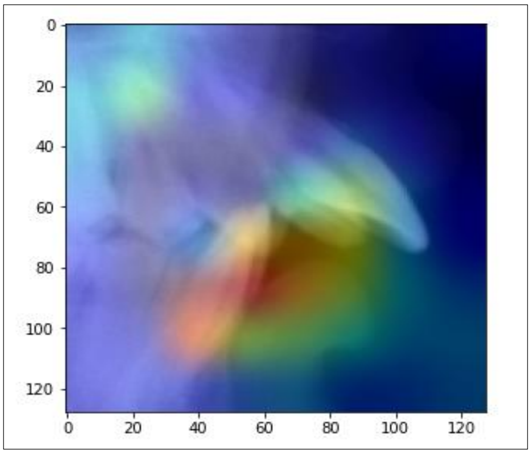

Figure 10. Heatmap output with cut-out method.

Both Figure 9 and 10 revealed significant differences between each other. As shown in Figure 9, without the cut-out method, the focus section was spread to regions that were irrelevant to the study's observation. Furthermore, the focus was not on a precise section, which has led to unfavourable outcomes. On the other hand, with the implementation of the cut-out method, Figure 10 shows that the focus was specific on the incisor areas. Thus, the results were more accurate and convincing, with the implementation of the cut-out method. It was also evident that the implementation of cut-out method and CAM validation has greatly assisted the highlighting and classifying of the incisor areas. Further validation analysis of the cut-out method using the heatmap was done on samples of each class. The results of the validation analysis are shown in Table 6.

As shown in Table 6, a sample from each class was validated with and without the cut-out method. For all classes, the implementation of the cut-out method resulted in the highest accuracy. For instance, a sample of Class II with ID 317 implemented a cut-out method yielded the highest accuracy of 0.827 of being classified as Class II. Similarly, the sample from Class I (sample ID of 46) and Class III (sample ID of 456) yielded the highest accuracy of 0.700 and 0.598, respectively of being classified as Class I and III.

Table 6. Heatmap Result Analysis.

Sample		Confident level		
		Class 1	**Class 2**	**Class 3**
Sample ID: 317	Without Cut-Out	0.330	0.332	0.336
[Class 2]	With Cut-Out	0.050	**0.827**	0.121
Sample ID: 46	Without Cut-Out	0.335	0.330	0.333
[Class 1]	With Cut-Out	**0.700**	0.172	0.127
Sample ID: 456	Without Cut-Out	0.331	0.333	0.334
[Class 3]	With Cut-Out	0.148	0.253	**0.598**

Both validation on a single sample and the three samples that represented each class demonstrated that the implementation of the cut-out method and CNN model have assisted in the identification and marking of specific incisor areas for specific samples, which further assisted in the classification of malocclusion. Satisfying results with the implementation of the computational deep learning-based model were obtained not only in this study but also in other malocclusion studies conducted by other researchers.

Table 7 summarised two malocclusion studies that have implemented deep learning methods for malocclusion classification, which are closely similar to our study. Results obtained by the research listed in Table 7 are more promising than ours and remarks in column four of Table 7 stated the justification of the higher results that were obtained in comparison to our study. These remarks have open room for further improvement to our study, to attain better results in the future.

Both malocclusion studies listed in Table 7 used different data types. Thus, different data pre-processing was implemented. This brings the importance of CNN modelling in better understanding and focusing on the features. In our study, the focus was more on the cut-out method to highlight the incisor feature. Although we have carried out several data augmentations, the focus should be on improving the image quality as it could help highlight the feature much better.

Both studies listed in Table 7 used customised CNN deep learning models to work better with specific samples. The results of our study may be further improved by integrating the cut-out method with a customised CNN model. As the current study is a stepping stone toward understanding more about the behaviour of CNN when classifying the incisor dataset, enhancement of the image quality and customisation of the CNN deep learning model are improvements that can be included in our future work. The proposed improvements will be deeply analysed as they could open more research opportunities in the future.

Table 7. Malocclusion Studies.

Malocclusion study	Dataset type	Overall Result	Remarks
Molar-incisor-hypomineralisation (MIH) [24].	Intraoral photographs. (Normal image)	The result has an overall accuracy of 95.5%.	1. The samples used are clear and high quality. 2. The feature in the samples is clearly defined. 3. Deep CNN architecture was used which improved the efficiency at the cost of longer training time.
incisors, canines, premolars, and molars [25].	Grey-level image	The best result obtained of accuracy is 87%.	1. Rigorous augmentation on a limited sample of the dataset. 2. Customised pooling architecture.

6. Conclusions

In this study, the main objective related to the development of an improved classification model based on incisor data using the cut-out method and CNN modelling to facilitate the classification of malocclusion has been achieved with satisfying results. Input images that have been restructured with a focus on the incisor section using the cut-out method produced higher accuracy and lower loss in classifying malocclusion as compared to images that have not been implemented with the cut-out method. Despite this achievement, few works can be carried out in the future to expand the research further and improve existing results. One of improvements is to automate the cut-out method which currently was implemented manually in this study. The implementation of the cut-out method manually may cause inconsistent sizes of incisor region in some of the samples/images. An automated and adaptive cut-out method is therefore suggested to generate better consistency in the sizes of the incisor region for every sample, which may produce more accurate focal region detection.

Author Contributions: Conceptualization, methodology, literature search, MIAS, MFH, AA; validation, NAI, MIAS, AA; data collection, MFH, SMA, AA; writing—original draft preparation, MIAS, HAM; reviewed, editing, proofread, AAS, HH.

Funding: This work was funded by the Malaysian Ministry of Higher Education under Fundamental Research Grant Scheme (Grant Number: FRGS/1/2021/ICT02/UTM/02/5).

Acknowledgments: The authors would like to express gratitude to the Malaysian Ministry of Higher Education for the financial sponsorship of this study through the Fundamental Research Grant Scheme

(FRGS/1/2021/ICT02/UTM/02/5). The study is also supported by the Faculty of Computing, Universiti Teknologi Malaysia and Faculty of Dentistry, Universiti Sains Islam Malaysia.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Koskela A, Neittaanmäki A, Rönnberg K, Palotie A, Ripatti S, Palotie T. The relation of severe malocclusion to patients' mental and behavioral disorders, growth, and speech problems. *Eur J Orthod*. 2021;43(2):159-164. doi:10.1093/ejo/cjaa028

2. Leavy KM, Cisneros GJ, LeBlanc EM. Malocclusion and its relationship to speech sound production: Redefining the effect of malocclusal traits on sound production. *Am J Orthod Dentofac Orthop*. 2016;150(1):116-123. doi:10.1016/j.ajodo.2015.12.015

3. Al-Qurashi, H., Al-Farea, M., Alshamrani, H., Almasoud, N. N., & Nazir MA. Orthodontic treatment needs and association between malocclusion and oral hygiene behaviors. *Pakistan Oral Dent J*. 2018;38(1):62-66. Accessed September 20, 2022. https://podj.com.pk/index.php/podj/article/view/137

4. Jung MH. Factors influencing treatment efficiency: A prospective cohort study. *Angle Orthod*. 2021;91(1):1-8. doi:10.2319/050220-379.1

5. Mavreas D, Athanasiou AE. Factors affecting the duration of orthodontic treatment: A systematic review. *Eur J Orthod*. 2008;30(4):386-395. doi:10.1093/ejo/cjn018

6. Zou J, Meng M, Law CS, Rao Y, Zhou X. Common dental diseases in children and malocclusion. *Int J Oral Sci*. 2018;10(1). doi:10.1038/s41368-018-0012-3

7. Pinto AS, Alves LS, Maltz M, Susin C, Zenkner JEA. Does the duration of fixed orthodontic treatment affect caries activity among adolescents and young adults? *Caries Res*. 2018;52(6):463-467. doi:10.1159/000488209

8. Moresca R. Orthodontic treatment time: can it be shortened? *Dental Press J Orthod*. 2018;23(6):90-105. doi:10.1590/2177-6709.23.6.090-105.SAR

9. Hengsheng L, Jun Y. Management of Malocclusion In Children and Adolescents. Decisions in Dentistry. Published 2020. Accessed September 20, 2022. https://decisionsindentistry.com/article/management-of-malocclusion-in-children-and-adolescents/

10. Guibas JT, Virdi TS, Li PS. Synthetic Medical Images from Dual Generative Adversarial Networks. *31st Conf Neural Inf Process Syst (NIPS 2017)*. 2017;(31):1-9. http://arxiv.org/abs/1709.01872

11. Alkhadra T. Characteristic of malocclusion among Saudi special need group children. *J Contemp Dent Pract*. 2017;18(10):959-963. doi:10.5005/jp-journals-10024-2156

12. Grawish ME, Grawish LM, Grawish HM. Permanent Maxillary and Mandibular Incisors. In: *Dental*

Anatomy. IntechOpen; 2018. doi:10.5772/intechopen.69542

13. Mageet AO. Classification of Skeletal and Dental Malocclusion: Revisited. *Stomatol EDU J*. 2016;3(3-4):205-211. doi:10.25241/stomaeduj.2016.3(3-4).art.11
14. Shin HC, Roth HR, Gao M, et al. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Trans Med Imaging*. 2016;35(5):1285-1298. doi:10.1109/TMI.2016.2528162
15. Lee C, Tanikawa C, Lim JY, Yamashiro T. Deep Learning based Cephalometric Landmark Identification using Landmark-dependent Multi-scale Patches. *arXiv*. Published online June 7, 2019. doi:10.48550/arxiv.1906.02961
16. Kang G, Dong X, Zheng L, Yang Y. PatchShuffle Regularization. *arXiv*. Published online July 22, 2017. doi:10.48550/arxiv.1707.07103
17. Xie L, Wang J, Wei Z, Wang M, Tian Q. DisturbLabel: Regularizing CNN on the loss layer. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Vol 2016-Decem. IEEE Computer Society; 2016:4753-4762. doi:10.1109/CVPR.2016.514
18. DeVries T, Taylor GW. Improved regularization of convolutional neural networks with cutout. *arXiv*. Published online 2017. Accessed November 12, 2020. https://github.com/
19. Cai S, Shu Y, Chen G, Ooi BC, Wang W, Zhang M. Effective and Efficient Dropout for Deep Convolutional Neural Networks. *arXiv*. Published online April 6, 2019. doi:10.48550/arxiv.1904.03392
20. Angle E. Classification of Malocclusion. *Dent Cosm*. 1899;41(3):246-264. Accessed November 20, 2020. http://name.umdl.umich.edu/acf8385.0041.001
21. Arvind T V, Baskaranarayanan B, Chidambaram S, Premkumar S. Association between the British Standards Institute's Incisor Classification of Malocclusion and Angle's Classification of Malocclusion-An Analytical Study. *RGUHS J Med Sci*. 2015;5(2):62-66. Accessed November 22, 2020. http://rjms.in/index.php/rjms/article/view/116882/80943
22. Shi J, Chen D, Wang M. Pre-impact fall detection with cnn-based class activation mapping method. *Sensors (Switzerland)*. 2020;20(17):1-12. doi:10.3390/s20174750
23. Wang H, Du M, Yang F, Zhang Z. Score-CAM:Improved Visual Explanations Via Score-Weighted Class Activation Mapping. *arxiv*. Published online 2019. Accessed September 20, 2022. http://arxiv.org/abs/1910.01279
24. Schönewolf J, Meyer O, Engels P, et al. Artificial intelligence-based diagnostics of molar-incisor-hypomineralization (MIH) on intraoral photographs. *Clin Oral Investig*. 2022;1:3. doi:10.1007/s00784-

022-04552-4

25. Li Z, Wang SH, Fan RR, Cao G, Zhang YD, Guo T. Teeth category classification via seven-layer deep convolutional neural network with max pooling and global average pooling. *Int J Imaging Syst Technol*. 2019;29(4):577-583. doi:10.1002/ima.22337