Hybrid HCNN-KNN Transfer Learning Model Enhances Age Estimation Accuracy in Orthopantomography

Fatemeh Sharifonnasabi [1], NZ Jhanjhi*[1], Jacob John [2], Peyman Obeidy [3], Shahab S Band* [4], Hamid Alinejad-Rokny* [5,6,7]

[1,2] School of Computing & IT (SoCIT), Taylor's University Subang Jaya, Malaysia

[2] Department of Restorative Dentistry, Faculty of Dentistry, University of Malaya, Kuala Lumpur, Malaysia

[3] Charles Perkins Centre, Faculty of Medicine and Health, University of Sydney, AUSTRALIA

[4] Future Technology Research Centre, College of Future, National Yunlin University of Science and Technology, 123 University Road, Yunlin 64002, Taiwan

[5] BioMedical Machine Learning Lab (BML), The Graduate School of Biomedical Engineering, UNSW SYDNEY, Sydney, 2052, AUSTRALIA

[6] UNSW Data Science Hub, The University of New South Wales (UNSW, SYDNEY), Sydney, 2052, AUSTRALIA

[7] Health Data Analytics Program, AI-enabled Processes (AIP) Research Centre, Macquarie University, Sydney, 2109, AUSTRALIA

Corresponding authors: shamshirbands@yuntech.edu.tw, h.alinejad@unsw.edu.au

Abstract: Age estimation in dental radiographs Orthopantomography (OPG) is a medical imaging technique that physicians and pathologists utilise for disease identification and legal matters. For example, for estimating post-mortem interval, detecting child abuse, drug trafficking, and identifying an unknown body. Recent development in automated image processing models improved the age estimation's limited precision to an approximate range of +/- one year. While this estimation is often accepted as accurate measurement, age estimation should be as precise as possible in most serious matters, such as homicide. Current age estimation techniques are highly dependent on manual and time-consuming image processing. Age estimation is often a time-sensitive matter in which the image processing time is vital. Recent development in Machine learning-based data processing methods has decreased the imaging time processing; however, the accuracy of these techniques remains to be further improved. We proposed an ensemble method of image classifiers and transfer learning techniques to enhance the accuracy of age estimation using OPGs from one year to a couple of months (1-3-6). This hybrid model is based on convolutional neural networks (CNN) and K nearest neighbours (KNN). The hybrid (HCNN-KNN) model was used to investigate 1,922 panoramic dental radiographs of patients aged 15 to 23. These OPGs were obtained from the various teaching institutes and private dental clinics in Malaysia. To minimise the chance of overfitting in our model, we used the principal component analysis (PCA) algorithm and eliminated the features with high correlation. To further enhance the performance of our hybrid model, we performed systematic image pre-processing. We applied a series of classifications to train our model. We have successfully demonstrated that combining these innovative approaches has improved the classification and segmentation and thus the age-estimation outcome of the model. Our findings suggest that our innovative model, for the first time, to the best of our knowledge, successfully estimated the age in classified studies of one year old, six months, three months and one-month-old cases with accuracies of 99.98, 99.96, 99.87, and 98.78 respectively.

Keywords

Dental age, estimation, Orthopantomogram, Convolutional neural network, K-nearest neighbour, Biomedical machine learning.

Introduction

Age estimation in living and deceased individuals has always been important in paediatric studies, pathological complications, and forensic medicine. Human bone age can be estimated from foot, shoulder, ankle, hip, elbow, cervix, ankle, and teeth bones [1-3]. The accuracy of the calculated age from a bone specimen depends on genetic and environmental factors such as age, race, smoking, lifestyle, famine and natural disasters [4]. While current age estimation methods, for example, Greulich-Pyle or Tanner–Whitehouse method, can accurately estimate age from radiological examination of the left hand in children, age assassination complexities significantly increase when comparing young adults to adolescents. The absence of formal age documentation makes the process even more complicated despite recent discoveries in anthropometric fields demonstrating the importance of other bones, such as the pelvis, in determining bone age [1, 2]. OPG or dental radiographs remained to be the primary forensic concern.

Orthopantomography, also known as an OPG or dental radiographs, provides a wide panoramic view of the lower face bones. OPG displays all the teeth on both jaws. OPG also depicts the jawbone and the temporomandibular joint (TMJ) on the same radiography film. The dental OPG images have been suggested to be more effective when compared to other bones of the body X-ray images. Teeth bones are long-lasting and resistant to high temperatures and organisms which decompose a dead body [5]. The Malaysian Institute of Forensic Medicine claimed that the X-ray images of teeth are the primary and the most used method in estimating age [6, 7]. However, the traditional image processing approach for OPG age estimation is still relatively expensive, laborious and requires long-term monitoring, increasing potential radiation exposure [8].

Traditional image processing for dental age measurement is a manual method that may include several steps such as segmentation, feature extraction, image pre-processing, classification, or regression. Each of these steps is error-prone and can induce variations in the outcome. For example, in an identical age range, the dry bone image will be different from the wet bone in radiography scan. In recent years, Machine Learning/Deep Learning techniques have been widely used to identify patterns in complex data such as clinical imaging, genomics, bioimaging, and phenotypic data [9-13]. Machine learning algorithms for example, NN, SVM, KNN, Decision Tree, displayed promising ability in prediction and classification [14-23] including estimation of bone age [24, 25]. The medical and biological datasets are increasing rapidly. To analyse such big and complex data, artificial intelligence and machine learning algorithms become most popular [26-35]. Therefore, it is important to implement novel techniques to uncover the medical and biological patterns. In particular machine learning and deep learning techniques have been widely used to analyse imaging data [19, 32, 36].

Deep learning-based methods, also known as end-to-end learning-based methods, such as convolutional neural networks, are either unsupervised that operate directly on the input images and generate the desired output without intermediate steps such as segmentation and extraction or supervised. Supervised learning (SL) models utilise a learning function that requires example input-output pairs for the model validation. The supervised approach runs the comparison between a radiographic image of each subject to an existing reference (i.e., labelled images), including

gender and age information [37]. Nevertheless, most of these methods require a large amount of labelled data.

Collecting and labelling large datasets are usually expensive, time-consuming, or sometimes impossible. Thus, in recent years, Semi-Supervised Learning (SSL) has emerged. Leveraging both labelled and unlabelled data, SSL has been proved to be a practical approach. Transfer learning (TL), an SSL-based model) aiming to utilise training knowledge gained from one data set to analyse another using a few-shot classification framework [38, 39]. Moreover, unprocessed image databases, which usually can be fed to conventional ML or DL models, are prone to inaccurate class recognition. Extracting specific characteristics that accurately represent behaviour in different environmental conditions is also a time-consuming task. [7, 40]. Hence, choosing a classifier that can differentiate behaviours in terms of diversity within and outside each class is one of the challenges in recognising behaviour. Therefore, converting raw data into attribute vectors systematically and using effective attribute extractors using an engineering approach is required for proper classification [41-43].

Despite these improvements developing and training deep neural networks remained challenging and time-consuming. Thus, Transfer learning (TL) can use the pre-trained deep network to perform data classification, attracting increasing attention [44, 45] this article proposed an innovative automated machine learning model approach to estimate bone age. This approach is based on a deep convolutional neural network [46]. We improved the accuracy of age determination in dental panoramic images with approximately six-month intervals. The proposed method determines the precise age in the range of 15 to 23 years old that is divided into nine age groups. Each age group includes subcategories of images with one year, six, three and one-month(s) intervals. This method combined the methodology of features extraction of the convolutional neural networks with the nearest neighbour feature to analyse and classify information and features in dental images.

More specifically, the architecture of the proposed HCNN-KNN model used in this study consists of four steps (Fig1). Initially, a convolutional neural network (CNN) model was trained on OPG images. Then, a fully connected layer generated from CNN was fed into a principal component analysis (PCA). The PCA performed data transformation and dimension reduction. In the last step, data were classified using the KNN algorithm. The details of the prediction process of the classes in the data set were explained in the following sections.

Related Literature

Age estimation based on dental X-ray images remained a critical field to develop. Tooth bones are immortal and can last more than 20 years through post-mortem decomposition. This made the tooth an essential part of the forensic investigation (Blatt, 2020). Although the number of bone age estimation models has increased, most remain in the research state with slow or no transition in the industry. Computerised methods for bone age estimation based on OPG images are currently limited to ± six months precision [46]. Moreover, current Deep neural network algorithms used to estimate the age using dental X-ray images are, for example, AlexNet and ResNet. Training these algorithms requires large imaging datasets. For instance, Houssein et al. used a dataset with 1,429 dental X-ray images to reach acceptable output parameters [7].

In another study, Avuçlu et al., (2020) used 1315 dental images and 162 different dental classes to morphologically predict the age and gender in dental X-ray images with 95% accuracy [47, 48]. Farhadian et al. (2019) applied Transfer learning to reduce using a neural network that relied on dental data to alter age assessment. The age group of between 14 and 60 was used, and the overall data sample size for the research was three hundred images. With an MAE of 4.12 years, the neural network methodology showed a reasonably accurate outcome [25]. The neural network errors were smaller than the regression model's estimation errors, with the test data set being RMSE of 10.26 years and MAE of 8.17 years.

Alkaabi et al. (2019) examined different Convolutional Neural Networks s AlexNet, VGGNet, and ResNet. Architectures for age estimation. They used common CNN architectures to perform forensic dentistry's automatic age calculation without any modifications. Using Capsule-Net, predicting age estimation from dental images was also performed, which depicted 36% higher accuracy than the CNNs model. The Capsule-Net model based on transfer learning reached a cumulative accuracy of 76% for $\pm$ 1-Year-old sample OPG images.

Tao et al. (2019) has introduced the Multilayer Perceptron Neural network to estimate dental age. The experiments are carried out on a dataset composed of 1636 samples. It was also experimentally confirmed that this latest feature set makes the dental age estimation more reliable for sample OPG images of $\pm$ 1-Year-old individuals' sample OPG images.

Kim et al. (2019) applied a deep learning algorithm, based on CNN neural, to X-ray images of 9435 cases (4963 male, 4472 female). Data in this cohort was sorted into three-age gatherings. Their study suggests that deep learning algorithms based on CNN neural networks show that the proposed approach functions evaluated based on a database of panoramic dental radiographs and worked well for accuracy.

Banar et al. (2020) also implemented a fully automatic method leveraging the total capacity of the deep learning approach on their study of a dataset of 400 OPGs x-ray images with 20 OPGs per category and per gender. To overcome the limitations of having a dataset with a limited number of embodiments, the Barnar group employed a transfer learning approach using pre-trained CNNs, and data augmentation. Their study significantly reduced the imaging assessment time over the current conventional method. For example, they have reported that the entire automatic workflow took 2.72s to compute on average. Given the small size of the dataset, this pilot study indicated the strength of transform learning and suggested a completely automated solution capable of demonstrating outcomes not inferior to manual assessing the images.

Tuan et al. (2017) introduced a semi-supervised fuzzy clustering algorithm for pre-image processing and segmentation of dental X-ray images. Their study showed that their recommended work has superior accuracy than the initial semi-supervised fuzzy clustering and several related approaches.

De Tobel J. (2017) developed machine learning algorithms to estimate the age based on bones of the hands, wrists, and dental on X-rays and magnetic resonance imaging (MRI) images selected from the World Dental Federation notation element (FDI) database 38. The photos initially were evaluated manually by two observers in a unanimous manner. A third observer acted as the referee

and proposed a reference point in case of significant discrepancies. Then, they used a classification method to place a rectangular box positioned at the ROI using Adobe Photoshop CC 2017. Consequently, they measured in a 5-fold reliability situation using several credibility criteria, including accuracy, degree-N detection speed, mean absolute difference, kappa linear coefficient. Using their machine learning algorithm, the mean accuracy was calculated at 51.0, the average unconditional disagreement was reposted to be 0.6, and the kappa average was 0.82 in $\pm$ 1-year-old sample OPG images.

In the Department of Dentistry and Study of University Sains Islam Malaysia, an age evaluation approach was tested on Malaysian adolescents between the ages of 1 to 17. Initially, the first to the third teeth were segmented, and the invariant deformation characteristics were then collected based on a deep learning method. The designed DCNN model was used to extract a broad range of features in the hierarchical layers, for example, invariance of size, rotation and deformation for sample OPG images of $\pm$ 1-Year-old individuals.

Material and Methods

Image Pre-Processing and Data Augmentation

Prior to utilising images in our model, we performed image processing, adjusted the contrast and highlighted the images, removed extra margins (cropping), and normalised the image pixel values at edges from zero to one. In this study, data augmentation was performed by mirroring and data duplication.

Data were classified before, Splitting for Training the model. To perform the classification, the data in the dataset were divided into two categories: train data and validation data. The cross-validation was performed with 50-50, 70-30, and 20-80 train and validation split. To take an example, 20-80 split the 20% of the data for validation and 80% for training the model. Obtaining acceptable results on various test data will indicate the correct performance of the proposed model. Thus, to further enhance the validation of the model, we perform additional validation of the trained model using a data set that the model was seeing for the first time. The additional validation dataset included various features such as OPGs of different races, with different genders and at different ages.

The Architecture of HCNN-KKN Proposed Model

In this study, we proposed a model based on a combination of the HCNN-KKN. Steps in our hybrid model depicted in Figure 2 for pre-image processing and the dental classification are as follows:
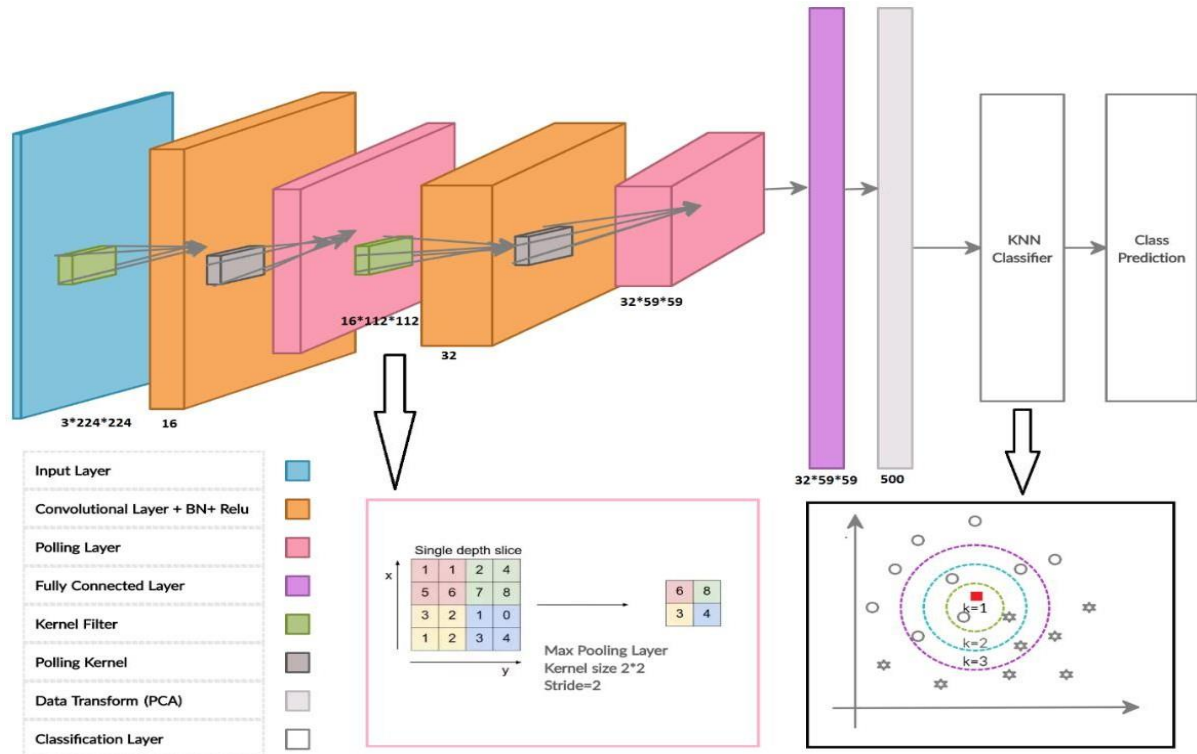
Figure 1. The Architecture of The Proposed HCNN-KNN Model

The architecture of the proposed HCNN-KNN model included an inputs layer, two Convolutional layers, a fully connected layer, a PCA transform data layer, and a KNN layer for classification (Figure 1). Using RGB images (224x224 pixels) as input, each image was fed to a convolutional layer (i.e., the second layer) containing 16 convolutional filters and a kernel size of 5 and 2 paddings. Low-level features like edges, blobs, shapes were obtained. In addition to the 16 convolutional normalisation filters, batch normalisation (16 layers), followed by a nonlinear activation function via rectified linear unit (ReLU) were performed. The convolutional layer is termed 'conv1'. Maxpooling layer downsampled the model with stride 2 and 2 paddings.

As a result, the Maxpooling layer reduced the dimensions of the images to 112×112 pixels, and the final output of this layer is equal to $112 \times 112 \times 16$. These images were fed to the second convolutional block with a similar structure to the first. The later conventional block has 32 filters with 5 and 2 paddings kernel, resulting in the $56 \times 56 \times 32$ images output. In the following step, all the fully connected layers were used to connect all the neurons in different layers (Table 1). To reduce the less important features to the next layer in the CNN model, the layers are connected with a 50% dropout between fully connected layers. We then applied a classification approach on a cohort of individuals between 15 to 23 years old and stratified all the cases with intervals of ± 6-months, ± 3-months, and ± 1-month. This resulted in 9 categories for each year (+_ Year), 18 categories for the six months intervals (each six months), 36 categories for each trimester of each year (each three months), 108 for every 12 months from each year (each one month). This resulted

in the last fully connected layer to modify nine classification tasks with parameters like bias learn rate factor and weight learn rate factor. Figure 1: Illustration of the CNN Network elements.
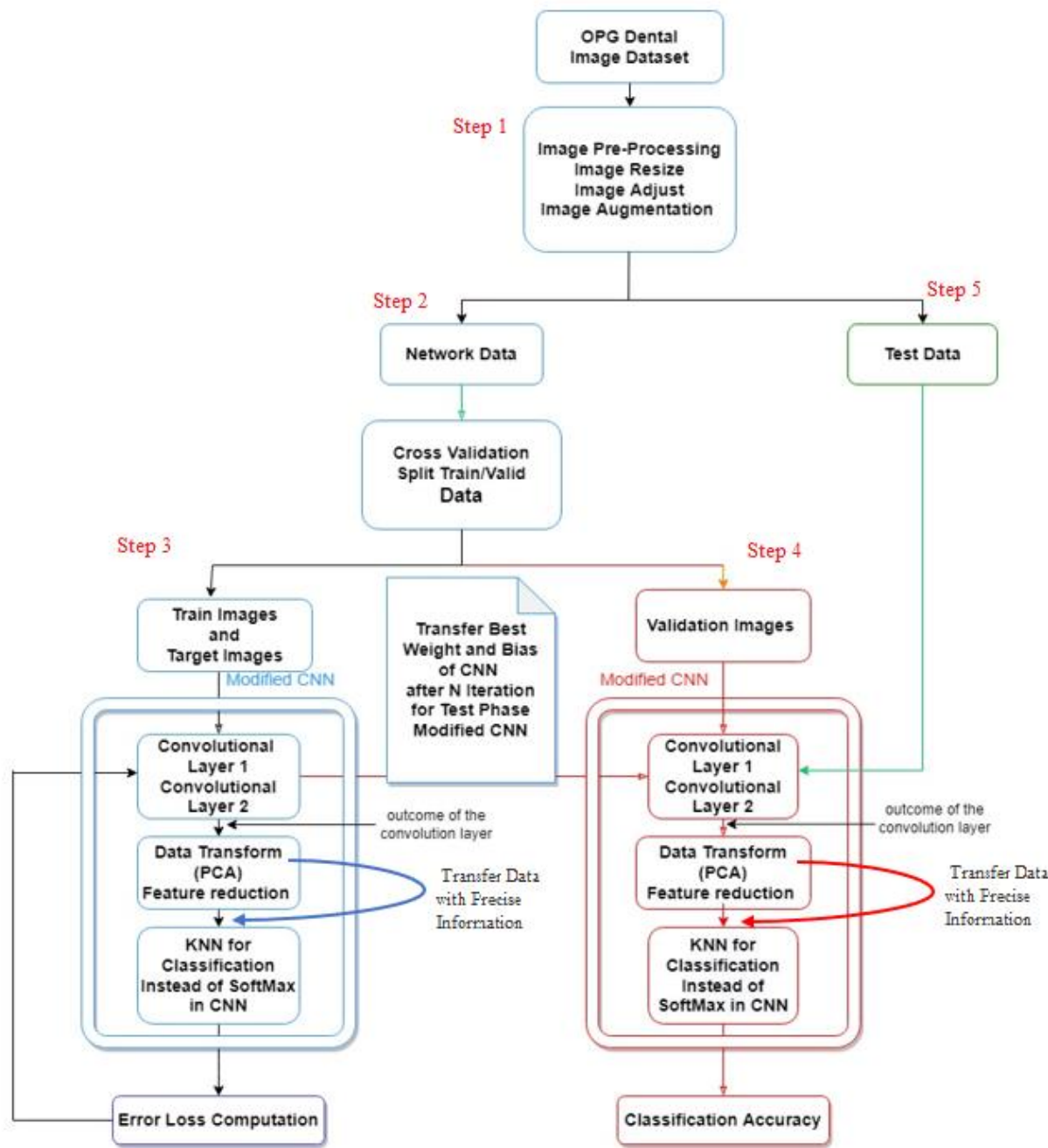

Figure 2. The Conceptual Framework for our hybrid model, dental Image Processing and Classification Processes.

In this step, a combined CNN model with KNN is used as a classifier to evaluate the class event instead of the SoftMax probability layer. Thus, before the data enters a fully connected layer of KNN layers, we proposed that in this model, the data were transferred in the space with more

precise information and smaller dimensions using the principal component analysis method so-called 'Fully_PCA'. Therefore, fully connected layer neurons after PCA data transformation and reductions are considered overlays of the KNN classification layer. In this case, the KNN input is assumed to be 500.

The HCNN-KNN model algorithm pseudo-code proposed in this bone age study is based on dental images. In the KNN network, the distance metric used is the Euclidean Distance metric and the "Ks" are 1, 2, 3. The model is trained using an SGD momentum optimiser with a 0.0001 learning rate. The model is then trained for ten epochs with a mini-batch size of 16 images. After learning the network and classification operations, the error loss rate is obtained to evaluate the proposed network model in subsequent epochs of network learning.

Evaluating the Proposed Model on Validation Data

The network neurons' optimal weight and bias were evaluated following the network training step. The validation data, segmented according to figure 2 (step 2), was then categorised on the CNN network with optimal parameters. The components of the classification algorithms were determined according to the training phase process and the criteria that allow the data to pass through the CNN-PCA-KNN layers. Finally, the identification rate was evaluated in which accuracy of the index was selected according to the Confusion Metrics. In brief, the sum of TP+TN (TP: the number of times the class is upbeat and correctly detected, FP: the number of times the class was positive and undiagnosed), divided with the sum of TP+FP+TN+FN. TP is the number of times the class is positive and correctly detected, FP abbreviation for the number of times the class was positive and undiagnosed. The TN, denotes the number of times the class is incorrect (negative) and correctly diagnosed. The FN represents the number of times each class is erroneous (negative) or not correctly recognised.

Validating the Proposed Model on a Test Dataset

To further validate the accuracy of the proposed model, we used test data which the model had not seen before, and was not part of the training dataset. This dataset was collected separately from the trained dataset and had different properties. This data combines X-ray OPG images of other races, with different genders and at different ages. This data enters the CNN network as a validation step and is categorised after passing through conv1, pooling1, conv2, pooling2, fully connected layer, PCA layer, and KNN layer. A vital feature of this step is the validation of the proposed model. Achieving acceptable results on various test data will indicate the satisfactory performance of the proposed model.

Computational Hardware Requirement

For this study, we used an NVIDIA GeForce GTX 1080 8GB, 8GB memory, Intel Core i7, and 3.40GHz. We also used Python libraries NumPy, Matplotlib, Sklearn, Metrics, PyTorch, Torchvision to perform the analyses.

Result and Discussion

We discussed the database used in this article, the tools, and the architecture introduced in the following. Finally, we evaluated our proposed model on the dataset.

Datasets Description

Dental OPG is a panoramic radiograph that scans the upper and lower jaws with a two-dimensional view that shows a semicircle from ear to ear. The used images in this study are dental OPG data, collected from dental teaching institutes and private dental clinics in Malaysia. These collected data with the different image sizes are resized before importing into the proposed model. The original image size for all data has been changed to 600×1024 pixels.

Subjects from this study were randomly selected from 1922 patients between 15 and 23 years of age, and this age range determines the age of minors which could help the model find more specific features and train the model more accurately, particularly for smaller time intervals. Moreover, to further validate the model, we used a dataset that hasn't been presented to the model during conventional training or the initial validation. Using an additional dataset including 130 random images, we showed that pre-training the model on such a prominent age recognition features (i.e. shorter time intervals, form of age range by year, age range by month, by season (3 months), and six months (Dataset on Year, on ± six months, on ± three months, and ± one month), allows us to evaluate our proposed method on a test database to enhance the performance of the proposed method (further details are provided in the experimental Result , Table 1, Figure 3).

Table 1: A Display of The Parameters Used in The CNN Network.

	Type	Filter size	Stride	# Filters	FC	Input
Layer1	Convolution	5×5	2×2	16	-	224×224×3
Layer2	Maxpooling	2×2	2×2	-	-	224×224×16
Layer3	Convolution	5×5	2×2	32	-	112×112×16
Layer4	Maxpooling	2×2	2×2	-	-	112×112×32
Layer5	Fully Connected	-	-	-	*	59×59×32

*9 for Dataset on Year, 18 for a dataset on ±6-month, 36 for a dataset on ±3 month and 108 for a dataset on ±1 month

15 years   16 years   17 years

18 years   19 years   20 years

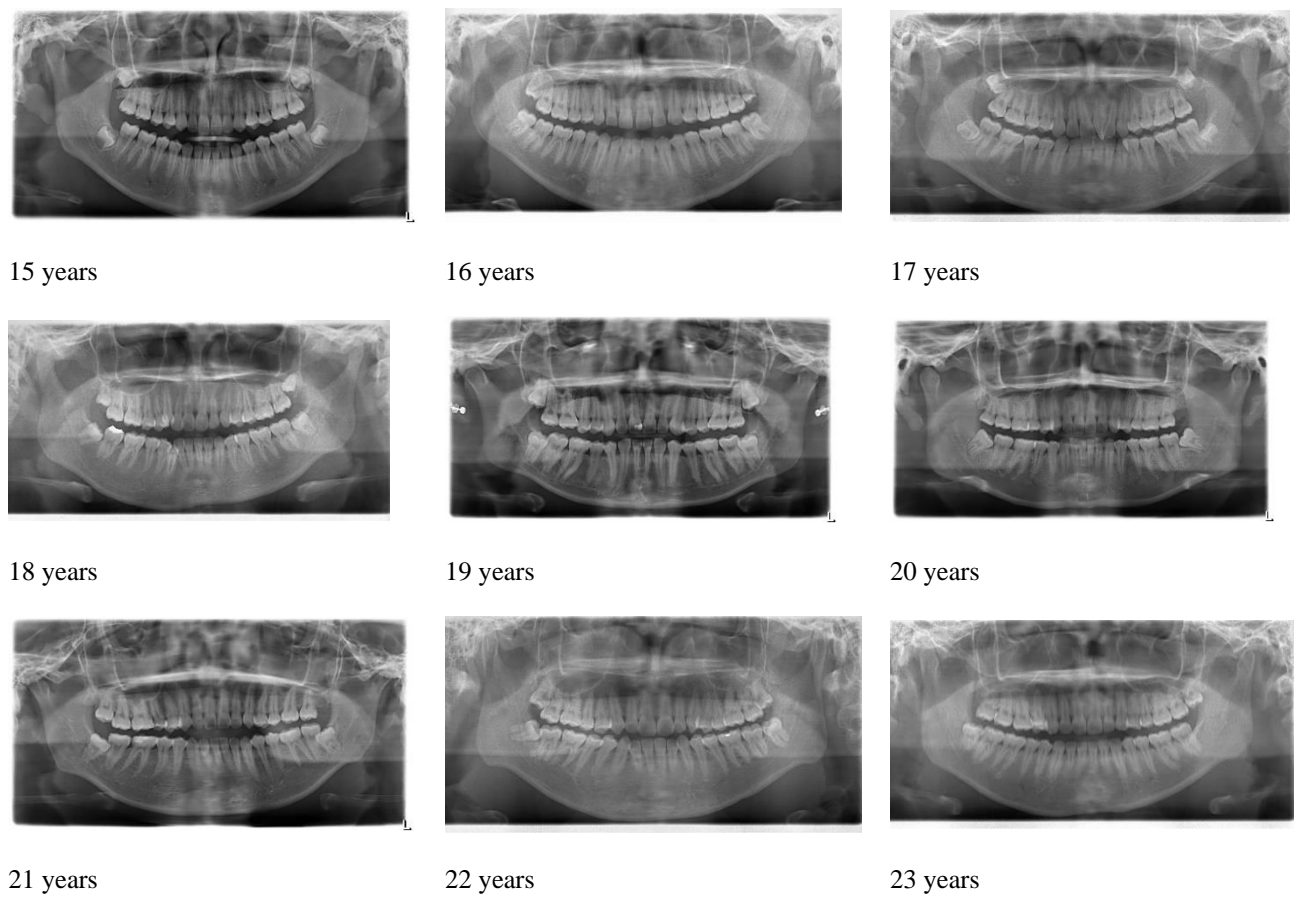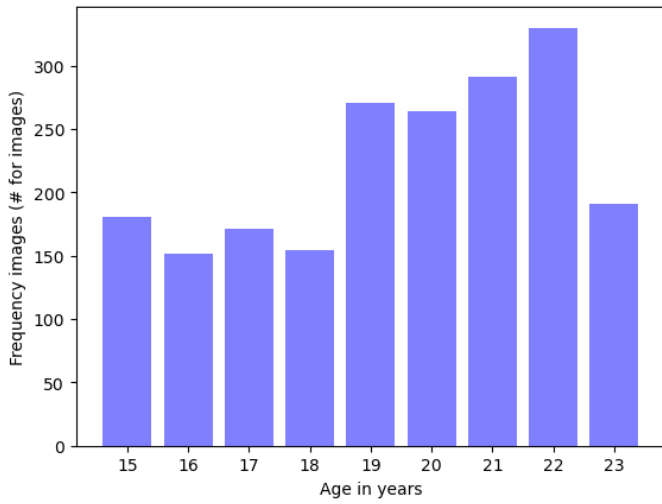21 years   22 years   23 years

Figure 3: Representation of Each Age Class

We elucidated the frequency of the dataset generated as the result of data augmentation when augmented more than 150, 68, 32 and 7 images in 1 year, 6 months, 3 months and 1-month categories, respectively (Figure 4).

a. Frequency of dental images for years dataset
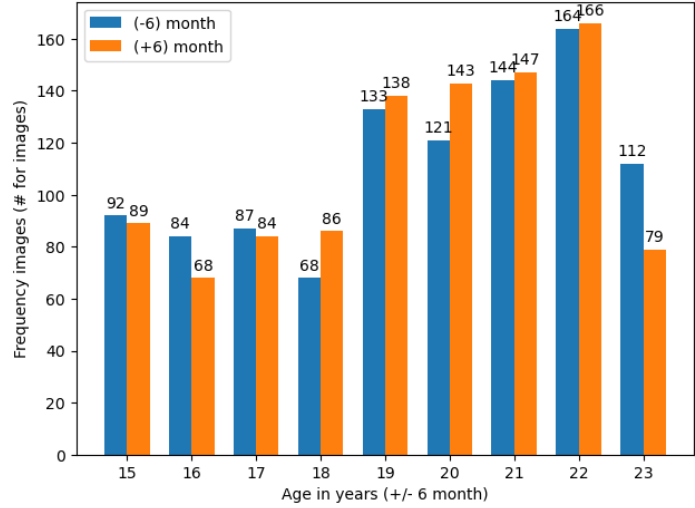
b. Frequency of dental images for ±6 months dataset
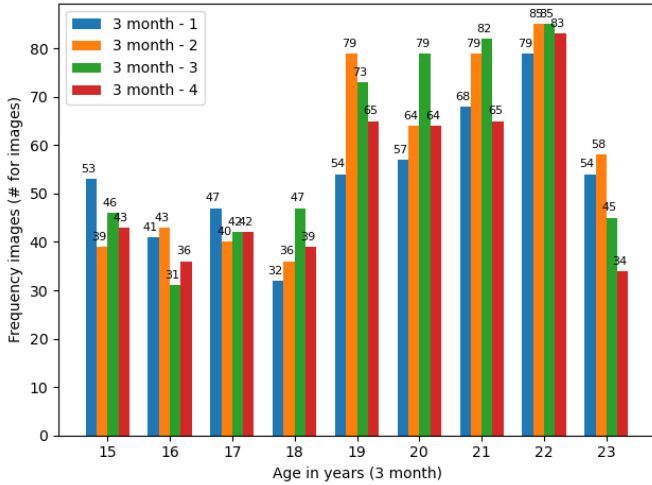
c. Frequency of dental images for ±3 months dataset
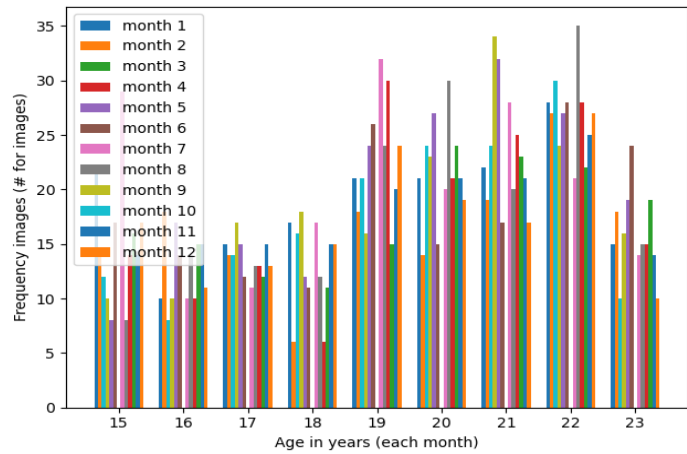
d. Frequency of dental images for ±1 months dataset

Figure 4: Frequency of Dataset Images Based on the Four States of the Dataset (A. Dataset on Year, B. On ±6-Month, C. On ±3 Month and D. On ±1 Month)

CNN Model Initial Accuracy on Four Different States of The Dataset

As the first experiment, we applied a CNN model on two other dataset states. In stage one, training and validation accuracy were evaluated from the original dataset without any pre-processing and data augmentation. In stage two, training and validation accuracy were assessed post-pre-processing operations on the data augmentation operations (Table 2).

Table 2: Experimental result on a dental dataset based on CNN.

Dataset	Train/Valid	Accuracy %			
		Dataset on Year	Dataset on ±6 month	Dataset on ±3 month	Dataset on ±1 month
Original	Train Accuracy	77.85	76.24	72.41	70.26
	Valid Accuracy	58.91	43.72	38.75	23.29
Augmented- pre-processed	Train Accuracy	98.84	98.58	98.75	96.99
	Valid Accuracy	97.43	98.44	98.13	95.62

We observed that pre-processing and data augmentation operations consistently improved the accuracy by approximately 20% and 40% for training and validation, respectively (Table 2).

We then evaluated the loss reduction for each step (i.e., four separate modes) of this data set (1 Year, ±6 months, ±3 months, and ±1 months) with 4000 repetitions (i.e., epoch equivalent to 10 with about 400 batch size 16). Note that the results are obtained on cross-validation with 80% training and 20% validation and segmentation on dataset images. We observe a sharp decay in loss reduction values for both the 1 year and ±3 months dataset, while the ±6 months and ±1 month dataset showed a more gradual decay (Figure 5).
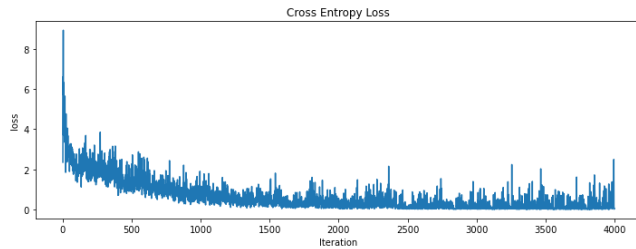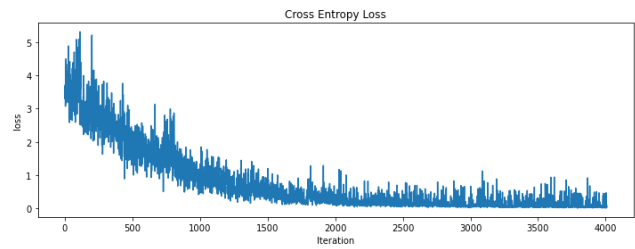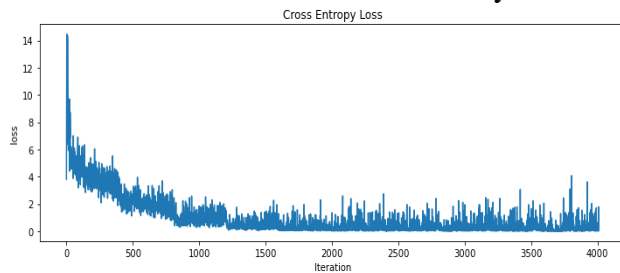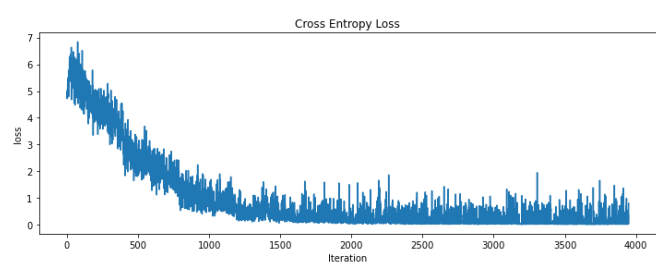

a. Error loss for dataset years

b. Error loss for the dataset with ±6 months

c. Error loss for the dataset with ±3 months

d. Error loss for the dataset with ±1 months

Figure 5: Error Loss for Training on A Dataset Based on CNN Model on the Four States of the Augmented Dataset (A. Dataset on Year, B. On ±6-Month, C. On ±3 Month and D. On ±1 Month)

We also evaluated the accuracy of the proposed model on train and validation data in 10 epochs. We observed that the accuracy of train data is generally slightly higher than validation data (Figure 6). Therefore, the upward trend inaccuracy at ten epochs indicates an increase in CNN learning.
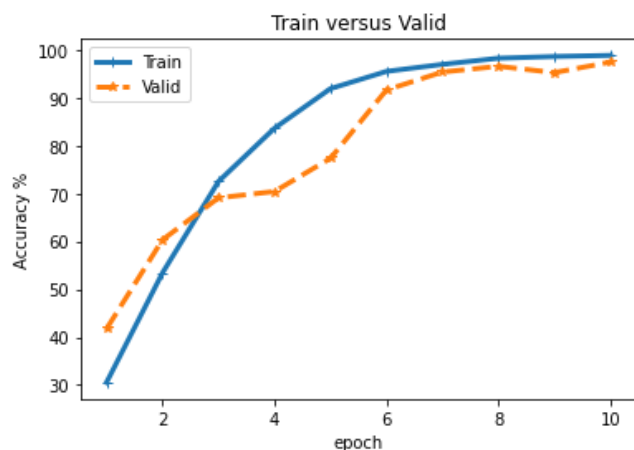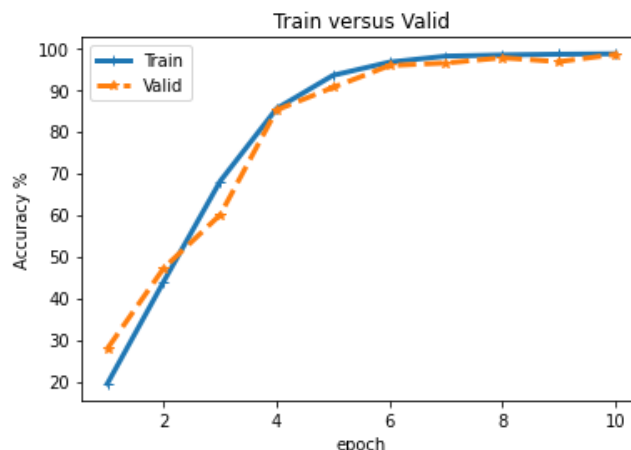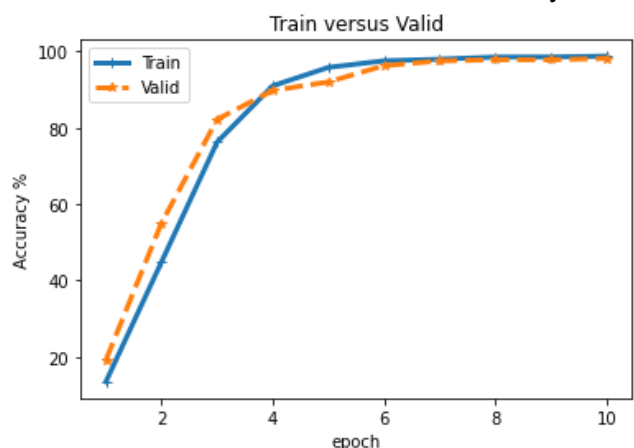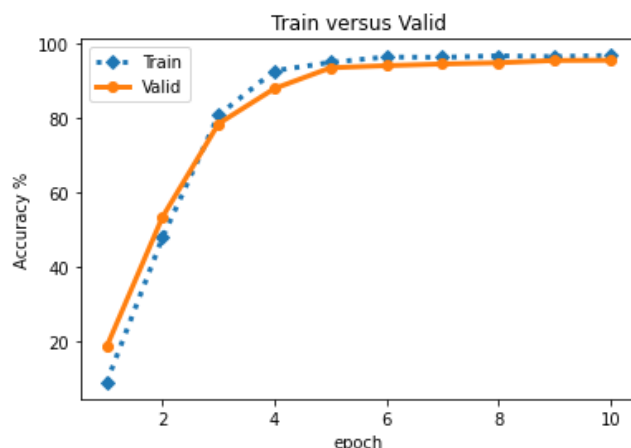
a. Train versus Valid for dataset years

b. Train versus Valid for the dataset with ±6 months

c. Train versus Valid for the dataset with ±3 months

d. Train versus Valid for the dataset with ±1 months

Figure 6: Evaluation of accuracy (Train versus Valid) in different epoch based on CNN model on four states of the dataset (a. Dataset on Year, b. on ±6-month, c. on ±3 month and d. on ±1 month)

Combining HCNN-KNN Enhanced the Accuracy of the Model.

We used a combined methodology of convolutional neural networks and the nearest neighbour (classification) to further enhance the model's accuracy. The resulting hybrid model (HCNN-KNN) contained three main CNN, PCA, and KNN layers. Our findings showed that the KNN classification with 500 features selected from 'Fully_PCA' (PCA data transform after fully connected in CNN) had increased the accuracy of the original data to approximately the same level of augmented data to 99.98, 99.96, 99.87, and 98.78 for one year, ± 6-months, ± 3-months, and on ± 1-month dataset, respectively. Intrinsically, we observed accuracy for train data to be equal to 100% (Table 3). This outcome indicated is fully trained using the training dataset. It is noteworthy that the results obtained for k-nearest neighbours value k = 1 in the proposed HCNN-KNN hybrid model. We have also evaluated the KNN method in the proposed HCNN-KNN hybrid model with a "k" value of 2 and 3 (Table 4). Model validation is essential for building a model and is susceptible to common pitfalls when splitting the training dataset [49, 50]. To evaluate the possibility of overfitting pitfall resulting from splitting the training dataset, particularly for small ks (e.g., k= 1-3), we further validate the model using a dataset of dental images separate from the data used for initial training

13

and validation; contained its special features. The unique features of these images are related to different races, different genders and are selected in the age range of 15 to 23. While obtaining acceptable results on these data indicates the proper performance of the KNN classifier even with k = 1.

Table 3: Experimental result on a dental dataset based on HCNN-KNN.

Dataset	Train/Valid	Accuracy %			
		Dataset on Year	Dataset on ±6 month	Dataset on ±3 month	Dataset on ±1 month
Original	Train Accuracy	100	100	100	100
	Valid Accuracy	99.93	99.84	99.74	98.15
Augmented	Train Accuracy	100	100	100	100
	Valid Accuracy	99.98	99.96	99.87	98.78

Table 4: Accuracy results of k = 1, 2, 3 on the validation data

K	Accuracy %			
	Dataset on Year	Dataset on ±6 month	Dataset on ±3 month	Dataset on ±1 month
K=1	99.98	99.96	99.87	98.78
K=2	59.26	55.61	52.64	51.20
K=3	43.89	41.64	38.02	33.77

Evaluation of Different Cross-Validation

Different cross-validation of datasets was applied to evaluate our method (e.g., 50-50, 70-30 and 20-80 train-valid cross-validation, Table 5).

Table 5: Evaluation of Accuracy on Different Cross-Validation on a Dataset

Dataset Type	Methods	Train/Valid	Accuracy on different Cross-Validation (valid-Train)		
			20-80	30-70	50-50
Dataset on Year	CNN	Train	98.84	98.16	97.20
		Valid	97.43	94.08	85.58
	HCNN-KNN	Train	100.00	100.00	100
		Valid	99.98	99.97	99.95
	Resnet (Deep model)	Train	99.04	98.83	98.80
		Valid	98.25	95.75	86.03
Dataset on ±6 month	CNN	Train	98.58	98.50	96.75
		Valid	98.44	94.72	86.66
	HCNN-KNN	Train	100.00	100.00	100
		Valid	99.96	99.78	99.78
	Resnet (Deep model)	Train	98.66	98.53	98.42
		Valid	98.69	95.47	87.43
Dataset on ±3 month	CNN	Train	98.75	98.48	97.77
		Valid	98.13	96.25	84.61

	HCNN-KNN	**Train**	**100.00**	**100.00**	**100.00**
		Valid	**99.87**	**99.82**	**99.68**
	Resnet (Deep model)	**Train**	98.69	98.74	98.70
		Valid	98.44	96.42	86.23
Dataset on ±1 month	CNN	**Train**	96.99	97.12	94.46
		Valid	95.62	92.40	80.32
	HCNN-KNN	**Train**	**100.00**	**100.00**	**100.00**
		Valid	**98.78**	**98.52**	**98.29**
	Resnet (Deep model)	**Train**	97.25	97.24	97.15
		Valid	95.63	93.16	90.27

Evaluation of Proposed Model on Second Dataset (Test Dataset)

One of this study's innovative and key features is the additional validation step in which a set of images from a diverse cohort has been used to further validate the proposed model. After the initial training and validation, the images are only available to the model. The dataset had a combination of 130 X-Ray OPG images with different races (e.g., Malay, Indian and Chinese), with distinct features, various age ranges and normal distribution of ordinary dentistry, orthodontic and malignant images (Figure 7).
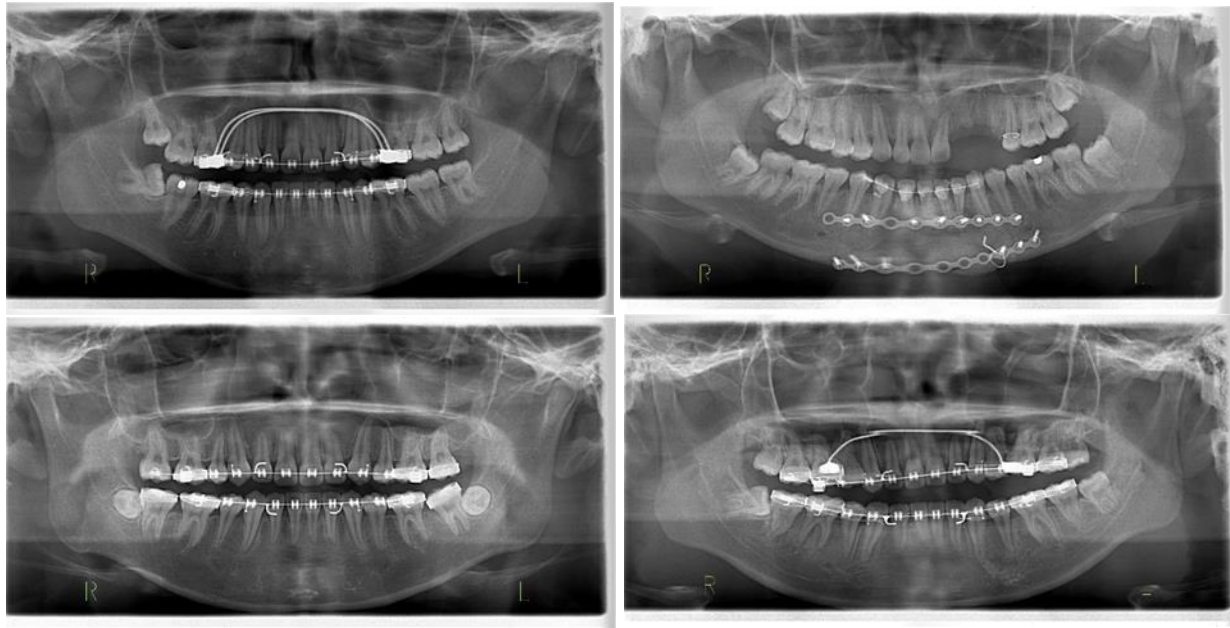

Figure 7: Different Images of Test Data to Evaluate the Orthodontics (Malignant) of the Proposed Model.

As Table 6 shows, the proposed HCNN-KNN model has obtained better results on the test data. Getting satisfactory results on a new dataset outside the dataset images indicates the performance of the proposed model. Thus, it seems that the proposed HCNN-KNN model has obtained superior results on the tested dataset. Obtaining satisfactory results on a new dataset outside the dataset images indicated the outstanding performance of the proposed model (Table 6).

Table 6: Accuracy Result of the Proposed HCNN-KNN Model on Test New Dataset

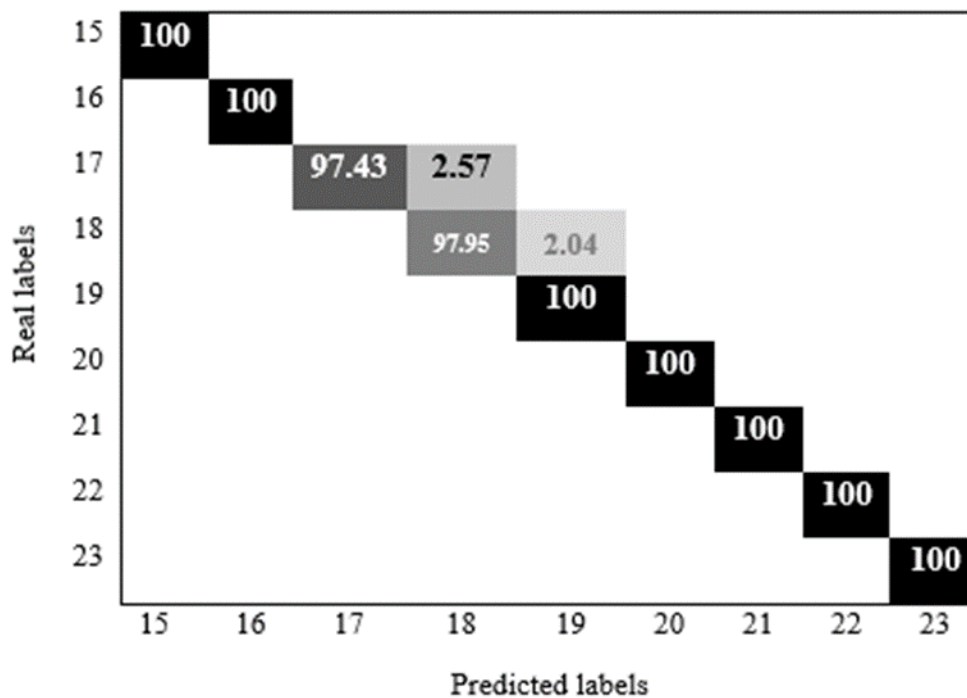|		CNN	HCNN-KNN	
All test Data	**Test data (130 dental images)**	22.93	**K=1**	**99.48**
			K=2	74.61
			K=3	70.22
Based on Different Races	**Malay (71 Images)**	25.23	**K=1**	**100**
			K=2	76.05
			K=3	73.23
	Indian (16 Images)	20.61	**K=1**	**93.75**
			K=2	75.00
			K=3	68.75
	Chinese (43 Images)	21.37	**K=1**	**97.67**
			K=2	72.09
			K=3	69.76

Figure 8: Confusion Matrix of Hybrid HCNN-KNN Proposed Model Application on Enhancement Dataset.

Figure 8 shows the confusion matrix of 97.95 % for the 18-year-old age range which was correctly estimated, and 2.04 % was incorrectly identified for the 19-year-old.

Models Accuracy

The collected data and the proposed model have been tested with other models such as ResNet, CNN, GoogLeNet Inception (Table 7). Our finding shows that the hybrid HCNN-KNN model has obtained much higher accuracy than different conventional classification algorithms tested.

Table 7: Comparison Between the Proposed HCNN-KNN Model and Other Studies in BAM Based on Dental Images.

Models	Accuracy	References
CNN	**97.43**	**Other Model Test Result**
HCNN-KNN	**99.98**	**The Proposed Hybrid Model Result**
ResNet	**98.25**	**Other Model Test Result**
GoogLeNet Inception	**55.69**	**Other Model Test Result**
SVM	92.1	[51]
FNN-TLBO	89.00	[52]
Multilayer Perceptron	90.00	[47]

The novelty of our study is that to the best of our knowledge, no other study reported measurements of ± six months, ± three months, and ± one month so far, but only compared by the year. However, the results of this proposed model are compared with previous studies by years of accuracy. For example, Avuçlu used the Multilayer Perceptron model with 1315 dental OPG images, and the highest accuracy was 90% just for accuracy of the years. Also, in 2020, Hussein used the SVM algorithm for a population of 1429 dental radiographs, and they only achieved 92% accuracy for years. Therefore, based on Table 7, the results of the comparative studies for accuracy by years are still below the accuracy achieved in this study which is 99.98.

Conclusions

For the first time, with precision to the range of +/- 6 months, this novel implementation developed an HCNN-KNN model, used for BAM to increase the existing model's accuracy and prevent CNN overload situations. We considered more specific cases of bone age measurement with the help of Dental X-Ray OPG images to solve the problem of bone age measurement in determining the age range of 15 to 23 years based on the Year, ± six months, ± three months, and ± one month. We integrated the methodology of convolutional neural networks for extraction and analysis of information and features in dental images) and the nearest neighbour (classification method). The primary purpose of this proposed model is to use KNN instead of SoftMax in the fully connected layer to increase the performance of convolutional networks in the classification phase. Using principal component analysis as data transform and feature dimensionality reduction in a fully connected layer before classifying KNN as Fully_PCA. The primary purpose of this method is to transfer data to space with more specific information to increase the performance of the classification phase.

Our proposed model achieved the accuracy of 99.98, 99.96, 99.87, and 98.78 in one year, ± 6-month, ±3-month, and ± 1-month range, respectively. Our proposed method evaluated different cross-validation of the dataset. 50-50 train-valid, 70-30 train-valid and 20-80 train-valid cross-validation on the dataset. Evaluating the proposed model on a new dataset with different races also proved the superior performance of the model. The benchmarking with current existing models also showed that the HCNN-KNN model is the best model for bone age measurement.

Author Contribution

Conflict of interests

Acknowledgement

Reference

1.   Wik, E.H., et al., *Skeletal maturation and growth rates are related to bone and growth plate injuries in adolescent athletics.* Scandinavian journal of medicine & science in sports, 2020. **30**(5): p. 894-903.
2.   McGill, D., *Bone Age*, in *Endocrine Conditions in Pediatrics*. 2021, Springer, Cham. p. 215-218.
3.   Dahlberg, P.S., et al., *A systematic review of the agreement between chronological age and skeletal age based on the Greulich and Pyle atlas.* European radiology, 2019. **29**(6): p. 2936-2948.
4.   Gurpinar, F., et al. *Kernel ELM and CNN based facial age estimation*. in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2016.
5.   Ezhil, I., N. Jagannathan, and M. Kumar, *Estimation of age from physiological changes of teeth.* Drug Invention Today, 2018. **10**.
6.   Čular, L., et al. *Dental age estimation from panoramic X-ray images using statistical models*. in *Proceedings of the 10th International Symposium on Image and Signal Processing and Analysis*. 2017.
7.   Houssein, E.H., N. Mualla, and M. Hassan, *Dental age estimation based on X-ray images.* Computers, Materials & Continua, 2020. **62**(2): p. 591-605.
8.   Shi, W., et al., *Fetal brain age estimation and anomaly detection using attention-based deep ensembles with uncertainty.* NeuroImage, 2020. **223**: p. 117316.
9.   Alinejad-Rokny, H., *Proposing on Optimized Homolographic Motif Mining Strategy Based on Parallel Computing for Complex Biological Networks.* Journal of Medical Imaging and Health Informatics, 2016. **6**(2): p. 416-424.
10.   Javanmard, A., A.J.I. Montanari, and I.A.J.o.t. IMA, *State evolution for general approximate message passing algorithms, with applications to spatial coupling.* Information and Inference: A Journal of the IMA, 2013. **2**(2): p. 115-144.
11.   Ahmadinia, M., et al., *Energy-efficient and Multi-stage Clustering Algorithm in Wireless Sensor Networks Using Cellular Learning Automata.* IETE Journal of Research, 2013. **59**(6): p. 774-782.

12. Mahmoudi, M.R., et al., *Principal component analysis to study the relations between the spread rates of COVID-19 in high risks countries.* Alexandria Engineering Journal, 2021. **60**(1): p. 457-464.

13. Hosseinpoor, M., et al., *New methodology to evaluate the Reynolds dilatancy of self-consolidating concrete using 3D image analysis-Coupled effect of characteristics of fine mortar and granular skeleton.* Cement and Concrete Composites, 2020. **108**: p. 103547.

14. Esmaeili, L., Behrouz Minaei-Bidgoli, and Mahdi Nasiri., *Hybrid recommender system for joining virtual communities.* Research Journal of Applied Sciences, Engineering and Technology, 2012. **4**(5): p. 500-509.

15. Hussein, I.J., et al., *Fully-automatic identification of gynaecological abnormality using a new adaptive frequency filter and histogram of oriented gradients (HOG).* Expert Systems, 2021: p. e12789.

16. KHAN, A.H., et al., *STATISTICAL FEATURE LEARNING THROUGH ENHANCED DELAUNAY CLUSTERING AND ENSEMBLE CLASSIFIERS FOR SKIN LESION SEGMENTATION AND CLASSIFICATION.* Journal of Theoretical and Applied Information Technology, 2021. **99**(5).

17. Mahmoudi, M.R., Akbarzadeh, H., Parvin, H., Nejatian, S., Rezaie, V., *Consensus function based on cluster-wise two level clustering.* Artificial Intelligence Review, 2021. **54**(1): p. 639-665.

18. Niu, H., Khozouie, N., Parvin, H., Beheshti, A., & Mahmoudi, M. R., *An ensemble of locally reliable cluster solutions.* Applied Sciences, 2020. **10**(5): p. 1891.

19. Parvin, H., Behrouz Minaei, Hosein Alizadeh, and Akram Beigi, *A novel classifier ensemble method based on class weightening in huge dataset*, in *In International Symposium on Neural Networks*. 2011, Springer. p. 144-150.

20. Parvin, H., Parvin, S., *A classifier ensemble of binary classifier ensembles.* International Journal of Learning Management Systems, 2013. **1**(2): p. 37-47.

21. Parvin, H., MirnabiBaboli, M. *Proposing a classifier ensemble framework based on classifier selection and decision tree.* Engineering Applications of Artificial Intelligence, 2015: p. 34-42.

22. Alinejad-Rokny, H., Pourshaban, H., Orimi, A. G., & Baboli, M. M., *Network motifs detection strategies and using for bioinformatic networks.* Journal of Bionanoscience, 2014. **8**(5): p. 353-359.

23. Bahrani, P., Minaei-Bidgoli, B., Parvin, H., Mirzarezaee, M., Keshavarz, A., *User and item profile expansion for dealing with cold start problem.* Journal of Intelligent & Fuzzy Systems, 2020. **38**(4): p. 4471-4483.

24. Tao, J., et al. *Dental age estimation in East Asian population with least squares regression*. in *International Conference on Advanced Machine Learning Technologies and Applications*. 2018. Springer, Cham.

25. Farhadian, M., et al., *Dental age estimation using the pulp-to-tooth ratio in canines by neural networks.* Imaging Science in Dentistry, 2019. **49**(1): p. 19-26.

26. Alinejad-Rokny, H., Anwar, F., Waters, S. A., Davenport, M. P., & Ebrahimi, D., *Source of CpG depletion in the HIV-1 genome.* Molecular biology and evolution, 2016. **33**(12): p. 3205-3212.

27. Alinejad-Rokny, H., Sadroddiny, E., & Scaria, V., *Machine learning and data mining techniques for medical complex data analysis.* Neurocomputing, 2018. **276**(1).

28. Bayati, M., Rabiee, H. R., Mehrbod, M., Vafaee, F., Ebrahimi, D., Forrest, A. R., *CANCERSIGN: a user-friendly and robust tool for identification and classification of mutational signatures and patterns in cancer genomes.* Scientific reports, 2020. **10**(1): p. 1-11.

29. Dashti, H., Dehzangi, A., Bayati, M., Breen, J., Lovell, N., Ebrahimi, D., *Integrative analysis of mutated genes and mutational processes reveals seven colorectal cancer subtypes.* bioRxiv, 2020. **2020**.

30. Ghareyazi, A., Mohseni, A., Dashti, H., Beheshti, A., Dehzangi, A., Rabiee, H. R., *Whole-genome analysis of de novo somatic point mutations reveals novel mutational biomarkers in pancreatic cancer.* Cancers, 2021. **13**(17): p. 4376.

31. Hosseinpoor, M., Parvin, H., Nejatian, S., Rezaie, V., Bagherifard, K., Dehzangi, A., *Proposing a novel community detection approach to identify cointeracting genomic regions.* Mathematical Biosciences and Engineering, 2020. **17**(3): p. 2193-2217.

32. Javanmard, R., JeddiSaravi, K., *Proposed a new method for rules extraction using artificial neural network and artificial immune system in cancer diagnosis.* Journal of Bionanoscience, 2013. **7**(6): p. 665-672.

33. Rajaei, P., Jahanian, K. H., Beheshti, A., Band, S. S., Dehzangi, A., *VIRMOTIF: A user-friendly tool for viral sequence analysis.* Genes, 2021. **12**(2): p. 186.

34. Alinejad-Rokny, H., Pedram, M. M., & Shirgahi, H., *Discovered motifs with using parallel Mprefixspan method.* Scientific Research and Essays, 2011. **6**(20): p. 4220-4226.

35. Ahmadinia, M., Ahangarikiasari, H., *Data aggregation in wireless sensor networks based on environmental similarity: A learning automata approach.* Journal of Networks, 2014. **9**(10): p. 2567.

36. Niu, H., Xu, W., Akbarzadeh, H., Parvin, H., Beheshti, A., *Deep feature learnt by conventional deep neural network.* Computers & Electrical Engineering, 2020. **84**: p. 106656.

37. Hemalatha, B. and N. Rajkumar, *A modified machine learning classification for dental age assessment with effectual ACM-JO based segmentation.* International Journal of Bio-Inspired Computation, 2021. **17**(2): p. 95-104.

38. Ouali, Y., C. Hudelot, and M.J.a.p.a. Tami, *An overview of deep semi-supervised learning.* arXiv preprint arXiv:2006, 2020: p. 05278.

39. Goldberg, A.B., *New directions in semi-supervised learning.* 2010, University of Wisconsin--Madison.

40. Spampinato, C., et al., *Deep learning for automated skeletal bone age assessment in X-ray images.* Medical Image Analysis,, 2017. **36**: p. 41-51.

41. Kalantari, A., et al., *Computational intelligence approaches for classification of medical data: State-of-the-art, future challenges and research directions.* Neurocomputing, 2018. **276**: p. 2-22.

42. Shamshirband, S., T. Rabczuk, and K.-W.J.I.A. Chau, *A survey of deep learning techniques: application in wind and solar energy resources.* IEEE Access 2019. **7**: p. 164650-164666.

43. Shamshirband, S., et al., *A review on deep learning approaches in healthcare systems: Taxonomies, challenges, and open issues.* Journal of Biomedical Informatics, 2021. **113**: p. 103627.

44. Qummar, S., et al., *A deep learning ensemble approach for diabetic retinopathy detection.* IEEE Access, 2019. **7**: p. 150530-150539.

45.   Li, J., et al., *Suphx: Mastering mahjong with deep reinforcement learning.* arXiv preprint arXiv 2013, 2020: p. 13590.

46.   Sharifonnasabi, F., et al., *A Review on Automated Bone Age Measurement Based on Dental OPG Images.* International Journal of Engineering Research and Technology, 2020. **13**(12): p. 5408-5422.

47.   Avuçlu, E. and F. Başçiftçi, *The determination of age and gender by implementing new image processing methods and measurements to dental X-ray images.* Measurement, 2020. **149**: p. 106985.

48.   Fatih, H. and G. Catalin, *RECREATION AND PHYSICAL ACTIVITY OF YOUNG GIRLS WITH INTELLECTUAL DISABILITY.* Science, Movement and Health, 2020. **20**(2 Supp): p. 237-241.

49.   Wasule, V. and P. Sonar. *Classification of brain MRI using SVM and KNN classifier*. in *Third International Conference on Sensing, Signal Processing and Security (ICSSS)*. 2017. IEEE.

50.   Greener, J.G., et al., *A guide to machine learning for biologists.* Nature Reviews Molecular Cell Biology, 2022. **23**(1): p. 40-55.

51.   Houssein, *Dental age estimation based on X-ray images.* Computers, Materials & Continua, 2020. **62**(2): p. 591-605.

52.   Hemalatha, *A versatile approach for dental age estimation using fuzzy neural network with teaching learning-based optimization classification.* Multimedia Tools and Applications, 2020. **79**(5): p. 3645-3665.