

Enhancing CNN for Forensics Age Estimation Using CGAN and Pseudo-Labeling

Sultan Alkaabi^{1,*}, Salman Yussof¹ and Sameera Al-Mulla²

¹Institute of Informatics and Computing in Energy, Universiti Tenaga Nasional, Malaysia

²College of Engineering and IT, University of Dubai, UAE

*Corresponding Author: Sultan Alkaabi. Email: PT20740@utn.edu.my

Received: 15 March 2022; Accepted: 13 June 2022

Abstract: Age estimation using forensics odontology is an important process in identifying victims in criminal or mass disaster cases. Traditionally, this process is done manually by human expert. However, the speed and accuracy may vary depending on the expertise level of the human expert and other human factors such as level of fatigue and attentiveness. To improve the recognition speed and consistency, researchers have proposed automated age estimation using deep learning techniques such as Convolutional Neural Network (CNN). CNN requires many training images to obtain high percentage of recognition accuracy. Unfortunately, it is very difficult to get large number of samples of dental images for training the CNN due to the need to comply to privacy acts. A promising solution to this problem is a technique called Generative Adversarial Network (GAN). GAN is a technique that can generate synthetic images that has similar statistics as the training set. A variation of GAN called Conditional GAN (CGAN) enables the generation of the synthetic images to be controlled more precisely such that only the specified type of images will be generated. This paper proposes a CGAN for generating new dental images to increase the number of images available for training a CNN model to perform age estimation. We also propose a pseudo-labelling technique to label the generated images with proper age and gender. We used the combination of real and generated images to train Dental Age and Sex Net (DASNET), which is a CNN model for dental age estimation. Based on the experiment conducted, the accuracy, coefficient of determination (R^2) and Absolute Error (AE) of DASNET have improved to 87%, 0.85 and 1.18 years respectively as opposed to 74%, 0.72 and 3.45 years when DASNET is trained using real, but smaller number of images.

Keywords: Dental forensics; age estimation; generative adversarial network; pseudo-labelling; convolutional neural network



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1 Introduction

Age is an essential factor in establishing the identity of an individual. Several techniques have been used in the literature for the problem of age estimation from images [1]. Estimation of the age of individuals is obtained considering many factors that may be looked at when it comes to the standing of teeth [2]. Thus, in this manner, dental factors such as teeth square measure the foremost sturdy and resilient elements of the skeleton assessed by radiography are an important source of information in forensic odontological age estimation and, with their physical variations, pathoses are considered very important which teeth are the least affected by the taphonomic process [3,4]. Their durability means that they are sometimes the only body part available for study [5,6]. Age estimation of orthopantomogram (OPG) images using the traditional techniques have not achieved accurate results as they cannot learn the complex and non-linear features of the dental images based on age and gender [7]. To address the issue, the authors proposed deep learning techniques especially convolutional neural network (CNN) for age estimation from dental x-ray images and outperformed state-of-the-art models. CNN performs feature extraction, classification and regression stages in an end-to-end manner and performs well to learn the features of the images [8]. In literature, Nolla, in the development of human dentition came up with a study that analyzes the growth of an individual tooth [8]. In the results, she shows the growth of the teeth takes place in 10 stages and a score is given at each stage then cumulatively the scores counted over the stages which finally decides the chronological age. In another paper, the researcher proposed a deep learning technique for estimation of age based on the images of teeth [9]. In their proposed model a set of 2676 dental x-ray images were used, each with the label of gender and age in the form of days. For example, image 1 with gender- female and age-(23 years 11 months 2 days). The authors proposed two architectures of the CNN where the first architecture DANET is used for age estimation and second architecture DASNET calculates the age estimation and incorporates the gender specific feature.

However, due to the issue of privacy of medical images [9], researchers have widely used the estimation of the medical image via ImageNet initialization even with various textures of natural medical images. A recent study found that such ImageNet-trained CNN is biased towards recognizing texture rather than shape and its performance degrades with having less samples. GAN is a data augmentation technique in which the generator and discriminator networks are used to generate more and accurate data samples [10]. Generative adversarial networks [11] helps in training generative models with neural networks. This helps overcome of data limitation that come about generative models and it adopts a stochastic training descent training method. The density model of a training data can be easily served by the generative model of a GAN. As the network allows some noise as inputs and outputs the new data as learned during training, this makes sampling easy and efficient. Conditional GAN (CGAN) is proposed by researchers in [12] serves as an extension of the Generative Adversarial Network allowing the model to apply a condition on data information. Hence, it works better when interacting with the generative model by providing external information. In another work which is considered as a benchmark paper. Using CGAN, synthetic images are generated but they have no labels and therefore cannot be used for training CNN. The concept of pseudo-labelling has been proposed for training with unlabeled or weakly labeled data. Pseudo-labelling is a Semi supervised learning method to involve feeding the model with both labelled and unlabeled data at the same time. In principle, this technique can combine almost all neural network models and training methods to overcome age regression problem as well as the gender classification problem [13].

The existing deep learning techniques proposed for the age estimation using dental images doesn't perform well due to the limited data samples. Thus, the objective of the research is to investigate the

role of GANs in the medical image data in terms of classification and regression and evaluating its performance without the pre-training.

The key contribution of the paper are as follows:

1. We propose CGAN for the data augmentation to overcome the data limitation.
2. We implement the pseudo-labelling to generate reliable training labels synthetic images that generated by CGAN
3. We build a new model using the combined original data, and synthetic data to generate reliable training labels, which results in a highly accurate and robust performance.

The paper is divided into 5 sections with Related work, being in Section 2. The proposed approach where we detail our CGAN and semi supervised learning in Section 3. Experimental studies and results are presented in Sections 4. Conclusions are given in Section 5.

2 Related Works

In this section, we discuss the relevant works on automated dental age estimation using CNN, the use of CGAN to generate new image samples and the use of pseudo-labelling to label the new images generated by CGAN.

2.1 CNN for Dental Age Estimation

The authors of [14] evaluated different CNN models for age estimation from dental OPG Images. Their dataset consists of around 2k images that are divided into seven different classes for network training. Their images are only labelled with age without gender. All models are evaluated with recall, F1-score and precision. At the end of the paper, they mentioned some of the issues that affected their results such as overlapping teeth, missing teeth and rotation of some of the images. These issues have caused a low accuracy result of not more than 40% [14].

The authors studied the age estimation of Malaysian children from age 1 to 17 years old. They proposed a deep convolutional neural network (DCNN) technique for the age estimation task. Feature extraction is done using the intensity projection method. This method provides high order features that are resilient to rotation and minor deformation changes. Their approaches were evaluated with 456 labelled OPGs. The result presented shows that the proposed method can perform automated age estimation with high level of accuracy [15]. An interesting study was proposed by Vila-Blanco et al. [16] with two different CNN approaches to perform age estimation from OPG dataset. The first is called DANET, which performs age estimation and the second one is called DASNET which incorporates the capability for gender identification in addition to age estimation. They trained 2289 OPG dataset which contains dental images of age 4.5 to 89.2. The images in the dataset are low quality images. Their results concluded that DASNET performs four months better than DANET in median Error. Their methods were evaluated using Absolute Error (AE), median Error (E) and coefficient of determination (R²). DASNET was compared to manual age estimation that has the under-estimation problem. Therefore, they concluded that the DASNET can perform automated age estimation with higher accuracy as compared to DANET, especially for young individuals with developing dentitions [16]. In another work, authors proposed a 3D watermarking technique based on the wavelet transform for protection of the medical data and simulation results show an improved robustness [17,18]. The authors have also utilized the concept of Convolutional long short-term memory (CLSTM) for the feature extraction of images and has show significant accuracy when the data samples are large in number [19,20].

2.2 Generative Adversarial Networks

Generative adversarial network (GAN) is a class of machine learning developed by Ian Goodfellow and his colleagues in 2014. The given a training set, GAN learns to propose generating data with same statistics in the training set [21]. As shown in Fig. 1, GAN trains two networks. The generator which inputs an unknown distribution and generates synthetic samples. The discriminators work is to distinguish whether the unknown distribution is synthetic or real.

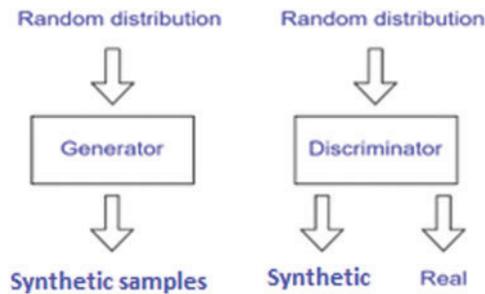


Figure 1: Generative adversarial network

The objective function of a GAN is:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim P_{data}(x)} [\log D(x)] + \mathbb{E}_{z \sim P_z(z)} [\log(1 - D(G(z)))] \quad (1)$$

The generator objective function used is shown below:

$$\mathcal{L}^G(\theta^{(G)}, \theta^{(D)}) = -\mathbb{E}_z [\log \mathcal{D}(G(z|y') | y')] \quad (2)$$

The loss function of the discriminator is as shown below.

$$\mathcal{L}^D(\theta^{(G)}, \theta^{(D)}) = -\mathbb{E}_{x \sim P_{data}(x)} \log \mathcal{D}(x|y) - \mathbb{E}_z \log(1 - \mathcal{D}(G(z|y'))) \quad (3)$$

GAN contains two networks: a generative network (G) and a discriminant network (D). Generative Adversarial Network (GAN) adopts a supervised learning approach to perform an unsupervised learning task by generating synthetic looking data. The discriminator later tries to determine if the generated sample is synthetic or real. GAN tries to understand the data, either its distribution or density estimation.

NVIDIA researchers were able to utilize the powerful functionalities of GAN. They generated realistic faces of people who do not even exist. Their system was so good such that it would generate small details such as pimples and skin pores. This was an improvement from 2017, where they tried the same functionalities using a neural network image generator. The results from the 2017 approach were not satisfying since the faces generated had some distortion [22]. The algorithm used by NVIDIA is StyleGAN. Generally, it can mimic any source. The researchers perform experiments using the same algorithm to generate other targets like graffiti. This shows the capabilities of GAN and how many big tech companies are adopting their functionalities.

ThisPersonDoesNotExist.com was created by an employee of Uber, Philip Wang. The website randomly generates synthetic faces using GAN. It was trained on lots of data and its results are impressive. Every time the refresh the webpage, a new face is generated. This will help researchers with more dataset especially if they want to do more research on the faces [22].

2.3 Conditional Generative Adversarial Networks

The Conditional Generative Adversarial Network is a variant of GAN where the generator generates a synthetic sample with a specific condition rather than a generic sample from unknown random noise vector distribution as shown in Fig. 2.

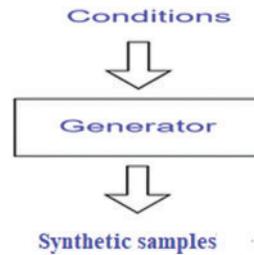


Figure 2: Conditional generative adversarial network

The objective function of a CGAN is:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim P_{data}(x)} [\log D(x)] + \mathbb{E}_{z \sim P_z(z)} [\log (1 - D(G(z), y))] \quad (4)$$

Conditional GAN [22,23] is a powerful class of generative models where the goal is to learn from input to output distributions conditioned on some auxiliary information. It consists of two networks, a generative network (G) and discriminant network (D). G takes input as x and the outputs to be generated the images $G(x, y)$ and y is class labels as conditional to receive the input and distinguish once the training generated image $G(x, y)$ and generative network (G).

Gauthier in his research proposed the first extension of GANs to a conditional setting [23]. In their work, by varying the conditional information in the GAN, they were able to generate faces with specific attributes from the generative model. They evaluated the chance in real face images to be implemented with generative model. CGAN has been used in [23] where experiments were conducted in which their model was able to generate realistic new images using color as the main attribute. They proposed a lightweight model which was small compared to the existing models which are larger in size and take time to train. Another contribution of their research work, which involves images of birds, is that they were able to generate the same bird with different colors and different backgrounds.

2.4 Pseudo-Labeling

Semi supervised algorithm acts as a bridge between the two main learning algorithms: supervised and unsupervised learning algorithm. It operates on data that is both labelled and unlabeled. The semi supervised learning uses both classification and clustering when it wants to label data and group data into specific clusters respectively. In a semi supervised algorithm, a model is trained partially using some of the labelled data. The resultant models then label the remaining unlabeled data. The labels that are attached to the previously unlabeled data are called pseudo-labelled data. Both the pseudo labels and the originally labelled data are now trained to create a unique model that can be used in prediction analysis. Pseudo-labelling is a concept which researchers are currently interested on. In the research paper [24], they implemented pseudo labelling to help in the semantic segmentation on their data. They used pseudo-labelling, a semi supervised method, together with deep neural networks for point cloud classification and segmentation. They were able to minimize the huge task that comes along with learning semantic segmentation tasks by implementing pseudo-labelling. This worked well considering they were in possession of a small number of labelled points [25]. Considering a complex relationship

between unlabeled and labelled training samples [26], a semi-supervised approach was used to re-identify a person with less data. A generative model via adversarial training approximates the actual data manifold. The adopted approach tries to distinguish the relationship between the unlabeled data and the already labelled data sample using intermediate feature representations in deep neural network [27]. In recent research [28,29] on person re-identification, an approach of multi-pseudo labels on the generated data was used to reduce the risk of overfitting. The existing techniques for the age estimation using the deep learning techniques such as CNN fails to perform well due to limited datasets and as a result a research gap exists to investigate the data augmentation techniques that can address the limited data samples issue and perform well.

3 Proposed Techniques

The main objective of this paper is to develop a method for generating new training data, together with their label, that can be used to train the dental age estimation CNN model. CGAN helps to overcome the challenges on the lack of training data. The original image dataset is labelled with age and gender. To use CGAN, the age label is important. The architecture of CGAN contains two models: a generator and a discriminator. The CGAN generator learns from the data presented in training and generates new data which is statistically or almost like the real data.

In our proposed method, the input to the generator is random dental x-ray images together with their age label. The age label acts as the condition to the GAN, hence making it CGAN. This condition helps it to generate new images with age consideration as opposed to a normal GAN which would generate images randomly. The discriminator receives the generated images that have been generated by the generator. As highlighted earlier, the generator produces new images, and in some cases may regenerate the same real images as the one presented during training. In GAN, the discriminator attempts to identify synthetic images from real images so that it would help the generator to produce images that look as real as possible.

Even though, data generated by the CGAN is unlabeled, labelled examples can be used to improve the learned features representation in a semi supervised fashion. In this section, the implementation of the CGAN and the semi supervised learning implementation will be discussed as shown in Fig. 3.

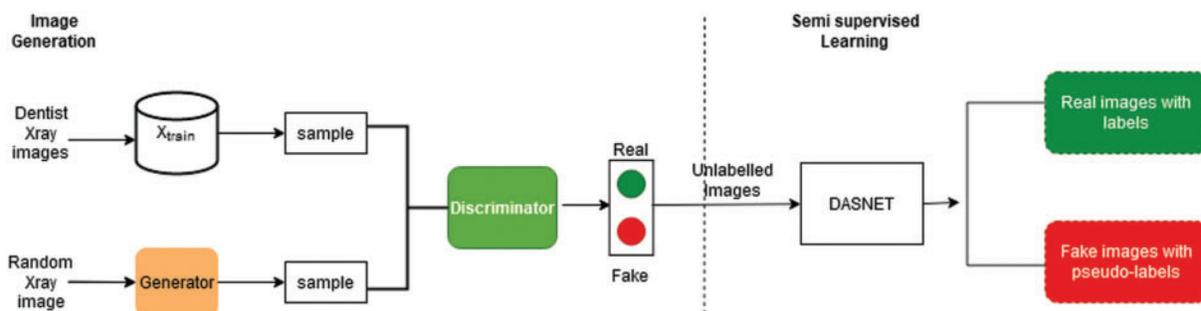


Figure 3: Overall architecture using CGAN and Pseudo labeling

The image generation was implemented by a CGAN. For semi-supervised learning, we had to train the labelled images with DASNET, then train it with our unlabeled images by pseudo labels to label the unlabeled images.

3.1 CGAN

In total, 2676 images from original dataset S_{all} and divided to 4 age groups: ($0 < \text{year} < 20$, < 20 years, 30 years, < 40 years). With age groups being a categorical feature, we applied one-hot encoding for it to be a one hot numeric array. The general architecture of our CGAN is depicted in Fig. 4.

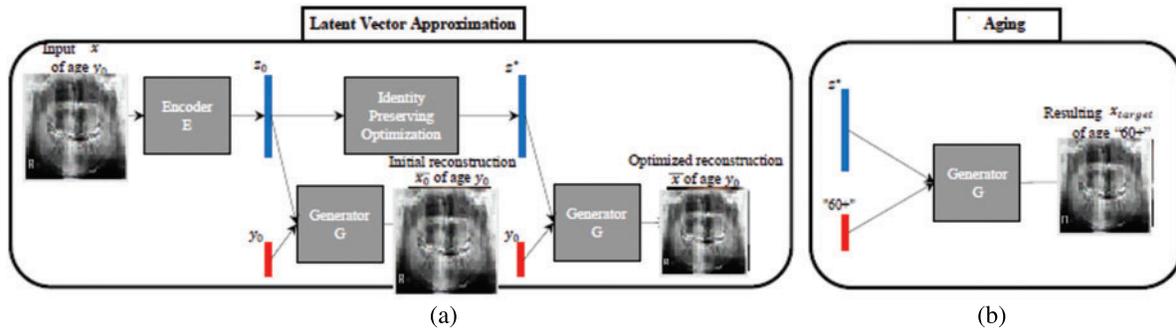


Figure 4: Details of CGAN model for age estimation

The proposed CGAN consists of two main parts: The generator Network and The Discriminator Network.

3.2 Generator Network

The primary objective is to generate an image having the dimension (128, 128, and 3). It takes a 100-dimensional latent vector. The architecture is a deep CNN made up of dense, convolution layers and activation layers. It takes 2 input values—a noise vector and a conditioning value. The conditional value provided to the network is the age group as shown in Fig. 4. On the other hand, the generator, a sequential model, consists of five hidden layers. Each layer consisted of Convolution layers and a Leaky RELU activation function. The network up samples the images of a dimension of 4×4 to 128×128 . Fig. 5 shows the detail architecture of the CGAN in which we have real dental images forwarded to the discriminator and second is the generator with synthetic random images. The Conv2D architecture based on multiple hidden layers in the generator is used to train the model by using the discriminator such that the generator generates images according to real image.

3.3 Discriminator Network

The discriminator identifies whether the provided image is synthetic or real. It passes the image through a series of classification layers and activation functions. The discriminator consists of five hidden layers. Each layer consists of Convolution layers and a Leaky RELU activation function. The network down samples the images received from the generator from 128×128 to 4×4 . The visualization of the generator and discriminator can be seen in Fig. 6.

3.4 Pseudo Labelling

As explained in section II, pseudo-labeling takes a semi-supervised learning training approach. The CGAN generated 1000 new images that were unlabeled and needed pseudo-labels to be attached.

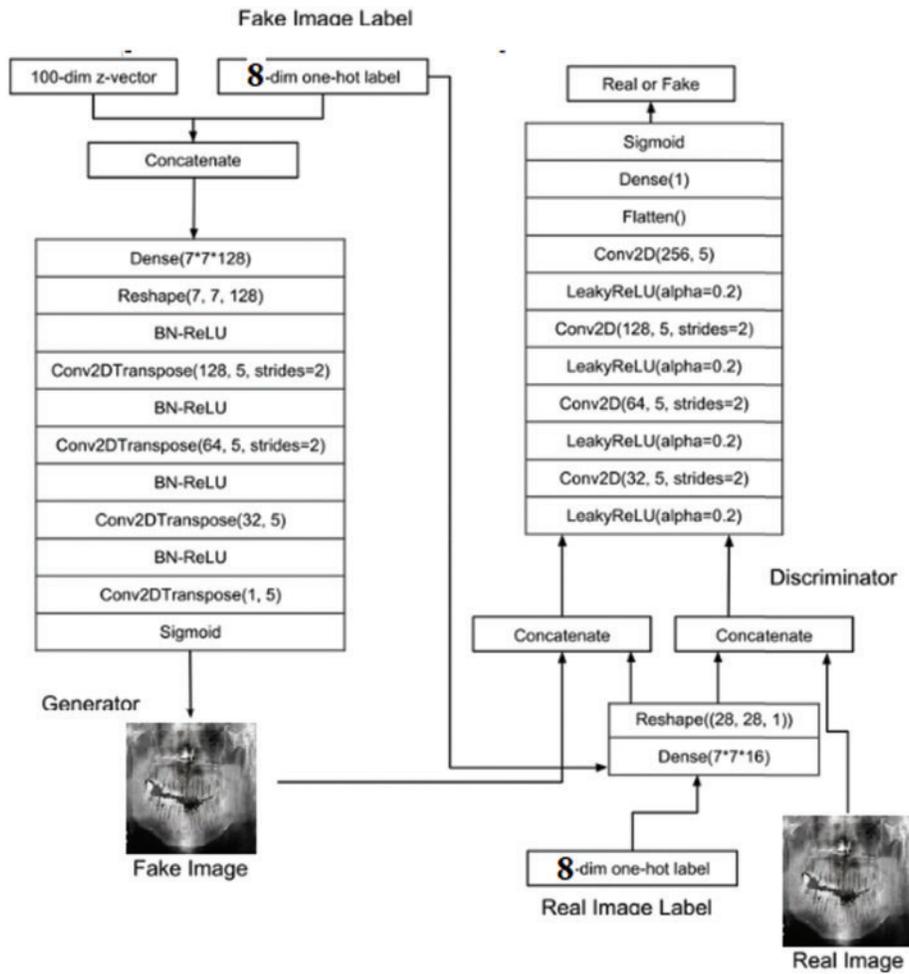


Figure 5: Detailed model of our CGAN

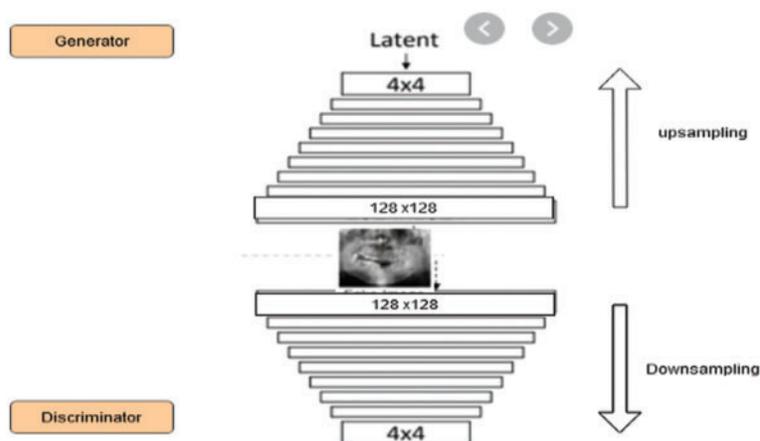


Figure 6: Generator and discriminator models

A more formalized description of pseudo-labelling is given below in Algorithm 1.

Input variable \mathbf{x} , output variable \mathbf{y}

A learning function $f: \mathbf{X} \rightarrow \mathbf{y}$

A combination of labelled input and the corresponding output will be represented as $(X_l, Y_l) = \{(x_{1:l}, y_{1:l})\}$

The unlabeled input will be represented as input $(X_u) = \{x_{l+1:l+u}, y_{1:l}\}$, without a corresponding output. $\{(x_{n+1...}, y_{n+1...})\}$ which is test data that has been reserved to test the generalization of the algorithm

We used the self-training algorithm of semi-supervised learning to train the model. The algorithm is as follows:

Input: Labelled data $\{(x_i, y_i)\}_{i=1}^l$, unlabeled data $\{(x_j)\}_{j=l+1}^{l+u}$

Initially, let $L = \{(x_i, y_i)\}_{i=1}^l$ and $U = \{(x_j)\}_{j=l+1}^{l+u}$

While True:

Train the learning function f from L in a supervised learning approach.

Apply the learning function f to the unlabelled instances in U

Separate subset S from U ; and compute $\{(x, f(x) | x \in S)\}$ to L

The main idea about pseudo labeling can be seen in Fig. 7. The process that we used to apply pseudo-labeling is as follows:

Step 1: We trained the DASNET model with the original labelled 2676 images. The labels are age and gender. We later saved the model.

Step 2: Using CGAN we generated 1000 images which are unlabeled, they have no age or gender labels attached to them.

Step 3: We used the pre-trained model saved in step 1 to predict the age and gender labels of the unlabeled images in step 2. The predicted labels are called pseudo-labels.

Step 4: Now that all our images (2676 original images + 1000 generated images) are labelled, we then retrain the DASNET model in step 1 with the labelled and pseudo labelled data together to generate a more accurate model.

4 Experiment

The experiment was conducted on a personal computer with Intel Core i7-4600U and 12 GB RAM running Linux Mint version 20. Python programming language was used with the TensorFlow framework to implement the proposed techniques.

4.1 Data Collection

For the experiment, a set of 2676 orthopantomogram images from Crescent Dental Maxillofacial and Orthodontic services [25] in Bangladesh were used. The age given in the image is calculated by subtracting the date of the image capture and the date of birth of the subject. To capture these images, dental implant CBCT T1 device is used to be computed tomography x-ray to produce panoramic images from different angles. In our experiment, we used 70% of the dataset as training set and the remaining as test set.

4.2 Training CGAN

We used the mini batch stochastic gradient descent training of GANs. This technique differs from the gradient descent algorithm in the sense that it splits the training data into small batches to calculate

the model errors and update the model coefficients. The algorithm of our training is detailed below, where k is a hyper parameter that states the number of steps to apply to our discriminator. After training, we generated 1000 images. A comparison between the real samples and the generated samples can be seen from Figs. 8 and 9.

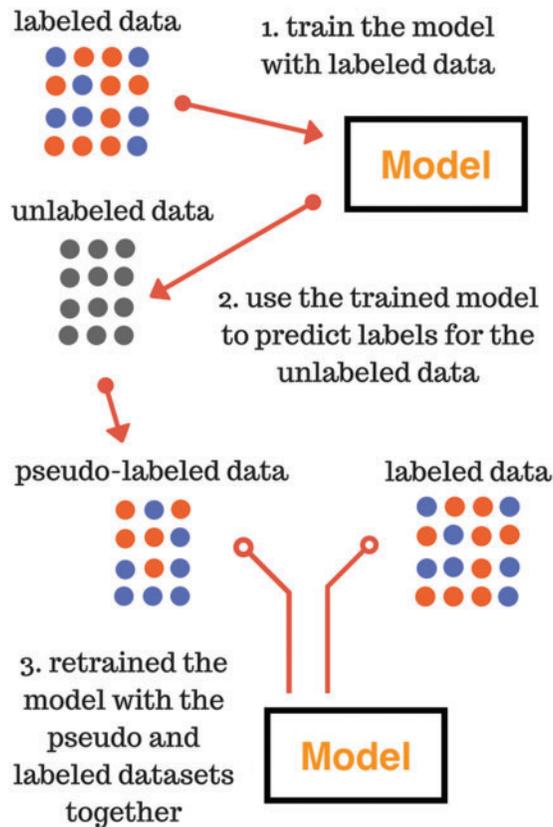


Figure 7: Pseudo-labelling technique (The dots with various colors represent different labels)

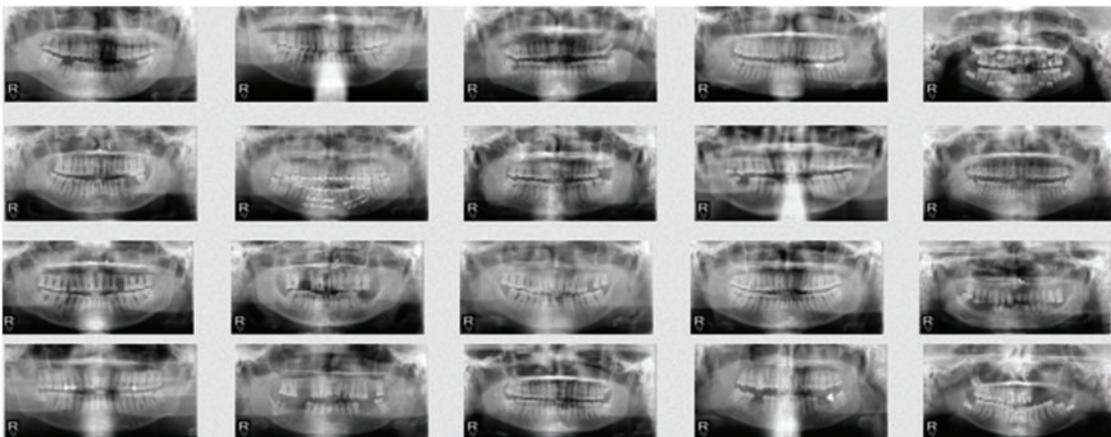


Figure 8: Sample of real images

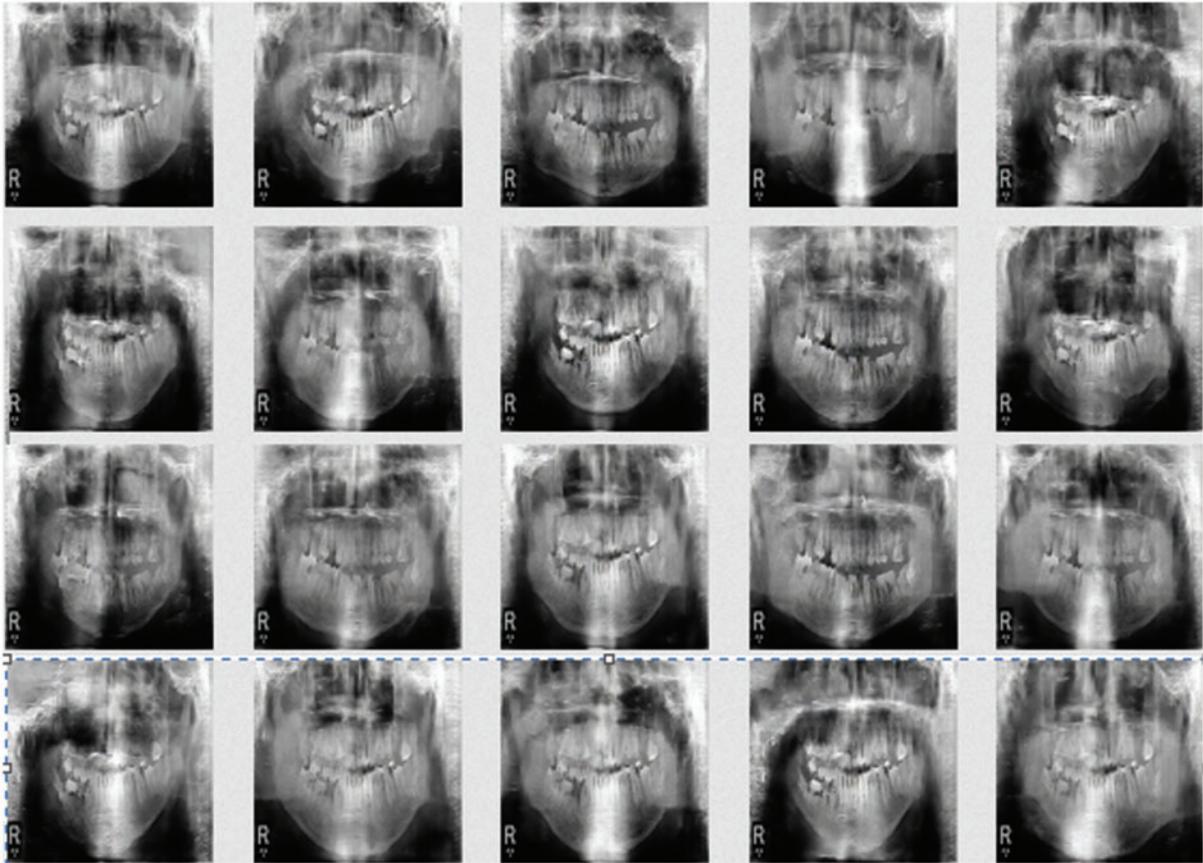


Figure 9: Sample of generated images after 500 epochs

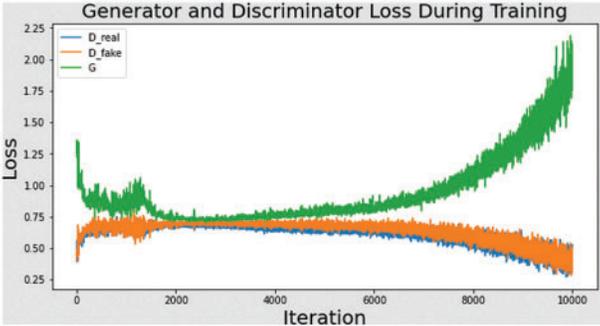


Figure 10: The discriminator and generator

5 Performance Metrics

Since the problem at hand requires regression analysis as well as classification analysis, the performance evaluation was done using several metrics.

5.1 Mean Absolute Error (MAE)

We took the average over the absolute residual error of the ages. The MAE is the average of absolute difference between the predicted age by the DASNET model and the real age.

$$MAE = \frac{1}{n} \sum_{j=1}^n |y_j - \hat{y}_j| \quad (5)$$

5.2 Root Mean Squared Error (RMSE)

We took the square root of the average of squared residual errors of the ages. The residual error is the difference between the predicted age by the DASNET model and the real age.

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2} \quad (6)$$

5.3 Classification Accuracy

The classification accuracy on test set is the percentage that of test set tuples to be correctly classified with the classifier.

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

5.4 True Positive Rate (Sensitivity)

This is calculated as the ratio of the true positive values to the number of actual positives. It is also called recall.

$$\text{Sensitivity} = \frac{\text{TruePositive}}{\text{FalseNegative} + \text{TruePositive}} \quad (8)$$

5.5 True Negative Rate (Specificity)

This is calculated as the ratio of the true negative values to the number of actual negatives.

$$\text{Specificity} = \frac{\text{TrueNegative}}{\text{TrueNegative} + \text{FalsePositive}} \quad (9)$$

6 Results and Discussions

The CGAN generated 1000 more images. Its performance was measured by logarithmic loss, as shown in Fig. 10 that shows the behavior of the discriminator loss and generator loss for the CGAN during the training. The prediction is done on the test data that was not available to the training algorithm. The algorithm does not learn much till around 3000 iterations where the adversarial technique commences. The discriminator loss decreases towards a constraint of $\ln\left(\frac{1}{2}\right)$, while the generator loss improves gradually.

In Fig. 11 the model converged after 500 epochs and was able to generate clear and acceptable images compared to the original images. Below is a sample of the new generated images after 500 epochs. After 500 epochs of training cycle, we were able to generate 1000 more images. Their distribution in terms of age group can be seen in Fig. 12 below.

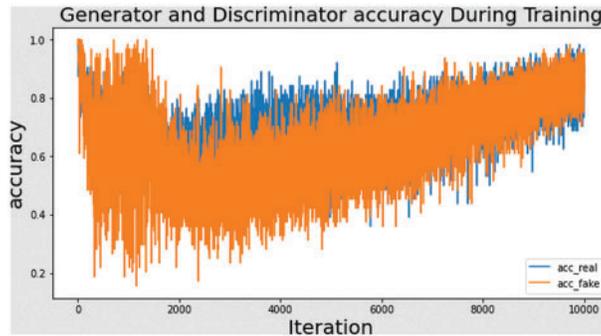


Figure 11: Accuracy during CGAN training

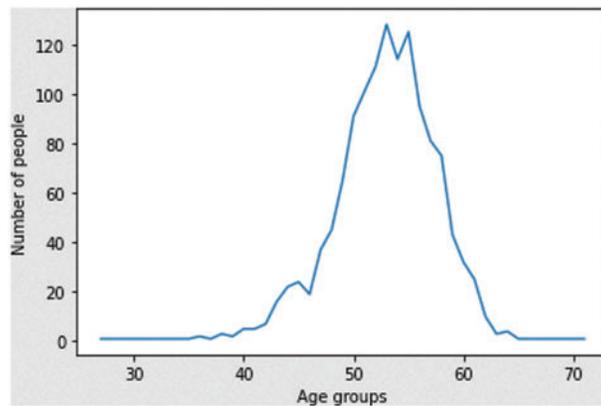


Figure 12: Generated images age group distribution count

Our experiments on the comparative studies between our research model when generated 1000 sample images as shown in Fig. 13 and DASNET [8]. Given the same advantage of training and resources, our research model had a better correlation with $R^2 = 0.85$ as compared to $R^2 = 0.72$ as shown in Figs. 14 and 15. In order to deduce conclusions between the proposed technique and DASNET, all resources and training factors were kept same. The results of training the DASNET as it is, without changing anything, produce the results as shown in Tab. 1.

The result of the experiment shows that the proposed technique of using CGAN to generate new images for training produces better results and validates our research objective goal. The results are as follows in Tab. 2.

The results from Tabs. 1 and 2 were compared in each category. The age and gender prediction for the age group 40, 30, 25, 20 and 15 was analyzed. The gender classification in S_{all} improved greatly in all aspects: Acc (Accuracy), Sens (Sensitivity), Spec (Specificity) and AUC score. The accuracy improved by 0.13 from 0.74 to 0.87 compared to the benchmark. This is due to the fact that proposed CGAN is able to generate more images similar to the original and model improves the generalization issue. Hence, the accuracy is improved compared to the benchmark. On the other hand, the state-of-the-art scheme cannot generate more data images and cannot achieve higher accuracy. The trend is similar across all age groups.

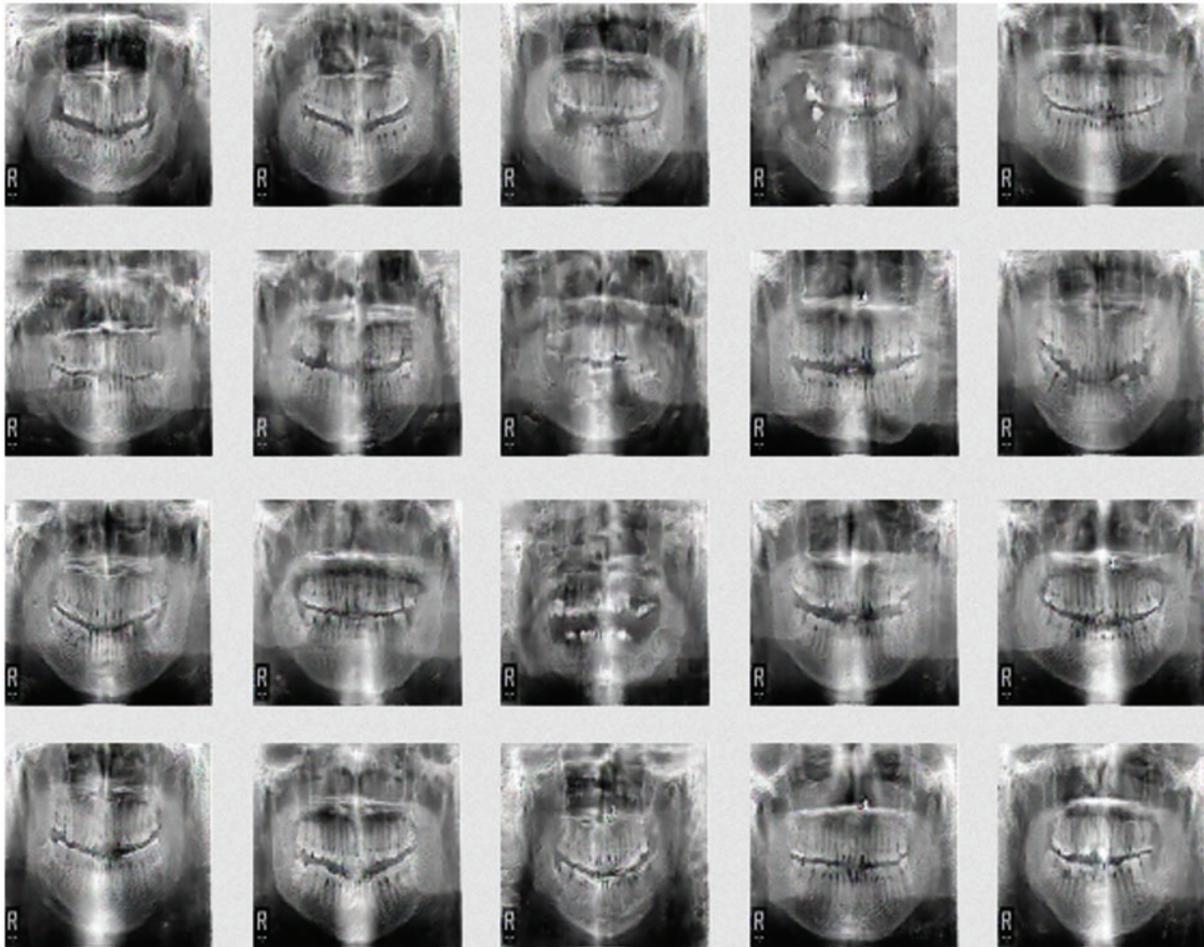


Figure 13: Sample of the generated images

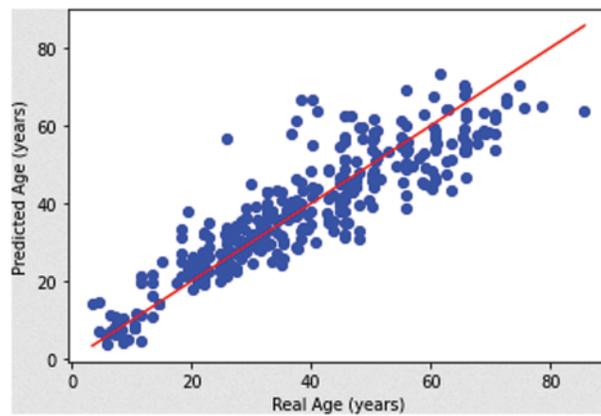


Figure 14: DASNET age regression plot

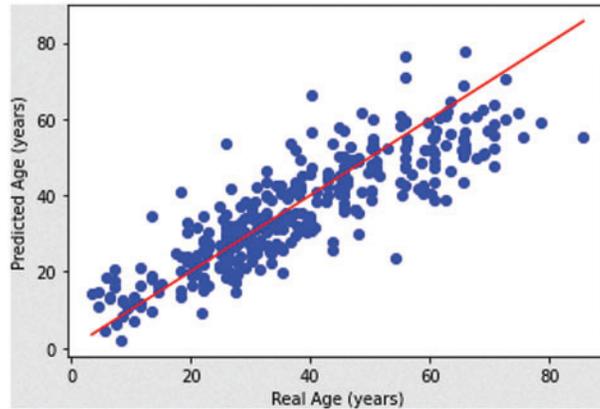


Figure 15: Proposed model age regression plot

Table 1: DASNET training statistics; IQR-inter quartile range

		CNN (DASNET) => 2676 images								
		Age regression					Gender classification			
Age gr.	#	Error Median	Absolute error		R ²	Accuracy	Sensitivity	Specificity	AUC score	
			$\mu \pm \sigma$	Median						IQR
S _{all}	2676	0.46	3.45 ± 5.45	1.75	4.06	0.72	0.74	0.76	0.73	0.72
S _{<40}	2317	0.24	2.23 ± 3.34	1.41	2.25	0.74	0.73	0.73	0.73	0.72
S _{<30}	1945	0.15	1.83 ± 1.96	1.24	1.87	0.69	0.64	0.74	0.72	0.71
S _{<25}	1810	0.12	1.48 ± 1.57	1.06	1.63	0.71	0.69	0.71	0.67	0.70
S _{<20}	1417	0.08	1.21 ± 1.34	0.75	1.14	0.72	0.71	0.69	0.69	0.69
S _{<15}	1078	0.06	1.03 ± 1.09	0.63	0.95	0.70	0.66	0.66	0.68	0.68

Table 2: Training statistics of research model; IQR-inter quartile range

		CGAN + CNN (DASNET) => 3676 images								
		Age regression					Gender classification			
Age gr.	#	Error Median	Absolute error		R ²	Accuracy	Sensitivity	Specificity	AUC score	
			$\mu \pm \sigma$	Median						IQR
S _{all}	3676	0.29	1.81 ± 4.5	1.12	3.41	0.85	0.87	0.89	0.83	0.88
S _{<40}	3304	0.11	1.44 ± 2.31	1.06	2.12	0.83	0.84	0.85	0.89	0.82
S _{<30}	2274	0.07	1.31 ± 1.24	0.94	1.74	0.84	0.85	0.80	0.83	0.83
S _{<25}	1756	0.01	1.12 ± 1.16	0.73	1.51	0.85	0.74	0.71	0.67	0.70
S _{<20}	1490	-0.09	0.85 ± 0.81	0.64	0.63	0.86	0.72	0.71	0.69	0.69
S _{<15}	1203	-0.16	0.67 ± 0.63	0.62	0.54	0.85	0.70	0.66	0.66	0.68

We also calculated the **E (median)** by taking the residual error and computing its median. It is worth noting that in the whole dataset, the research model improved by 0.17 which translates to 2 months. The absolute error (AE), which is the absolute difference of the measured value and the true value, was summarized with the mean and standard deviation $AE(\mu \pm \sigma)$. The Mean Absolute Error and the Standard Deviation Absolute Error in the research model performs better than DASNET by 1.64 and 0.95 respectively. This translates to 19 months and 12 months respectively. The coefficient of determination (R^2) as shown in the equation below, was calculated which is mathematically written as, $R^2 = 1 - \frac{\sum \text{of residual errors}}{\sum \text{of total errors}}$, the proposed model score improves from that of [8] in almost all categories. Finally, we calculated the area under the curve (AUC) for testing the robustness of the classification methods. It was used for gender classification, with S_{all} DASNET and proposed CGAN model improves by 0.16 compare to the benchmark scheme.

7 Conclusion

Dental age estimation is significant for determining the actual age of an individual such as criminal responsibility, detecting legal and other social events. The use of deep learning techniques such as CNN have been proposed by other researchers. However, it is difficult to obtain sample dental images due to privacy issue. As a result, the accurate age estimation using the dental images is not very good because CNN does not work well with small dataset. Hence, this paper proposed the use of Conditional GAN to generate new images to increase the number of images available to train CNN. This paper also proposed the use of pseudo-labeling to automatically label the images produced by CGAN with the proper age and gender. Based on the experiment conducted, our proposed technique outperformed the DASNET and achieved an accuracy of 0.87, an improvement of 0.17 and 0.16 in terms of E-median and AUC respectively compared to the benchmark scheme. In future work, for incorporating more images from a certain group or almost the same number of images in all age groups, a Balancing GAN (BAGAN) will be investigated and will increase the data samples of the underrepresented classes. Moreover, the transfer learning and few shot learning techniques will also be explored to increase the generalization capability of the model.

Acknowledgement: This work has been partially supported by Institute of Informatics and Computing in Energy, Universiti Tenaga Nasional. We gratefully acknowledge the support of Crescent dental maxillofacial and orthodontic services from Bangladesh for providing the dataset.

Funding Statement: The publication of this paper is funded by the UNITEN BOLD Refresh publication fund.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] M. Farhadian, F. Salemi, S. Saati and N. Nafisi, "Age estimation by using dental radiographs," *Imaging Science in Dentistry*, vol. 49, no. 1, pp. 19–26, 2019.
- [2] H. Alsaffar, W. Elshehawi, G. Roberts, V. Lucas, F. McDonald *et al.*, "Dental age estimation of children and adolescents: Validation of the Maltese reference data set," *Journal of Forensic and Legal Medicine*, vol. 45, pp. 29–31, 2017.

- [3] P. Limdiwala and J. Shah, "Precision and reliability of pulp/tooth area ratio (RA) of second molar as indicator of adult age," *Journal of Forensic Dental Sciences*, vol. 5, no. 2, pp. 118, 2013.
- [4] R. Scendoni, M. Cingolani, A. Giovagnoni, M. Fogante, P. Fedeli *et al.*, "Analysis of carpal bones on MR images for age estimation: First results of a new forensic approach," *Forensic Science International*, vol. 313, pp. 110341, 2020.
- [5] S. Kahaki, M. J. Nordin, N. S. Ahmad, M. Arzoky and W. Ismail, "Deep convolutional neural network designed for age assessment based on orthopantomography data," *Neural Computing and Applications*, vol. 32, no. 13, pp. 9357–9368, 2020.
- [6] P. Baumann, T. Widek, H. Merkens, J. Boldt, A. Petrovic *et al.*, "Dental age estimation of living persons: Comparison of MRI with OPG," *Forensic Science International*, vol. 253, pp. 76–80, 2015.
- [7] K. Reppien, B. Sejrnsen and N. Lynnerup, "Evaluation of post-mortem estimated dental age versus real age: A retrospective 21-year survey," *Forensic Science International*, vol. 159, pp. S84–S88, 2006.
- [8] I. Goodfellow, J. Pouget-Abadie, M. Mirza, X. Bing, D. Warde-Farley *et al.*, "Generative adversarial nets," in *Proc. Int. Conf. on Neural Information Processing Systems*, Montreal, Canada, pp. 1–9, 2014.
- [9] X. Yi, E. Walia and P. Babyn, "Generative adversarial network in medical imaging: A review," *Medical Image Analysis*, vol. 58, pp. 101552, 2019.
- [10] L. Guarnera, O. Giudice, C. Nastasi and S. Battiato, "Preliminary forensics analysis of DeepFake images," arXiv Preprint arXiv, pp. 12626, 2004.
- [11] M. D. Kohli, R. M. Summers and J. R. Geis, "Medical image data and datasets in the era of machine learning," whitepaper from the 2016 C-MIMI meeting dataset session. *Journal of Digital Imaging*, vol. 30, no. 4, pp. 392–399, 2017.
- [12] C. Han, L. Rundo, R. Araki, Y. Nagano, Y. Furukawa *et al.*, "Combining noise-to-image and image-to-image GANs: Brain MR image augmentation for tumor detection," *IEEE Access*, vol. 7, pp. 156966–156977, 2019.
- [13] R. Togashi, M. Otani and S. Satoh, "Alleviating cold-start problems in recommendation through pseudo-labelling over knowledge graph," in *Proc. of the 14th ACM Int. Conf. on Web Search and Data Mining in Virtual Event Israel*, pp. 931–939, 2021.
- [14] H. Zhang, "3D model generation on architectural plan and section training through machine learning," *Technologies*, vol. 7, no. 4, pp. 82, 2019.
- [15] H. Mansourifar and W. Shi "One-shot GAN generated fake face detection," arXiv Preprint arXiv:2003.12244, 2020.
- [16] N. Vila-Blanco, M. J. Carreira, P. Varas-Quintana, C. Balsa-Castro and I. Tomas, "Deep neural networks for chronological age estimation from opg images," *IEEE Transactions on Medical Imaging*, vol. 39, no. 7, pp. 2374–2384, 2020.
- [17] X. Zhang, W. Zhang, W. Sun, X. Sun and S. Kumar, "A robust 3-D medical watermarking based on wavelet transform for data protection," *Computer Systems Science and Engineering*, vol. 41, no. 3, pp. 1043–1056, 2022.
- [18] X. Zhang, W. Zhang, W. Sun, X. Sun and S. Kumar, "Robust reversible audio watermarking scheme for telemedicine and privacy protection," *Computers, Materials & Continua*, vol. 71, no. 2, pp. 3035–3050, 2022.
- [19] S. Kwon, "CLSTM: Deep feature-based speech emotion recognition using the hierarchical ConvLSTM network," *Mathematics*, vol. 8, no. 12, pp. 2133, 2020.
- [20] M. Mustaqem and S. Kwon, "1D-CNN: Speech emotion recognition system using a stacked network with dilated CNN features," *Computers, Materials & Continua*, vol. 67, no. 3, pp. 4039–4059, 2021.
- [21] M. Mirza and S. Osindero, "Conditional generative adversarial nets," arXiv:pp.1411.1784, 2014.
- [22] L. Guarnera, O. Giudice, C. Nastasi and S. Battiato, "Preliminary forensics analysis of deepfake images," in *2020 AEIT Int. Annual Conf. (AEIT)*, pp. 1–6, 2020.
- [23] J. Gauthier, "Conditional generative adversarial nets for convolutional face generation," *Class Project for Stanford CS231N: Convolutional Neural Networks for Visual Recognition, Winter Semester*, vol. 2, no. 5, pp. 1–9, 2014.

- [24] Y. Yasuhiro, K. Xu, K. Murasaki, S. Ando and A. Sagata, "Pseudo-labelling-aided semantic segmentation on sparsely annotated 3D point clouds," *IP SJ Transactions on Computer Vision and Applications*, vol. 12, no. 1, pp. 1–13, 2020.
- [25] X. Katie, Y. Yasuhiro, K. Murasaki, S. Ando and A. Sagata, "Semantic segmentation of sparsely annotated 3D point clouds by pseudo-labelling," in *2019 Int. Conf. on 3D Vision (3DV)*, Canada, IEEE, pp. 463–471, 2019.
- [26] H. Shim, S. Luca, D. Lowet and B. Vanrumste, "Data augmentation and semi-supervised learning for deep neural networks-based text classifier," in *Proc. of the 35th Annual ACM Symp. on Applied Computing*, Brno Czech Republic, pp. 1119–1126, 2020.
- [27] Y. Yao, K. Xu, K. Murasaki, S. Ando and A. Sagata, "Pseudo-labelling-aided semantic segmentation on sparsely annotated 3D point clouds," *IP SJ Transactions on Computer Vision and Applications*, vol. 12, no. 1, pp. 1–13, 2020.
- [28] L. Huang, Q. Yang, J. Wu, Y. Huang, Q. Wu *et al.*, "Generated data with sparse regularized multi-pseudo label for person re-identification," *IEEE Signal Processing Letters*, vol. 27, pp. 391–395, 2020.
- [29] "Crescent dental," 2020. [Online]. Available: <http://crescentdental.com.bd/>. [Accessed 02 01 2022].