

Dataset Augmentation for Machine Learning Applications of Dental Radiography

Shahid Khan¹, Altaf Mukati²

Shaheed Zulfikar Ali Bhutto Institute of Science and Technology
Karachi, Pakistan

Abstract—The performance of any machine learning algorithm heavily depends on the quality and quantity of the training data. Machine learning algorithms, driven by training data can accurately predict and produce the right outcome when trained through enough amount of quality data. In the medical applications, being more critical, the accuracy is of utmost importance. Obtaining medical imaging data, enough to train machine learning algorithm is difficult due to a variety of reasons. An effort has been made to produce an augmented dental radiography dataset to train machine learning algorithms. 116 panoramic dental radiographs have been manually segmented for each tooth producing 32 classes of teeth. Out of 3712 images of individual tooth, 2910 were used for machine learning through general augmentation methods that include rotation, intensity transformation and flipping of the images, creating a massive dataset of 5.12 million unique images. The dataset is labeled and classified into 32 classes. This dataset can be used to train deep convolutional neural networks to perform classification and segmentation of teeth in x-rays, Cone-Beam CT scans and other radiographs. We retrained AlexNet on a subset of 80,000 images of the entire dataset and obtained classification accuracy of 98.88% on 10 classes. The retraining on original dataset yielded 88.31%. The result is evident of nearly a 10% increase in the performance of the classifier trained on the augmented dataset. The training and validation datasets include teeth affected with metal objects. The manually segmented dataset can be used as a benchmark to evaluate the performance of machine learning algorithms for performing tooth segmentation and tooth classification.

Keywords—Data augmentation; Cone-Beam Computed Tomography; dental X-Rays; panoramic; dataset; classification; deep convolutional neural network; benchmark

I. INTRODUCTION

The machine learning, especially deep convolutional neural network has been playing a key role in the advancements in the medical imaging field [1]. Today, more than ever medical imaging applications are powered by artificial intelligence. Medical practitioners use computer-based systems for automatic diagnosis, analysis, planning and simulation of diseases and treatment planning [1].

As the name indicates, data-driven methods depend on the training data [2]. Popular image classification methods [3] [4] [5] are trained on millions of images to be able to produce necessary outcome. The large datasets used to train these models are produced manually by teams which consist of many human resources. Some datasets with millions of images are produced by crowd sourcing. Enormous resources and

man hours are spent to produce such datasets. The accuracy of applications of convolutional neural networks is more critical in medical applications. The accuracy is directly related to the amount and quality of the training data. Generally, the more training data we provide, the better results we get. Unlike AlexNet [4] and Keras where most images are downloaded from the internet, it is hard to put together a quality training dataset consisting millions of dental images. In addition, the number of instances in each class in dental radiographs are naturally imbalanced. A normal subject has only 4 canines and 0 – 4 wisdom teeth as compared to 8 instances in each of the rest of classes. The absence of wisdom tooth is also common, the ratio of subjects with no wisdom tooth is even higher in the youth. To solve this issue of lack of training dataset and natural imbalance, data augmentation is proposed to generate synthetic radiography datasets for training deep convolutional neural networks to perform classification and segmentation in CBCT, X-Rays and panoramic radiographs.

We manually segmented 116 panoramic radiographs to obtain 2,910 individual teeth. Fig. 1 shows an instance from the source panoramic dental radiographs. We applied common image augmentation technique such as rotation, resizing, flipping and intensity transformation to produce a synthetic dataset of total 5.12 million images containing 160,000 images of each of 32 teeth of humans. The dataset is labeled by directory names. We retained AlexNet [4] on a subset from this dataset containing 80,000 images and obtained 98.88% classification accuracy on 10 classes. The dataset we produced is not only a useful resource for training deep neural networks for tooth segmentation, classification and labeling but it can also be used as a benchmark for evaluating the performance of deep learning models.

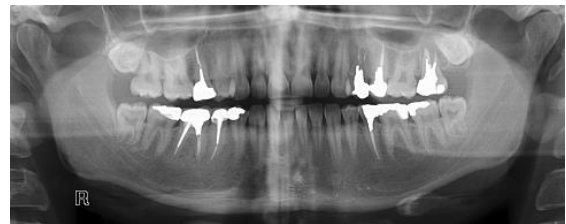


Fig. 1. Source Panoramic Dental Radiograph.

II. RELATED WORK

To maximize the generalization capability of the deep learning models, we require a lot of training data. State-of-the-art deep neural networks for example has millions of parameters which requires huge amount of data to achieve

good results. In the field of image classification, researchers have been using a variety of dataset augmentation techniques to supplement the training and validation data. The augmentation of adding random noise in the original data points has been used in [2] to automatically generate augmented datasets. Classifiers trained on this generated dataset outperformed the classifier trained on the original network. In another example of dataset augmentation [6] in the field of medical image processing application of deep learning where a limited dataset of only 182 liver lesion cases is enlarged first using classical augmentation techniques and further enlarged using Generative Adversarial Networks (GAN). The reported that augmentation through GAN improved the classifier performance by nearly 10%. The classical as well as GAN based image dataset augmentation techniques helped in improving the image classification in underwater images. The scientists in the research [7] reported that they improved the classification confidence in the submarine and sonar image classification through classical and GANs based techniques. Image data augmentation also helps in improving classification confidence in classifying images from 3D volumetric radiographs. A novel approach [8] of image data augmentation is proposed in this research where low-resolution images are generated from high resolution volumetric CT scans. It is reported the significant improvement in classification has been achieved through their novel augmentation approach.

III. MATERIALS AND METHODOLOGY

The dataset used in [9] has been used as the starting point. This dataset contains 116 volumes of panoramic dental radiographs. We manually segmented all individual teeth present in the 116 volumes, of which 2910 individual teeth qualified for inclusion into our dataset. The 2910 teeth included the instances where there were external bodies such as implants and metallic fillings present. The dataset also included rotten and overlapping teeth.

Table I shows the distribution of 2910 images of individual teeth in 32 classes. It is evident that the ratio of wisdom teeth at number 1, 16, 17 and 32 is relatively low.

The distribution of the dataset in 10 classes is given in Table II.

Common image data augmentation techniques, as used in [10], such as rotation, resizing, intensity transformation, and flipping are applied in a sequence to produce exponential results.

A. Rotate

Each 400 x 400 image containing single tooth is first rotated to the right side one degree per iteration of a loop and the rotated image is captured as a unique image. The process is performed 10 times to produce 10 unique images. The original source image that has no rotation is used in the next step to perform the same operation in the opposite direction to produce 10 more unique images. It is considered to keep the rotation rate minimum to avoid the potential confusion between mandibular and maxillary teeth.

TABLE. I. ORGANIC/NATURAL DATA ITEM COUNT IN DATASET FOR EACH TOOTH TYPE

Tooth	1	2	3	4	5	6	7	8
Count	41	102	94	96	93	104	95	97
Tooth	9	10	11	12	13	14	15	16
Count	99	95	97	90	91	90	96	51
Tooth	17	18	19	20	21	22	23	24
Count	60	95	76	97	103	100	102	102
Tooth	25	26	27	28	29	30	31	32
Count	99	98	107	104	99	76	91	70

TABLE. II. DISTRIBUTION OF ORGANIC/NATURAL DATASET IN 10 CLASSES

Class	Count	Class	Count
Maxillary Incisors	386	Mandibular Incisors	401
Maxillary Canine	201	Mandibular Canine	208
Maxillary Premolar	370	Mandibular Premolar	403
Maxillary Molar	382	Mandibular Molar	338
Maxillary Wisdom	91	Mandibular Wisdom	130

B. Resize

The second manipulation is done by resizing. The resizing is performed on each of the images that we produced in the previous step. To perform the resizing, each 400 x 400 image is enlarged 1 pixel per iteration diagonally. The operation is performed on the original image in reverse order by producing additional 10 images by reducing the size 1px per iteration.

C. Intensity Transformation

The next manipulation is achieved by increasing and decreasing grey level values of the pixels. The input for this step is the collection of all images that we produced in the last step. Grey level augmentation is very significant because the radiography equipment manufacturers do not follow common standard. To produce augmented data of various contrasts, we increased grey level by a random number between 10 and 40, a total of ten times on each input image. We then performed the same process for another ten times by decreasing the grey level by a random number between 10 and 40.

D. Horizontal Flip

In the final step, we horizontally flipped all the images that we obtained in the last step. This only doubles the amount of data that we have obtained so far. We did not perform the vertical flip because it will create a confusion between maxillary and mandibular teeth.

Fig. 2 shows a set of organic teeth followed by instances of augmented images.

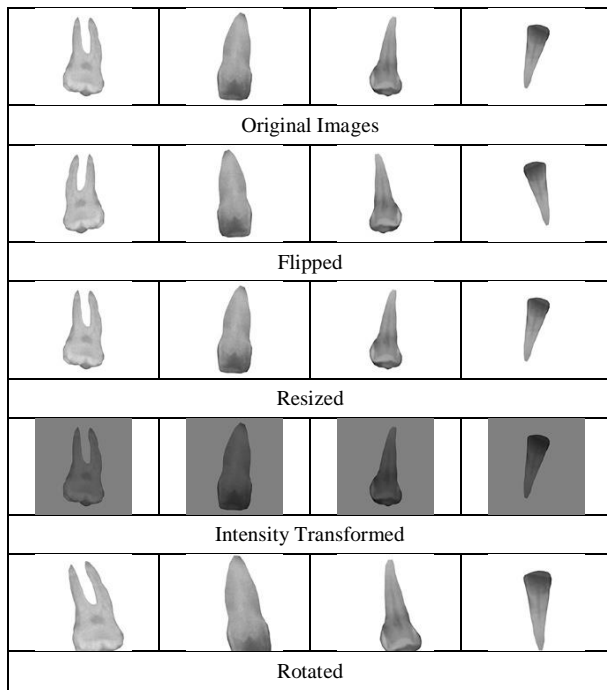


Fig. 2. Organic and Augmented Teeth Images.

IV. EXPERIMENT AND EVALUATION

We retrained AlexNet on the original dataset on 10 classes which yielded classification accuracy of 88.31%. Although the performance on original dataset is relatively low but for a very small dataset of only 2910 images and having imbalanced dataset, it is still favorable. The retraining of AlexNet with similar parameters and factors on the augmented and balanced dataset yielded 98.88% classification accuracy. The results are evident that the classifier trained on the augmented dataset outperformed the classifier trained on the original dataset with a considerable margin. In this experiment, we used a subset of 80,000 images from the dataset. Example images are shown in Fig. 2. We divided this subset into training and validation subsets by random sampling with ratio of 70% for training and 30% for validation. The training process converged after 1200 iterations. As training options, we used “Stochastic Gradient Descent with Momentum” at learning rate of 0.0001 and mini batch size of 4. Fig. 3 and 4 show the training progress.

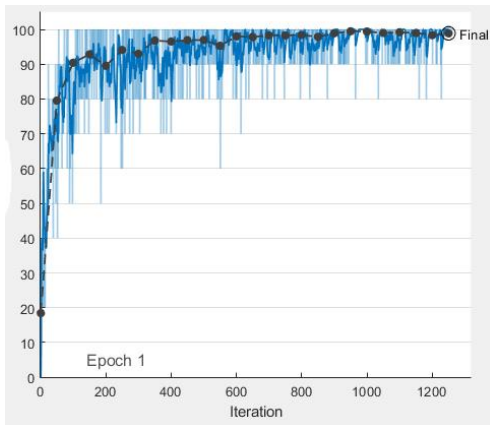


Fig. 3. Training Progress – Accuracy (%).

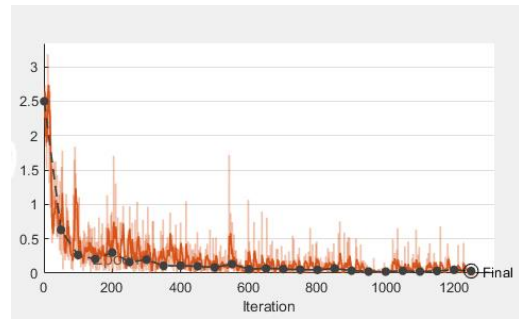


Fig. 4. Training Progress – Loss.

A copy of the pretrained neural network and a subset of the dataset can be downloaded from the website <https://shahidsci.com/dataset> for evaluation purpose.

V. LIMITATION

The source dataset in our work is from a single imaging facility. Although the outcome of source dataset is evident that several different radiography imaging machines have been used to produce this source dataset, but it is worthwhile to have a diverse dataset that is produced on a broader range of equipment. Further, the source radiography is conducted in single locality, which limits the diversity of the dataset.

VI. CONCLUSION

An effort has been made to solve the deficiency and imbalance of dental radiography data through proposed data augmentation method. We produced a massive dataset of 5.12 million images from dental radiographs. This dataset can be used to train ever-craving deep convolutional neural networks to perform segmentation and classification on CBCT dental images, x-rays and other dental radiographs. The dataset can be downloaded from the website mentioned in the previous section. We performed transfer learning from AlexNet on a subset of 80,000 images from the dataset and obtained accuracy of 98.88% on 10 classes which on nearly 10% more than the accuracy of the classifier trained on the original dataset.

VII. FUTURE WORK

In the future work, we plan to apply augmentation through Generative Adversarial Networks (GANs) alongside the classical augmentation techniques that we used. It is anticipated that it will further increase the classification confidence in the training of the classifiers. We also plan to diversify the dataset by adding more original X-Ray and CBCT radiography images from different parts of the world. We also plan to train a novel machine learning model on the produced dataset that can be used to classify teeth in radiography images from a verity of different sources.

REFERENCES

- [1] X. Xu, C. Liu, and Y. Zheng, “3D Tooth Segmentation and Labeling using Deep Convolutional Neural Networks,” *IEEE Trans. Vis. Comput. Graph.*, vol. XX, no. XX, pp. 1–13, 2018.
- [2] K. Fujita, “Data Augmentation using Evolutionary Image Processing,” *2018 Digit. Image Comput. Tech. Appl.*, pp. 1–6, 2018.
- [3] C. Szegedy et al., “Going Deeper with Convolutions,” in *Computer Vision and Pattern Recognition (CVPR)*, 2015.

- [4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.
- [5] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in *International Conference on Learning Representations*, 2015.
- [6] M. Frid-Adar, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "Synthetic data augmentation using GAN for improved liver lesion classification," *Proc. - Int. Symp. Biomed. Imaging*, vol. 2018-April, no. Isbi, pp. 289–293, 2018.
- [7] Y. Xu, Y. Zhang, H. Wang, and X. Liu, "Underwater image classification using deep convolutional neural networks and data augmentation," *2017 IEEE Int. Conf. Signal Process. Commun. Comput. ICSPCC 2017*, vol. 2017-January, pp. 1–5, 2017.
- [8] M. Farhadi and A. H. Foruzan, "Data Augmentation of CT Images of Liver Tumors to Reconstruct Super-Resolution Slices based on a Multi-Frame Approach," *ICEE 2019 - 27th Iran. Conf. Electr. Eng.*, pp. 1783–1786, 2019.
- [9] A. H. Abdi, "Automatic segmentation of mandible in panoramic x-ray."
- [10] Y. Miki et al., "Classification of teeth in cone-beam CT using deep convolutional neural network," *Comput. Biol. Med.*, vol. 80, pp. 24–29, 2017.