2020 IEEE REGION 10 CONFERENCE (TENCON)
Osaka, Japan, November 16-19, 2020

Anomaly detection in panoramic dental x-rays using a hybrid Deep Learning and Machine Learning approach

Dhruv Verma, Dr Sunaina Puri†, Dr Srikanth Prabhu‡ and Dr Komal Smriti§*

*Research and Development Engineer, 3PAR File Persona, Hewlett Packard Enterprise, Bangalore, Karnataka, India
†Former Dental Intern, Manipal College of Dental Sciences, Manipal Academy of Higher Education, Manipal, Karnataka, India
‡Associate Professor, Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka, India
§Assistant Professor, Department of Oral Medicine and Radiology, Manipal College of Dental Sciences, Manipal Academy of Higher Education, Manipal, Karnataka, India
Email: dhruv.verma110597@gmail.com, sunaina_puri2@yahoo.in, srikanth.prabhu@manipal.edu, komal.smriti@manipal.edu

***Abstract*—Automated anomaly detection in panoramic dental x-rays is a crucial step in streamlining post diagnosis treatment. It can reduce clinical time for a patient and also aid in giving them faster access to medical care. In this paper, we propose a hybrid deep learning and machine learning based approach to detect evident dental caries/periapical infection, altered periodontal bone height, and third molar impactions using panoramic dental radiographs. We use a Convolutional Neural Network as a feature extractor for an input image and use a Support Vector Machine to classify the image as either "Normal" or "Anomalous" based on the extracted features. We compare the performance of this model with the performance of a Convolutional Neural Network and a Support Vector Machine for the same classification task. We also compare our best model with other existing models trained to detect carries and periodontal bone loss. The results obtained with the hybrid deep learning and machine learning approach outperformed the existing methods in the literature.**

***Index Terms*—Panoramic dental x-rays, biomedical imaging, deep learning, machine learning, image processing**

I. Introduction

Advancements in the field of deep learning have shown the effectiveness of Convolutional Neural Networks (CNNs) in solving tasks such as image segmentation, image classification and object detection. They outperform traditional algorithms and are considered state-of-the-art. The primary reason for the success of CNNs can be attributed to their ability to learn the features from an input automatically [1]. Apart from CNNs, traditional machine learning algorithms are also well suited for classification tasks. In particular, support vector machines (SVMs) perform well when it comes to classification. This is because SVMs tend to avoid overfitting. Additionally, SVMs come with a kernel which allows a mapping from the original feature space to a higher dimensional feature space. This can increase classification accuracy if the classes are more easily separable in a higher dimension. The soft margin SVM optimization constraint for all m inputs x, all outputs y, the weight vector w and bias b is defined in (1). ζ is a slack variable introduced to allow soft margin optimization. The regularized optimization problem for this classifier is defined in (2). C is a regularization constant to prevent the misclassification of each training example.

$$y_{\mathrm{i}}(wx_{\mathrm{i}}+b) \geqslant 1-\zeta_{\mathrm{i}},\ i=1..m,\ \zeta_{\mathrm{i}}\ \geqslant\ 0 \tag{1}$$

$$\min_{\mathrm{w,\,b,}\,\zeta} \frac{1}{2}||w||^2 + C\sum_{i=0}^{m}\zeta_{\mathrm{i}}\ subject\ to\,(1) \tag{2}$$

Classification is an important task in medical images particularly for automatic detection of diseases, assisted diagnosis, and prognosis. Classifications not only play an important role in identifying any deviation from apparently normal physiology and functioning, but also facilitate communication between medical professionals by standardizing pathological data. This in turn saves both time and effort taken in arriving at a diagnosis, thereby magnifying the efficiency of and streamlining post diagnostic treatment planning. Automated classification of images into broad diagnostic categories can reduce the preoperative time that a patient spends at the hospital as compared to manual diagnosis. The level of human error is also minimized during this process, thereby enhancing the accuracy. However, automated classification of pathology is a rather challenging task particularly due to artifacts and imaging errors. Orthopantomography or panoramic imaging, provides us with a wide diagnostic visual field which aids a holistic diagnosis. There are certain challenges that may be encountered during this process, as panoramic images are likely to contain image overlaps, ghost images, and other artifactual information which could lead to errors in image processing. Other roadblocks that are common to all imaging modalities would include an inadequate contrast, digital graininess or noise, blurred boundaries and superimposition. [2]. In order to overcome some of these issues, we propose an algorithm using image processing techniques such as thresholding and morphological dilation and erosion. We use this to enhance the tooth boundaries and increase the contrast between the teeth and the rest of the image. This is used to pre-process the images in order to make the task of feature engineering more efficient.

While SVMs are good at classification tasks, they need hand-labelled input features which are then used to map an input to an output. In such a task where features are not hand-labelled and need to be extracted from an image, the task of feature engineering can be far more tedious. We explore feature extraction methods for images such as the Hu moments algorithm and Haralick texture extraction. Hu moments are a set of seven numbers calculated using central moments which are invariant to certain image transformations [3]. While the first six moments are invariant to scale, translation, reflection and rotation, the seventh moment changes in case of image reflection. Haralick measures are used to extract texture features from a Gray-Level Co-occurrence Matrix (GLCM) [4]. A GLCM P is a square matrix with dimension N_{g}, where N_{g} is the number of intensity levels in the image. Each element $P(i, j)$ is generated by counting the frequency of occurrence of a pixel with value i next to pixel with value j [5]. This helps characterize certain features of the image such as contrast, structure and roughness. These features were used to train a baseline SVM model.

In order to hasten the process of manually extracting features, we made use of CNNs to classify the images. As stated previously, CNNs can auto-learn features to map an input to an output. The stacked architecture of CNNs allows the shallow layers to learn local features and information, while the deeper layers are able to retain global features and information. CNNs made the task much more accurate.

Our final approach used a hybrid CNN-SVM architecture for the classification task. Using the features extracted by the CNN, we trained an SVM to classify images. Thereby combining the efficiency of a CNN in the task of feature extraction and the classification efficiency of an SVM. This model gave the highest accuracy on our dataset.

In this paper, using a CNN-SVM hybrid architecture we demonstrate an effective, accurate and robust method to detect evident dental caries/periapical infection, altered periodontal bone height, and third molar impactions. Our method achieved state-of-the-art results on our dataset and also effectively demonstrated the ability of CNNs to extract features which can further be used to train an SVM. We also compared our results to other models used to detect caries and altered periodonal bone height in panoramic dental x-rays.

II. Related work

In this paper, we present three different machine learning and deep learning approaches to classify panoramic dental x-rays. Additionally, we also worked on using computer vision techniques to enhance the images and extract features. Hence, we will primarily focus on techniques of classifying medical images and methods used to enhance medical images for better feature extraction.

A. Image enhancement and feature extraction

The primary objective of image pre-processing is to make it more suitable for human viewers or to enhance the image to aid algorithms in image analysis and object detection [6]. There are a few conventional methods that have been used in panoramic dental x-rays to define a region of interest, remove unwanted regions and to clarify the jaw and teeth.

One such method was proposed by Yusra Y. Amer et al. [7]. The authors primarily worked on tooth segmentation from panoramic dental x-rays. Their image processing steps began with contrast enhancement to help distinguish teeth region from other tissues by increasing tooth visibility. This was followed by applying Otsu's threshold to the image to remove any unwanted regions. The binary image was then morphologically dilated to connect the teeth with the jaw and also to smoothen any boundaries. The connected components were all segmented using Connected Component Labeling (CCL), which also helped in distinguishing the unwanted region from the region of interest. Finally, this processed image was multiplied with the original image to get the pre-processed image.

There has also been extensive research about feature extraction for medical image analysis and classification. One such method was proposed by Ehsani Rad et al. [8]. This method involved texture extraction from dental x-rays using a GLCM. The authors proposed this method for feature extraction as texture information of a tooth was an essential factor in analyzing its characteristics. The features measured include the contrast, correlation, entropy, homogeneity and energy. Another method for feature extraction was proposed by Pattanachai, Nakintorn et al. [3] . This paper proposes using Hu moments invariants as a shape descriptor for each tooth. Hu moment invariants have been extensively used as the measures for image recognition for their invariance to image translation, scaling and rotation. These have also been used for real-time eye detection using SVMs [9].

B. Classification approaches using machine learning and deep learning

Machine learning techniques such as SVMs and Naïve Bayes are commonly used in classification problems. One such automatic classification approach was proposed by Oliveira, João et al. [10]. The authors tested multiple approaches to detect caries in a panoramic dental x-ray. The process involved tooth segmentation, feature extraction and classification. Multiple classifiers were evaluated, these included artificial neural networks (ANNs), the K-Nearest-Neighbors (KNN) algorithm, Naïve Bayes and SVMs.

Deep learning classification approaches with CNNs have become popular [11]. However, these do not generalize well in case the dataset is not large enough. A deep learning approach for the detection of periodontal bone loss in panoramic dental x-rays was proposed by Krois, Joachim et al. [12]. In this paper, the authors proposed a CNN based classification. The CNN trained by the authors showed at least similar discrimination ability as dentists for assessing periodontal bone loss. This approach was further enhanced by Kim, Jaeyoung et al. [13] using a model known as DeNTNet to detect lesions and provide the corresponding tooth numbers as well. The DeNTNet model applied CNNs using transfer learning to overcome issues of morphological variation of lesions and the issue of an imbalanced dataset. Another such deep learning approach to detect and diagnose dental caries using CNNs was proposed by Jae-Hong Lee et al. [14].

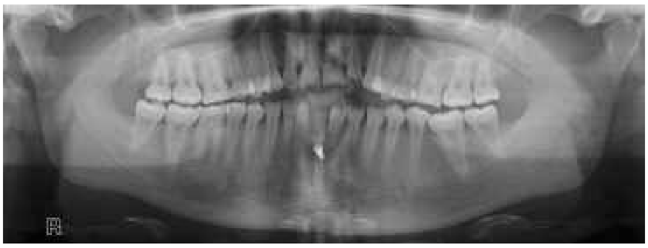

Fig. 1. A sample belonging to the "Normal" class

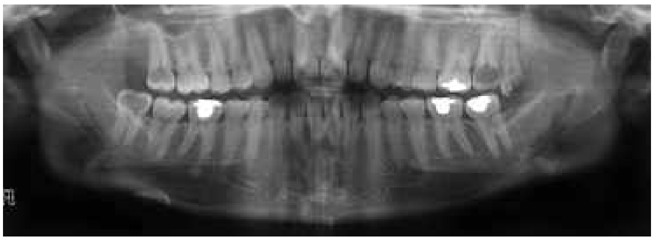

Fig. 2. A sample belonging to the "Anomalous" class

A hybrid machine learning and deep learning approach was proposed by X. Sun et al. [15]. Their research used a CNN-SVM hybrid architecture for the recognition of functional magnetic resonance images. Their proposed fusion achieved an accuracy of over 99.5% using the synergy effect of a CNN as a feature extractor and an SVM as a classifier. These promising results inspired us to try using a hybrid architecture for our classification task.

III. Methodology

In this section, we outline our dataset, image pre-processing algorithm and training procedure in detail.

A. Dataset

The dataset used by us consisted of 250 anonymized panoramic dental x-ray images provided by the Manipal College of Dental Sciences, Manipal and an open-source dataset consisting of 116 panoramic dental x-ray images [16]. The images were labelled by us before the training process. Two class labels were considered – "Normal" and "Anomalous". "Anomalous" in this context was a term extended to evident caries, periapical infections, pathological migration, altered bone height as well as bony impactions of third molars – all of which require clinical care over the course of time. Images suggesting the need of orthodontic intervention were excluded as a criterion as panoramic radiographs only provide a two dimensional view.

Fig. 1 shows a sample belonging to the "Normal" class. Fig. 2 shows a sample from our dataset belonging to the "Anomalous" class due to internal carries showing up in the upper molars due to radiolucency. Since the dataset was small and CNNs require a large number of training samples to converge successfully without overfitting; we make use of image generation and data-augmentation to increase the size of the training data artificially. The following augmentation steps were used:

- Horizontal flips
- Zoom up to 1.3 times of the image
- Random changes to the contrast of the image

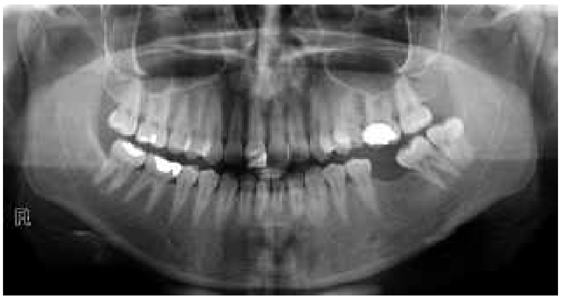

Fig. 3. Panoramic dental x-ray image

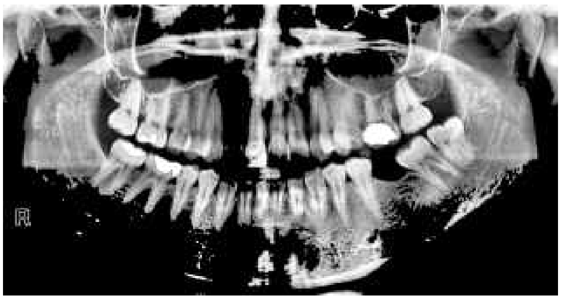

Fig. 4. Pre-processed panoramic dental x-ray image

- Random changes to the brightness of the image

These augmentation techniques were applied to generate two images for each image in the dataset. Therefore, the dataset size was increased to 1098 images. The training data consisted of 878 images, and the testing data consisted of 220 images. In addition to this, the training data was further split into a validation dataset of 87 images. The image-augmentation steps were also applied to an input image during run-time while training the CNN.

B. Image pre-processing

All the images in the dataset were pre-processed to enhance the contrast between the jawbones, spine and the teeth and also to sharpen tooth boundaries. This was done to improve image feature extraction for the SVM and by the CNN. The approach we followed was an iterative algorithm. The number of iterations was empirically obtained based on experimental results.

Each iteration involved carrying out a set of operations. The first step involved converting the image to grayscale and applying Contrast Limited Adaptive Histogram Equalization (CLAHE) to equalize the image. This step was followed by applying Otsu's threshold on the adaptive image. The binary image obtained was morphologically dilated to accentuate tooth features using a 7x7 kernel. The final step of the iteration involved masking the CLAHE image with the morphologically dilated image using a bitwise AND operation. The resulting image was divided by 255 to normalize the input. Fig. 3 shows a sample input image and Fig. 4 shows the same image after pre-processing.

C. Training

As part of the training process, three approaches were tried. The first involved training a baseline SVM. The second approach involved training a vanilla CNN on pre-processed images. The third involved using the trained CNN as a feature extractor. The extracted features were used to train an SVM.

1) SVM: An SVM is a supervised learning method used for problems of classification, regression and outlier detection. As part of the baseline model, an SVM was trained on the dataset. We tried to find the best hyperplane dividing the two classes by using a soft margin classifier. Feature extraction from the pre-processed images was an essential aspect of training this classifier. While extracting features, the key points considered included number of teeth, tooth shape, bone loss and the texture of the tooth. The features were extracted from the images using two methods:

- Hu moments: Hu moments invariants can be used as a shape descriptor in computer vision. This concept was used by computing the Hu moments for all the teeth in the image. These are used as input features to the SVM [3].
- Haralick texture: Haralick texture consists of fourteen features extracted from a GLCM [5]. μ_{x}, μ_{y}, σ_{x} and σ_{y} represent the means and standard deviations of the marginal probability matrix P_{x} (3) and P_{y} (4). We used four features out of the fourteen features - contrast (5), correlation-linear dependency (6), entropy (7) and angular second momentum (8). This was done as the other ten features did not have significant differences in their values between the two classes.

All features extracted for each image using the methods above were stacked together, each feature vector was normalized between 0 and 1 and then used as an input into an SVM classifier with a linear kernel and a square hinge loss (l2 loss). The best base model was obtained by applying grid search to the soft margin regularisation parameter.

$$P_{\mathrm{x}}(i) = \sum_{j=0}^{N_{\mathrm{g}}-1} P(i,j) \tag{3}$$

$$P_{\mathrm{y}}(j) = \sum_{i=0}^{N_{\mathrm{g}}-1} P(i,j) \tag{4}$$

$$f1 = \sum_{n=0}^{N_{\mathrm{g}}-1} n^2 \{\sum_{i=1}^{N_{\mathrm{g}}} \sum_{j=1}^{N_{\mathrm{g}}} P(i,j)\},\ where\, n = |i-j| \tag{5}$$

$$f2 = \sum_{i=0}^{N_{\mathrm{g}}-1} \sum_{j=0}^{N_{\mathrm{g}}-1} \frac{(i*j)*P(i,j) - \mu_{\mathrm{x}} * \mu_{\mathrm{y}}}{\sigma_{\mathrm{x}} * \sigma_{\mathrm{y}}} \tag{6}$$

$$f3 = -\sum_{i=0}^{N_{\mathrm{g}}-1} \sum_{j=0}^{N_{\mathrm{g}}-1} P(i,j) * \log(P(i,j)) \tag{7}$$

$$f4 = \sum_{i=0}^{N_{\mathrm{g}}-1} \sum_{j=0}^{N_{\mathrm{g}}-1} \{P(i,j)\}^2 \tag{8}$$

2) Vanilla CNN: To try and enhance the performance in comparison with an SVM a CNN was used as a classifier for the two classes. The model architecture we used for this task consists of multiple stacked layers, each comprising of a two-dimensional convolutional layer with a rectified linear unit (ReLU) activation and a two-dimensional max-pooling layer. Each stacked-layer had an increasing filter size in the convolutional layer while having a fixed kernel size over the input. Each max-pooling layer had a constant pool size throughout the network. The final output of these combinations of convolutional and max-pooling layers was flattened and fed forward into fully-connected layers. These fully-connected layers eventually condensed down to give the final output layer consisting of two units with a softmax activation function. Each output vector was converted to a one-hot encoded vector y of size 2, where $y_{\mathrm{i}} \,\epsilon\, \{0,1\}$. 1 for the “Anomalous” class and 0 for the “Normal” class. The softmax activation was chosen as a final activation due to the task being a binary classification problem where each class is mutually exclusive. Fig. 5 shows the model architecture.

The input data is given a mini-batch size of 5 images per batch due to the small size of the dataset. Each image in the mini-batch undergoes image augmentation as per the steps defined previously. Another aspect that was handled here was having higher augmentations for images belonging to the “Normal” class. Every image belonging to the “Normal” class in a mini-batch was augmented twice while an image belonging to the “Anomalous” class was augmented once. This was done to try and overcome the class imbalance present in the dataset. The Adam optimizer with a learning rate of 10^{-4} and a decay of 10^{-5} was used for the model along with a categorical-crossentropy loss function (9). n is the number of output classes, y is the ground truth vector for an image while $\hat{y}$ is the predicted vector. The model was trained for over 50 epochs with an early stopping callback to stop training when the model started overfitting the training data.

$$L_{\mathrm{CCE}} = -\sum_{i=0}^{n} y_{\mathrm{i}} \log \hat{y}_{\mathrm{i}} \tag{9}$$

3) CNN as a feature extractor for an SVM classifier: One major advantage of a CNN is that it is able to learn features from an input automatically. Keeping this in mind, we implemented a hybrid deep learning and machine learning approach for the classification task. The previously trained CNN was used for this task. The final layer of the network was removed, and the second last layer consisted of all the output features. This modified model was used to extract the features for the train, test and validation datasets. These extracted features were then inputted to an SVM classifier with a linear kernel and a square hinge loss (l2 loss). In order to fine-tune the model's hyperparameters, grid search was applied to the soft margin regularisation parameter of the SVM.

D. Testing

The test set comprised of 220 labelled images. Each image was pre-processed using the image processing algorithm defined previously before being used as an input to the model. Each model was individually evaluated using three evaluation metrics:

- Accuracy: The fraction of predictions a model got right.

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions\ made} \tag{10}$$

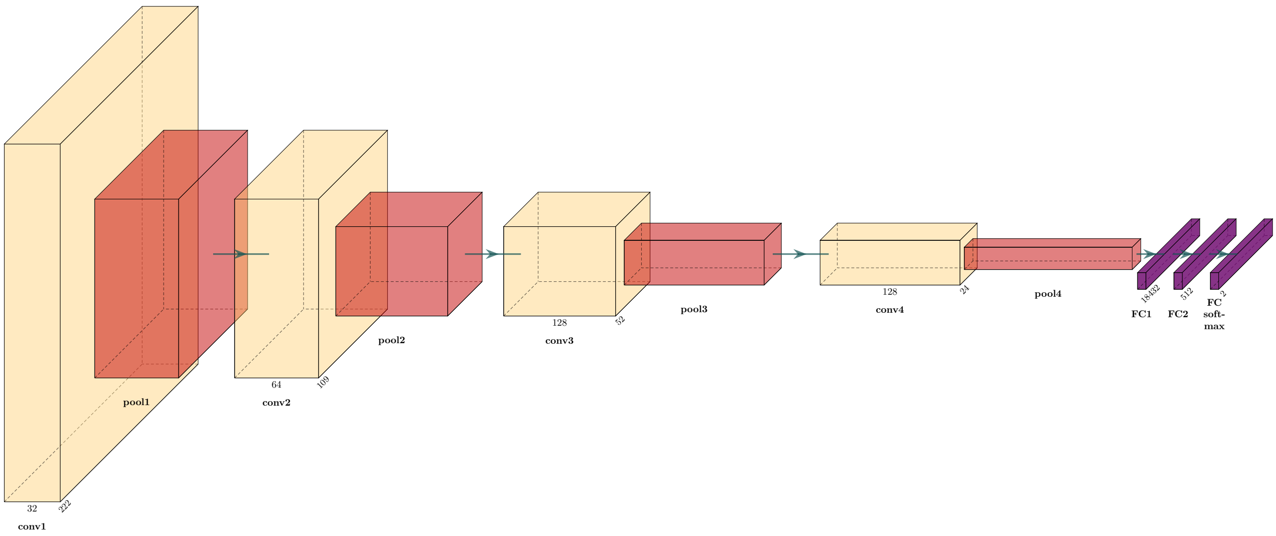

Fig. 5. CNN architecture

- Specificity: The true negative rate or the proportion of negatives that were correctly identified.

$$Specificity = \frac{True\ negatives}{True\ negatives\ +\ False\ positives} \tag{11}$$

- Sensitivity: The true positive rate or the proportion of positives that were correctly identified.

$$Sensitivity = \frac{True\ positives}{True\ positives\ +\ False\ negatives} \tag{12}$$

Sensitivity and specificity were chosen as metrics keeping in mind the class imbalance in the dataset. Apart from these metrics being evaluated on the test set, both the SVMs were evaluated using 10-fold cross-validation on the concatenated train and validation datasets. The mean accuracy and standard deviation of the accuracy were used as a metric to evaluate their performance on the training and validation data. The vanilla CNN was evaluated during training by carrying out validation at the end of every epoch.

IV. Results

We followed a two-fold method to evaluate the performance of the proposed models. Firstly, we tested how our models performed in comparison with each other on our test set. Secondly, we compared our best model's results on our test set with other existing models trained on similar datasets. All of the results achieved were averaged over five runs to ensure their credibility.

A. Comparison of proposed models

Table I shows the performance of the proposed models on our dataset. As evident from the results, the SVM trained on image features extracted using Hu moments and Haralick textures was outperformed by the vanilla CNN and the CNN-SVM model. The main reason behind this is that the CNN automatically learns features from the input domain, thereby being much more efficient in feature extraction. This, in turn, results in better classification.

TABLE I
Test set performance comparison

Model	**Accuracy**	**Specificity**	**Sensitivity**
SVM	0.6507	0.6371	0.7939
Vanilla CNN	0.9385	0.9612	0.9301
CNN-SVM	**0.9869**	**0.9857**	**0.9795**

TABLE II
Best model comparison with other models

Author	**Accuracy**	**Specificity**	**Sensitivity**
Krois, Joachim et al.	0.81	0.81	0.81
Kim, Jaeyoung et al.	-	0.96	0.87
Oliveira, João et al.	**0.9870**	-	-
Best model	0.9869	**0.9857**	**0.9795**

The performance of the CNN-SVM is even better than the Vanilla CNN model. The primary reason for this is a combination of CNNs outstanding local and global feature extraction ability and an SVMs classification ability. The features extracted by the CNN were more accurate than the features extracted using the manual methods described above.

B. Comparison of the best model with other models

Table II shows the results of the comparisons between our best model and other models. As evident from the results, our model outperforms all others in the fields of specificity and sensitivity.

V. Conclusion

In this paper, we provide a hybrid deep learning and machine learning approach using a CNN-SVM to classify panoramic dental x-rays as “Normal” or “Anomalous”. The images in the dataset were acquired from the Manipal College of Dental Sciences, Manipal. These were labelled by Dr Sunaina Puri on the basis of the prescence of evident caries, periapical infections, pathological migration, altered bone height as well as bony impactions of third molars. We used a pre-processing algorithm to enhance contrast, tooth boundaries and make the process of feature extraction more efficient. We compare our results with previous attempts at classifying panoramic dental x-rays. Experimental results showed that the proposed network achieved state-of-the-art results on our dataset. In the future, we plan on working on tooth boundary segmentation and classification of the affected teeth using state-of-the-art models such as the UNet or DeepLabV3 [1].

ACKNOWLEDGMENT

We would like to thank Dr Komal Smriti from the Manipal College of Dental Sciences, Manipal for providing the dataset and, Google Colab for providing the compute resources. We sincerely thank the efforts of Dr Sunaina Puri for labelling the entire dataset.

REFERENCES

[1] B. J. Bhatkalkar, D. R. Reddy, S. Prabhu, and S. V. Bhandary, "Improving the performance of convolutional neural network for the segmentation of optic disc in fundus images using attention gates and conditional random fields," *IEEE Access*, vol. 8, pp. 29 299–29 310, 2020.

[2] M. Dhillon, S. Raju, S. Verma, D. Tomar, R. Mohan, M. Lakhanpal, and B. Krishnamoorthy, "Positioning errors and quality assessment in panoramic radiography," *Imaging science in dentistry*, vol. 42, pp. 207–12, 12 2012.

[3] N. Pattanachai, N. Covavisaruch, and C. Sinthanayothin, "Tooth recognition in dental radiographs via hu's moment invariants," 05 2012, pp. 1–4.

[4] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610–621, 1973.

[5] N. Vamsha Deepa, N. Krishna, and G. Hemanth Kumar, "Feature extraction and classification of x-ray lung images using haralick texture features," in *Smart and Innovative Trends in Next Generation Computing Technologies*, P. Bhattacharyya, H. G. Sastry, V. Marriboyina, and R. Sharma, Eds. Singapore: Springer Singapore, 2018, pp. 899–907.

[6] G. Dougherty, *Digital Image Processing for Medical Applications*. Cambridge University Press, 2009.

[7] Y. Y. Amer and M. J. Aqel, "An efficient segmentation algorithm for panoramic dental images," *Procedia Computer Science*, vol. 65, pp. 718 – 725, 2015, international Conference on Communications, management, and Information technology (ICCMIT'2015). [Online]. Available: http://www.sciencedirect.com/science/article/pii/S187705091502846X

[8] A. Ehsani Rad, M. Shafry, M. Rahim, and A. Norouzi, "Digital dental x-ray image segmentation and feature extraction," *TELKOMNIKA Indonesian Journal of Electrical Engineering*, vol. 11, pp. 3109–3114, 06 2013.

[9] Guang-Yuan Zhang, Bo Cheng, Rui-Jia Feng, and Jia-Wen Li, "Real-time driver eye detection method using support vector machine with hu invariant moments," in *2008 International Conference on Machine Learning and Cybernetics*, vol. 5, 2008, pp. 2999–3004.

[10] J. Oliveira and H. Proença, *Caries Detection in Panoramic Dental X-ray Images*. Dordrecht: Springer Netherlands, 2011, pp. 175–190. [Online]. Available: https://doi.org/10.1007/978-94-007-0011-6_10

[11] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. [van der Laak], B. [van Ginneken], and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60 – 88, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1361841517301135

[12] J. Krois, T. Ekert, L. Meinhold, T. Golla, B. Kharbot, A. Wittemeier, C. Dörfer, and F. Schwendicke, "Deep learning for the radiographic detection of periodontal bone loss," *Scientific Reports*, vol. 9, 06 2019.

[13] J. Kim, H.-S. Lee, I. S. Song, and K.-H. Jung, "Dentnet: Deep neural transfer network for the detection of periodontal bone loss using panoramic dental radiographs," *Scientific Reports*, vol. 9, 12 2019.

[14] J.-H. Lee, D.-H. Kim, S.-N. Jeong, and S.-H. Choi, "Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm," *Journal of Dentistry*, vol. 77, pp. 106 – 111, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0300571218302252

[15] X. Sun, J. Park, K. Kang, and J. Hur, "Novel hybrid cnn-svm model for recognition of functional magnetic resonance images," in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2017, pp. 1001–1006.

[16] A. H. A. D.D.S., S. Kasaei, and M. Mehdizadeh, "Automatic segmentation of mandible in panoramic x-ray," *Journal of Medical Imaging*, vol. 2, no. 4, pp. 1 – 8, 2015. [Online]. Available: https://doi.org/10.1117/1.JMI.2.4.044003