Pulp Stone Detection Using Deep Learning Techniques

Amal Selmi, Liyakathunisa Syed, and Bashaer Abdulkareem(✉)

Taibah University, Medina, Saudi Arabia
amal.alrehily@gmail.com, lansari@taibahu.edu.sa,
b.alhejaily@gmail.com

Abstract. Today, the aid of modern technology in the medical field can be seen in every respect. In the field of radiology, many deep learning and image processing techniques have been applied for timely and better analysis, as well as conclusive results. However, dental radiographs are detailed, and some of these details are fine and vague, making them difficult to interpret. With the help of deep learning techniques, the automated uncovering of these fine details has great potential. In this paper, we aim to detect pulp stones in dental radiographs using Convolutional Neural Network (CNN)-based feature extraction followed by multiple classifiers. We conclude that the Residual Network 50 (ResNet-50) achieves an accuracy of 76.4% with the Medium Gaussian Support Vector Machine (SVM), while Inception v3 reaches an accuracy of 73.1% with the same classifier. Further, the ResNet-50's false positive rate is less than the Inception v3's by 7%, giving it the potential for further experiments.

Keywords: Deep learning · Dental radiographs · Pulp stones · Convolutional Neural Networks · Medium Gaussian Support Vector Machine

1 Introduction

Image processing is one of the most important and utilised concepts of computer science and modern technology. Moreover, it is used in multiple applications such as industrial inspection, the military, remote sensing, and medical imaging [1]. Moreover, researchers began to focus on automated analysis for medical images as soon as images became digitised. Initially, the analysis of medical images was accomplished using sequential applications of mathematical modelling and low-level pixel processing to build compound systems that were rule-based and had the ability to solve specific tasks. Furthermore, Artificial Intelligence (AI) analogy using expert systems with a large number of if–then–else conditional statements were popular from the 1970s to the 1990s [2]. However, we witnessed a shift from human-designed systems to computer-trained systems that use example data to train and extract feature vectors. Training systems with data given in pairs of images and that correlate outcomes for the images are known as supervised learning [3]. Thus, supervised techniques gained increasing popularity in the analysis of medical images by the end of the 1990s [2].

S. Spinsante et al. (Eds.): HealthyIoT 2021, LNICST 432, pp. 113–124, 2022.
https://doi.org/10.1007/978-3-030-99197-5_10

In the medical field, AI techniques, and computer vision in particular, are applied widely. Additionally, deep learning using Convolutional Neural Networks (CNNs) is known as a type of the many models of machine learning and has proven to have astounding potential to assist doctors in the detection of pathologies. Moreover, it has the potential to detect bodily-related structures from medical images with high accuracy that can surpass that of medical professionals. CNNs also have the ability to detect structures (e.g., decide if a tooth is on an image), segment them (e.g., identify the precise shape of a tooth on an image) and classify them (e.g., label every tooth in a dentition) [3].

Undoubtedly, medical imaging plays a critical role in fields such as dentistry. Today, dentists diagnose patients by analysing the patient's panoramic radiographs. After that, they observe and register tooth conditions, such as those of implants, dental bridges etc. Additionally, the Federation Dentaire International notation system is used to register tooth symptoms, a process that uses Arabic numerals to clarify the positioning of the tooth. The dentition is then accurately divided into four quadrants, with eight teeth in each quadrant, and each tooth is identified by its two-digit number. The first of the two digits indicates the quadrant in which the tooth is located, while the second digit indicates the tooth's number in this quadrant. Overall, the assessment of the tooth conditions is through observation, and the recording is done manually in a text file. As a result, the full process costs considerable effort and time, with delays in diagnostic results [1] and obstructions to the clinic's workflow. Furthermore, some studies clarify that dentistry students' traditional way of interpreting radiographic images is inaccurate. Thus, we cannot depend on observation only to identify complex and vague cases [4]. Therefore, for a clinical decision support system that is based on imaging, medical image classification is a crucial step. However, the existing techniques for X-ray image classification present the image using only its generic features, such as its shape and colour. Hence, the ability to measure the image's characteristics is limited [5].

A large volume of images is acquired every year in the dental field. For example, in European countries, dental radiography such as panoramic and cephalometric radiography is likely the most taken type of radiograph in the medical field. For instance, it was roughly calculated that between 250 and 300 dental images for every 1000 individuals were taken in 2010. Additionally, CNNs in medicine have been successfully used for object detection in image segmentation and recognition, and recently, CNNs have been used in the field of dentistry to reveal periodontal bone loss [6]. Thus, the application of CNNs has much potential in the dentistry field, even though which CNN applications are most focused upon or which techniques (imagery types, outcome metrics etc.) are most popular has not been systematically assessed. Additionally, it is not clear how CNNs perform in comparison to human experts. Thus, further research and implementation could help to evaluate how advantageous the application of CNNs is for dental clinical practice [3].

Even though multiple CNN applications have been used in the dental field, such as cavity classification [8], oral cancer detection [9], teeth segmentation [10] etc., the detection of fine and vague objects such as pulp stones using radiography is yet to be explored. Therefore, in this paper, we aim to detect pulp stones in X-ray images using CNN-based feature extraction and the application of different classifiers.

The paper is organised as follows: background, methodology, experimentation, results and discussion and conclusion.

2 Background

2.1 Convolutional Neural Networks

CNN is a deep learning technique that implicitly extracts image data from deeper networks through feature extraction. Specifically, the CNN model consists of several layers of convolution and pooling that are stacked on top of each other. Additionally, several weights are shared by the convolutional layer, while the pooling layer sub-samples the convolutional layer output and reduces the data rate [15]. Additionally, CNN is commonly used in the area of image processing. Moreover, it has been inspired by biological processes such as the human brain's ability to distinguish between various objects through visuals only and the use of previous observations. This phenomenon can be used in the processing of medical images to classify various images and identify diseases [15].

2.2 Deep Learning in Dentistry

In 2017, research started to use CNN to detect teeth. In a study, the automated identification of teeth and their numbering was performed using CNN, and a heuristic approach was used to detect the teeth [16]. Structure identification and segmentation are the main tasks that CNNs have been tested for so far. Surprisingly, even with the given difficulties in dental image diagnostics, a few studies inspected the detection of pathologies resulting in limited accuracy and reliability. Seemingly, the focus of researchers has been on basic tasks such as identification (e.g., mostly tooth identification) because it is considered the foundation of more complex detection systems. Then, as a subsequent step, pathologies are detected and allocated to the teeth [3].

Schwendicke et al. (2019) the authors studied a collection of existing literature and concluded that almost half of the studies have used CNN architectures that are individually constructed. Furthermore, CNNs performed remarkably well when applied to medical imagery. Their application could allow for more thorough, precise and reliable disease detection and image evaluation. Notably, the use of CNNs has been continuously increasing in the field of dental research. However, even though the first published application for CNN in dentistry was in 2015, there is no significant leap in sight. On the other hand, medicine had 42 CNN entries available on PubMed in 2015 and over 800 in the first half of 2019. Therefore, it is likely that dentistry is following medicine in general in terms of the application of CNNs. Thus, this may allow for the clarification of methodological standards and early strictness, but different disciplines possess large amounts of evidence, which may lead to bias [3].

Many deep learning algorithms are composed of several layers that convert input data, such as images, into outputs, such as existing diseases. The use of CNNs is broadening presently in the analysis of medical images [2]. Lee et al. (2018) the detection and diagnosis of dental caries were done using CNNs for tooth classifications and diagnoses. Moreover, in this study, they aimed to evaluate the efficiency of deep CNNs because panoramic radiographs have different distortions according to the region being photographed [11].

Today, dental cavities usually occur due to the consumption of sugary drinks, food particles etc. When they are left in a tooth and some time passes, bacteria are generated. After this, plaque forms because of the mixed bacteria, saliva and acid. Soon after, the tooth enamel targets the plaque and results in the appearance of the familiar blackish holes. Hence, early diagnosis helps to avoid the development of dental illnesses, and the analysis of dental images allows for the accurate detection of dental disease in its early stages [8]. The method that Sukegawa et al. (2020) aimed to detect the early stages of cavities through Sobel Edge detection and deep CNN. Furthermore, contrast enhancement, histogram equalisation and feature selection were applied in the pre-processing stage. Additionally, diverse techniques of segmentation were used for comparison, namely Watershed and Otsu's threshold. The proposed method detected the edges in the dental images using the Sobel method and was applied in the gradient measurement of the intensity values of the pixel in the Gx and Gy directions. They concluded that the proposed method accomplished a 96.08% accuracy for its rate of prediction in comparison to other methods [8].

Notably, most oral lesions can develop into oral cancer. Moreover, the primary diagnosis of oral cancer involves examining the ocular regions accurately and recording the patient's oral cavity with true-colour digital images. Following this, the choice of further treatment for the patient with oral cancer mostly relies on the appearance of the lesion [9]. In Prabhakar and Rajaguru (2017), it was suggested that oral cancer is classified into two types: pathological and clinical. The goal was to assist the diagnosis of oral cancer patients using neural network classifiers. In the study, 75 oral cancer patients were examined with the help of Tumour Node Metastasis (TNM). The TNM variables were regarded as the input variables, and the results showed that the average accuracy of the classification was 100% for stage 1, 85.19% for stage 2, 84.21% for stage 3 and 94.12% for stage 4. Notably, the TNM accuracy results were compared to the linear layer neural network. Thus, these mentioned diagnostic tests can help to determine the stages of oral cancer, as well as help to make a treatment decision [9].

In the dental field, X-ray image techniques are classified into two categories: One technique is intraoral radiography, which is obtained from within the mouth of the patient, and the other technique is extraoral radiography, which is obtained from outside the mouth of the patient. In 2017, researchers proposed applying deep learning methods, such as tooth segmentation, to dental X-ray images depending on the detection and segmentation of every tooth in the panoramic images. Moreover, they used 1500 panoramic images in their study, even though it is considered a difficult way to isolate teeth through panoramic images because they show some parts of the patient's body, such as their jaw and chin, as well. Thus, the researchers suggested an automated method of segmentation for isolating. Consequently, this technique may be a start to assist dentists with their diagnoses through panoramic images [10].

In 2020, researchers in Japan performed a study on dental implants. They presented a classification model of various dental implant brands based on using the deep learning method. Moreover, it focused on using CNNs with panoramic X-rays for classification. Therefore, they applied and evaluated five models of CNN for implant classification.

However, among the five models, the finely tuned VGG16 model displayed the best performance in classifying implant brands in panoramic X-rays [12].

2.3 The Importance of Pulp Stone Detection

In dentistry, dental pulp is living tissue that survives due to the steady flow of blood. Pulp stones, often observed in bitewing and periapical radiographs, are separate calcified bodies in the dental pulp of healthy, diseased and unerupted teeth. Further, pulp stones vary in size, ranging from small, microscopic particles to large masses that almost occlude the pulp chamber. Additionally, they are more common in the coronal locations than in the pulp's radicular portions [13].

Memon et al. (2018), the authors conducted a study aiming to radiologically detect pulp stones and investigate any association between the occurrence of pulp stones and age, gender, the type of tooth, the dental arch and the status of the tooth. It was discovered that 21.7% of the patients suffering from renal stones had pulp stones and that a local or systematic pathology may increase the number or size of pulp stones. Hence, a correlation may exist between local or systemic pathologies and the occurrence of pulp stones. Moreover, the detection of vague objects is integral for the future of dental clinical support systems. In Fig. 1, two images from this research's dataset are displayed. The left side of the image shows a healthy tooth, while the right side of the image shows a tooth with a pulp stone.

Fig. 1. A healthy tooth (left), and a tooth with a pulp stone (right).

3 Methodology

In the following section, we will discuss our methodology to detect and classify pulp stones. Figure 2 shows the framework for the proposed methodology.

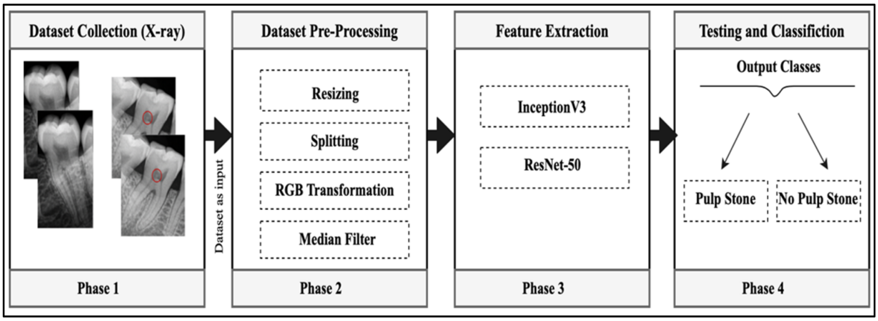

Fig. 2. Framework for the proposed methodology.

3.1 Dataset

The Information Technology department at Taibah University Dental Hospital in Al-Medina, KSA, archived and collected the dataset. A total of 212 periapical radiographs of female patients from 2019 to 2021 were selected and converted from RVG format to PNGs. Additionally, the dataset was divided into two folders: The first was labelled 'Stone' and contained 106 periapical radiographs showing pulp stones, and the other was labelled 'No Stone' and contained 106 periapical radiographs showing no pulp stones. The dataset was divided with the help of a radiology specialist.

3.2 The Pre-processing Stage

Since we aimed to classify X-ray images based on the detection of small and fine objects in the tooth pulp, CNN was chosen. The choice of CNN was mainly because of its capability to classify images, as well as its models' reduced computational time. Image pre-processing is a crucial technique that handles discrepancies and faults in images and prepares them for the upcoming procedures within the CNN models. Often, X-ray images contain some unclear boundaries and other defects such as poor exposure. In our case, we resized the dataset images for each model: The images were resized to 299 × 299 for the Inception v3 model and 224 × 224 for the Residual Network 50 (ResNet-50) model. The images were also converted from greyscale to RGB scale. Additionally, due to the targeting of fine and vague objects in the X-rays, we applied a median filter to remove the noise from the images. After that, the following techniques were applied to these pre-processed images.

3.3 Feature Extraction

In our method, we considered two different pre-trained CNN models, namely the Inception v3 and ResNet-50. The ResNet-50 was chosen because of its capability to handle complex problems while maintaining good performance, while the Inception v3 model was selected for its efficiency in detecting specific features that are inconsistent in size. To follow is a brief description of the models, along with their architecture.

ResNet-50: One of the known CNNs is the ResNet-50. It is a 50-layer deep model that was trained using a large number of images from the ImageNet database. The ResNet-50 can classify images into 1000 categories based on objects [22]. Additionally, the model contains five stages that are residual blocks, each of which contains layers and convolutions [20]. The overall architecture can be seen in Fig. 3.

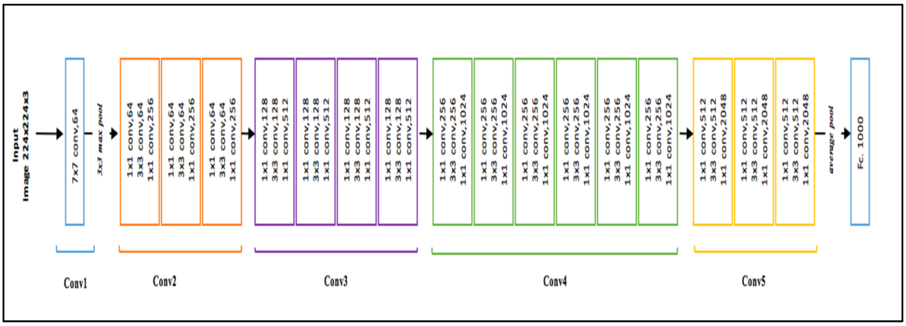

Fig. 3. The ResNet-50 [20].

Inception v3: Szegedy et al. first introduced the Inception v3 model in 2016. It is a CNN classification architecture that takes three convolutions with different window sizes and joins them all together [18]. Thus, the joining is done by chaining the three outputs into a single value. Then, the single-value output becomes the input of the next stage. Further, max pooling is connected with the convolution layers as an assessment additional to the classification process. Moreover, another 1×1 convolutional layer is added to the overall architecture. Hence, the overall architecture can be seen in Fig. 4.

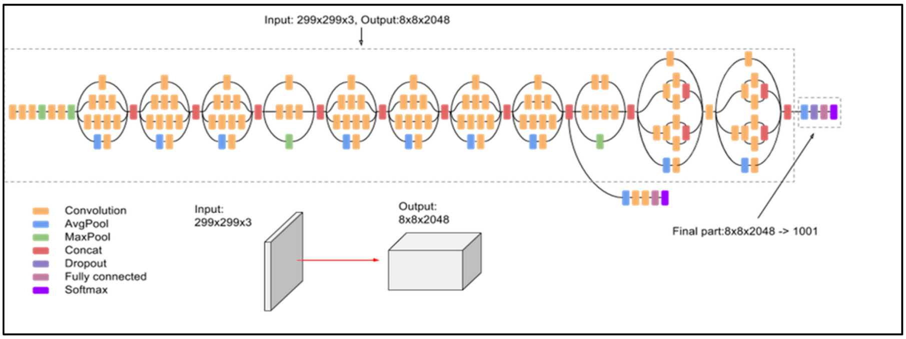

Fig. 4. The Inception v3 model [19].

3.4 The Classification Stage

After we extracted the features, the classification task was accomplished using different classifiers, namely the Support Vector Machine (SVM), Logistic Regression (LR) and K-Nearest Neighbours (KNN).

4 Experimentation

This section represents the experiments we performed to accomplish our goal of detecting pulp stones accurately. We used MATLAB (Ver 9.4.0) and uploaded our dataset to the MATLAB platform. The experiment consisted of two phases that differed in the size of their datasets but were identical in the rest of the processes. The dataset of the first phase contained 106 images in total, while the dataset of the second phase was expanded to 212 images. First, the dataset was randomly split into two sets: 70% for training and 30% for testing. Further pre-processing was performed, including resizing the images, transforming greyscale to RGB and applying the median filter to the dataset. Then, each of the two models, the Inception v3 [19] and ResNet-50 [7], was tested with the pre-processed dataset. By the end of the compilation, the feature labels, in addition to the testing and training, were extracted for each model. Furthermore, we evaluated the models' performance and applied different classifiers to each. In Table 1, the accuracy that the multiple classifiers of the two models obtained is listed.

Table 1. Obtained results from applying multiple classifiers to the pre-trained convolutional neural network models.

Phases	Classifier	Feature extractor	
		Inception v3	Resnet50
'One'	Cubic SVM	82.1%	81.1%
	Medium Gaussian SVM	81.1%	84%
	Logistic Regression	62.3%	58.5%
	Fine KNN	75.5%	78.3%
'Two'	Cubic SVM	72.6%	75.5%
	Medium Gaussian SVM	73.1%	76.4%
	Logistic Regression	51.4%	56.1%
	Fine KNN	67.5%	67%

5 Results and Discussion

In the experiment, we started with 106 images equally divided into stone and non-stone radiographs as our dataset. As a result, the ResNet-50 model reached an accuracy of 84% with the Medium Gaussian classifier, while the Inception v3 reached an accuracy of 81.1% with the same classifier. However, with the expansion of the dataset to reach a total of 212 images, it was noted that the accuracy of the two models decreased. In the second phase, the ResNet-50 reached an accuracy of 76.4% with the Medium Gaussian SVM classifier, while the Inception v3's accuracy with the same classifier was 73.1%. Notably, the ResNet-50 still excelled in comparison to the Inception v3 model. Ultimately, the accuracy results of the SVM classifiers for the two models were considered close. Figure 5 shows the results of the first phase, while Fig. 6 shows the results of the second phase.

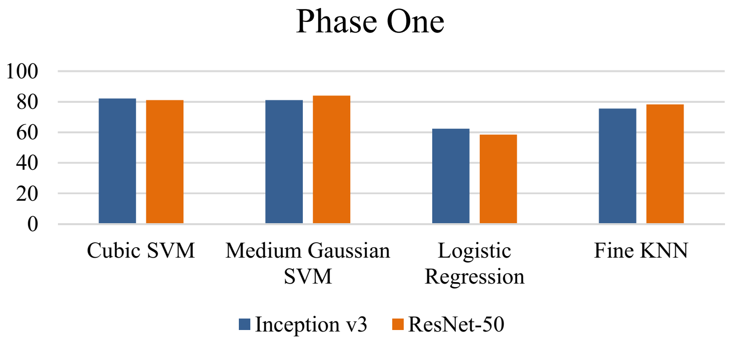

Fig. 5. A bar graph of the area under the curve obtained from applying multiple classifiers to the pre-trained convolutional neural network models—Phase one.

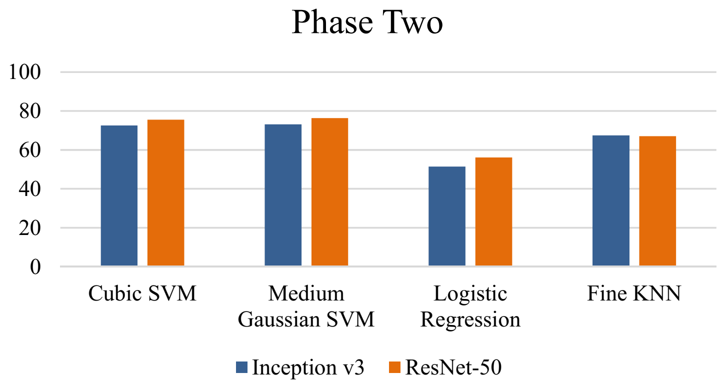

Fig. 6. A bar graph of the area under the curve obtained from applying multiple classifiers to the pre-trained convolutional neural network models—Phase two.

The area under the Receiver Operating Characteristic (ROC) curve summarised the performance of the two models after the use of multiple classifiers. We were interested in the ROC curve of the second phase since it represented a challenge due to its larger dataset. In Fig. 7, the ROC curve on the left side of the image shows that the ResNet-50 model incorrectly assigned 34% of the observations to the positive class. However, the model also correctly assigned 70% of the observations to the positive class. Overall, the total accuracy of the performance reached 0.70 for the ResNet-50 model. On the other hand, in Fig. 7, the ROC curve on the right side of the image shows that the Inception v3 model incorrectly assigned 41% and correctly assigned 58% of the observations to the positive class. Thus, the Inception v3 reached a total performance accuracy of 0.64.

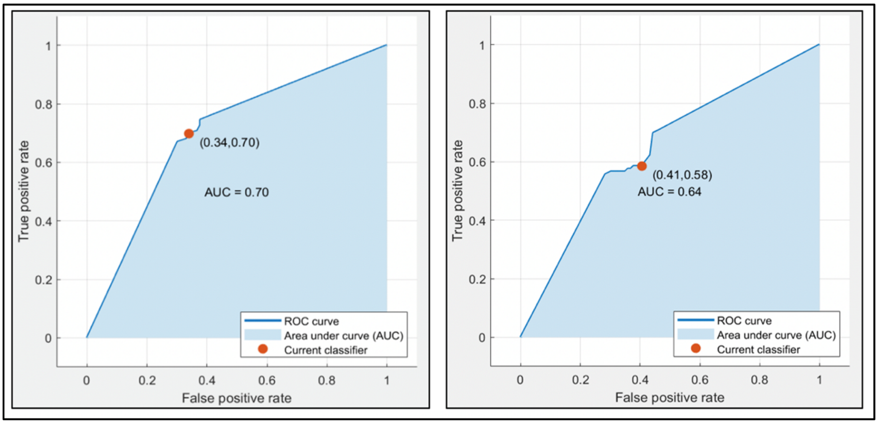

Fig. 7. The receiver operating characteristic curve for the ResNet-50 model (left), and Inception v3 model (Right).

Hence, these results suggest that the ResNet-50 is the best model for detecting pulp stones in radiographs since the false positive rate of the ResNet-50 was lower than that of the Inception v3, and the accuracy of the classifiers was better with the ResNet-50 than with the Inception v3.

It was noted that the flexibility of the chosen technique was advantageous because the parallel implementation of the different models from the same categories provided space for observation and alteration. Thus, the enhancements of the models and the choice of the classifiers were easier. However, the needed computational power and the complexity of the models were limited. Moreover, it was concluded that while the ResNet-50 model had the ability to classify pulp stones in radiographs, the decrease in accuracy when the dataset was expanded shows that it needs improvement. Additionally, the fuzziness of some of the radiographs, along with the vagueness of the object to be detected, made the classification more challenging and may have increased the false positive rate.

6 Conclusion

CNN usage in the medical field is increasing, and different applications are appearing every day. This paper aimed to assess the potential that CNN models have to detect fine and vague objects in dental radiographs. The fine objects we aimed to detect were pulp stones. Notably, pulp stones are objects that may be unclear for the observer to notice instantly, especially for dental students who are not yet used to the details of radiographs. We tested and compared the performance of two CNN models, namely the ResNet-50 and Inception v3, with multiple classifiers. The ResNet-50 model reached an accuracy of 76.4% with the Medium Gaussian SVM, while the Inception v3 reached an accuracy of 73.1% with the same classifier. Importantly, this experiment clarified the potential for uncovering the fine and vague details of dental radiographs accurately. It also clarified that the application of CNNs is promising in terms of uncovering the vague details of dental radiographs and can be used in clinical decision support systems.

7 Future Work

We aim to continue in the same domain and improve our methodology for detecting pulp stones in X-ray images. Further, we aim to increase the accuracy of our methodology by applying enhancements to the model architectures used.

Acknowledgements. The authors would like to express their appreciation to Dr Ebtihal H. Zain Alabdeen, assistant professor of oral radiology dentistry at the College of Dentistry at Taibah University for her great help and effort during the data preparation.

References

1. Karthick, K., Premkumar, M., Manikandan, R., Cristin, R.: Survey of image processing based applications in AMR. Rev. Comput. Eng. Res. **5**(1), 12–19 (2018)
2. Litjens, G., et al.: A survey on deep learning in medical image analysis. Med. Image Anal, **42**, 60–88 (2017)
3. Schwendicke, F., Golla, T., Dreher, M., Krois, J.: Convolutional neural networks for dental image diagnostics: a scoping review. J. Dent. **91**, 103226 (2019)
4. Soltanimehr, E., Bahrampour, E., Imani, M., Rahimi, F., Almasi, B., Moattari, M.: Effect of virtual versus traditional education on theoretical knowledge and reporting skills of dental students in radiographic interpretation of bony lesions of the jaw. BMC Med. Educ. **19**(1) (2019)
5. Ahn, E., Kumar, A., Kim, J., Li, C., Feng, D., Fulham, M.: X-ray image classification using domain transferred convolutional neural networks and local sparse spatial pyramid. In: IEEE 13th International Symposium on Biomedical Imaging (ISBI), pp. 855–858 (2016)
6. Privado, M., Villalón, J., Martínez, C., Ivorra, C.: Dental images recognition technology and applications: a literature review. Appl. Sci. **10**(8), 2856 (2020)
7. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
8. Muthu Lakshmi, M., Chitra, P.: Classification of dental cavities from X-ray images using deep CNN algorithm. In: 4th International Conference on Trends in Electronics and Informatics (ICOEI) (48184), pp. 774–779 (2020)
9. Prabhakar, S., Rajaguru, H.: Performance analysis of linear layer neural networks for oral cancer classification. In: 6th ICT International Student Project Conference (ICT-ISPC), pp. 1–4 (2017)
10. Jader, G., Fontinele, J., Ruiz, M., Abdalla, K., Pithon, M., Oliveira, L.: Deep instance segmentation of teeth in panoramic X-ray images. In: 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), IEEE, Parana (2018)

11. Lee, J.-H., Kim, D.-H., Jeong, S.-N., Choi, S.-H.: Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. J. Dent. **77**, 106–111 (2018)
12. Sukegawa, S., et al.: Deep neural networks for dental implant system classification. Biomolecules **10** (2020)
13. Galav, A., Vyas, T., Kaur, M., Chauhan, M., Satija, N.: Association of pulp stones & renal stones- a clinical study. J. Oral. Biol. Craniofac. Res. **5**(3), 189–192 (2018)
14. Memon, M., Kalhoro, F.A., Shams, S., Arain, S.: A study on radiographic assessment of pulp stone. J. Endod. **25**, 992–996 (2018)
15. Khan, S., Yong, S.-P.: A Deep learning architecture for classifying medical images of anatomy object. In: 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC). IEEE, Kuala Lumpur (2017)
16. oktay, A.B.: Tooth detection with convolutional neural networks. BioMed. Res. Int. **2021** (2017)
17. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016)
18. Xia, X., Xu, C., Nan, B.: Inception-v3 for flower classification. In: 2017 2nd International Conference on Image, Vision and Computing (ICIVC) (2017)
19. Papers with code - Inception-v3 explained: Paperswithcode.com. https://paperswithcode.com/method/inception-v3. Accessed: 12 Dec 2020
20. Almabdy, S., Elrefaei, L.: Feature extraction and fusion for face recognition systems using pre-trained convolutional neural networks. Int. J. Comput. Digit. Syst. **9** (2020)
21. Rezende, E., Ruppert, G., Carvalho, T., Ramos, F., De Geus, P.: Malicious software classification using transfer learning of resnet-50 deep neural network. In: 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA), pp. 1011–1014. IEEE (2017)
22. Meghana, A.S., Sudhakar, S., Arumugam, G., Srinivasan, P., Prakash, K.B.: Age and gender prediction using convolution, ResNet50 and inception ResNetV2. Int. J. Adv. Trends Comput. Sci. Eng. **9** (2020)