Histopathologic Image Classification of Benign Fibro-Osseous Lesions of the Jaws Using Deep Convolutional Neural Network

Natchanapat Promana*
Piyatida Meesatean*
natchanapat.kkwan@mail.kmutt.ac.th
piyatida.kutara@mail.kmutt.ac.th
Computer Engineering Department
King Mongkut's University of Technology Thonburi
Princess Srisavangavadhana College of Medicine, Chulabhorn Royal Academy
Bangkok, Thailand

Paniti Achararit
Princess Srisavangavadhana College of Medicine, Chulabhorn Royal Academy
Bangkok, Thailand
paniti.ach@cra.ac.th

Phond Phunchongharn
Computer Engineering Department
King Mongkut's University of Technology Thonburi
Bangkok, Thailand
phond.p@mail.kmutt.ac.th

Kraisorn Sappayatosok
College of Dental Medicine, Rangsit University
PathumThani, Thailand
kraisorn.s@rsu.ac.th

Ekarat Phattarataratip†
Oral and Maxillofacial Pathology at the Faculty of Dentistry, Chulalongkorn University
Bangkok, Thailand
ekarat.P@chula.ac.th

ABSTRACT

Fibro-osseous lesions of the jaws are a group of diseases that are challenging to diagnose in general there is a diverse and high overlap in the histopathological features. It is necessary to consult with a pathologist and dentist who specializes in diagnosis. In this research, we present a comparison of different digital image preprocessing techniques to enhance the quality of histopathological images and find the appropriate solution to elevate the classification capability of the deep convolutional neural network: InceptionResNetV2. To achieve this, we experimented with various techniques, including grayscale, contrast limited adaptive histogram equalization (CLAHE), global histogram equalization (GHE), piecewise linear contrast stretching (P)+CLAHE, and P+GHE. The performance of the techniques was evaluated based on their ability to enhance image quality and the accuracy of the model. The results show that P+GHE images achieved the highest accuracy, followed by CLAHE with accuracy rates of 89% and 87%, respectively. We also conducted an additional experiment to compare the order of technique and found that using P before CLAHE/GHE can reduce noise in the image. Overall, our findings suggest that image preprocessing techniques can improve the quality and accuracy of model classification.

*Both authors contributed equally to this research.
†Corresponding author: Ekarat Phattarataratip, ekarat.P@chula.ac.th

ICBBT 2023, May 26–28, 2023, Xi'an, China

ACM ISBN 979-8-4007-0038-5/23/05…$15.00
https://doi.org/10.1145/3608164.3608165

CCS CONCEPTS

• **Applied computing → Imaging**; **Computational biology**.

KEYWORDS

classification, convolutional neural networks; deep learning, fibro-osseous lesions, histopathological image

ACM Reference Format:
Natchanapat Promana, Piyatida Meesatean, Paniti Achararit, Phond Phunchongharn, Kraisorn Sappayatosok, and Ekarat Phattarataratip. 2023. Histopathologic Image Classification of Benign Fibro-Osseous Lesions of the Jaws Using Deep Convolutional Neural Network. In *2023 15th International Conference on Bioinformatics and Biomedical Technology (ICBBT 2023), May 26–28, 2023, Xi'an, China.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3608164.3608165

1 INTRODUCTION

Fibro-osseous lesions (FOLs) of the jaws constitute a unique group of pathologic conditions that exhibit an admixed production of fibrocollagenous tissue and mineralized materials microscopically [1]. Their classification has been a subject of constant modification, with the 2022 World Health Organization (WHO) classification of head and neck tumors including cemento-osseous dysplasia, fibrous dysplasia, cemento-ossifying fibroma, juvenile ossifying fibroma, familial gigantiform cementoma, and segmental odontomaxillary dysplasia into this group [2]. By far the most common FOLs of the jaws are cemento-osseous dysplasia, fibrous dysplasia, and cemento-ossifying fibroma, respectively [3].

The diagnosis of these lesions can be challenging due to their diverse and high degree of overlap in the histopathological features, which pose challenges to diagnosis. Pathologists need to be skilled in the diagnosis, and microbiological characteristics are used in combination with clinical characteristics such as the patient's age, symptoms, affected location, and accompanying radiographic features to diagnose the disease more specifically [4].

Artificial intelligence (AI) has been intensively studied in recent literature and could offer possibilities to revolutionize the diagnostic paradigm in medicine and dentistry [5]. Deep neural network is a type of AI algorithm that utilizes multi-layered neural networks to compute the dataset, and it shows strong potential in disease diagnosis, predicting prognosis, and helping develop personalized treatment planning [6]. Currently, AI is in the developmental phase, and results are promising in various fields of dentistry, particularly in the diagnosis of oral diseases using radiographic and/or clinical datasets. Several research developments focus on the assessment of deep learning models in disease classification using digital histopathology images. [7] [8]

In this research, the challenge of diagnosing FOLs was addressed by introducing a deep neural network model to classify the three types of lesions belonging to the group of lesions of the jaw using the histopathological image dataset. Due to the nature of the photographic dataset, it was difficult to use in model development, and the researchers emphasized the process of improving the quality of image data by doing image preprocessing using the OpenCV library, such as adjusting image color, contrast, and image detail with various techniques to develop a deep neural network model that classified the diseases with highly accurate and reliable prediction results.

Overall, the goal of this research is to increase the accuracy and help the pathologist diagnose FOLs of the jaws more efficiently. An accurate diagnosis is necessary for effective treatment planning and management of the disease. Moreover, we conducted a comprehensive comparison of image processing techniques using various experimental designs and some novel combinations to find the techniques and parameters that are most effective for our dataset by using the grayscale image, global histogram equalization (GHE), contrast limited adaptive histogram equalization (CLAHE) and piecewise linear contrast stretching. Additionally, we also use convolutional neural networks for image classification to further enhance the accuracy of our analysis. We hope to make significant strides in the field of medical imaging analysis.

2 PRELIMINARIES

This section is divided into two parts: a review of the literature and basic knowledge used in this research.

2.1 Related work

There are several pieces of research on image classification histopathological images with the deep convolutional neural network but it is unused to fibro-osseous lesions.

H. El Attar M. El Khannoussi H. El Agouri, M. Azizi, and A.Ibrahimi presented the histopathological image of a breast cancer prediction model by using deep learning algorithms: ResNet50 and Xception. The model accuracy was 84.5% and 88%respectively [9]. Chao Liu Haosheng Tang, Guo Li, and Donghai Huang studied the lymph node metastasis in head and neck squamous cell carcinoma diagnosis using different deep learning models with 98.7% accuracy from MobileNet-V2. However, it had only 135 records for training models that affected model reliability and it could be biased from the limited dataset [10]. Şaban Öztürk and Bayram Akdemir studied the effect of histopathological image preprocessing on convolutional neural networks. This paper compared the efficiency of pre-processing techniques in which image sets were grouped into 4 groups, original images and enhance images (3 groups). Each group used different image preprocessing techniques. They reported that the second normal preprocessing algorithm using histogram equalization had the highest accuracy with 94.7% by the AlexNet model. [11]

2.2 Deep convolutional neural network

A deep convolutional neural network (CNN) is a type of neural network that is commonly used for image recognition and analysis. It is designed to automatically learn and extract features from images by applying a series of convolutional and pooling layers. These layers allow the network to identify patterns and relationships in the image data, such as edges, textures, and shapes.

It typically consists of multiple layers, each with a specific function in the image processing pipeline. The first layers are responsible for low-level feature extraction, such as detecting edges and corners, while the later layers focus on higher-level features, such as object recognition and classification.

The architecture of a convolutional neural network (CNN) typically consists of the following layers: Convolutional layer, Activation layer, Pooling layer, Fully connected layer, Dropout layer, and Output layer.

2.3 Image Processing

2.3.1 Contrast Limited Adaptive Histogram Equalization (CLAHE). Contrast limited adaptive histogram equalization (CLAHE) is a digital image processing technique that enhances the contrast of images by redistributing the pixel intensities of an image using histogram equalization. However, standard histogram equalization can result in the amplification of noise in the image, leading to the over-amplification of small-scale features. CLAHE is a modified version of histogram equalization that overcomes this limitation by limiting the contrast amplification in local regions of an image [12].

The technique is implemented by dividing the image into small, non-overlapping regions called tiles. Each tile has its own histogram, and the contrast amplification is limited by clipping the histogram of each tile at a specified clip limit value. The clipped portions of the histogram are then redistributed uniformly over the histogram range. This process ensures that the contrast amplification is limited and helps to prevent the over-amplification of small-scale features.

2.3.2 Global Histogram Equalization (GHE). Global Histogram Equalization (GHE) is a widely used technique in digital image processing to enhance the contrast of images. The goal of GHE is to distribute the intensity values of an image uniformly over the entire range of intensities. This is achieved by computing the cumulative distribution function (CDF) of the intensity values in the image and then transforming the intensity values based on the CDF [12, 13].

2.3.3 Piecewise linear contrast stretching (P). Piecewise linear contrast stretching is a technique used to enhance the contrast of an image by mapping the pixel values to a new range of values. In this technique, the input image is first divided into several regions,

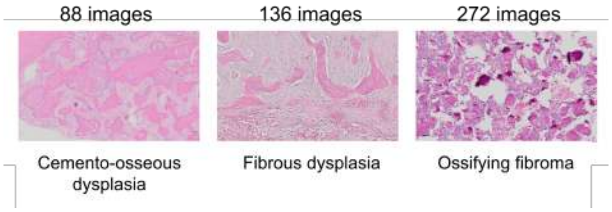

Figure 1: Examples and number of pathological images using in our research.

and then a linear mapping function is applied to each region. This function is defined by two points, the start point, and the endpoint, and it maps the pixel values in that region to a new range of values [12].

2.4 Grad-CAM Visualization

Grad-CAM (Gradient-weighted Class Activation Mapping) is a technique for visualizing and understanding the decisions made by deep convolutional neural networks (CNNs) in computer vision tasks. It uses the gradient information flowing into the last convolutional layer of the CNN to produce a coarse localization map of the important regions in the input image that contribute to the final classification decision. Grad-CAM can be used to identify the regions of the input image that are most important for a given class label, providing insight into how the CNN is making its decision [14].

3 MATERIAL AND METHODS

This section includes the details about our dataset, the overall framework of our proposed framework, image preprocessing, image classification using a CNN model, and understanding of the model classification by using Grad-CAM, respectively.

3.1 Dataset

All cases are retrieved from the archive of the Department of Oral Pathology, Faculty of Dentistry, Chulalongkorn University during 2012-2021. All microscopic slides will be reviewed, and the diagnoses are confirmed by experienced pathologists based on the 2022 World Health Organization criteria. Histopathologic images of H&E-stained tissues from all included cases are obtained using Olympus CX31 microscope equipped with Canon EOS 600D EOS Digital SLR Camera at 100X magnification. Only images showing characteristic lesions are included in the respective diagnostic groups, which are Cemento-osseous, Fibrous dysplasia, and Ossifying fibroma. The number and example of images in each group is illustrated in Figure 1.

3.2 Pathological Images Classification

Our proposed framework consists of two major parts as shown in Figure 2 which are i) image preprocessing for preparing image dataset before training deep neural network model and ii) image classification based on a CNN model. Finally, we use gradient-weighted class activation Mapping (Grad-CAM) for visual explanation.

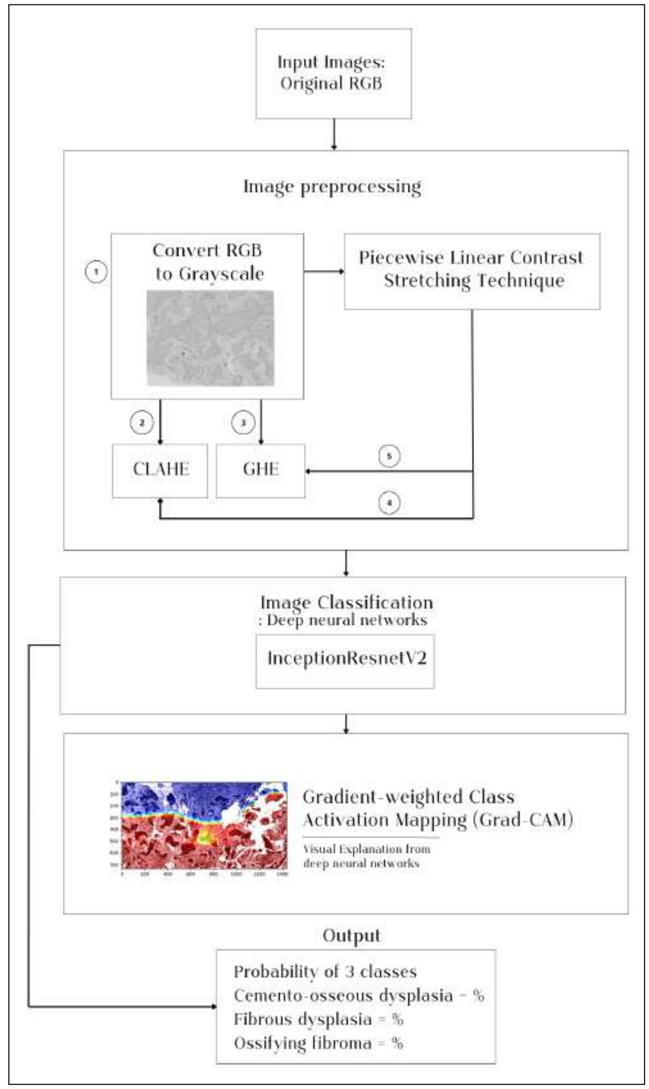

Figure 2: Framework

3.2.1 Image Preprocessing. Image preprocessing is to prepare the dataset for training the deep neural network model. The model aims at classifying the FOLs of the jaw. Various digital image preprocessing techniques are used to enhance the quality of image data by adjusting the image color and contrast. In this study, three image preprocessing techniques are considered: I) piecewise linear contrast stretching (P), II) global histogram equalization (GHE), and III) contrast limited adaptive histogram equalization (CLAHE). Note that the OpenCV library is used to perform image preprocessing.

I) Contrast Limited Adaptive Histogram Equalization (CLAHE) We aim to enhance histopathology details in images dataset. CLAHE is particularly useful in enhancing the contrast of medical images. Before using this technique we convert the original RGB image to grayscale. We use clipLimit = 12 and tileGridSize = (8,8) to apply CLAHE to all images.

II) Global Histogram Equalization (GHE) We apply GHE to the grayscale images by converting the original RGB images to grayscale. We set the parameters of GHE to achieve the best results for our dataset. The performance of the GHE technique is evaluated by comparing it with other techniques, such as CLAHE and piecewise linear contrast stretching, using different experimental designs.

III) Piecewise Linear Contrast Stretching (P) The goal of using piecewise linear contrast stretching is to further enhance the contrast of the images before applying the histogram equalization techniques. The parameters used in this technique are setting r1 = 105, s1 = 0, r2= 210, and s2 = 255 which determine the range of pixel intensities to be stretched and the new intensity values for those pixels. To avoid over-enhancement, we apply a combination of piecewise linear contrast stretching with other techniques, such as CLAHE or GHE.

3.2.2 Image Classification Based on CNN. In the image classification stage, we use the popular CNN model: InceptionResNetV2 to classify the enhanced images. We compare the performance of this model using the different pre-processed image sets to identify the most effective model for histopathology image analysis. This model was chosen based on their strong performance in related image classification tasks.[16]

InceptionResNetV2 is a deep convolutional neural network architecture developed by Google for image classification tasks. It is an extension of the original InceptionNet architecture, which introduced the concept of using filters with different sizes in parallel to capture features at different scales. The model architecture uses a combination of 1x1, 3x3, and 5x5 convolutional filters in parallel to extract features from the input image. It also includes multiple Inception modules, which are composed of multiple parallel branches that process the input data in different ways. In addition, this model uses batch normalization, dropout, and max pooling layers to improve performance and reduce overfitting. The final layer of the network is a softmax activation layer that outputs the predicted class probabilities and the optimizer is an adam. We set batch size = 4, epoch = 100, and image size to 224x224 before training with this model.

Before training the model we used ImageDataGenerator for the augmentation of image data. It applies various transformations to the training images to increase the size of the dataset, and avoid overfitting. The data augmentation techniques that we used with ImageDataGenerator include horizontal_flip, vertical_flip, width_shift_range, and height_shift_range.

The data was split into training, validation, and test dataset. The validation dataset was composed of 5% of the images from the cemento-osseous dysplasia dataset, 45% of the images from the fibrous dysplasia dataset, and 75% of the images from the ossifying fibroma dataset. The test dataset contains 45 images divided into 15 images for each class.

3.3 Understanding Model By Using Grad-CAM

To understand how the CNN model interpretes the classification decision, Grad-CAM (Gradient-weighted Class Activation Mapping) is used to illustrate the important region in the image for a given class label. The results from Grad-CAM can be later used for improving accuracy and precision of the image classification model.

4 EXPERIMENTS

To compare the results of digital image processing techniques, we designed five experiments that used different combinations of image processing techniques as shown below:

(1) Classification on Grayscale image
(2) Classification on Grayscale image with CLAHE
(3) Classification on Grayscale image with GHE
(4) Classification on Grayscale image with piecewise linear contrast stretching and CLAHE
(5) Classification on Grayscale image with piecewise linear contrast stretching and GHE

The trained models were tested on the test dataset to evaluate the performance and study the features used by the model to classify histopathological images of 3 diseases using the Grad-CAM, which is used to visualize the features that provide insight into how the CNN is making decisions. To evaluate the model, we used accuracy which has the following equation:

$$Accuracy = (TP + TN)/(TP + FP + FN + TN) \qquad (1)$$

Where, True Positive (TP): the number of correctly predicted positive instances. False Positive (FP): the number of instances that were predicted as positive but are actually negative. True Negative (TN): the number of correctly predicted negative instances. False Negative (FN): the number of instances that were predicted as negative but are actually positive.

Lastly, Grad-CAM technique was used to provide insight into how the InceptionResNetV2 model made a decision on a class of lesions.

4.1 Results

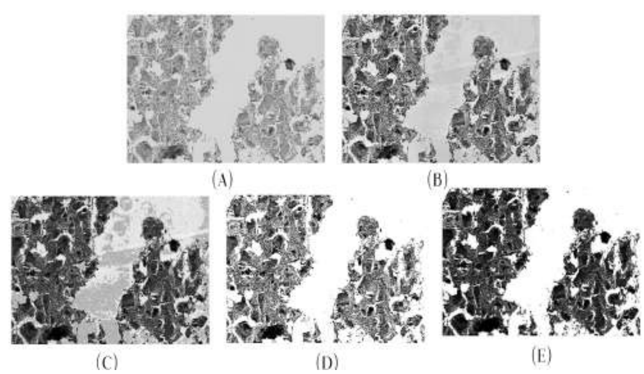

Figure 3: Examples of a histopathological image from five preprocessing techniques. (A) Grayscale conversion (B) CLAHE (C) GHE (D) P+CLAHE (E) P+GHE

The result of five experiments with different image preprocessing techniques is shown in Figure 3. Moreover, the comparison between image which applied piecewise linear contrast stretching before histogram equalization to solve the image problem of pixel breakage is shown in Figure 4.

These preprocessed images were used to perform image classification by using a deep convolutional neural network model: InceptionResNetV2 to evaluate

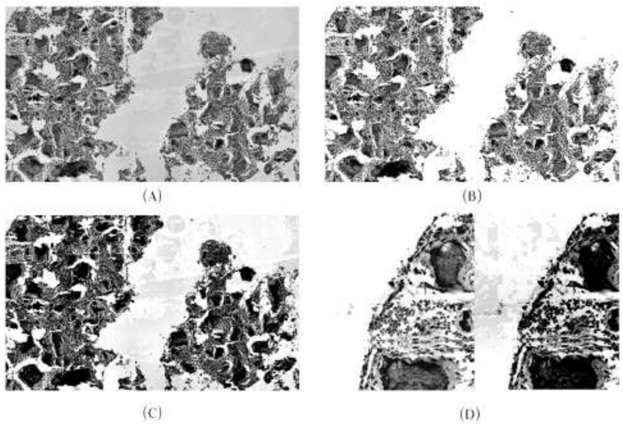

Figure 4: Compare the order of technique using piecewise linear contrast stretching before and after CLAHE. (A) Image using CLAHE. (B) Image using Piecewise linear contrast stretching before CLAHE. (C) Image using Piecewise linear contrast stretching after CLAHE. (D) Zoom image from (B) and (C)

Table 1: Comparison of accuracies obtained by different combinations of image techniques using InceptionResnetV2.

Preprocessing Techniques	Accuracy
Grayscale	0.84
CLAHE	0.87
GHE	0.80
P+CLAHE	0.82
P+GHE	**0.89**

their effectiveness based on classification accuracy. Table 1 shows that images with P+GHE techniques had the highest accuracy, followed by CLAHE, Grayscale, P+CLAHE, and GHE. The Inception-ResNetV2 performed best in image training using P+GHE, CLAHE, and Grayscale techniques, with accuracy rates of 89%, 87%, and 84%, respectively. Additionally, Grad-CAM was used to test the model's ability to identify disease types in the images, and it was found that the model could accurately identify the types of diseases in the images with about 80% accuracy based on the bottom part of the images (as shown in Figure5). Note that the result, when comparing Grad-CAM with the diagnosis of expert pathologists is consistent.

In addition, we further discuss image preprocessing techniques and their order in histopathology images. Image preprocessing techniques are essential for improving the quality of images for further analysis. In this experiment, we discuss the order of applying two specific techniques: piecewise linear contrast stretching and histogram equalization. According to Figure 4 we found that applying piecewise linear contrast stretching before histogram equalization can help reduce noise in the images. Therefore, we suggest that piecewise linear contrast stretching should be applied before histogram equalization in histopathology images.

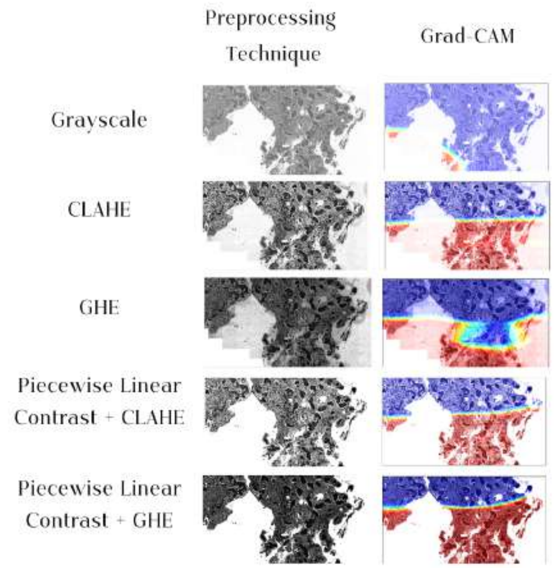

Figure 5: Example of Grad-CAM visualization by Inception-ResNetV2 model

Nevertheless, it is important to note that while these techniques can enhance contrast and reduce noise, they may not be beneficial for other image types.

Furthermore, when we use Grad-CAM to illustrate the important region in an image for a given class label is shown in Figure 5, we discovered that the model training with grayscale and GHE image sets only considers a small area of the image, and some areas do not have a tissue. Therefore, we assume that it may be a result of pixel breakage, causing model confusion. So, when we compared the image set from the fifth experiment using the P+GHE technique, which had the highest accuracy of 89%, we discovered that the model uses 80% of the image area to identify the types of diseases from grad-CAM visualization, which we discuss makes sense because all of the areas in the image are a lesion. Moreover, we consider the image area in which the model may be unused to decisions as a result of weight initialization.

5 CONCLUSION

Our study aims to find the appropriate solution to elevate the classification capability of the deep convolutional neural network. Although this research has limitations in terms of the size and quality of the histopathology images dataset for the trained model, we demonstrated that the limitations that affected the development and accuracy of the model can be overcome by using image preprocessing and image augmentation techniques. From the experiments, we found that piecewise linear contrast stretching combined with global histogram equalization technique produced the highest accuracy, followed by contrast limited adaptive histogram equalization and normal grayscale image, respectively. Interestingly, cemento-osseous dysplasia images are lacking in our dataset, but the disease was the most accurately achieved. This indicates the usefulness of image preprocessing combined with image augmentation. Moreover, when using the Grad-CAM technique for understanding the region of the image where the model is making decisions, the image characteristics extracted by the InceptionResNetV2 model were found to be consistent with the diagnostic of expert pathologists.

ACKNOWLEDGMENTS

This paper was supported by the Department of Oral Pathology, Faculty of Dentistry, Chulalongkorn University for image dataset; Big Data Experience (BX) Center, King Mongkut's University of Technology Thonburi (KMUTT), Thailand; and Princess Srisavangavadhana College of Medicine, Chulabhorn Royal Academy, Thailand.

REFERENCES

[1] Waldron, C. A., "Fibro-osseous lesions of the jaws". J Oral Maxillofac Surg 51, 828-835, doi:10.1016/s0278-2391(10)80097-7 (1993).
[2] Vered, M. & Wright, J. M., " Update from the 5th Edition of the World Health Organization Classification of Head and Neck Tumors Odontogenic and Maxillofacial Bone Tumours", Head Neck Pathol 16, 63-75, doi:10.1007/s12105-021-01404-7 (2022).
[3] Brannon, R. B. & Fowler, C. B., " Benign fibro-osseous lesions: a review of current concepts, " Adv Anat Pathol 8, 126-143, doi:10.1097/00125480-200105000-00002 (2001).
[4] karat Phattarataratip, Chanajit Pholjaroen and Prawwalee Tiranon, 2014, "A Clinicopathologic Analysis of 207 Cases of Benign Fibro-Osseous Lesions of the Jaws," Int J Surg Pathol, vol. 22, no. 4, pp. 326-333, 2014.
[5] Patil, S. et al.," Artificial Intelligence in the Diagnosis of Oral Diseases: Applications and Pitfalls," Diagnostics (Basel) 12, doi:10.3390/diagnostics12051029 (2022).
[6] Wu, Y. et al.," Recent Advances of Deep Learning for Computational Histopathology: Principles and Applications," Cancers (Basel) 14, doi:10.3390/cancers14051199 (2022).
[7] El Agouri, H. et al.," Assessment of deep learning algorithms to predict histopathological diagnosis of breast cancer: first Moroccan prospective study on a private dataset," BMC Res Notes 15, 66, doi:10.1186/s13104-022-05936-1 (2022).
[8] Yang, S. Y. et al., "Histopathology-Based Diagnosis of Oral Squamous Cell Carcinoma Using Deep Learning," J Dent Res, 220345221089858, doi:10.1177/00220345221089858 (2022).
[9] H. El Attar M. El Khannoussi H. El Agouri, M. Azizi and A.Ibrahimi, "Assessment of deep learning algorithms to predict histopathological diagnosis of breast cancer: first Moroccan prospective study on a private dataset," BMC Res Notes, doi: 10.1186/s13104-022-05936-1 (2022).
[10] Chao Liu Haosheng Tang, Guo Li and Donghai Huang, "Diagnosis of lymph node metastasis in head and neck squamous cell carcinoma using deep learning," Laryngoscope Investig Otolaryngol, doi: 10.1002/lio2.742 (2022).
[11] Şaban Öztürk and Bayram Akdemir, "Effects of Histopathological Image Preprocessing on ConvolutionalNeural Networks," Procedia Computer Science, vol. 132, pp. 396-403 (2018).
[12] Rafael C. Gonzalez, "Digital ImageProcessing Third Edition," University of Tennessee.
[13] Scharon Harding,What Is Grayscale? A Basic Definition, howpublished [online]. Available: https://www.tomshardware.com/reviews/glossary-grayscale-tracking-definition,5885.html
[14] Rachel Draelos, MD, PhD, Grad-CAM: Visual Explanations from Deep Networks [online]. Available: https://glassboxmedicine.com/2020/05/29/grad-cam-visual-explanations-from-deep-networks/
[15] Bharathi, Confusion Matrix for Multi-Class Classification [online]. Available: https://www.analyticsvidhya.com/blog/2021/06/confusion-matrix-for-multi-class-classification/
[16] Mohammad Reza Abbasniya, Sayed Ali Sheikholeslamzadeh a, Hamid Nasiri and Samaneh Emami, " Classification of Breast Tumors Based on Histopathology Images Using Deep Features and Ensemble of Gradient Boosting Methods," Computers and Electrical Engineering (2022).