International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7384
***Article Received**: 02 April 2023 **Revised**: 18 May 2023 **Accepted**: 20 June 2023*

Progressive Faster Residual Convolutional Neural Network for Improving Osteoarthritis of the Temporomandibular Joint Detection

Vijaya Kumar Krishnamoorthy[1], Santhi Baskaran[2]
[1]Department of Computer Science and Engineering
Puducherry Technological University
Puducherry, India.
mkvijayakumaramphd@gmail.com
[2]Department of Computer Science and Engineering
Puducherry Technological University
Puducherry, India
santhibaskaran@ptuniv.edu.in

Abstract— Osteoarthritis of the Temporomandibular Joint (TMJ-OA) is a chronic condition that affects the TMJ and is characterized by the progressive degeneration of the internal surfaces of the joint. Several deep learning models were adopted for identifying the TMJ-OA from the panoramic dental X-ray scans. Amongst, an Optimized Generative Adversarial Network (OGAN) with Faster Residual Convolutional Neural Network (FRCNN) produces more synthetic images to train the FRCNN for recognizing TMJ-OA cases. But, its accuracy was comparatively low while recognizing Region-of-Interest (RoI) from the panoramic scans that have analogous objects. Hence in this paper, an OGAN with a Progressive FRCNN (OGAN-PFRCNN) model is proposed, which enhances the FRCNN by integrating the Feature Pyramid Network (FPN) and RoI-grid attention strategy for TMJ-OA identification. First, the training images are fed to the ResNet101 for feature mining, which provides Multi-Scale Feature Map (MSFM) from the dental panoramic scans. Those features are then passed to the FPN with the RoI-grid attention strategy, which encodes richer characteristics by considering standard attention and graph-based point functions into a combined formulation. Then, those characteristics are fused at various levels to get a useful MSFM, which increases the network efficiency significantly. Moreover, such a Feature Map (FMap) is used to train the PFRCNN model, which is later applied to recognize the test scans into either healthy or TMJ-OA. At last, the testing outcomes show that the OGAN-PFRCNN attains 96.2% accuracy on the panoramic dental X-ray database compared to the FRCNN model.

Keywords- Temporomandibular joint, Osteoarthritis, Panoramic dental image, OGAN-FRCNN, Feature pyramid network, RoI-grid attention.

I. INTRODUCTION

The foremost recurrent form of osteoporosis, which impacts the TMJ is osteoarthritis (OA). It alters the osteoarticular area of the mandibular condyle and fossa. The extreme motorized strain on joint tissue is the common reason for OA. Subarticular bone resorption develops if concurrent stress is supplied to the articular surface (chondromalacia). Radioactivity osteoarthritis is triggered by the regular bone shift that causes the decrease of the subchondral cortical layer and bone degradation [1-3].

TMJ-OA is investigated using health records, clinical diagnostics, and radiographic evaluation. TMJ-OA manifests in the form of limited lower jaw mobility due to anxiety, crepitus, and local paraspinal ache in the joint promotion. When radiography exposes functional bone shift, OA is recognized [4-6]. Orthopantomography can be utilized to quantify the condylar and Ramal asymmetry of the mandible in people with Juvenile Idiopathic Arthritis (JIA). Multiple wide-ranging TMJ illnesses and analyses have occurred that might be hurting or not [7, 8].

When bone anomalies in the TMJ are discovered, panoramic radiographs are engaged in the beginning levels of diagnosis. Nonetheless, as the TMJ surrounds minuscule bone structures at the joint site and the joint is concealed by the big skull, simple imagery is challenging to reveal bone alterations [9-11]. Secondly, anatomical variations or pathologies in the TMJ are recurrently ignored in uncomplicated examinations owing to poorly dematerialized bone tissue in the beginning phases of OA. As a result, assessing panoramic radiographs needs the use of competent specialists with substantial quantifiable skill, and additional radiography should be obtained as obligatory. Distributing panoramic pictures to skilled diagnostic experts and acquiring outcomes is time-consuming while health practitioners who can accurately diagnose OA solely on panoramic radiographs are not around.

Besides, the process of submitting panoramic radiographs and waiting for diagnostic results must be repeated since a doctor cannot rapidly assess the treatment condition of osteoarthritis.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7384
***Article Received**: 02 April 2023 **Revised**: 18 May 2023 **Accepted**: 20 June 2023*

To address these problems, an AI-based methodology for autonomously treating TMJ OA has been developed [12–14]. Numerous advancements in the AI paradigm have offered several X-ray image analysis algorithms [15]. The use of CNN technology has been demonstrated for a variety of tasks, including the extraction of specific X-ray scan regions and the detection of anomalies [16–18]. Numerous studies on dental X-ray image analysis have been conducted, including those on osteoporosis, sinusitis, and tooth detection using panoramic picture analysis. Investigations that use panoramic X-ray analysis are useless, nevertheless. Thus, by classifying panoramic dental X-ray images using image recognition techniques, Kim et al. [19] created a model to identify the mandibular condyle. The panoramic images were first collected retrospectively, and two different models were used: (1) the first model, called FRCNN, has been adopted to identify the TMJ-OA and adjacent structural forms (such as joint fossa and condyle), and (2) the other model, called CNN, has been adopted to predict whether the identified structure covers any abnormality based on the contour of the TMJ.

Additionally, it was done to fine-tune CNN models like VGG16, ResNet, and Inception to predict the presence and absence of TMJ-OA. The number of samples was quite small, which harmed the model's ability to accurately recognize objects. So, an OGAN was designed [20], which generates synthetic dental panoramic X-ray scans for TMJ-OA detection. This GAN model utilizes the generator to create synthetic panoramic dental images while training the discriminator to determine if the produced images were fake or real. Also, the most effective hyperparameters of this GAN were chosen using the Elephant Herding Optimization (EHO) technique according to the clan and separation factors. The real database was then updated with the newly developed synthetic panoramic photos to obtain the training and test collections. Later, the learning pictures were given to the FRCNN to extract the condylar region and identify anomalies from those images. The trained FRCNN model was further validated using test pictures for TMJ-OA identification. On the other hand, the FRCNN has relatively poor detection accuracy when used to identify targets in panoramic images that contain identical objects. Also, the detection efficiency was influenced by the noise and identical background textures in the images.

Therefore, this article develops an OGAN-PFRCNN model for TMJ-OA identification. In this model, the standard FRCNN model is enhanced by adopting the FPN and RoI-grid attention strategy to obtain MSFM with rich contextual data. The attention strategy is utilized to enhance the FPN and recalibrate the contribution units of feature channels. First, the training images are fed to the ResNet101 for feature mining, which provides MSFM from the dental panoramic scans. Those features are then passed to the FPN with the RoI-grid attention strategy, which encodes richer characteristics by considering standard attention and graph-based point functions into a combined formulation. Then, those characteristics are fused at various levels to get a useful MSFM, which increases the network efficiency significantly. Moreover, such an FMap is used to train the PFRCNN model, which is later applied to recognize the test scans into either healthy or TMJ-OA. Thus, the recognition accuracy is increased by solving noise and complex background textures from the panoramic scans.

The residual portion of this article is planned as follows: Section II reviews the prior studies on the identification of TMJ-OA. The OGAN-PFRCNN model is explained in Section III and its effectiveness is demonstrated in Section IV. Section V provides a summary of the work and provides recommendations for further work

II. TYPE LITERATURE SURVEY

With the help of the DetectNet structure and the DIGITS, a deep learning framework [21] was created to recognize and classify radiolucent cancers in the mandible on panoramic radiography images. The first step was to gather panoramic images of people with mandibular radiolucent tumors. Next, RoI coordinates and lesion tags were created. As well, the DetectNet was fed training data with relevant tags to identify and classify radiolucent lesions of the mandible. But, more samples were needed to accurately recognize tiny tumors.

Al Kheraif et al. [22] presented data mining schemes to recognize dental disorders from radiographic 2D dental scans. In the pre-processing, a histogram using a dynamic scheme was utilized to stretch the contrast and equalize the brightness of the images. Then, those images were split into the oral cavity and its tissues using a hybrid graph cut method for differentiating the foreground teeth and the areas of background bones. Also, statistical characteristics were obtained and learned by CNN to classify dental disorders. But, it needs advanced deep learning schemes to enhance the automatic classification of dental disorders.

Using artificial intelligence, a computer vision system [23] was modeled to identify and categorize various dental restorations in panoramic photographs. The RoI with maxillary and mandibular alveolar ridges was initially found by cropping panoramic pictures. The reconstructions were then divided by a local dynamic threshold, and statistical characteristics were produced according to the contour and grayscale distribution. To further categorize the features into the various restoration kinds, a cubic SVM with Error-Correcting Output Code (ECOC) was used. Conversely, the training efficiency was less because of inadequate images.

Muresan et al. [24] developed a new automated teeth identification and dental symptoms categorization from panoramic X-ray images. Initially, the CNN was trained by the

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7384
***Article Received**: 02 April 2023 **Revised**: 18 May 2023 **Accepted**: 20 June 2023*

labeled samples to get the semantic segregation data. Then, several image processing methods were applied to segregate and refine the bounding boxes related to teeth recognition. Further, all tooth samples were annotated and the symptom impacting it was recognized by the histogram-based majority voting within the recognized RoI. But, it needs more semantic classes to further increase the accuracy.

Viloria et al. [25] presented the prediction of mandibular morphology based on the Artificial Neural Network (ANN) utilizing Cranio-maxillary measures in posterior-anterior radiographs. But, its complexity was high while increasing the number of layers for large-scale databases. Leo and Reddy [26] developed a novel Hybrid Neural Network (HNN) model to recognize dental caries based on their class and categorize the caries regions. First, the dental images were pre-processed and segregated into their constituent pixels. After that, various characteristics were obtained and learned by the HNN to categorize the caries classes. But, the number of images was very limited, which impacts the training efficiency.

Using Surface Electromyography (SEMG), a variety of machine learning classifiers [27] was suggested to diagnose TMJ disorder. The SEMG records were gathered, and various common traits were retrieved. Additionally, the most important variables were chosen and fed to several machine-learning categorizers to divide them into normal and TMJ categories. But, it needs deep learning algorithms to increase efficiency while classifying more images. Several deep learning architectures, including DenseNet121, VGG16, InceptionV3, and ResNet50, were suggested [28] for classifying the kind of canine impaction from panoramic dental pictures. However, the number of labeled images was insufficient. Also, low-resolution and poor-quality images were needed to preprocess for high accuracy.

To predict TMJ-OA of Magnetic Resonance Imaging (MRI), Ito et al. [29] created a computerized articular disc detection and partition model. The Detection for Displaced Articular DISC based on CNN (3DiscNet), U-Net, and SegNet-based semantic partition techniques were used to separate the TMJ articular discs from various MRI images. But, robustness was ineffective and only a small group of specialists created the Ground Truth (GT) scans.

III. PROPOSED METHODOLOGY

In this section, the OGAN-PFRCNN model is explained in brief for TMJ-OA identification. A pipeline of this study is illustrated in Fig.1. Initially, a panoramic dental X-ray image corpus is acquired and expanded using the OGAN. Afterward, such scans are split into learning and test sets. The learning sets are taken to train the new PFRCNN categorizer, whereas the test sets are taken to analyze the efficiency of a trained PFRCNN for mandibular condylar detection and TMJ-OA recognition.

A. Progressive FRCNN Model

To increase the accuracy of recognizing the TMJ-OA, a new FRCNN model called PFRCNN is designed, whose architecture is depicted in Fig. 2.

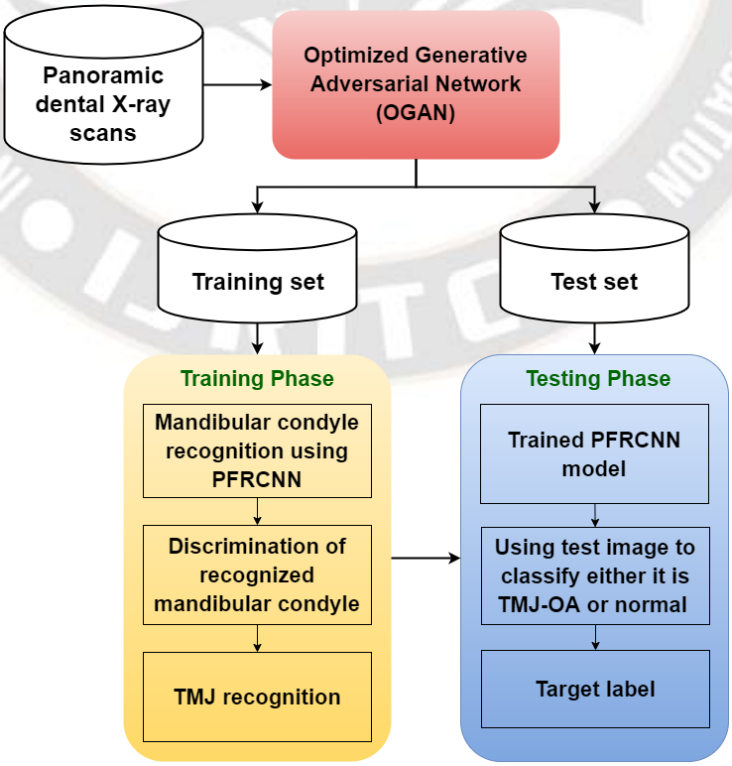

Figure 1. Pipeline of the Presented Study

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7384
***Article Received**: 02 April 2023 **Revised**: 18 May 2023 **Accepted**: 20 June 2023*

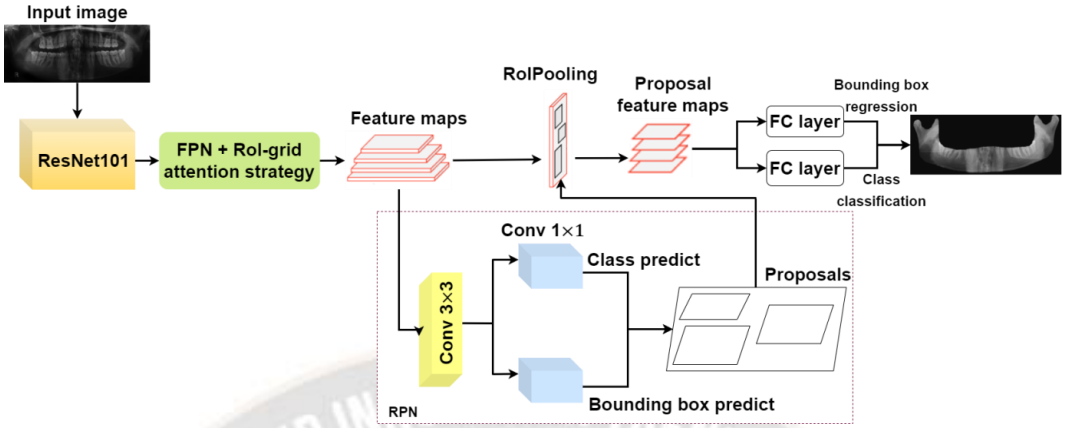

Figure 2. Structure of PFRCNN Model

In this model, after augmentation, the panoramic X-ray scans are given to the ResNet101 for capturing multilevel characteristics. Those multilevel characteristics are fed to the FPN with the RoI-grid attention strategy for merging characteristics and creating MSFM with rich contextual data. Then, the FMap is given to the Region Proposal Network (RPN) to produce Region Proposals (RPs). Further, the FMaps and RPs are transferred to the RoIPooling to produce final proposal FMaps, which are passed to the Fully Connected (FC) to recognize the mandibular condyle and get the accurate location of the target bounding boxes (i.e., to decide whether the abnormal condyle is properly recognized or not).

This new model demonstrates the below enhancements over the standard FRCNN:

- ResNet101 is employed as the feature mining block to improve the ability to capture characteristics from the dental X-ray scans and the FPN is used to merge the multilevel characteristics from the ResNet101, which creates the FMaps with rich semantic and position data;

To enhance the FPN, the RoI-grid attention strategy is introduced, which encodes rich data from sparse positions by combining standard attention and graph-based point functions into a single model.

1) RoI-Grid Attention Strategy

RoI-grid attention is adopted, which integrates the graph and attention-based point functions into the single model. This strategy can act as a good alternate for usual pooling-based functions in the FRCNN. Consider p_{grid} is the coordinate of a RoI-grid point and p_i, f_i are the coordinate and the related feature vector of i^{th} Point of Interest (PoI) near p_{grid}. RoI feature mining intends to capture the relevant feature vector f_{grid} of the RoI-grid point p_{grid} utilizing the data of adjacent p_i and f_i.

a) Pooling-based functions: These are widely used for RoI feature mining. The adjacent feature f_i and the relative position $p_i - p_{grid}$ initially go through a Multi-Layer Perceptron (MLP) layer to get the converted feature vector: $V^i = MLP([f_i, p_i - p_{grid}])$, where $[\cdot]$ is the fusion operation and then a max-pooling process is employed on each converted feature V to find the RoI-grid feature f_{grid}^{pool}:

$$f_{grid}^{pool} = \max_{i \in \Omega(r)} pool(V^i) \quad (1)$$

In (1), $\Omega(r)$ defines PoI within the predetermined radius r of p_{grid}. The pooling-based functions merely concentrate on the highest channel response and this loses a great amount of semantic and position data.

b) Graph-based functions: These map the grid points and PoI as a graph. The graph node i denotes the converted feature of f_i: $V^i = MLP(f_i)$ and the edge Q_{pos}^i is devised as a linear estimation of the position variances between 2 nodes: $Q_{pos}^i = Linear(p_i - p_{grid})$. For the graph node of p_{grid}, the feature f_{grid}^{graph} is gathered from neighboring nodes by a weighted fusion process. The typical formula is defined as the similar notations in (1).

$$f_{grid}^{graph} = \sum_{i \in \Omega(r)} W(Q_{pos}^i) \odot V^i \quad (2)$$

In (2), the function $W(\cdot)$ estimates the graph boundary inserting into the scalar or vector weight space and $\odot$ is the Hadamard product product between trained weights and nodes.

c) Attention-based functions: These are employed on the grid points and PoI. Q_{pos}^i in (2) is considered as the query embedding from p_{grid} to p_i. Also, V^i represents the value embedding acquired from f_i as Equation 4. The key embedding K^i is defined by $K^i = Linear(f_i)$. So, typical attention is defined by

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7384
***Article Received**: 02 April 2023 **Revised**: 18 May 2023 **Accepted**: 20 June 2023*

$$f_{grid}^{atten} = \sum_{i\in\Omega(r)} W\left(Q_{pos}^{i}K^{i}\right)\odot V^{i} \quad (3)$$

Additional regularization operation i.e., softmax is employed in $W(\cdot)$. This typical attention and (3) are extended by considering the point transformer as:

$$f_{grid}^{tr} = \sum_{i\in\Omega(r)} W\left(K^{i} + Q_{pos}^{i}\right)\odot\left(V^{i} + Q_{pos}^{i}\right) \quad (4)$$

d) RoI-grid attention: In this method, the structural similarity of (2)– (4) is analyzed. It is observed that such formulas contain familiar fundamental components and functions. So, it is usual to combine such formulas into a single model with gated functions. This novel formula is called RoI-grid attention:

$$f_{grid} = \sum_{i\in\Omega(r)} W\left(\sigma_k K^{i} + \sigma_q Q_{pos}^{i} + \sigma_{qk} Q_{pos}^{i} K^{i}\right)\odot\left(V^{i} + \sigma_v Q_{pos}^{i}\right) \quad (5)$$

In (5) , σ_* denotes a learnable gated function that is executed by the linear estimation of the relevant embedding with a sigmoid activation result.

Fig. 3 portrays the design of RoI-grid attention, which adopts trainable gated functions σ_* to adaptively choose the attention elements and creates a single solution that involves the standard graph and attention functions.

RoI-grid attention is a global model integrating graph and attention-based processes. The graph function, i.e. (2) is determined from Equation 5 if $\sigma_q = 1, \sigma_k = 0, \sigma_{qk} = 0$ and $\sigma_v = 0$. Likewise, the typical attention, i.e. (3) is determined if $\sigma_q = 0, \sigma_k = 0, \sigma_{qk} = 1$ and $\sigma_v = 0$ or point transformer, i.e. (4) is determined if $\sigma_q = 1, \sigma_k = 1, \sigma_{qk} = 0$ and $\sigma_v = 1$.
It is a promising process for RoI feature mining. Using σ_*, RoI-grid attention can capture which point is relevant to the RoI-grid points, from the position data Q_{pos} and the semantic data K, as well as, their mixture $Q_{pos}K$, dynamically. Using σ_v, RoI-grid attention learns to balance the fraction of position characteristics Q_{pos} and semantic characteristics V utilized in the feature fusion. In contrast to the pooling-based techniques, merely some linear estimation units are stacked in RoI-grid attention that ensures the generalization ability. Substituting pooling-based functions with RoI-grid attention substantially increases the recognition efficiency.

B) FPN Module

The classical FRCNN obtains deep characteristics, which

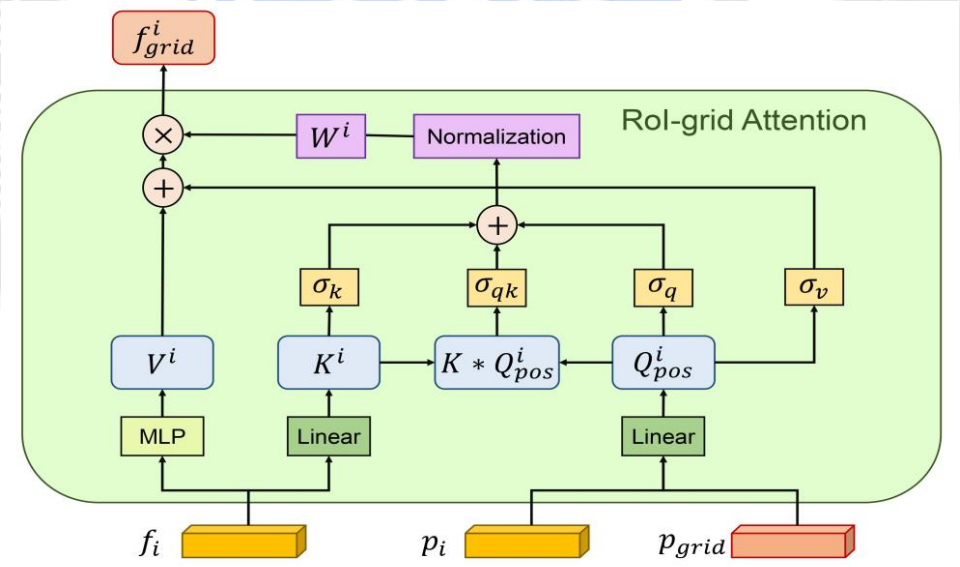

Figure 3. Design of RoI-grid Attention

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7384
***Article Received**: 02 April 2023 **Revised**: 18 May 2023 **Accepted**: 20 June 2023*

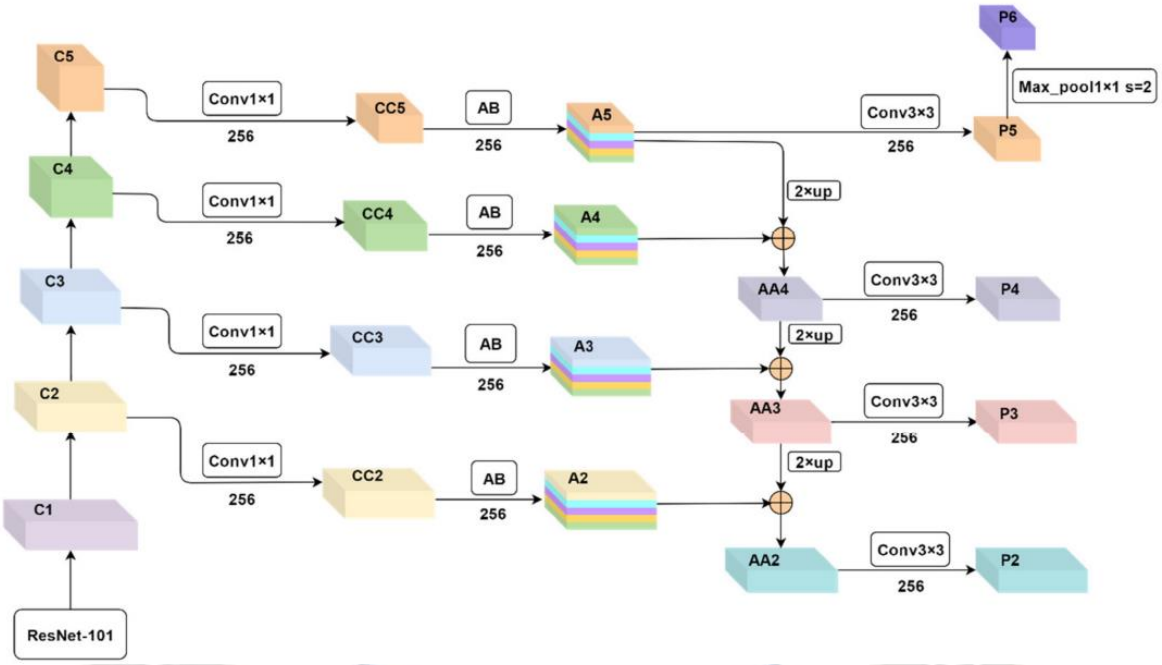

Figure 4. Structure of IFPN (AB is the channel attention strategy module, 2×up is 2-times upsampling, 256 is the quantity of final channels and ⊕ is element-wise summation)

contain rich semantic data but position data is lost. In mandibular condyle region identification, position data is vital. On the other hand, shallow characteristics contain weak semantic data but are prone to position data. So, the FPN is applied to merge deep and shallow MSFM to completely use the feature semantics and position data so that the model efficiency is further increased. Performing ResNet101 to enhance the feature mining abilities, a channel attention strategy is also introduced and an Improved FPN (IFPN) is developed that merges characteristics at various levels to get a useful MSFM; thus significantly increasing the recognition accuracy of the FRCNN. An IFPN structure is depicted in Fig. 4.For example, in Fig. 4, ResNet101 is applied as the feature mining network. $C1, C2, C3, C4$, and $C5$ in the IFPN are applied to mine various levels of characteristics wherein $C2 - C5$ performs feature concatenation. The quantity of channels is 256, 512, 1024 and 2048, correspondingly. So, 1×1 convolution process is utilized to minimize the size of $C2 - C5$. The resultant outcomes are $CC2 - CC5$ and the quantity of channels is 256. The channel attention strategy is applied to determine the contribution weight of all channels of $CC2 - CC5$ that are reassigned based on their weight. Therefore, the contributions of relevant characteristic channels are improved. The resultant outcomes are $A2 - A5$ and the quantity of channels is 256. While executing feature concatenation through IFPN, pixels related to the characteristics of various levels are merged.

The amount of rows and columns in the feature unit should match the number of feature channels. So, the nearest interpolation scheme is employed to conduct 2-times upsampling on $A3, A4$, and $A5$. After that, element-wise summation is implemented with $A2, A3$, and $A4$ to achieve the level-by-level feature concatenation, which provides $AA2, AA3$, and $AA4$, as well as the quantity of channels, is 256. A 3×3 convolution process is applied to $A5$, which provides $P5$ and the quantity of channels is 256. Max-pool of 1×1 is conducted on $P5$, the stride is allocated to 2, the result is $P6$ and the quantity of channels is 256. A 3×3 convolution process is conducted on $AA2, AA3$, and $AA4$, which provides $P2, P3$, and $P4$ and the quantity of channels is 256. The outcome through FPN is the MSFM, i.e. $\{P2, P3, P4, P5, P6\}$.

C) RPN Module

The most well-known contribution of FRCNN is the RPN that utilizes a CNN rather than the classical discriminatory search scheme to produce candidate areas, so considerably enhancing the model accuracy. The RPN is utilized to create RPs. In this work, the MSFM $\{P2, P3, P4, P5, P6\}$ from the FPN are utilized to produce RPs. The regions of anchors (ARs) for MSFM are assigned to $\{32^2, 64^2, 128^2, 256^2, 512^2\}$ and the AR aspect proportions are assigned to $\{1{:}\,2, 1{:}\,1, 2{:}\,1\}$.

The FMaps are fed to the RoIPooling with RPs such as $\{P2, P3, P4, P5\}$. The selection of the FMap with the most proper scale for RP is based on the below formula:

$$k = k_0 + \log_2\left(\sqrt{wh}/_H\right) \qquad (6)$$

In (6), k is the level of FMap related to the RP, k_0 is assigned to the maximum level of FMaps, w and h are the width

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7384
***Article Received**: 02 April 2023 **Revised**: 18 May 2023 **Accepted**: 20 June 2023*

and height of the RP, correspondingly, and H is the input height after resizing the image slices. This is highly effective since a high-dimension RP can relate to a high-level FMap and produce a RP that may recognize large targets effectively. Likewise, a low-dimension RP relates to the low-level FMap and produces the RP that may recognize tiny targets efficiently.

D) Loss Factor

The loss factor is determined by

$$L(p_i,t_i) = \frac{1}{N_{cls}}\sum_i L_{cls}(p_i,p_i^*) + \alpha\frac{1}{N_{reg}}\sum_i p_i^* L_{reg}(t_i,t_i^*) \quad (7)$$

In (7), N_{cls} is the amount of ARs in the mini-batch, N_{reg} is the amount of AR positions, α is the weight balance variable assigned to 10, i is the index of an AR in a mini-batch. Also, p_i is the estimated categorization possibility of the AR.

Particularly, if the AR is positive, $p_i^* = 1$ and if it is negative $p_i^* = 0$. Further, ARs that satisfy the below criteria are termed positive:

- The AR takes the maximum Intersection-Over-Union (IOU) similarity with the GT box; or
- The IOU similarity of the AR with the GT box is greater than 0.7.

On the other hand, if the IOU similarity of the AR with the GT box is < 0.3, the AR is termed negative (8) - (10)

.

$$L_{cls}(p_i,p_i^*) = -\log[p_i p_i^* + (1-p_i)(1-p_i^*)] \quad (8)$$

$$L_{reg}(t_i,t_i^*) = \sum_{i\in\{x,y,w,h\}} Smooth_{L1}(t_i,t_i^*) \quad (9)$$

$$Smooth_{L1}(x) = \begin{cases} 0.5x^2, & if\ |x| < 1 \\ |x| - 0.5, & Or\ else \end{cases} \quad (10)$$

The parameterization of 4 coordinates for the bounding box regression is described by (11a) – (12b).

$$t_x = \frac{(x-x_a)}{w_a}, t_y = \frac{(y-y_a)}{h_a} \quad (11a)$$

$$t_w = \log\left(\frac{w}{w_a}\right), t_h = log\left(\frac{h}{h_a}\right) \quad (11b)$$

$$t_x^* = \frac{(x^*-x_a)}{w_a}, t_y^* = \frac{(y^*-y_a)}{h_a} \quad (12a)$$

$$t_w^* = log\left(\frac{w^*}{w_a}\right), t_h^* = log\left(\frac{h^*}{h_a}\right) \quad (12b)$$

Here, x and y indicate the coordinates of the centroid of the bounding box, w and h are the width and height of the bounding box, correspondingly. Also, x, x_a and x^* are the estimated box, AR box and GT box, correspondingly, equal to y, w and h.

E. Learning and Fine-tuning

Because the FRCNN is applied as the baseline network, the hyperparameters are assigned. In this model, a transfer learning method is adopted and the base network is ResNet101, which is set with its pre-learned weights on ImageNet. Also, each additional layer is set with k aiming normal. This PFRCNN has trained 80 epochs of the learning scans.

The Stochastic Gradient Descent (SGD) is applied as the fine-tuner, the primary training rate is assigned to 0.01, momentum is 0.09, weight decay is 0.0001 and the batch size is 4. Thus, the new PFRCNN is trained and utilized to categorize the test photographs into healthy and TMJ-OA with proper identification of mandibular condyle regions.

IV. EXPERIMENTAL RESULTS

In this part, the efficiency of the OGAN-PFRCNN model is assessed by executing it in MATLAB 2019b. In this scrutiny, a total of 116 panoramic dental X-ray scans are acquired from https://data.mendeley.com/datasets/hxt48yk462/2. Those scans are augmented into 6000 scans by the OGAN to create the training and test collections. Amongst, 65% of scans are applied for learning and the residual 35% are applied for testing. Also, the obtained efficiency is compared with the existing models: FRCNN [19], OGAN-FRCNN [20], DetectNet [21], ANN [25], HNN [26] and DenseNet121 [28] regarding the below metrics:

- Accuracy: It is the proportion of exact recognition over the total samples tested (13) .

$$Accuracy = \frac{True\ Positive\ (TP) + True\ Negative\ (TN)}{TP+TN+False\ Positive\ (FP)+False\ Negative\ (FN)} \quad (13)$$

In (13), the quantity of normal samples properly categorized as normal is TP, while the quantity of TMJ-OA

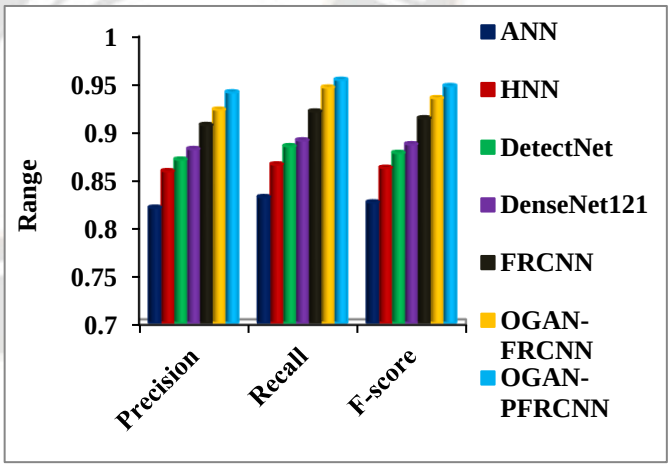

Figure 5. Assessment of Precision, Recall & F-measure for Proposed and Existing TMJ-OA Recognition Model

samples properly categorized as TMJ-OA is TN. Also, FP is the quantity of TMJ-OA samples improperly categorized as normal, whereas FN is the quantity of normal samples improperly categorized as TMJ-OA.

- Precision: It determines the correctly recognized labels at TP and FP rates (14)

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7384
***Article Received**: 02 April 2023 **Revised**: 18 May 2023 **Accepted**: 20 June 2023*

$$Precision = \frac{TP}{TP+FP} \quad (14)$$

- Recall: It is the proportion of labels, which are correctly recognized at TP and FN rates (15)

$$Recall = \frac{TP}{TP+FN} \quad (15)$$

- F-score (F): It is measured as (16)

$$F = \frac{2\times Precision\times Recall}{Precision+Recall} \quad (16)$$

Fig. 5 shows the effectiveness of the different classification models implemented on the dental panoramic X-ray corpus to recognize the mandibular condyle. It states that the success rate of the OGAN-PFRCNN model regarding precision, recall and f-score is higher than that of other existing models due to the capturing both semantic and position data during the feature mining process. As a result, it defines that the precision values achieved by the OGAN-PFRCNN are 14.6% superior to the ANN, 9.6% superior to the HNN, 8% superior to the DetectNet, 6.7% superior to the DenseNet121, 3.8% superior to the FRCNN and 2% superior to the OGAN-FRCNN. The recall values achieved using the OGAN-PFRCNN are 14.7% greater than the ANN, 10.2% greater than the HNN, 7.8% greater than the DetectNet, 7.1% greater than the DenseNet121, 3.6% greater than the FRCNN and 0.8% greater than the OGAN-FRCNN. Likewise, the f-score values achieved using the OGAN-PFRCNN are 14.7% higher than the ANN, 9.9% higher than the HNN, 7.9% higher than the DetectNet, 6.8% higher than the DenseNet121, 3.7% higher than the FRCNN and 1.3% higher than the OGAN-FRCNN models.

Fig. 6 demonstrates the accuracy of distinct models executed on the panoramic dental X-ray corpus to recognize the mandibular condyle region and categorize the TMJ-OA. It notices that the accuracy of the OGAN-PFRCNN is 16.8%

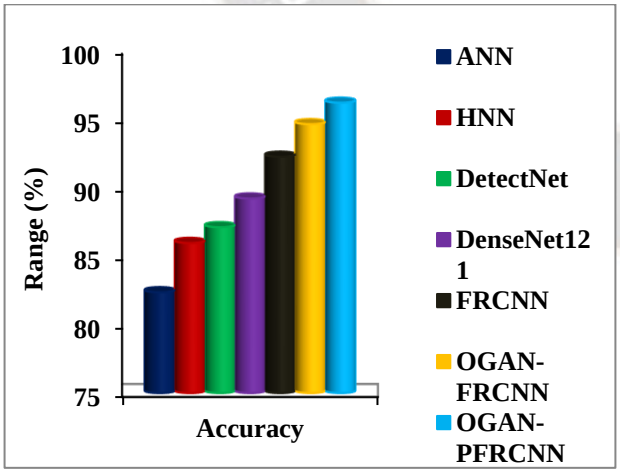

Figure 6. Assessment of Accuracy for Proposed and Existing TMJ-OA Recognition Models

greater than the ANN, 11.9% greater than the HNN, 10.4% greater than the DetectNet, 7.8% greater than the DenseNet121, 4.3% greater than the FRCNN and 1.6% greater than the OGAN-FRCNN models. From these analyses, it is realized that the OGAN-PFRCNN model can increase the accuracy of recognizing the mandibular condylar and classifying TMJ-OA effectively.

V. CONCLUSION

In this study, the OGAN-PFRCNN model was developed for recognizing the mandibular condyle regions and TMJ-OA from the panoramic dental X-ray scans. Primarily, the dental X-ray scan corpus was obtained and extended by the OGAN to form learning and test sets. Afterward, the learning sets were utilized to train the PRFCNN, which involves the FPN with RoI-grid attention strategy and RPN modules to extract MSFM and RPs. Such FMaps and RPs were further learned by the FC unit to recognize the target class and the accurate mandibular condyle region (i.e., the location of the target bounding box). Finally, the investigational results proved that the OGAN-PFRCNN model on panoramic dental X-ray scans has 96.2% accuracy for TMJ-OA classification and mandibular condyle region recognition contrasted to the stand-of-the-art models.

REFERENCES

[1] A. Kui, S. Buduru, M. Pacurar, I. Socaciu, A. Berar, S. Balhuc, A. Ciurea and M. Negucioiu, “Prevalence of disc displacements with reduction in temporo-mandibular joint in young people – a preliminary study,” Romanian Journal of Oral Rehabilitation, vol. 12, no.2,pp.198-204,2020.

[2] B. Lund, M. Ulmner, T. Bjørnland, T. Berge, H. Olsen-Bergem and A. Rosèn, “A disease-focused view on the temporomandibular joint using a Delphi-guided process,” Journal of Oral Science, vol.62,no.1,pp. 1-8,2020.

[3] E. Sofyanti, T. Boel and D. Satria, “Special investigation procedure of postural disorder related to developmental mandibular asymmetry: a review,” In Proceedings of the 2nd International Conference on Tropical Medicine and Infectious Disease, pp. 75-79,2020.

[4] E. A. Al-Moraissi, P. C. R. Conti, A. Alyahya, K. Alkebsi, A. Elsharkawy and N. Christidis, “ The hierarchy of different treatments for myogenous temporomandibular disorders: a systematic review and network meta-analysis of randomized clinical trials, “Oral and Maxillofacial Surgery, pp. 1-15, 2021.

[5] R. Rongo, E. Ekberg, I. M. Nilsson, A. Al-Khotani, P. Alstergren, P. C. R. Conti and A. Michelotti, “ Diagnostic criteria for temporomandibular disorders (DC/TMD) for children and adolescents: An international Delphi study—Part 1-Development of Axis I,” Journal of Oral Rehabilitation, vol.48, no, 7, pp. 836-845,2021.

[6] R. Rongo, E. Ekberg, I. M. Nilsson, A. Al-Khotani, P. Alstergren, P. C. Rodrigues Conti and A. Michelotti, “ Diagnostic criteria for temporomandibular disorders in children and adolescents: An international Delphi study-Part 2-Development of Axis II,” Journal of Oral Rehabilitation,vol. 49, no. 5, pp. 541-552 , 2022.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7384
***Article Received**: 02 April 2023 **Revised**: 18 May 2023 **Accepted**: 20 June 2023*

[7] D. T. S. Li and Y. Y. Leung, "Temporomandibular Disorders: Current Concepts and Controversies in Diagnosis and Management," Diagnostics,vol. 11,no. 3, pp. 1-15, 2021.

[8] Q. Auh and Y. H. Lee, "Can the arthralgia of temporomandibular joint cause referred pain?," Journal of Oral Medicine and Pain, vol.47,no 1, pp. 72-73.2022.

[9] C. Lee, K. J. Jeon, S. S. Han, Y. H. Kim, Y. J. Choi, A. Lee and J. H. Choi, " CT-like MRI using the zero-TE technique for osseous changes of the TMJ," Dentomaxillofacial Radiology, vol. 49, no 3,pp. 1-7, 2020.

[10] D. Schnabl, A. K. Rottler, W. Schupp, W. Boisserée and I. Grunert, "CBCT and MRT imaging in patients clinically diagnosed with temporomandibular joint arthralgia," Heliyon. vol. 4, no. 6, p.e00641, 2018.

[11] Sarangi, D. P. K. . (2022). Malicious Attacks Detection Using Trust Node Centric Weight Management Algorithm in Vehicular Platoon. Research Journal of Computer Systems and Engineering, 3(1), 56–61. Retrieved from https://technicaljournals.org/RJCSE/index.php/journal/article/view/42

[12] S. M. Gharavi, Y. Qiao, A. Faghihimehr and J. Vossen, "Imaging of the temporomandibular joint," Diagnostics, vol. 12,no.4, p. 1006, 2022.

[13] A. Thurzo, W. Urbanová, B. Novák, L. Czako, T. Siebert, P. Stano, S. Mareková, G. Fountoulaki, H. Kosnácová and I. Varga, "Where is the artificial intelligence applied in dentistry? systematic review and literature analysis," Healthcare, vol. 10, no.7 p.1269, 2022.

[14] S. B. Khanagar, A. Al-Ehaideb, P. C. Maganur, S. Vishwanathaiah, S. Patil, H. A., Baeshen, S. Bhandi, "Developments, application, and performance of artificial intelligence in dentistry–a systematic review," Journal of dental sciences,vol. 16, no. 1, pp. 508-522, 2021.

[15] M. M. Meghil, P. Rajpurohit, M. E. Awad, J. McKee, L. A. Shahoumi and M. Ghaly, "Artificial intelligence in dentistry," Dentistry Review, vol. 2, no.1, p. 1100009, 2022.

[16] K. Panetta, R. Rajendran, A. Ramesh, S. P. Rao and S. Agaian, "Tufts dental database: a multimodal panoramic X-ray dataset for benchmarking diagnostic systems," IEEE Journal of Biomedical and Health Informatics,vol. 26,no. 4, pp. 1650-1659, 2021.

[17] S. Patil, S. Albogami, J. Hosmani, S. Mujoo, M. A. Kamil, M. A. Mansour and S. S. Ahmed, "Artificial intelligence in the diagnosis of oral diseases: applications and pitfalls," Diagnostics, vol.12, no. 5, p.1029, 2022.

[18] R. H. Putra, C. Doi, N. Yoda, E. R. Astuti and K. Sasaki, "Current applications and development of artificial intelligence for digital dental radiography," Dentomaxillofacial Radiology,vol. 51, no. 1, p.20210197, 2022.

[19] M. T. G. Thanh, N. Van Toan, V. T. N. Ngoc, N. T. Tra, C. N. Giap and D. M. Nguyen, "Deep Learning Application in Dental Caries Detection Using Intraoral Photos Taken by Smartphones," Applied Sciences, vol. 12, no. 11, pp. 1-10, 2022.

[20] D. Kim, E. Choi, H. G. Jeong, J. Chang and S. Youm, "Expert system for mandibular condyle detection and osteoarthritis classification in panoramic imaging using R-CNN and CNN," Applied Sciences, vol. 10,no.21, pp. 1-10, 2020.

[21] V. K. Krishnamoorthy and S.Baskarn, "Optimized adversarial network with faster residual deep learning for osteoarthritis classification in panoramic radiography," International Journal of Intelligent Engineering & Systems, vol. 15, no. 6 , pp. 191-200, 2022.

[22] Gaikwad, R. S. ., & Gandage, S. . C. (2023). MCNN: Visual Sentiment Analysis using Various Deep Learning Framework with Deep CNN. International Journal of Intelligent Systems and Applications in Engineering, 11(2s), 265 –. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2625

[23] Y. Ariji, Y. Yanashita, S. Kutsuna, C. Muramatsu, M. Fukuda, Y. Kise and E. Ariji, "Automatic detection and classification of radiolucent lesions in the mandible on panoramic radiographs using a deep learning object detection technique," Oral Surgery, Oral Medici, Oral Pathology and Oral Radiology, vol. 128, no.4, pp. 424-430, 2019.

[24] Mr. Nikhil Surkar, Ms. Shriya Timande. (2012). Analysis of Analog to Digital Converter for Biomedical Applications. International Journal of New Practices in Management and Engineering, 1(03), 01 - 07. Retrieved from http://ijnpme.org/index.php/IJNPME/article/view/6

[25] A. A. Al Kheraif, A. A. Wahba and H. Fouad, "Detection of dental diseases from radiographic 2d dental image using hybrid graph-cut technique and convolutional neural network," Measurement. vol. 146, pp. 333-342,2019.

[26] R. Abdalla-Aslan, T. Yeshua, D. Kabla, I. Leichter and C. Nadler, "An artificial intelligence system using machine-learning for automatic detection and classification of dental restorations in panoramic radiography," Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology,vol. 130,no. 5 , pp. 593-602,2020.

[27] M. P. Muresan, A. R. Barbura and S. Nedevschi, Teeth detection and dental problem classification in panoramic X-ray images using deep learning and image processing techniques, In IEEE 16th International Conference on Intelligent Computer Communication and Processing, pp. 457-463,2020.

[28] Thomas Wilson, Andrew Evans, Alejandro Perez, Luis Pérez, Juan Martinez. Machine Learning for Anomaly Detection and Outlier Analysis in Decision Science. Kuwait Journal of Machine Learning, 2(3). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/207

[29] A. Viloria, M. Mendinueta, L. A. Borrero and O. B. Pineda, "Prediction of mandibular morphology through artificial neural networks," Procedia Computer Science, vol.170, pp. 370-375, 2020.

[30] L. M. Leo and T. K. Reddy, "Learning compact and discriminative hybrid neural network for dental caries classification," Microprocessors and Microsystems, vol. 82, p. 103836, 2021.

[31] D. S. Bormane and R. B. Kakkeri, " Detection of temporomandibular joint disorder using surface electromypgraphy by supervised classification models," Materials Today: Proceedings, 2021.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7384
***Article Received**: 02 April 2023 **Revised**: 18 May 2023 **Accepted**: 20 June 2023*

[32] M. Aljabri, S. S. Aljameel, N. Min-Allah, J. Alhuthayfi, L. Alghamdi, N., Alduhailan and W. Al Turki, "Canine impaction classification from panoramic dental radiographic images using deep learning models," Informatics in Medicine Unlocked. 30, p.100918, 2022.

[33] S. Ito, Y. Mine, Y. Yoshimi, S. Takeda, A. Tanaka, A. Onishi and K. Tanimoto, Automated segmentation of articular disc of the temporomandibular joint on magnetic resonance images using deep learning, Scientific Reports,vol. 12, no.1, p.221, 2022.