RESEARCH ARTICLE

Interpretation of Magnetic Resonance Images of Temporomandibular Joint Disorders by Using Deep Learning

SIFA OZSARI [ID]1, **FATIMA RABIA YAPICIOGLU** [ID]2, **DILEK YILMAZ** [ID]3, **KIVANC KAMBUROGLU**4, **MEHMET SERDAR GUZEL** [ID]1, **GAZI ERKAN BOSTANCI** [ID]1, **KORAY ACICI** [ID]5, **AND TUNC ASUROGLU** [ID]6

1Department of Computer Engineering, Ankara University, 06830 Ankara, Turkey
2Department of Information Engineering, Università degli Studi di Padova, Padova, 35100 Veneto, Italy
3Department of Dentomaxillofacial Radiology, Faculty of Dentistry, Başkent University, 06830 Ankara, Turkey
4Department of Dentomaxillofacial Radiology, Ankara University, 06830 Ankara, Turkey
5Department of Artificial Intelligence and Data Engineering, Ankara University, 06830 Ankara, Turkey
6Faculty of Medicine and Health Technology, Tampere University, 33720 Tampere, Finland

Corresponding author: Tunc Asuroglu (tunc.asuroglu@tuni.fi)

ABSTRACT In recent years, Machine Learning (ML), especially Deep Learning (DL) approaches, has attracted great attention in medical field. In this study, we proposed a deep learning-based approach in order to automatically diagnose Temporomandibular Disorder (TMD) on Magnetic Resonance (MR) images. 2576 MR images of 200 patients diagnosed with and without TMD were collected. These images were classified as 8 groups. First of all, a basic Convolutional Neural Network (CNN) was used for the problem. After that, 6 different fine-tuned pre-trained convolutional neural network models, Xception, ResNet-101, MobileNetV2, InceptionV3, DenseNet-121 and ConvNeXt were applied on data set. Finally, the accomplishment of Vision Transformer (ViT) in task solving was also discussed. Performances of the approaches were evaluated by metrics such as accuracy rate, precision, sensitivity, F1-score, Negative Predictive Value (NPV), specificity, Area Under Curve (AUC) and kappa coefficient. Grad-CAM results of the best architectures for diagnostic examination were obtained. Intraclass Correlation Coefficients (ICC) value was computed to assess correlation between the models. According to the test results, deep learning-based architectures assessed were found to be successful in the diagnosis of TMD.

INDEX TERMS Deep learning, magnetic resonance imaging, temporomandibular joint disorders.

I. INTRODUCTION

Temporomandibular Joint (TMJ), located between mandibular condyle and temporal bone, is the most complex joint of the human body. Temporomandibular joint consists of condyle, articular tubercle, articular disc, glenoid fossa, joint capsule, retro discal tissue and synovial membrane. Clinical problems of the temporomandibular joint and joint structures are called temporomandibular joint disorders [1].

The associate editor coordinating the review of this manuscript and approving it for publication was Marco Giannelli [ID].

Temporomandibular joint disorders do not only affect TMJ but also, they influence the masticatory muscles and other components of stomatognathic system [1]. Pain symptoms frequently arise in patients with TMJ Internal Derangement (ID). The articular disk of the TMJ is composed of a biconcave fibrocartilaginous structure. Internal irregularity, manifested as disc displacement, is a common form of TMDs. TMJ ID is defined as an abnormal positional relationship between the articular disc and the mandibular condyle and the articular eminence. Anterior disc displacement is discovered more often than medial, posterior and lateral displacements [2].

TMJ effusion, an inflammatory response, is the accumulation of excess fluid in and around the temporomandibular joint. This reaction emerges in consequence of internal disorder, arthritis, trauma, and inflammatory changes associated with rheumatoid diseases. Degenerative bony changes are more frequent in the mandibular condyle than in the mandibular fossa or the articular eminence and are characterized by the development of pathological bony changes (erosion, osteophytes and deformity) and adaptive bony changes (marginal proliferation, flattening, concavity, sclerosis and subchondral cysts) [3], [4]. These abnormalities are considered to be radiological signs of OsteAarthritis (OA) and are frequently observed in joints with long-standing anterior disc displacement without reduction [5]. TMD, involving clinical symptoms, such as; TMJ pain, joint sounds and limited jaw function, affects 28% of the world's population [6]. These disorders include those associated with articular disc structure and position, as well as alterations in synovial fluid and soft tissue [7].

Diagnosis of TMJ disorders is made by clinical examination of the patient, along with patient history and assessment of diagnostic images including Magnetic Resonance Image (MRI) when necessary. MRI allows excellent depiction of the TMJ anatomy and abnormalities because of its inherent tissue contrast and high resolution. Magnetic resonance imaging is considered as reference standard for diagnostic imaging of some TMDs including those associated with articular disc structure and position, as well as alterations in synovial fluid and soft tissue [8]. However, this procedure is considered as costly and time consuming. Therefore, it is worth introducing an automatic inference system to facilitate diagnosis of disease for physicians. Artificial Intelligence (AI), machine learning, especially deep learning-based models, are highly effective approaches in order to achieve such tasks. The thought of artificial intelligence was first mentioned at a computer science conference held in Dartmouth College, Hanover, New Hampshire in 1956 [9], [10]. Artificial intelligence is described as "A system's ability to interpret external data correctly, to learn from such data, and to use that information to achieve specific goals and tasks through flexible adaptation." [11]. Artificial intelligence, which is essentially based on the way of human thinking, has subfields. Machine learning is one of these subsets. In machine learning, a mathematical model, which is dependent on training data, is designed in order to make predictions or decisions [12]. Machine learning algorithms have received much attention in recent years, and they have been used in many applications. However, these approaches have not always shown effective results as desired, especially on image and sound data. Krizhevsky et. al., achieved an 86.3% success in "Deep Learning's ImageNet Large-Scale Visual Recognition Competition (ILSVRC)" in 2012 [13] and studies based on deep learning were carried out by other researchers in the following years. In recent years, deep learning-based models performed better than conventional machine learning based methods in most computer vision and medical image processing problems. Accordingly, they have been utilized in several diverse fields, such as classification, disease diagnosis, face recognition, and image enhancement [14], [15], [16], [17], [18].

There are published studies regarding assessment of TMJ by using machine learning, artificial intelligence, or deep learning in previous literature. Application of deep learning approaches related to different TMJ diseases and diagnostic problems were studied in [19], [20], [21], [22], [23], [24], [25], [26], [27], and [28].

In this study, our aim was to interpret TMJ disorders displayed on MRI by using deep learning approaches and to assess its effectiveness. A basic CNN architecture, 6 different fine-tuned pre-trained models, Xception [29], Residual Neural Network (ResNet)-101 [30], MobileNetV2 [31], InceptionV3 [32], Dense Convolutional Network (DenseNet)-121 [33] and ConvNeXt [34] were applied on magnetic resonance images of temporomandibular joint patients to predict whether they have disc displacement, effusion or condylar bone degeneration or not. In addition, the performance of ViT, which has been very popular lately, was examined.

II. MATERIALS AND METHODS

Sample images were taken from Ankara University, Faculty of Dentistry and a data set was created. Ethical approval for the data set was also obtained from Ankara University, Faculty of Dentistry, Ethics Committee (36290600/34/2021). Informed consent was obtained from the patients.

A. MR IMAGES DATA SET

The present study included MR images of 200 patients who were referred to Ankara University Faculty of Dentistry, Dentomaxillofacial Radiology and Oral Surgery Clinics for TMJ Disorders between 2015-2021. Patients who had surgery including chin and face and patients with facial syndrome were excluded. MR images were obtained by utilizing 1.5 Tesla machine in both closed and open mouth positions. T1 W (T1 weighted), T2 W images, Multiple Echo Recombined Gradient Echo (MERGE) and Proton Density (PD), series images of bilateral TMJs (3 mm thick) were obtained with sagittal reconstructions. In order to increase the number of MR images, T1 and PD images were used for articular disc position, T2 and MERGE for effusion, and T1 and PD series for condylar degeneration assessment. Screenshots were taken as shown in Figure 1 (a), Figure 1 (b) for articular disc position, Figure 1 (c) for effusion, and Figure 1 (d) condylar degeneration.

MR images used in the present study were collected by an experienced dentomaxillofacial radiology specialist in MRI assessment. In cases where a decision could not be made by the dentomaxillofacial radiology specialist a consensus was achieved by consulting a more experienced dentomaxillofacial specialist (15 years of experience in reading MR images). 2576 images were obtained from diagnostic MR images of 200 patients. At least 3 images of the right and left regions
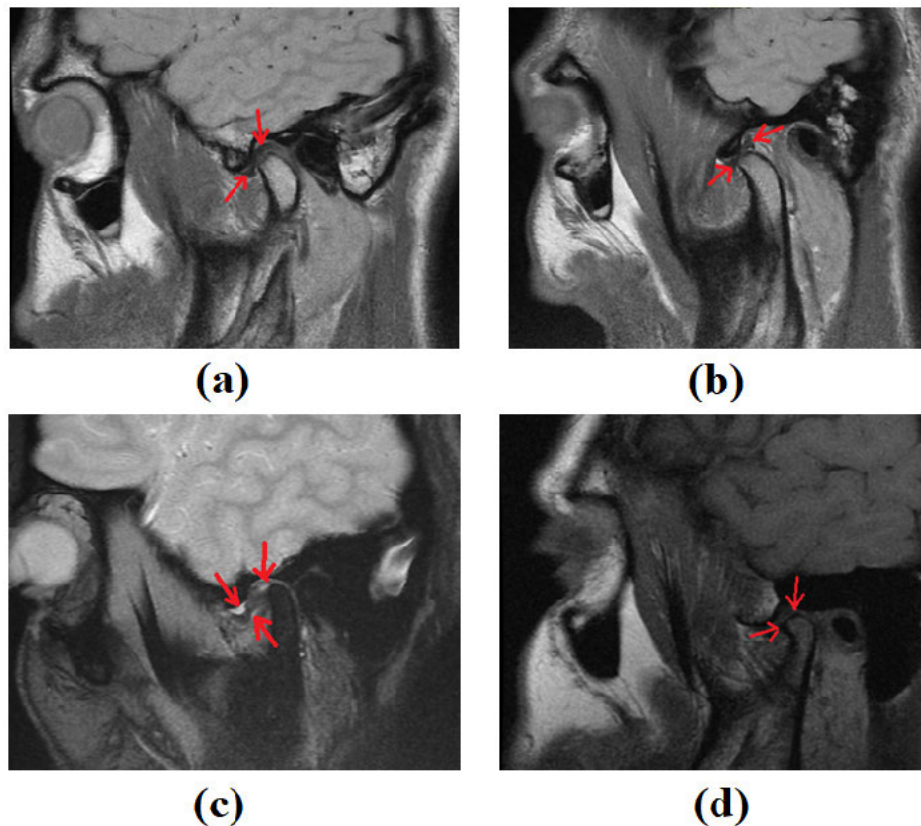
FIGURE 1. Right TMJ in sagittal oblique plane, in PD series: (a) closed mouth, (b) open mouth position, (red arrow) normal disc position view, (c) right mouth closed TMJ effusion (red arrows) in sagittal oblique plane, MERGE series, (d) right mouth closed TMJ in sagittal plane, in T1 series; view of condylar degeneration (red arrows).

TABLE 1. Data set properties.

Category	Normal	Positive
At mouth closed position, articular disc	369	357
At mouth open position, articular disc	253	274
Synovial membrane fluid level	479	476
Mandibular condyle	178	190

of each patient were transferred to the data set. The image set we obtained was divided into 4 groups and 8 classes:

1) At closed mouth, articular disc position normal
2) At closed mouth, articular disc position in front
3) At open mouth, articular disc position normal
4) At open mouth, articular disc position in front
5) Joint cavity without effusion
6) Joint cavity with effusion
7) There is no mandibular condyle degeneration
8) There is mandibular condyle degeneration

Table 1 shows the distribution of number and percentage of images for all groups from a total of 2567 data images.

In Table 1, for T1 and Proton density sequences "Positive" corresponds to the "disc in anterior or disc anteriorly displaced" for "closed mouth disc position" and "open mouth disc position" groups. For "joint cavity effusion" group and for the "mandibular condyle degeneration" group, Class 5, 6, 7 and 8 were determined according to following criteria:

- For 5: In the T2 and MERGE sequence, no increase in intensity is observed in the TMJ interval,
- For 6: There is hyperintensity in the TMJ space on the T2 and MERGE sequence,
- For 7: The appearance of a smooth and rounded line on the mandibular condyle surface at T1 and Proton density sequence,
- For 8: It is the appearance of irregularity, flattening and osteophyte (bird beak appearance) on the mandibular condyle surface at T1 and Proton density sequence.

B. TRANSFER LEARNING

In broad terms, Transfer Learning (TL) can be defined as applying the knowledge obtained while solving a problem to a different but relevant task. It is a very strong learning approach in order to solve problems where there are not enough training examples in order to train a model from scratch. Instead of starting the learning from scratch, transfer learning allows to create accurate models in a time saving
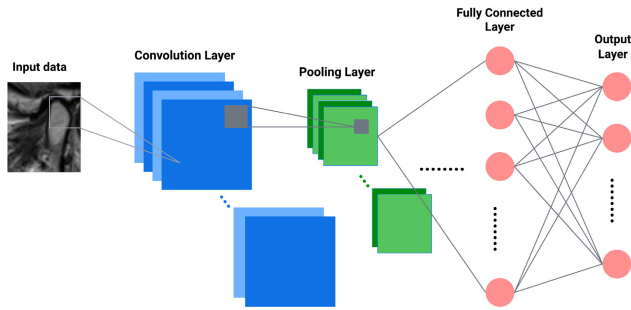
FIGURE 2. Basic layers in a convolutional neural network.

way, starting from learned patterns while solving a task [35], [36]. In this study, transfer learning was employed as we did not have adequate sample images to train an architecture from scratch. When a pre-trained model is fine-tuned for a new task, either the entire model is trained or some layers are trained and others are left frozen, or the convolutional structure is left in its original state.

In deep learning, a CNN is a type of artificial neural network in which mathematical operations are performed. It has widespread use on image data. The primary layers of a CNN architecture are shown in Figure 2. The general description of layers in Figure 2 is as follows: The convolution layer is the first layer of a CNN. In this layer, features are extracted from the image using some filters. The pooling layer, which has a different number of functions, is a layer added between convolutional layers in a network. The aim here is to reduce the parameters and the number of calculations in the network. The Fully Connected (FC) layer is one of the last layers of a convolutional neural network. The neurons in this layer are connected to all neurons before and after them. Finally, the output layer is the layer where the result is produced. In the present study, the last layer of pre-trained models, namely, Xception [29], ResNet-101 [30], MobileNetV2 [31], InceptionV3 [32], DenseNet-121 [33] and ConvNeXt [34] were fine-tuned. The details of each DL based architecture are as follows:

1) XCEPTION

Xception is a popular CNN architecture proposed by Francois Chollet [29]. It is based on depth wise separable convolution layers with residual connections. The researcher performs a theory which is a more robust version of the basic hypothesis of the Inception model. According to this theory, the mapping of cross-channels correlations and spatial correlations in the feature maps of CNNs can be completely separable. In the Xception model, there are 36 convolution layers structured into 14 modules. Except for the first and last modules, all these modules have linear residual connections around them. The data goes through the input flow, middle flow, and output flow, respectively [29].

2) ResNet-101

ResNet architecture is one of the most popular Deep Neural Networks (DNN) available in many varieties with different

numbers of layers. ResNet was introduced by He et al. [30]. They presented a residual learning framework to ease the training of networks and they reformulated the layers as learning residual functions. Due to the vanishing gradient problem, deep networks are difficult to train. Therefore, increasing the depth of the network by stacking more layers may not be effective. As the size of the net gets bigger, the accuracy becomes saturated and then reduces rapidly [30]. He et al. presented a deep residual learning framework for the problem of degradation. They introduced "identity short-cut connections" that skip one or more layers with identity mapping [30].

3) MobileNetV2

MobileNetV2 [31] is a convolutional neural network architecture specifically designed for mobile devices and environments with constrained resources. MobileNetV2 is based on ideas taken from MobileNetV1 [37], however; it is a significant improvement over MobileNetV1. Deep separable convolution is used as efficient building blocks in MobileNetV2. However, MobileNetV2 has the following two new features: Linear bottlenecks between layers and shortcut connections [38]. In the first place, the MobileNetV2 model has a fully convolution layer, involving 32 filters. After this layer, there are 19 bottleneck layers. ReLU is used as non-linearity, in addition kernel size is 3×3 and dropout and batch normalization are used during the training process. A constant rate of expansion is utilized throughout the network, except for the first layer [31].

4) InceptionV3

InceptionV3 [32], which started as a model for GoogLeNet, is a CNN to aid in image analysis and object detection [39]. It is the 3rd version from the Inception family and was first introduced during the ImageNet Recognition Challenge. InceptionV3 includes convolutions, average pooling, max pooling, concats, dropouts, and fully connected layers [32]. This model aims to prevent the number of parameters from increasing too much while the network gets deeper. Thereby, it has lower computational cost.

5) DenseNet-121

DenseNet connects each layer to every other layer in a feed-forward form. At each layer, features from all previous layers are taken as input and their own features are passed to all subsequent layers. DenseNets provide significant advantages such as; they make feature propagation robust, support feature reuse, lessen the vanishing-gradient problem, and significantly decrease the number of parameters [33]. The model starts with the convolution pooling blocks, continues with the dense block transition layer, and finally ends with the global average pool and a fully-connected block.

6) ConvNeXt

ConvNeXT [34] inspired from vision transformers is a convolutional model proposed by Liu et al. They gradually

"modernized" a standard ResNet towards a vision trans-formers design. They found some key components and these components contribute to the performance difference. Deep convolution which is a special case of grouped convolution was used in ConvNeXt. Accordingly, the number of groups is equal to the number of channels. In depth-wise convolution, a process similar to the weighted sum in self-attention is performed. This process works on a channel basis, in other words it only mixes spatial information.

7) ViT

ViT is an architecture with transformer. Transformer was firstly used in Natural Language Processing (NLP) tasks [40]. Basically, a transformer includes a self-attention structure. This mechanism weights the importance of each part of the input data. The main purpose of this method is to focus on significant point in data. This approach has attracted great attention due to its success in NLP applications. Accordingly, it started to find a place in image applications [41].

8) METRICS

In order to examine the performance of models, metrics such as accuracy rate, precision (Positive Predictive Value/PPV), recall (sensitivity), F1-score, (Negative Predictive Value/NPV), specificity, AUC and kappa score were used. Descriptions of aforementioned metrics are given at equations 1, 2, 3, 4, 5 and 6.

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall(sensitivity) = \frac{TP}{TP + FN} \tag{3}$$

$$F1 - score = \frac{2 \times precision \times recall}{precision + recall} \tag{4}$$

$$NPV = \frac{TN}{TN + FN} \tag{5}$$

$$Specificity = \frac{TN}{TN + FP} \tag{6}$$

For equations 1, 2, 3, 4, 5 and 6. TP corresponds to true positives, FP false positives, TN true negatives and FN false negatives. TP means that the test sample has disease and is predicted to have disease. FP is the example that is healthy but classified as diseased. TN is healthy and classified as healthy. FN is the sample that has disease but is classified as not healthy.

Kappa score is a measure of reliability of a model in classification. That is, it expresses the agreement between model and real classes. Its interpretation is as follows: Kappa values from 0.0 to 0.2 shows slight agreement, 0.21 to 0.40 indicates fair agreement, from 0.41 to 0.60 is moderate agreement, 0.61 to 0.80 indicates substantial agreement, and from 0.81 to 1.0 gives almost perfect or perfect agreement.

In addition, Gradient-weighted Class Activation Mapping (Grad-CAM) [42] images of best architectures were examined. Grad-CAM is a heat map image of the tested data. It shows from which regions the architectures were inferred. Red color in the heat map image is the area where the models predict most intensely. In this way, we observed whether the models' learned parts related to the disease or not.

Correlation between architectures was also examined by using ICC value. The correlation between the metric results of models was statistically evaluated by using ICC. The relation was measured separately for each success metric. Intraclass correlation coefficients were computed for each class according to the following criteria: < 0.40 = poor agreement; 0.40-0.59 = fair agreement; 0.60-0.74 = good agreement; 0.75-1.0 = excellent agreement. Thereby, the correlation of architectures with each other in correct diagnosis was assessed.

III. RESULTS

A series of experiments were carried out to evaluate and to confirm the performance of a CNN structure, different TL based models (Xception, ResNet-101, MobileNetV2, InceptionV3, DenseNet-121 and ConvNeXt) and ViT in terms of inferencing TMJ disorders automatically. All experiments were run in Google Colaboratory (Colab) [43]. Colab is a product developed by Google Research, where python codes can be written and executed online. It allows the use of GPU and to import many libraries automatically, therefore it is very practical for studies on machine learning.

The images were divided into 4 groups and 8 different classes and training, and testing processes were carried out separately for each group. This allowed observation and evaluation of performance analysis of deep learning approaches in a more efficient way. Considering that the jaw structure could be in different shapes and sizes depending on gender and age groups, the data set comprised images of patients from different gender and age groups. In the first place, the data set was divided as training, validation, and test groups so as to provide a better validation processes. For each category, training data was 80% of the total number of data and test data was 20%. Also, 20% of the training data was employed for the validation process. Moreover, data augmentation techniques were applied to increase the amount of data. Data augmentation is the artificial expansion of available data by making some random (but realistic) changes to the existing images. These changes are minor ones such as rotating, increasing brightness, zooming etc. Each operation increases the number of data as much as itself. For example, if three different augmentations are applied, the number of data will triple. In this experimental study, we used 3 different augmentation methods: contrast, flip and rotation. In order to obtain more reliable results, the region of disc was cut out from MR images. In other words, the data set was cropped manually. The reason why images were cut this way is because a huge amount of data was needed for architectures to learn from the full version of the MR images (as in Figure 1). MR images as in Figure 1 included parts that were not related to the TMJ, such as the brain, eye, mouth, etc. These parts caused the
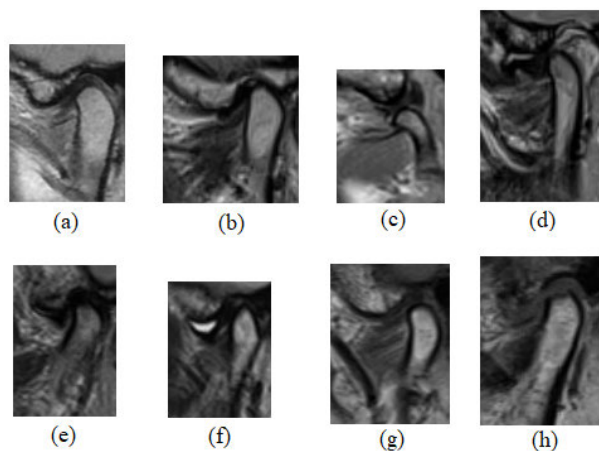
FIGURE 3. The examples for cropped images: (a) at closed mouth, articular disc position normal, (b) at closed mouth, articular disc position in front, (c) at open mouth, articular disc position normal, (d) at open mouth, articular disc position in front, (e) joint cavity without effusion, (f) joint cavity with effusion, (g) there is no mandibular condyle degeneration, (h) there is mandibular condyle degeneration.

network to make inferences from the wrong parts. Figure 3 represents an example for cropped versions of images.

Firstly, classification was performed with a basic CNN network. The network consists of the Convolution, Pooling and Dense layers. After that, models were applied separately for each class on the data set during the training process. Finally, experiments with ViT were carried out.

In order to obtain a fair comparison between DL architectures, parameter values were set the same for all these models. For instance, the number of epochs was 100, loss function was binary cross-entropy, optimizer was ADAM [44] and learning rate was 1×10^{-6}. However, in some trainings, the loss value of architectures increased after the 50th iteration. Therefore, the training phase of these models was discontinued at 50. After the training process was completed, the success of models was tested through the use of test data. Accuracy rate, precision, sensitivity, F1-score, NPV, specificity, AUC and kappa score were calculated. According to these results, which architectures Grad-CAM images would be produced were determined. Finally, correlation between the metric results of architectures was examined by using ICC. Figure 4 illustrates the flow chart of steps followed to solve the problem.

Tables 2, 3, 4 and 5 present the results for "closed mouth disc position", "open mouth disc position", "joint cavity effusion" and "mandibular condyle degeneration" respectively. According to our findings, it can be said that MobileNetV2 was the best architecture for "closed mouth disc position" and Xception for "open mouth disc position" groups, whereas ResNet-101 was considered as the most effective model for "joint cavity effusion" and MobileNetV2 was the most effective model for "mandibular condyle degeneration" groups, respectively. The best architectures were determined by considering all metrics. Training and

validation accuracy graphics of successful architectures were presented in Figure 5.

We also mentioned that experiments were carried out with ViT. We conducted the experiments, but unfortunately the results were 50% and below. These outcomes are much lower than expected and not acceptable. Therefore, it was not included in the tables and a comparison with other models was not made.

Grad-CAM results of best architectures were also produced. The Grad-CAM results of architectures are presented in Figure 6. In Figure 6, the first row shows diseased areas marked by the dentomaxillofacial radiologist. The MRI images used in this study were images of patients diagnosed by consensus of dentomaxillofacial specialists. Therefore, it allows an accurate assessment of the success of architectures. As mentioned earlier, the red parts are the areas where architectures were heavily focused for learning. It was not expected to obtain a precise boundary drawing from those images. The main purpose was to observe whether inferences were made from the right regions. As can be seen from the Grad-CAM images, models mostly colored (selected) the correct regions. This means that the networks were quite successful in classification. Finally, the ICC value for metrics is presented in Table 6.

The ICC values of models were 0.58 for accuracy, 0.57 for precision, 0.63 for sensitivity, 0.64 for F1-score, 0.28 for NPV, 0.79 for specificity, 0.74 for AUC and 0.67 for kappa score. While the correlation between models for the NPV value was low, the results for other metrics were found to be acceptable.

IV. DISCUSSION

The TMJ is among one the most complex joints of the human body. Temporomandibular joint and its related structures have an important role in distributing the stresses produced by frequently performed tasks like speaking, chewing and swallowing, as well as directing the jaw movement [45]. Temporomandibular joint disorders correspond to the clinical problems of the temporomandibular joint and joint structures as well as the masticatory muscles [1]. TMD can be the cause of psychological diseases such as depression and inferiority complex or vice versa anxiety and depression may cause TMJ complaints [46]. According to epidemiological studies, the rate of adults showing at least one TMD symptom during their examination may rise to 75% [47]. The MR images of TMJ are interpreted and reported by dentomaxillofacial radiologists. Therefore, observer performance and experience are important factors in the diagnosis of TMJ disease.

Translation research, which has developed in recent years, aims to integrate new methods into medical applications and accelerate the transition to clinical practice. Various studies have been carried out in this context [48], [49]. The aim of this study was to automatically report TMJ MR images in a short time and to prevent the interpretation difference depending on observer performance and experience. In addition, this method can be used for evaluation and follow-up after
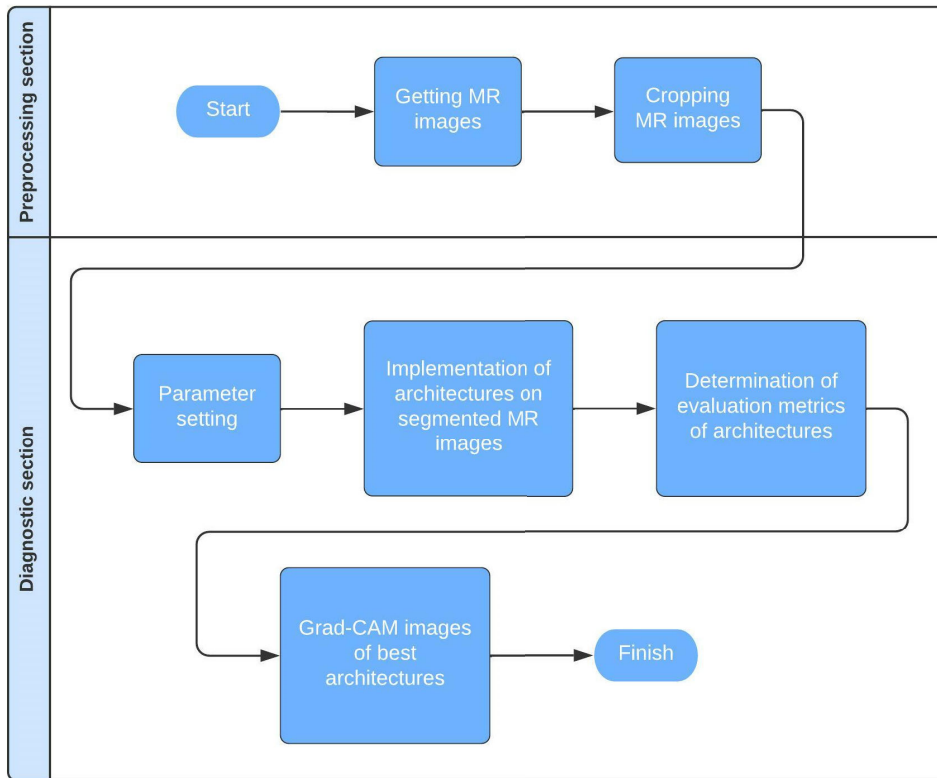
FIGURE 4. Schematic chart of the proposed method.

TABLE 2. Performances of networks for closed mouth disc position.

Model	Accuracy rate	Precision	Sensitivity	F1 − score	NPV	Specificity	AUC	Kappa score
CNN	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.05
Xception	0.88	0.88	0.88	0.87	0.89	0.82	0.85	0.72
ResNet-101	0.91	0.91	0.91	0.91	0.91	0.90	0.90	0.81
MobileNetV2	0.97	0.97	0.97	0.97	0.98	0.94	0.95	0.93
InceptionV3	0.84	0.84	0.84	0.84	0.84	0.83	0.84	0.68
DenseNet-121	0.84	0.87	0.84	0.85	0.77	0.87	0.86	0.66
ConvNeXt	0.87	0.88	0.87	0.87	0.87	0.88	0.88	0.75

TABLE 3. Performances of networks for open mouth disc position.

Model	Accuracy rate	Precision	Sensitivity	F1 − score	NPV	Specificity	AUC	Kappa score
CNN	0.68	0.60	0.68	0.64	0.46	0.34	0.51	0.03
Xception	0.81	0.85	0.81	0.79	0.92	0.64	0.72	0.52
ResNet-101	0.77	0.82	0.81	0.80	0.84	0.68	0.75	0.54
MobileNetV2	0.75	0.70	0.72	0.70	0.57	0.54	0.63	0.29
InceptionV3	0.66	0.69	0.66	0.60	0.71	0.52	0.59	0.20
DenseNet-121	0,69	0.64	0.69	0.66	0.50	0.40	0.55	0.11
ConvNeXt	0.78	0.78	0.78	0.77	0.76	0.67	0.72	0.48

treatment. To our knowledge, our study is one of the leading works to diagnose TMD by using different architectures and to interpret the success of models according to images diagnosed by dentomaxillofacial radiologists. In our opinion, use of this methodology will enable higher accuracy in diagnosis while consuming less time and therefore we believe that a significant contribution will be made to current TMJ related literature.

TABLE 4. Performances of networks for joint cavity effusion.

Model	Accuracy rate	Precision	Sensitivity	F1 − score	NPV	Specificity	AUC	Kappa score
CNN	0.53	0.53	0.53	0.53	0.53	0.53	0.53	0.07
Xception	0.81	0.82	0.81	0.81	0.82	0.80	0.81	0.62
ResNet-101	0.84	0.86	0.84	0.84	0.84	0.86	0.85	0.69
MobileNetV2	0.75	0.78	0.75	0.75	0.72	0.78	0.77	0.5
InceptionV3	0.72	0.77	0.72	0.73	0.88	0.76	0.74	0.42
DenseNet-121	0.81	0.82	0.81	0.81	0.80	0.82	0.82	0.62
ConvNeXt	0.69	0.69	0.69	0.69	0.66	0.66	0.68	0.35

TABLE 5. Performances of networks for mandibular condyle degeneration.

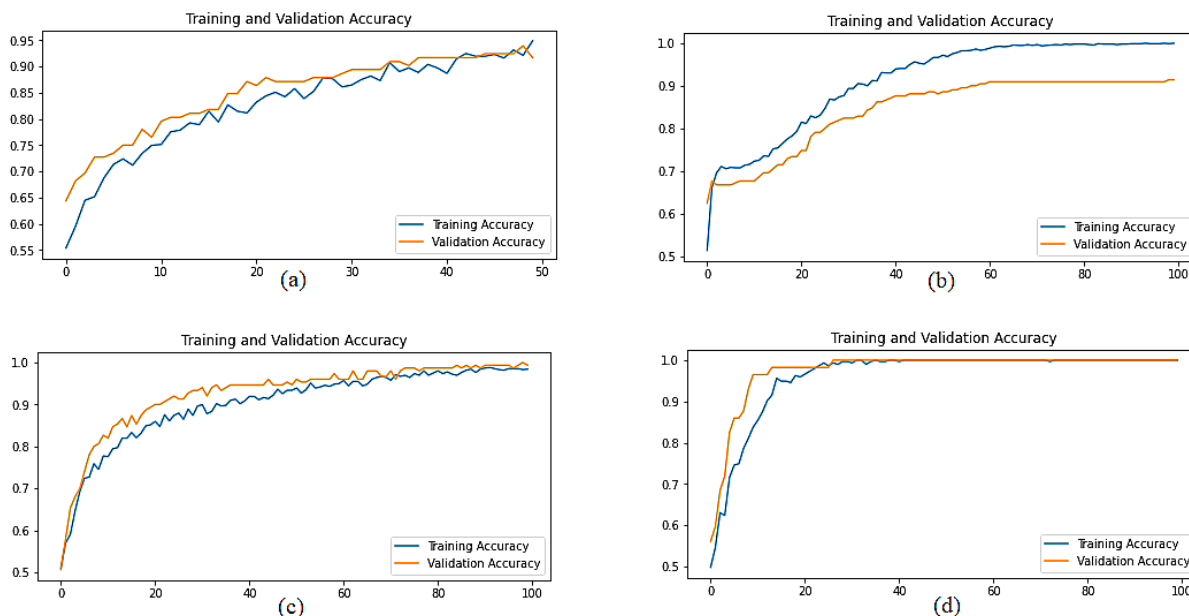| Model	Accuracy rate	Precision	Sensitivity	F1 − score	NPV	Specificity	AUC	Kappa score
CNN	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.02
Xception	0.8	0.89	0.88	0.88	0.77	0.89	0.88	0.73
ResNet-101	0.94	0.94	0.94	0.94	0.95	0.92	0.93	0.85
MobileNetV2	0.97	0.97	0.97	0.97	0.96	0.97	0.97	0.94
InceptionV3	0.78	0.78	0.78	0.78	0.77	0.78	0.78	0.56
DenseNet-121	0.94	0.94	0.94	0.94	0.93	0.93	0.94	0.87
ConvNeXt	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.87

FIGURE 5. Accuracy graph of the best architectures: (a) closed mouth disc position MobileNetV2, (b) open mouth disc position Xception, (c) joint cavity effusion ResNet-101, (d) mandibular condyle degeneration MobileNetV2.

The implementation that is analogous to our study and that we can compare is the publication of Kao et al. [28]. They performed DL-based models (InceptionResNetV2, InceptionV3, DenseNet169, and VGG16) on total 300 images of 32 healthy and 52 TMD patients. As a first step, they detected the articular space between the temporal bone and the mandibular condyle using the U-Net architecture from 100 sagittal MRI images of the TMJ. Then, they made classification using the specified architectures. Recall, precision, accuracy, and F1 score values for InceptionV3 were 1.0, 0.81, 0.85, 0.9, respectively, and 0.92, 0.86, 0.85, 0.89 for DenseNet169.
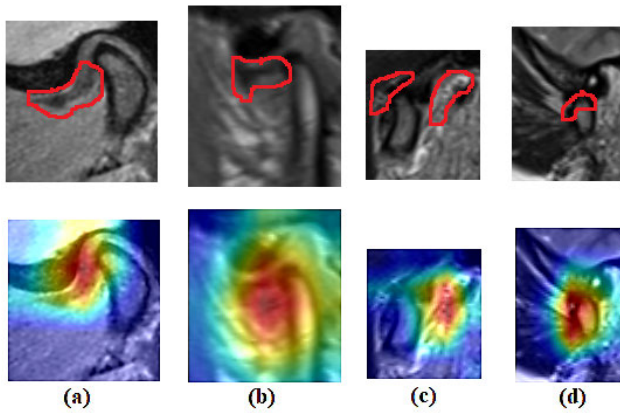
FIGURE 6. The Grad-CAM images of best architectures: (a) closed mouth disc position MobileNetV2, (b) open mouth disc position Xception, (c) joint cavity effusion ResNet-101, (d) mandibular condyle degeneration MobileNetV2.

TABLE 6. Performances of networks for mandibular condyle.

Metric	ICC value
Accuracy rate	0.58
Precision	0.57
Recall	0.63
F1-Score	0.64
NPV	0.28
Specificity	0.79
AUC	0.74
Kappa score	0.67

Previous studies other than [28] generally focused on different diseases and inferences. For instance, Kim et al. [25] applied random forest and MultiLayer Perceptron (MLP) methods by using MR images for the detection of TMJ disc perforation. They studied 299 joints belonging to 289 patients. They divided these joints into two groups as perforated and non-perforated. This separation was determined according to the presence of disc perforation detected during surgery. They compared the performance of models by using AUC. MLP performed best with AUC of 0.940, followed by random forest with AUC of 0.918, and disk shape alone resulted in AUC of 0.791. There are some published studies regarding TMJ segmentation in the literature. Authors of a previous research, [26] proposed an all-automatic articular disc detection and segmentation system. In this system, a DL-based semantic segmentation approach was proposed. Within the proposed system, authors aimed to support the diagnosis of TMD in MRI. Two hundred and seventeen (217) MR images were used. These images were images of patients with displaced or normal articular discs. Three DL-based semantic segmentation approaches were used. The first one, an encoder-decoder CNN model named 3DiscNet (Detection for Displaced articular DISC using convolutional neural NETwork), proposed as a new approach within the study. This

was compared with U-Net and SegNet-Basic architectures. Another interesting study on temporomandibular joint segmentation was carried out by Liu et al. [27]. They introduced an automated segmentation algorithm based on deep learning followed by a post processing stage. First, a U-Net model was applied to separate images into 3 categories (glenoid fossa, condyles, and background). In the post-processing stage, the internal force constraint of a snake model was used to renew the integrity of the fracture boundary for structural fractures in these split images. Based on the tracking concept, the initial boundary of the snake was obtained. A total of 206 low-dose Computed Tomography (CT) cases were used and compatibility between the experimental results and the gold standard was evaluated with indicators such as the Dice Coefficient (DC) and the Mean Surface Distance (MSD).

In the current study, a CNN network, 6 different pre-trained, fine-tuned models (Xception, ResNet-101, MobileNetV2, InceptionV3, DenseNet-121 and ConvNeXt) were applied on the MR images (2576 MR images of 200 patients) prepared together with Ankara University, Faculty of Dentistry. In addition to these, the performance of ViT was also examined. It is an important point that the data used in our study was not obtained from the public data set but collected and created by the researchers. A pre-processing on the images was also carried out by the researchers. Before applying the models, the region related to the disc was cut out from images. Afterwards, experiments were carried out on these cropped images. Once the experimental results were examined, we observed that the TL-based architectures provided effective results. We figured out from Tables 2, 3, 4 and 5 that the best accuracy rates, F1 scores and AUC values were promising with values above 0.75. Considering all metrics in Table 2, MobileNetV2 provided the highest value. It produced a 97% accuracy rate, 0.97 F1-score and 0.95 AUC value. As a result of the execution of architectures on ''open mouth disc position'' MR images, Xception yielded the best accuracy with 81%. It was seen that the Xception architecture was the most successful model for precision, sensitivity and NPV values as well. The F1-score of this architecture was very effective at 0.79, but here, ResNet-101 presented the highest precision value with 0.80. In addition, the AUC value of ResNet-101 was 0.75, which was higher than Xception. According to the results in Table 4, in terms of all metrics, ResNet-101 provided the best results, hence it can be accepted as the best DL architecture. When we analyzed Table 5, it could be seen that MobileNetV2 was the most effective architecture in consideration to all findings. Since deep webs are black boxes, it is not clear how they learn and why different architectures are successful for each disease is unknown. But, we could state that the MobileNet is a model developed for limited data, this feature has enabled the architecture to be the most successful one in two classes and produce effective results in other groups.

Nevertheless, high success rates do not guarantee that architectures have the capability to diagnose diseases in the right regions. It is possible that models can make

predictions over wrong areas. Therefore, we also examined Grad-CAM images. Assessment of results was based on the regions marked by the dentomaxillofacial radiology specialist. In other words, the images generated by coloring with Grad-CAM were compared with the version of the same images marked by the specialist. In Figure 6, the red parts in the second rows show the regions from which the models infer. We observed that the models mostly colored the correct region, when the Grad-CAM images of the architectures were thought to be the best for each group examined. We would also like to state our interpretation on ConvNeXt. ConvNeXt had decent outcomes, but when Grad-CAM images were examined, it was seen that the architecture made inferences from wrong points. In our interpretation, the main reason for this is that the network focuses on wrong area.

When ICC values are analyzed, in general, they revealed good to excellent agreement values suggesting acceptable correlation among the models assessed.

We would like to emphasize again that the process of creating a dataset is costly and time consuming. Experimental results showed that our findings were encouraging. Provided models were successful in the assessment of diagnostic images of TMD patients. However, we believe that future studies should include a higher number of images for all groups incorporating images of patients with more than one disease. This is very important for the results to be more reliable and robust. Because the results obtained with little data are instructive but not sufficient for full practical application. In this regard, we plan to increase the number of data, to examine and to apply different methods mostly used in medical diagnosis.

V. CONCLUSION

For the dentistry literature, a CNN network, 6 different deep learning architectures were applied on MR images for TMD diagnosis. In addition to these, the effectiveness of ViT, which has been very popular nowadays, on MR images was investigated. We observed whether learning was carried out in part related to the disease with Grad-CAM images. Also, ICC was utilized to analyze the correlation between the models. The ICC value shows the agreement between models in correct diagnosis. Considering the results, the reliability between architectures was interpreted as satisfactory. Within the limitations of the present research, architectures were found to be successful in terms of metrics, such as accuracy rate, precision, sensitivity, F1-score, specificity, AUC and kappa score, examining the Grad-CAM images and calculating the ICC. Therefore, this study has the potential to make a significant contribution to the literature and encourage researchers to apply different DL based approaches to solve the diagnostic problem.

REFERENCES

[1] D. Yılmaz and K. Kamburoğlu, "Comparison of the effectiveness of high resolution ultrasound with MRI in patients with temporomandibular joint disorders," *Dentomaxillofacial Radiol.*, vol. 48, no. 5, Jul. 2019, Art. no. 20180349.

[2] R. Emshoff, I. Brandlmaier, S. Bertram, and A. Rudisch, "Relative odds of temporomandibular joint pain as a function of magnetic resonance imaging findings of internal derangement, osteoarthrosis, effusion, and bone marrow edema," *Oral Surg., Oral Med., Oral Pathol., Oral Radiol., Endodontol.*, vol. 95, no. 4, pp. 437–445, Apr. 2003.

[3] J. Koyama, H. Nishiyama, and T. Hayashi, "Follow-up study of condylar bony changes using helical computed tomography in patients with temporomandibular disorder," *Dentomaxillofacial Radiol.*, vol. 36, no. 8, pp. 472–477, Dec. 2007, doi: 10.1259/dmfr/28078357.

[4] W. Talaat, S. Al Bayatti, and S. Al Kawas, "CBCT analysis of bony changes associated with temporomandibular disorders," *CRANIO*, vol. 34, no. 2, pp. 88–94, Mar. 2016, doi: 10.1179/2151090315y.0000000002.

[5] M. I. G. Campos, P. S. F. Campos, M. C. T. Cangussu, R. C. Guimarães, and S. R. P. Line, "Analysis of magnetic resonance imaging characteristics and pain in temporomandibular joints with and without degenerative changes of the condyle," *Int. J. Oral Maxillofacial Surg.*, vol. 37, no. 6, pp. 529–534, Jun. 2008, doi: 10.1016/j.ijom.2008.02.011.

[6] L. LeResche and M. V. Korff, "Research diagnostic criteria for temporomandibular disorders: Review, criteria, examinations and specifications, critique," *J. Craniomandib Disord*, vol. 6, no. 4, pp. 301–355, 1992.

[7] Y.-W. Jung, S.-H. Park, S.-W. On, and S.-I. Song, "Correlation between clinical symptoms and magnetic resonance imaging findings in patients with temporomandibular joint internal derangement," *J. Korean Assoc. Oral Maxillofacial Surgeons*, vol. 41, no. 3, p. 125, 2015, doi: 10.5125/jkaoms.2015.41.3.125.

[8] X. Tomas, J. Pomes, J. Berenguer, L. Quinto, C. Nicolau, J. M. Mercader, and V. Castro, "MR imaging of temporomandibular joint dysfunction: A pictorial review," *RadioGraphics*, vol. 26, no. 3, pp. 765–781, May 2006, doi: 10.1148/rg.263055091.

[9] M. Copeland, "The difference between AI, machine learning, and deep learning?" NVIDIA Blog, NVIDIA Corp., Santa Clara, CA, USA, Tech. Rep., 2021. Accessed: Feb. 9, 2023. [Online]. Available: https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/

[10] A. Akay and H. Hess, "Deep learning: Current and emerging applications in medicine and technology," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 3, pp. 906–920, May 2019, doi: 10.1109/jbhi.2019.2894713.

[11] A. Kaplan and M. Haenlein, "Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence," *Bus. Horizons*, vol. 62, no. 1, pp. 15–25, Jan. 2019, doi: 10.1016/j.bushor.2018.08.004.

[12] X. D. Zhang, *Matrix Algebra Approach to Artificial Intelligence*. Singapore: Springer, 2021.

[13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.

[14] N. Kussul, M. Lavreniuk, S. Skakun, and A. Shelestov, "Deep learning classification of land cover and crop types using remote sensing data," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 5, pp. 778–782, May 2017, doi: 10.1109/lgrs.2017.2681128.

[15] S. L. Oh, Y. Hagiwara, U. Raghavendra, R. Yuvaraj, N. Arunkumar, M. Murugappan, and U. R. Acharya, "A deep learning approach for Parkinson's disease diagnosis from EEG signals," *Neural Comput. Appl.*, vol. 32, no. 15, pp. 10927–10933, Aug. 2020, doi: 10.1007/s00521-018-3689-5.

[16] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, and G. J. Soufi, "Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning," *Med. Image. Anal.*, vol. 65, Oct. 2020, Art. no. 101794, doi: 10.1016/j.media.2020.101794.

[17] W. Wang, J. Yang, J. Xiao, S. Li, and D. Zhou, "Face recognition based on deep learning," in *Proc. HCII*. Cham, Switzerland: Springer, 2014, pp. 812–820.

[18] M. Gharbi, J. Chen, J. T. Barron, S. W. Hasinoff, and F. Durand, "Deep bilateral learning for real-time image enhancement," *ACM Trans. Graph.*, vol. 36, no. 4, pp. 1–12, Aug. 2017, doi: 10.1145/3072959.3073592.

[19] P. de Dumast, C. Mirabel, L. Cevidanes, A. Ruellas, M. Yatabe, M. Ioshida, N. T. Ribera, L. Michoud, L. Gomes, C. Huang, H. Zhu, L. Muniz, B. Shoukri, B. Paniagua, M. Styner, S. Pieper, F. Budin, J. B. Vimort, L. Pascal, and J. C. Prieto, "A web-based system for neural network based classification in temporomandibular joint osteoarthritis," *Comput. Med. Imag. Graph.*, vol. 67, pp. 45–54, Jul. 2018, doi: 10.1016/j.compmedimag.2018.04.009.

[20] N. T. Ribera, P. De Dumast, M. Yatabe, A. Ruellas, M. Ioshida, B. Paniagua, M. Styner, J. R. Gonalves, J. Bianchi, L. Cevidanes, and J.-C. Prieto, "Shape variation analyzer: A classifier for temporomandibular joint damaged by osteoarthritis," *Proc. SPIE*, vol. 10950, Mar. 2019, Art. no. 1095021.

[21] K. S. Lee, H. J. Kwak, J. M. Oh, N. Jha, Y. J. Kim, W. Kim, U. B. Baik, and J. J. Ryu, "Automated detection of TMJ osteoarthritis based on Artificial Intelligence," *J. Dental Res.*, vol. 99, no. 12, pp. 1363–1367, 2020, doi: 10.1177/0022034520936950.

[22] J. Bianchi, A. C. de Oliveira Ruellas, J. R. Gonalves, B. Paniagua, J. C. Prieto, M. Styner, T. Li, H. Zhu, J. Sugai, W. Giannobile, E. Benavides, F. Soki, M. Yatabe, L. Ashman, D. Walker, R. Soroushmehr, K. Najarian, and L. H. S. Cevidanes, "Osteoarthritis of the temporomandibular joint can be diagnosed earlier using biomarkers and machine learning," *Sci. Rep.*, vol. 10, no. 1, 2020, Art. no. 8012, doi: 10.1038/s41598-020-64942-0.

[23] E. Choi, D. Kim, J.-Y. Lee, and H.-K. Park, "Artificial intelligence in detecting temporomandibular joint osteoarthritis on orthopantomogram," *Sci. Rep.*, vol. 11, no. 1, pp. 1–7, May 2021, doi: 10.1038/s41598-021-89742-y.

[24] D. Perpetuini, N. Trippetti, D. Cardone, L. Breda, M. D'Attilio, and A. Merla, "Detection of temporomandibular joint disfunction in juvenile idiopathic arthritis through infrared thermal imaging and a machine learning procedure," in *Proc. 8th EMBEC*, 2020, pp. 372–381, doi: 10.1007/978-3-030-64610-3_43.

[25] J. Y. Kim, D. Kim, K. J. Jeon, H. Kim, and J. K. Huh, "Using deep learning to predict temporomandibular joint disc perforation based on magnetic resonance imaging," *Sci Rep.*, vol. 11, no. 1, 2021, Art. no. 6680, doi: 10.1038/s41598-021-86115-3.

[26] S. Ito, Y. Mine, Y. Yoshimi, S. Takeda, A. Tanaka, A. Onishi, T. Y. Peng, T. Nakamoto, T. Nagasaki, N. Kakimoto, T. Murayama, and K. Tanimoto, "Automated segmentation of articular disc of the temporomandibular joint on magnetic resonance images using deep learning," *Sci Rep.*, vol. 12, no. 1, 2022, Art. no. 221, doi: 10.1038/s41598-021-04354-w.

[27] Y. Liu, Y. Lu, Y. Fan, and L. Mao, "Tracking-based deep learning method for temporomandibular joint segmentation," *Ann. Transl. Med.*, vol. 9, no. 6, p. 467, Mar. 2021, doi: 10.21037/atm-21-319.

[28] Z.-K. Kao, N.-T. Chiu, H.-T.-H. Wu, W.-C. Chang, D.-H. Wang, Y.-Y. Kung, P.-C. Tu, W.-L. Lo, and Y.-T. Wu, "Classifying temporomandibular disorder with artificial intelligent architecture using magnetic resonance imaging," *Ann. Biomed. Eng.*, vol. 51, no. 3, pp. 517–526, Mar. 2023, doi: 10.1007/s10439-022-03056-2.

[29] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. CVPR*, 2017, pp. 1251–1258, doi: 10.1109/CVPR.2017.195.

[30] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. CVPR*, 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[31] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. CVPR*, 2018, pp. 4510–4520, doi: 10.1109/CVPR.2018.00474.

[32] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. ICVPR*, 2016, pp. 2818–2826, doi: 10.1109/CVPR.2016.308.

[33] W. Huang, G. Peng, and X. Tang, "A limit of densely connected convolutional networks V1," Tech. Rep., 2019, doi: 10.17504/protocols.io.8j6hure.

[34] Z. Liu, H. Mao, C. Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A ConvNet for the 2020s," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 11966–11976, doi: 10.1109/CVPR52688.2022.01167.

[35] *Towards Data Science, Medium*. Accessed: Feb. 18, 2022. [Online]. Available: https://towardsdatascience.com/transfer-learning-from-pre-trained-models-f2393f12475

[36] E. Sivari, M. S. Güzel, E. Bostanci, and A. Mishra, "A novel hybrid machine learning based system to classify shoulder implant manufacturers," *Healthcare*, vol. 10, no. 3, p. 580, Mar. 2022, doi: 10.3390/healthcare10030580.

[37] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.

[38] Google AI Blog. *MobileNetV2: The Next Generation of On-Device Computer Vision Networks*. Accessed: Feb. 9, 2022. [Online]. Available: https://ai.googleblog.com/2018/04/mobilenetv2-next-generation-of-on.html

[39] Wikimedia Foundation. *InceptionV3(2020)Wikipedia*. Accessed: Feb. 4, 2022. [Online]. Available: https://en.wikipedia.org/wiki/Inceptionv3

[40] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. NIPS*, vol. 30, 2017, pp. 5998–6008.

[41] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An image is worth 16×16 words: Transformers for image recognition at scale," 2020, *arXiv:2010.11929*.

[42] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. ICCV*, 2017, pp. 618–626, doi: 10.1109/ICCV.2017.74.

[43] Google. *Google Colab*. Accessed: May 1, 2023. [Online]. Available: https://research.google.com/colaboratory/faq.html

[44] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[45] M. K. Murphy, R. F. MacBarb, M. E. Wong, and K. A. Athanasiou, "Temporomandibular disorders: A review of etiology, clinical management, and tissue engineering strategies," *Int. J. Oral Maxillofacial Implants*, vol. 28, no. 6, p. e393, 2013, doi: 10.11607/jomi.te20.

[46] Y. Feng, J. Shu, Y. Liu, T. Zheng, B. Shao, and Z. Liu, "Biomechanical analysis of temporomandibular joints during mandibular protrusion and retraction motions: A 3D finite element simulation," *Comput. Methods. Programs. Biomed.*, vol. 208, Sep. 2021, Art. no. 106299, doi: 10.1016/j.cmpb.2021.106299.

[47] B. Shao, H. Teng, S. Dong, and Z. Liu, "Finite element contact stress analysis of the temporomandibular joints of patients with temporomandibular disorders under mastication," *Comput. Methods. Programs. Biomed.*, vol. 213, Jan. 2022, Art. no. 106526, doi: 10.1016/j.cmpb.2021.106526.

[48] M. Mediouni, D. R. Schlatterer, H. Madry, M. Cucchiarini, and B. Rai, "A review of translational medicine. The future paradigm: How can we connect the orthopedic dots better?" *Current Med. Res. Opinion*, vol. 34, no. 7, pp. 1217–1229, Jul. 2018, doi: 10.1080/03007995.2017.1385450.

[49] M. Mediouni, R. Madiouni, M. Gardner, and N. Vaughan, "Translational medicine: Challenges and new orthopaedic vision (Mediouni-Model)," *Current Orthopaedic Pract.*, vol. 31, no. 2, pp. 196–200, 2020, doi: 10.1097/bco.0000000000000846.

SIFA OZSARI received the B.Sc. and M.Sc. degrees from the Computer Engineering Department, Selçuk University, Turkey, in 2016. She is currently pursuing the Ph.D. degree with the Computer Engineering Department, Ankara University. She is also a Research Assistant with the Computer Engineering Department, Ankara University. Her research interests include optimization, machine learning, and computer vision.

FATIMA RABIA YAPICIOGLU received the joint B.Sc. degree in computer engineering from the Engineering Faculty (Exchange Studies), Ankara University, and the School of Computing, University of Eastern Finland, in 2021. She is currently pursuing the M.Sc. degree with the University of Padova, Italy. She received the Ministry of Foreign Affairs and International Cooperation Scholarship for two consecutive years, and has been awarded with Mille e una Lode honor degree and best machine learning project from Galileo Galilei Physics with the Data Department, UNIPD, during her M.Sc. studies. She is currently doing research in collaboration with Kings College London on dopamine hormone brain lateralization in healthy controls and schizophrenia patients compromising explainable artificial intelligence (XAI). Her past experiences include contributing to medical applications and artificial intelligence convergence with the ICT Medical Development Center, Kuopio University Hospital, Finland. She is also an Artificial Intelligence Specialist Intern with Baker Hughes American company, which has subsidiaries GE and Bently Nevada. Her research interests include artificial intelligence for life sciences and health applications, explainable artificial intelligence, computer vision, time series analysis, and machine learning.

DILEK YILMAZ received the B.S. degree in dentistry from the Dentistry Faculty, Ankara University, Turkey, in 1993. She is currently an Assistant Professor with the Department of Oral and Maxillofacial Radiology, Faculty of Dentistry, Başkent University. She has published multiple academic papers about temporamandibular joint MRI, ultrasound, CBCT, chewing muscles, styloid process, and similar topics. She has made multiple presentations in national and international congresses. She received her expertise in oral and maxillofacial radiology from the Dentistry Faculty, Ankara University, in 2018.

KIVANC KAMBUROGLU received the D.D.S. degree from the Faculty of Dentistry, Ankara University, Ankara, Turkey, in 1999, the M.Sc. degree from the Oral Diagnosis and Radiology Department, Faculty of Dentistry, Ankara University, in 2002, and the Ph.D. degree the Faculty of Dentistry, Ankara University, in 2007. He was a Research Fellow with the Department of Oral Pathology and the Oral Medicine and Oral Radiology Department, Maurice and Gabriela Goldschleger School of Dental Medicine, Tel Aviv University, from 2005 to 2007. From December 2007 to December 2008, he was a second Lieutenant Doctor with the Turkish Army; Gülhane Military Medical Academy, Dental Sciences Center, Ankara. From April 2011 to April 2012, he was a Visiting Postdoctoral Researcher with the Oral and Maxillofacial Radiology and Imaging Department, Faculty of Dentistry, Louisville University, Louisville, Kentucky, USA. He is currently a full-time Faculty Member with the Department of Dentomaxillofacial Radiology, Faculty of Dentistry, Ankara University, where he was a Professor in dentomaxillofacial radiology, in 2016. He is well published in the field of dentomaxillofacial radiology and serves as a reviewer and a board member for several national and international journals. He is a member of IADMFR, EADMFR, AAOMFR, and Asian AOMFR.

MEHMET SERDAR GUZEL received the B.Sc. degree from the Computer Engineering Department, Başkent University, in 2004, the M.Sc. degree from the Computer Engineering Department, Ankara University, Turkey, and the Ph.D. degree from the Mechanical and Systems Engineering Department, Newcastle University, U.K., in 2012. He is currently an Associate Professor with the Computer Engineering Department, Ankara University. He has contributed to several international projects and published many academic papers in his research fields. His research interests include machine learning, swarm intelligence, computer vision, and robotics.

GAZI ERKAN BOSTANCI received the B.Sc. degree from the Computer Engineering Department, Ankara University, Turkey, in 2007, the M.Sc. degree in real-time battlefield simulation from Ankara University, in 2009, and the Ph.D. degree from School of Computer Science and Electronic Engineering, University of Essex, U.K., in 2014. His Ph.D. thesis on real-time user tracking for augmented reality. He joined as a Research Assistant with Ankara University. He was with the Gendarmarie School Command as a Planning Officer Designate, in June 2014, where he conducted the research for developing a vision-based system for analyzing crime scenes. He has been promoted to a second lieutenant, in January 2015. Having completed his military service, he currently continues his post with Ankara University as an Assistant Professor. In 2018, he was promoted to an Associate Professor. His research interests include the different yet closely related aspects of computer science from image processing, computer vision, and graphics to artificial intelligence and fuzzy logic, and mathematical modeling and statistical analysis. He recently developed a vision-based user tracking system for various augmented reality applications for cultural heritage in particular. He setup the SAAT Laboratory, department for conducting research with the main aim for incorporating AI approaches, such as fuzzy decision making and evolutionary computation for solving a wide range of real-world problems from dental trauma diagnosis to finding optimal parameters in high energy physics problems. He has been involved in technical committees for several conferences and organizing an international conference for several years and acted as a reviewer for various journals.
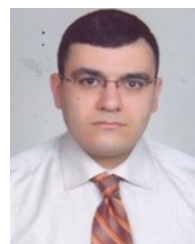
KORAY ACICI received the B.S., M.S., and Ph.D. degrees in computer engineering from Başkent University, Turkey, in 2005, 2015, and 2021, respectively. He was an Assistant Professor with the Department of Computer Engineering, Başkent University. He is currently an Assistant Professor with the Department of Artificial Intelligence and Data Engineering, Ankara University, Ankara, Turkey. His research interests include bioinformatics, medical informatics, artificial intelligence, and machine learning.

TUNC ASUROGLU received the B.S. degree in computer engineering from the TOBB University of Economics and Technology, Turkey, in 2013, and the M.S. and Ph.D. degrees in computer engineering from Başkent University, Turkey, in 2015 and 2020, respectively. He was an Assistant Professor with the Department of Computer Engineering, Başkent University. He was also a Guest Researcher with the Faculty of Computer Sciences, Østfold University College, Norway. He worked on many international projects related to health informatics. He is currently a Postdoctoral Research Fellow with the Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland. His research interests include the applications of computational intelligence in health informatics and wearable sensor systems.

● ● ●