2022 IEEE North Karnataka Subsection Flagship International Conference (NKCon)

Prediction of severity of Knee Osteoarthritis on X-ray images using deep learning

[1] Sajeev Ram Arumugam
Department of AI&DS
Sri Krishna College of Engineering and Technology
Coimbatore, India
imsajeev@gmail.com

[2] Balakrishna R
Department of AI&DS
Vel Tech High Tech Dr. Rangarajan Dr. Sakunthala Engineering College,
Chennai, India
rbalakrishna@velhightech.com

[3] V. Rajeshram
Department of Computer Science and Engineering
M.Kumarasamy College of Engineering,
Karur, India
rajeshram107@gmail.com

[4] Sheela Gowr
Department of CSE
Vels Institute of Science Technology and Advanced Studies(VISTAS)
Chennai, India
sheela.se@velsuniv.ac.in

[5] Sankar Ganesh Karuppasamy
Department of CSE
Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology
Chennai, India
prof.sankarganesh@gmail.com

[6] S.P. Premnath
Department of ECE
Sri Krishna College of Engineering and Technology
Coimbatore, India
premnathsp@yahoo.com

2022 IEEE North Karnataka Subsection Flagship International Conference (NKCon) | 978-1-6654-5342-4/22/$31.00 ©2022 IEEE | DOI: 10.1109/NKCON56289.2022.10126934

***Abstract*— The most prevalent kind of arthritis is osteoarthritis (OA). Radiologists use to employ the Kellgren–Lawrence (KL) grading system to identify the aggressiveness of OA based on the information shown on the pair of knee joints. Computer-assisted strategies have recently been proposed to enhance the accuracy of OA diagnosis. Previous semiautomatic segmentation approaches, on the other hand, required human interaction, limiting their use on huge datasets. Furthermore, CNN is used to quantify OA rigorousness to investigate the relationships among distinct local regions. SSD reduces human interaction and provides a back-to-back approach to computerized osteoarthritis detection by incorporating the object detection model. The rating is based on X-ray scans from the Osteoarthritis Initiative (OAI) dataset. At the cost of training on a huge dataset with over 8260 knee joint samples, our method accurately segments 96.37% of data.**

Keywords*— *Knee osteoarthritis, Kellgren–Lawrence (KL) grading system), CNN (Convolution Neural Network), SSD (Single-Shot Multibox Detector)

I. INTRODUCTION

Osteoarthritis is a disorder in which the cartilage that acts as a cushion among joints wears away. As a result, the bones of the joints push closer together, and cartilage shock-absorbing properties are lost. The rubbing causes discomfort, distension, painfulness, limited mobility, and the production of bone spurs in some cases [1]. The disorder can distress some joints in the body, but the knee is primarily influenced. Knee arthritis affects ordinary actions like walking and mounting difficult. For many people, it is a major source of time off work and serious impairment. While arthritis is mostly an adult disease, it does impact youngsters in some cases. Even though there is no treatment for arthritis, there are a variety of therapeutic choices that can assist people to manage their discomfort and stay active. While aging is a significant threat factor for knee osteoarthritis, it can also affect young people. It could be hereditary in some cases [2]. Other knee osteoarthritis develops due to wound, infection, or being overweight. A physical examination by the physician will be the first step in diagnosing knee osteoarthritis. The physician will also take your health records and make a note of any signs you are experiencing. Additional tests, such as X-rays, which can reveal bone and cartilage deterioration and the existence of bone spurs, and magnetic resonance imaging scans, as suggested by the doctor.

The knee is the body's longest and most powerful joint. The patella(kneecap) is the combination of the lower end of the femur and the upper end of the tibia. Articular cartilage is a smooth, flexible structure that surrounds the extremities of the three bones that make up the knee joint. It supports and guards the bones when you turn and raise your knee. In the middle of femur and tibia, two wedge-shaped pieces of cartilage known as meniscus operate as shock absorbers. They are robust and stretchy to support the joint [3]. The synovial membrane is a narrow layer that protects the knee joint. It secretes a lubricating liquid that coats the cartilage. Osteoarthritis is a deteriorating, "wear-and-tear" kind of arthritis that mostly influences persons fifty and above, however, it can influence adolescent people. The cartilage in the knee joint finally erodes in osteoarthritis. The cartilage becomes jagged and sharp as it wears away, and the protected area among the bones shrinks. This could cause severe pain bone spurs because bone is rubbing up on bone. Osteoarthritis typically develops over time, and the pain it causes intensifies. Fig.1 shows the difference between normal knees affected by osteoarthritis [4].

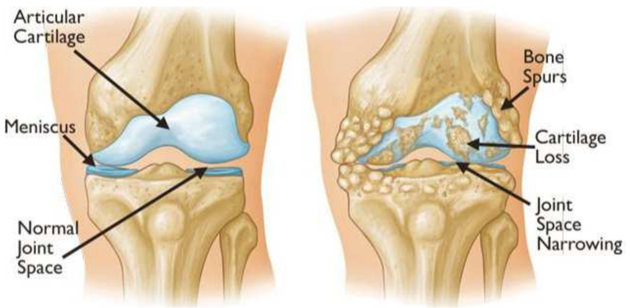

Fig. 1. Comparison between knee and knee osteoarthritis

Osteoarthritis (OA) is the second leading rheumatologic disorder and perhaps the most common joint illness in India, with an incidence of 22% to 39%. Women are more likely than men to have OA, although the prevalence rises drastically with age [5]. Nearly half of all women above the age of 65 have signs, and 70% of those over the age of 65 show radiographic evidence. Knee osteoarthritis (OA) is a prevalent reason for the low mobility, especially in women. The 10th greatest cause of nonfatal burden was projected to be OA. Clinical factors, which make up the inclusion criteria for most clinical trials in this sector, can be used to classify knee OA. National clinical recommendations in Sweden and Denmark likewise support using clinical criteria to diagnose OA, stating specifically that radiography is not required for an OA diagnosis. The focus should be on OA to develop clinically relevant studies and treatment techniques that can help to alleviate the disease's enormous burden [6]. To that purpose, studies of the symptoms of knee OA, frequently in conjunction with structural changes, are becoming increasingly popular.

II. LITERATURE SURVEY

Each section of the image was utilized to extract time-series features, which were then used to train, test, and validate the system. Data from four lower limb joints collected using a four-channel Surface Electromyography (SEMG) system were combined with goniometer signals to determine knee range of motion. Statistical features were collected from the EMG and goniometer data and used to train, test, and validate the system [7]. SEGM and ANN algorithms can be used to diagnose knee disease. Deep convolutional neural network models that have been trained on ImageNet and calibrated on knee OA images can dramatically enhance classification accuracy [8]. The precision of computerized knee OA severity prognosis should be evaluated using a continuous distance-based evaluation measure such as MSE rather than classification accuracy. As a result, KL grade prediction is framed as a regression problem, which enhances accuracy.

The proposed method may be designed to automatically differentiate between persons who will get osteoarthritis and those who will not by detecting precise biochemical changes in the middle of the knee's cartilage. The output provide pre-symptomatic individuals with baseline imaging as well as a decrease in liquid concentration [9]. The method employs a fully convolutional neural network to identify the knee joints automatically. Convolutional neural networks (CNN) are being utilized to train from the ground up to estimate the knee OA severity [10]. This joint training enhances the complete assessment of knee OA severity while also delivering instantaneous multi-class classification and regression outputs naturally.

The Kellgren-Lawrence grading scale is used to grade knee OA seriousness using the Deep Siamese Convolutional Neural Network [11]. The system was trained entirely using data from the Multicenter Osteoarthritis Study dataset from the Osteoarthritis Initiative. As a result of this knowledge, the practitioner's decision-making process becomes more visible, resulting in increased faith in automated approaches. Given similar inputs, the DNN beats existing classification algorithms in OA classification. Local prediction models will receive more attention, as they are trained on data segments characterized by metrics like BMI paired with demographics and socioeconomic indices [12]. Open data and scientific tools for OA diagnostics that use machine learning and deep learning techniques are extremely auspicious and should be actively promoted within the OA exploration community. The relationships between distinct local regions are not explored by the CNN architectures used to determine OA severity. YOLO provides a back-to-back method to diagnose osteoarthritis automatically by combining the object detection model with the visual transformer[13]. The classification performance is improved by the visual transformer blocks with CNN architecture.

III. PROPOSED METHOD

A. Dataset

We used data from the National Institutes of Health-sponsored Osteoarthritis Initiative (OAI) (https://nda.nih.gov/oai/), which is a multi-centre, 10-year experimental study of males and females. The mission of OAI is to provide tools that will help people understand how to prevent and manage knee osteoarthritis, which is the primary reason of debility in adults. The Osteoarthritis Initiative (OAI), which contains 4796 individuals ranging in age from 45 to 79 years old, provided the knee X-ray pictures used in the study. There are 8260 knee joints in 4130 X-ray images. The following Kellgren-Lawrence category [14] was allotted to the knees in this dataset:

Category 0: There are no radiological signs of OA.
Category 1: probable osteophytes lipping and dubious joint space narrowing (JSN).
Category 2: osteophytes with a chance of JSN.
Category 3: Multiple osteophytes, obvious JSN, sclerosis, and potential bone deformity.
Category 4:Large osteophytes, JSN, severe sclerosis, and evident bone deformity.

We divided every knee X-rays into 3105:173:173 training, validation, and testing sets at random. To sustain constant grade distribution among training, testing and verification sets, this split-up is done grade-by-grade based on the KL grade of the knee joint in an X-ray image.

B. Knee Joint Recognition

Computer vision approaches for locating and labelling things are referred to as object detection. Object detection algorithms can be used to detect objects in both static and moving images. Computer vision techniques are widely used in medical fields, as they may provide useful insight into a variety of disorders and help detect them quickly. Detection is the technique of identifying certain objects in the OA affected place from X-Ray images of the knee. Classification is the process of determining if the output image is OA impacted or not. Object detection is commonly used in the medical field to detect disorders. Other object detection approaches are used in conjunction with neural network-based classifiers. An SSD (Single-Shot Multibox Detector) using Mobilenet-model was trained and tested using TensorFlow Object Identification API, an open-source framework for activities that involve object detection[15]. A dataset of OA X-Ray images was used to test the model as pre-trained and with fine-tuning. Fig.2 shows the four categories of images.

This proposed work is based on single-shot detection (SSD). A VGG model is usually used to start the SSD, which is then turned into a FCN. We include more convolutional layers to help with huge objects. A 38x38 feature map is produced by the VGG network. The further layers provide feature maps that are 19x19, 10x10, 5x5, 3x3, and 1x1. At various sizes, all of the feature maps are utilised to forecast bounding boxes. A sub-network that works as a classifier and a localizer receives some of the activations.

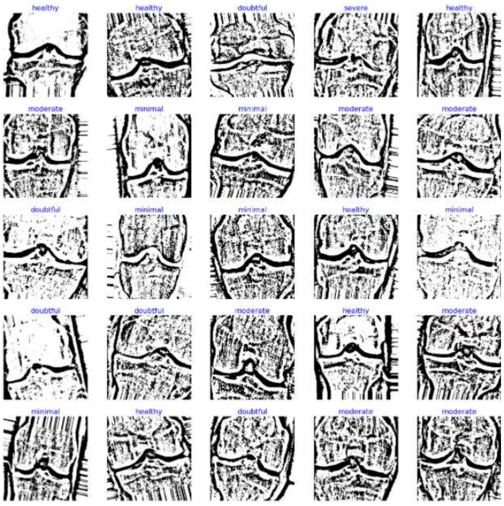

Fig. 2. Knee Osteoarthritis Images

C. *Classification*

Convolution is a mathematical procedure on two legitimate argument functions. The input is referred to as the first function, while the kernel is referred to as the second function in imaging language. A feature map is the result of this procedure [16]. CNN models are deep neural network hierarchies, with hidden layers between the general layers to support the system's weights in learning more about the features detected in the input image. A pooling operation is performed by another sort of layer that is widely utilised in CNN designs. The spatial resolution is lowered as a result, and only the most important elements are retained. This is necessary to keep the network size manageable. There are four major steps in the construction of a Convolutional Neural Network. Fig.3 demonstrates the proposed work.

- Convolution
- Pooling
- Flattening
- Full connection

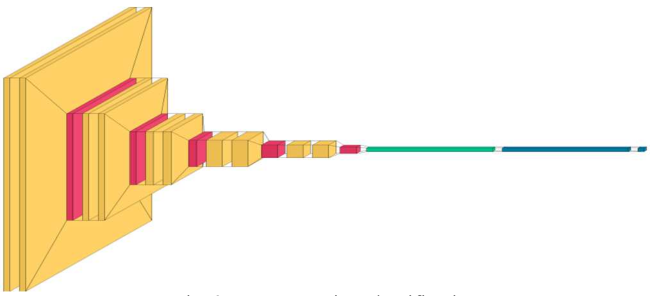

Fig. 3. Knee Severity Classification

Because these operations form the cornerstone of any Convolutional Neural Network, we must properly understand how they operate to gain a thorough understanding of how these ConvNets work. ConvNets get their name from the operator 'Convolution.' The extraction of input picture features is the main goal of this operator. Convolution uses small squares of input data to learn visual attributes and functions in tandem with pixels. The convolution of the image can be calculated by performing element-wise exponentiation among the two matrices for each point and summing the exponentiation outputs to reach the concluding element of the output matrix. Filters are used to detect features in the input image. Every convolution operation in a typical ConvNet is followed by ReLU. ReLU is a feature map procedure that works element by element. T the maximum real-life data we feed into our ConvNet is nonlinear, it serves the aim of bringing nonlinearity into ConvNet. Non-linear functions, such as sigmoid, can be used in place of ReLU for operation.

The key purpose of pooling is to decrease the structural size of the input. Pooling achieves the following results in particular:

1. Creates input representations, i.e. features with modest dimensions that are easier to manage.
2. Moderates the size of the parameter used and calculations performed in the network to more effectively control overfitting.
3. Makes the network robust to distortions, translations, and minor modifications in the input image. Even if there are minor distortions in the input, the pooling output will not be distorted because the average/maximum value is used in the local area.
4. It assists in the creation of an equivariant representation of the input image. This is highly significant since it allows us to recognize little items in our image regardless of their location.

The essential structure of any CNN is made up of the workings of Convolution, ReLU, and Pooling up to layers. There are two sets of Convolution, ReLU, and Pooling layers; the second convolution layer employs six filters to convolution the output of the first Pooling layer, resulting in six feature maps. The ReLU layer is applied to every six feature maps distinctly. Coordination between these layers aids in the extraction of meaningful features from the images fed then provides non-linearity to the network and decreases feature dimension, making the features scale and translate equivariant [17]. The out-turn of the second pooling layer provides input to the Fully Connected Layer. The fully connected layer is defined as a Multilayer Perceptron that uses the output layer's softmax activation function. Each neuron in the following layer is related to each neuron in the preceding layer. The outputs of the pooling and convolutional layers show high-resolution aspects of the input image. These features are used by the fully connected layer, which is based on the training dataset, to categorize incoming photos into different classes. Instead of utilizing classification, introducing a fully-connected layer offers a more general technique of learning non-linear groupings of such features. The convolutional and pooling layers features are useful for classification and the Softmax activation function in the output layer. This function turns an arbitrary real-valued vector into a vector with a value between one and zero, yielding a sum of one. The Softmax function computes the event's probability distribution over 'n' distinct events. Overall potential target classes, this function will determine the odds of each target class.

IV. RESULTS AND DISCUSSIONS

In this study, we used the Kellgren- Lawrence grading scheme to construct a deep learning model for the computerized evaluation of knee osteoarthritis extremity. We evaluated our model to the OAI dataset's KL score evaluation. The Single Shot Multibox Detector (SSD) is a fast enough detector for real-time applications. However, this comes at the expense of precision. SSD with MobileNet denotes a model having SSD as the model meta-architecture and MobileNet as the feature extractor type. SSD achieves a great balance of speed and precision. SSD computes a feature map by running a neural network on input images only once. Fig.2 provides an example of the knee detection results.

An epoch is a unit of measurement for the amount of time it takes to train a neural network utilizing the training data. We utilize all the data just once. One pass consists of a forward and backward pass. Since one epoch is too enormous to send to the computer, we split it into several smaller batches. We employed 20 epochs in this scenario.

Fig. 4(a) Model Loss

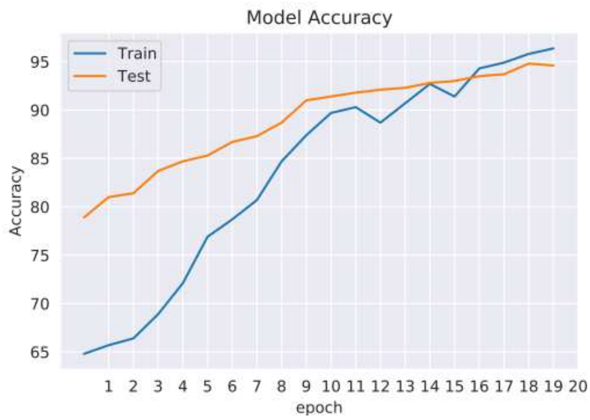

Fig. 4(b) Model Accuracy

Loss functions define how similar an expected value is to the actual one. A loss function associates assessments to the costs connected with them. Loss functions are dynamic, changing based on the job at hand and the desired outcome. It is a way of determining how well the algorithm models the data. Fig. 3.a shows model loss at every epoch and Fig 3.b shows model accuracy at every epoch.

In classification problems, the error can be measured as the number of actual and inaccurate positivity compared to the number of false detections. Positivity and negativity are the basis for two popular information retrieval techniques (1) and (2):

$$Precision = \frac{Actual\ positivity}{Actual\ Positive + Inaccurate\ Positivity} \quad \ldots (1)$$

$$Recall = \frac{Actual\ Positivity}{Actual\ Positivity + Inaccurate\ Negativity} \quad \ldots (2)$$

The F-score is a measurement of accuracy. It is determined by splitting the number of accurate positivity results to the number of all positivity results (3), which include those that were inaccurately recognized, and the recall by the number of true positivity findings divided by the number of all samples that should have been detected as positivity.

$$F1 - score = \frac{2}{recall^{-1} + precision^{-1}} \quad \ldots (3)$$

Classification Report:

	precision	recall	f1-score	support
minimal	0.95	0.95	0.95	38
healthy	0.93	0.95	0.94	42
moderate	0.95	0.95	0.95	39
doubtful	0.98	0.95	0.97	44
severe	0.90	0.90	0.90	10
accuracy			0.95	173
macro avg	0.94	0.94	0.94	173
weighted avg	0.95	0.95	0.95	173

Fig. 5. Classification Report of The Proposed System

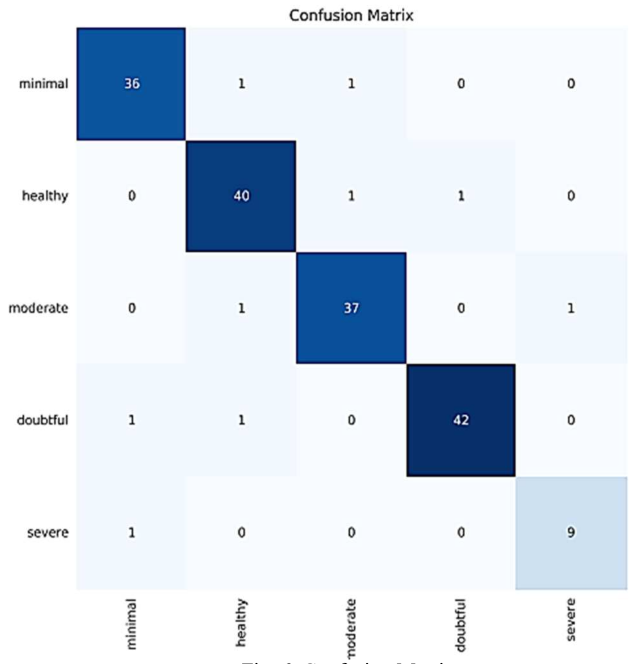

Fig. 6. Confusion Matrix

The proposed method classification result is shown in Fig.4. The system attains better precision, recall, and f1-score value compared to existing methods. A confusion matrix is a process of describing the classification efficiency. If there is an uneven number of instances in every class or if the dataset has more than two classifications, accuracy can be misleading. Generating a confusion matrix helps in determining when the classification model is succeeding and how it is failing. It not only notifies about the errors produced by the classifier but also about the errors types made by the classifier. Fig.5 shows the confusion matrix of the proposed system.

V. CONCLUSION AND FUTURE SCOPE

In this research, we proposed a deep learning-based highly automated technique for diagnosing knee osteoarthritis. We proposed a fully automated deep learning-based method for diagnosing knee osteoarthritis in this research. The segmentation of the knee joint area was done using transfer learning. We classified the severity of OA in different regions of the X-ray images using the SSD model. The proposed solution beats earlier methods in terms of OA severity categorization, according to the results of the experiments. Different phases of knee osteoarthritis necessitate different treatments, determining the grade of osteoarthritis affecting a patient is critical. By reliably allocating KL grades to knee radiographs, the suggested technique helps in the reduction of human perception in grading radiographs, lessens radiologist labour stress, and improves reporting times.

REFERENCES

[1] Y. Y. Chow and K. Y. Chin, "The Role of Inflammation in the Pathogenesis of Osteoarthritis," *Mediators Inflamm.*, vol. 2020, 2020, doi: 10.1155/2020/8293921.

[2] H. H. Dragan Primorac , Vilim Molnar , Eduard Rod Željko Jeleˇc , Fabijan ˇCukelj , Vid Matiši´c , Trpimir Vrdoljak , Damir Hudetz and and I. Bori´c, "Knee Osteoarthritis : A Review of Pathogenesis and," *Genes (Basel).*, vol. 11, no. 8, pp. 854–889, 2020.

[3] C. Kokkotis, S. Moustakidis, E. Papageorgiou, G. Giakas, and D. E. Tsaopoulos, "Machine learning in knee osteoarthritis: A review," *Osteoarthr. Cartil. Open*, vol. 2, no. 3, p. 100069, 2020, doi: 10.1016/j.ocarto.2020.100069.

[4] "Arthritis of the Knee - OrthoInfo - AAOS." https://orthoinfo.aaos.org/en/diseases--conditions/arthritis-of-the-knee/ (accessed Oct. 24, 2021).

[5] C. P. Pal, P. Singh, S. Chaturvedi, K. K. Pruthi, and A. Vij, "Epidemiology of knee osteoarthritis in India and related factors," *Indian J. Orthop.*, vol. 50, no. 5, pp. 518–522, 2016, doi: 10.4103/0019-5413.189608.

[6] E. M. Roos and N. K. Arden, "Strategies for the prevention of knee osteoarthritis," *Nat. Rev. Rheumatol.*, vol. 12, no. 2, pp. 92–101, 2016, doi: 10.1038/nrrheum.2015.135.

[7] J. de D. Uwisengeyimana and T. Ibrikci, "Diagnosing Knee Osteoarthritis Using Artificial Neural Networks and Deep Learning," *Biomed. Stat. Informatics*, vol. 2, no. 3, pp. 95–102, 2017, doi: 10.11648/j.bsi.20170203.11.

[8] S. Ram Arumugam, S. Ganesh Karuppasamy, S. Gowr, O. Manoj and K. Kalaivani, "A Deep Convolutional Neural Network based Detection System for Autism Spectrum Disorder in Facial images," 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2021, pp. 1255-1259, doi: 10.1109/I-SMAC52330.2021.9641046.

[9] R. K. Et.al, "Analysis On Deep Learning Approaches For Timely Detection Of Osteoarthritis," *Turkish J. Comput. Math. Educ.*, vol. 12, no. 3, pp. 4396–4405, 2021, doi: 10.17762/turcomat.v12i3.1818.

[10] J. Antony, K. McGuinness, K. Moran, and N. E. O'Connor, "Automatic detection of knee joints and quantification of knee osteoarthritis severity using convolutional neural networks," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10358 LNAI, pp. 376–390, 2017, doi: 10.1007/978-3-319-62416-7_27.

[11] A. Tiulpin, J. Thevenot, E. Rahtu, P. Lehenkari, and S. Saarakkala, "Automatic Knee Osteoarthritis Diagnosis from Plain Radiographs : A Deep Learning-Based Approach," *doi 10.1038/s41598-018-20132-7. PMID 29379060; PMCID PMC5789045.*, pp. 1–10, 2018, doi: 10.1038/s41598-018-20132-7.

[12] E. Christodoulou, "Exploring deep learning capabilities in knee osteoarthritis case study for classification," *2019 10th Int. Conf. Information, Intell. Syst. Appl.*, vol. 00, no. c, pp. 1–6, 2019.

[13] Y. Wang, X. Wang, T. Gao, L. Du, and W. Liu, "An Automatic Knee Osteoarthritis Diagnosis Method Based on Deep Learning : Data from the Osteoarthritis Initiative," *J. Healthc. Eng.*, vol. 2021, 2021.

[14] H. Knipe and V. Pai, "Kellgren and Lawrence system for classification of osteoarthritis," *Radiopaedia.org*, Jan. 2014, doi: 10.53347/RID-27111.

[15] S. R. Arumugam, R. Balakrishna, R. Khilar, O. Manoj and C. S. Shylaja, "Prediction of Autism Spectrum Disorder in Children using Face Recognition," 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC), 2021, pp. 1246-1250, doi: 10.1109/ICOSEC51865.2021.9591679.

[16] H. Chen *et al.*, "A deep learning CNN architecture applied in smart near-infrared analysis of water pollution for agricultural irrigation resources," *Agric. Water Manag.*, vol. 240, no. April, p. 106303, 2020, doi: 10.1016/j.agwat.2020.106303.

[17] A. Pingjun, C. Linlin, G. Xiaoshuang, S. Kyle, and L. Yang, "Ac ce p te d us ip t," *Comput. Med. Imaging Graph.*, 2019, doi: 10.1016/j.compmedimag.2019.06.002.