Journal of Information Systems Engineering and Management
2025, 10(5s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Detection of Knee Osteoarthritis grade using Convolutional Neural Networks

Annu James1*, C. Beulah Christalin Latha2, Arun Thachil3, J. Macklin Abraham Navamani4
[1]R1Research Scholar, Karunya Institute of Technology and Sciences, Coimbatore, Tamilnadu, India.
2Faculty, Karunya Institute of Technology and Sciences, Coimbatore, Tamilnadu, India.
3Professor, College of Physiotherapy, Medical Trust Institute of Medical Sciences, Kerala, India.
4Faculty, Karunya Institute of Technology and Sciences, Coimbatore, Tamilnadu, India.
*Corresponding Author-Annu James

ARTICLE INFO

Received: 10Oct 2024
Revised: 12Dec 2024
Accepted:24Dec 2024

ABSTRACT

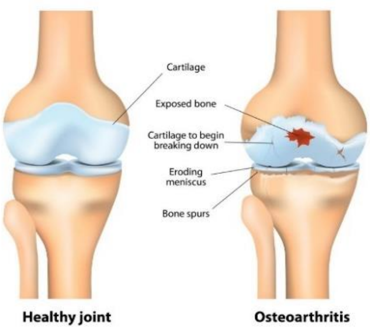
Knee OsteoArthritis (KOA) is a disease that affects a person's quality of life. Early detection and monitoring of KOA progression is essential for effective therapy and quick recovery. A survey of the recent literatures indicates that deep learning methods can effectively assess KOA severity with improved accuracy and efficiency. Convolutional Neural Networks (CNN) help us to classify the levels of severity of Knee Osteoarthritis. The present study proposes a deep learning method in classification of osteoarthritis using Convolution Neural Networks. The study focuses on predicting the grades of input images with KL grades of Knee osteoarthritis. The study employs convolutional neural networks with the Rectified Linear Unit function (ReLU), activation function and Adam optimization algorithm to achieve high performance. The study evaluates 10 performance measures and the results indicate an improvement in performance measures when compared with existing techniques.

Keywords: Knee Osteoarthritis (KOA), kellgren-lawrence classification, convolutional neural network, medical image processing, deep learning.

INTRODUCTION

Knee Osteoarthritis (OA) is a wear &tear disease that is commonly occurring in people above the age of 60 due to aging. Primary OA is seen in old people, mainly on the weight bearing joints (hip and knee). Secondary OA refers degeneration of the joints. According to pathology, knee osteoarthritis primarily affects the articular cartilage as shown in Fig 1(a). Formation of subchondral cysts and sclerosis is common. The articular cartilage having loose flakes will cause synovial (fluid inside the joint) inflammation. In the knee joint the medial compartment (inside of knee) is affected more than the lateral (outside of the knee) [1] refer Fig 1(b). In the diagnosis of osteoarthritis using radiological methods, we may note features like Joint Space Narrowing (JSW), sclerosis, cysts, formation of osteophyte, loosely packed bodies and deformity of joints.

KOA contributes to 83% of the total osteoarthritis. OA is commonly observed as age advances. Knee Osteoarthritis affects 33%-10% men and 13% of women in the United States who are aged over 60 [2]. More than 27 million are affected in the United States by KOA. KOA has a prevalence of more than 33% in India. Knee Osteoarthritis is common in obese individuals especially females above 60 years of age. KOA is ranked three among musculoskeletal disease and ranked 11th highest global disability factor affecting more than 300 million people in 2017 [3]. The early symptoms of Osteoarthritis include knee pain, knee instability, and difficulty in performing activities of daily living.



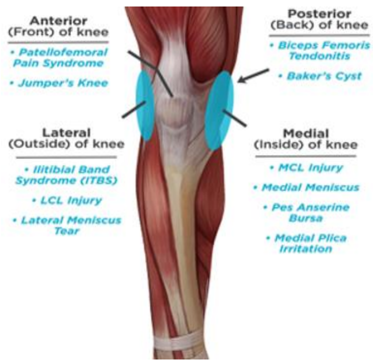


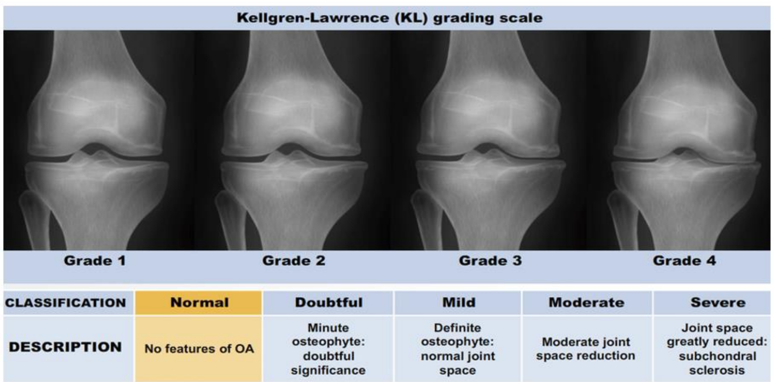
Fig1 (a) Images of healthy joint and Osteoarthritis   Fig1 (b) Images of Knee compartments

The impact of KOA is classified using the Kellgren and Lawrence (KL) scaling system which classifies Knee osteoarthritis in a grading scale, ranging from KL 0-4. KL was originally used for antero-postero knee radiographs. Severity of osteoarthritis using 5 grades as shown in Table 1 and pictorially represented in Fig 2.

Table 1: Different KL grades

KL Grade	Severity of KOA
KL-0(None)	Absence of KOA
KL-1(Doubtful)	possible joint space decrease, chance of osteophytic kissing
KL-2(Mild)	osteophyte and possible joint space decrease
KL-3(Moderate)	Many osteophytes, considerable decease of joint space, sclerosis and chance of deformity.
KL-4(Severe)	Many osteophytes, severe decrease in joint space, severe sclerosis and definite irregularity

Fig 2: Different classification of Knee Osteoarthritis: Source: http://www.adamondemand.com/clinical-management-of-osteoarthritis/



2. Literature Review

In the traditional approach, radiography, clinical characteristics, and physical findings were employed to assess the severity of KOA. Orthopaedists and physiotherapists classified it. Due to human mistake, a wrong diagnosis could occur. The Kellgren & Lawrence (KL) grading system assigns a grade or category to the severity of knee OA [4]

Machine learning techniques were applied to different features of knee and was used to predict pain in knee osteoarthritis patients [5]. Automated system with active shape model and bone textures detected OA has improved accuracy [6]. In order to detect KOA in machine learning, calculated Joint space width is compared to a standard JSW value specified [7]. Gait features also help in detecting KOA in machine learning [8]. KL grades can be better than techniques using gait data analysis [9,10] because there is no need of physical analysis of gait patterns.

Deep learning techniques has more accuracy than manual techniques [11]. Deep learning-based algorithm matches radiologists' performance in assessing the severity of knee osteoarthritis [12]. Developing Automatic KOA Assessment Methods using Highly Confident Samples was more effective than those without it [13].

CNNs are more effective than fully connected networks and more like brain neurons. The flexibility of CNN has shown to be able to overcome the flaws of earlier techniques. CNN shows more accuracy in diagnosing breast cancer than MLP according to paper [24] when compared to other methods. A well-known method for classifying images is CNN [14]. CNN can be used to localize the knee joint in knee X-ray image [15]. A new way using the CNN-LSTM algorithm to classify the grades of knee osteoarthritis from radiographic images [16].

CNN were used to overcome the shortcomings of the preceding methods, such as poor test phase optimization, particular denoising models, manual parameter settings, etc [17]. Lots of advantages of CNN are listed in paper [18]. CNN works well in robotics [19], image recognition [20], facial expression [21], self-driving [22], etc.

3. Methodology

CNN will be used in computer vision so that we may decrease the number of features. In this study, CNN with ReLU, MaxPooling, and fully connected layers are used. A convolutional neural network has four layers. The layers are explained as follows.

Input Layer: Takes input images then scale them for feature extraction.

Convolution Layer: It will filter images which helps us in finding out features and for calculating the matching features at the time of testing. Here absence of pixel intensity is represented by -1 and presence is represented by 1. Convolution can be done by finding the features in the image multiplying and adding the value of the respective feature pixel and finding the average. Example is shown in Fig 3

1	-1	-1
-1	1	-1
-1	-1	1

-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	-1	-1	1	-1	-1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1

1	1	1
1	1	1
1	1	1

(1+1+1+1+1+1+1+1+1)/9=1

Fig 3: Convolution

Activation function(ReLU): Activation function add nonlinearity in the network. In this study we will be using ReLU. The ReLU layers deletes every negative integer of the convolutional layer and replace it with 0 which makes CNN stable as in Fig 4.

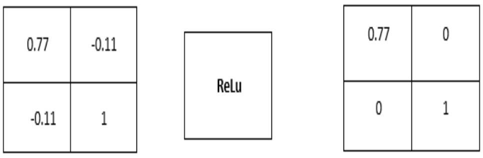


Fig4: ReLU

Pooling Layer: Pick a window and stride usually of size 2 and slide the window extracted features are given to pooling layer that shrink the larger images down at the mean time it preserves the most important information. In our study it takes only the maximum value from each window as it using Maxpool. It helps in reducing the dimensionality of input features as well as number of parameters. Example is shown in Fig 5.

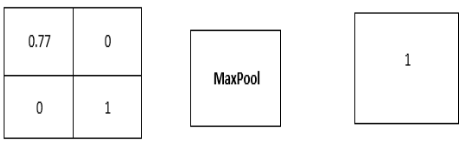


Fig 5: MaxPool

Fully Connected Layer: Each neuron in this layer is connected to all other neuron in the previous layer. It takes higher level filtered images and rename them with their corresponding labels.

The steps in the study is represented in Fig6

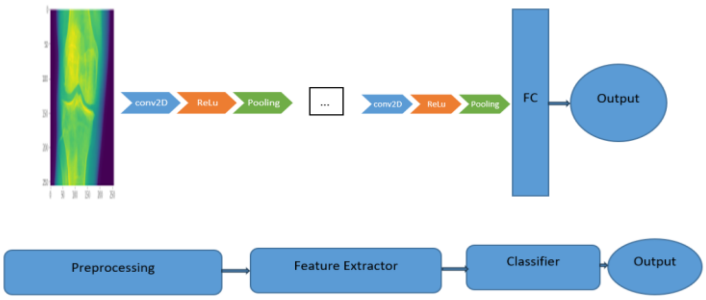


Fig 6 Processing steps in CNN

3.1 Dataset

The study employed the Mendeley database "https://data.mendeley.com", which has images of knees categorised into 5 classes based on the Kellgrance-Lawrence scale normal, Doubtful, mild, moderate, and severe.

3.2 Data Pre-processing

In this study data pre-processing is done by assigning labels to different categories as 0normal, 1Doubtful, 2mild, 3moderate, and 4severe as shown in Fig 7. In this study to attain a common size for all of the images, we pre-processed the 1650 data by scaling the image into 256*256 and converting RGB into grey scale. The data has been scaled back by 255. To fit the data into the 0 to 1 range the study has normalise the data. The data are being normalised to make them roughly the same scale.

Fig7: Label of images

3.3 Convolution Neural Network

CNN in conjunction with ReLU and MaxPooling will be used to classify various grades in Kellgrance-Lawrence.

Convolution layers are created using Conv2D. The number of features is reduced using the pooling function, and a dropout is applied using the dropout function. Each image is assigned to one of the five KL grades in the final, dense layer.

Activation function used in this study is Rectified linear unit (ReLU). The ReLU function has the benefit of not activating every neuron at once. Using the activation functions, we can preserve the most relevant features and suppress the irrelevant features. The neuron does not get activated for the negative input values and the result will be zero which makes it nonlinear. ReLU will avoid vanishing gradient

Loss function used in this study uses Categorical-Cross Entropy which will minimize loss. Because the derivative of the sigmoid function only has a big value in a very limited space of it, it will aid in preventing gradient vanishing. Loss function will provide a greater value if your forecasts are incorrect, just like the weighing machine. Your loss function will show if you are making any progress (or failing to make any progress) when you experiment with your method to try to improve your model.

The study's optimization approach is Adaptive moment estimation (Adam) [23] which adjusts network weights during training period. Adam extends stochastic gradient descent. This has high performance compared to other optimization algorithms. It combines the adaptive learning rate and momentum methods.
In this investigation, over fitting was avoided by applying dropouts to each hidden layer. Dropouts will be used for regularization. It will deactivate certain neurons and thus avoid overfitting
The sigmoid function is essentially generalised by the softmax function. It is typically used for multi-class classification tasks and applied to the network's top layer. Applying softmax in multi class classification will find the class with maximum probability.

The Fig 8 below shows the model summary. Sequential layer groups layer into a linear stack. The study has 3 convolutional layers followed by activation function ReLU and Maxpooling. and 3 dense layer. We have also applied dropout layer Sequential layer groups layer into a linear stack. The study has 3 convolutional layers followed by activation function ReLU and Maxpooling. and 3 dense layer. We have also applied dropout layer.

Out of the total params (3,788,6450) all are trainable.

Model: "sequential_2"	
Layer (Type)	Output Shape
conv2d_4 (Conv2D)	(None, 254, 254, 128)
activation_4 (Activation)	(None, 254, 254, 128)
max_pooling2d_4 (MaxPooling2)	(None, 127, 127, 128)
conv2d_5 (Conv2D)	(None, 125, 125, 64)
activation_5 (Activation)	(None, 125, 125, 64)
max_pooling2d_5 (MaxPooling2)	(None, 62, 62, 64)
conv2d_6 (Conv2D)	(None, 60, 60, 32)
activation_6 (Activation)	(None, 60, 60, 32)
max_pooling2d_6 (MaxPooling2)	(None, 30, 30, 32)
flatten_2 (Flatten)	(None, 28800)
dropout_3 (Dropout)	(None, 28800)
dense_4 (Dense)	(None, 128)
dropout_4 (Dropout)	(None, 128)
dense_5 (Dense)	(None, 64)
dense_6 (Dense)	(None, 5)

Fig 8 Model Summary

3.4 Splitting data for training, testing and fitting the model

The study has split the data into training 90% and test 10%. We have also done validation based on 20% of data. The current study has done validation split based on 20% of the data which means 20% of our training data will be used as validation samples and the rest for training. The current study has fit the model for 90 epochs since the current study is training with only few images. The functions evaluate the neural network's performance on the test set after training it using the training set. The accuracy of predictions made in the training set and the accuracy gained in the test set, respectively, are shown for each epoch by the variables "acc" and "val acc."

The study generated a training loss and accuracy graph as shown in Fig 9.

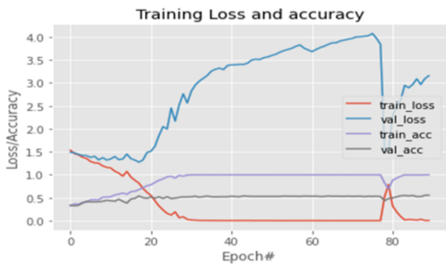


Fig 9: Training loss and accuracy graph

3.5 Prediction

The study will be predicting the output as follows in Fig 10

A.I predicts: 1Doubtful

Correct prediction for label 1 is 1Doubtful

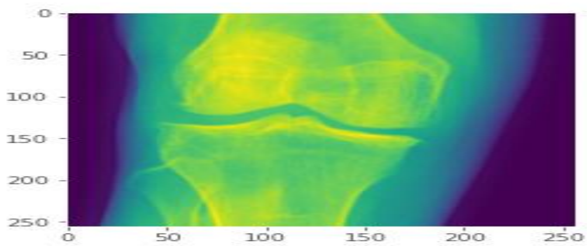


Fig 10 Predicted output

3.6 Confusion matrix

The investigation produced a confusion matrix, which is displayed in Fig 11.

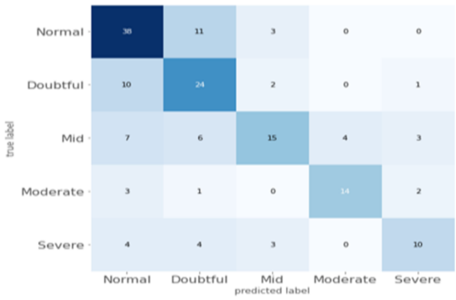


Fig 11 Confusion Matrix

Classification performance measurement

True positive (TP) - number of peoples with absence particular grade of KOA (grade) predicted as absence of corresponding KOA(grade).

False positive (FP) - number of peoples with presence of KOA (grade) predicted as absence of corresponding KOA (grade).

True negative (TN) - number of peoples with presence of KOA (grade) predicted as presence of corresponding KOA (grade).

False negative (FN) - number of peoples actually have absence of KOA (grade) predicted as presence of corresponding KOA (grade).

Fig 12 shows the equations of various performance measures.

Measure	Derivations
Sensitivity	TPR = TP / (TP + FN)
Specificity	SPC = TN / (FP + TN)
Precision	PPV = TP / (TP + FP)
Negative Predictive Value	NPV = TN / (TN + FN)
False Positive Rate	FPR = FP / (FP + TN)
False Discovery Rate	FDR = FP / (FP + TP)
False Negative Rate	FNR = FN / (FN + TP)
Accuracy	ACC=(TP+TN)/(TP+TN+FP+FN)
F1 Score	F1 = 2TP / (2TP + FP + FN)
Matthews Correlation Coefficient	TP*TN - FP*FN / sqrt((TP+FP)*(TP+FN)*(TN+FP)*(TN+FN))

Fig 12 Performance measurements

Results

We were able to categorise knee images using Kellgrance-Lawrence in this work. Images were classified to KL-0 (Normal), KL-1(Doubtful), KL-2(Mild), KL-3(Moderate) and KL-4(Severe) based on severity of knee osteoarthritis. Based on a confusion matrix, the study examined performance metrics. The study is showing 77% accurate for normal, 78% accurate for dubious, 83% accurate for mild, 93% accurate for moderate, and 89% accurate for severe. The study also discovered that, according to the confusion matrix, patients with moderate knee osteoarthritis had the highest accuracy (93%), recall (70%), specificity (97%), and F1 score (71%). In Table3, we have provided the whole results. If we train the model with more Knee images, we can further increase accuracy and other performance metrics

Table 2: Values based on the confusion matrix for several performance measures.

Measure	Normal	Doubtful	Mild	Moderate	Severe
Sensitivity	0.7308	0.6486	0.4286	0.7	0.4762
Specificity	0.7876	0.8281	0.9385	0.9724	0.9583
Precision	0.6129	0.5217	0.6522	0.7778	0.625
Negative Predictive Value	0.8641	0.8908	0.8592	0.9592	0.9262

False Positive Rate	0.2124	0.1719	0.0615	0.0276	0.0417
False Discovery Rate	0.3871	0.4783	0.3478	0.2222	0.375
False Negative Rate	0.2692	0.3514	0.5714	0.3	0.5238
Accuracy	0.7697	0.7879	0.8303	0.9394	0.897
F1 Score	0.6667	0.5783	0.5172	0.7368	0.5405
Matthews Correlation Coefficient	0.4972	0.4435	0.4332	0.7039	0.4894

Our research model accuracy average is 84.6%.

Conclusion

Comparing CNN to other machine learning, deep learning methods , this study offer great accuracy. Convolution and Maxpooling are the techniques used in this study to extract features. CNN has the advantage of sparse connectivity because every node in the convolution layer receives input from small number of nodes and it also use parameter sharing. Pooling reduces dimensionality.

To improve precision and avoid overfitting, this work makes use of ReLU and Adam optimization. ReLU which is the best activation function currently available is used in this study. It avoids overfitting and loss function. ReLU also speeds up training. For a better-optimized gradient, the Adam optimizer that employs stochastic gradient descent is used in this study. Adam optimizer outperforms other optimizers by a significant margin. Inorder to reduce loss this study uses the loss function-Categorical-Cross Entropy. This approach has the benefit of using dropouts for regularisation. This study has helped to classify Knee Ostearthritis to different classes by using Convolutional neural network (CNN) algorithm and also proposed that due to increase in the volume of health care data it is very good practice to train our model at large scale and fast processing by using cloud platform Google colab and GPU. The current study generates 10 performance measures, which show great specificity, accuracy, F1 Score, Negative Predictive Value, Precision, False Positive Rate, False Discovery rate, False Negative Rate and Matthews Correlation Coefficient given in Table 2. When compared to other CNN-based approaches, this study's KOA severity classification of knee osteoarthritis performs remarkably well.

References

[1] Du, Y., Almajalid, R., Shan, J., & Zhang, M. (2018). A novel method to predict knee osteoarthritis progression on MRI using machine learning methods. IEEE transactions on nanobioscience, 17(3), 228-236.

[2] Zhang, Y., & Jordan, J. M. (2010). Epidemiology of osteoarthritis. Clinics in geriatric medicine, 26(3), 355-369.

[3] Safiri, S., Kolahi, A. A., Hoy, D., Smith, E., Bettampadi, D., Mansournia, M. A., ... & Cross, M. (2019). Global, regional and national burden of rheumatoid arthritis 1990–2017: a systematic analysis of the Global Burden of Disease study 2017. Annals of the rheumatic diseases, 78(11), 1463-1471.

[4] Kellgren, J. H., & Lawrence, J. (1957). Radiological assessment of osteo-arthrosis. Annals of the rheumatic diseases, 16(4), 494.

[5] Alexos, A., Kokkotis, C., Moustakidis, S., Papageorgiou, E., & Tsaopoulos, D. (2020, July). Prediction of pain in knee osteoarthritis patients using machine learning: Data from Osteoarthritis Initiative. In 2020 11th International Conference on Information, Intelligence, Systems and Applications (IISA (pp. 1-7). IEEE.

[6] Podsiadlo, P., Wolski, M., & Stachowiak, G. W. (2008). Automated selection of trabecular bone regions in knee radiographs. Medical physics, 35(5), 1870-1883.

[7] Ntakolia, C., Kokkotis, C., Moustakidis, S., & Tsaopoulos, D. (2021). Prediction of joint space narrowing progression in knee osteoarthritis patients. Diagnostics, 11(2), 285.

[8] Kwon, S. B., Han, H. S., Lee, M. C., Kim, H. C., & Ku, Y. (2020). Machine learning-based automatic classification of knee osteoarthritis severity using gait data and radiographic images. IEEE Access, 8, 120597-120603.

[9] Leporace, G., Gonzalez, F., Metsavaht, L., Motta, M., Carpes, F. P., Chahla, J., & Luzo, M. (2021). Are there different gait profiles in patients with advanced knee osteoarthritis? A machine learning approach. Clinical Biomechanics, 88, 105447.

[10] Kwon, S. B., Ro, D. H., Song, M. K., Han, H. S., Lee, M. C., & Kim, H. C. (2019). Identifying key gait features associated with the radiological grade of knee osteoarthritis. Osteoarthritis and Cartilage, 27(12), 1755-1760.

[11] Indolia, S., Goswami, A. K., Mishra, S. P., & Asopa, P. (2018). Conceptual understanding of convolutional neural network-a deep learning approach. Procedia computer science, 132, 679-688.

[12] Swiecicki, A., Li, N., O'Donnell, J., Said, N., Yang, J., Mather, R. C., ... & Mazurowski, M. A. (2021). Deep learning-based algorithm for assessment of knee osteoarthritis severity in radiographs matches performance of radiologists. Computers in biology and medicine, 133, 104334.

[13] Wang, Y., Bi, Z., Xie, Y., Wu, T., Zeng, X., Chen, S., & Zhou, D. (2021). Learning From Highly Confident Samples for Automatic Knee Osteoarthritis Severity Assessment: Data From the Osteoarthritis Initiative. IEEE Journal of Biomedical and Health Informatics, 26(3), 1239-1250.

[14] Antony, J., McGuinness, K., O'Connor, N. E., & Moran, K. (2016, December). Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks. In 2016 23rd International Conference on Pattern Recognition (ICPR) (pp. 1195-1200). IEEE.

[15] Nebauer, C. (1998). Evaluation of convolutional neural networks for visual recognition. IEEE transactions on neural networks, 9(4), 685-696.

[16] Wahyuningrum, R. T., Anifah, L., Purnama, I. K. E., & Purnomo, M. H. (2019, October). A new approach to classify knee osteoarthritis severity from radiographic images based on CNN-LSTM method. In 2019 IEEE 10th International Conference on Awareness Science and Technology (iCAST) (pp. 1-6). IEEE.

[17] Lucas A, Iliadis M, Molina R, Katsaggelos AK (2018) Using deep neural networks for inverse problems in imaging: beyond analytical methods. IEEESignal Process Mag 35(1):20–36

[18] Khan A, Sohail A, Zahoora U et al (2020) A survey of the recent architectures of deep convolutional neural networks. Artif Intell Rev 53:5455–5516

[19] Paolini, R., Rodriguez, A., Srinivasa, S. S., & Mason, M. T. (2014). A data-driven statistical framework for post-grasp manipulation. The International Journal of Robotics Research, 33(4), 600-615.

[20] Zhu, P., Isaacs, J., Fu, B., & Ferrari, S. (2017, December). Deep learning feature extraction for target recognition and classification in underwater sonar images. In 2017 IEEE 56th annual conference on decision and control (CDC) (pp. 2724-2731). IEEE.

[21] Wu, H., Liu, Y., Liu, Y., & Liu, S. (2019). Efficient facial expression recognition via convolution neural network and infrared imaging technology. Infrared Physics & Technology, 102, 103031.

[22] Ramos, S., Gehrig, S., Pinggera, P., Franke, U., & Rother, C. (2017, June). Detecting unexpected obstacles for self-driving cars: Fusing deep learning and geometric modeling. In 2017 IEEE Intelligent Vehicles Symposium (IV) (pp. 1025-1032). IEEE.

[23] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

[24] Desai, M., & Shah, M. (2021). An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN). Clinical eHealth, 4, 1-11.