

## Feature Extraction Of Ultrasound Prostate Image Using Modified Vgg-19 Transfer Learning

<sup>1</sup>Mr.J.Ramesh, <sup>2</sup>Dr. R.Manavalan, <sup>3</sup>Dr.K.M. Prabusankarlal

<sup>1</sup>Assistant Professor, Department of Computer Applications, K.S.Rangasamy College of Arts and Science (Autonomous), Tiruchengode,

<sup>2</sup>Assistant Professor & Head, Department of Computer Applications, Arignar Anna Government Arts College, Villupuram,

<sup>3</sup>Assistant Professor & Head,, Department of Electronics & Communication, K.S.Rangasamy College of Arts and Science (Autonomous), Tiruchengode,

**Article History:** Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

### ABSTRACT

Feature extraction plays a vital role in classification, clustering, diagnosis, and recognition. For the classification process, feature extraction is very important to embody the content of images accurately. The extraction of features from Transrectal Ultrasound (TRUS) is complex as it contains speckle noise, low dissimilarity, the fuzzy region between the object and background. To solve this problem, Ultrasound (US) prostate images are preprocessed and segmented by the Ant Colony Optimization-Boundary Complete Recurrent Neural Network (ACO-BCRNN) method. To extract the relevant features many techniques are applied by the researchers. Recently, transfer learning methods are used for feature extraction. Transfer learning is a machine learning technique in which a trained model in one problem is used in the development of another related problem. Transfer learning consists of various pre-trained models such as VGG16 (Visual Geometry Group), VGG19, Resnet50, InceptionV3, Xception, MobileNetV2, DenseNet, and ResNetV2. The modified VGG-19 model is proposed to extract the features of the prostate ultrasound image. The proposed method was compared with other pre-trained models and its performance is evaluated by using Support Vector Machine (SVM), Extreme Learning Machine (ELM), and Grid Search. The experimental results demonstrated that the pre-trained model VGG19 is superior to other models in terms of Precision, Recall, Accuracy, and F1 Score.

**Keywords:** Prostate US image, Pre-trained Model, Modified VGG-19 model, SVM, Grid Search, ELM

### INTRODUCTION

Prostate Cancer is the second leading common disease of men in the United States. Prostate cancer is more likely to develop in older men with an average age at diagnosis is about 66. The American cancer society estimated about 1,91,930 new cases of prostate cancer in 2020 with a fatality rate of 17.3% [1]. Prostate cancer is produced by the continuous deposition of protein in the prostate gland. Ultrasound (US) is the most widely used technique for prostate biopsy. Usually, the symptoms of prostate cancer can be identified only by the extension of the prostate gland. Prostate cancer is treatable if it is diagnosed early. It becomes life-threatening if the tumor has extended beyond the prostate gland [2]. The computer-based automated image analysis method involves various steps such as image enhancement, segmentation, feature extraction, selection, and classification [3]. To distinguish between cancer and normal image, it is important to explore the feature, or set of features to obtain accurate quantifications of the characteristics of tumors[4]. Since, each image has its unique features such as shape, color, texture, edges of the image, etc.[5], the extraction and selection of effective features play a vital role in classification [6]. The feature extraction process acquires essential as well as relevant information to achieve the highest accuracy.

The traditional quantitative techniques such as GLCM, GLRLM, GLDM, Histogram, Shape, etc, are applied to extract features from prostate US images [7, 8]. In the past decade, Artificial Neural Network (ANN), used to extract features has shown good classification performance with Machine Learning-based classifiers [9, 10, 11,12]. The Artificial Neural Network (ANN), consists of several inputs, multiple hidden and output layers. The primary drawback associated with ANN is that for every interconnected layer, the number of weights is increased which in turn affects the learning rate. For processing images, the arrangement of several filters on the input layer which activates high-level features is called Convolutional Neural Network (CNN)[13]. Currently, it is considered an effective approach for image recognition and classification[14]. The CNN, used for large-scale image classification achieved superior performance with image databases such as ImageNet [15, 16]. Girshick et al. [17] applied supervised pre-training for large data set with tuning in a specific domain. Donahue et al. applied a pre-trained CNN, to extract the feature and used it for various image classification techniques [18]. Based on the transfer learning, pre-trained CNNs are utilized by using previously learned knowledge for acquiring high accuracy [19]. In our approach, eight existing pre-trained CNN models are used on ImageNet to extract the features from prostate tumors [20].

By applying the Transfer Learning method, the system can utilize the knowledge learned from a previous errand to a new errand [21]. ImageNet is the largest image database in which the images are arranged

according to the WordNet hierarchy[22]. Currently, the database consists of around 14 million images of 1000 object categories. Several CNN architectures that are already trained on the subset of the ImageNet database are utilized on the small dataset available to train a CNN that can have millions of weights to learn. While using CNN as a feature extractor, the last fully connected layer; the output layer has been removed. The extracted feature values contain raw values which are then transformed by a ReLU (Rectified Linear Unit). Both preReLU features (deep features) and postReLU features are obtained from the last hidden layer. The extracted deep features are useful for training and classification.

The CNN is a kind of multilayer feed-forward neural network that contains several layers that can be classified into 3 categories. The convolutional layer is utilized to compute the output for the connected input, the max-pooling layer is used for subsampling the inputs and fully connected layers are used to compute class activation [ 23, 24, 25]. In a CNN, an input image can be applied in the format  $h \times w \times c$ , where  $h$  is the height,  $w$  is the width, and  $c$  is the number of channels (3 for RGB and 2 for binary). There could be  $k$  convolutional filters of size  $n \times n$  in the convolutional layer, where  $n < h$ . In this article, the pre-trained CNN architectures used to extract the feature from the prostate image are VGG16, VGG19, ResNet50, InceptionV3, Xception, MobileNet, DenseNet, and ResNetV2.

The rest of the paper is organized as follows: Section 2 depicts the various transfer learning models for feature extraction. The comparative experimental analysis using various evaluation metrics is presented in Section 3 and the conclusion is drawn in Section 4 along with future directions.

## 2. TRANSFER LEARNING

Generally, Transfer Learning is a process of utilizing a model trained in one problem to solve another related problem. In this method, a neural network is used for training and is applied to a similar problem. From the trained model, one or more layers are used in the new model. The benefits of transfer learning are to reduce the training time and also decrease the error rate. The weights of the reused layers are used as an initial point at the training process and applied to a new problem. Several high-performing models have been developed and used on Imagenet large-scale visual recognition challenge (ILSVRC) for image classification. Thus the challenge is referred to as ImageNet and it gives the source of the image and weights. It has resulted from several innovations in the development of architecture and training of convolutional neural networks. The pre-trained model can be used for classification, stand-alone or integrated feature extraction, and weight initialization. It has three popular models namely VGG(VGG16, VGG19), GoogLeNet (InceptionV3), and Residual Network (ResNet50). A brief description of the various models are hereunder

**VGG-16:** Karen Simonyan and Andrew Zisserman introduced VGG16 architecture in 2014 [26]. VGG is a successor of AlexNet. The 16 in VGG16 referred to the 16 layers with weights. It has been created with a 16-layer network consists of convolutional and fully connected layers. It has 13 convolutional layers, 2 fully connected layers, and 1 Softmax classifier layer. The default size of the image input of this model is 224 x 224.

**Resnet50:** Residual Network which consists of 50 layers is named ResNet50. The input size of the image is 224 x 224. Initially, convolution with the kernel size of 7x7 and 64 different kernels with a stride size of 2, gives 1 layer as output. It is max-pooled with a stride size of 2. In the next convolution, 1x1 64 kernels are followed by 3x3 64 kernels.

**InceptionV3:** An Inception V3 model has been introduced by Szegedy et al. in 2014 [27] with new factorization ideas. The concept of factorization is to reduce the number of parameters without decreasing the network efficiency.

**Xception:** Xception model was proposed by Francois Chollet in 2017 [28]. The term Xception stands for Extreme Inception. This model has 36 convolutional layers, which are structured into 14 modules, all of which have linear residual connections except for the first and last modules.

**MobileNetV2:** Andrew G. Howard et. al [29] introduced the MobileNet model in 2017. This model uses depthwise separable convolutions and a single filter is applied to each input channel in MobileNetV2. To combine the outputs of the depthwise convolution, a 1x1 convolution is applied. In standard convolutions filtering and combining processes are the inputs to the next layer. However, in depthwise separable convolution, a separate layer is used for filtering and combining. This factorization reduces computational time as well as the model size.

**DenseNet:** DenseNet is a logical extension of ResNet, introduced by Gao Huang et al. in 2016 [30]. In this model, each layer connects to every other layer in a feed-forward fashion. Moreover, the layers are very

narrow with a small set of feature maps to the collective knowledge of the network while the remaining feature maps remain unchanged. Based on all feature maps, the final classifier makes a decision.

**ResNet152:** Residual Network consists of 152 layers is named ResNet152. The input size of the image is 224 x 224. In this model, batch normalization and the ReLU activation function are applied before the multiplication of convolution operation (weight matrix). Then, the value is given as input to the next block for further processing. In this work, the modified VGG-19 pre-trained models are implemented to extract the features from the prostate image.

### 2.1 The Basic Architecture of VGG19 Model

The VGG19 is a successor of the VGG16 model which contains 19 layers. The input size of the image and architecture is the same as that of the VGG16 model with few exceptions. The VGG19 has four convolution layers from the third block to the fifth block. The output size and activation functions are the same. The architecture model of VGG19 is shown in Figure-1. The increased number of convolution layers leads to greater accuracy and efficient performance.

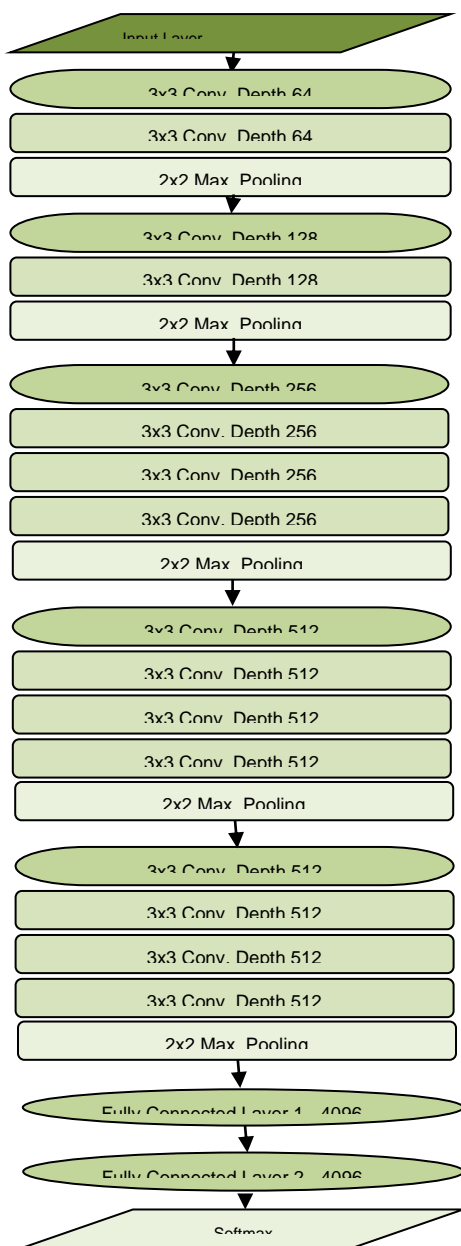


Figure 1: The architecture of VGG-19 with 16 Convolutional Layers using 3x3 Kernel Filters, 2 fully connected layers, and 1 Softmax Classifier Layer

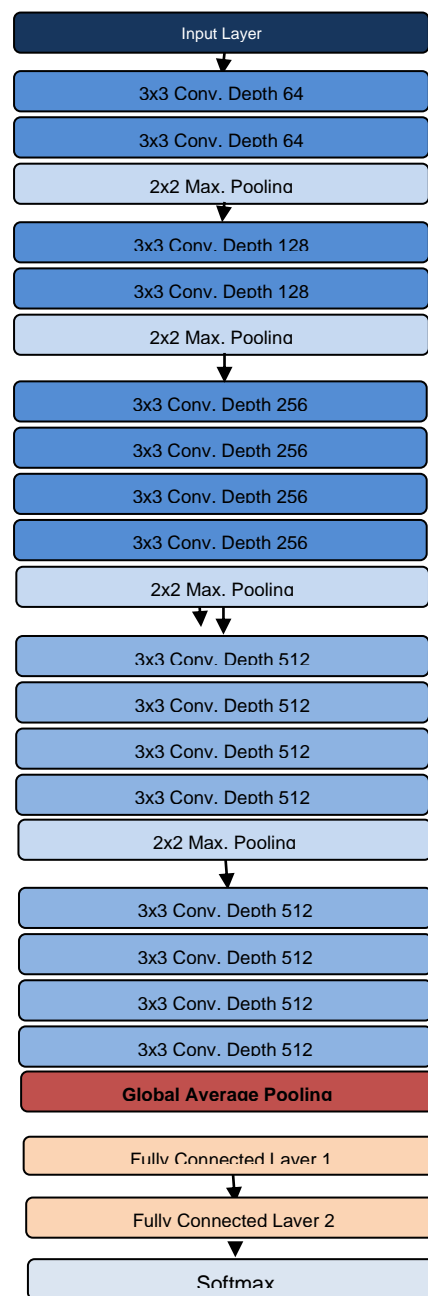


Figure 1: The modified architecture of VGG-19 model, only change in global average pooling in the final layer

## 2.2 The Proposed Model

The modified architecture of the VGG-19 model is shown in Fig.2. The proposed model is similar to the original VGG-19, modification in the final block instead of max pooling, global average pooling is used. Modified VGG-19 model consists of 16 convolutional layers, 4 max-pooling layers, 1 global average pooling layer, 2 fully connected layers, and an output layer. The model uses 512, 128, and 64 neurons on its hidden dense layers. Each dense layer is followed by a dropout. Dropout is used to reduce the capability of the model while training and guides against over fitting. ReLu activation functions are used in dense layers and the sigmoid activation function is used for classification. The training procedure uses learning rate decay by multiplying the previous learning rate with 0.5 for every four epochs having constant validation accuracy. Finally, the feature vector is collected from the fully connected layer 1.

## 3. EXPERIMENTAL RESULTS

A database consists of 4500 segmented prostate ultrasound US images are used to access the efficiency of the proposed methods and 70% of the dataset is used for training, the remaining 30% for testing. The various transfer learning models such as VGG16, modified VGG19, ResNet, Inception, Xception, MobileNet, DenseNet, and ResNetV2 are implemented in Python 3.7 for extraction of features from the prostate image. The performance of each model is evaluated by using classifiers such as Support Vector Machine (SVM), Grid Search, Extreme Learning Machine (ELM), and F1 Score.

### 3.1 Evaluation parameters

The Precision, Recall, Accuracy, and F1 Score are used as metrics for accessing the performance of the proposed methods. Precision is a measure of quality, it is the fraction of relevant instances among the retrieved instances. The recall is the fraction of the total amount of relevant instances, which are retrieved. The formula for the statistical parameters is given in Table 1.

**Table 1: The statistical parameters**

| Metrics   | Formula   | Descriptions  |
|-----------|---|---|
| Precision | $\frac{TP}{TP + FP}$  | TP-True Positive, TN-True Negative, FP-False Positive, and FN-False Negative. F1-Score is the harmonic mean of precision and recall |
| Recall    | $\frac{TP}{TP + FN}$  |   |
| F1 Score  | $2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$ |   |
| Accuracy  | $\frac{TP + TN}{TP + TN + FP + FN}$                         |   |

The overall performance measures of the feature extraction using pre-trained models are depicted in Table 2.

**Table 2: Overall Performance of Feature Extraction using the proposed pre-trained models(Acc.-Accuracy, Prec.-Precision, Rec.-Recall, and F1-F1 score)**

| Models          | SVM  |      |      |      | GridSearch |      |      |      | ELM    |      |      |      |
|-----------------|------|------|------|------|------------|------|------|------|--------|------|------|------|
|                 | Acc  | Prec | Rec  | F1   | Acc        | Prec | Rec  | F1   | Acc    | Prec | Rec  | F1   |
| <b>VGG16</b>    | 0.98 | 0.98 | 0.98 | 0.98 | 1.00       | 1.00 | 1.00 | 1.00 | 0.9976 | 0.98 | 0.96 | 0.98 |
| <b>VGG19</b>    | 0.99 | 0.99 | 0.99 | 0.99 | 1.00       | 1.00 | 1.00 | 1.00 | 0.9976 | 0.99 | 0.98 | 0.98 |
| Resnet50        | 0.98 | 0.98 | 0.98 | 0.98 | 0.99       | 0.99 | 0.99 | 0.99 | 0.9856 | 0.97 | 0.92 | 0.96 |
| Inception_V3    | 0.95 | 0.95 | 0.95 | 0.95 | 0.99       | 0.98 | 0.99 | 0.98 | 0.9856 | 0.98 | 0.91 | 0.94 |
| <b>Xception</b> | 0.58 | 0.83 | 0.58 | 0.57 | 0.92       | 0.92 | 0.92 | 0.92 | 0.6004 | 0.59 | 0.62 | 0.58 |

|           |      |      |      |      |      |      |      |      |        |      |      |      |
|-----------|------|------|------|------|------|------|------|------|--------|------|------|------|
| MobileNet | 0.95 | 0.95 | 0.95 | 0.95 | 0.99 | 0.98 | 0.99 | 0.98 | 0.9856 | 0.97 | 0.98 | 0.92 |
| DenseNet  | 0.67 | 0.80 | 0.67 | 0.66 | 0.94 | 0.94 | 0.94 | 0.94 | 0.4330 | 0.42 | 0.65 | 0.61 |
| ResNetV2  | 0.21 | 0.24 | 0.20 | 0.07 | 0.21 | 0.24 | 0.20 | 0.07 | 0.2057 | 0.23 | 0.32 | 0.05 |

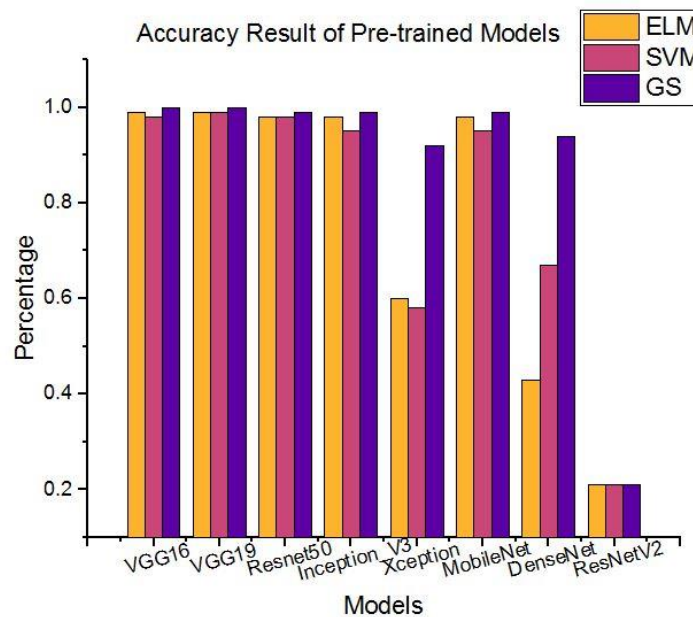
From the table, it is observed that VGG16, VGG19, and Xception models acquired better performance than other models. While considering the ELM accuracy, VGG16 and VGG19 models provide a better result of 99.76%.

Table 3 shows the accuracy obtained by the pre-trained models. The graphical representation is depicted in Fig. 3.

**Table 3: The performance of Accuracy all pre-trained models**

| Models          | Accuracy    |             |             |
|-----------------|-------------|-------------|-------------|
|                 | ELM         | SVM         | GS          |
| <b>VGG16</b>    | <b>0.99</b> | <b>0.98</b> | <b>1.00</b> |
| <b>VGG19</b>    | <b>0.99</b> | <b>0.99</b> | <b>1.00</b> |
| Resnet50        | 0.98        | 0.98        | 0.99        |
| Inception_V3    | 0.98        | 0.95        | 0.99        |
| <b>Xception</b> | 0.60        | 0.58        | 0.92        |
| MobileNet       | 0.98        | 0.95        | 0.99        |
| DenseNet        | 0.43        | 0.67        | 0.94        |
| ResNetV2        | 0.21        | 0.21        | 0.21        |

The accuracy produced by the pre-trained models is shown in Fig. 3. VGG16 and VGG19 models obtained an accuracy of 99% using ELM, SVM, and Grid Search methods. Resnet50 and InceptionV3 models obtained an accuracy of 98%, which is almost 1% lesser than the VGG16 and VGG19 models. The accuracy achieved by the DenseNet model using ELM, SVM, and Grid Search is 43%, 67%, and 94% respectively. ResNetV2 model obtained an accuracy of just 20% using all the methods which makes it not suitable for feature extraction.

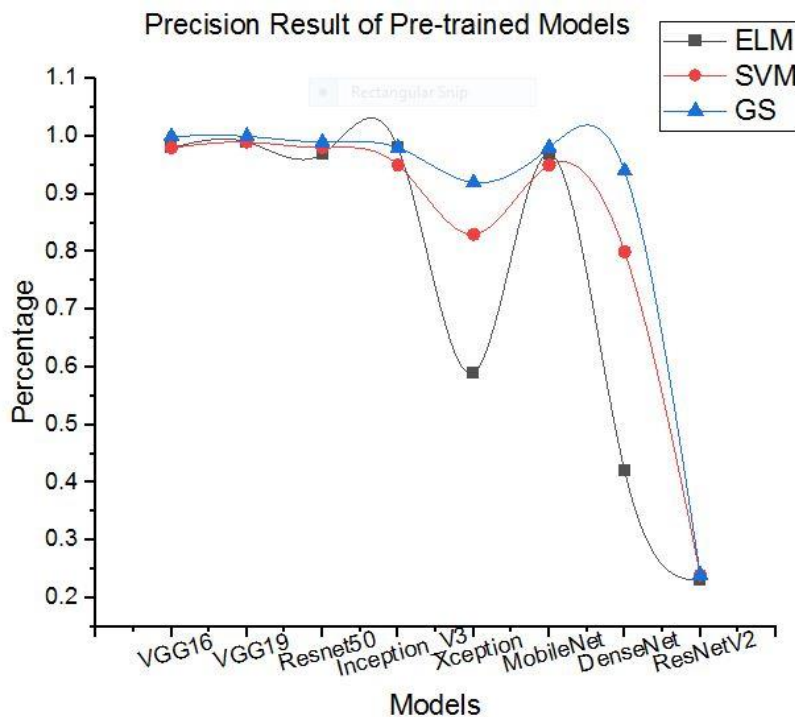


**Fig. 3: Accuracy Results**

The quantitative performance of the Precision and pictorial representation is shown in Table 4 and Fig. 4 respectively.

**Table 4: Precision Results**

| Models          | Precision   |             |             |
|-----------------|-------------|-------------|-------------|
|                 | SVM         | GridSearch  | ELM         |
| <b>VGG16</b>    | <b>0.98</b> | <b>1.00</b> | <b>0.98</b> |
| <b>VGG19</b>    | <b>0.99</b> | <b>1.00</b> | <b>0.99</b> |
| Resnet50        | 0.98        | 0.99        | 0.97        |
| Inception_V3    | 0.95        | 0.98        | 0.98        |
| <b>Xception</b> | 0.83        | 0.92        | 0.59        |
| MobileNet       | 0.95        | 0.98        | 0.97        |
| Densenet        | 0.80        | 0.94        | 0.42        |
| ResNetV2        | 0.24        | 0.24        | 0.23        |



**Fig. 4: Precision Results**

From Table 4, it is observed that the precision obtained by the modified VGG19 model using SVM, Grid Search, and ELM are 99%, 100%, and 99% respectively which are comparatively 1% higher than the VGG16 and ResNet50 models. The Xception model achieved 83%, 92%, and 59% using SVM, Grid Search, and ELM respectively. However, the ResNetV2 model acquired a very low precision of 24%.

The recall results of the proposed methods are shown in Table 5 and depicted in Fig. 5. From the table, it is noticed that the recall values of the modified VGG19 model obtained through SVM, Grid search, and ELM are 99%, 100%, and 98% respectively. The VGG16 model achieved a 100% recall rate through Grid search and 98% through SVM which is 1% lesser than that of VGG19. The Resnet50, Inception\_V3, and MobileNet model achieved 99% recall using Grid search and the recall rate using SVM is poor. ResNetV2 model does not seem to be appropriate since it achieved only a 20% recall value.

**Table 5: Recall Result**

| Models       | Recall      |             |             |
|--------------|-------------|-------------|-------------|
|              | SVM         | GridSearch  | ELM         |
| <b>VGG16</b> | <b>0.98</b> | <b>1.00</b> | <b>0.96</b> |
| <b>VGG19</b> | <b>0.99</b> | <b>1.00</b> | <b>0.98</b> |

|                 |      |      |             |
|-----------------|------|------|-------------|
| Resnet50        | 0.98 | 0.99 | 0.92        |
| Inception_V3    | 0.95 | 0.99 | 0.91        |
| <b>Xception</b> | 0.58 | 0.92 | 0.62        |
| MobileNet       | 0.95 | 0.99 | <b>0.98</b> |
| Densenet        | 0.67 | 0.94 | 0.65        |
| ResNetV2        | 0.20 | 0.20 | 0.32        |

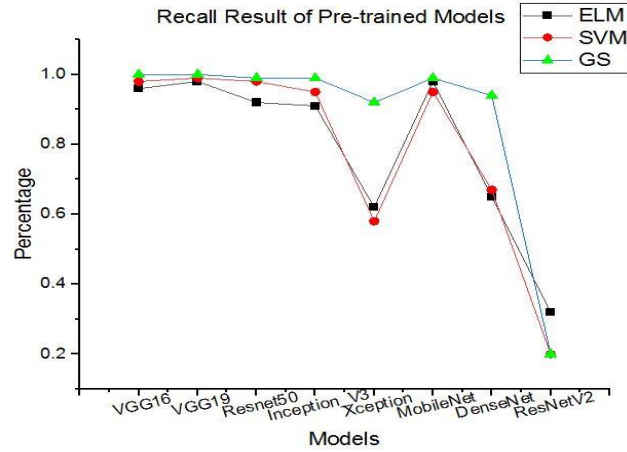
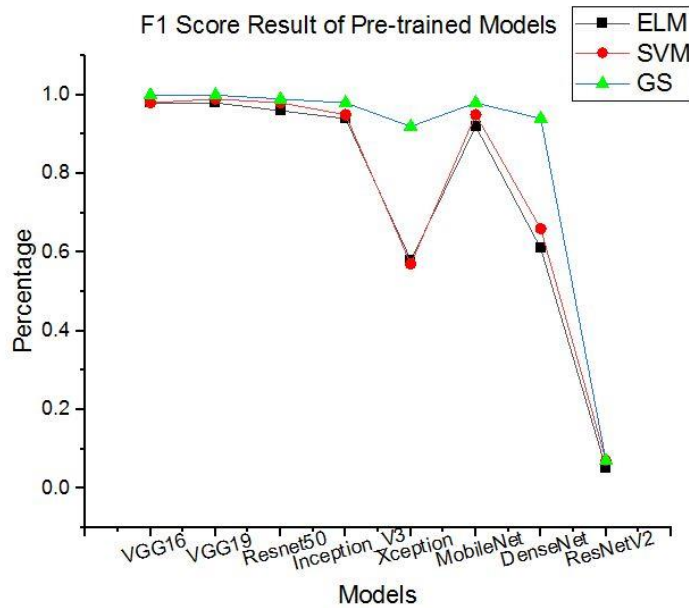


Fig. 5: Recall Result

Table 6: F1-Score Result

| Models       | F1-Score    |             |             |
|--------------|-------------|-------------|-------------|
|              | SVM         | GridSearch  | ELM         |
| <b>VGG16</b> | <b>0.98</b> | <b>1.00</b> | <b>0.98</b> |
| <b>VGG19</b> | <b>0.99</b> | <b>1.00</b> | <b>0.98</b> |
| Resnet50     | 0.98        | 0.99        | 0.96        |
| Inception_V3 | 0.95        | 0.98        | 0.94        |
| Xception     | 0.57        | 0.92        | 0.58        |
| MobileNet    | 0.95        | 0.98        | 0.92        |
| Densenet     | 0.66        | 0.94        | 0.61        |
| RenseNetV2   | 0.07        | 0.07        | 0.05        |



**Fig. 6: F1-Score Result**

The F1-score performance of the proposed methods is given in Table 6. The comparative performances are also depicted in the graph in Fig. 6. From the results given in Table 6, it is evident that the modified VGG19 model achieved superior performance to other models. The model acquired an F1 score of 100%, 99%, and 98% using Gridsearch, SVM, and ELM respectively. The VGG16 and Resnet50 models achieved 100%, 98%, and 98% using Gridsearch, SVM, and ELM respectively. The Inception\_V3 and MobileNet model obtained an F1 score of almost 4% and 2% less than that of VGG19 using SVM and Gridsearch respectively. The ResnetV2 model achieved highly inferior poor performance compared to other models.

The comparative performances of the proposed models in terms of Accuracy, Precision, Recall, and F1-Score indicate that the modified VGG19 is the most appropriate model for extraction of features from prostate ultrasound images.

#### 4. CONCLUSION

Feature extraction methods play a significant role to extract the features from the segmented image. However, choosing a suitable algorithm is not a minor task. Many researchers applied various methods to extract the features, the appropriate method is not identified for the prostate dataset. And, to improve the efficiency of the result, the pre-trained models of transfer learning are used for feature extraction of the prostate image in the proposed method. Initially, the TRUS prostate image has been enhanced with the ACO and segmented by ACO-BCRNN methods. The informative features from normal, benign, and malignant TRUS prostate images are extracted by various pre-trained models such as VGG16, modified VGG19, ResNet50, Inception\_V3, Xception, MobileNet, Densenet, and RenseNetV2. Finally, The proposed method is evaluated using classifiers like SVM, Grid Search, and ELM. The experimental results reveal that the feature extraction based on the modified VGG-19 method attained better accuracy compared to other methods. The modified VGG19 model attained superior performance over all other models with an accuracy of 99% and an F1-Score of 100% and it is effective and appropriate for the extraction of features in a computer-aided diagnosis(CAD) system for clinics. In future works, the proposed algorithm is applied for signal processing, galaxy image classification, and video processing.

#### REFERENCES

- [1] Cancer Facts and Figures. American Cancer Society. [Online] <http://www.cancer.org>, (2020).
- [2] P. B. Dhir, R. Landsitte., M. R. Hitchens and R.H. Getzenberg, "Detection of Prostate Cancer with a Blood-Based Assay for Early Prostate Cancer Antigen", *Cancer Research*, vol. 65, May 2005.
- [3] Revathi, R., & Hemalatha, M. An Emerging Trend of Feature Extraction Method In Video Processing. *Cs & It-Cscp*, 2012 69–80.
- [4] Pereira WCA, Alvarenga AV, Infantosi AFC, Macrini L, Pedreira CE (2010) A non-linear morphometric feature selection approach for breast tumor contour from ultrasonic images. *Comput Biol Med* 40(11):912–918
- [5] Dong ping Tian, "A Review on Image Feature Extraction and Representation Techniques", *International Journal of Multimedia and Ubiquitous Engineering* Vol. 8, No. 4, July 2013, pp 385-396.



- [6] Prabusankarlal KM, Thirumoorthy P, Manavalan R (2014) Computer-Aided Breast Cancer Diagnosis Techniques in Ultrasound: A Survey. *J Med Imaging Health Inform* 4(3):331–349
- [7] R.Manavalan, K.Thangavel (2012) Comparative Analysis of Feature Extraction Methods for the Classification of Prostate Cancer from TRUS Medical Images. *IJCSI Int J Comput Sci Issues* 9(1):1694–0814
- [8] Prabusankarlal, K. M., Thirumoorthy, P., & Manavalan, R. (2015). Assessment of combined textural and morphological features for diagnosis of breast masses in ultrasound. *Human-centric Computing and Information Sciences*, 5(1), 1-17.
- [9] H. Y. Priyanga, & D. Ruliandi, "Application of Pattern Recognition and Classification Using Artificial Neural Network in Geothermal Operation," *PROCEEDINGS, 43rd Workshop on Geothermal Reservoir Engineering Stanford University, Stanford, California, February 12-14, 2018*
- [10] C. Shang, A. Palmer, J. Sun, K.S. Chen, J. Lu, & J. Bi, "VIGAN: Missing view imputation with generative adversarial networks," In 2017 IEEE International Conference on Big Data (Big Data) (pp. 766-775). IEEE.
- [11] R. Bala, D. & Kumar, "Classification Using ANN: A Review," *International Journal of Computational Intelligence Research*, (2017). 13(7), 1811-1820
- [12] S. Knerr, L. Personnaz, & G. Dreyfus, "Handwritten digit recognition by neural networks with single-layer training," *IEEE Transactions on neural networks*, (1992), 3(6), 962-968.
- [13] Sakib, S.; Ahmed, N.; Kabir, A.J.; Ahmed, H. An Overview of Convolutional Neural Network: Its Architecture and Applications. *Preprints 2018*, 2018110546 (DOI: 10.20944/preprints201811.0546.v1).
- [14] Xin, M., Wang, Y. Research on image classification model based on deep convolution neural network. *J Image Video Proc.* 2019, 40 (2019). <https://doi.org/10.1186/s13640-019-0417-8>
- [15] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: *Neural Information Processing Systems 2012*. Lake Tahoe, NV: Advances in Neural Information Processing Systems; 2012.
- [16] J. Deng, W. Dong, R. Socher, L. Li, Kai Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, 2009, pp. 248-255, DOI: 10.1109/CVPR.2009.5206848.
- [17] Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: *Computer Vision and Pattern Recognition Conference 2014*. Columbus, OH; 2014
- [18] Donahue J, Jia Y, Vinyals O, Hoffman J, Zhang N, Tzeng E, Darrell T. DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition. In: *International Conference on Machine Learning*. Beijing, China. International Machine Learning Society; 2014
- [19] Raina R, Battle A, Lee H, Packer B, Ng AY. Self-taught learning: transfer learning from unlabeled data. 24th International Conference on Machine Learning. Corvallis, OR. The International Machine Learning Society; 2007.
- [20] Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. ImageNet: a large-scale hierarchical image database. In: *IEEE Conference on Computer Vision and Pattern Recognition*. Miami, FL. Institute of Electrical and Electronics Engineers (IEEE); 2009
- [21] Paul R, Hawkins SH, Hall LO, Goldgof DB, Gillies RJ. Combining deep neural network and traditional image features to improve survival prediction accuracy for lung cancer patients from diagnostic CT. *IEEE International Conference on Systems, Man, and Cybernetics*. October 2016
- [22] J. Deng, W. Dong, R. Socher, L. Li, Kai Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, 2009, pp. 248-255, DOI: 10.1109/CVPR.2009.5206848.
- [23] Fischer P, Dosovitskiy A, Brox T. Descriptor matching with convolutional neural networks: a comparison to SIFT. *arXiv preprint arXiv:1405.5769*. 2014.
- [24] Oquab M, Bottou L, Laptev I, Sivic J. Learning and transferring mid-level image representations using convolutional neural networks. In: *IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, OH; 2014
- [25] Zeiler MD, Fergus R. Visualizing and understanding convolutional networks. In: *13th European Conference on Computer Vision*. Zurich, Switzerland; 2014.
- [26] Karen Simonyan and Andrew Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [27] Christian Szegedy, Wei Liu, Yangqing Jia, et al., Going Deeper with Convolutions. *arXiv:1409.4842*, 2014.
- [28] Francois Chollet, Xception: Deep Learning with Depthwise Separable Convolutions. *arXiv:1610.02357v3*, 2017.
- [29] Andrew G. Howard, Menglong Zhu, Bo Chen, et al., MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *arXiv:1704.04861v1*, 2017.
- [30] Gao Huang, Zhuang Liu, and Kilian Q. Weinberger, Densely Connected Convolutional Networks. *arXiv: 1608.06993v3*, 2016.

- [31]Huang, Guang-Bin, Zhu, Qin-Yu, Siew, Chee-Kheong, "Extreme learning machine: theory and applications", *Neuro Computing*, 70(1): 489-501, 2006. [CiteSeerX 10.1.1.217.3692](#)
- [32]Huang G, Liang N, Rong HJ et al. On-line sequential extreme learning machine. In: *The IASTED international conference on, computational intelligence*, pp. 232–237, 2005