Multiple Scaling Based EfficientNet Modelling for Liver Tumor Classification on CT Images

Bilga Jacob*, R. S. Vinod Kumar, S. S. Kumar

Department of Electronics and Communication Engineering
Noorul Islam Center for Higher Education
Kumarakoil, Kanyakumari, Tamil Nadu, India
E-mails: bilgajacob@gmail.com, rsvinodkumar@niuniv.com, kumarss@live.com

**Corresponding author*

Received: April 20, 2024 **Accepted: August 15, 2024**

Published: September 30, 2024

Abstract: *For both women and men over 60, liver cancer is the primary cause of cancer-related deaths. To help physicians to diagnose patients more accurately, computer-assisted imaging techniques have become increasingly important in recent years. Recent, deep Convolutional Neural Network (CNN) research has produced amazing improvements in image segmentation and classification. The same issue of diagnosing liver nodules in computed tomography (CT) scans is addressed in this research by introducing a novel Computer-Aided Detection (CAD) system that makes use of an Efficient Network (EfficientNet) image classification algorithm. Unlike CNN, which adjusts its network parameters arbitrarily, a set of predetermined scaling coefficients is used in the EfficientNet scaling technique to reliably scale the network's breadth, depth, and resolution. Here the EfficientNet models are assessed by varying the input dimensions of the CT scans from The Liver Tumor Segmentation (LiTs) dataset. Finally, the performance evaluation shows that the input dimension 224×224 effectively classified the images and is superior to the other models evaluated with 0.991 AUC and 99.37% F1-Score, precision 99.44%, recall 99.30%, specificity 99.43%, and accuracy 99.36% for Kaggle datasets.*

Keywords: *Deep learning, Convolution Neural Network, Computed tomography, Liver tumor classification, The Liver tumor segmentation.*

Introduction

Among all cancers, liver cancer has one of the lowest five-year survival rates. Once exclusive to the elderly, modern lifestyle changes have made liver disorders, especially liver cancer, more common in younger people. One important strategy for reducing the fatality rate is early detection of liver nodules that may turn into cancer [8]. The increased detection rate of abdominal computerized tomography screening makes it a more effective means of reducing deaths from liver cancer. For more than 20 years, research has been conducted to create Computer-Aided Detection (CAD) systems that will aid radiologists in their search for malignant lesions in computed tomography (CT) scans. Convolution Neural Network (CNN) was made well-known as the answer to this problem by Yann [3]. The main problem with employing CNNs to detect liver tumors is the variety of purposes for nodule appearance and the volume of information found in CT scans. This technique has certain limitations, such as a high false-positive rate, variable sensitivity leading to missed diagnoses and undue anxiety in patients.

Deep learning based systems [5] have made machine vision tasks more advanced and adaptable to assist in medical diagnosis. These networks are preferable because of the valuable and strong semantic properties they extract from the input material. A more effective image identification technique called the EfficientNet [17] has piqued the interest of scientists all over the world.

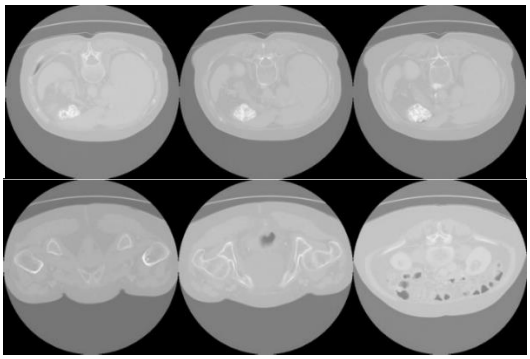
Prior to the introduction of EfficientNet, the domain of deep learning saw the creation of several important designs, especially in the area of CNNs, which transformed a number of applications, including object detection and envision categorization. This includes LeNet-5, the first CNN followed by Alex Net [2], VGG [4], Inception Net [16], Dense Net [12], Mobile Net [15], etc. Each of these designs pushed the envelope of what was possible and brought novel concepts and techniques that had a big impact on the deep learning community. Building on these developments, EfficientNet presented a methodical approach to scaling up CNNs in a more balanced fashion across many dimensions (depth, width, and resolution), which enhanced accuracy and efficiency even further. For a given level of accuracy, EfficientNets use less computing power than previous models like VGG, ResNet, or Inception. EfficientNets remains a competitive alternative to more recent architectures such as Vision Transformers (ViT) [9], particularly in situations where model efficiency is crucial.

The majority of studies on the categorization of liver cancer have focused on models trained and tested on the publicly available Kaggle Dataset from The Liver Tumor Segmentation Benchmark. In their study, Kumar et al. [11] focused on the statistical texture descriptor's capacity to automatically differentiate between benign and aggressive liver tumors, achieving 96.7% accuracy, 97.3% sensitivity, and 96% specificity, respectively. Additionally, a hybrid hash-based CNN model [13] that extracts features from images using a hash function was developed. The model's ability to categorize tumors into benign and malignant groups showed promise. Tumor classification was achieved using VGG-16 and a SegNet-based deep convolutional encode-decoder model [1]. The liver lesions were also classified into benign and malignant groups using a SegNet-based deep learning model. For the dataset, the Dice index, correlation coefficient, and Jaccard index were 0.96, 0.968, and 0.962, respectively.

Chang et al. [6] suggested using a completely automated CAD system to identify hepatocecullar cancer. Support Vector Machine (SVM) yielded 98.7% accuracy compared to 98.4% accuracy for Artificial Neural Network (ANN). To extract liver tumors from CT scans, a Computer-Aided Diagnosis approach was presented. Using Binary Logistic Regression Analysis based on Leave-One-Out Cross-Validation Strategy, the model yielded an accuracy of 81.2% [6]. A classification model for diagnosing liver cancers is suggested, utilizing pix2pix generation adaptive modules and Deep Lab V-3. The model performed better in classifying cancers [6]. Like this, a hybrid cascade segmentation network built on 2D and 3D neural network models was used to identify liver cancers using The Liver Tumor Segmentation (LiTs) CT image dataset. 130 CT images were subjected to histogram segmentation, and the model yielded encouraging results [19]. As was noted in the literature previously examined, an effective and trustworthy method for early detection of liver cancers is required. This will make it easier for doctors to promptly and correctly diagnose liver cancers.

Materials and methods

Advanced and versatile deep learning based tasks are now possible to aid in medical diagnosis. The LiTs dataset [21] was the source of the liver datasets, which are freely available to the public. The LiTs dataset was created to evaluate the effectiveness of automated segmentation algorithms and assist in the identification and demarcation of liver tumors (hepatocellular carcinoma) from CT scans. The collection consists of 131 3D CT scans from a variety of individuals. It shows a variety of liver disorders, sizes, and forms, including both healthy livers and those with tumors of different sizes and types. The NIfTI file format of the images is used, and they are sliced to separate the component DICOM before being preprocessed and saved as PNG files. The collection contains 14 124 pictures of both benign and malignant groups. Some exemplary cases from the database are shown in Fig. 1.



A)

B)

Fig. 1 Samples of CT DICOM images from LiTs dataset:
liver slices without tumor (A), liver slices with tumor (B)

Methodology of the proposed liver tumor classification

Neural architecture search, which employs a controller network like Recurrent Neural Network (RNN) and sample network topologies from a search space with probability "*p*" is used to construct the EfficientNet models [20]. Using grid search, the six additional models are generated, ranging from EfficientNet B1 to EfficientNet B7, by scaling up the depth, width, and picture quality of their original model, which they named EfficientNet B0 [17]. EfficientNet V2, which surpasses EfficientNet in terms of training speed and parameter efficiency, adds new convolutional blocks like fused-MBConv [18]. The enhanced generalization capabilities of EfficientNet V2 over EfficientNet V1 can be attributed to modifications in data augmentation methodologies, model design, and training protocols.

The proposed architecture for EfficientNet modelling for liver tumor classification is shown in Fig. 2. Here, liver tumor classification is implemented using the EfficientNet V2-S model, which only requires 3×3 filters and has a reduced expansion ratio than EfficientNet V1. The network begins with a stem made out of a conventional convolution and then moves through several phases that gradually boost the model's capacity and complexity. Multiple fused-MBConv or MBConv blocks make up each step. The fused-MBConv blocks are depicted in Fig. 3.

Tan and Le [18] combine depthwise and pointwise convolutions into a single convolution, in contrast to the original MBConv blocks. Depending on the stage and the EfficientNet V2 variant, there are differences in the number of blocks and their configurations, such as expansion ratios, kernel sizes and filter counts. One or more fully connected layers are incorporated into the process at the output stage to classify the incoming data. The network's final stages include the application of global average pooling, a dropout layer, and a fully linked layer that generates the final predictions.

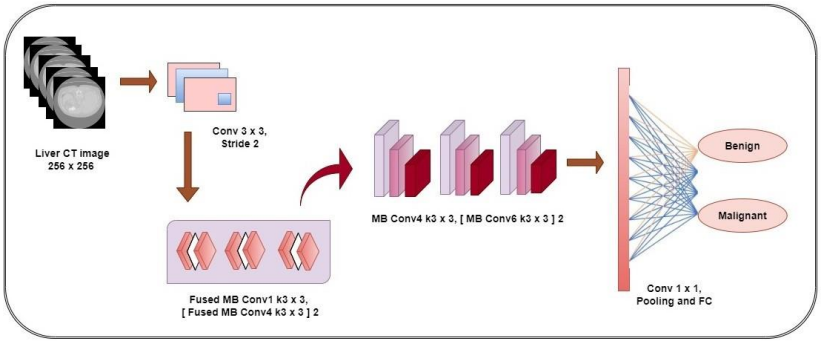


Fig. 2 An architecture of the proposed EfficientNet modelling for liver tumor classification

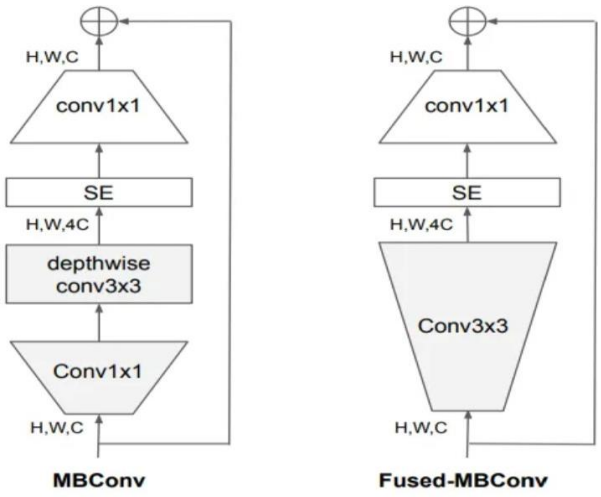


Fig. 3 MBConv and fused-MBConv layers of EfficientNet V2

Experimental setup

The models were implemented using Python 3 and the Keras library throughout. These were performed using a Google Colab pro+, which had a P100 GPU processor, 52 GB of RAM, and 2 TB of storage. Before being used, all the input pictures were resized and normalized using Kera's Image DataGenerator class. The preprocessed images were loaded into the suggested deep learning model for binary classification. Each model was trained and confirmed for around 50 epochs initially at 188 iterations per epoch using an Adam optimizer and the appropriate fit parameters. The batches had a size of 32. This research investigated how the architecture of deep learning was affected by different input image dimensions, namely 224×224, 229×229, and 256×256.

The pseudocode for the proposed liver tumor classification using EfficientNet is depicted in Algorithm 1 (Fig. 4).

Input: Input Datasets → D
Model → [EfficientNet]
Learning Rate = 0.001, Optimizer → ADAM, Epochs = 50
Output: Image Classification Determined
Select the Liver Datasets and Split the datasets into three sections → 60% Training, 20% each for Validation and Testing
Create Pretrained EfficientNet Architecture
Define Image Sizes $S = \{224, 229, 256\}$
for s in S **do**
    Input Initial Image Size $S0$
    for i in Models **do**
        Initialize training of EfficientNet architecture with image size $S0$
        while epoch ≤ Epochs **do**
            Execute backpropagation and optimization algorithms to minimize the loss function.
        end while
    end for
    Utilize test data to forecast classes.
    Display the anticipated class as the result.
end for

Fig. 4 Algorithm 1: Pseudocode for the proposed liver tumor classification using EfficientNet

Results and discussion

The model assessments make use of a variety of metrics. This study uses the confusion matrix and related metrics including accuracy, specificity, precision, recall and F1-score for evaluation. False Positive (FP), False Negative (FN), True Positive (TP), and True Negative (TN) are the four possible outcomes. A TP result indicates the presence of cancer in the patient since the result is classified as positive and the real value is positive. When a test indicates there is no malignancy when the patient actually has the condition being tested for is referred to a FN, and it suggests that the lab misread the result and it was positive. FP are results of cancer that are classified as positive but have a real value of negative when they do not actually exist. If the test yields a TN, it indicates that the subject is healthy or that the true value is negative. The following formulae, which range from Eq. (1) to Eq. (5), represent the performance measures.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \times 100; \quad (1)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \times 100; \quad (2)$$

$$\text{Precision} = \frac{TP}{TP+FP} \times 100; \quad (3)$$

$$\text{Recall} = \frac{TP}{TP+FN} \times 100; \quad (4)$$

$$\text{F1-Score} = \frac{2\times TP}{2\times TP+FP+FN} \times 100. \quad (5)$$

The confusion matrices of the models were used to analyze digital images to classify liver tumor as either normal or abnormal. The generated confusion matrices for the initial training of 50 epochs are displayed in Fig. 5. The results demonstrated that, with an accuracy of 99.07%, there was a predicted disagreement in 12 out of 2 825 images for the input image dimensionof 224×224. With an accuracy of 97.98%, the prediction disparity in the case of input image dimension 229×229 was in 27 digital images. The diagnosis accuracy was 85.62% with a predicted discrepancy of 89 images for an input image dimension of 256×256. The results show that input dimension and accuracy are negatively correlated, with input 224 exhibiting the maximum accuracy. On the other hand, a drop in performance has been noted for 256×256 input dimensions. The 99.07% accuracy rate of the EfficientNet model is an exceptional performance, and all other metrics are similarly good. Table 1 presents analysis's findings.

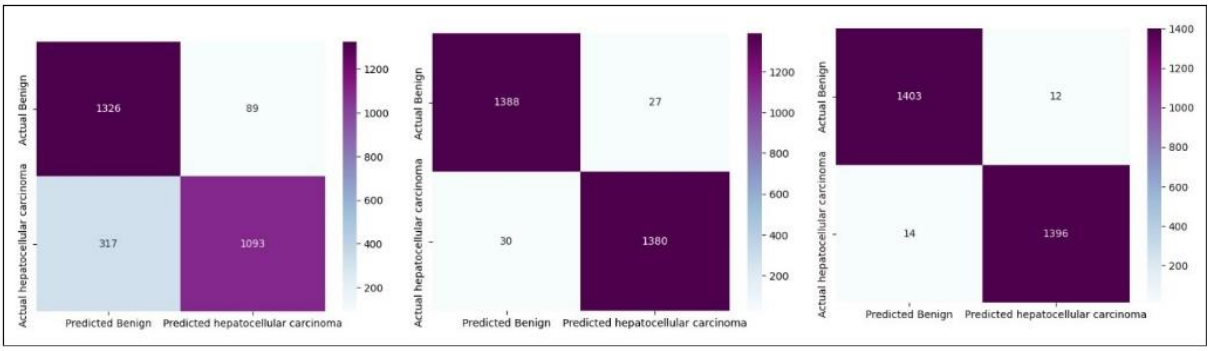


Fig. 5 Confusion matrices of the proposed work for various input image dimension for 50 epochs: 256×256 (left), 229×229 (middle), 224×224 (right).

Table 1. Performance analysis of the proposed EfficientNet architecture based on various input image dimensions for 50 epochs

Input image dimension	Accuracy, (%)	Specificity, (%)	Precision, (%)	Recall, (%)	F1-Score, (%)
256×256	85.62	92.47	93.71	80.70	86.72
229×229	97.98	98.08	98.09	97.88	97.98
224×224	**99.07**	**99.14**	**99.15**	**99.22**	**99.08**

The classification training and validation accuracy for the suggested model at each epoch is displayed in Fig. 6. The experiment was initially carried out for 50 epochs in total. The accuracy curves for input dimensions 256×256 and 229×229 show that the validation accuracy is slightly higher than the training accuracy which is a sign of underfitting. To better understand the implemented model in this situation, the training experiments are conducted for 80 epochs which required more computational resources. Even though performance improvement is visibly pronounced in the accuracy curves for 229×229, the performance dynamics of the input image dimension of 256×256 was not very effective. The classification training and validation accuracy for the suggested model at each epoch is displayed in Fig. 7 and the generated confusion matrices are displayed in Fig. 8.

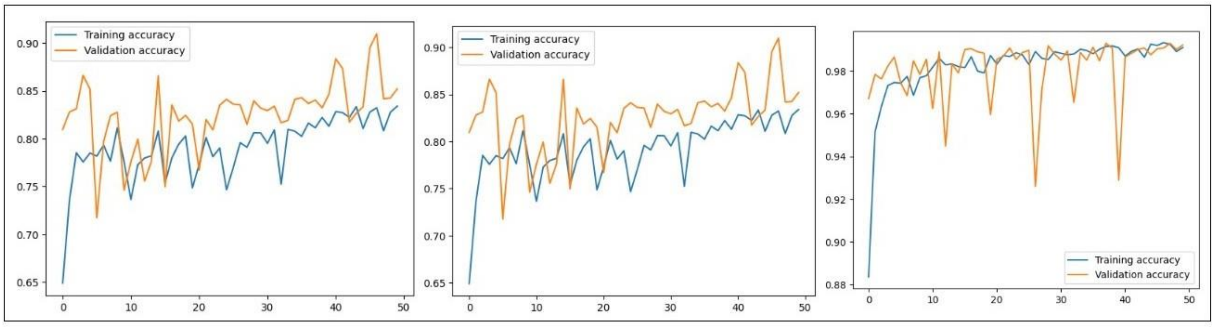


Fig. 6 Training and validation accuracy curves for liver tumor classification for 50 epochs based on input image dimension: 256×256 (left), 229×229 (middle), 224×224 (right).

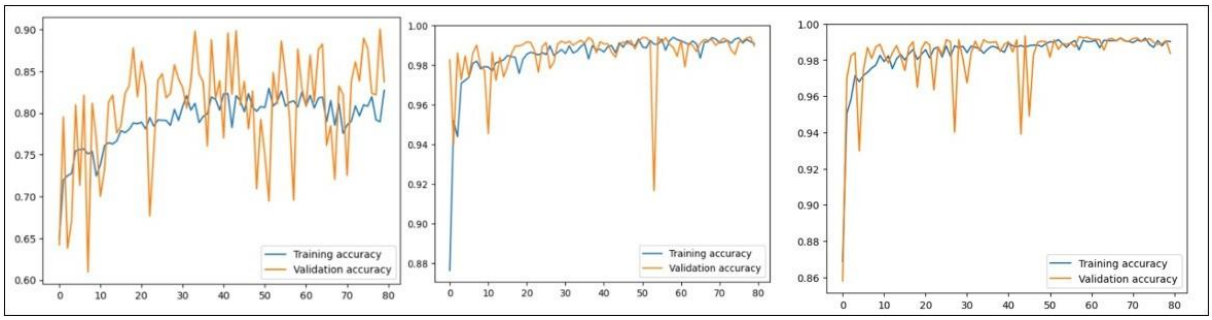


Fig. 7 Training and validation accuracy curves for liver tumor classification for 80 epochs based on input image dimension: 256×256 (left), 229×229 (middle), 224×224 (right).

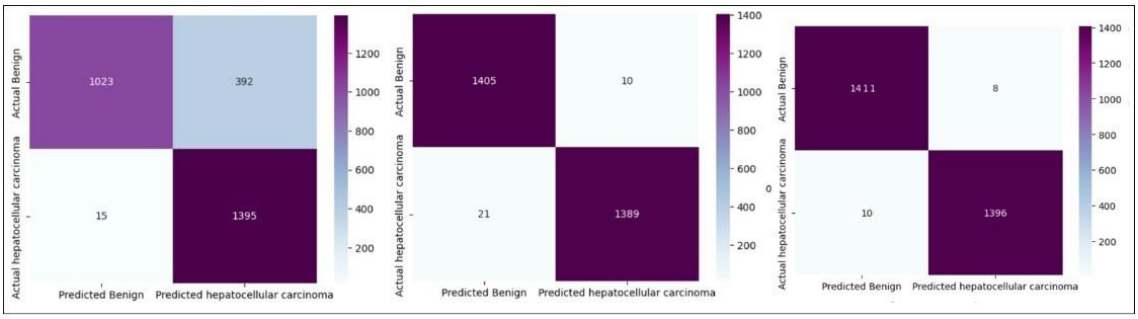


Fig. 8 Confusion matrices of the proposed work for various input image dimensions for 80 epochs: 256×256 (left), 229×229 (middle), 224×224 (right).

The input image dimension of 224×224 shows improved metrics in terms of accuracy, specificity, precision, recall and F1-Score (Table 2). The results demonstrated that with an accuracy of 99.36%, the predicted disagreement is reduced from 12 to 8 out of 2 825 images for the input image dimension of 224×224 which shows the robustness of the model. Additionally, the best specificity (99.43%), precision (99.44%), recall (99.30%), and F1-Score (99.37%) were obtained with the input image dimension of 224×224. The results reveal that input dimension and accuracy are negatively correlated, with input 224×224 exhibiting the maximum accuracy. On the other hand, a drop in performance has been noted for 256×256 input dimensions. The accuracy rate of 99.36% of the EfficientNet model is an exceptional performance, and all other metrics are similarly good.

Table 2. Performance analysis of the proposed EfficientNet architecture based on various input image dimensions for 80 epochs

Input image dimension	Accuracy, (%)	Specificity, (%)	Precision, (%)	Recall, (%)	F1-Score, (%)
256×256	85.59	78.06	72.3	98.55	83.41
229×229	98.90	99.29	99.29	98.53	98.91
224×224	**99.36**	**99.43**	**99.44**	**99.30**	**99.37**

A binary classification model's effectiveness is assessed using the area under the curve (AUC). Greater discrimination and overall model performance are indicated by a higher AUC value. With an input dimension of 224×224, an AUC of 0.991 is attained, indicating that the model is discriminating well at that moment in time. Fig. 9 displays the AUC values for the model's receiver operating characteristic (ROC) curves for various input image dimensions.

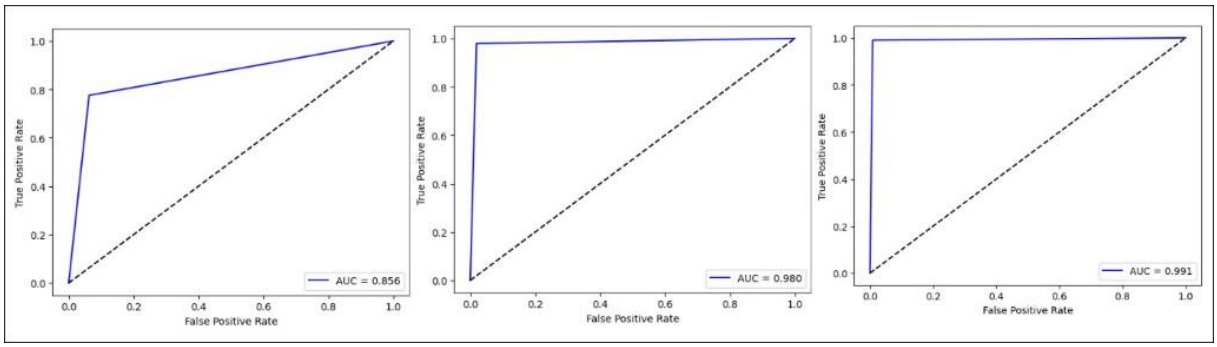


Fig. 9 AUC curves for various input image dimensions for liver tumor classification: 256×256 (left), 229×229 (middle), 224×224 (right).

The performance of many cutting-edge deep learning methods for liver tumor classification using LiTs dataset is compared in Table 3.

Table 3. Performance comparison of the proposed EfficientNet architecture with the state of art methods using LiTs database

State of art networks	Accuracy, (%)	Specificity, (%)	Precision, (%)	Recall, (%)
Chlebus [7]	87.00	-	-	92.00
Rela [14]	91.75	95.38	90.00	-
Kolli [10]	99.25	97.83	-	98.63
Proposed work	**99.36**	**99.43**	**99.44**	**99.30**

Table 3 displays the accuracy of the classification of liver tumors, which was reported by recent research [10] using improved probabilistic neural networks achieving an accuracy of 99.25%. Furthermore, opposition based hyena optimization method was used to classify liver tumors with great results in another work [14]. The accuracy of the investigation was 91.75%. The suggested EfficientNet design outperformed the other models in terms of accuracy, specificity, precision and recall.

Conclusion

In this work, a deep learning model for liver tumor classification was created and evaluated on medical envision data. Using volumetric imagery and the EfficientNet architecture, liver tumors may be effectively classified. Features are taken out of the images and employed in model training so that each image's class can be accurately predicted. The suggested model outperformed the other architectures, according to extensive testing using medical imagery. Pathologists should expect substantial advice from the process of using computer-assisted technologies to assess clinical specimens for efficiently categorising tumors, based on the results of current research as well as previously reported remarks.

Future research efforts will focus on segmenting the lesion's component, which will help the pathologist even more. By increasing the amount of images in the datasets being used, increasing the number of training epochs, and utilizing additional deep learning techniques like GanNet and MobileNet, research is being done to enhance the usefulness of the suggested model.

Acknowledgements

The Liver Tumor Segmentation (LiTs) database is freely available to the public and has been contributed by KAGGLE, which the authors thank for this. There was not a single government, private, or charity grant given specifically for this research.

References

1. Almotairi S., G. Kareem, M. Aouf, B. Almutairi, et al. (2020). Liver Tumor Segmentation in CT Scans using Modified SegNet, Sensors, 20(5), 1516.
2. Alom M. Z., T. M. Taha, C. Yakopcic, S. Westberg, et al. (2018). The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches, arXiv Preprint, arXiv:1803.01164.
3. Araújo J. D. L., L. B. Da Cruz, J. O. B. Diniz, J. L. Ferreira, et al. (2022). Liver Segmentation from Computed Tomography Images Using Cascade Deep Learning, Computers in Biology and Medicine, 140, 105095.

4. Bushara A. R., R. V. Kumar, S. S. Kumar (2023). An Ensemble Method for the Detection and Classification of Lung Cancer Using Computed Tomography Images Utilizing a Capsule Network with Visual Geometry Group, Biomedical Signal Processing and Control, 85, 104930.
5. Bushara A. R., R. V. Kumar, S. S. Kumar (2024). Classification of Benign and Malignancy in Lung Cancer Using Capsule Networks with Dynamic Routing Algorithm on Computed Tomography Images, Journal of Artificial Intelligence and Technology, 4(1), 40-48.
6. Chang C. C., H. H. Chen, Y. C. Chang, M. Y. Yang, et al. (2017). Computer-aided Diagnosis of Liver Tumors on Computed Tomography Images, Computer Methods and Programs in Biomedicine, 145, 45-51.
7. Chlebus G., A. Schenk, J. H. Moltz, B. van Ginneken, et al. (2018). Automatic Liver Tumor Segmentation in CT with Fully Convolutional Neural Networks and Object Based Postprocessing, Scientific Reports, 8(1), 15497.
8. Dorgham O., M. A. Naser, M. H. Ryalat, A. Hyari, et al. (2022). U-NetCTS: U-Net Deep Neural Network for Fully Automatic Segmentation of 3D CT DICOM Volume, Smart Health, 26, 100304.
9. Han K., Y. Wang, H. Chen, X. Chen, et al. (2022). A Survey on Vision Transformer, IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(1), 87-110.
10. Kolli S., B. R. Parvathala, A. P. Krishna (2024). A Novel Liver Tumor Classification Using Improved Probabilistic Neural Networks with Bayesian Optimization, e-Prime-Advances in Electrical Engineering, Electronics and Energy, 8, 100514.
11. Kumar S. S., R. Moni, J. Rajeesh (2013). An Automatic Computer Aided Diagnosis System for Liver Tumors on Computed Tomography Images, Computers Electrical Engineering, 39(5), 1516-1526.
12. Mienye I. D., Y. Sun, Z. Wang (2020). Improved Predictive Sparse Decomposition Method with DenseNet for Prediction of Lung Cancer, International Journal of Computers, 1, 533-541.
13. Pal M. (2005). Random Forest Classifier for Remote Sensing Classification, International Journal of Remote Sensing, 26(1), 217-222.
14. Rela M., S. Nagaraja Rao, P. Ramana Reddy (2021). Optimized Segmentation and Classification for Liver Tumor Segmentation and Classification Using Opposition Based Spotted Hyena Optimization, International Journal of Imaging Systems and Technology, 31(2), 627-656.
15. Sinha D., M. El-Sharkawy (2019). Thin Mobilenet: An Enhanced Mobilenet Architecture, IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), 0280-0285.
16. Szegedy C., S. Ioffe, V. Vanhoucke, A. Alemi (2017). Inception-V4, Inception-ResNet and the Impact of Residual Connections on Learning, Proceedings of the AAAI Conference on Artificial Intelligence, 31(1).
17. Tan M. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, arXiv Preprint, arXiv:1905.11946.
18. Tan M., Q. Le (2021). EfficientNetV2: Smaller Models and Faster Training, International Conference on Machine Learning, PMLR, 10096-10106.
19. Zheng Z., X. Zhang, H. Xu, W. Liang, et al. (2018). A Unified Level Set Framework Combining Hybrid Algorithms for Liver and Liver Tumor Segmentation in CT Images, BioMed Research International, 2018(1), 3815346.
20. Zoph B. (2016). Neural Architecture Search with Reinforcement Learning, arXiv Preprint, arXiv:1611.01578.
21. https://www.kaggle.com/datasets/andrewmvd/liver-tumor-segmentation (Liver Tumor Segmentation Challenge Dataset, LiTs) (Access date 19 September 2024).

Bilga Jacob, Ph.D. Student
E-mail: bilgajacob@gmail.com

Bilga Jacob received her B.Tech. Degree in Electronics and Communication Engineering from Calicut University, Kerala in 2006. She received her M.Tech. Degree in Electronics and Communication specialization in VLSI and Embedded System from Mahatma Gandhi University, Kerala in 2012. She is pursuing her Ph.D. Degree in Medical Image Processing and has contributed technical papers in various journals and international conferences. Her research interests are in the fields of biomedical signal processing, digital image processing and artificial intelligence.

R. S. Vinod Kumar, Ph.D.
E-mail: rsvinodkumar@niuniv.com

R. S. Vinod Kumar received his B.Eng. Degree in Electronics and Communication Engineering from Madurai Kamaraj University, Tamil Nadu in 1991. He received his M.Sc. Degree in Electronics and Control from Birla Institute of Technology and Science, Pilani in 1995. He is a gold medalist in his M.Eng. Degree in Communication Systems from Madurai Kamaraj University, Tamil Nadu, in 2000. He pursued his Ph.D. Degree from Anna University, Chennai in 2015 under the supervision of Dr. S. Arivazhagan., M.Eng., Ph.D., Principal, Mepco Schlenk Engineering College, Sivakasi. He has contributed 35 technical papers in various journals, and 45 papers in various international conferences. His areas of interest include digital image processing, analogue modulation systems, digital communications, linear integrated circuits, and digital electronics.

S. S. Kumar, Ph.D.
E-mail: kumarss@live.com

S. S. Kumar received his B.Eng. Degree in Electronics and Instrumentation Engineering from Manonmaniam Sundaranar University, Tamil Nadu in 2002. He received his M.Eng. Degree in Applied Electronics from Anna University, Tamil Nadu in 2004. He pursued his Ph.D. degree from Anna University, Chennai in 2013. He has more than 20 years of teaching experience and 14 years of research experience. His areas of interest include medical image processing, artificial intelligence, satellite image processing, and embedded robotics. He is an active supervisor for many Ph.D. scholars, postgraduate and undergraduate students. He has contributed 50 technical papers in various peer-reviewed journals, and 30 papers in various international conferences.