RESEARCH ARTICLE | JULY 09 2025

An extensive examination of gastrointestinal abnormalities utilizing endoscopic imaging and deep learning

S. Srinivasulu Raju; S. Hariprasad; R. Swapna; D. Bhanuchandra Rao

AIP Conf. Proc. 3298, 020053 (2025)
https://doi.org/10.1063/5.0279397

An Extensive Examination of Gastrointestinal Abnormalities Utilizing Endoscopic Imaging and Deep Learning

S. Srinivasulu Raju [1, a)], S. Hariprasad [2, b)], R. Swapna [3, c)], and D. Bhanuchandra Rao [4, d)]

[1]Department of Electronics and Instrumentation Engineering, VR Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India
[2]Department of Physics, Aditya Institute of Technology and Management (AITAM), Tekkali, Srikakulam, Andhra Pradesh-532201, India
[3]Department of Physics, Basic Science & Humanities, Avanthi Institute of Engineering and Technology, Tagarapuvalasa-531162, Vizianagaram, Andhra Pradesh, India
[4]Department of Mechanical Engineering, Aditya Institute of Technology and Management (AITAM), Tekkali, Srikakulam, Andhra Pradesh-532201, India

[b)] Corresponding author: sh.prasad5@gmail.com
[a)] srinu85raju@gmail.com
[c)] kotharu.swapna@gmail.com
[d)] bhanurch@gmail.com

Abstract. Gastroenterologists perform exams of aberrant tissue in the gastrointestinal tract, during which they consider polysegmentation to be a challenging task. Although it offers crucial support to gastroenterologists. Polyps, which are anomalous proliferations of tissue, predominantly arise in the colorectal portion of the gastrointestinal tract. Moreover, it is primarily situated in the mucous membrane, which already has multiple protrusions of the micro aberrant tissue, greatly increasing the likelihood of disease progression. Identifying polyps at an early stage helps prevent their transformation into dangerous tissue, specifically adenomas, which have the potential to grow into cancer. The identified anomalies consist of raised polyps that have been stained, a normal cecum, a normal pylorus, a normal z-line, esophagitis, polyps, ulcerative colitis, and stained resection margins. The identification of these irregularities is accomplished through the utilization of the pre-processed en-coder and decoder phases. The implementation of the preprocessing phase can greatly improve the quality of colonoscopy pictures. The gastrointestinal tract primarily shows abnormalities such as polyps and colitis. The VGG convolutional model has superior accuracy in comparison to other contemporary deep learning models. In this study article, we employed the Kvasir data collecting. The Kvasir dataset consists of 4,000 annotated photographs of the gastrointestinal tract, which are divided into 8 separate categories. Every class has roughly 500 photos. Data sets primarily contain information specifically related to the bowel portion of the gastrointestinal system that is obtained during colonoscopy. The VGG model demonstrates exceptional performance on the Kvasir dataset, achieving accuracy levels ranging from a minimum of 87% to a maximum of 97%.
Keywords: *Segmentation, VGG CNN (Convolutional Neural Network), R-CNN, ResNet, AlexNet, UNet, Inception Net, Colorectal, Polyps*.'

INTRODUCTION

The gastrointestinal system comprises the stomach, large intestine, small intestine (including the colon), rectum, and anus, which are the primary constituents of our digestive tract. The gastrointestinal system is the primary organ of the human body's digestive tract, distinguished by mucosal enlargements that can vary in severity from mild to chronic illnesses [1]. Polyps are anomalous proliferations that usually develop beneath the mucosal layer. The colorectal portion of the gastrointestinal tract is primarily composed of abnormal tissue called polyps. Polyps that typically develop in the colorectal area can be classified into two primary categories: neoplastic and non-neoplastic. There are five primary types of non-neoplastic polyps, namely hyperplastic polyps, hamartomatous polyps, and inflammation in the inner lining of tissue [2]. Neoplastic polyps have the potential to progress into cancerous forms through malignant transformation as time passes. The probability of this change is estimated based on the condition and size of the polyp. The colorectal area is the main location where polyps and other important gastrointestinal disorders arise. This area consists of the main regions of tissue growth in the innermost layer. This type of proliferation typically arises after the progression of colorectal cancer, a highly perilous and potentially lethal condition.

AI Powered Technology Integration for Sustainability (AI-PTIS-2024)
AIP Conf. Proc. 3298, 020053-1–020053-7; https://doi.org/10.1063/5.0279397
Published under an exclusive license by AIP Publishing. 978-0-7354-5217-6/$30.00

The motivation for this study is to investigate the impact of colorectal cancer stage at diagnosis on mortality rates. The overall five-year survival rate for colorectal cancer, regardless of the stage, is approximately 65.1%. Therefore, the survival percentage for colorectal cancer patients five years after diagnosis is 65.1%, whereas the mortality rate stands at 34.9%. Early detection of anomalies is crucial for enhancing outcomes and decreasing mortality rates. The incitements are. Colonoscopy is a frequently performed medical technique utilized to detect problems in the colorectal region. The procedure entails the inspection of the colon and rectum by means of a flexible tube equipped with a camera. Laser endomicroscopy is a microscopic technology that enables easy visualization of subcellular structures during biopsy operations. Convolutional neural networks (CNNs) are essential in computer vision research. They have been utilized to identify abnormalities such as polyps in images from colonoscopies. This paper explores the process of identifying, categorizing, and understanding the structure of polyps through the utilization of the VGG model. The next section showcases all of the experimental findings and offers a thorough analysis. The discussion section also articulates the future prospects of the research endeavor in a concise manner.

LITERATURE REVIEW

The mortality rate for colorectal problems is a mere 10%. A colonoscopy procedure is conducted to excise the colon upon identification of abnormalities in the colorectal area, with the objective of mitigating premature mortality rates. Various endoscopic techniques can be used to identify anomalies in the gastrointestinal system, particularly in the conical focal region. Laser endomicroscopy is an advanced microscopic technique utilized in biopsy procedures, enabling the convenient capture of images at the subcellular level [3, 4]. Gastroenterologists perform histopathology manually to examine tumors or aberrant tissue growth. By identifying deviations at an early stage, we can minimize the mortality rate resulting from abnormal tissue and preserve lives. Endocytic scopy is a method used by endoscopists to acquire images of microvascular tissue utilizing the NBI (Narrow Band Imaging) mode. This approach enables the retrieval of high-resolution images with a maximum scale of 520. The article has been organized in the following manner: The introduction is succeeded by a section on the related research carried out. The next section outlines the materials and methods used for the proposed VGG convolutional network. Banik et al. [5], Sun et al. [6], and Sanjana et al., [7] have discussed in recent decades, there has been a significant focus on researching automatic disease diagnosis and segmentation, which has shown to be quite intriguing. There are already various appropriate techniques or algorithms available for detecting gastrointestinal anomalies. Over time, several development approaches and algorithms were considered to enhance the texture and color scheme of the polyp. In this research study, we have referred to handcrafted learning characteristics as descriptors. Mori Y et al., [8] discussed about the Convolutional neural networks play a crucial role in the research field, particularly in addressing impending issues. They have proven to be useful for mankind, namely in the domain of computer vision. Polyp and other anomalies are typically detected using colonoscopy photos and movies. CNN transfer learning is mostly utilized for the Inception and ResNets models. The framework for disease identification and segmentation challenges has been developed using Generative Adversarial Networks. An advanced and highly sensitive object detection algorithm, known as YOLO (You Only Look Once), has previously been created for segmentation methods. Kara et al., [9] proposed a methodology to transfer Learning assesses the degree of specificity and sensitivity in the context of anomalous polyp segmentation. Kronbrog et al., [10] introduced data-driven methodologies to incorporated segmentation elements that are beneficial for the detection of polyps. Object segmentation employs down sampling and up sampling algorithms to classify pixels. Long has proposed the utilization of a fully convolutional network for the purpose of dissecting anomalous polyps. UNet architectures are a type of convoluted networks. Mekala et al. [11], Razman et al. [12], Rasheed et al., [13] discussed about the UNets consist of two main pathways: an encoding pathway and a decoding pathway, which are responsible for synthesizing and classifying information. The analysis path discusses the characteristics of deep learning algorithms, while the synthesis component introduces a technique for segmenting features in the learning paradigm. UNet and FCN both utilize an encoder and decoder as essential components in semantic segmentation. The neural network utilizes encoder-decoder convolutional layers, with the encoder extracting important semantic characteristics across different levels.

The present study seeks to enhance colonoscopy by employing deep learning techniques to automatically detect and segment polyps during the process. This method has the potential to decrease mortality rates associated with colorectal cancer. The section examines multiple established ways for identifying polyps, such as colonoscopy with different endoscopic procedures and manually designed learning algorithms. Nevertheless, the attention is redirected towards the capability of Convolutional Neural Networks (CNNs) for this particular undertaking. Scientists are investigating well-known Convolutional Neural Network (CNN) structures such as VGG, Inception, and ResNets, in addition to transfer learning methods, in order to create a system that can automatically identify and separate polyps in colonoscopy pictures. Implementing this could potentially alleviate the burden on gastroenterologists and enhance the precision of polyp identification.

Table 1 presents a condensed representation of the features utilized by the VGG net model. The VGG model employed the Kvasir dataset for the purpose of detecting anomalies in polyps. Here, we have included photos of all the potential gastrointestinal abnormalities, each exhibited separately.

TABLE 1. *Displays the parameters of the vgg sequential model*

Type of Layering	Output Shaping Values	Parater Values
VGG 19	None 3,3,512	232134368
Flatten	4609	0
Dense	1034	4178269
Dense_1	484	520824
Dense_2	224	131326
Droup Out	224	0
Dense_3	112	35436
Dense_4	8	1243

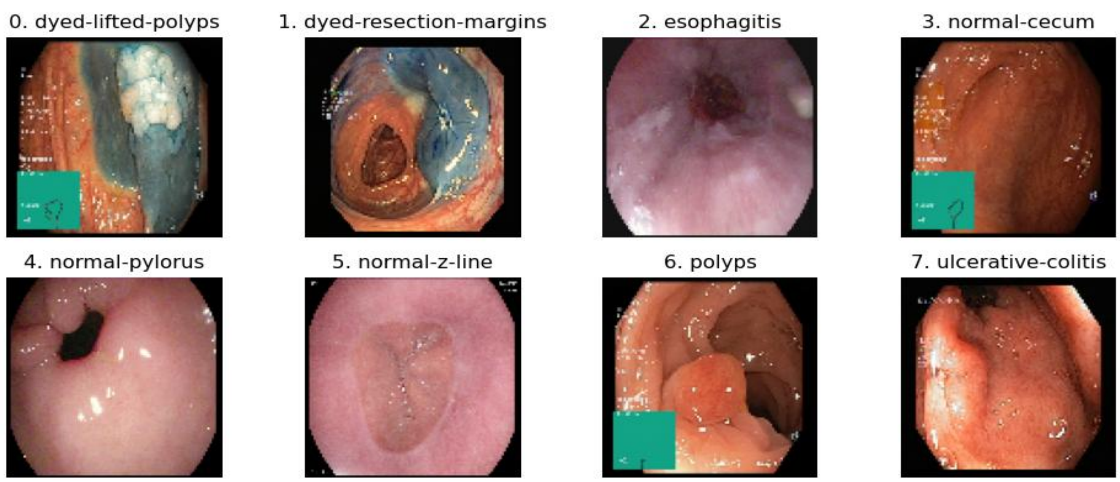

FIGURE 1. *Displays photos of each category of gastrointenstial abnormality*

METHODOLOGY

Colorectal cancer is a prominent contributor to global mortality, although the timely identification of the disease with colonoscopy can greatly enhance patient prognosis. Nevertheless, conventional colonoscopy depends on the expertise and proficiency of the gastroenterologist in detecting polyps, which are abnormal growths that may develop into cancer. This procedure is susceptible to subjectivity and human fallibility, which may result in the overlooking of polyps or the performance of unneeded biopsies.

In this present study investigates the capacity of deep learning, particularly convolutional neural networks (CNNs), to overcome these constraints and enhance colonoscopy. CNNs are a very effective category of machine learning algorithms that demonstrate exceptional performance in tasks related to image recognition and segmentation. By training a CNNs on a substantial dataset of colonoscopy pictures that include both polyps and healthy tissue, the model can acquire the ability to autonomously identify and separate polyps. This can aid gastroenterologists in the process of diagnosing and treating colorectal cancer. The suggested method employs a deep learning framework featuring an encoder-decoder structure. The encoder's role is to extract features from the colonoscopy image. It usually comprises a sequence of convolutional layers accompanied by pooling procedures. These layers acquire the ability to recognize basic characteristics such as edges, textures, and colors, and gradually integrate them to create more complex characteristics that indicate the existence or non-existence of polyps. The decoder utilizes the extracted features obtained from the encoder to produce a segmentation mask. The mask is a binary image in which each pixel is categorized as either part of a polyp (foreground) or healthy tissue (background). The decoder commonly utilizes upsampling layers to enhance the resolution of the feature maps and convolutional layers to improve the precision of the segmentation boundaries. The architecture includes skip connections, which are direct connections between corresponding layers in the encoder and decoder that have the same spatial resolution. These connections enable the decoder to directly access the low-level features from the encoder. This is critical because low-level features provide vital information about the spatial relationships between pixels, which is necessary for precise segmentation. The decoder utilizes the high-level semantic characteristics from the encoder along with the low-level spatial data from the skip connections to produce segmentation masks that are more precise and accurate.

The study used a pre-trained VGG16 network as the encoder. VGG16 is a widely recognized CNNs structure that has been effectively utilized in many image recognition assignments. By utilizing a pre-trained network, the researchers may take advantage of the considerable knowledge the model has previously acquired from a large dataset of genuine photos. The pre-existing information serves as a solid base for extracting features in the context of colonoscopy pictures, hence decreasing the amount of time and computational resources needed to train the model from the beginning. Transfer learning is a method in which a model that has already been trained is adjusted to perform a new task. Here, a pre-trained ResNet50 model could be modified to accurately detect polyps in colonoscopy pictures. The lower layers of the network, responsible for learning general image characteristics, will remain unchanged, while the upper levels will be trained again using the colonoscopy dataset to focus on identifying polyps.

Rasheed S et at., [14, 15] Residual networks, such as ResNet50, are utilized for disease detection and localization in transfer learning. The residual network employs identity mapping with 3*3 convolutional layers. Identity mapping involves the elimination of deeper neural networks due to the occurrence of vanishing gradients and expanding gradients. The model evaluation in this study utilizes the Kvasir data set. The model evaluation in this work employs the Kvasir dataset. This dataset comprises a publicly accessible compilation of colonoscopy pictures encompassing both normal tissue and polyps. Through the evaluation of the model's performance on this dataset, the researchers can gauge its capacity to precisely segment polyps and distinguish them from healthy tissue

RESULT AND DISCUSSIONS

The KVASIR collection comprises 4,000 annotated photographs of the gastrointestinal tract, which are classified into 8 main groups that indicate different irregularities. Every class consists of 500 images. Here, we have created instances of classes based on the names of the folders containing the datasets. The dataset consists of 8 folders, each representing a separate class. Every folder contains a grand total of 500 photographs. We have assigned 30% of the training set to the validated data set. Data augmentation has been executed. Data augmentation is a method used to increase the size of a training dataset by creating modified images derived from the original datasets. The altered photographs are rotated at a randomly selected angle ranging from 0 to 360 degrees. We created instances of the Image Data Generator objects to implement augmentation techniques and then applied scaling to the training, testing, and validation datasets. This section presents an elucidation of the Kvasir dataset and examines the metrics used to evaluate performance. The Kvasir dataset is detailed in the data methodology part of the visual geometry group. We performed training and testing on an Nvidia GTX 1070 GPU setup for the data component of the visual geometry group. We employed a Windows 10 operating system on a machine that featured a core i5 processor and 8GB of inbuilt RAM. Colab and Py-thon Spyder IDE are employed for visualizing and doing model evaluation. This section includes the evaluation methodology, performance outcomes, and analytical presentation of graphs for the Visual Geometry Group model.

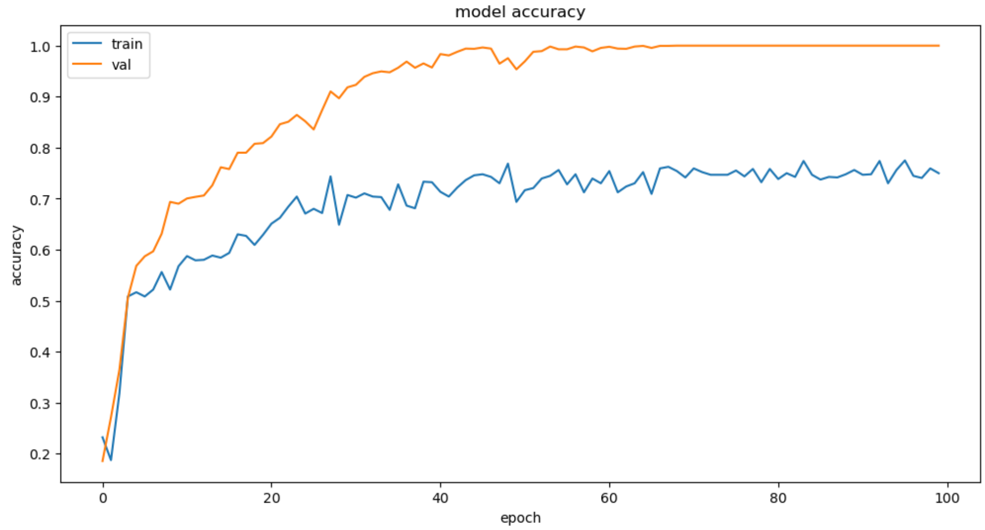

FIGURE 2. *Displays the accuracy of the training and validation dates*

In this area, we have provided the model's performance evaluation procedures and features visualization and shown in figure 2. We have initialized the VGG 19 model, a pre-trained convolutional neural network, which performed exceptionally well as a transfer learning model. VGG 19 has a significant advantage in that it has undergone pre-training on millions of photos from the ImageNet collection. This pre-trained network can classify photos into thousands of object categories. We have implemented the VGG 19 model by incorporating additional layers to accommodate the 8 classes. A pre-trained network accurately classified a wide range of object categories. We employed a sequential paradigm to construct this dense network. The addition of layers is accomplished by the utilization of batch activation and normalizing techniques. Hyper parameters are adjustable parameters that can be used to modify the modeling process. Hyper parameters play a crucial role in determining the performance of the model. Compilation hyper parameters re applied after the initialization procedure.

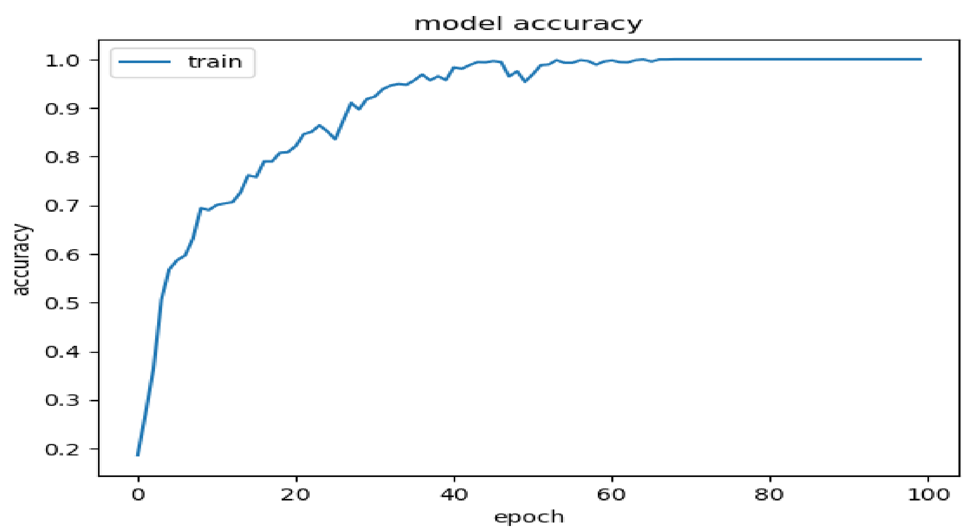

FIGURE 3. *Displays the accuracy of the training data*

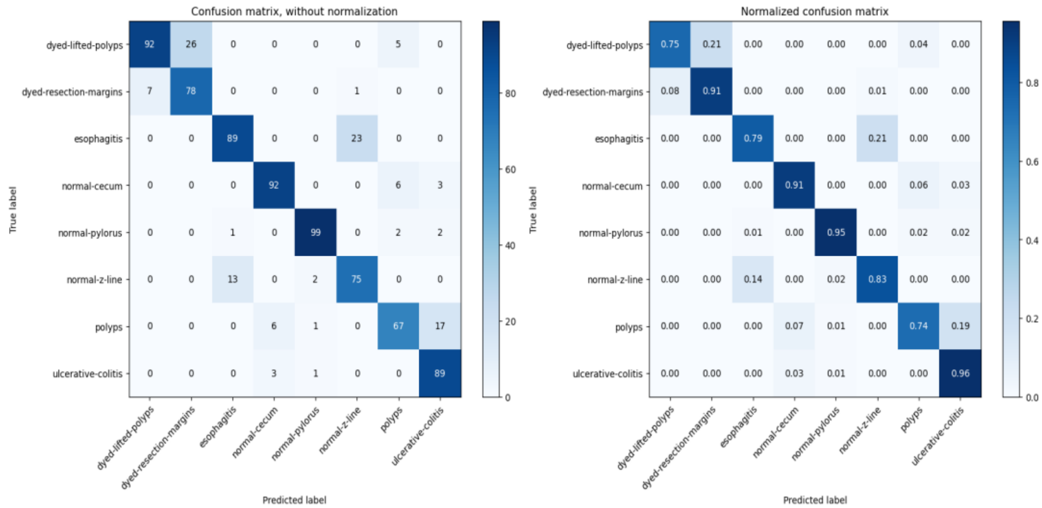

FIGURE 4. *Displays the classification performance using a confusion matrix*

In this part, we have generated and displayed both the normalized and unnormalized confusion matrix. We have examined the accurate and inaccurate categorizations of numbers using both normalized and normalized confusion matrices. From figure 3, it is evident that we have achieved a minimum accuracy of 79% and a maximum accuracy of 97%. We have categorized all the photos into eight primary classes. As shown in figure 4, we have optimized our training settings to enhance the likelihood of retaining and expanding our operations.

CONCLUSION

The implemented vgg net model demonstrates greater accuracy in semantic segmentation when compared to other convolutional neural network models. Vgg employs the kva-sir datasets to achieve precise segmentation of colorectal polyps. The clahe method is commonly used to enhance the intensity level of the segmented frames in the kvasir dataset during the initial step of preprocessing. The vgg convolutional neural network consists of encoding and decoding blocks. In this study, we have created five double-stranded dna (dsb) universal serial bus (usb) blocks. To optimize the feature mapping for all eight categories, a graft network is implemented in each dsb block of the encoder. The vgg model, as proposed, yields a minimum accuracy of 85% and a maximum accuracy of 97%. Resnet, inception nets, and alexnets can be used as substitutes for the encoder. The vgg model can be used as a reference point to achieve wide-ranging goals and showcase advanced establishment approaches.

FUTURE SCOPE OF WORK

Future research in this field should prioritize the acquisition of more extensive and varied datasets to enhance the models' ability to generalize. Furthermore, the investigation of various architectural designs and the integration of attention mechanisms may result in more accurate polyp segmentation. Models that can operate in real-time and can be easily incorporated into clinical workflows are essential for practical use. It is crucial to examine ethical factors such as mitigating bias, obtaining regulatory approval, and promoting collaboration between humans and ai in order to ensure responsible deployment. Through surmounting these obstacles, deep learning possesses the capacity to emerge as a valuable instrument for gastroenterologists, resulting in earlier and more precise identification of polyps, ultimately resulting in life-saving outcomes.

ACKNOWLEDGEMENTS

The authors are thankful to the director, Aditya institute of technology and management (AITAM), Tekkali, A.P., India for the constant support and encouragement. The authors are also thankful to the principle, VR Siddhartha Engineering College, Vijayawada, A.P., India for carrying out this research investigation. The authors are also thankful to the Director, Avanthi Institute of Engineering and Technology, Vizianagaram, A.P., India for the constant support.

REFERENCES

1. W. M. Grady, adv. Cancer. Res. **151**, 425-468 (2021).
2. J. Naz, M. Sharif, M. Raza, J. H. Shah, M. Yasmin, s. Kadry, S. Vimal. *Neural process*, **4**(1), 1-26 (2021).
3. A. Nagahara, A. Shiotani, K. Iijima, T. Kamada, Y. Fujiwara, K. Kasaugi, M. Kato, *Dig. Endosc.* **34**(1), 63-72 (2021).
4. J. Naz, M. Sharif, M. Yasmin, M. Raza, M.A. khan. *Curr. Med. Imaging rev*, **17**(1), 479-490 (2021).
5. S. S. Raju, M. Srikanth, K. Guravaiah, P. Pandiyaan, B. Teja, & K.S. Tarun, International conference on innovative trends in information technology (ICITIIT) IEEE 2023, pp.568-575.
6. D. Banik, k. Roy, d. Bhattacharjee, m. Nasipuri, o. Krejcar, *IEEE transactions on instrumentation and measurement* **70**(1), 1-12 (2020).
7. S, sun, J. Guo, M.S. Bhutani, M. Giovanini, Z. Li, Z. Jin, A. Yang, G. Xu, G. Wang, *Front. Surg.* **6**(1), 376-381 (2017).
8. C. V. Sanjana, S. S. Raju, M. Anusha, & B. Sivani, classification and detection of skin cancer using deep learning methods. International conference on computer communication and informatics 2023 (ICCCI), pp.15-93.
9. Y. Mori, S. Kudo, N. Ikehara, K. Wakamura, Y. Wada, M. Kutsukwa, M. Misawa, T. Kudo, Y. Kobayanshi, H. Miyanchi, *Endoscopy*, **45**(1), 98-105 (2013).
10. S. T. Kara, S. Loya, S. S. Raju, N. Vanteru, B. Rajulapati, detection of fake indian currency using deep convolutional neural network. Ieee 3rd mysore sub section international conference 2023. (mysurucon), pp.1-8.
11. O. Kronbrog, J. Regula, *Dig. Dis.* **25**(3), 1069-1077 (2007).
12. S. Mekala, S. S. Raju, M. Gomathi, K. N. V. Rao, D. Kothandaraman, S. Sharma, human breast cancer classification employing the machine learning ensemble. Human cancer diagnosis and detection using exascale computing, pp.19-34. (2024).
13. M. Razman, M. Raza, M. Sharif, M. A. Khan, Y. Nam, *Comput. Mater. Contin.* **69**(1), 3239-3257 (2021).
14. S. Rasheed, M. Raza, M. Sharif, S. Kardry, A. Alharbi, *J. Pers. Med.* **12**(1), 12-32 (2022).
15. S. S. Raju, B. Teja, R. Bhuvaneswari, G. U. Stella, machine learning algorithms for prediction of stock aarket: a systematic literature review. Ieee international conference on computing, power and communication technologies (IC2PCT), 6(1), 96-102 (2024).