Pakistan Journal of Engineering and Technology (PakJET)
Multidisciplinary & Peer Reviewed
Volume: 5, Number: 1, Pages: 1- 10, Year: 2022

Machine Vision based Intelligent Breast Cancer Detection

Nof Yasir*, Shahzad Anwar and Muhammad Tahir Khan
Department of Mechatronics Engineering, University of Engineering Technology, Peshawar,Pakistan
*Corresponding author: Nof Yasir (e-mail: nof_yasir@uetpeshawar.edu.pk)

Abstract- **Artificial intelligence, especially deep learning, has sparked a great deal of interest in bioinformatics, particularly complications in clinical imaging. It has achieved great success by helping the CAD system achieve high-precision results. Despite this, detecting breast cancer on mammography images is still considered a critical challenge. The work aims to decrease False-positive rate (FPR) and False-negative rate (FNR) and increase the value of Mathews correlation coefficient (MCC). To achieve this goal, two state-of-the-art object detection models are used, YOLOv5 and Mask RCNN.YOLOv5 detects and classifies the mass as benign or malignant. Due to the spatial limitations of YOLOV5, the original model is modified to achieve the desired results. Mask RCNN detects the edges of tumours invading the breast parenchyma and also detects the size of the tumours. The size of the tumours defines the stage of cancer. The model was trained on the INbreast dataset with YOLOv5+Mask RCNN. The performance of the proposed model was evaluated compared to the original version of YOLOv5. The proposed technique achieves higher performance with a lower False-positive rate of 0.05 and False-negative rate of 0.03 and a high MCC value of 92.02%. The experiments performed show that the accuracy of YOLOv5 in combination with Mask RCNN is 0.06 higher than that of YOLOv5 alone. Additionally, this work could help determine the patient's prognosis and allow physicians to be more accurate and predictable at early-stage breast cancer detection**.

Index Terms—**Machine learning, Biomedical Engineering, Clinical image processing**

I. INTRODUCTION

Breast cancer is one of the common invasive cancers and the most liable reason for the cause of cancer deaths in women in the world. In 2018, 2.1 million new cases were reported, with 627,000 losses [1]. Although lung, cervical, and thyroid cancers are also common among women, breast cancer represents 1 in 4 cancers diagnosed worldwide. According to the World Health Organization (WHO), 19.3 million cancer cases are expected in 2025 [2]. Several studies show a global increase in mortality due to breast cancer in most regions and age groups. Though there is a global increase in Breast cancer incidence in all age groups however is highest in women below the age of 50 i.e. before menopause followed by a gradual increase [3]. Among Asian countries, Pakistan has one of the highest rates of age-standardized incidence of breast cancer[4] . It is the highest among Asian countries, every one in nine women is at risk of being diagnosed with this disease during their lifetime [5]. Unfortunately, due to delayed diagnosis and regional/cultural restrictions, insufficient diagnostic equipment, and treatment facilities, the fatality rate of breast cancer patients in the country is high and moving towards increased incidence. Worldwide research and studies conclude early management is a key solution to cure breast cancer. A patient with early-stage diagnosed cancer has a greater recovery possibility and survival than the one diagnosed with a later or metastatic stage. Hence, awareness in the public about regular screening and early diagnosis in case of any symptoms and its treatment is emphasized to significantly reduce the morbidity and mortality rates in the long term [6]. Mammograms, being one of the best screening tests have limitations, such as they are not 100% accurate in detection if a woman has false negative or false positive breast cancer. Studies have also proved that all breast cancer detected on mammograms is not exactly detected by radiologists. Due to the elusive and complex nature of the radiographic findings, some of the small calcifications and low contrast image features could be misinterpreted in the radiological diagnosis making it difficult [7]. Computer-Aided Diagnosis (CAD) hence provides a second opinion for the accurate diagnosis. It could spot changes that are invisible to the human eye. Furthermore, this study also allows the detection of other breast-related diseases and may recommend the tumours as benign or malignant.

Breast tumours are broadly classified into two categories, benign and malignant [8]. Benign or non-cancerous tumours may be a cyst or fibroadenoma. They have smooth well defined borders. They are also called carcinoma in situ which means they do not spread to other parts of the body. The malignant or invasive cancerous tumours are found in the milk duct. They grow abnormally invade into the breast parenchyma and gradually to other parts of the body [9]. Early diagnosis and in time treatment could stop the uncontrollable abnormal cell division. The percentage of patients who were diagnosed and treated at Stage-I, II, and III cancers had survived in the next five years are shown in Table I [10]. The ones who were treated at a premature stage had

a 100% survival rate than the patients who were diagnosed and treated at a later stage A. Hence regular screening and early diagnosis can prevent the initial stage of cancer to invade other parts of the body.

Table I: The 5-year survival rate

Breast Cancer Stage	5-year breast cancer-specific survival
I	98-100 %
II	90-99%
III	66-98%

This work focus on the research gap associated with Metastatic Merkel cell carcinoma (MCC) values. MCC is a more reliable and robust metric than accuracy as it summarizes the classifier performance in a single value [11]. This study aims to achieve a high MCC value by reducing the false-negative and false-positive rates. The study proposes the following contributions:

- A model that detects the breast mass on mammogram images and classifies them as benign or malignant.
- The misclassified data is again subjected to another model which precisely detects the borders of the tumours and determines the size of the tumour indicating the stage of cancer.
- Increases the MCC value by lowering FNR and FPR without reducing the accuracy.
- Compares the proposed model with related works to evaluate the performance of the system.

II. LITERATURE REVIEW

Image analysis and evaluation play an important role in the field of medical science. It helps the radiologists and practitioners to visualize, recognize and identify the abnormalities in that area. Mammograms are an impressive tool for the diagnosis of breast cancer [12]. However, they miss about 20% of cancers giving false-negative results. False-negative diagnosis leads to delayed treatment and in some cases even lethal consequences. Artificial intelligence (AI) based techniques have been contributing successfully to predicting cancer. Various AI diagnostic CAD tools are established to assist clinicians in the detection of breast cancer with high accuracy. Overall, the tool detects cancer just as well as an average radiologist, however, the advantage is the CAD tool works together with a human radiologist to produce accurate results. Frequent effective and efficient efforts have been made to develop innovative CAD systems for the detection and classification the breast mass as malignant or benign. All the works aim to achieve high accuracy with efficient time and cost management. The work proposed in [13] was the feature extraction using VGG-16 incorporated by SSD. They employ Convolution Neural Network (CNN) on a large data of Digital Database for Screening Mammography (DDSM) containing 2620 cases. The results generated an accuracy of 96.2%. For an increased number of cases, the procedure was limited to perform efficiently in noisy images. Authors in [14] applied a Decision tree induction algorithm to train the model on the data set of mammograms having cancerous and noncancerous cells. Adopting CART as a white-box method, being transparent and explainable, Linear projections were employed to explore and visualize the data after applying preprocessing techniques. Features were selected using the RFE feature selection method incorporated with hyperparameters. Although the results obtained by CART were not as accurate as of the results of the Blackbox model, there could be a tradeoff between accuracy and transparency.

A new approach in [15] was proposed to reduce the false-positive cases. The experiments were conducted on ultrasound breast images. The features were first extracted from both the ROI and background. For the detection of ROI from the background, the images were subjected to be scanned by a fixed size window. Out of 250 images, 150 were benign and 100 were malignant. The proposed method achieved an accuracy of 95.4%. A study in [16] proposed that machine learning was a decent approach for breast cancer detection, performing efficiently on liner data, however, when the data was in form of images deep learning was an innovative technique. Convoluted Neural Networks(CNN) were one of the best methods for breast cancer classification [17]. Wisconsin database for breast cancer with 569 instances and 32 variable attributes was employed. The methodology was divided into two stages. The first stages involved feature extraction using Principal-component-analysis (PCA) and reduction using Linear-discriminant-analysis (LDA). The second stage involved the testing and training of the reduced dataset over three different classifiers namely support vector machine (SVM), Adaptive neural network-based fuzzy-rule interference system (ANNFIS), and multi-layer perceptron (MLP). Among all classifiers, the proposed method of joint PCA and LDA showed the highest accuracy of 98.6%. The method could also be employed for real-time multidimensional images. To develop a CAD system based on feature extraction using a well-known Deep Convolutional Neural Network (DCNN) architecture named Alex Net [2] was proposed. Support Vector Machine (SVM) was connected to the last layer to obtain better accuracy. A digital database for screening mammography (DDSM) and the Curated Breast Imaging Subset of DDSM (CBIS-DDSM) were employed. The accuracy of SVM was 87.2% with 0.94(94%) of the area under the curve (AUC). The method was implemented to a relatively small number of samples of the biomedical datasets. However, training on a big set of data may offer better accuracy. A machine learning approach using the feature selection method was proposed [18]. The features extracted were the morphology of breast cancer cells, their molecular and clinical features of histology. The

resulting computational multiplex-histology analysis resulted in an accuracy of 95%. According to the study the most frequently employed methods support vector machines employed were 51.6%, artificial neural networks 58.1%, decision trees 61.3%, and ensemble learning 32.3%. 25 studies showed a sensitivity ranging from 0.037 to 1. About 24 studies exhibited specificity from 0.008 to 0.993. 20 studies revealed the AUC from 0.500 to 0.972. And six studies had precision ranging from 0.549 to 1 [19]. All the models except one were internally validated. Hybrid techniques of Artificial Intelligence (AI) could be employed to obtain better accuracy of classification of benign and malignant tumours [20]. It suggested the development of CAD tools having time-efficient and improved accuracy features. The classification techniques for the detection of breast tumours of various classifiers were compared in terms of accuracy, sensitivity, and specificity of diagnosis of breast cancer. The method proposed in [21] employed five learning algorithms SVM, Naïve Bayes, k nearest neighbour (K-NN), Random Forest, and logistic regression. Data set from the University of Wisconsin Hospitals Madison Breast Cancer Database was employed. Method of feature selection was employed for the classification of breast tumours. SVM outperformed the other methods in terms of sensitivity, specificity, and precision of classification with the highest accuracy of 97.9%.%. A novel approach named BCD-WERT was proposed employing Whale Optimization Algorithm (WOA) and Extremely Randomized Tree for feature selection and classification of breast tumours. The results were compared with eight Machine learning algorithms which included SVM, logistic regression, Kernel support machine, Random Forest, k-Nearest neighbour, Gradient descent, and Gaussian Naïve Bayes. The proposed method for feature selection outperformed all other methods with the highest prediction accuracy of 99.30% [22].

The method proposed in [23] employed the prediction power of neural networks of 5-year survival of breast cancer patients. An Artificial Neural Network (ANN) was trained and tested using the Wisconsin Breast Cancer Dataset. Using Just Neural Network (JNN) environment and testing the network on Haberman's Breast Cancer Survival dataset which was collected from the Center for Machine Learning and Intelligent Systems, University of California, showed outperformance of the work. The accuracy achieved was 88.24%. An algorithm called delay-multiply-and -sum (DMAS) [24] was proposed which involved the technique of implementing ultra-wideband confocal microwave imaging. This method resulted in improved accuracy as compared to the DAS imaging algorithm. The conformal predictors' method was proposed using a rule-based genetic algorithm [25]. The method was implemented on two datasets, one on the breast cancers dataset and one on the dataset gathered for ovarian cancer. The method successfully displayed the predictive areas. The readability of the rules made this method more efficient than other conformal predictors. A study in [26] compared YOLO series and proved ResNet and Inception to perform better with an accuracy of 91% and 95.5%.

Numerous research and several conventional cancer detection methods employed in the recent studies achieved accuracies even to 99%. However, the emphasis on a crucial aspect misclassification ratio and MCC score is still lacking. Computing accuracy on confusion matric is a useful option in binary classifications for balanced datasets. But when there is an imbalanced dataset, MCC is a more reliable statistical measure. It shows a high value only when all the four categories in the confusion matrix predict good results. Moreover, when FNR and FPR are low the MCC score is high [27]. Adopting only high accuracy as a measure of good prediction may lead to dangerous results in clinical diagnosis. If the model predicts a false positive tumour, the patient would have to suffer mental stress along with painful procedures like biopsy or surgery until the tumour is declared benign by the pathologist. On the other hand, if the model predicts false-negative tumours, the results would even be worse and life-threatening due to delay and false diagnosis. Hence, it is necessary to build an effective and authentic model which not only reduces the false classification ratio but also lowers FNR and FPR and boosts MCC score.

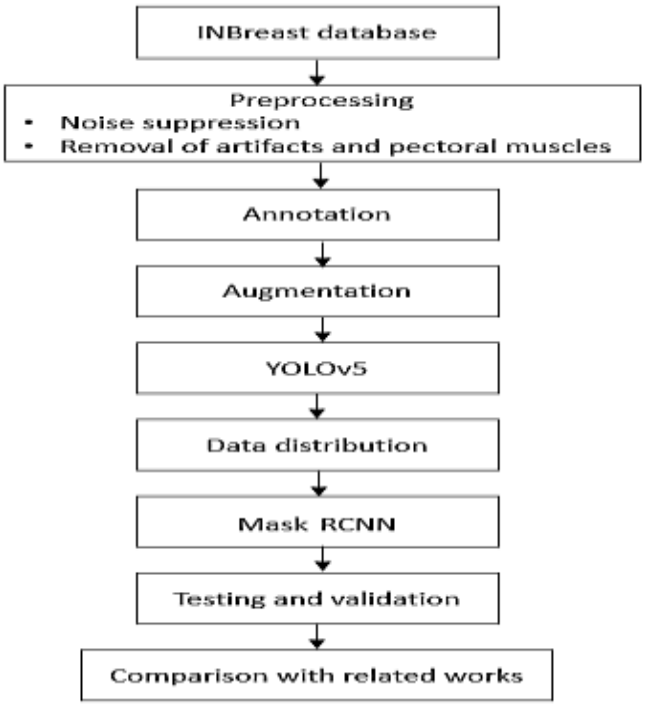
III. METHODOLOGY

The objective of this work is early detection of breast mass abnormality in a mammogram and classify it as malignant or benign without human involvement. The overall structure of the proposed methodology is illustrated in Fig. 1.

Mammograms are preprocessed thus removing noise and enhanced via CLACHE. The artefacts and pectoral muscles are removed for better results. Then the images are annotated introducing BI-RADS. To increase the number of images for model training, an augmentation technique was employed. The data was then fed to the state of art model YOLOv5 (You Only Look Once) and Mask RCNN for detection and classification. Semantic segmentation was employed in the Mask RCNN to characterize the prominent aspects of the tumours in the mammographic scans. The original model of YOLOv5 was modified to reduce the computational parameters. The improved version was next trained on the same dataset during experimentation. Reducing the complexity of the model and due to lightweight attributes, the model shows a better performance than the original version. The results of both models are compared and analyzed. As the study aims at predicting and classifying benign and malignant tumours furthermore also focuses to reduce the FPR and FNR to ensure authentic results. The size of the tumour is also determined.. The results are tested

and validated. To evaluate the efficiency of the proposed model, a comparative study with the previous related works is performed proving that the proposed methodology outperforms the results in measures of MCC[28]. The proposed methodology provides a practical approach for the radiologists to diagnose the tumour, its class, size and ultimately the suspected stage of cancer. The details of the methodology are discussed in the upcoming sections.

FIGURE 1: Methodology Employed



A. DATABASE

One of the public INbreast, accessible with ground truth annotations was incorporated in experimentation. The dataset contains 410 Full Field Digital Mammograms (FFDM) acquired from screening, diagnosis, or follow-up cases of 115 patients [29]. During experimentation, patient ID was unidentified for the privacy of the patient. The dataset has three classes: normal, malignant and benign. It incorporates perspectives on both Medio-Lateral Oblique (MLO) and Craniocaudal (CC) views. Identification of the tumour in the breast as malignant or benign is the core process of diagnosis on the mammograms. The malignant tumours are characterized by having rough boundaries, irregular margins, large size, and appearing dense and whiter than other tissue encompassing it. The benign tumours on the other hand are well-circumscribed with smooth boundaries having a round or oval shape as shown in Fig. 2.

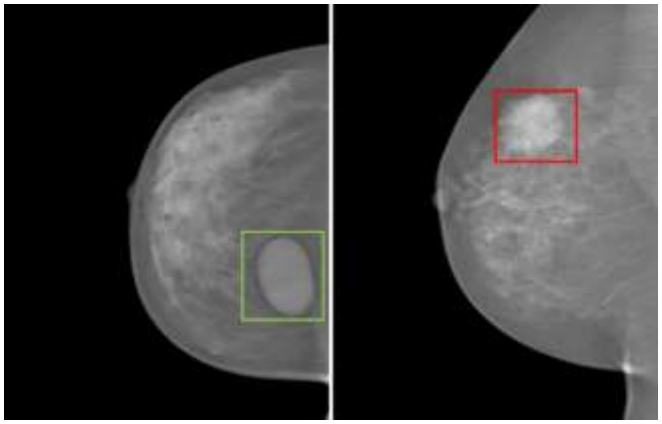


FIGURE 2: (a) Benign Tumor (b) Malignant Tumor

B. DATA PREPROCESSING

To enhance the image quality by removing the unwanted distortions and enhancing desired features so that it can be evaluated in a better way. The images are preprocessed using the following steps:

1. Denoising: For the denoising step, the median and mean filters are employed Initially, the median filter is utilized to suppress the salt & pepper noise and preserve the edges present in mammograms. The equation for a median filter is given by:

$$\text{Img}o(xi, yi) = med\{\text{Img}i(xi - \hat{j}, yi - k)j, k\epsilon T\} \quad (1)$$

where Img*o* (*xi*,*yi*) and Img*i* (*xi*,*yi*) are the output and input operated image represents the 2-D mask of size n x n and ^j and ^k denote pixels in the image. Then the mean filter is applied to remove the artificial contours generated by the median filter in some cases.

2. Removal of Pectoral muscles: Pectoral muscles can be seen in the mediolateral oblique view (MLO) of the mammographic images. They occur like the breast parenchyma in the mammographic images as both have the same intensity values as shown in Fig. 3.

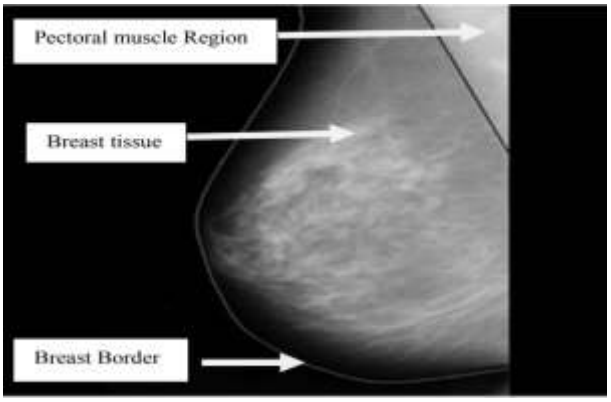


FIGURE 3: pectoral muscle and breast tissue

This may lead to false detection of the tumour site. Therefore, removing artefacts and pectorals muscles from MLO views becomes important for the correct identification of the lesion [30]. The results before and after the preprocessing of mammographic images are shown in Fig. 4.

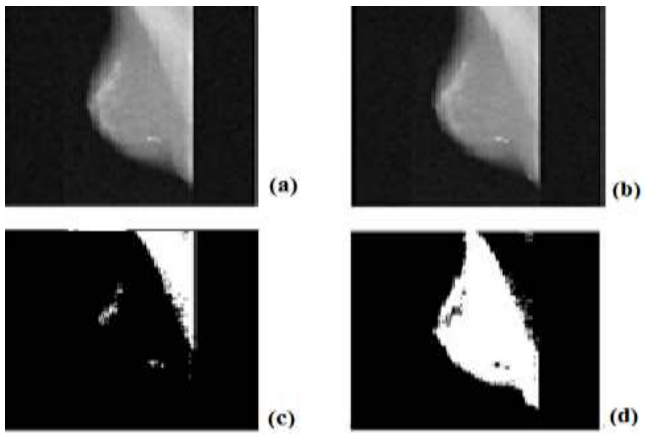


FIGURE 4: (a) original image (b) denoised image (c) identified pectoral region using global threshold (d) removal pectoral muscle using a global and grey level threshold.

C. ANNOTATION

An annotation is a process of labelling. Mammographic images were annotated incorporating make sense an online tool. Data is annotated for both algorithms. The Mask RCNN semantic segmentation is employed. It is employed to differentiate between the two classes of tumours. The benign tumours have smooth and well-circumscribed edges whereas the malignant tumours have irregular and spiky edges showing the aggressive nature of the tumour. Two files are created for the annotations of the data. The image file contains the mammographic images and the text file(.txt) contains bounding box dimensions f the suspected region having mass. The image file for annotation is shown in Fig. 5.

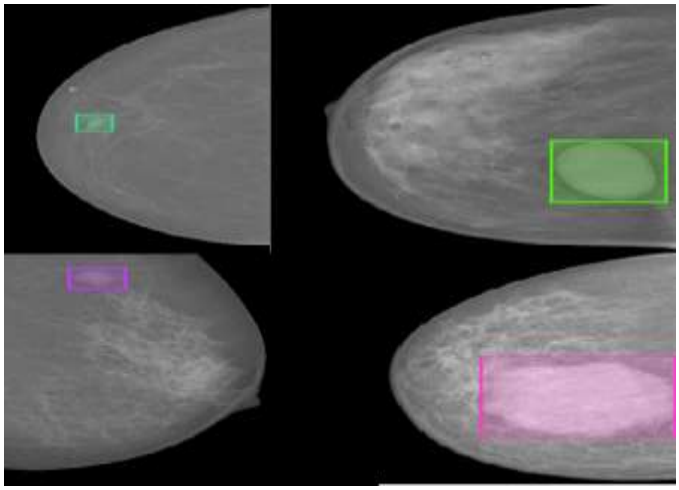


FIGURE 5: Annotated image

The histopathological evidence for the lesions is available in the XML file for each case. They are categorized as BI-Rads score, the standard system employed by radiologists to describe cancer. Out of 107 cases, 41 masses are assigned to BI-RAD 2 and 3 while the other 75 are categorized as malignant assigned to BI-RAD 4,5, and 6.

D. DATA AUGMENTATION

Data augmentation is a powerful technique of modifying the existing data by artificially creating variations in the images to create a larger dataset and generalized training model. It enables training the model to learn the features more deliberately for better differentiation of the objects. In this study, 107 mammograms having lesions are selected for training and evaluation of the model. Some images had more than one mass in different regions of a breast. Hence, a total of 116 masses are obtained. The number of images is increased using data augmentation for better training the model as shown in Fig. 6.

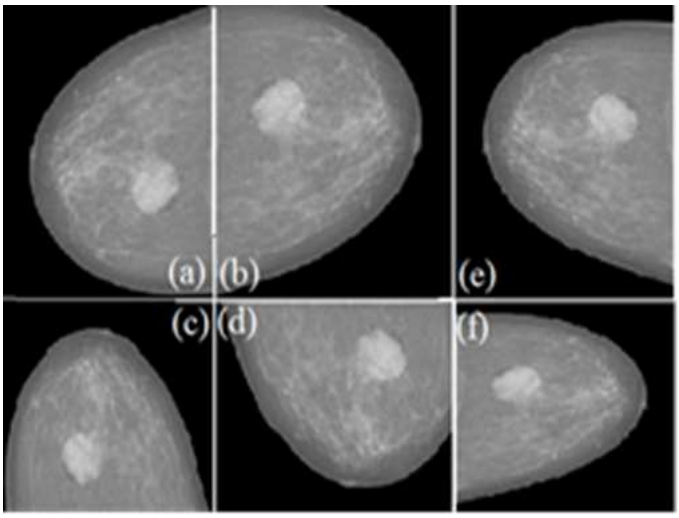


FIGURE 6: (a) Original image (b) 90̊ left rotation (c) 90̊ right rotation (d) flip and 90̊ rotation (e) vertical flip (f) horizontal flip

E. YOLOv5 ARCHITECTURE

YOLOv5 a deep learning convolutional neural network that performs object detection and classification. In this study, it is employed for the detection of benign and malignant tumours in mammographic images. It has also been employed for the detection of melanoma and melanocytic naevus, a condition of skin cancer [31]. It has fast speed as it employs an end-to-end method instead of a pipeline [32]. Moreover, it uses global features and exhibits good results in unexpected and new inputs [33]. This makes YOLO a good candidate to be employed in this work. The mass detection task comprises of identifying the location of the abnormality using a bounding box bordering the mass on the input image and then categorizing those lesions as benign or malignant.

F. YOLOv5 ARCHITECTURE

YOLO architecture consists of three major blocks [34]. The Backbone, Neck, and Head, as shown in figure7.

1. YOLOv5 Backbone: BottleneckCSP is the model backbone of YOLOv5 designed for feature extraction.

- The primary layer of the backbone increases the training speed due to fewer computations. It employs the slicing methods splitting the trio image each into four slices of 3x320x320.
- The second layer of the backbone concatenates the four segments to the output feature map of 12x320x320 and then with 32x320x320 using a 32-convolution kernel layer. Using the Batch Normalization function (BN) the results are fed to the subsequent layers.
- The third layer of the backbone BottleneckCSP combines the convolution layer of the size of 1x1(Conv2d layer, Batch Normalization and ReLu) with another convolution layer of size 3x3. Both the findings are then summed up giving the output of Bottleneck as follows:

$$A_1 = B_1 \times A_0, \quad (2)$$
$$A_{2=}B_2 \times [A_0, A_1], \quad (3)$$
$$A_{k=}B_k \times [A_0, A_1, \dots, A_k - A_1]. \quad (4)$$

where $[A_0, A_1 \dots .]$ shows the concatenation, B_k are the weights and A_k is the output of the kth layer. Hence the size of the output feature map turns out like the size of input.

2. YOLOv5 Neck: PANet is employed by the neck of the model to generate feature pyramids to perform buildup on the features making it faster and simpler [35]. It aids the model to work well on unknown data.
3. YOLOv5 Head: The head of the model is designed for final detection. It has three layers. The size of each layer is 80x80, 40x40, and 20x20. The images of different sizes are detected using anchor boxes with class probabilities, objectless scores and bounding boxes.

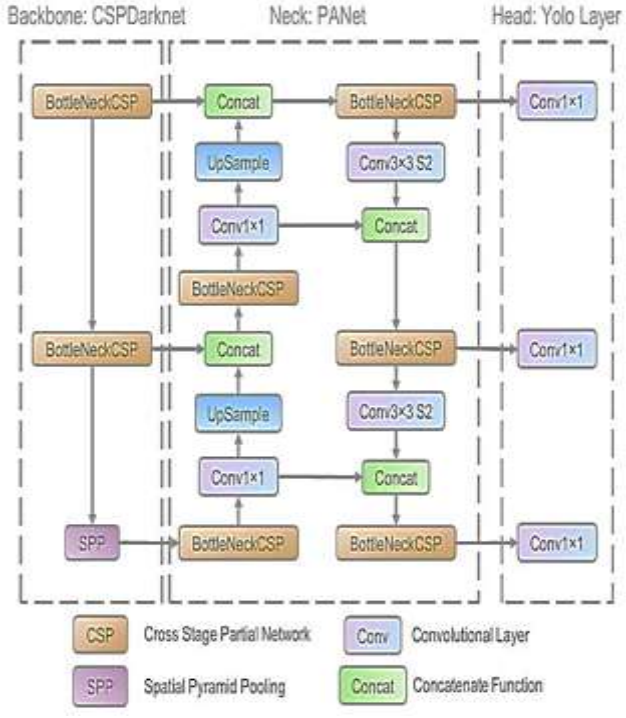


FIGURE 7: YOLOv5 Architecture

G. MODIFICATION AND IMPROVEMENTION IN YOLOv5

The YOLOv5 model is fast and efficient, however when it comes to images where tumors and the breast parenchyma is challenging to differentiate, the model limits in prediction and classificationthe lesions correctly. The basic model having four BottleneckCSP module is complex with numerous convolutional layers. It may extract features precisely however, due to increase in the number of parameters, the size of the model also increases, which make it hard for hardware implementation. Hence the basic model is improved in this study. To reduce parameters and model size, input feature map of BottleneckCSP module was directly linked to the output feature map. The convolutional layer on this branch was removed. Figure 8 shows the original and modified version of YOLOv5 backbone module.The modification of Bottleneck CSP modules reduces parameters however also reduces the efficiency to extract deep features. The loss is calculated using Equation 5.

$$C = P(tumor) \times loss_p \quad (5)$$

where C is the confidence, P(tumour) symbolizes the likelihood of the presence of tumor. When the midpoint of the tumour lies in the cell grid, P (tumour) has value of 1 else it is 0, $loss_p$ calculates the difference between predicting bounding box and ground truth box. The IOU is defined as the measure of the degree of overlap between two detection frames and is given by (6),

$$\text{IOU} = \text{area}(B_P \cap B_{gt}) / \text{area}(B_P \cup B_{gt}) \quad (6)$$

where Bp is the predicted frame and Bgt is the ground truth.

H. MASK RCNN

Mask-RCNN is developed on the working of Faster-RCNN with predicting segmentation masks on each Region of Interest (RoI) [36], uses pixel to pixel approximation to detect the shape of the object.

It is a region-based convolutional neural network the output of which is bounding boxes for each object and its class label with a confidence score. Here mask RCNN works in two stages.

Stage1-the proposal: The first stage consists of two networks, backbone (Inception, ResNet, etc.) and region proposal network (RPN). The backbone algorithm creates the region of interest (ROI) and RPN generates proposals based on where the object (tumour) is in the region. Anchor boxes are employed to detect the objects. It runs one time for each image to give a set of region proposals. Region proposals are regions in the feature map which have the object.

Stage2-the prediction: RCNN detection is employed two times to predict the bounding boxes and object class for each proposed

region. The size of the region can be made fixed by using RoI Align technique or RoI pool.

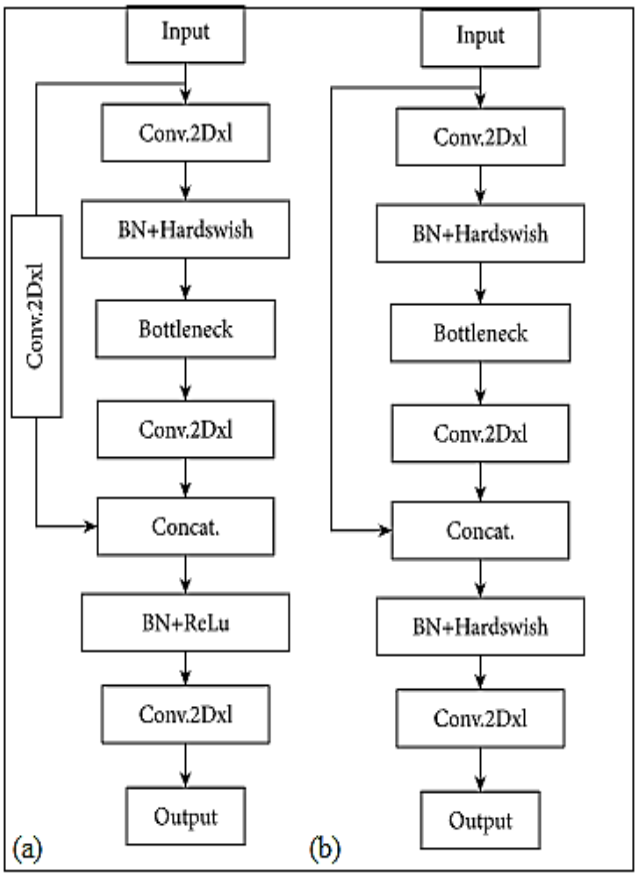


Figure 8: (a) Original Bottleneck CSP model (b) Modified BottleneckCSP model

I. THE TUMOR SIZE PREDICTION

As mentioned earlier, early diagnosis is the key to preventing the disease. Initially, the tumour is very small and cannot be detected by the naked eye. Eventually, it starts growing, invading the skin parenchyma acquiring a spiky feature. As shown in Table II, the tumours greater than or equal to 20mm and less than 50mm have higher chances of treatment while the lesion greater than 50mm moves towards the metastatic stage of cancer. Once the tumours are detected and classified as benign or malignant, they measured according to the breast cancer staging devised by the American Joint committee on breast cancer staging [10]. This was performed by training the model on mask RCNN.

Firstly, Mask RCNN is employed for detecting all the tumours to be measured and segmenting each tumour in the mammogram. The dataset has two classes, so semantic segmentation is employed as the best option.

TABLE II: American Joint Committee on breast cancer staging

Stage	Primary tumor	Nodes
Stage IA	≤20mm	None
Stage IB	≤20mm	Nodal metastasis
Stage IIA	≤20mm	N1/ None
Stage IIB	>20mm ≤50mm >50mm	N1/None
Stage IIIA	≤50mm >50mm	N2/N1 or N2
Stage IIIB	Extension to chest wall/skin	N0-N2
Stage IIIC	Any size	N3
Stage IV	Any size	Any involvement

Then the minimum bounding box of each mass is taken edge contour extraction on all the regions of the mass. Finally, the size of the bounding box is calculated to determine the size of the tumour. This method is effective even for highly irregular. shaped tumours. size is predicted using the python code.

IV. EXPERIMENTATION

Out of a dataset of 2120 mammogram images, 60% are employed for training the model, 30% for validating and 10% for testing purposes. The images are labelled using makesense.ai, a web-based markup tool, including two classes, benign and malignant, and a bounding box for each. The coordinate of the lesion's bounding box is labelled by an expert in reading mammographic images. Figures 9 (a) and 9(b) show the example of input images and output images.

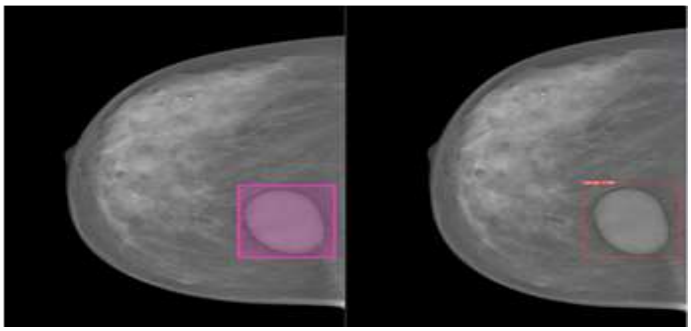


Figure 9 (a): Benign input and output image

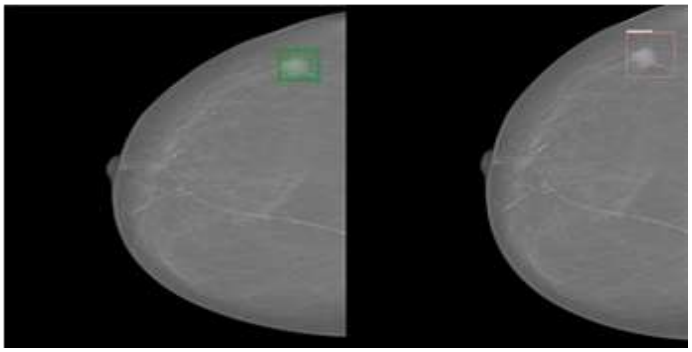


FIGURE 9 (b): Malignant input and output image

The pixel values of the mammograms are 14-bit contrast (0 to 214) representing 16,384 different values so the images are normalized from 0–255. The ground truth annotations and coordinates of the lesion extracted from the XML file are normalized comparative to the height and width of the image to be from [0 to 1] for better experimentation. The parameters are saved to be employed for input images during the training of the model. The mammograms are converted into different sizes to fit into the model. The work uses 440x448, 640x648, and 32x832 resolutions for model training. To evaluate the performance of the modified version, first, the original model is trained, then the improved version is trained on the same dataset. Then the results of both the models are compared. In this study, YOLOv5 model is trained with the following experimental setup and parameters. The experimentational setup parameters are listed in Table III.

TABLE III: Experimental setup

GPU	COLAB
Python	3.6
Libraries	CUDA, open CV
Operating System	Windows 10
Deep Learning Framwork	YOLO v5+ MASK RCNN
Optimization	Stochastic gradient descent (SGD) Rectified Linear Unit (ReLU)

V. RESULTS

Evaluation metrics is the measure of evaluating the performance of the algorithm in terms of accuracy [37], recall or sensitivity, F1-score [38], MCC [39], specificity, precision, ROC curve and log loss [40]. MCC is measured using TP (true positive), FP(false positive), TN(true negative) and FN(false negative). They are defined as, True positive (TP): the number of true predictions same as Ground Truth. False-positive (FP): the condition when the ground truth class belongs to false, but the predicted class is true. False-negative (FN): the condition when the ground truth is true, and the model prediction shows false. Similarly FPR (false positive rate): FNR (false negative rate): the results show the absence of the tumor as per equatoipn listed in (7) to (11).

$$\text{sensitivity} = \text{TP}/(\text{TP} + \text{FN}) \quad (7)$$

$$\text{specificity} = \text{TN}/(\text{TP} + \text{FN}) \quad (8)$$

$$\text{precision} = \text{TP}/(\text{TP} + \text{FN}) \quad (9)$$

$$\text{accuracy} = (\text{TN} + \text{TP})/(\text{FP} + \text{TN} + \text{FN} + \text{TP}) \quad (10)$$

$$\text{MCC} = \big((\text{TP})(\text{TN}) - (\text{FP})(\text{FN})\big)\Big/ \sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})} \quad (11)$$

Matthews's correlation (MCC) is the most reliable performance parameter in binary classification. It has the highest value only when all the four parameters of the confusion matrix have good prediction performance. The other vital parameter is the mean average precision of the model. It is the measure of the evaluation of the training of the system.

VI. PRACTICAL RESULTS

The model was well trained, it efficiently detects and classifies the breast mass as shown in Fig. 10.

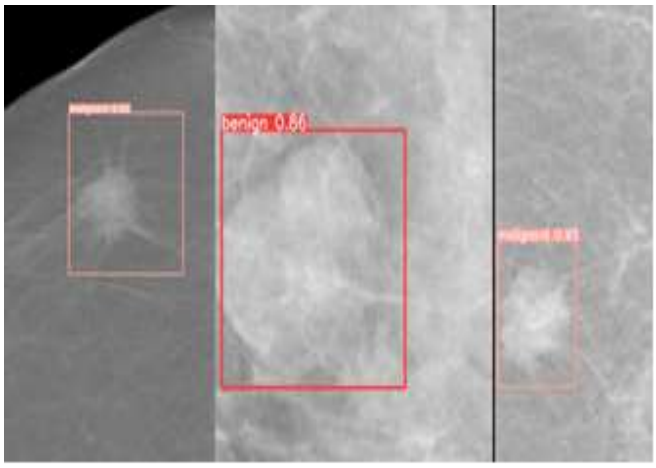


FIGURE 10: Breast Masses detection and classification

The size of the tumours is also well predicted. The practical results of three malignant cases are shown in figure 11 show the size of the breast mass. The size of the tumours determines the size of the tumour. Figure 11 shows a malignant tumour of 20mm, this means it is in its initial stage. Figure 11 show two tumours of size 40mm and 39mm. the third stage of cancer.

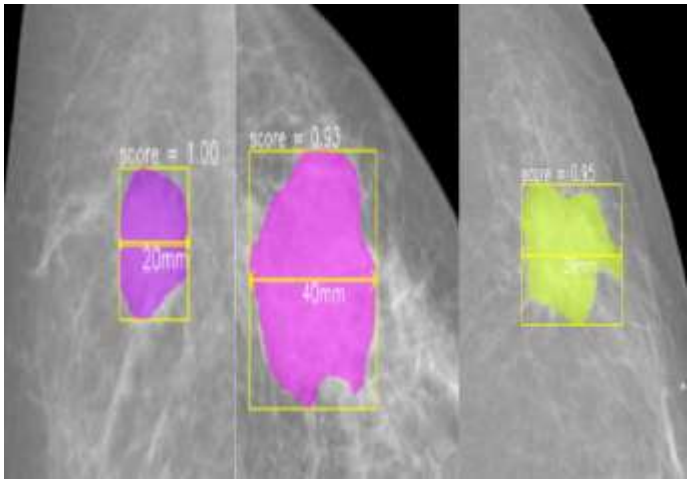
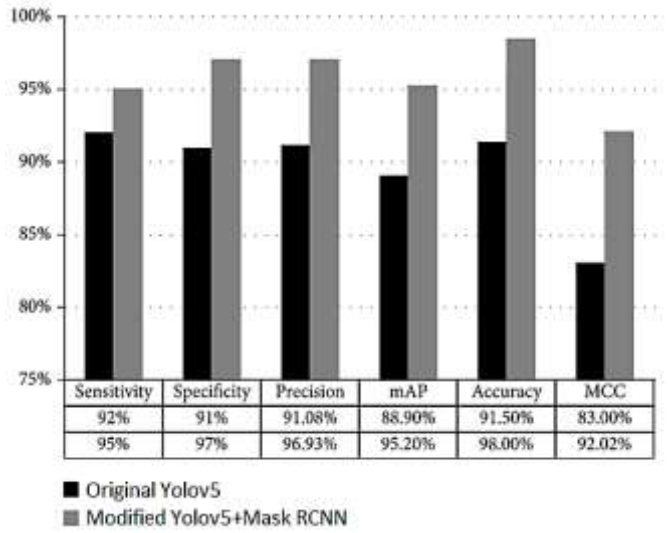


FIGURE 11: The predicted tumour size.

The comaprision results in Fig. 12 show that using modified YOLOv5 combined with Mask RCNN improves the MCC value from 83% to 92%. It also lowers the FPRvalue. The FPR value for original and modified versions is 0.06 and 0.05 respectively, whereas the FNR value for both is 0.03. The accuracy also increases from 91.50% to 98%.



Sensitivity	Specificity	Precision	mAP	Accuracy	MCC
92%	91%	91.08%	88.90%	91.50%	83.00%
95%	97%	96.93%	95.20%	98.00%	92.02%

FIGURE 12: A Comparision

VII. CONCLUSION

The goal of this work was the development of a breast mass detection and classification of and subsequently, determine the aggression of the cancerous tumours employing a deep learning approach. INbreast dataset was introduced and data preprocessing was incorporated prior to the model training. The proposed methodology contains two steps, first is a modification of the detection algorithm. To reduce the computational complexities and size of the model the original version was modified. The second step was to determine the aggression of the tumour from images. YOLOv5+Mask RCNN detects the location of the tumour deploying anchor boxes from the mammographic images to predict the class of the mass. The breast masses detected are either malignant or benign, with defined tumour size. The results exhibit an improvement in accuracy and MCC value. It is evident from the results that the two-step novel method (YOLOv5 +Mask RCNN) is better and more efficient compared to a single detection technique. The results were cross-validated with clinician consultation thus assisting radiologists in establishing their final decisions efficiently.

In future, it would be interesting to work on a combination of other models with YOLOv5 to achieve high MCC value.

REFERENCES

[1] "GLOBOCAN 2020: New Global Cancer Data | UICC." https://www.uicc.org/news/globocan-2020-new-global-cancer-data (accessed Sep. 2021).

[2] D. A. Ragab, M. Sharkas, S. Marshall, and J. Ren, "Breast cancer detection using deep convolutional neural networks and support vector machines," *PeerJ*, vol. 2019, no. 1, pp. 1–23, 2019,

[3] S. M. Lima, R. D. Kehm, and M. B. Terry, "Global breast cancer incidence and mortality trends by region, age-groups, and fertility patterns," *The Lancet*, vol. 38, no. 100685, 2021.

[4] S. Zaheer, N. Shah, S. A. Maqbool, and N. M. Soomro, "Estimates of past and future time trends in age-specific breast cancer incidence among women in Karachi, Pakistan: 2004-2025," *BMC Public Health*, vol. 19, no. 1, Jul. 2019.

[5] M. Arif, M. Javed, and H. Raza, "Letter to the Editor Breast Cancer in Pakistan: Alarming Situation of Breast Cancer in Near Future Dear Editor-in-Chief," 2020.

[6] N. H. Khan, S. F. Duan, D. D. Wu, and X. Y. Ji, "Better reporting and awareness campaigns needed for breast cancer in Pakistani women," *Cancer Manag. Res.*, vol. 13, pp. 2125–2129, 2021.

[7] "Limitations of Mammograms," *Am. CancerSociety*, pp. 2–3, 2020.

[8] A.-K. V. West *et al.*, "Division induced dynamics in non-Invasive and invasive breast cancer," *Biophys. J.*, vol. 112, no. 123a, 2017.

[9] K. Shaikh, S. Krishnan, and D. R. Thank, *Artificial Intelligence in Breast Cancer Early Detection and Diagnosis*. Springer, 2021.

[10] W. Anna *et al.*, "Validation study of the American joint committee on cancer eighth edition prognostic stage compared with the anatomic stage in breast cancer," *JAMA Oncol.*, vol. 4, no. 2, pp. 203–209, 2018.

[11] D. Chicco, N. Tötsch, and G. Jurman, "The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation," *BioData Min.*, vol. 14, no. 1, p. 13, 2021.

[12] J. Bailar, "Mammography: a contrary view. Annals of Internal Medicine, vol. 84, no. 1, pp. 77–84.

[13] M. Aaqib, M. Tufail, and S. Anwar, "A Novel Deep Learning Based Approach for Breast Cancer Detection," *nternational Conf. Math. Actuar. Sci. Comput. Sci. Stat.*, vol. 13, 2019.

[14] T. Brito-Sarracino, M. R. Santos, E. F. Antunes‡, I. B. A. Santos, J. C. Kasmanas, and Andr´e C.F. Carvalho, "Explainable Machine Learning for Breast Cancer Diagnosis," in *2019 8th Brazilian Conference on Intelligent Systems (BRACIS)*, 2019, pp. 681–686.

[15] D. Q. Zeebaree, H. Haron, A. M. Abdulazeez, and D. A. Zebari, "Machine learning and Region Growing for Breast Cancer Segmentation," 2019.

[16] Priyanka and K. Sanjeev, "A Review Paper on Breast Cancer Detection Using Deep Learning," 2020.

[17] R. Preetha and S. V. Jinny, "Early diagnose breast cancer with PCA-LDA based FER and neuro-fuzzy classification system," *J. Ambient Intell. Humaniz. Comput.*, 2020.

[18] A. Binder *et al.*, "Morphological and molecular breast cancer profiling through explainable machine learning," *Nat. Mach. Intell.*, vol. 3, pp. 355–366, 2021.

[19] J. Li *et al.*, "Predicting breast cancer 5-year survival using machine learning: A systematic review," *journal.pone.0250370*, 2021.

[20] G. Meenalochini and S. Ramkumar, "Survey of machine learning algorithms for breast cancer detection using mammogram images," *Mater. Today Proc.*, vol. 37, no. Part 2, pp. 2738–2743, Jan. 2021.

[21] Y. Khourdifi and M. Bahaj, "Applying Best Machine Learning Algorithms for Breast Cancer Prediction and Classification," in *IEEE*, 2018, pp. 1–5.

[22] S. Abbas *et al.*, "BCD-WERT: a novel approach for breast cancer detection using whale optimization based efficient features and extremely randomized tree algorithm," *PeerJ Comput. Sci.*, 2021.

[23] M. Shawarib, A. A. Latif, B.-D. Al-Zatmah, and S. S. Abu-Naser, "Breast Cancer Diagnosis and Survival Prediction Using JNN," *Int. J. Eng. Inf. Syst.*, vol. 4, pp. 23–30, 2020.

[24] H. B. Lim, N. T. Nhung, and E.-P. Li, "Confocal Microwave Imaging for Breast Cancer Detection: Delay-Multiply-and-Sum Image Reconstruction Algorithm," *IEEE Trans. Biomed. Eng.*, vol. 55(6), pp. 1697–1704.

[25] A. Lambrou, H. Papadopoulos, and A. Gammerman, "Reliable Confidence Measures for Medical Diagnosis With Evolutionary Algorithms," *IEEE Trans. Inf. Technol. Biomed.*, vol. 15, 2011.

[26] A. G. Hamed, M. Marey, S. A. El-Sayed, and M. F. Tolba, "YOLO Based Breast Masses Detection and Classification in Full-Field Digital Mammograms," *Comput. Methods Programs Biomed.*, 2020.

[27] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary

classification evaluation," *BMC Genomics*, vol. 21, no. 1, p. 6, 2020.
[28] T. FF, T. YJ, and S. KS., "Convolutional neural network improvement for breast cancer classification," *Expert Syst. with Appl*, vol. 120, pp. 103-115., 2019.
[29] M. MJ, and C. JS, "INbreast: toward a full-field digital mammographic database.," *Natl. Libr. Med. Publ. by Elsevier*, 2012.
[30] M. Mehmood *et al*., "Machine Learning Enabled Early Detection of Breast Cancer by Structural Analysis of Mammograms," *Computers, Materials \& Continua* , vol. 67, no. 1. 2021.
[31] Ü. HM and A. E, "Skin Lesion Segmentation in Dermoscopic Images with Combination of YOLO and GrabCut Algorithm.," *MDPI*, 2019.
[32] J. Redmon, A. Farhadi, and IEEE, "YOLO9000: Better, Faster, Stronger," *30th Ieee Conference On Computer Vision And Pattern Recognition (CVPR 2017)*, 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Univ Washington, Seattle, WA 98195 USA, pp. 6517–6525, 2017.
[33] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, and IEEE, "You Only Look Once: Unified, Real-Time Object Detection," *2016 IEEE Conference On Computer Vision And Pattern Recognition (Cvpr)*, no. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Univ Washington, Seattle, WA 98195 USA, pp. 779–788, 2016.
[34] M. Kasper-Eulaers, N. Hahn, S. Berger, T. Sebulonsen, O. Myrland, and P. E. Kummervold, "Short Communication: Detecting Heavy Goods Vehicles in Rest Areas in Winter Conditions Using YOLOv5," *Algorithms*, vol. 14, no. 4, 2021.
[35] A. Kuznetsova, T. Maleva, and V. Soloviev, "Detecting Apples in Orchards Using YOLOv3 and YOLOv5 in General and Close-Up Images BT - Advances in Neural Networks – ISNN 2020," 2020, pp. 233–243.
[36] M. Namjoshi and K. Khurana, "A Mask-RCNN based object detection and captioning framework for industrial videos," *Int. J. Adv. Technol. Eng. Explor.*, vol. 8, Nov. 2021.
[37] R. Tkachenko, I. Izonin, N. Kryvinska, I. Dronyuk, and K. Zub, "An Approach towards Increasing Prediction Accuracy for the Recovery of Missing IoT Data based on the GRNN-SGTM Ensemble," *Sensors*, vol. 20, 2020.
[38] I. Izonin, R. Tkachenko, N. Kryvinska, P. Tkachenko, and M. Greguš, "Multiple Linear Regression Based on Coefficients Identification Using Non-iterative SGTM Neural-like Structure," 2019, pp. 467–479.
[39] D. Chicco, V. Starovoitov, and G. Jurman, "The Benefits of the Matthews Correlation Coefficient (MCC) Over the Diagnostic Odds Ratio (DOR) in Binary Classification Assessment," *IEEE Access*, vol. 9, pp. 47112–47124, 2021.
[40] P. Henderson and V. Ferrari, "End-to-End Training of Object Class Detectors for Mean Average Precision," in *13th Asian Conference on Computer Vision*, 2016, pp. 198–213.