ISSN: 0976-9102 (ONLINE)
DOI: 10.21917/ijivp.2025.0502

IMPROVED ATTENTION-BASED MBCONVBLOCK-EFFICIENTDET NETWORK BASED CUCKOO SEARCH ALGORITHM FOR OSTEOSARCOMA NODULE DETECTION ENHANCEMENT

A. Nandhini and M. Sengaliappan
Department of Computer Applications, Nehru College of Management, India

Abstract

Osteosarcoma is a malignant bone tumor that is extremely dangerous to human health. Not only does it require a large amount of work, it is also a complicated task to outline the lesion area in an image manually, using traditional methods. With the development of computer-aided diagnostic techniques, more and more researchers are focusing on automatic segmentation techniques for osteosarcoma analysis. However, existing methods ignore the size of osteosarcomas, making it difficult to identify and segment smaller tumors. In this research work, initially Pre-processing with Chebyshev Filter Combined with Kalman Filtering (HF) approach is done to remove the noise and enhance the image for further processing. Then the preprocessing images undergone Segmentation using Multi-tier Otsu Thresholding (MOT) algorithm. After the process of segmentation, the wavelet and GLCM based feature extraction is executed. Furthermore, introduce a hybrid Attention-based MBConvBlock-EfficientDet (A-MBConvBlock-EfficientDet) model for classification. Specifically, the MBConvBlock is reconstructed to enable the exchange of information between the channels of the feature layer. The fully connected layer of the attention module is removed and convolution is used to cut down the amount of network parameters. Here the Cuckoo Search (CS) Algorithm is used to optimize the A-MBConvBlock-EfficientDet model, potentially enhancing its performance in identifying osteosarcoma lung nodules accurately. The proposed methodology is evaluated through experimental studies. These experiments validate the efficiency of the system in achieving precise osteosarcoma LND. Metrics like accuracy, sensitivity, specificity, and F1 score may be used to assess the performance of the suggested method against existing approaches.

Keywords:

Osteosarcoma, Lung Nodule (LN), Cuckoo Search (CS) Algorithm, Attention based EfficientDet and Adaptive Kalman Filter

1. INTRODUCTION

Osteosarcoma (OS) is the most common primary malignant bone tumour in children and adolescents [1]. OS is characterised by rapid hematogenous spread, with the lung being the most common site [2]. According to the literature, 10–25% of OS patients present with detectable metastases at the time of initial diagnosis, of which 85–90% have lung metastases [3]. Tumour cells in OS metastases produce bone and this potential may be apparent on imaging. Moreover, the recurrence of OS is predominantly located in the lung (~ 80% of cases) [1]. For the assessment of pulmonary metastases, chest computed tomography (CT) has remained for years the gold imaging standard, as also suggested by the Children's Oncology Group (COG) [4]. With the continuous advances in multi-row-detector computed tomography (MDCT) scanners, the sensitivity in detecting small lung nodules has improved. Nonetheless, the distinction between malignant and benign pulmonary lesions on CT scans in paediatric patients with sarcomas is still below expectations [5], and the correct classification of pulmonary nodules on imaging as either malignant or benign is a clinical dilemma even for radiologists experienced in the field.

The presence of metastases has a significant impact on survival in OS patients [6]. Therefore, early diagnosis and appropriate treatment is an interdisciplinary challenge for the entire team involved, including the oncologist, surgeon, radiologist and pathologist. All OS metastases must be resected completely, regardless of their number and site, if the patient is treated with curative intent [6]. Computer-aided diagnosis (CAD) [7]-[8] is an advanced technology to help radiologists and oncologists to read and identify the nodules from the CT-scanned image of the patient. CAD can be categorized into image processing and machine learning approaches. An image processing approach used the static mathematic model to enhance image, feature extraction, and mathematically categorize the results while a machine learning approach uses a dynamic mathematic model that can be trained by using humans' intelligence and experiences to mimic human decisions for specific situations. Convolutional Neural Networks or CNNs [9], [10] are one of the most popular machine learning techniques, especially for image-related applications.

The object detection CNNs were widely used in many applications such as faces or human detection. the RegionCNNs (R-CNNs) [11], You only look once (Yolo) [12] and Single Shot Detection (SSD) [13] were the most popular frameworks for object detection network because they could pinpoint the location of the multiple interested objects within an image and classify these found objects with confidences. R-CNNs found the interesting objects by searching the image using different size bounding boxes and found the possibility of class that can represent each search box identity. Because R-CNNs had to search the objects with all possible boxes, the calculation time was very large and impractical for low calculation power computer such as a normal computer in the government hospital. Convolutional Neural Networks (CNNs) were used in another research's method, which demonstrated the ability to effectively gather data on a mean of 92% of clinically important coronary artery segments [14]. Thus, this research set out to determine how well the AI diagnostic DCNN framework performed in assessing pulmonary nodules in patients with osteosarcoma who were adolescents or young adults. A comparison of its clinical value against a different manual approach was a further objective of the current investigation.

Medical professionals have to work a lot of hours for evaluating lung CT images visually. AI-assisted diagnosis technology has been increasingly applied in clinical fields in the past few years. AI imaging technology can increase doctors' productivity and has a high accuracy rate, according to several research [15]. As the number of precautionary and early detection methods rises as predicted, researchers are developing automated

solutions that minimize the workload of physicians, boost the accuracy of diagnostics by lowering subjectivity, expedite analysis, and lower clinical expenses. DL has the benefit of enabling end-to-end detection in CAD programs by teaching it to recognize the most important features throughout training. As a result, the network can withstand fluctuations since it may detect the characteristics of nodules in different CT images with different parameters [16]. Better outcomes are made possible by the framework's innate ability for invariant features learning from malignant nodules due to a training set rich in variability. Because no features are designed, the network can use the given ground-truth to independently understand the relationship among features and cancer. In this research work, introduce an A-MBConvBlock-EfficientDet model for classification which enhancing its performance in identifying osteosarcoma lung nodules accurately.

The research is prearranged as follows: The related work section 2 reviews relevant literature, highlighting the uniqueness of the proposed hybrid approach. The methodology section 3 describes in detail the integration of the proposed algorithm for nodule detection. The outcomes and analysis section 4 offers the findings, comparing the optimized EfficientDet model's performance with existing methods. The conclusion summarizes the key findings and emphasizes the potential clinical applications in section 5.

2. RELATED WORKS

Several researchers have proposed and implemented various approaches for detecting lung cancer using machine learning and image processing techniques.

Monica Ramakrishnan et al. [17] developed a technique to detect the Lung Cancer Nodule using CT images in 2022, offering a new method of implementing CNN using the pre trained VGG model for feature extraction and RNN for feature classification to identify pulmonary nodules in lung cancer detection. Machine learning, data mining, and image processing methods are used in this study to predict lung cancer nodules in high-risk patients. Using a publicly available data set of lung CT images, a model for lung cancer nodule detection was developed based on the research and analysis conducted for this work. By combining image processing and classification techniques, an end-to-end method for detecting lung cancer nodules with 70% accuracy was developed in this study. Because of the large amount of data included in each patient's CT scan, processing the images takes a long time.

Aggarwal et al. [18] proposed a model in 2015 that differentiates between nodules and the structure of normal lung anatomy. Grey levels, statistics, and geometry are used to extract properties. The best thresholding and segmentation classifier is LDA. The system has 84% accuracy and 53.33% specificity. This methodology does detect the cancer nodule with an high accuracy, but itis still inadequate.

In 2015, Roy.T.et al. [19] developed a method for detecting lung cancer nodules that employs an active contour model and a fuzzy interference approach. To improve visual contrast, this method employs a grey transformation. Before segmentation, an image is binarized, and the resulting image is segmented using an active contour model. The fuzzy inference method is used to classify cancer. To train the classifier, features such as area, entropy, mean, correlation, main axis length, and minor axis length are extracted. The system's overall accuracy is 94.12%. This method has the disadvantage of being unable to distinguish between benign and malignant tumours.

An elaborated review of various techniques that constitute a conventional CAD system has been tabulated by Bhavanishankar et al. and Li et al. [20, 21]. Three different approaches for detecting the pulmonary nodules in CT scans were implemented by Camarlinghi et al. [22] resulting in an overall sensitivity of 80%. Kuruvilla et al. [23] proposed a method which used six statistical features extracted from the segmented slices of two lungs and feed-forward back-propagation neural networks to classify the cancerous nodules resulting in sensitivity ranging from 82% to 91.4%. Torres et al. [24] proposed a CAD system based on Channeler Ant model for segmentation and a feed-forward neural network for classification of nodules, and this approach resulted in a sensitivity of about 80%.

Devinder et al. [25] proposed a methodology wherein deep features are extracted using a two-layered network called autoencoder and the nodules are classified as benign or malignant. This approach resulted in an average sensitivity of 83.25%. Fakoor et al. [26] proposed an unsupervised and deep learning approach (stacked autoencoder) to classify and diagnose different types of cancers including lung cancer. Gruetzemacher et al. [27] used a deep learning approach for the classification of pulmonary nodules which resulted in a sensitivity of 78.2% with five convolution layers. Ginneken et al. [28] proposed a Convolutional Neural Network based approach to detect the pulmonary nodules in CT scans which generated the sensitivity of 78%. Most of the methods in the literature tested their algorithms on data sets obtained from LIDC-IRDI repository.

Although several researchers have contributed significantly towards reduction of false positives and achieving better classifications, they lack to deliver better accuracy with large data sets. This paper addresses this issue.

3. PROPOSED METHODOLOGY

The proposed methodology diagram is shown in Fig.1. In this research work, Pre-processing with Chebyshev Filter Combined with Kalman Filtering approach is done to remove the noise. Then the preprocessing images undergone Segmentation using Multi-tier Otsu Thresholding (MOT) algorithm. Finally, a hybrid Attention-based MBConvBlock-EfficientDet (A-MBConvBlock-EfficientDet) model is proposed for Classification and osteosarcoma lung nodule detection.

3.1 DATASET DESCRIPTION

The dataset described comprises CT-scanned PNG format images derived from DICOM files obtained from Siemens Somatom Definition 64 machines at Lerdsin Hospital in Thailand [29]. The Human Research Ethics Committee of Thammasat University (Science) (HREC-TUSc) and the RECof Lerdsin Hospital, Department of Medical Services, Ministry of Public Health in Thailand gave their approval for these photos to be used for study. Images from 202 patients at Lerdsin Hospital who were given an osteosarcoma diagnosis are included in the dataset. After performing lung scans, the DICOM files were converted into CT-scanned PNG format images for examination. Out of the 269,025

image files that were taken from the patients' DICOMs, 2,212 files show that there are anomalous nodules.

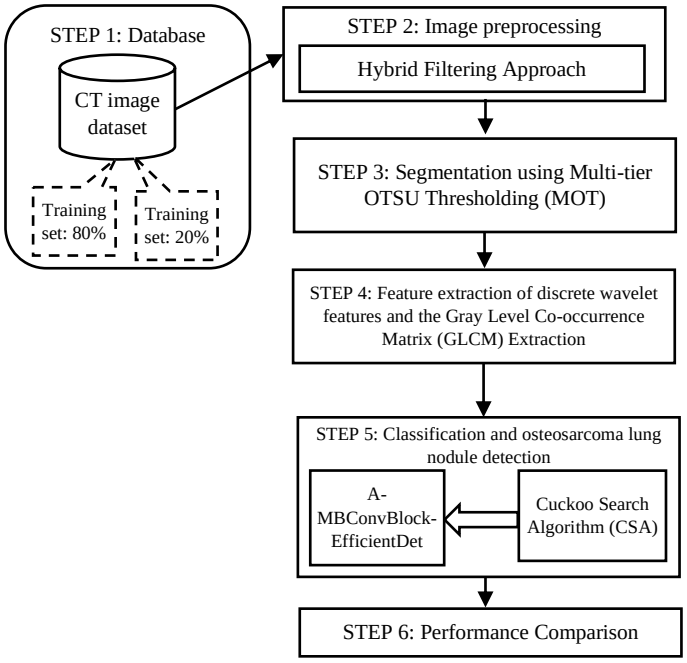

Fig.1. Architecture Diagram of Proposed Osteosarcoma Lung Nodule Detection Model

3.2 PREPROCESSING USING HYBRID FILTERING APPROACH

The input images are undergone preprocessing using hybrid filtering approach. A Hybrid Chebyshev Filter typically refers to the combination of the Chebyshev filter with In this approach, the static nature of the Chebyshev filter is combined with adaptive filtering method called Kalman filter to create a system that can adapt to changing signal characteristics over time. This can be useful in applications where the signal's noise characteristics are not constant.

Chebyshev filters are a kind of filter normally employed in (IP) Image Processing and signal processing for tasks such as noise reduction [30]. They are designed to minimize the ripple in the passband or stopband of a filter's frequency response while still achieving a sharp cutoff. In the context of image processing, Chebyshev filters can be used for noise reduction by selectively filtering out noise while preserving important image features. The general equation for a Chebyshev filter's frequency response is:

$$H(\omega) = \frac{1}{\sqrt{1+\partial^2 T_n^2\left(\frac{\omega}{\omega_c}\right)}} \quad (1)$$

where, the frequency filter response can be denoted as $H(\omega)$, ϵ is the ripple factor, controlling the amount of allowed ripple in the passband, $T_n(x)$ is the nth order Chebyshev polynomial, ω is the angular frequency and ω_c is the cutoff angular frequency. For noise reduction in images, can apply Chebyshev filter in the frequency domain using the Eq.(1).

The Kalman filter operates in the time domain to estimate the current state of a system by fusing noisy measurements with predicted states. The Kalman filter assumes a linear dynamic system and updates the state estimates based on the process and measurement models. The Kalman filter consists of two steps: prediction and update.

The state Prediction is defined as

$$\hat{x}_{k|k-1} = F_k \hat{x}_{k-1|k-1} + B_k u_k \quad (2)$$

where the $\hat{x}_{k|k-1}$ is a prediction state vector at time k, F_k is a state transition matrix, B_k is a control input matrix and u_k is a control input at time k.

Next Error covariance prediction is defined as

$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_k \quad (3)$$

where the $P_{k|k-1}$ is a predicted covariance matrix and Q_k is the process noise covariance matrix.

Then the update step is done by kalman filter is explained below. The kalman gain is defined as

$$K_k = P_{k|k-1} H_k^T (H_k P_{k|k-1} H_k^T + R_k)^{-1} \quad (4)$$

where, K_k is a kalman gain, H_k refers to measurement matrix and R_k is a measurement noise covariance matrix. The state update is defined as:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k (z_k - H_k \hat{x}_{k|k-1}) \quad (5)$$

where, z_k is a measurement at time k and $H_k \hat{x}_{k|k-1}$ is a predicted measurement. The error covariance update is defined as

$$P_{k|k} = (I - K_k H_k) P_{k|k-1} \quad (6)$$

3.3 COMBINING CHEBYSHEV AND KALMAN FILTERS

Step 1: Preprocess with Chebyshev Filter: The signal is first passed through the Chebyshev filter to remove unwanted frequency components (e.g., high-frequency noise). This stage helps clean the signal for further processing, making it more suitable for state estimation.

$$y_{\text{filtered}}(t) = H_{\text{Chebyshev}}(s) \cdot y(t) \quad (7)$$

where $y_{\text{filtered}}(t)$ is the signal after Chebyshev filtering.

Step 2: Kalman Filter for State Estimation: The filtered signal is then passed to the Kalman filter to estimate the state of the system. The Kalman filter uses the cleaned signal as its observation and applies its prediction-update mechanism to estimate the current and future states of the system.

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k (y_{\text{filtered}}(t) - H_k \hat{x}_{k|k-1}) \quad (8)$$

The Chebyshev-Kalman hybrid filter combines the strengths of both filters, handling frequency-domain noise removal through the Chebyshev filter and real-time state estimation via the Kalman filter.

3.4 SEGMENTATION USING MULTI-TIER OTSU THRESHOLDING (MOT)

In the segmentation step of this study, Multi- tier Otsu Thresholding (MOT) technique is used to segment the dataset based on thresholding. Segmentation is a crucial process that extracts the Osteosarcoma diseased portions by considering the

intensity of pixels. This allows for the extraction of essential features from the input images for further analysis.

Thresholding is a technique that converts a gray-level image into a binary image based on the intensity levels of pixels and specific threshold values. Selecting optimal threshold values is a significant challenge in image segmentation. The traditional Otsu method is an effective segmentation approach that provides better results for ordinary images when there is a clear distinction between foreground and background objects.

However, in cases where there is an overlap of boundaries in the images and the gray value distribution of pixels cannot be easily differentiated, multiple-level thresholding is utilized for segmenting the images. In the Otsu approach, the image is divided into two regions: object and background. The optimal thresholding aims to enhance the contrast between these two regions for both bi-level and multiple-level segmentation.

To achieve the optimal threshold level in gray-scale images, the Otsu function is employed. The Otsu function calculates the threshold value that maximizes the contrast between the object and background regions, leading to more accurate segmentation results. By utilizing the MOT technique, the segmentation process enables the identification and extraction of Osteosarcoma regions from the images, which is crucial for subsequent analysis and classification of different Osteosarcoma.

In multi-level thresholding, there requires a t threshold level $(l_1, l_2, ..., l_t)$ to split the image into t classes. Therefore, the class c_0 having gray levels in the range 0 to $l_1 - 1$, c_1 having gray levels in the range l_1 to l_2-1, and c_t with enclosed levels from l_t to N-1 are considered. Therefore, the process of segmentation is expressed by an objective function as follows in Eq.(9)

$$F(l)_{\max} = \sigma_1 + \sigma_2 + ... + \sigma_t \tag{9}$$

From the above equation, the values for σ is expressed in the following Eq.(10):

$$\begin{aligned} \sigma_0 &= \omega_0 (\eta_0 - \eta_l)^2, \\ \sigma_1 &= \omega_1 (\eta_1 - \eta_l)^2, ..., \\ \sigma_t &= \omega_t (\eta_t - \eta_l)^2 \end{aligned} \tag{10}$$

Thus, the segmented output is given as the input for feature extraction

3.5 FEATURE EXTRACTION OF DISCRETE WAVELET FEATURES AND THE GRAY LEVEL CO-OCCURRENCE MATRIX (GLCM) FEATURES

After the segmentation process, the resulting segmented output is then used as input for the DWT (Discrete Wavelet Transform) and GLCM (Gray Level Co-occurrence Matrix) are both widely used methods for feature extraction, especially in image processing applications. These methods help in extracting important characteristics from images for purposes of classification. Texture feature extraction can be described as a statistical technique that discloses specific properties about the spatial distribution of gray levels in image texture, considering the spatial connection of pixels. The tumor and normal regions have specific textures and spectral information that differentiate normal and abnormal tissue growth. The DWT decomposes an image into sub-bands at different resolutions, effectively capturing both spatial and frequency information. This multi-resolution approach is useful for analyzing the localized changes in frequency and texture in an image. The Steps in DWT Feature Extraction:

3.5.1 Decomposition:

The image is decomposed into four sub-bands at each level:

- **LL (Low-Low)**: Low-frequency approximation image.
- **LH (Low-High)**: Horizontal detail information (edge details).
- **HL (High-Low)**: Vertical detail information.
- **HH (High-High)**: Diagonal detail information.

3.5.2 Multi-Level Decomposition:

The LL sub-band is further decomposed to capture more details. This can be done recursively, providing more levels of resolution.

3.5.3 Feature Extraction:

From each sub-band (LL, LH, HL, HH), statistical features are extracted, such as:

3.5.4 Mean:

The average pixel intensity.

$$\text{Mean} = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} I(i, j) \tag{11}$$

where $I(i,j)$ is the pixel value at position (i,j), and M,N are the dimensions of the image.

3.5.5 Variance:

The spread of intensity values.

3.5.6 Energy:

The sum of squared values.

$$\text{Energy} = \sum_{i=1}^{M} \sum_{j=1}^{N} I(i, j)^2 \tag{12}$$

3.5.7 Entropy:

A measure of randomness in the sub-band.

$$\text{Entropy} = -\sum_{i=1}^{M} \sum_{j=1}^{N} P(i, j) \log_2 (P(i, j)) \tag{13}$$

where $P(i,j)$ is the probability of a pixel intensity occurring.

For feature extraction, GLCM was employed, which is a widely used statistical technique developed [31] for processing of remote sensing data. In the first step, the original image was converted to the grayscale. The next step was to extract spatial features from the gray-scale images based on the relationship of brightness values to the center pixel with its neighborhood defined by a kernel or window size. The relationship of the brightness values was represented in the form of a matrix. The matrix was made up of the frequent occurrence of the sequential pair of the pixel values along with a defined direction. The relationship helps GLCM to generate a different set of texture information based on gray-scale, kernel size, and direction. In [31] defined fourteen

textural features, which provide redundant spatial context information which was an overhead in classification.

GLCM is a statistical method that considers the spatial relationship between pixels in an image. It generates a matrix that counts how often a pixel with a particular intensity is adjacent to a pixel with a different intensity. From this matrix, various texture features are calculated.

Define the spatial relationship between pixel pairs (e.g., adjacent horizontally, vertically, or diagonally). Compute the co-occurrence matrix by counting how often pairs of pixels with specific values (gray levels) occur. Normalization: Normalize the GLCM by dividing each element by the total number of occurrences. This gives a probability matrix. Extract Texture Features: Calculate texture features from the GLCM, such as:

- Mean: It represents the average intensity value of the plaque tissue.
- Standard deviation: It measures the variation or spread of intensity values within the plaque tissue.
- Skewness: It quantifies the asymmetry of the intensity distribution. Positive skewness indicates a longer tail on the right side, while negative skewness indicates a longer tail on the left side.
- Kurtosis: It measures the peakedness or flatness of the intensity distribution. Higher kurtosis values indicate a sharper peak and heavier tails.
- Contrast: It measures the local variations in intensity values, reflecting the presence of edges or boundaries within the plaque tissue.
- Energy: It represents the sum of squared intensity values, indicating the overall magnitude or strength of the plaque tissue texture.
- Entropy: It quantifies the randomness or uncertainty of the intensity distribution within the plaque tissue.
- Homogeneity: It measures the similarity or uniformity of neighboring intensity values, indicating the smoothness of the plaque tissue texture.
- Correlation: It captures the linear dependency between pairs of intensity values within the plaque tissue.
- Inverse different moments: They represent the inverse of the differences between intensity values and their mean, providing information about the distribution of intensity variations.

Then, these extracted features are given in the classification phase. By combining DWT and GLCM, can effectively capture both frequency (via DWT) and texture (via GLCM) characteristics of an image. The process typically involves:

- Apply DWT to the image to decompose it into different frequency sub-bands (LL, LH, HL, HH).
- Calculate GLCM on the relevant sub-bands, typically LL or other sub-bands depending on the application.
- Extract features from both DWT and GLCM for each sub-band.
- Combine the features into a single feature vector for classification or analysis tasks. This information is highly useful to identify whether the segmented regions belong in the nodule or non- nodule detection category.

3.6 CLASSIFICATION AND OSTEOSARCOMA NODULE DETECTION USING AN OPTIMIZED ATTENTION-BASED MBCONVBLOCK-EFFICIENTDET MODEL

In this section, the features extracted from the images are given as a input to the optimized Attention-based MBConvBlock-EfficientDet model. Here the CNN contains convolutional computation and has a deep structure. It includes convolutional layers, pooling layers and fully connected layers. The connections between the convolutional layers are called sparse connection which is used to reduce the connections between the network layers and the amounts of parameters and make the operation easy and efficient. The nature of weight sharing in CNN improves the stability and generalization ability of the network structure, avoids overfitting and enhances the learning effect [32].

EfficientDet is one of the most advanced object detection algorithms, which has a simple structure and excellent performance. It is available in seven versions from D0 to D6. And the resolution, depth and width of the model can be scaled simultaneously by it according to resource constraints to meet the detection requirements under different conditions. By using the EfficientNet as the backbone network, BiFPN as the feature network, and the shared class/box prediction network, EfficientDet balances the speed and accuracy in the object detection task well.

Enhancing the depth of the neural network, adding the width of the feature layer, and increasing the resolution of the input image can improve the detection accuracy of the network [33], but also lead to more network parameters and higher computational costs. To improve the detection efficiency, balancing the dimensions of network width, depth and resolution is crucial during CNN scaling. EfficientNet combines these three characteristics and puts forward a new model scaling method that uses an efficient composite coefficient to simultaneously adjust the depth, width, and resolution of the net- work. Grounded on the neural structure search technology [35], the optimal composite coefficient can be obtained. As shown in equations.

$$\text{Depth: } d = \alpha^{\phi} \tag{14}$$

$$\text{Width } w = \beta^{\phi} \tag{15}$$

$$\text{Resolution: } r = \gamma^{\phi} \tag{16}$$

$$\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2 \tag{17}$$

where α, β, γ represent the weight of depth, width and resolution, respectively, and are constants that can be determined by small network search, and ϕ is a user-specified factor that controls the number of resources used for model scaling. Under this constraint, α=1.2, β=1.1 and γ=1.15 are obtained. When ϕ=0, an optimal base model EfficientNet-B0 is obtained; when the ϕ is increased, it is equivalent to expanding the three dimensions of the base model at the same time, and the model becomes larger, the performance also improves, and the resource consumption also becomes larger. The structure of EfficientNet is shown in Fig.2.

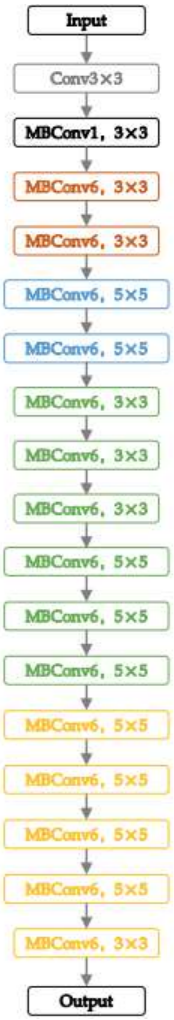

Fig.2. EfficientNet is composed of 16 large Blocks stacked, 16 large Blocks can be divided into 1, 2, 2, 3, 3, 4, 1 Blocks. 3 × 3/5 × 5 represent the convolutional kernel size respectively. The different Blocks are distinguished by color and convolutional kernel size.

The general structure of Block is shown in Fig.3, and its design idea is Inverted Residuals, using 1 × 1 convolution to up-dimension before the 3 × 3 or 5 × 5 DWConv, adding a channel attention mechanism after the DWConv, and finally adding a large residual edge after using 1 × 1 convolution to down-dimension. BN is the BatchNorm.

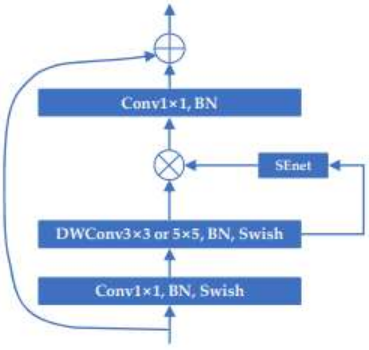

Fig.3. The structure of MBConvBlock

The Fig.3 shows the structure of MBConvBlock. SENet is an attention mechanism. BN is the BatchNorm, which serves to normalize the data. Swish is an activation function. DWConv is the Depthwise convolution. The top feature map has rich semantic information but low resolution, while the bottom feature map has low-level semantic information but higher resolution. Multi-scale feature fusion is the aggregation of features with different resolution semantic information, so that the network has the ability to detect features of different scales. FPN, NAS-FPN, PANet, etc., [34, 35] have been widely used in multi-scale feature fusion. However, their direct combinations of the feature maps in different layers and ignores the contribution of different resolution features to the output features. The authors proposed BiFPN, which is simple and efficient. As shown in the formula, the importance of different input features to feature fusion is expressed by calculating the weights of different layers.

There are 3 methods to calculate the weights assigned to the features. “Unbounded fusion” is the name of the initial one, and it is stated as

$$O = \sum_i w_i I_i \tag{18}$$

where, the w_i, for each feature, channel, or pixel, is a learnable weight that might be a scalar, vector, or multidimensional tensor. The fact that the scalar weight in this technique is unbounded is a drawback that could make training more difficult. In the second, a “Softmax fusion” is created as given below:

$$O = \sum_i \frac{e^{w_i}}{\sum_j e^{w_j}} I_i \tag{19}$$

Although this approach has a considerably larger latency cost, it ensures feature contribution preservation and bounds the weight within [0, 1]. A “fast normalized fusion” is provided in the third instance as.

$$O = \sum_i \frac{w_i}{\varepsilon + \sum_j w_j} I_i \tag{20}$$

Here, the input features can be denoted as I_i, ε is a tiny positive constant that is set at 0.0001 for numerical stability, and $w_i \geq 0$ is verified for each w_i using a ReLU function. Because it guarantees a bounded weight and a quicker response, this method is preferred. The EfficientDet architecture (Fig.2) uses a 1-stage detection approach. The foundation is assumed as EfficientDet. Subsequently, BiFPN is added as a feature network, obtaining features from the backbone’s 3rd through 7th tiers. A class and bounding box prediction network receive the fused output. Incorporating attention mechanisms into the EfficientDet model for osteosarcoma lung nodule detection offers compelling advantages. Firstly, attention mechanisms enable selective focus, allowing the model to prioritize relevant features in CT scans crucial for nodule identification. This selective attention also aids in noise reduction, filtering out irrelevant information to improve detection accuracy. Additionally, attention mechanisms provide the model with contextual understanding, crucial for distinguishing nodules amidst complex anatomical structures. Moreover, the interpretability offered by attention mechanisms fosters trust in the model’s predictions, facilitating collaboration with medical experts. Overall, integrating attention mechanisms enhances the EfficientDet model’s performance in osteosarcoma

nodule detection, making it a promising approach for automated medical image analysis.

3.7 CUCKOO SEARCH (CS) ALGORITHM TO OPTIMIZE ATTENTION-BASED MBCONVBLOCK-EFFICIENTDET MODEL

Cuckoo Search (CS) is a nature-inspired optimization algorithm that mimics the brood parasitic behavior of certain species of cuckoo birds. It's a metaheuristic optimization technique commonly used for solving complex optimization problems. In the context of EfficientDet, which is a neural network architecture for object detection, cuckoo search optimization can be used to optimize the hyperparameters, architecture, or specific components of the model (e.g., feature fusion, attention mechanisms, learning rates).

The CSO algorithm follows the next three idealized rules:

- One egg is laid by each cuckoo at a time, and it deposits its egg in a randomly selected nest;
- The best nests with the highest-quality eggs will be passed down to the following generations.
- The host bird finds the cuckoo's egg with a probability $p_a \in [0, 1]$, and the counts of host nests that are accessible is fixed. The host bird in this situation has two options: it can discard the egg or leave the nest and construct a brand-new one.

This likelihood represents the impact of each generation's replacement of cuckoo eggs, eggs found by the host bird, with fresh eggs. Keep in mind that an egg symbolizes a solution. As a selection procedure for the optimization algorithm, these presumptions guarantee the best solutions will endure from generation to generation. Therefore, the objective of the CSO algorithm is to swap out the low-quality solutions in the nests with higher-quality ones. For cuckoos, a novel solution X_i^{t+1} is provided by:

$$X_i^{t+1} = X_i^t + \alpha \otimes \text{Levy}(\lambda) \tag{21}$$

$$\alpha = \alpha_0 \otimes (X_j^t - X_i^t) \tag{22}$$

where α represents step sizes (α > 0) with dimensions equal problems dimensions; products $\otimes$ represent entry-wise multiplications; X_i^t represents randomly selected solutions; and *Lévy*(λ) imply Lévy flights random walks. The value of α_0 is set to 0.01 as suggested in order to improve the search efficiency. One type of random walk where the step length is derived from the Lévy distribution is called a Lévy flight. A sequence of instantaneous jumps produced by a probability density function with a power law tail provided by characterizes this distribution:

$$\text{Levy}(s, \lambda) = \frac{\lambda \Gamma(\lambda) \sin(\pi\lambda/2)}{\pi s^{1+\lambda}} \tag{23}$$

The step length *S* of Lévy flights is drawn from a uniform distribution that obeys *Lévy* distribution. Furthermore, the algorithm used a balanced combination of a local random walk and the global explorative random walk, controlled by a switching parameter p_a. The local random walk can be written as

$$X_i^{(t+1)} = X_i^t + \alpha s \otimes H(p_a - \varepsilon) \otimes (X_j^t - X_k^t) \tag{24}$$

where X_j^t and X_k^t signify two different solutions selected randomly using permutations, *H* represents Heaviside functions, ε represents random numbers drawn from uniform distributions, and *s* represents step sizes.

On the other hand, the global random walk is carried out by using Levy flights:

$$X_i^{(t+1)} = X_i^t + \alpha \oplus L'evy(s, \lambda) \tag{25}$$

where, $\alpha > 0$ represents step size scaling factors; Levy(s, λ) represents step-lengths distributed based on probability distributions of Eq.(25) which have infinite variances with infinite means:

$$L'evy(s, \lambda) = \frac{\lambda \Gamma(\lambda) \sin(\pi\lambda/2)}{\pi} \frac{1}{s^{(1+\lambda)}} \tag{26}$$

Algorithm: CSO

Input: Initialize the population of CS and maximum iterations.

Output: Global Best Solution

Step 1: Fitness-function $f(x)=(x_1,x_2,x_3,\ldots,x_d)^T$

Step 2: Initializing *n* nests, where $x_i (i \le n)$.

Step 3: While(t<Max-generation-value) do

Step 4: Levy flight use to update x_i

Step 5: Get its Fitness F_i

Step 6: Pick a nest x_j at random

Step 7: If ($F_i > F_j$) then

Step 8: Replace x_j by the new solution

Step 9: Rest worse nest use Eq.(21)

Step 10: Rank the nest (solution) and keep the best nest

Step 11: Final result output and presentation

Reassess the fitness metric to determine the optimal individual and overall solutions: Using the CSO Lévy flight rule, the search space's solutions are updated, and the top global solutions are identified. Once the maximum counts of iterations have been completed, the worldwide best solutions are announced.

Hence By applying CS for hyperparameter tuning, A-EfficientDet can be fine-tuned to achieve better performance and adaptability across different object detection tasks and datasets. The pseudocode outlines the main steps involved in hyperparameter tuning using CS for A-EfficientDet as shown below:

Algorithm: Hyperparameter tuning using CS for A-EfficientDet

Input: Initialize the hyperparameters of A-EfficientDet model, population of CS and maximum iterations.

Output: Hyperparameters of A-EfficientDet model

Step 1: Define Hyperparameters and Objective Function

hyperparameters = {'learning_rate': (0.001, 0.01), 'batch_size': (16, 64),

Define other hyperparameters and their ranges

- def objective function(hyperparameters): Train A-EfficientDet with given hyperparameters, evaluate

performance using validation dataset, Return performance metric (e.g., accuracy)

- return performance metric

Step 2: Initialization, population size = 10, max iterations = 100

bats = initialize virtual bats (population size, hyperparameters)

Step 3-7: Cuckoo Search Algorithm

- Initialize N nests (solutions) randomly with different hyperparameters
- Evaluate the fitness (loss or mAP) of each nest
- For each iteration:
- For each nest i:
- Generate a new solution (cuckoo's egg) using Lévy flight
- Evaluate the fitness of the new solution
- If the new solution is better than a randomly selected nest:
- Replace the randomly selected nest with the new solution
- Abandon a fraction of the worst nests and replace them with new random solutions
- Find and store the best-performing nest (solution)
- If convergence criteria are met, stop
- Return the best solution (optimized hyperparameters)

Step 8: Final Model Selection

- Best hyperparameters = best performing nest (solution)

Step 9: Evaluation on Test Set

- Final model performance = objective function (best hyperparameters)

End for

After optimization, finally, through the Softmax function, the model will output the prediction probabilities, which represent the probability value of the input rock image belonging to each type. The type corresponding to the maximum probability is the final classification result.

4. EXPERIMENTAL RESULTS AND DISCUSSION

In this research, preprocessing of the CT scan images was conducted to enhance their quality for deep learning analysis. This included image normalization, illumination and contrast correction, noise reduction, and background removal. As original CT images from hospitals may contain artifacts and undesirable elements such as low-intensity levels, noise, and textual information, preprocessing steps were essential to ensure better model learning, generalization, and robustness. Additionally, specific filters like high-pass filters were applied to extract relevant information, such as clear lung images.

To validate the accuracy of ground truths, a panel consisting of 1 radiologist and 2 oncology experts independently examined the CT images for abnormal nodules detection associated with Osteosarcoma metastatic disease. Consensus was reached if at least two out of the three experts agreed on the presence of nodules. These labeled CT images were then separated into training and validation sets, comprising 80% (1,769 images) and 20% (443 images) of the dataset, respectively. The dataset primarily focuses on the detection of Osteosarcoma nodules, ensuring the model's training and validation on relevant pathological features. In traditional medical screening methods, accuracy is typically calculated based on the number of patients rather than individual images. If a patient has at least one correctly detected abnormality (such as a nodule), the case is considered positive. However, in engineering-oriented approaches like the one described in this section, a different method is employed to provide deeper analysis focusing on both quantity and quality.

Instead of aggregating results at the patient level, the engineering standard method evaluates accuracy based on the number of images. This proposed optimized A-MBConvBlock-EfficientDet approach allows for a more detailed analysis of each individual image, considering every detail in the dataset. While using the number of images to calculate accuracy may result in a lower overall accuracy compared to patient-level aggregation, it provides insights into the execution of the procedure on a per-image basis. By adopting the number of image approach, the evaluation can capture finer details and nuances in the detection process, which may be missed when considering only patient-level outcomes. This approach enables a more granular assessment of the algorithm's performance, facilitating improvements in both detection quantity and quality across the entire dataset.

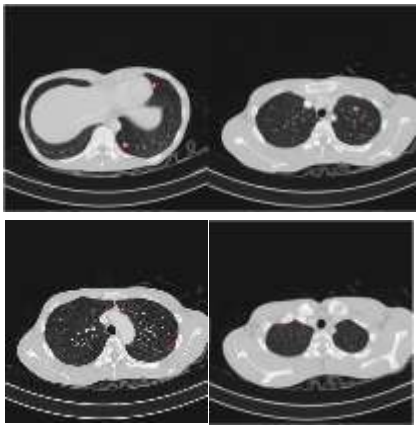

Fig.4. Example of the labelled CT-Scanned nodule images.

The proposed optimized A-EfficientDet is compared with exiting methods such as MSANet [15], TSCNN [16] and SSD-VGG16 [23] depends on performance matrix such as precision, recall, f-measure and accuracy. In the evaluation process, a corrected predicted nodule is considered a True Positive (TP) outcome. To compute performance scores, the True Negative (TN) value must also be obtained. Since the evaluated images did not originally contain non-nodule CT-scanned images, these non-nodule images were added to the evaluation set to obtain the False Positive (FP) value. These non-nodule CT-scanned images are referred to as "no-class" images. Given the absence of a typical technique for adding no-class images, one non-nodule image was employed in this study as one object in the image. The amount of TP discovered from the first evaluation of the photos was equal to the number of non-nodule photographs added during the review procedure. The common performance scores used in this evaluation are the F1-score and accuracy. The F1-score is calculated using the harmonic mean of Precision and Recall, which are calculated as follows:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{27}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{28}$$

The F1-score is then calculated as:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{29}$$

Additionally, accuracy is calculated using the formula:

$$\text{Accuracy}(\%) = \frac{TP + TN}{TP + FP + TN + FN} \times 100 \tag{30}$$

These performance metrics provide an inclusive assessment of the procedure's detection performance, considering both the precision and recall of nodule detection as well as the total accuracy of the framework.

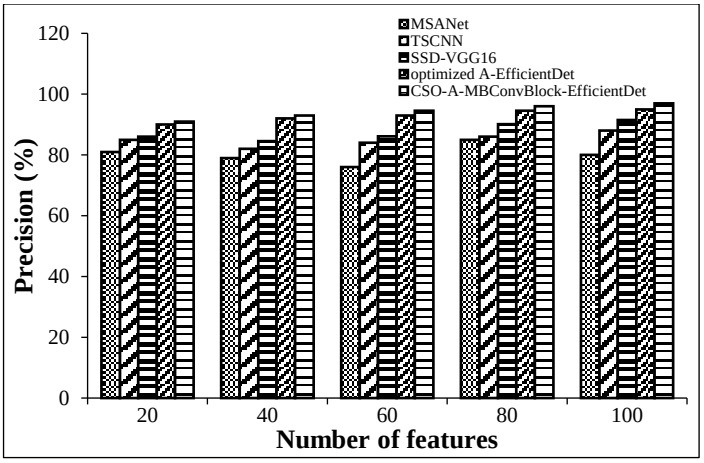

Fig.5. Precision performance comparison

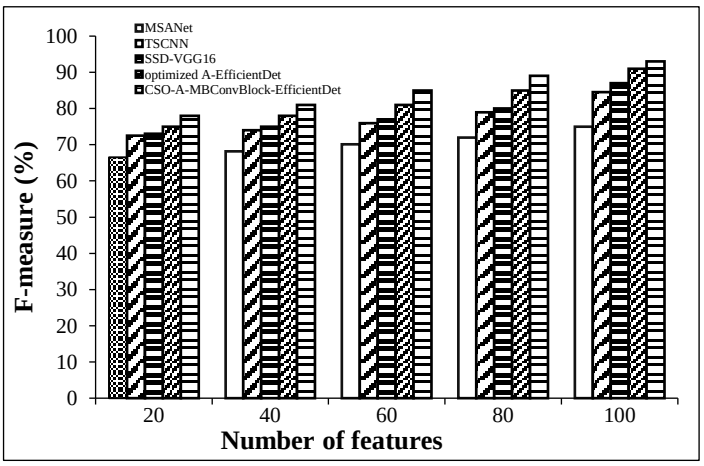

Fig.6. F-measure performance comparison

The Fig.5 presents the outcomes of a precision comparison among the proposed CSO-A-MBConvBlock-EfficientDet and the traditional MSANet, TSCNN, SSD-VGG16 and optimized A-EfficientDet classifiers. Comparing the suggested strategy to other current techniques, the graph indicates that it gives a high rate of accuracy. It is a highly effective technique that has a 97% precision percentage. In evaluating the accuracy of current methods, MSANet, TSCNN, SSD-VGG16 and optimized A-EfficientDet offer notable accuracy rates of 80%, 88.5%, 91.5% and 95%, in that order. The high precision attained by the proposed methodology can be attributed to meticulous preprocessing, including contrast correction and noise reduction, which enhances image clarity. Additionally, accurate labelling by multiple experts and systematic handling of false positives contribute to the reliability of nodule detection.

The Fig.6 presents the F-measure comparison outcomes of the SSD-VGG16, MSANet, TSCNN, and recommended CSO-A-MBConvBlock-EfficientDet classifiers. It is a highly effective technique that has a 93% of F-measure percentage. The suggested CSO-A-MBConvBlock-EfficientDet has an extremely good F-measure rate of 91%, based on the data MSANet, TSCNN, SSD-VGG16 and optimized A-EfficientDet yield lower rates of 73.15 %, 84.5 %, 87%, and 91% accordingly, as compared with the rate of the F-measure among the current techniques. This indicates that the recommended strategy can produce better attack detection outcomes than the prior strategies. The high F-measure is achieved due to the balanced optimization of both precision and recall, facilitated by accurate ground truth labelling and robust object detection algorithms, resulting in effective identification of nodules while minimizing false positives and false negatives.

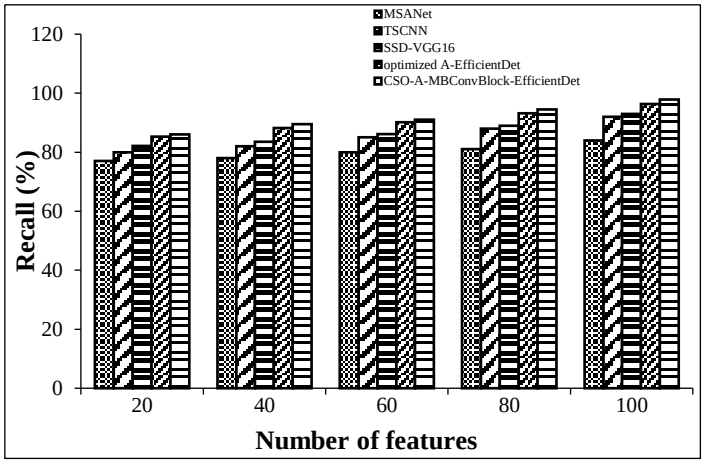

Fig.7. Recall performance comparison

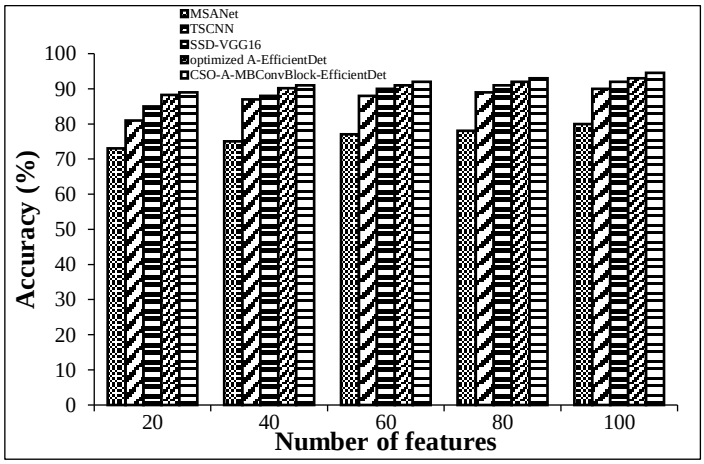

Fig.8. Accuracy performance comparison

The Fig.7 displays the recall comparison outcomes for the SSD-VGG16, TSCNN, MSANet, and recommended optimized A-EfficientDet classifiers. The 97.80% recall rate offered by the

suggested CSO-A-MBConvBlock-EfficientDet method is incredibly high. These results indicate that the CSO-A-MBConvBlock-EfficientDet that has been recommended has a high memory rate and an average attack detection rate. By contrasting the recall rates of the current methods, it can be seen that the recommended system can produce better attack detection outcomes than the earlier methods. MSANet, TSCNN, SSD-VGG16 and optimized A-EfficientDet yield recall rates of 84%, 92%, and 93%, accordingly. High recall is attained through meticulous preprocessing to enhance image quality, coupled with the utilization of advanced object detection networks trained on diverse datasets, ensuring comprehensive coverage of nodule variations.

The attack detection accuracy comparison can be observed in the graph in Fig.8 above. Techniques like CSO-A-MBConvBlock-EfficientDet multiclass classifiers, TSCNN, Optimized AE+DNN, and MSANet are applied. The optimized A-EfficientDet has a high accuracy rate of 94.5%, making it a great way to get precise predictions. In evaluating the accuracy of earlier methods like MSANet, TSCNN, and SSD-VGG16, the rates are 80%, 90%, and 92%, in that order. Higher accuracy is possible while resolving the local optima problem due to enhanced A-EfficientDet learning algorithms, which are comparatively robust to noise in training data. The suggested methodology achieves high accuracy through rigorous preprocessing techniques that enhance image quality, coupled with the utilization of advanced object detection networks trained on diverse datasets. Additionally, accurate ground truth labeling by multiple experts and systematic handling of false positives contribute to the reliability of nodule detection, resulting in overall superior performance.

5. CONCLUSION FUTURE WORK

Thus, the hybrid cuckoo search Algorithm with a hybrid Attention-based MBConvBlock-EfficientDet (A-MBConvBlock-EfficientDet) model for classification presents a promising approach for precise detection of osteosarcoma lung nodules. The Chebyshev Filter Combined with Kalman Filtering (HF) approach is act as a pre-processing pipeline and optimizing the Attention-based MBConvBlock-EfficientDet model with the CSO optimization algorithm, significant improvements in detection accuracy and performance are achieved. The experimental results demonstrate remarkable outcomes: a precision rate of 97%, F-measure rate of 93%, recall rate of 97.80%, and accuracy rate of 94.5% in osteosarcoma lung nodule detection. These findings underscore the effectiveness and reliability of the proposed methodology in enhancing MI (Medical Image) analysis and facilitating early detection of critical pathologies. For future work, further exploration and refinement of the proposed hybrid approach could focus on the following areas:

- Exploring the integration of multi-modal MI data, such as combining CT scans with MRI or PET scans, to deliver a more inclusive and accurate prediction and classification of osteosarcoma and other pulmonary pathologies.

Integration of feature selection approaches to select the best features for further classification improvement.

REFERENCES

[1] K.M. Ingley, A. Maleddu, F.L. Grange, C. Gerrand, A. Bleyer and E. Yasmin, “Current Approaches to Management of Bone Sarcoma in Adolescent and Young Adult Patients”, *Pediatric Blood and Cancer*, Vol. 69, No. 2, pp. 1-7, 2022.

[2] A.S. Mohammed and A. Neravetla, “An Analysis of Security Protocols for Cloud Computing Algorithms in Mobile Ad Hoc Networks”, *Proceedings of IEEE World Conference on Applied Intelligence and Computing*, pp. 1316-1321, 2024.

[3] A.M. Chiesa, P. Spinnato, M. Miceli and G. Facchini, Radiologic Assessment of Osteosarcoma Lung Metastases: State of the Art and Recent Advances”, *Cells*, Vol. 10, No. 3, pp. 1-7, 2021.

[4] D. Carrle and S. Bielack, “Osteosarcoma Lung Metastases Detection and Principles of Multimodal Therapy”, *Cancer Treatment and Research*, Vol. 152, pp. 165-184, 2009.

[5] K.M. Ghosh, L.H. Lee, T.B. Beckingsale, C.H. Gerrand and K.S. Rankin, “Indeterminate Nodules in Osteosarcoma: What’s the Follow-Up?”, *British Journal of Cancer*, Vol. 118, No. 5, pp. 634-638, 2018.

[6] P. Picci, “Osteosarcoma (osteogenic sarcoma)”, *Orphanet Journal of Rare Diseases*, Vol. 21, No. 2, pp. 1-6, 2007.

[7] K. Doi, ‘‘Computer-Aided Diagnosis in Medical Imaging: Historical Review, Current Status and Future Potential”, *Computerized Medical Imaging and Graphics*, Vol. 31, pp. 198-211, 2007.

[8] Y.J. Chen, K.L. Hua, C.H. Hsu, W.H. Cheng and S.C. Hidayati, ‘‘Computer-Aided Classification of Lung Nodules on Computed Tomography Images via Deep Learning Technique’’, *OncoTargets Therapy*, Vol. 8, pp. 1-6, 2015.

[9] T.Y.A. Naseer, A. Azhar, T. Shakeel and K. Zafar, ‘‘Computer-Aided Brain Tumor Diagnosis: Performance Evaluation of Deep Learner CNN using Augmented Brain MRI”, *International Journal of Biomedical Imaging*, Vol. 14, No. 2, pp. 1-7, 2021.

[10] Z. Lia, M. Dong, S. Wen, X. Hu, P. Zhou and Z. Zeng, ‘‘CLU-CNNs: Object Detection for Medical Images”, *Neurocomputing*, Vol. 350, pp. 53-59, 2019.

[11] S. Ren, K. He, R. Girshick and J. Sun, ‘‘Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks’’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, pp. 1137-1149, 2016.

[12] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, ‘‘You Only Look Once: Unified, Real-Time Object Detection’’, *Proceedings of International Conference on Computer Vision and Pattern Recognition*, Vol. 5, pp. 779-788, 2016.

[13] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.Y. Fu and A.C. Berg, ‘‘SSD: Single Shot Multibox Detector’’, *Proceedings of International Conference on Computer Vision*, pp. 21-37, 2016.

[14] J.M. Wolterink, R.W. van Hamersvelt, M.A. Viergever, T. Leiner and I. Isgum, “Coronary Artery Centerline Extraction in Cardiac CT Angiography using a CNN-based Orientation Classifier”, *Medical Image Analysis*, Vol. 51, pp. 46-60, 2019.

[15] Y.L. Ni, X.C. Zheng, X.J. Shi, Y.F. Xu and H. Li, “Deep Convolutional Neural Network based on CT Images of Pulmonary Nodules in the Lungs of Adolescent and Young

Adult Patients with Osteosarcoma", *Oncology Letters*, Vol. 26, No. 2, pp. 1-8, 2023.

[16] C. Loraksa, S. Mongkolsomlit, N. Nimsuk, M. Uscharapong and P. Kiatisevi, "Effectiveness of Learning Systems from Common Image File Types to Detect Osteosarcoma based on Convolutional Neural Networks Models", *Journal of Imaging*, Vol. 8, No. 1, pp. 1-7, 2021.

[17] N. Sivakumar and A. Shankar, "The Speech-Language Processing Model for Managing the Neuro-Muscle Disorder Patients by using Deep Learning", *NeuroQuantology*, Vol. 20, No. 8, pp. 918-932, 2022.

[18] T. Aggarwal, A. Furqan and K. Kalra, "Feature Extraction and LDA based Classification of Lung Nodules in Chest CT Scan Images", *Proceedings of International Conference on Advances in Computing, Communications and Informatics*, pp 1189-1193, 2015.

[19] T. Roy, N. Sirohi and A. Patle, "Classification of Lung Image and Nodule Detection using Fuzzy Inference System", *Proceedings of International Conference on Computing, Communication and Automation*, pp. 1204-1207, 2015.

[20] K. Bhavanishankar and M.V. Sudhamani, "Techniques for Detection of Solitary Pulmonary Nodules in Human Lung and their Classifications-A Survey", *International Journal on Cybernetics and Informatics*, Vol. 4, No. 1, pp. 27-40, 2015.

[21] Q. Li, "Recent Progress in Computer-Aided Diagnosis of Lung Nodules on Thin-Section CT", *Computerized Medical Imaging Graphics*, Vol. 31, No. 4, pp. 248-257, 2007.

[22] M.E. Fantacci, "Algorithms for Automatic Detection of Lung Nodules in CT Scans", *Proceedings of International Symposium on Medical Measurements and Applications*, pp. 623-627, 2011.

[23] J. Kuruvilla and K. Gunavathi, "Lung Cancer Classification using Neural Networks for CT Images", *Computer Methods Programs in Biomedicine*, Vol. 113, No. 1, pp. 202-209, 2014.

[24] E.L. Torres, E. Fiorina, F. Pennazio, C. Peroni, M. Saletta, N. Camarlinghi, M.E. Fantacci and P. Cerello, "Large Scale Validation of the M5L Lung CAD on Heterogeneous CT Datasets", *Medical Physics*, Vol. 42, No. 4, pp. 1477-1489, 2015.

[25] D. Kumar, A. Wong and D.A. Clausi, "Lung Nodule Classification using Deep Features in CT Images", *Proceedings of International Conference on Computer and Robot Vision*, pp. 133-138, 2015.

[26] R. Fakoor, F. Ladhak, A. Nazi and M. Huber, "Using Deep Learning to Enhance Cancer Diagnosis and Classification", *Proceedings of the International Conference on Machine Learning*, Vol. 28, pp. 1-9, 2013.

[27] R. Gruetzemacher and A. Gupta, "Using Deep Learning for Pulmonary Nodule Detection and Diagnosis", *Proceedings of International Conference on Information Systems*, pp. 1-9, 2016.

[28] B. Van Ginneken, A.A. Arnaud Setio, Colin Jacobs and Francesco Ciompi, "Off-the-Shelf Convolutional Neural Network Features for Pulmonary Nodule Detection in Computed Tomography Scans", *Proceedings of International Symposium on Biomedical Imaging*, pp. 286-289, 2015.

[29] C. Loraksa, S. Mongkolsomlit, N. Nimsuk, M. Uscharapong and P. Kiatisevi, "Development of the Osteosarcoma Lung Nodules Detection Model based on SSD-VGG16 and Competency Comparing with Traditional Method", *IEEE Access*, Vol. 10, pp. 65496-65506, 2022.

[30] P. Bolourchi, H. Demirel and S. Uysal, "Target Recognition in SAR Images using Radial Chebyshev Moments", *Signal, Image and Video Processing*, Vol. 11, pp. 1033-1040, 2017.

[31] R.M. Haralick, K. Shanmugam and I.H. Dinstein, "Textural Features for Image Classification", *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 6, pp. 610-621, 1973.

[32] X. Li; H. Hu, L. Zhao, H. Wang, Y. Yu, L. Wu and T. Liu, "Polarimetric Image Recovery Method Combining Histogram Stretching for Underwater Imaging", *Scientific Reports*, Vol. 24, No. 1, pp. 1-7, 2018.

[33] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition", *Proceedings of International Conference on Computer Vision and Pattern Recognition*, pp. 770-778, 2016.

[34] G. Ghiasi, T.Y. Lin, R. Pang; Q.V. Le, "NAS-FPN: Learning Scalable Feature Pyramid Architecture for Object Detection", *Proceedings of International Conference on Computer Vision and Pattern Recognition*, pp. 7029-7038, 2019.

[35] S. Liu, L. Qi, H. Qin, J. Shi and J. Jia, "Path Aggregation Network for Instance Segmentation", *Proceedings of International Conference on Computer Vision and Pattern Recognition*, pp. 8759-8768, 2018.