

3D visualization diagnostics for lung cancer detection

Rana M. Mahmoud^{1,2}, Mostafa Elgendy¹, Mohamed Taha¹

¹Department of Computer Science, Faculty of Computer and Artificial Intelligence, Benha University, Benha, Egypt

²Faculty of Computers and Information Technology, The Egyptian E-learning University, Giza, Egypt

Article Info

Article history:

Received Oct 11, 2023

Revised Feb 22, 2024

Accepted Mar 21, 2024

Keywords:

3D reconstruction

Computed tomography

Deep learning

Lung nodule detection

Lung nodule segmentation

ABSTRACT

Lung cancer is the primary contributor to cancer-related deaths globally, accounting for approximately 2 million new diagnoses and resulting in 1.76 million deaths yearly. Early detection can improve survival, and computerized tomography (CT) scans are a precise imaging technique to diagnose lung cancer. However, analyzing hundreds of 2D CT slices is challenging and can cause false alarms. 3D visualization of lung nodules can aid clinicians in detection and diagnosis. The MobileNet model integrates multi-view and multi-scale nodule features using depthwise separable convolutional layers. These layers split standard convolutions into depthwise and pointwise convolutions to reduce computational cost. Finally, the 3D pulmonary nodular models were created using a ray-casting volume rendering approach. Contrasted to other cutting-edge deep neural networks, this factorization enables MobileNet to achieve a much lower computational cost while maintaining a decent degree of accuracy. The proposed approach was tested on a dataset comprising 986 nodules from lung image database consortium (LIDC). Experiment findings reveal that MobileNet achieves outstanding performance in segmenting the LIDC dataset with an accuracy of 93.3%. Conclusion: The study demonstrates that the MobileNet detects and segments lung nodules somewhat better than other older technologies. As a result, the proposed system proposes an automated 3D lung cancer tumor visualization.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Rana M. Mahmoud

Faculty of Computers and Information Technology, The Egyptian E-learning University

33 El Mesaha Street, Dokki, Giza, Egypt

Email: rmohamedali@eelu.edu.eg

1. INTRODUCTION

Lung cancer originates in the lungs, which are vital organs for breathing. It stands as the foremost contributor to global cancer-related fatalities. Lung cancer is classified into two types: small-cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). SCLC consists of cancer cells that are small and viewed under a microscope. In contrast, in NSCLC, the cancer cells are more extensive. Generally, SCLC progresses faster than NSCLC. NSCLC represents the prevailing manifestation of the disease. It happens when abnormal cells in the lungs grow and divide uncontrolled, causing a tumor or cancer that begins in the lungs, which are the organs responsible for breathing that can interfere with normal lung function. Lung cancer is most often caused by smoking; however, it can also be caused by other environmental factors such as air pollution and radon gas exposure. Tobacco use is the primary cause of lung cancer, accounting for around 80% of all occurrences [1]. Common risks include being subjected to air pollution, secondhand smoke, radon gas, and specific chemicals and substances encountered in certain occupational settings. According to the World Health Organization, lung cancer will account for around 1.8 million deaths globally in 2020, accounting for nearly 18% of all cancer fatalities, as shown in Figure 1 [2]. According to data from 2019, 324,949 patients in Egypt were undergoing treatment for malignant neoplasms at the expense of the state [3]. As per the December 2020

estimates from the Global Cancer Observatory (GLOBOCAN), the most prevalent cancers in Egypt, with a 5-year prevalence across all age groups, include breast, lung, colorectal, prostate, stomach, liver, and cervical cancer, totaling 8,879,843 cases for all cancers, as depicted in Figure 2 [4].

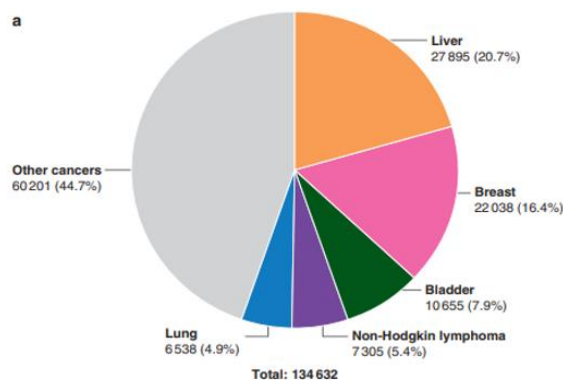


Figure 1. Statistics for cancers on the globe [2]

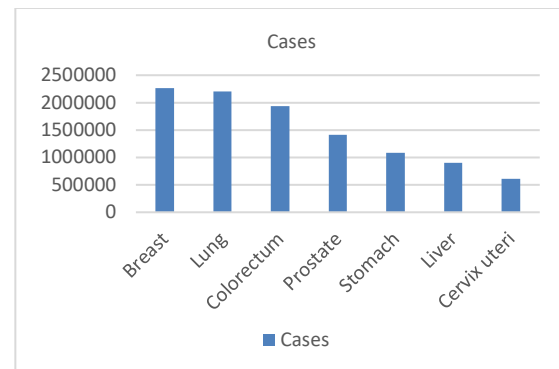


Figure 2. Statistics prevalent cancers in Egypt [4]

The primary goal nowadays is to detect and forecast cancer to begin treatment as quickly as possible [5]. Screening and incidental findings are the two most prevalent methods in identifying lung cancer. Similar to breast cancer screening, the primary approach for detecting lung cancer should be through screening [6]. Despite this, most nations, including the UK, lack a screening program, and most cases are detected by coincidence. Lung nodules are discovered unexpectedly while examining a different organ from the one being examined. Nodules, for example, may be discovered during a computerized tomography (CT) scan of the heart or liver. The main challenge is that radiologists lack the necessary skills to distinguish between benign and malignant nodules [7]. These patients should be sent to a pulmonologist, who will perform various tests to check for malignant cells and rule out additional medical conditions. In this case, clinicians order an X-ray of the lungs and look for signs of a tumor, scarring, or a buildup of fluid because it may indicate a suspicious lump or nodule [8].

Even minute lung abnormalities that an X-ray might miss can be detected by a CT scan. A biopsy may be executed via bronchoscopy, among other techniques. The physician examines aberrant regions of the lungs by inserting a lit instrument down the patient's esophagus and into the lungs. Mediastinoscopy is an alternative technique that collects lymph node tissue samples through the insertion of surgical instruments through an incision made at the base of the neck and behind the breastbone. An alternative method is a needling biopsy, during which the physician inserts a needle into the lung tissue via the chest wall and utilizes X-ray or CT imaging to identify abnormal cells. An additional possibility is the collection of a biopsy sample from lymph nodes or other metastases of the malignancy, including the liver [9].

Utilizing three-dimensional visualization diagnostics in lung cancer detection enables radiologists and oncologists better to comprehend lung lesions' spatial distribution and properties. This aids in precise diagnosis, treatment strategizing, and ongoing monitoring of lung cancer patients can result in improved patient prognoses over time [10]. The application of three-dimensional visualization diagnostics in lung cancer detection allows radiologists and oncologists to have a more comprehensive understanding of lung lesions' spatial distribution and characteristics. This can assist in accurate diagnosis, treatment planning, and monitoring of lung cancer patients, leading to improved patient outcomes. Volumetric reconstruction is performed in 2D using the MobileNet approach by using the result as beginning data and propagating the result. It is accomplished by deploying machine learning on massive healthcare image collections.

Liu *et al.* [11] proposed a data-driven method known as the cascaded dual-pathway residual network (CDP-ResNet), leveraging ResNet to improve the segmentation of lung nodules in CT images. This model subsequently calculates the probability of voxel membership within the nodule and provides a 3D visualization of the final segmentation outcome. Four radiologists extensively evaluated that method on a lung image database consortium (LIDC) dataset containing 986 annotated nodules. The results indicate that the CDP-ResNet model outperforms assessments conducted by four different radiologists. Additionally, the dice similarity coefficient (DSC) between CDP-ResNet and each radiologist averages 82.69%, slightly surpassing the average variability observed among radiologists on the LIDC dataset, which stands at 82.66% [11].

Liu *et al.* [12] used the context attention network (CA-Net). It extracts both contextual features and nodules, effectively integrating them during benignity/malignancy classification. To precisely capture

contextual features influenced by or associated with the nodule, an attention mechanism incorporates the nodule's information as a reference. Besides, contextual features' effect on classification can vary across nodules. The data science bowl 2017 (DSB) is a dataset for lung cancer prediction used for testing and evaluation. The experimental results show that CA-Net reaches an accuracy of 83.79% [12].

Mei *et al.* [13] proposed the Mask region-based convolutional neural network (R-CNN) using ResNet-50 as the backbone and applied feature pyramid network (FPN) to explore multi-scale feature maps fully. Using region proposal network (RPN) to propose candidate bounding boxes. The lung nodule analysis (LUNA) challenge is tested and evaluated on the publicly accessible LUNA16 dataset provided as part of ISBI 2016. Experimental findings indicate that Mask R-CNN achieves a sensitivity of 88.1% [13].

Cai *et al.* [14] proposed an artificial intelligence lung imaging analysis system (ALIAS) featuring lung nodule detection and segmentation networks. The three-dimensional rectified linear unit (ReLU) cascade FPN is used for nodule detection. In nodule-based analysis, features such as histograms of hounsfield units (Hus) and radionics features are extracted. These features allow for the comparison of discrepancies observed between malignant and benign nodules, of various sizes. The ALIAS undergoes testing and evaluation on images obtained from collaborative institutions, following the institutional review board protocols and respective material transfer agreements. This evaluation encompasses a total of 8540 pulmonary CT images from 7,716 patients. Within the testing set, there are 138 malignant nodules (positive) and 91 benign nodules (negative), resulting in an accuracy of 83.8% [14].

Chen *et al.* [15] introduced the slice-aware network (SANet) for the detection of lung nodules. This network utilizes a slice-grouped non-local (SGNL) module and U-Net-like structure and ResBlock to generate nodule candidates. The performance of the SANet is assessed and validated using the pulmonary nodule dataset (PN9), which comprises 40,439 annotated and nodules 8,798 thoracic CT scans. The experimental results show that SANet reaches a precision of 35.92% and a recall of 70.20%, which means the accuracy is 87% [15].

Mkindu *et al.* [16] proposed a computer-aided detection (CAD) scheme incorporating a 3D multi-scale vision transformer (MSViT). This architecture employs a local-global transformer block structure, with the local transformer stage processing each scale patch independently before merging them into the global transformer level to incorporate multi-scale features. The CAD scheme proposed was tested on 888 CT images from the publicly available LUNA16 dataset. The 3D-MSViT algorithm achieved an accuracy rate of 91.1% [16].

Zhang and Zhang [17] proposed that a 3D selective kernel residual network (SK-ResNet) based on the selective kernel network and three-dimensional residual network is located. A 3D RPN, employing SK-ResNet, is developed for the identification of lung nodules, accompanied by a multi-scale feature fusion network tailored for nodule classification. The effectiveness of the method is examined and assessed using the publicly available LUNA16 dataset. The SK-ResNet algorithm achieved an accuracy rate of 91.75% [17].

This paper uses the MobileNet model that integrates multi-view and multi-scale characteristics of various nodules from CT images. Then, the three-dimensional pulmonary nodular models were created using a ray-casting volume rendering approach. The creation of the three-dimensional model will offer lossless reconstruction, providing a more realistic depiction of the tumor compared to the wired model, utilizing the three-dimensional reconstruction of the tumor cells. It will allow the physician can more effectively examine the relationship between the tumor and its surroundings, even before surgery. This proposed system will analyze lung damage globally: the full CT image, using deep learning approaches. This research has the following contributions: i) the system suggests a three-dimensional visualization technique for computer-aided lung nodule identification based on MobileNet and the ray-casting volume rendering methodology and ii) classify pulmonary nodules. The experimental results indicate that MobileNet is useful for segmenting and classifying lung nodules.

2. METHOD

Recent studies have underscored significant advancements in medical image segmentation enabled by the application of deep learning methodologies. The improvement is clear, especially When comparing the outcomes of traditional segmentation approaches with those of deep learning techniques in liver delineation. The proposed system detects lung tumors automatically. It segments the CT images of the lung to recognize. The method reconstructs the tumor in three-dimensional images with tumor volumetry. The system includes three phases, as shown in Figure 3.

- The pre-processing module (PrM) phase converts lung images from world coordinates to image coordinates, splitting images into slices and normalizing images.
- The detection module (DM) phase detects lung tumors.
- The three-dimensional reconstruction module (3DRM) phase visualizes the lung tumors in three-dimensional.

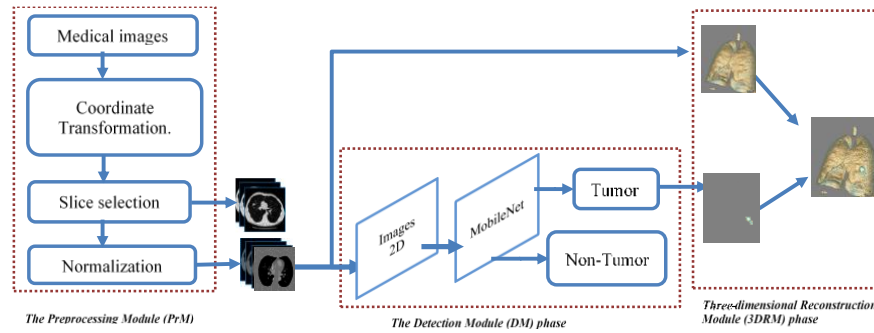


Figure 3. The proposed model-based segmentation and detection framework for pulmonary nodule three-dimensional visualization diagnosis

2.1. Pre-processing module

The input data CT images are represented by dividing three-dimensional scans into 2D slices. They were resampled and normalized to each component. Finally, each slice picture is converted into a binary image. The blobs related to the image's border are removed as a pre-processing step to remove any artifacts or noise that may be present in the data. This phase assists in increasing the accuracy and dependability of analytic processes like segmentation or feature extraction. The two biggest portions are preserved and utilized to divide the two lungs. Holes filled the mask, separating the lungs and revealing each lung's convex hull. The two lungs are joined to keep nodules fixed to the lung wall, as shown in Figure 4.

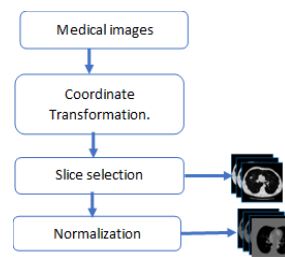


Figure 4. Pre-processing module

2.1.1. Coordinate mapping components

The raw coordinates of medical image data are expressed in the global coordinate system. In (1) represents transforming the global coordinate framework to the image-based coordinate structure. Figure 5 visually depicts this conversion process for better understanding [18].

$$\text{Image_coordinates} = \frac{|\text{world_coordinates} - \text{origin}|}{\text{Spacing}} \quad (1)$$

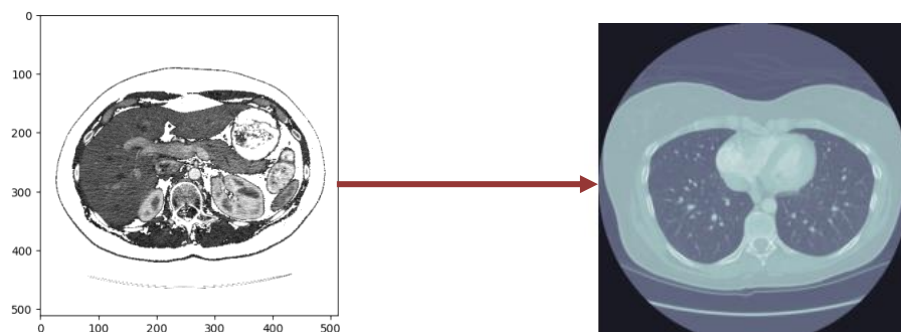


Figure 5. Medical image-to-image coordinates

2.1.2. Slice selection

The Z-axis in the image coordinate system extends from lower to upper. The logical equation represented for the diameter, slice thickness, and coordinates in the annotated file define the position, written selected z , and the number of chosen slices, written as in (2) [19], [20].

$$EN = \begin{cases} \text{True} & (d - n * s \geq 3) \text{ and } (c - n \leq s_z) \text{ and } (c + n \geq s_z) \\ \text{False} & \text{otherwise} \end{cases} \quad (2)$$

Where EN is effective nodules, d is diameter, s is slice thickness, s_z is selected z , and c is coordinate z may be derived from the annotated file. Effective nodules were defined as nodules greater than 3 mm in diameter, as shown in Figure 6.



Figure 6. Image coordinates into slices of an image

2.1.3. Normalization

The relative density of tissues and organs in CT scans is represented by the Hu. The Hounsfield value of bones in a chest CT scan is larger than 400, whereas the Hounsfield value of the lung ranges from 1000 Hu to 400 Hu. The final picture pixel values were windowed in the [1000, 400] range before being normalized to [0, 1], as shown in Figure 7.

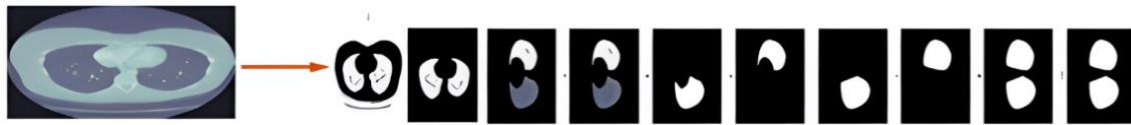


Figure 7. Normalization

2.2. Detection module

This part will discuss the MobileNet core layers, which are depth-wise separable filters. The subsequent section 2.2.2 elaborates on the overall structure of the MobileNet network. A visual representation of this structure is presented in Figure 8 for better conceptualization.

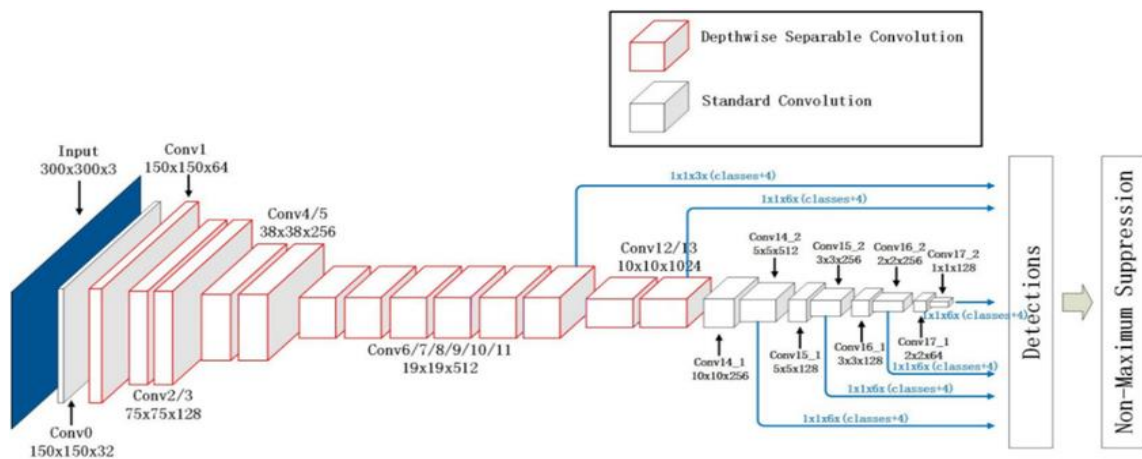


Figure 8. MobileNet architecture [21]

2.2.1. Depth-wise separable convolution

An example of a factorized convolution is the depth-wise separable convolution, which takes a regular convolution and splits it into two: depth-wise and pointwise. As such, the MobileNet model is based on this design. With depth-wise convolution, MobileNet processes each input channel with a single filter. Eleven convolutions are executed by merging the results of depth-wise and pointwise convolutions. Combines inputs with ordinary convolution filters to produce more outputs all at once. A depth-wise separable convolution creates a filtering layer and a combining layer. Computing and model size are both significantly reduced using the factorization technique. The square input feature map F 's spatial width and height, denoted as $DF \times DF \times N$ (DF), and the feature map G that results from it are represented as $DF \times DF \times N$. In a square output feature map, dynamic group (DG) denotes the spatial extent and height, and N the number of output channels (input depth). One million indicates the density of input channels in a given spatial context. The conventional convolutional layer uses a convolution kernel K with parameters $DK * DK * M * N$, where DK is the spatial dimension of the kernel, M is the number of input channels, and N is the number of output channels. This is in keeping with what was previously described. Feature maps produced by conventional convolution with stride one and padding look like this:

$$G_{k,l,n} = \sum_{i,j,m} F_{k+i-1,i+j-1,m} \cdot K_{i,j,m,n}$$

The computational cost of standard convolutions is:

$$D_f \cdot D_f \cdot D_k \cdot D_k \cdot M \cdot N$$

If the computational cost depends on the size of the feature map, kernel, input channels, and output channels, then...We cover all these keywords and their linkages in MobileNet models. To decouple the kernel size from the number of output channels, we begin with depth-wise separable convolutions. To create a new representation, the conventional convolution method uses convolutional kernels to filter and combine data. The use of depth-wise separable convolutions allows for a significant reduction in computation costs by splitting the filtering and combination stages into two parts. Two layers make up depth-wise separable convolutions: The two main types of convolutions are pointwise and depth-wise. Applying depth-wise convolutions (input depth) with a single filter per input channel. Applying a basic 11 convolution follows the linear combination of the depth-wise layer's output using pointwise convolution. For both layers, MobileNet employs batch normalization and ReLU activations. Here is an example of how to express normal convolution using one filter for each input channel (input depth),

$$G_{k,l,n} = \sum_{i,j,m} F_{k+i-1,i+j-1,m} \cdot K_{i,j,m,n}$$

If we take the $DK \times DK \times M$ depth-wise convolutional kernel K and apply it to the m th channel in F , we get the m th channel of the filtered output feature map G .

2.2.2. Network structure and training

In order to discover a good network topology, we may quickly assess them by giving a clear description of the network. For nonlinear classification, all layers except the last completely linked one feed into a SoftMax layer. A batch norm and ReLU nonlinearity follow each layer. When comparing two layers, one uses conventional ReLU nonlinearity, batch-norm, and convolutions, while the other is a factorized layer that uses depth-wise convolution, 1×1 pointwise convolution, batch-norm, and ReLU after each convolutional layer [22]. Is the first layer and the depth-wise convolutions use stride convolution to handle down-sampling. Using a last average pooling procedure reduces the spatial resolution to 1 before the fully linked layer. There are a total of 28 layers in MobileNet when you separate the depth-wise and pointwise convolutions. It is inadequate to describe networks using just a small number of multiply-adds. Efficiently implementing these activities is of the utmost importance. As an example, unless the sparsity level is really high, dense matrix operations will often outperform unstructured sparse ones. Deep 1×1 convolutions include all calculations inside the system structure. Very efficient generic matrix multiply (GEMM) routines make this possible. The initial memory reordering called im2col is necessary to align it with a GEMM, which is typically used to conduct convolutions [23]. The popular caffe package, for instance, makes use of this technique. One of the best numerical linear algebra methods, GEMM, allows you to easily implement 1×1 convolutions without reordering memory.

a. Three-dimensional reconstruction module

Three-dimensional reconstruction creates a 3D model of a scene from a set of two-dimensional images or an object or video frame. This technology is widely used in entertainment, medical imaging, and computer vision. A 3D reconstruction module typically consists of the following steps:

- The initial stage is to collect a collection of two-dimensional lung images. These images are created by dividing 3D images into slices. Each slice is a 2D image obtained from medical imaging equipment.
- In the next step, images are processed to extract features that may be used to detect common points between them. These extracted features are crucial in detecting common points between the images. They contribute to the algorithm's ability to identify and match key points effectively.
- The retrieved features calculate the corresponding points between the images in this stage. This process involves comparing the features of each image and selecting the best match to establish correspondence. The matching of features enhances the precision of identifying corresponding points across the images [24].
- Identifying the intrinsic and extrinsic properties of the cameras used to collect images is known as camera calibration. These parameters determine the camera's position and orientation in three-dimensional space. Establishing these properties enhances the accuracy of spatial relationships and measurements in the captured imagery [25].
- A 3D model of the object or scene is reconstructed using the estimated correspondences and camera parameters. This reconstruction process typically employs triangulation, stereo reconstruction, or structured light scanning techniques. These methods contribute to accurately capturing the three-dimensional representation of the observed object or scene [26].
- Once the 3D model is reconstructed, textures can be applied to the model's surface to give it a realistic appearance. This process involves projecting textures from 2D images onto the corresponding surfaces of the model. The 3D model gains a lifelike appearance by seamlessly integrating these textures, enriching visual fidelity.

The marching cubes algorithm is widely used for extracting an iso-surface (a surface representing a The marching cubes technique, introduced by Lorensen and Cline [27] in 1987, is a widely used method for extracting polygonal mesh representations of iso-surfaces from 3D scalar fields or voxel data, which is often used for medical imaging data such as magnetic resonance imaging (MRI) scans [27], [28]. The fundamental concept is to divide 3D space into cubes, with the vertices of each cube corresponding to 2D neighboring slices. The algorithm moves over the scalar field, considering eight neighboring points simultaneously, generating an imaginary cube, as illustrated in Figure 9. This approach finds the polygons needed to depict the iso-surface passing through this cube. The individual polygons are merged to generate the required surface. The merging method is aided using an index of a precomputed array of 256 potential polygon configurations. The number of different instances has been reduced to 15 due to symmetry and rotation procedures, as illustrated in Figure 10.

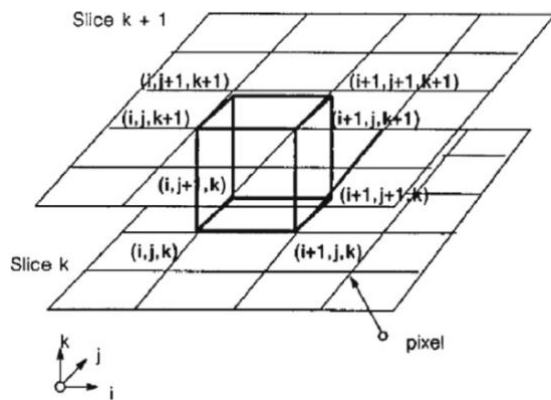


Figure 9. Cube delimited by 2D adjacent slices

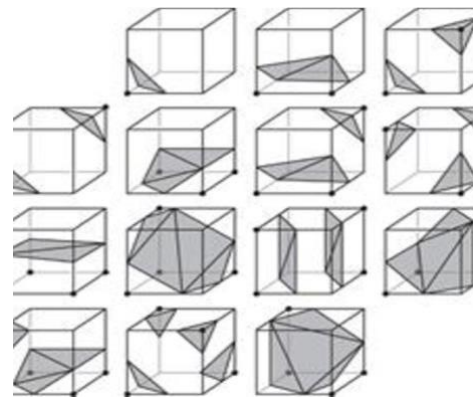


Figure 10. 15 unique cube configurations generated by marching

Each cube's eight scalar values correspond to a bit in an 8-bit integer. If the value of a scalar exceeds the iso-value (showing that it is inside the surface), the associated bit is set to one; otherwise, it is set to zero. The final 8-bit integer, resulting from checking all eight scalars, is used as an index to access the polygon indices array. Each vertex of the created polygons is placed at the correct location along the cube's edge by linearly interpolating between the two scalar values connected by that edge to construct the final mesh. The `marching_cubes_lewiner` function accepts a three-dimensional binary image p as input and outputs a triangular mesh representation of the binary image's surface. The indices of the three vertices that make up a face are stored in each array row. We can see the result in Figure 11 [29].

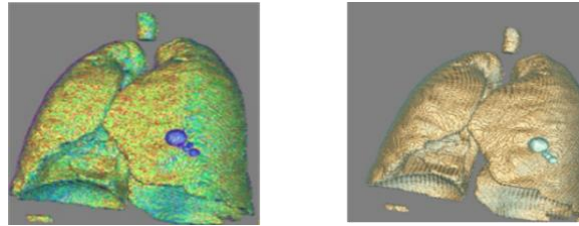


Figure 11. Lung three-dimensional with tumor

3. RESULT AND DISCUSSION

3.1. Experiment result and analysis

We use the size measurement of nodules, an essential factor for radiologists to classify benign and malignant nodules. Additionally, we use 2D texture features, statistical and geometric, which give us improved results. To analyze the performance of the proposed algorithm qualitatively and quantitatively, we select a MobileNet framework to implement the corresponding network model. This model is beneficial for tasks in which complicated evaluative speed is critical. The experimental hardware platform is the Nvidia Tesla K80 GPU (12GB) and Nvidia Tesla K80(12GB). We now present this section in detail.

3.1.1. Data and analysis

The LUNA16 dataset challenge is part of ISBI 2016. It is a publicly available medical imaging dataset that contains CT scans of the chest from 1,018 patients. It excluded scans with slice thickness greater than 2.50 mm (about 0.1 in) from lung image database consortium and image database resource initiative (LIDC-IDRI). The dataset provides a valuable resource for researchers working on lung nodule detection and classification, which are early signs of lung cancer. The images are in RAW format and labeled with nodule annotations from multiple radiologists. The dataset also includes a CSV file with additional patient and nodule metadata. The LUNA16 dataset is widely used in machine learning research to develop nodule methods for detection and classification [30].

3.1.2. Experimental data sets and evaluation criteria

This paper utilizes the LUNA16 dataset for both testing and training of the network. The dataset comprises 1,018 patient case images, divided into two sections. We partitioned the LUNA16 datasets randomly and equally into ten subsets. Nine of these subsets, totaling 800 CT scans, were utilized for training, while one subset containing 88 CT scans was reserved for testing. Tenfold cross-validation was performed on the test model. Among the 888 CT images, a total of 36,378 nodules were identified, with 1186 nodules selected for analysis. Precision and sensitivity metrics were employed in this study to assess the localization performance of the proposed system. We utilized accuracy to assess how many expected positives were successfully recognized, sensitivity to establish how many expected positives were recognized, and sensitivity to establish how many genuine positives were accurately detected. The F1 score was also used to compute the harmonic mean of sensitivity and accuracy, as well as the harmonic mean of sensitivity and accuracy. Because the F1 score measures accuracy and sensitivity, it always assigns equal weight to both measures. This is because there is always a trade-off between the two since the mean value is always strongly influenced at the expense of either value.

3.1.3. Training, testing, and implementation details

The image size of the network input is $224 \times 224 \times 3$. The MobileNet architecture uses depthwise separable convolutions, combining a pointwise (which applies a 1×1 filter to the output of the depthwise convolution) and a depthwise convolution (which applies a single filter to each input channel). The depthwise separable convolution reduces the number of parameters and the amount of computation needed compared to a standard convolutional layer. This also includes a custom implementation of the depthwise convolution, which applies a separate 2D convolution to each input channel using a shared kernel. This implementation allows for more efficient computation than a standard depthwise convolution, which applies a single filter to each input channel. The MobileNet architecture consists of a series of depthwise separable convolutional layers and an initial convolutional layer. The output of the last convolutional layer is fed into a global average pooling layer, which computes the average value of each feature map, and a fully connected output layer. Then the output show with voxel size is $32 \times 32 \times 32$. The model uses the loss function and the stochastic gradient descent (SGD). It uses the SGD optimizer, with the learning rate set to 0. decay, momentum to 0.9, and weight decay.

3.1.4. Performance metrics

Performance metrics are essential tools that help organizations measure and evaluate how well they are achieving their goals. They provide a quantitative basis for assessing the effectiveness of various processes, systems, or teams. In this discussion, we will focus on specific performance metrics, such as sensitivity, specificity, accuracy, and mean squared error (MSE).

– Sensitivity

Sensitivity measures the ability to detect the true positives from the test. It is calculated by considering the detected true positives in relation to the total number of actual positives. In (3) provides the specific mathematical representation for sensitivity in this context [31].

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (3)$$

– Specificity

Specificity measures the ability to determine benign cases correctly. It is quantified by (4), which defines the mathematical representation for specificity. This equation captures the ability of a test to correctly classify true negatives among the total number of actual negatives [31].

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (4)$$

– Accuracy

Accuracy measures the ability to differentiate the malignant and benign cases correctly. In (5) provides a mathematical representation for accuracy, capturing the ratio of correctly identified cases to the total number of cases. This equation serves as a quantitative measure of the overall performance of the classification or diagnostic process [31].

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

– Mean squared error

MSE is a statistical measure used to quantify the average squared deviation between predicted and observed values in regression tasks. It quantifies the average squared error or variance of the model's predictions. In (6) represents MSE [32].

$$MSE = \frac{\sum(\text{actual}-\text{prediction})^2}{\text{Number of observations}} \quad (6)$$

In this context, TP represents the count of true positives, FN represents the count of false negatives, and FP represents the count of false positives. Recall, precision, accuracy, F-score, and MSE got the values of 0.5, 1.0, 0.93, 0.67, and 0.067, respectively. There are opportunities for further improvement of these evaluation criteria, which we are currently working on. We can see this in the confusion matrix. The proposed methodology performs better than the previous approaches, as shown in Figure 12.

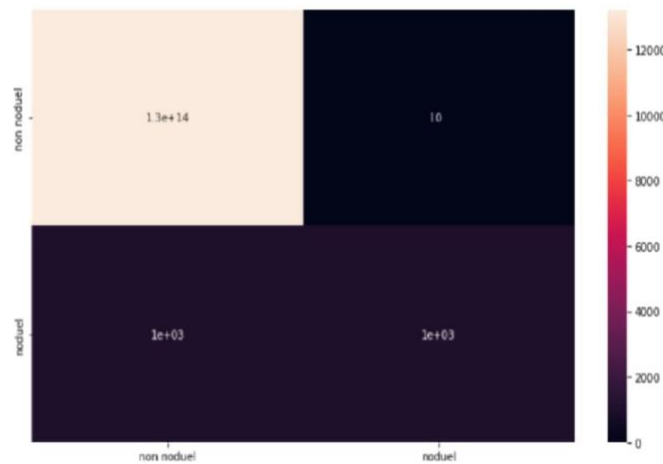


Figure 12. Confusion matrix

3.1.5. Three-dimensional reconstruction results

Examine the medical records of a single patient. The original CT scans of him had a size of $224 \times 224 \times 133$, 133 with the size 224×224 photos were acquired by pre-processing as indicated in section 3. 1, the MobileNet retrieved some images with nodules, and nodule normalization was achieved. Normalizing the original image yields the nodule sequence. We kept the nodule sequence size of $224 \times 224 \times 133$ to ensure that nodules and lungs are in the same coordinate. Lastly, we obtained three-dimensional models of pulmonary nodules (as shown in Figure 12) and lungs using ray-casting volume rendering. Figure 12 depicts three-dimensional representations of lung and pulmonary nodules. Curves of color and opacity values in ray-casting volume rendering are shown in the image's lower right corner. In terms of overall system performance, the majority of the time is spent obtaining nodule masks through the MobileNet, with a speed of 1.29 seconds required per scan a basic PC equipped with an Nvidia Tesla K80 GPU (12GB) and RAW (13GB), which takes 142 seconds in total. The three-dimensional reconstruction step occupies the majority of memory, with a total capacity of 6.081 GB for three-dimensional models of the lung and three nodules. Overall, as long as the nodule masks are precise enough, we can obtain three-dimensional models of the projected nodules and lungs using our approach. Also, we can observe more about lung and nodules tissues by altering the color and opacity of the ray-casting, which is incredibly useful for diagnosis and subsequent therapy.

3.2. Discussion

The proposed system is used for pulmonary nodule segmentation, detection, and three-dimensional visualization system. The system using the Luna dataset is a public and available dataset. It is used for detecting tumors by the MobileNet model and then visualized in three-dimensional based on the volume rendering algorithm of ray-casting. This system can render the three-dimensional models of the detected lungs and nodules, as shown in Figure 13. Note that three-dimensional reconstruction modules and the detection are designed to be separated. Therefore, we can enhance the segmentation network to get designed to be separated and a dedicated contour to detect other kidney, cancer, or lesions tumors. Also, the volume rendering algorithm can be replaced and optimized to meet practical needs. In the detection and segmentation module, we used MobileNet. As shown in Table 1, MobileNet outperforms previous systems evaluated under similar environments, which indicates that it has promise for two-stage object recognition in medical image analysis. The MSE for this model is equal to 0.067.

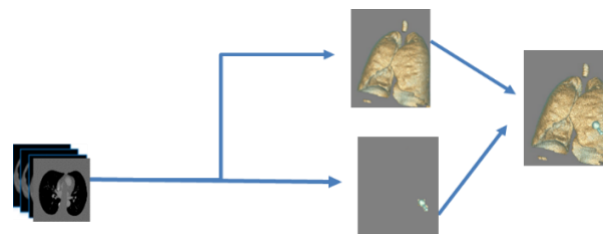


Figure 13. Intermediate and results of 3D reconstruction

Table 1. Comparison of accuracy algorithms

Reference	Model	Dataset	Training set	Testing set	Accuracy (%)
[11]	CDP-ResNet	LIDC	447	539	82
[12]	CA-Net	DSB 2017	1595	506	83.79
[15]	ALIAS	CT images	23979	7989	83.8
[13]	SANet	PN9	6707	2091	87.4
[14]	Mask R-CNN	LUNA	800 cases	88 cases	88.1
[16]	3D-MSViT	LUNA	800 cases	88 cases	91.1
[17]	3D SK-ResNet	LUNA	800 cases	88 cases	91.75
Proposed alg.	MobileNet	LUNA	800 cases	88 cases	93.3

Our network utilized only a single two-dimensional view image, whereas Setio *et al.* [33] employed multi-view 2D images, and Farhat *et al.* [34] incorporated three-dimensional spatial data. This suggests that incorporating more spatial information may enhance performance. Regarding three-dimensional visualization, there's room for optimizing rendering effects. For instance, distinct volume rendering techniques or parameters could be tailored for lungs and pulmonary nodules. Additionally, memory optimization poses a challenge in practical implementations. As depicted in Figure 12, we can discern the relative positions and sizes of pulmonary nodules within the lung. The development of a segmentation, detection, and three-dimensional

visualization system for assisting in the diagnosis of pulmonary nodules holds promise for advancing diagnostic techniques, surgical research, and practical applications.

4. CONCLUSION

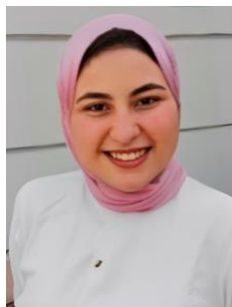
This study proposed system segmentation and detection methods for the three-dimensional visualization diagnosis of pulmonary nodules utilizing the MobileNet and ray-casting method for volume rendering algorithm to help radiologists identify pulmonary nodules with greater accuracy. The system conducted experiments related to the publicly available LUNA16 dataset to evaluate the proposed method in this paper. MobileNet model is used to detect the tumor. It extracts features by a depth-wise convolutional layer and a pointwise convolutional layer. The MobileNet model performed well regarding segmentation accuracy, with 93.3%. The system can successfully segment challenging cases. Experiment results indicate that our proposed approaches may segment and identify pulmonary nodules more accurately, allowing patients and radiologists to evaluate easily diagnosis results. In future work, we plan to develop an algorithm dedicated to identifying lung nodules. This algorithm will then be integrated into our methodology to create a completely automated system for segmenting lung nodules. Furthermore, we propose including the MobileNet architecture in the FCN network to accelerate training and prediction. Furthermore, aims to enhance the specificity of three-dimensional models depicting lungs and pulmonary nodules. We will employ various parameters and rendering techniques to optimize the rendering effects in the 3-D reconstruction of pulmonary nodules and lungs.




REFERENCES

- [1] S. A. Bialous and L. Sarna, "Lung cancer and tobacco: what is new?," *Nursing Clinics of North America*, vol. 52, no. 1, pp. 53–63, 2017, doi: 10.1016/j.cnur.2016.10.003.
- [2] WHO, "Cancer," *World Health Organization*, 2020, [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/cancer>.
- [3] A. H. Ibrahim and E. Shash, "General oncology care in Egypt," *Cancer in the Arab World*, pp. 41–61, 2022, doi: 10.1007/978-981-16-7945-2_4.
- [4] R. Sharma *et al.*, "Mapping cancer in Africa: a comprehensive and comparable characterization of 34 cancer types using estimates from GLOBOCAN 2020," *Frontiers in Public Health*, vol. 10, 2022, doi: 10.3389/fpubh.2022.839835.
- [5] Y. Bengio, "Learning deep architectures for AI," *Foundations and Trends in Machine Learning*, vol. 2, no. 1, pp. 1–27, 2009, doi: 10.1561/2200000006.
- [6] I. Bush, "Lung nodule detection and classification," *The Report of Stanford Computer Science*, vol. 20, pp. 196–209, 2016.
- [7] W. Chen *et al.*, "Cancer statistics in China, 2015," *CA: A Cancer Journal for Clinicians*, vol. 66, no. 2, pp. 115–132, 2016, doi: 10.3322/caac.21338.
- [8] M. G. Krebs *et al.*, "Analysis of circulating tumor cells in patients with non-small cell lung cancer using epithelial marker-dependent and -independent approaches," *Journal of Thoracic Oncology*, vol. 7, no. 2, pp. 306–315, 2012, doi: 10.1097/JTO.0b013e31823c5c16.
- [9] "Lung cancer - diagnosis and treatment," *Mayo Clinic*. [Online]. Available: <http://www.mayoclinic.org/diseases-conditions/lung-cancer/basics/tests-diagnosis/con-20025531>.
- [10] P. Papadimitroulas *et al.*, "Artificial intelligence: Deep learning in oncological radiomics and challenges of interpretability and data harmonization," *Physica Medica*, vol. 83, pp. 108–121, 2021, doi: 10.1016/j.ejmp.2021.03.009.
- [11] H. Liu *et al.*, "A cascaded dual-pathway residual network for lung nodule segmentation in CT images," *Physica Medica*, vol. 63, pp. 112–121, 2019, doi: 10.1016/j.ejmp.2019.06.003.
- [12] M. Liu, F. Zhang, X. Sun, Y. Yu, and Y. Wang, "CA-Net: leveraging contextual features for lung cancer prediction," *Lecture Notes in Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, pp. 23–32, 2021, doi: 10.1007/978-3-030-87240-3_3.
- [13] J. Mei, M. M. Cheng, G. Xu, L. R. Wan, and H. Zhang, "SANet: a slice-aware network for pulmonary nodule detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 8, pp. 4374–4387, 2022, doi: 10.1109/TPAMI.2021.3065086.
- [14] L. Cai, T. Long, Y. Dai, and Y. Huang, "Mask R-CNN-based detection and segmentation for pulmonary nodule 3D visualization diagnosis," *IEEE Access*, vol. 8, pp. 44400–44409, 2020, doi: 10.1109/ACCESS.2020.2976432.
- [15] L. Chen *et al.*, "An artificial-intelligence lung imaging analysis system (ALIAS) for population-based nodule computing in CT scans," *Computerized Medical Imaging and Graphics*, vol. 89, 2021, doi: 10.1016/j.compmedimag.2021.101899.
- [16] H. Mkindu, L. Wu, and Y. Zhao, "3D multi-scale vision transformer for lung nodule detection in chest CT images," *Signal, Image and Video Processing*, vol. 17, no. 5, pp. 2473–2480, 2023, doi: 10.1007/s11760-022-02464-0.
- [17] H. Zhang and H. Zhang, "LungSeek: 3D selective Kernel residual network for pulmonary nodule diagnosis," *Visual Computer*, vol. 39, no. 2, pp. 679–692, 2023, doi: 10.1007/s00371-021-02366-1.
- [18] F. Shariaty and M. Mousavi, "Application of CAD systems for the automatic detection of lung nodules," *Informatics in Medicine Unlocked*, vol. 15, 2019, doi: 10.1016/j.imu.2019.100173.
- [19] G. Y. Zheng, X. B. Liu, and G. H. Han, "Survey on medical image computer aided detection and diagnosis systems," *Ruan Jian Xue Bao/Journal of Software*, vol. 29, no. 5, pp. 1471–1514, 2018, doi: 10.13328/j.cnki.jos.005519.
- [20] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," *Proceedings of the IEEE International Conference on Computer Vision*, vol. 2015 Inter, pp. 1026–1034, 2015, doi: 10.1109/ICCV.2015.123.
- [21] S. Kanimozhi, G. Gayathri, and T. Mala, "Multiple real-time object identification using single shot multi-box detection," *ICCIDS 2019 - 2nd International Conference on Computational Intelligence in Data Science, Proceedings*, 2019, doi: 10.1109/ICCIDS.2019.8862041.
- [22] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *32nd*




- International Conference on Machine Learning, ICML 2015*, vol. 1, pp. 448–456, 2015.
- [23] L. Sifre and S. Mallat, "Rigid-motion scattering for texture classification," *arXiv-Computer Science*, pp. 1–19, 2014, doi: 10.48550/arXiv.1403.1687.
 - [24] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004, doi: 10.1023/B:VISI.0000029664.99615.94.
 - [25] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000, doi: 10.1109/34.888718.
 - [26] J. Salvi, S. Fernandez, T. Pribanic, and X. Llado, "A state of the art in structured light patterns for surface profilometry," *Pattern Recognition*, vol. 43, no. 8, pp. 2666–2680, 2010, doi: 10.1016/j.patcog.2010.03.004.
 - [27] W. E. Lorensen and H. E. Cline, "Marching cubes," *Seminal Graphics*, pp. 347–353, 1998, doi: 10.1145/280811.281026.
 - [28] N. S. Zghal and N. Derbel, "MRI images segmentation and 3D reconstruction for cerebral cancer detection," *Journal of Testing and Evaluation*, vol. 46, no. 6, pp. 2707–2717, 2018, doi: 10.1520/JTE20170309.
 - [29] D. I. H. Farías, R. G. Cabrera, T. C. Fraga, J. Z. H. Luna, and J. F. G. Aguilar, "Modification of the marching cubes algorithm to obtain a 3D representation of a planar image," *Programming and Computer Software*, vol. 47, no. 3, pp. 215–223, 2021, doi: 10.1134/S0361768821030051.
 - [30] A. A. A. Setio *et al.*, "Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge," *Medical Image Analysis*, vol. 42, pp. 1–13, 2017, doi: 10.1016/j.media.2017.06.015.
 - [31] N. G. N. Gupta, "Accuracy, sensitivity and specificity measurement of various classification techniques on healthcare data," *IOSR Journal of Computer Engineering*, vol. 11, no. 5, pp. 70–73, 2013, doi: 10.9790/0661-1157073.
 - [32] S. Prayudani, A. Hizriadi, Y. Y. Lase, Y. Fatmi, and Al-Khowarizmi, "Analysis accuracy of forecasting measurement technique on random k-nearest neighbor (RKNN) using MAPE and MSE," *Journal of Physics: Conference Series*, vol. 1361, no. 1, 2019, doi: 10.1088/1742-6596/1361/1/012089.
 - [33] A. A. A. Setio *et al.*, "Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1160–1169, 2016, doi: 10.1109/TMI.2016.2536809.
 - [34] Q. Dou, H. Chen, L. Yu, J. Qin and P. -A. Heng, "Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection," in *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 7, pp. 1558–1567, July 2017, doi: 10.1109/TBME.2016.2613502.

BIOGRAPHIES OF AUTHORS






Rana M. Mahmoud    received her B.Sc. from the Department of Computer Science, Faculty of Computers and Artificial Intelligence, Benha University, Egypt, 2020. She was a demonstrator at the Faculty Modern Academy for Computer Science and Management Technology, Egypt, from February 2021 to February 2022. Now, she is a demonstrator in the Department of Computer Science, The Egyptian E-Learning University, Egypt. Her research interests include artificial intelligence, machine learning, and computer vision. She can be contacted at email: rmohamedali@eelu.edu.eg



Mostafa Elgendy    received his M.Sc. degree from the Department of Computer Science, Faculty of Computers and Artificial Intelligence, Benha University, Egypt, 2015. He received his Ph.D. degree from the Department of Electrical Engineering and Information Systems, University of Pannonia, Veszprem, Hungary, in 2021. He worked as a demonstrator and assistant lecturer in the Faculty of Computers and Informatics, Benha University, Egypt, from May 2009 to 2021. Now, he is an Assistant professor in the Department of Computer Science, Faculty of Computers and Artificial Intelligence, Benha University, Egypt. His research interests include cloud computing, assistive technology, and machine learning. He can be contacted at email: mostafa.elgendy@fci.bu.edu.eg



Mohamed Taha    is an Assistant Professor at Benha University, Faculty of Computers and Artificial Intelligence, Department of Computer Science, Egypt. He received his M.Sc. degree and his Ph.D. in computer science at Ain Shams University, Egypt, in February 2009 and July 2015. His research interests concern computer vision (object tracking-video surveillance systems), digital forensics (image forgery detection – document forgery detection - fake currency detection), image processing (OCR), computer networks (routing protocols - security), augmented reality, cloud computing, and data mining (association rules mining-knowledge discovery). He has contributed more than 20+ technical papers to international journals and conferences. He can be contacted at email: mohamed.taha@fci.bu.edu.eg