IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 25, 2021, accepted April 20, 2021, date of publication April 28, 2021, date of current version June 28, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3076158

Classification of Thyroid Carcinoma in Whole Slide Images Using Cascaded CNN

AHMED S. EL-HOSSINY[1,2], WALID AL-ATABANY[2,3], OSAMA HASSAN[4], AHMED M. SOLIMAN[2], AND SHERIF A. SAMI[1,5]

[1]Biomedical and Systems Engineering Department, Higher Institute of Engineering, Cairo 11837, Egypt
[2]Department of Biomedical Engineering, Faculty of Engineering, Helwan University, Cairo 11421, Egypt
[3]Information Technology and Computer Science School, Nile University, Giza 11311, Egypt
[4]Department of Lab, International Medical Center, Cairo 21451, Egypt
[5]Department of Biomedical Engineering and Systems, Faculty of Engineering, Cairo University, Giza 12613, Egypt

Corresponding author: Ahmed S. El-Hossiny (a.elhossiny@sha.edu.eg)

ABSTRACT The objective of this research is to build a ''Whole Slide Images'' classification system using Convolutional Neural Network (CNN). This system is capable of classifying Thyroid tumors into three types: Follicular adenoma, follicular carcinoma, and papillary carcinoma. Furthermore, the cascaded CNN technique is additionally employed to classify the classified follicular carcinoma into four subclasses: follicular carcinoma, papillary follicular variant, well-differentiated follicular carcinoma, and Poorly-differentiated follicular carcinoma. Results of the proposed CNN architecture showed effective classification of Thyroid carcinoma in the whole slide images with an overall accuracy of 94.69%. In the first classification stage, the images are classified into either one of three main types with an overall accuracy of 98.74%, while in the second classification stage, using the cascaded CNN, accuracy was 95.90% for further sub-classification into four sub-classes. Our cascaded CNN outperformed the accuracy of other studies due to splitting classification process of the thyroid into two stages which reduces the number of classes in each stage.

INDEX TERMS Whole slide images, convolutional neural network, thyroid carcinoma, data augmentation, histopathology whole slide imaging.

I. INTRODUCTION

The thyroid gland plays a critical role in regulating multiple body functions such as the metabolic rate, energy expenditure, and the function of organs like the heart and the brain. It is in the lower part of the anterior neck made of two lobes. Each lobe is about 3–4 cm long and 2 cm wide and a few millimeters (mm) thick [1].

Thyroid diseases are among the most prevalent medical conditions. According to the World Health Organization (WHO) 2017, there are three main classifications for thyroid diseases [2] Goiter, Tumors, and Thyroiditis, in which tumors are classified into benign neoplasms (adenomas) and Malignant (carcinomas).

Thyroid Carcinoma is the most common endocrine malignancy, accounting for 2.1% of all new malignancies (excluding skin cancer and in situ carcinomas) diagnosed annually worldwide. The annual incidence of thyroid cancer has kept rising for the last decades, and if the current trends persist, it may become the fourth most common cancer in 2030 [3].

Thyroid carcinoma is diagnosed by a pathologist based on the visual examination of tissue samples prepared on microscopic slides. These samples may be obtained by two different techniques: Fine-Needle Aspiration (FNA) technique or surgical biopsy. Samples obtained by the FNA technique are composed of a relatively small number of cells and may be affected by the nature of the targeted area. Meanwhile, samples obtained as a tissue slice by surgical biopsy contain a large arrangement of tissue cells. Therefore, a pathologist examines samples obtained by FNA as Cytopathology slide, which of limited classification efficacy because it depends only on cell features, while those obtained by surgical biopsy as Histopathology slides, will have more classification power since they have two parameters for classification which are the cell features and the distribution pattern in the tissue.

Microscopic digital images of sample slides are acquired by either an ordinary microscope equipped with a high-resolution digital camera, with limited dimension and resolution or by a ''Whole Slide Imaging system'', which

The associate editor coordinating the review of this manuscript and approving it for publication was Kin Fong Lei.

allows the pathologist to examine the whole slide in a single image with a high resolution and a whole detail of the tissue sample.

Although the ''Whole Slide Image'' of samples reveals valuable information for tissue abnormalities classification, it represents a real challenge to pathologists due to examination burden and resource consumption. Therefore, this necessitates the use of automatic classification systems based on computers and AI technology.

The classification of thyroid tumors has been accomplished by using several different techniques of imaging modalities and machine learning over the years. For Thyroid Cytopathology classification, a Deep Semantic Mobile Application was built in 2016 by Kim *et al.* [4], four machine learning algorithms used Random Forest (RF), Linear Support Vector Machines (SVM), K-nearest neighbor (KNN), and Convolutional Neural Networks (CNN). The experiment was completed on a dataset of 459 labeled cytology images, consisting of 391 training images and 68 testing images. RF, SVM, and CNN all give the same accuracy on the test set, about 70%. Two of the convolutional neural networks VGG-16 and Inception-v3, were used in 2019 by Guan *et al.* [5]. The training was done on 279 cytological images of the thyroid. In new images, the accuracy of the VGG-16 network and Inception-v3 were 95% and 87.5%, respectively.

In 2017, Wang *et al.* [6] built an algorithm for the classification of thyroid tissue in three phases: (1) Image pre-processing and segmentation, (2) feature extraction, and (3) perform the predictive model. The trained data involve 14 H&E Tissue Microarray from 153 patients and 13 BRAF IHC Tissue Microarray from 140 patients. The results demonstrated that the prediction accuracy is still away from the accurate thyroid tissue classification in the microarray. Deep convolutional neural networks Inception-ResNet-v2 and VGG-19 were used to multi-classify thyroid carcinoma using histopathology images in 2019 by Wang *et al.* [7]. They have used 11,715 images from 806 patients in their study. The test set onVGG-19 gives a better accuracy than Inception-ResNet-v2 (97.34% vs. 94.42%, respectively).

Dov *et al.* [8] used a deep learning algorithm based on a cascade of two convolutional neural networks to classify thyroid carcinoma in whole-slide cytopathology images. Both networks share the same architecture based on VGG11. The proposed algorithm used 799 slides in the training phase. Experimental results show that the area under the curve of ROC was 0.932. In 2020 Elliott Range *et al.* [9] built a system to predict malignancy based on two convolutional neural networks (CNNs). They have used 908 whole slides from 659 patients of FNABs images. The algorithm predicted a malignancy with sensitivity and specificity of 92.0% and 90.5%, respectively. Tao *et al.* [10] developed a system for follicular segmentation of thyroid cytopathological based on WSI's. Firstly, they used a hybrid segmentation architecture, then applied the ResNet 101 for the classification phase. The dataset used in their study contains 15 WSI's from the FNAB images dataset. They achieved a segmentation accuracy of 53.4%.

Deep learning has been successfully used for different types of cell classification [11]. Convolutional Neural Networks (CNN's) showed an ability to differentiate healthy and malignant cell samples [12].

In this paper, an automatic classification system for thyroid carcinoma types from WSI's is presented based on cascaded CNN. The architecture of the cascaded CNN is evolved using different training parameters to acquire the most appropriate structure.

The manuscript is organized as follows; in section 2, the proposed methodology is discussed in detail, starting from the original WSI's and how handling occurs to be suitable for the system and how the adaptation of training parameters of the cascaded CNN model to the inputs, tools and hardware resources used in this research. Section 3 and 4 are dedicated to results and discussion, respectively, followed by a conclusion in Section 5.

The main contribution of this study is to classify thyroid carcinoma from whole slide histology image using deep convolutional neural network through the following steps:

1) Generating a dataset for the thyroid carcinoma whole slide images.

2) Patching each Whole slide image into small patches with an overlap of 5% to overcome the computation process.

3) Constructing our cascaded custom Convolutional Neural Network architecture and test it for ensuring its performance quality.

II. METHOD

This research aims to build a ''Whole Slide Images'' classification system using a particular type of deep learning neural network called Convolutional Neural Network (CNN). The proposed system consists of two main phases, as shown in Fig 1; the first phase aims to differentiate among three main types of thyroid tumors: Follicular adenoma, Follicular carcinoma, and papillary carcinoma. In the second phase, a cascaded CNN is used to further classify the classified follicular carcinoma into four subclasses: Follicular carcinoma, papillary follicular variant, Well-differentiated follicular carcinoma, and Poorly-differentiated follicular carcinoma.

A. DATABASE AND PRE-PROCESSING

Twenty-four medical Whole slide images were collected from the International Medical Center (IMC) – Egypt. For privacy, first, patient personal and medical information should be removed. Then, we prepare the samples for scanning by Whole Slide Imaging scanner ''Panoramic SCAN II'' [13] with a magnification of 20X in automatic imaging mode. Samples were cut at 5–6 μm thickness and stained with Hematoxylin and Eosin stain (H&E). Fig 2 shows samples of WSI's from the database.

The obtained images were classified and diagnosed based on examining expert pathologists using OpenEV [14], which is an application for analyzing and viewing raster vector

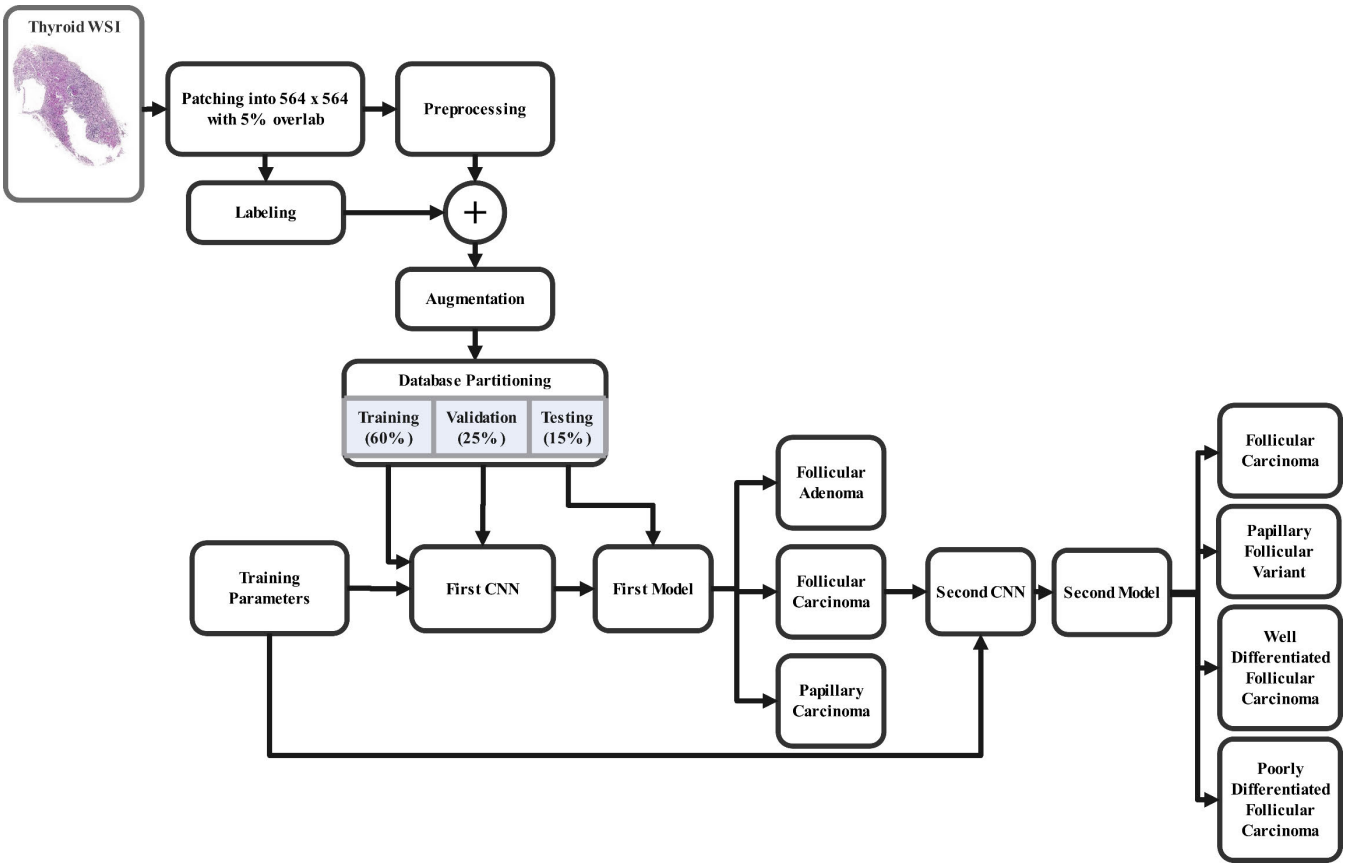

FIGURE 1. Flowchart of the whole system.

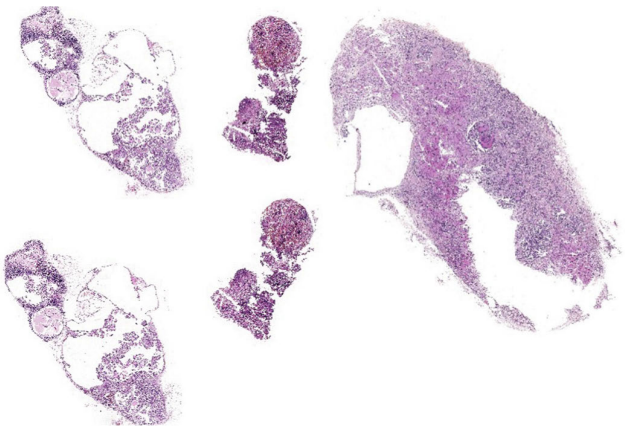

FIGURE 2. Sample of acquired database (down sampled).

TABLE 1. Database classes.

Main Class	Cases	Sub Class	Cases
Follicular adenoma	5	Follicular carcinoma	4
Follicular carcinoma	10	Papillary Follicular variant	4
Papillary carcinoma	9	Well-differentiated follicular carcinoma	1
		Poorly-differentiated Follicular carcinoma	1

geospatial images. Table 1 shows the dataset cases number and diagnosis by our consultant decision.

As WSI's have a huge size of dimension, so the classification approach cannot be directly applied. It is necessary to pre-process these WSI's before applying the classification stage.

The WSI's should be divided into small patches [15] with size (512 × 512) with 5% overlapping that yields a final size of 564 × 564 pixels. The overlapping ensures that the division does not lead to loss or damage of any critical information in the image. The patching is done using Large TIFF Tools v.1.3.6 [16]; it is an application dedicated to managing (very) large TIFF files, hard to fit entirely into the computer's memory. Generally, the acquired whole slide image consists of nearly 70% of tissue, and the remaining 30% is a background. These background images are removed by applying thresholds as they are irrelevant in the classification process.

Expert pathologists carefully classify small patches into groups to represent each class individually to be ready for training and testing. The total number of labeled patches out from this process is 18,653 images.

Algorithm 1 . Pseudocode of the Whole System

1) Star
2) For I=1: 2
   Input Whole Slide Image (I).
   Patching into small images with 576 × 576 with overlap 5%.
   Automatic threshold for removing total weight background patches.
   Random selection and manually labeling.
   Collect dataset for training.
3) End for
4) Dataset augmentation
5) Images Patches pre-processing.
6) Define training options.
7) Define CNN structure.
8) Start training and validation for the main classes by the first model.
9) End training and validation for the main classes.
10) Export true positive value images from Follicular carcinoma dataset.
11) Start training and validation for the sub-classes by the second model.
12) End training and validation for the sub-classes.
13) Validate each patch from each whole slide with the cascaded models.
14) If patch (Follicular adenoma
    Color patch = green.
15) Elseif (Follicular carcinoma
    Color patch = blue.
16) Elseif (Papillary carcinoma)
    Color patch = yellow.
17) Elseif (Papillary follicular variant
    Color patch = cyan.
18) Elseif (well-differentiated follicular carcinoma
    Color patch = black.
19) Elseif (poorly-differentiated follicular carcinoma
    Color patch = magenta.
20) Group each patch for each whole slide image.
21) End

To reduce processing time and complexity, the 564 × 564 pixel patches were furtherly resized by the Gaussian pyramid approach [17] to 282 × 282 pixels.

Training of the CNN by using the raw data without reinforcing the desirable features leads to non-acceptable training accuracy. Therefore, the special features of the thyroid carcinoma in the patches should be enhanced in the pre-processing stage before starting the training phase. To do that, it is necessary to use an un-sharp mask to enhance cell edges (desirable features) and consistency of subtracting an un-sharp (smoothed) version of a patch from the original patch to get enhanced edges in the patches [18].

In trying to generalize the features of images to cover each case in various circumstances, the database was augmented by four different techniques [19] as follows:

- Adding Salt and Pepper noise to the original image with noise density 0.05.
- Rotating the original image by 45° counterclockwise around its center point.
- Flipping the original image in the left-right direction around the vertical axis.
- Flipping the original image in the up-down direction around the horizontal axis.

Fig 3 shows a sample of the augmentation process.

After augmenting the database, it is necessary to split it into three parts, 60% for training, 25% for validation, and 15% for testing.

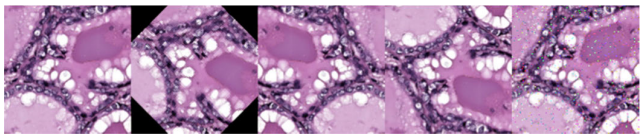

FIGURE 3. Example of an image after augmentation, from left to right, original image, rotation, horizontal flip, vertical flip, and salt and pepper noise.

B. PROPOSED CNN ARCHITECTURE

Our proposed CNN consists of the following layers:

1) The first layer is the input stage for the convolution neural network.

2) The convolution layer; in this layer, features extraction is done by different filters. Sizes of filters, number filters, Stride, and zero paddings are determined in this layer and subsequently illustrated.

3) Batch normalization layer (Batch Norm.).

4) Rectified linear unit.

5) Max pooling.

6) From layers 2 to 5, represent the feature extraction group. The feature extraction level depends on the number of repetitions of this group; for example, if the feature is superficial, it does not require to be repeated. However, if the feature is deep like our case, it should be repeated. In our proposed network, this group is repeated three times.

7) Full connected layer.

8) Rectified linear unit.

9) DROPOUT LAYER.

There are 24 layers in our proposed network, and its complete construction is shown in Fig 4. Each layer in the CNN has several learnable parameters to solve while training. The first layer starts to collect the 282 × 282 colored images from the database.

Extracting the input images' features is the next stage done by the Convolution layers (2, 6, and 10) by applying several kernels (filters). Each kernel has the same size and moves across the image with steps horizontally and vertically which are called stride. The low-level features are detected only by the Kernels in the early layers, while Kernels detect complex features in the network's advanced convolution layers. [20]. So, our convolution group is repeated three times to extract the deep features in the patches of the WSI's.

FIGURE 4. The full construction of the CNN.

Each convolution layer is followed by an activation ReLU function that is generally used to minimize the time of training compared to other activation functions [21]–[23].

Batch Normalization is performed after the activation layer to remove covariate shift and minimize training time. Batch Normalization minimizes the effects of various gradients because every data becomes nearly normally distributed [24].

Down sampling image size is the following function performed by the max-pooling layer. The max-pooling outputs the maximum value in the input region [20].

Preparing the data for classification is the penultimate function of CNN that is performed by a fully connected layer. The output feature maps of the final convolution or pooling layer are typically flattened a one-dimensional (1D) array of numbers (or vector). The final fully connected layer typically has the same number of output nodes as the number of classes. Classification is the final function of the CNN that is performed by the SoftMax Classifier layer takes an input vector (fully connected layer), and it generates the output vector in the range [0-1], where its elements add up to one. SoftMax function must follow the final fully connected layer, and it is given by [20]:

$$S(y_i) = \frac{e^{y_i}}{\sum_j y_j} \tag{1}$$

where y_i is the output of the classifier, e is the error.

TABLE 2. Training parameters.

Parameter(s)	Values	Best Values
Number of convolutional + ReLU layers + pooling layer	1, 2, 3	**3**
Batch normalization layers	I, 2, 3	**3**
Number of drop out layers	1, 2, 3	**2**
Maximum epochs	10,16,32,64,128	**32**
Number of fully connected layers	1, 2, 3	**2**
Number of convolutional kernels	16, 32, 64, 128	**---**
Kernel size	3, 5, 7, 9, 10,11	**---**
Pooling layer window size	2, 3, 4, 5	**3**
Optimizers	SGD, Adam, RMSprop	**SGD**
Mini-batch size	16, 32, 64, 128	**32**
Initial learning rate	0.01, 0.001, 0.0001	**0.001**
Learning rate drop factor	**0.1, 0.2, 0.3,0.5**	**0.5**

TABLE 3. Confusion matrix for testing.

True Class \ Predicted Classes	Glioma	Meningioma	Pituitary
Glioma	1034	32	4
Meningioma	12	512	6
Pituitary		10	685

TABLE 4. Comparing the testing result with related works.

	Model	Accuracy	Method
1	Cheng, et al. [27]	91.28%	SVM and KNN
2	Paul, et al. [28]	91.43%	CNN
3	Afshar, et al. [29]	90.89%	CNN
4	Anaraki, et al. [30]	94.2%	GA-CNN
5	**Proposed structure**	**98.14%**	**CNN**

C. CASCADED CONVOLUTIONAL NEURAL NETWORK

In this research, two cascaded CNN are used having the same architecture as shown in Fig 1. The first CNN is to classify the main three thyroid carcinoma (Follicular adenoma, Follicular carcinoma, and papillary carcinoma), then the second CNN classifies the follicular carcinoma into four sub-classes (Follicular carcinoma, Papillary-follicular variant, Well-differentiated follicular carcinoma, and Poorly-differentiated follicular carcinoma).

D. TRAINING AND OPTIMIZATION ALGORITHM

Training of Convolutional neural networks involves solving model parameters using the labeled database to allow the network to map an input image to a class label [25].

TABLE 5. Main classes confusion matrix.

True Class \ Predicted Classes	FC	FA	PC
FC	6480	62	68
FA	32	1468	
PC	103		5777

TABLE 6. Sub classes confusion matrix.

True Class \ Predicted Classes	FC	PFV	PDFC	WDFC
FC	2043	65	12	15
PFV	124	2654		12
PDFC	5		599	1
WDFC	26	12		982

A cost function is used to calculate the error (the difference between ground truth label and prediction) during a convolutional neural network training process. The Mean Squared Error (MSE) loss function is suggested in this work as it calculates the squared average error E of all the individual errors, and it is given by [25]:

$$E = \frac{1}{n}\sum_{i=1}^{n} e_i^2 \tag{2}$$

where e_i is:

$$e_i = target\ (i) - output(i) \tag{3}$$

The training is trying to optimize an equation parameter of the classification problem.

Gradient descent is an optimization technique used to minimize the cost function by determining gradients necessary to update the parameter values of the convolution neural network. Stochastic Gradient Descent (SGD) is known as incremental gradient descent, where the gradient is calculatedPm at a time followed by parameter values updating. It applies one update at a time [25].

$$w = w - \mu.\nabla E(W; x\ (i)\ ; y(i)) \tag{4}$$

where: ∇E (W;x(i); y(i)) is the gradient of loss function. W is the weight for the training example (x(i). y(i)). μ is the learning rate.

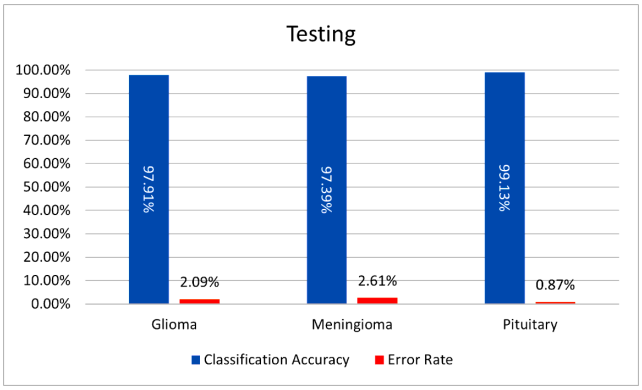

FIGURE 5. Testing accuracy.

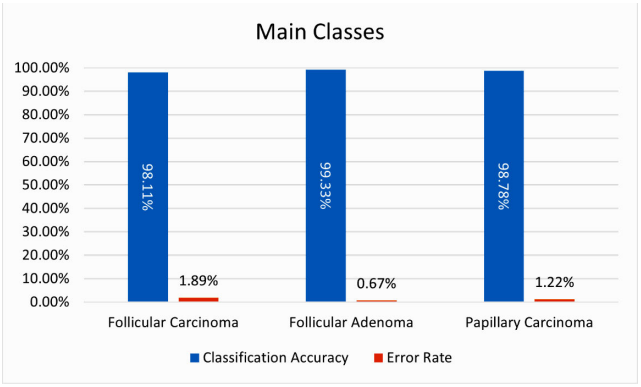

FIGURE 6. Accuracy of main-classes classification.

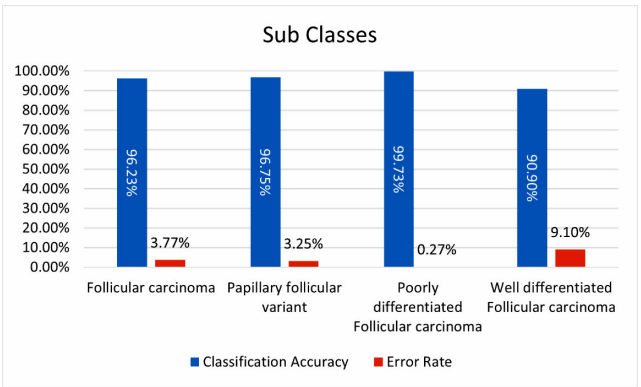

FIGURE 7. Accuracy of subclasses classification.

Training is used to update CNN parameters and minimize loss function to reach the global minimum in the ideal case by taking small steps to the negative gradient direction [26]. Table 2 shows all the tested and best parameters which are used to construct our proposed CNN.

III. EXPERIMENT'S RESULTS

Training and testing of the proposed CNN are done on a workstation computer with Intel®Xeon®Processor W3690 (12M Cache, 3.46 GHz, 6.40 GT/s Intel® QPI) with RAM

TABLE 7. Result analysis.

	Main Classes			Sub Classes			
Tumor Type	Follicular Carcinoma	Follicular Adenoma	Papillary Carcinoma	Follicular carcinoma	Papillary follicular variant	Poorly-differentiated Follicular carcinoma	Well-differentiated Follicular carcinoma
TP	6480	1468	5777	2043	2654	599	982
TN	7245	12428	8042	4260	3683	5933	5502
FP	135	62	68	155	77	12	611
FN	130	32	103	92	136	6	38
Accuracy	98.11%	99.33%	98.78%	96.23%	96.75%	99.73%	90.90%
Error Rate	1.89%	0.67%	1.22%	3.77%	3.25%	0.27%	9.10%
Sensitivity	98.03%	97.87%	98.25%	95.69%	95.13%	99.01%	96.27%
Specificity	98.17%	99.50%	99.16%	96.49%	97.95%	99.80%	90.00%
False Positive Rate	1.83%	0.50%	0.84%	3.51%	2.05%	0.20%	10.00%
False Negative Rate	1.97%	2.13%	1.75%	4.31%	4.87%	0.99%	3.73%
Precision	97.96%	95.95%	98.84%	92.95%	97.18%	98.04%	61.64%
Negative Predictive Value	98.24%	99.74%	98.74%	97.89%	96.44%	99.90%	99.31%
F1-Score	98.00%	96.90%	98.54%	94.30%	96.14%	98.52%	75.16%
Total no.	6610	1500	5880	2135	2790	605	1020

TABLE 8. Proposed system comparison with other related work.

No.	Method	Accuracy		Sensitivity	Specificity	Tissue type
1	Guan, et al. [5]	97.66%		100%	94.91%	Cytopathology Images
2	Wang, et al. [7]	88.33%		--	--	Histopathology images
3	Dov, et al. [8]	93.20%		--	--	WSI cytopathology
4	Elliott Range, et al. [9]	92.00%		--	--	WSI cytopathology
5	Tao, et al. [10]	53.40%		--	--	WSI cytopathology
6	Jamil Ahmed Chandio [32]	74%		72.53%	--	Cytopathology Images
7	J and V [33]	98.18%		98.75%	91.42%	Histopathology images
8	Halicek, et al. [34]	96.4		96.80%	96.10%	Histopathology images
9	Purposed Method Main Classes	**98.74%**	**94.69%**	98.05%	98.94%	WSI Histopathology
10	Purposed Method Sub-Classes	**95.90%**		96.53%	96.06%	WSI Histopathology

of 16GB and operating system Ubuntu 18.04.5 LTS 64-bit and MATLAB R2019b. The average training time for the proposed CNN architecture of the training and validation parts of the whole Database is 50 hours and 30 minutes.

Since our generated database and the proposed CNN were not previously tested, it is suggested to start to test the proposed CNN on "Brain Tumor Dataset v5" which contains 3064 T1-weighted contrast-enhanced images with three kinds of a brain tumor [27].

The obtained results are shown in Table 3 and Fig 5 achieving an overall accuracy of 98.14%. Table 4 compares the performance of the proposed CNN with different related methods from the literature showing that it has the best performance.

In our experiments, it is suggested to split the classification process into two cascaded CNN, and the first one classifies the dataset into three types of thyroid carcinoma (Follicular Adenoma, Follicular Carcinoma, and Papillary Carcinoma). Then the true positive output of the follicular carcinoma class is further classified into four sub-classes (Follicular Carcinoma, Follicular Papillary Variant, Well Differentiated Follicular Carcinoma, and Poorly-differentiated Follicular Carcinoma). The total number of labeled patches used for this training set is 18,653 images with a dimension of 564 × 564.

The overall accuracy of the main classes over the first CNN is 98.74%, as shown in Fig 6 and Table 5. The Follicular carcinoma is classified with an overall accuracy of 98.0%, and furtherly classified into four sub-classes through the second CNN with an overall accuracy of 95.90% as shown in Fig 7 and in Table 6. This yields an overall accuracy of 94.69% for the cascaded CNN.

A confusion matrix is a representation table for evaluating classifier performance. The matrix consists of the following terms:

1) True positives (TPs): Number of instances identified as positive from the classifier.

2) False positives (FPs): Number of instances identified as positive from the classifier, but they are negative.

3) True negatives (TNs): Number of instances identified as negative from the classifier, and they are negative as well.

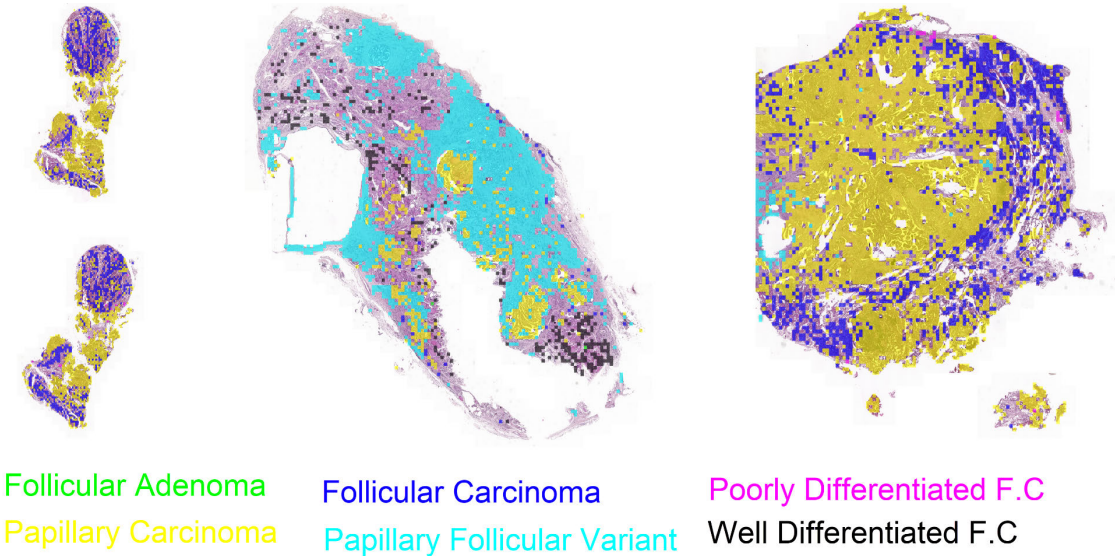

FIGURE 8. Result of the validation (down sampled), three different WSI with different color areas.

4) False negatives (FNs): Number of instances identified as negative from the classifier, and they are positive.

In general, the confusion matrix of a classifier should have a greater number of TP and TN and a smaller number of FP and FN.

Using the confusion matrix, other metrics can be used in the evaluation of the model. These metrics are defined as in Table 7 and according to the following equations [31]:

$$
\begin{aligned}
Accuracy &= \frac{(TP+TN)}{(TP+FN+FP+TN)} \\
Error\ Rate &= \frac{(FP+FN)}{(TP+FN+FP+TN)} \\
Sensitivity &= \frac{TP}{(TP+FN)} \\
Specificity &= \frac{TN}{(TN+FP)} \\
False\ Positive\ Rate &= \frac{FP}{(FP+TN)} \\
False\ Negative\ Rate &= \frac{FN}{(FNTP)} \\
Precisionor\ Positive\ Predictive\ Value &= \frac{TP}{(TP+FP)} \\
Negative\ Predictive\ Value &= \frac{TN}{(TN+FN)} \\
F1-Score &= \frac{(2\times(Sensitivity\ x\ Precision))}{(Sensitivity+Precision)}
\end{aligned}
\tag{5}
$$

IV. DISCUSSION

Some considerations should be taken seriously in building an automated classification system for thyroid carcinoma in histology whole slide images. Firstly, images should be captured by 40× magnification because the features of the cells will be clearer that minimizes the use of image pre-processing techniques. Secondly, the patching process allows us to computationally process the whole slide image. This process should be carefully done to prevent data loss or damage, so the patching process is suggested to be done/managed/fixed by 512 × 512, overlapping 5%.

Reinforcing the desired features is an important stage for training the CNN, so it is suggested to enhance the cells' edges using an Un-sharp filter. This process reflected on our achieved results.

As the CNN is built from scratch with 24 layers, it is tested on a brain MRI database image "brain tumor dataset v5" that gives an accuracy of 98.0%. Hence, it helps us to be sure of using this proposed CNN to classify our 24 whole slide images.

Moreover, the classified patches are validated from our proposed CNN with the original consultant decision for the whole slide images of the 24 cases as follows; each class of the classified patches is given a color code, then it is suggested to resize the patch to 50 × 50. The downsized and colored patches are recombined again into new down-sampled WSI. Consequently, this downscaled version of the WSI became colored according to each available class's relative location in the original WSI, as shown in Fig 8. According to the dominant area of the major class in the slide, it can be deduced that this matches with the consultant diagnosis opinion by 100% for the whole 24 WSI's.

V. CONCLUSION

In this work, a cascaded CNN system is presented to classify thyroid carcinoma based on whole slide images. In the first CNN, thyroid carcinoma is classified into three main types: Follicular adenoma, Follicular carcinoma, and papillary carcinoma. In the second CNN, it is suggested to further classify the follicular carcinoma into four sub-classes which are: Follicular carcinoma, papillary follicular variant, well-differentiated follicular carcinoma, and

poorly-differentiated follicular carcinoma. We have used a custom deep neural network structure. The proposed network is constructed from 24 layers. To assess the proposed CNN, we have tested it on a brain MRI database image, ''brain tumor dataset v5'', which achieved an overall accuracy of 98.0%. Although the dataset is relatively not big, data augmentation helped well to show better results and hence overcome this problem. Our proposed architecture has achieved the highest accuracy of 94.69%.

FUTURE WORK

Although we have included six subclasses of thyroid carcinoma in our proposed system, we aim in the future to extend our dataset to include other thyroid carcinoma classes which did not appear in the current dataset (e.g., Medullary thyroid carcinoma, Hürthle cell thyroid carcinoma, and Anaplastic thyroid carcinoma).

REFERENCES

[1] N. Stathatos, ''Anatomy and physiology of the thyroid gland,'' in *The Thyroid its Diseases: A Comprehensive Guide for Clinician*, M. Luster, L. H. Duntas, and L. Wartofsky, Eds. Cham, Switzerland: Springer, 2019, pp. 3–12.

[2] L. H. Duntas and S. Tseleni-Balafouta, ''Classification of thyroid diseases,'' in *The Thyroid its Diseases: A Comprehensive Guide for Clinician*, M. Luster, L. H. Duntas, and L. Wartofsky, Eds. Cham, Switzerland: Springer, 2019, pp. 87–99.

[3] M. P. J. Vanderpump, ''Epidemiology of thyroid disorders,'' in *The Thyroid its Diseases: A Comprehensive Guide for Clinician*, M. Luster, L. H. Duntas, and L. Wartofsky, Eds. Cham, Switzerland: Springer, 2019, pp. 75–85.

[4] E. Kim, M. Corte-Real, and Z. Baloch, ''A deep semantic mobile application for thyroid cytopathology,'' *Proc. SPIE*, vol. 9789, Apr. 2016, Art. no. 97890A.

[5] Q. Guan, Y. Wang, B. Ping, D. Li, J. Du, Y. Qin, H. Lu, X. Wan, and J. Xiang, ''Deep convolutional neural network VGG-16 model for differential diagnosing of papillary thyroid carcinomas in cytological images: A pilot study,'' *J. Cancer*, vol. 10, no. 20, p. 4876, 2019.

[6] C.-W. Wang, Y.-C. Lee, E. Calista, F. Zhou, H. Zhu, R. Suzuki, D. Komura, S. Ishikawa, and S.-P. Cheng, ''A benchmark for comparing precision medicine methods in thyroid cancer diagnosis using tissue microarrays,'' *Bioinformatics*, vol. 34, no. 10, pp. 1767–1773, May 2018.

[7] Y. Wang, Q. Guan, I. Lao, L. Wang, Y. Wu, D. Li, Q. Ji, Y. Wang, Y. Zhu, H. Lu, and J. Xiang, ''Using deep convolutional neural networks for multi-classification of thyroid tumor by histopathology: A large-scale pilot study,'' *Ann. Transl. Med.*, vol. 7, no. 18, p. 468, Sep. 2019.

[8] D. Dov, S. Z. Kovalsky, J. Cohen, D. E. Range, R. Henao, and L. Carin, ''Thyroid cancer malignancy prediction from whole slide cytopathology images,'' presented at the Proc. 4th Mach. Learn. Healthcare Conf., Mach. Learn. Res., 2019. [Online]. Available: http://proceedings.mlr.press

[9] D. D. Elliott Range, D. Dov, S. Z. Kovalsky, R. Henao, L. Carin, and J. Cohen, ''Application of a machine learning algorithm to predict malignancy in thyroid cytopathology,'' *Cancer Cytopathology*, vol. 128, no. 4, pp. 287–295, Apr. 2020, doi: 10.1002/cncy.22238.

[10] S. Tao, Y. Guo, C. Zhu, H. Chen, Y. Zhang, J. Yang, and J. Liu, ''Highly efficient follicular segmentation in thyroid cytopathological whole slide image,'' in *Precision Health and Medicine: A Digital Revolution in Healthcare*, A. Shaban-Nejad, and M. Michalowski, Eds. Cham, Switzerland: Springer, 2020, pp. 149–157.

[11] A. Gupta, P. J. Harrison, H. Wieslander, N. Pielawski, K. Kartasalo, G. Partel, L. Solorzano, A. Suveer, A. H. Klemm, O. Spjuth, I. Sintorn, and C. Wählby, ''Deep learning in image cytometry: A review,'' *Cytometry A*, vol. 95, no. 4, pp. 366–380, Apr. 2019.

[12] G. Forslid, H. Wieslander, E. Bengtsson, C. Wahlby, J.-M. Hirsch, C. R. Stark, and S. K. Sadanandan, ''Deep convolutional neural networks for detecting cellular changes due to malignancy,'' in *Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2017, pp. 82–89.

[13] *Pannoramic SCAN II 3DHISTECH Ltd.* Accessed: Nov. 2, 2020. [Online]. Available: https://www.3dhistech.com/products-and-software/hardware/pannoramic-digital-slide-scanners/pannoramic-scan-2/

[14] S. Erle, R. Gibson, and J. Walsh, *Mapping Hacks: Tips & Tools for Electronic Cartography*. Newton, MA, USA: O'Reilly Media, 2005.

[15] D. Komura and S. Ishikawa, ''Machine learning methods for histopathological image analysis,'' *Comput. Struct. Biotechnol. J.*, vol. 16, pp. 34–42, Jan. 2018. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6158771/pdf/main.pdf

[16] C. Deroulers, D. Ameisen, M. Badoual, C. Gerin, A. Granier, and M. Lartaud, ''Analyzing huge pathology images with open source software,'' *Diagnostic Pathol.*, vol. 8, no. 1, p. 92, Dec. 2013, doi: 10.1186/1746-1596-8-92.

[17] A. Distante and C. Distante, ''Data organization,'' in *Handbook of Image Processing and Computer Vision: From Energy to Image*, vol. 1. Cham, Switzerland: Springer, 2020, pp. 317–339.

[18] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. London, U.K.: Pearson, 2018.

[19] S. C. Wong, A. Gatt, V. Stamatescu, and M. D. McDonnell, ''Understanding data augmentation for classification: When to warp?'' in *Proc. Int. Conf. Digit. Image Comput., Techn. Appl. (DICTA)*, Nov. 2016, pp. 1–6.

[20] M. A. Wani, F. A. Bhat, S. Afzal, and A. I. Khan, ''Basics of supervised deep learning,'' in *Advances in Deep Learning*. Singapore: Springer, 2020, pp. 13–29.

[21] A. Krizhevsky, I. Sutskever, and G. E. Hinton, ''ImageNet classification with deep convolutional neural networks,'' in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2012, pp. 1097–1105.

[22] S. Ioffe and C. Szegedy, ''Batch normalization: Accelerating deep network training by reducing internal covariate shift,'' in *Proc. Int. Conf. Mach. Learn.* PMLR, 2015, pp. 448–456.

[23] V. Nair and G. E. Hinton, ''Rectified linear units improve restricted Boltzmann machines,'' in *Proc. ICML*, 2010, pp. 1–8.

[24] M. A. Wani, F. A. Bhat, S. Afzal, and A. I. Khan, ''Supervised deep learning in fingerprint recognition,'' in *Adv. Deep Learning. Singapore: Springer Singap.*, vol. 2020, pp. 111–132.

[25] M. A. Wani, F. A. Bhat, S. Afzal, and A. I. Khan, ''Training supervised deep learning networks,'' in *Advances in Deep Learning*. Singapore: Springer, 2020, pp. 31–52.

[26] L. Bottou, ''Large-scale machine learning with stochastic gradient descent,'' in *Proc. COMPSTAT*. Springer, 2010, pp. 177–186.

[27] J. Cheng, W. Huang, S. Cao, R. Yang, W. Yang, Z. Yun, Z. Wang, and Q. Feng, ''Enhanced performance of brain tumor classification via tumor region augmentation and partition,'' *PLoS ONE*, vol. 10, no. 10, Oct. 2015, Art. no. e0140381.

[28] J. S. Paul, A. J. Plassard, B. A. Landman, and D. Fabbri, ''Deep learning for brain tumor classification,'' *Proc. SPIE*, vol. 10137, Mar. 2017, Art. no. 1013710.

[29] P. Afshar, K. N. Plataniotis, and A. Mohammadi, ''Capsule networks for brain tumor classification based on MRI images and coarse tumor boundaries,'' in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 1368–1372.

[30] A. Kabir Anaraki, M. Ayati, and F. Kazemi, ''Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms,'' *Biocybern. Biomed. Eng.*, vol. 39, no. 1, pp. 63–74, Jan. 2019.

[31] M. Z. Hasan, K. T. Rahman, R. I. Riya, K. M. Z. Hasan, and N. Zahan, ''A CNN-based classification model for recognizing visual bengali font,'' in *Proc. Int. Joint Conf. Comput. Intell.* Singapore: Springer, 2020, pp. 471–482.

[32] J. Ahmed and M. Abdul, ''Intelligent diagnostic system for nuclei structure classification of thyroid cancerous and non-cancerous tissues,'' *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 7, 2017, doi: 10.14569/ijacsa.2017.080746.

[33] A. A. J. Jothi and M. A. Rajam, ''Automatic classification of thyroid histopathology images using multi-classifier system,'' *Multimedia Tools Appl.*, vol. 76, no. 18, pp. 18711–18730, Sep. 2017, doi: 10.1007/s11042-017-4363-0.

[34] M. Halicek, G. Lu, J. V. Little, X. Wang, M. Patel, C. C. Griffith, M. W. El-Deiry, A. Y. Chen, and B. Fei, ''Deep convolutional neural networks for classifying head and neck cancer using hyperspectral imaging,'' *J. Biomed. Opt.*, vol. 22, no. 6, 2017, Art. no. 60503, doi: 10.1117/1.Jbo.22.6.060503.

AHMED S. EL-HOSSINY was born in Cairo, Egypt, in 1984. He received the B.Sc. degree from the Biomedical and Systems Engineering Department, Higher Institute of Engineering, in 2006, and the M.Sc. degree from the Department of Biomedical Engineering, Cairo University, in 2015. He is currently pursuing the Ph.D. degree with the Department of Biomedical Engineering, Helwan University, Egypt. He is currently a Teacher Assistant with the Higher Institute of Engineering. His research interests include image processing and machine learning for medical diagnostics.

WALID AL-ATABANY received the B.Sc. and M.Sc. degrees from the Department of Biomedical Engineering, Cairo University, in 1999 and 2004, respectively, and the Ph.D. degree in biomedical engineering from Imperial College London, in 2010. In 2011, he was a Research Associate with Newcastle University for two years. He is currently an Associate Professor with the Department of Biomedical Engineering, Helwan University, Egypt. His research interests include signal and image processing (particularly for the visually impaired), medical imaging, and modeling retinal processing. He received the second Price Award from the 2nd Symposium of the Neuroscience Technology Network (NTN2009), the ARVO 2010 Travel Grant from the AFER/National Institute for Health Research Biomedical Research Centre for Ophthalmology, and the Grant from the Newton Institutional Link Grant from the British Council, in 2015 and 2016.

OSAMA HASSAN received the M.B.B. degree from the Faculty of Medicine, Cairo University, in 1967, and the Ph.D. degree from Manchester University, England, in 1978. He is currently a National Expert of surgical and diagnostic pathology and the Director of pathology services with the International Medical Center (IMC), Cairo. He is the formal Head of Biomedical and Systems Engineering Department, Higher Institute of Engineering, El Shorouk Academy. His research interests include image analysis and quantitative pathology, digital pathology and telemedicine, quality assurance, and safety in healthcare.

AHMED M. SOLIMAN received the B.Sc. (Hons.), M.S., and Ph.D. degrees in biomedical engineering from Helwan University, Cairo, Egypt, in 2003, 2010, and 2017, respectively, and the second M.S. degree in biotechnology engineering from the University of Chemical Technology and Metallurgy (UCTM), Sofia, Bulgaria, in 2009. From 2010 to 2012, he was an Exchange Ph.D. Student with Atomic Physics Division, Lund Medical Laser Centre, Biophotonics Group, Physics Department, Faculty of Engineering (LTH), Lund University, Lund, Sweden.

He is currently an Assistant Professor with the Department of Biomedical Engineering, Faculty of Engineering, Helwan University. He is also a Technical Examiner Member with Egyptian Patent Office. His research interests include surface acoustic wave devices, microfluidics, biosensors, biotechnology engineering, medical optics, biomedical devices, neural networks, deep learning, and modeling and simulations in biomedical applications.

Dr. Soliman received the fellowship for his M.Sc. degree and exchange Ph.D. from Erasmus Mundus External Cooperation Window (EMECW) Program, in 2007 and 2010, respectively. In 2012, he received a fellowship for his Ph.D. studies from the Swedish Institute Scholarship, Guest Scholarship Program, Sweden.

SHERIF A. SAMI received the B.Sc., M.Sc., and Ph.D. degrees in systems and biomedical engineering from the Systems and Biomedical Engineering Department, in 1988, 1993, and 2000, respectively.

He is the formal Assistant Director of Center of Advanced Software and Biomedical Engineering, Cairo University, from 2000 to 2004, the formal Director of Medical Equipment Calibration Laboratory (MECL), Cairo University, from2000 to 2014, and a Technical Consultant for Ministry of Health and Ministry of Justice, Egypt. He is currently an Assistant Professor with Systems and Biomedical Engineering Department, Faculty of Engineering, Cairo University, and the Head of systems and biomedical engineering, Higher Institute of Engineering, El Shorouk Academy. His research interests include testing, evaluation, and development of medical equipment, medical image processing and analysis, and computer vision.