Smart Endoscope System

Ahmed Abdel Fattah El-said, Ahmed Abdel Hakim Abdel Zahir, Ahmed Reda Abdel Rashid, Ali Mohammad Ali El-Shamy, Aida El-Morsi Farhat, and Shrouk Mahmoud Fathy Mahmoud

Zagazig University, Egypt,ahmedabdoahmed007@gmail.com, ahmedhakimpro@gmail.com, redaa1906@gmail.com, alielshamy6000@gmail.com ,aidanoser1999@gmail.com, shroukmahmoud0co@gmail.com.

Supervisor: Nissreen El-Saber, PhD

Zagazig University, Egypt, naelsaner@fci.zu.edu.eg

Abstract*– As a result of COVID-19 pandemic, doctors face great challenges regarding the operations long duration (e.g. the digestive system ones) and patients face the long waiting list. Unfortunately, most digestive system diseases take time to diagnose. We propose a solution to help clinicians and their patient's diagnosis processes by integrating Artificial Intelligence into medical devices like the endoscope, i.e. a device that enters the human body through its opening ends (e.g. mouth) to allow for a more detailed internal investigation. Smart endoscopes are supported by cameras which are designed to recognize the abnormal tissues in the stomach. Cameras can record certain scenes or take photos of tissues directly to the application. This allows the doctor to explain the observations to the patient and even colleagues. The proposed system supports automatic detection of abnormalities reducing the diagnosis error rate. Therefore, the medical team will be able to conduct a more precise assessment of the internal organs and contribute toward a more accurate diagnosis for the patient's condition. The proposed system takes one step closer towards the advancement of medical technology allowing faster diagnosis reducing the number of misdiagnosed patients or late diagnose due to the long waiting lists.

Keywords—Medical, Endoscope, AI, Deep Learning, U-Net

I. Introduction

Following the government view and strategies to shorten the surgeries waiting lists, the operations rooms are facing great pressure in all hospitals in Egypt. The situation is worsening during the COVID-19 pandemic. As a result, doctors face great challenges regarding the long duration of the operation (especially the digestive system operations " gastrointestinal"). Unfortunately, most digestive system diseases take time to diagnose.

An endoscope is a device that enters the human body through its opening ends (e.g. mouth) to allow for a more detailed internal investigation in digestive system diseases diagnose and treatment. We propose an intelligent update for the endoscope for automatic detection of diseases based on deep learning algorithms. It allows the doctors and researchers to develop systems that improve the health-care system in the context of disease detection in images or video. The ability to detect diseases varies between doctors. Accurate grading of diseases is important since it may influence decision-making on treatment and follow-up. The doctor controls the endoscope to take images and videos during the endoscopy.

Proposing this work, we are looking forward to propose helpful, and supportive intelligent tool to doctors and medical authorities to use worldwide. The system flow chart is presented in Fig. 1 below.

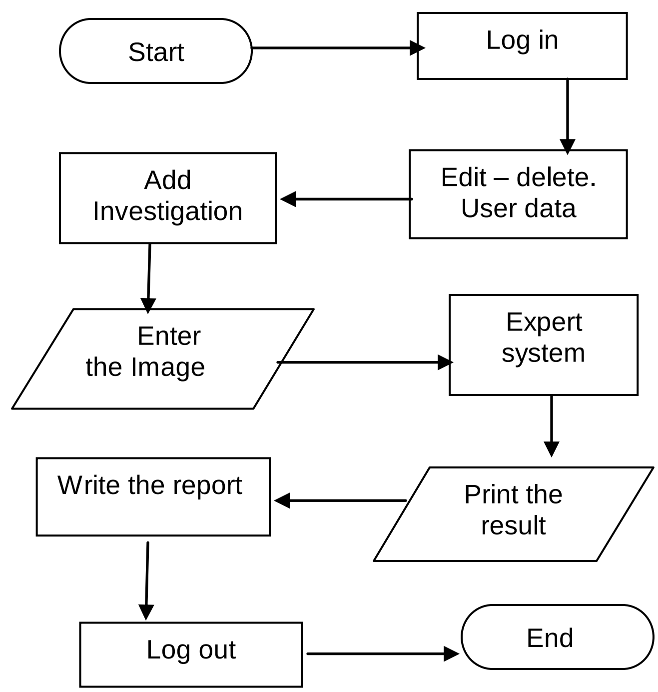


Fig. 1 Proposed System Flow Chart.

The rest of the paper is organized as follows: Section II presents essential preliminaries, while Section III introduces the proposed system requirement elicitation process, followed by the main design used in Section IV. In Section V shows the used tools and technologies are introduced with essential details on artificial intelligence used. Section VI shows sample of our results and discussion. Finally, the conclusion and future work are highlighted in Section VII.

II. PRELIMINARIES

In this section, we introduce some basic concepts for our work. First, the System Development Life Cycle (SDLC) and second the Deep Learning technique.

A. SDLC

The System Development Life Cycle (SDLC) is a widely used concept for projects steps. Building an information system using the SDLC follows a set of four fundamental phases: planning, analysis, design, and implementation. Each phase is itself composed of a series of steps, which rely on techniques that produce deliverables (specific documents and files that explain various elements of the system). The output of one phase is and input to the next. For space reasons, we will introduce a brief of the Analysis, Design and Implementation phases in Sections III, IV and V.

B. Deep Learning

Deep learning is an artificial intelligence (AI) function that imitates the workings of the human brain in processing data and creating patterns for use in decision making. Deep learning is a subset of machine learning in artificial intelligence that has networks capable of learning unsupervised from data that are unstructured or unlabeled [1]. It utilizes a hierarchical level of artificial neural networks to carry out the process of machine learning. The artificial neural networks are built like the human brain, with neuron nodes connected together like a web. While traditional programs build analysis with data in a linear way, the hierarchical function of deep learning systems enables machines to process data with a nonlinear approach. The used semantic segmentation architecture in this work is U-Net [1]. It consists of a contracting path and an expansive path. The contracting path follows the typical architecture of a convolutional network.

III. SYSTEM REQUIREMENTS ELICITATION

The SDLC Analysis is used in requirements gathering and documenting. Doctors, professional and medical consultant have been interviewed to gather the knowledge base and the most suitable way of providing the intelligence to the endoscopes. The gathered requirements are divided into functional and non-functional requirements. As names imply, functional requirements are the main functions the system will perform either process oriented or information oriented functions, while non-functional requirements are the characteristics the system should have in order to perform its functions (i.e. operational, performance, security, and cultural and political characteristics).

The process-oriented functional requirements include: (i) The system allows administrative support to login using username and password, (ii) The system allows administrative support to add user data, (iii) The system allows doctors to access the endoscope images, and (iv) The system can document the diagnosis report and offer more insights about a patient's status and make better data-driven decisions using machine learning algorithms and AI. The information-oriented functional requirements can be summarized as: (i) The system has dataset for gastrointestinal diseases, (ii) The system must retain history for patient, and (iii) The system dataset is updated every three years.

Operationally, the non-functional requirements are: (i) The system runs web and desktop applications and (ii) The system can integrate with hospitals Systems. On performance, (i) The system should be fast and accurate in processing and presenting the results, and (ii) The system can handle many datasets simultaneously. On security, (i) The system must be safe and trusted for login and logout, and (ii) Database, the information and the reports must be secured and not compromised. On cultural and political side: (i) The system operates in English language, and (ii) Patient information is not allowed for anyone and following Data Privacy laws.

Use Case Diagrams provides overview of the functional requirements for a system. They are useful for presentations to management, project stakeholders and/or developers. The figure is represented in Fig. 2.

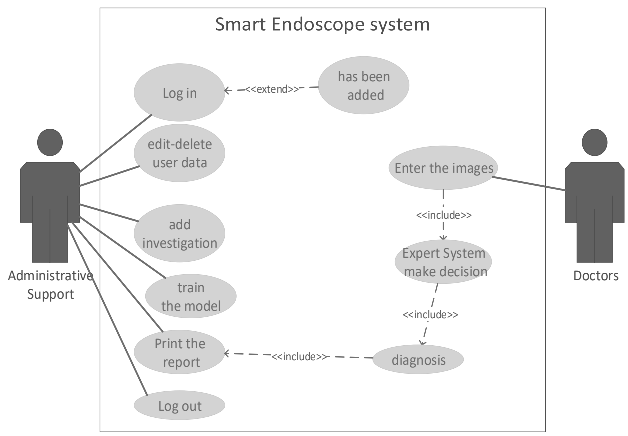


Fig. 2 Proposed System Use Case Diagram.

For a more detailed process model to represent the main processes, their inputs and outputs, we used Data Flow Diagrams (DFD) where the main data flows are depicted as in Fig. 3.

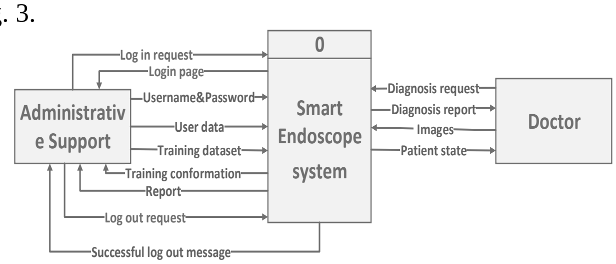


Fig. 3 Proposed System Context DFD.

DFD are based on decomposition, where the system can be decomposed into hierarchical levels of process details giving a higher level of details. Level 0 DFD is presented in Fig. 4.The consistency in systems analysis can be seen in the resemblance between use cases and processes in the DFD Level 0.

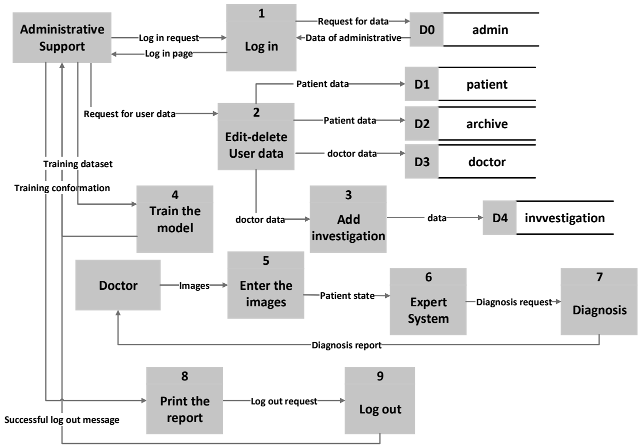


Fig. 4 Proposed System Level 0 DFD

IV. Proposed System Design

The purpose of the analysis phase in SDLC is to figure out what the business needs. The purpose of the design phase is to decide how to build it [2]. For this, the following is divided into two sections: first, the deep learning model design, and second, the application layout screens.

A. Deep Learning Model Design

The deep learning model follows the proess in Fig. 5. The four steps are data preparation, model specifications, train the network algorithm, and predict the new image matches.

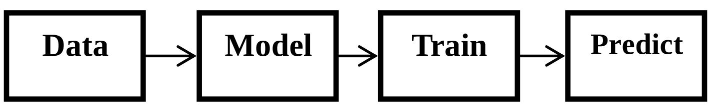


Fig. 5 Deep Learning main steps

Data: is a set of data that contains images of the disease and through which these images are passed to the algorithm for testing. Model: The algorithm we followed, which then performs operations on the image, such as resizing it. Train: is the process that checks the efficiency of the algorism and gives sufficient information to teach the machine. Predict: is the final process, in which the god has been taught and has the experience to identify the disease. After the picture is cleared, the place of injury that the patient suffers from appears

B. Application Layout

The following are the screens layout for the application. The application contains a homepage, shown in Fig. 6, which includes logging into the system by entering the user's username and the password correctly because if any of them are entered an error, error message will appear.

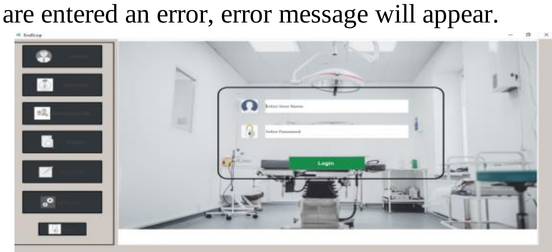


Fig. 6 Homepage screen

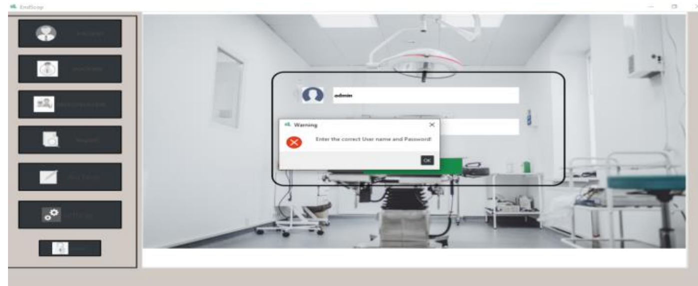


Fig. 7 Patient Screen

In Fig. 7, the patient information screen is shown. Through it, one can search for a specific patient by the ID number. More than one patient can be retrieved in an Excel file. Patient details can be exported in an Excel file as shown in Fig. 8. Patient information are entered as shown in Fig. 9.

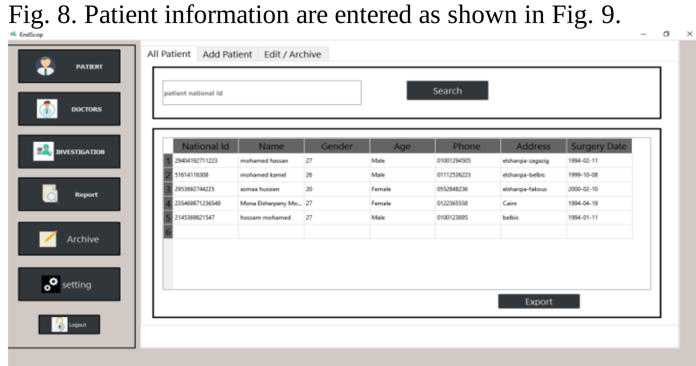


Fig. 8Retrieve Patients Screen

Fig. 9 Add Patient Screen

Fig. 10 shows the achieving patient screen, where patient information are not deleted, but they are archived for future retrieval. Patient national ID is used to look patients up and then information are retrieved from the database and can be updated. The deletion will take patient data to the archive.

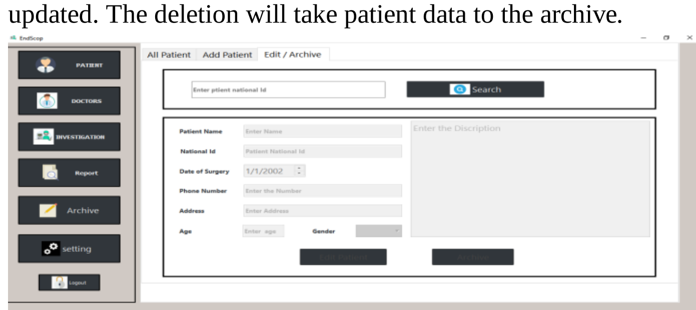


Fig. 10Archiving Patient Data Screen

For doctors information, Fig. 11 shows the Add New Doctor screen. Through it, I can search for a specific doctor by their ID number, add new doctor, delete (i.e. archive) doctors in case they left the medical institution. A list of available doctors can be retrieved and saved in an Excel file.

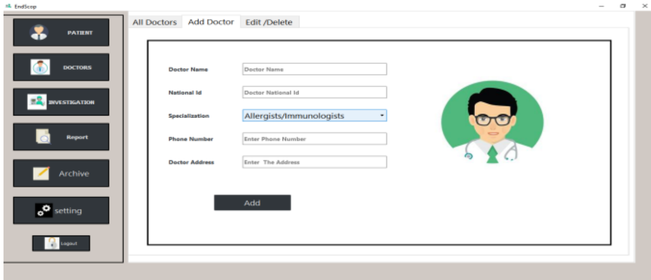


Fig. 11 Add and Modify Doctor Screen

Investigation screen, shown in Fig. 12, uses patient and doctor details to add a new investigation record. The entered information includes the surgery date and notes by the doctors. On the surgery, the pictures taken by the endoscope are imported preparing them for prediction step.

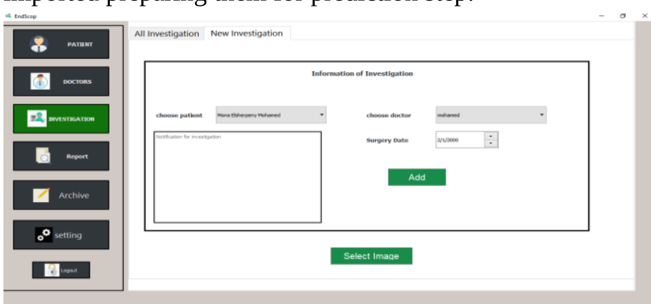
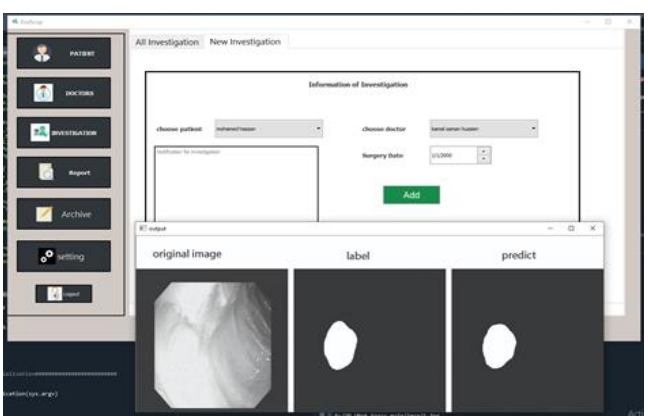


Fig. 12 Investigation Screen

To use the model, Predict screen is used. As shown in Fig. 13, the picture taken from the endoscope camera is entered to the model to predict the similarity between the input picture



and the trained images in our data set. The similarity results are shown and labeled to the doctors to decide.

Fig. 13 Predict Screen

V. Tools, Technologies and Deep Learning Model

In this section, the used tools and technologies are introduced in subsection A. This is followed by presenting the U-Net model used for deep learning model.

A. *Tools and Technologies*

For designing the web application, we used Qtdesigner, CSS, HTML, Python, and OOP. Cascading Style Sheets (CSS) is a style sheet language used for describing the presentation of a document written in a mark-up language such as HTML. CSS is a cornerstone technology of the World Wide Web, alongside HTML and JavaScript.

For database, MySQL database is used. MySQL is an Oracle-backed open source relational database management system (RDBMS) based on Structured Query Language (SQL). It runs on virtually all platforms, including Linux, UNIX and Windows. Although it can be used in a wide range of applications, MySQL is most often associated with web applications and online publishing.

For deep learning model, Numpy and CV2 libraries, Glob module, and TensorFlow.

The Numpy (i.e. Numerical Python) library is one of the most important libraries; i.e. first, many data science and machine learning libraries rely heavily on it, and second, it gives you the ability to deal with arrays in a better way than the existing Lists automatically as a form of data syntax in Python. It provides support for large multidimensional array objects and various tools to work with them. Various other libraries like Pandas, Matplotlib, and Scikit-learn are built on top of Numpy library.

The OpenCV library (Cv2) is one of the most powerful libraries in the field of image and video processing serving in the medical, industrial, artificial intelligence and machine learning fields. And It is used in operations on the image such as resizing, resizing, reading it, and scaling.

Glob Modle is a useful part of the Python standard library. Glob (short for global) is used to return all file paths that match a specific pattern.

TensorFlow is an open source customizable software library for performing numerical and graphical computations using data flow graphs and it contains all the libraries for deep learning. TensorFlow supports both CPU's and GPU's computing devices for distributed computing.

B. *Deep Learning Model*

As a subset of AI, deep learning has networks that are capable of unsupervised learning from data which are unstructured or unlabeled. U-Net is architecture for semantic segmentation [1] in deep learning used in this work. It consists of a contracting path and an expansive path. The contracting path follows the typical architecture of a convolutional network. It consists of the repeated application of two 3x3 convolutions (unpadded convolutions), each followed by a rectified linear unit (ReLU) and a 2x2 max pooling operation with stride 2 for downsampling. The model is depicted in Fig. 14.

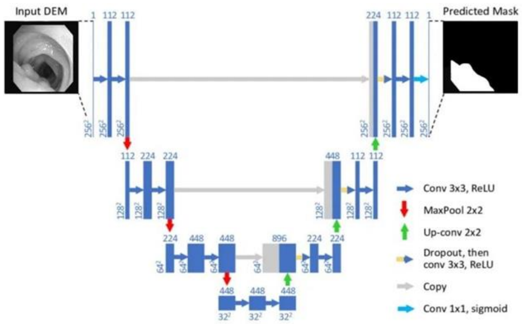


Fig. 14 U-Net Model.

At each downsampling step we double the number of feature channels. Every step in the expansive path consists of an upsampling of the feature map followed by a 2x2 convolution ("up-convolution") that halves the number of feature channels, a concatenation with the correspondingly cropped feature map from the contracting path, and two 3x3 convolutions, each followed by a ReLU. The cropping is necessary due to the loss of border pixels in every convolution. At the final layer a 1x1 convolution is used to map each 64-component feature vector to the desired number of classes. In total the network has 23 convolutional layers.

VI. Results and Discussion

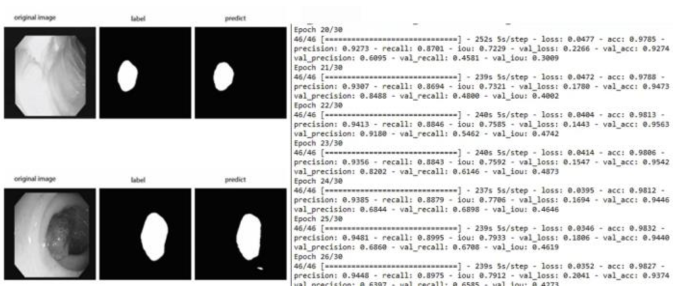


Fig. 15 Results Screen

As shown in Fig. 15, the doctor can take a picture through the surgical endoscope and compare it in our program, and our program can determine the affected area. At first, we used a dataset consisting of 600 images, the accuracy rate reached 93%, then we provided the machine with a larger dataset, which reached 80,000 images, and the accuracy of identification reached disease up to 98%. Resolution is the similarity between the image input to and output from the algorithm. In this case the accuracy has become 98%.

VII. Conclusions and Future Work

A smart endoscope system is introduced based on deep learning algorithms. The proposed endoscope is more accurate and intelligent so that it is able to respond to the voice of the user (the doctor). Thus, the doctor can give instructions during the surgery regarding taking pictures and/or recording videos. The proposed system allows the doctor to read the patient's medical tests and the ratios given to it allowing for comparison between the optimal condition and the patient's condition during the surgery.

Therefore, the proposed smart endoscope can diagnose diseases by analyzing the images given to it and comparing them to the pictures it takes inside the patient's body. This results in reducing the time of the surgery and alerts the doctor during the operation if there is a possible defect in certain area. Moreover, doctor can record and stop the video and take pictures by sound signals anytime during the procedure.

For future improvements we plan to add the functionality of producing automatic smart reports after comparing the pictures taken by the smart endoscope with the dataset. Moreover, a robot can be added to help the doctor perform the endoscope surgery. Furthermore, sound-enabled system instructions can be improved.

Acknowledgment

We would like to thank Dr. Nissreen El-Saber, our supervisor for her continuous support and guidance. A great thank you goes to Eng. Hossam Reda who has been our help during toughest of days.

References

[1] Khan, M. A., Kadry, S., Alhaisoni, M., Nam, Y., Zhang, Y., Rajinikanth, V., & Sarfraz, M. S. (2020). Computer-Aided Gastrointestinal Diseases Analysis from Wireless Capsule Endoscopy: A Framework of Best Features Selection. IEEE Access, 8, 132850– 132859. https://doi.org/10.1109/ACCESS.2020.3010448

[2] Dennis, A., Wixom, B. H., & Roth, R. M. (2019). System Analysis and Design 5thEdition. In *John Wiley & Sons, Inc*.