National University of Science and Technology
POLITEHNICA Bucharest
DOCTORAL SCHOOL ELECTRONICS
TELECOMMUNICATIONS AND INFORMATION
TECHNOLOGY

Decision No. 73 from 04.07.2024

Ph.D. THESIS
SUMMARY

SISTEM INFORMATIC MEDICAL CU MODUL AUTOMAT DE CLASIFICARE PATOLOGII

MEDICAL INFORMATION SYSTEM WITH AUTOMATIC PATHOLOGY CLASSIFICATION MODULE

Ph.D. Candidate: **Eng. Ana-Maria Ștefan**

THESIS COMMITTEE

Prof. Dr. Ing. Ion Marghescu National University of Science and Technology POLITEHNICA Bucharest	**President**
Prof. Dr. Ing. Mihai Ciuc National University of Science and Technology POLITEHNICA Bucharest	**PhD Supervisor**
Prof. Dr. Ing. Marius Dan Dobrea "Gheorghe Asachi" Technical University of Iași	**Referee**
Prof. Dr. Ing. Cristian Grava The University of Oradea	**Referee**
Conf. Dr. Ing. Alina Sultana National University of Science and Technology POLITEHNICA Bucharest	**Referee**

BUCHAREST 2024

ACKNOWLEDGEMENTS

The development of this doctoral thesis would not have been possible without the support of the scientific supervisor of the thesis, Professor Dr. Eng. Mihai Ciuc. I thank Professor Ciuc for granting me the freedom to approach this difficult topic in a personal way and for giving me the opportunity to pursue research in any direction necessary to achieve my goals.

I also wish to express my gratitude to Conf. Dr. Eng. Elena Ovreiu for guiding and supporting me on the path to obtaining my Doctorate degree, as well as for the personal support without which I would not have embarked on this journey, and whom I can now call a friend after many years.

I want to thank the guidance team for their scientific support throughout the doctoral school: Prof. Dr. Eng. Constantin Vertan, Dr. Eng. Elena Ovreiu, Dr. Eng. Alina Oprea, and the members of the thesis evaluation committee.

I am grateful to my colleagues Dr. Eng. Silvia Ovreiu and Eng. Elena Paraschiv for their collaboration in drafting the article we worked on together.

I am especially thankful and express my gratitude to my colleague and friend, MSc. Răzvan-Nicu Rusu, for the immense help he provided me unconditionally, for all his advice, criticisms, and ideas, and for his infinite patience, infectious optimism, and confidence in me. All of these things have contributed to and influenced the course of developing this doctoral thesis and obtaining the results presented in this work.

I also want to thank him for his ideas and involvement in drafting the articles we worked on together, but especially for never giving up, even when it seemed we had reached a dead end, to my colleague and friend Dr. Eng. Hassan El-khatib.

I want to express my gratitude to those who laid the foundation for my career in Digital Development at MEDIST Imaging & POC. I thank the CEO, Dr. Bogdan Radu Niculescu, for the opportunity he provided me to lead an entire department and to support the education of the new generation of colleagues and researchers in the field. I thank MSc, MBE, Eng. Daniel Pop, for being my mentor and role model, and for directing me in this direction, having an extraordinary confidence in me - that there is nothing I cannot do.

Last but not least, I want to thank my mother, without whom I would not have reached this point, who has always been by my side, supporting, and understanding me.

TABLE OF CONTENTS

Chapter 1

Introduction

Doctoral Thesis entitled "Medical Information System with Automatic Pathology Classification Module" was developed at the Faculty of Electronics, Telecommunications, and Information Technology of the University of Science and Technology POLITEHNICA Bucharest. It contains original contributions in the fields of automatics and computers science, digital development, electronics, and information technology. The doctoral thesis addresses two main subjects. The first subject deals with medical image analysis systems based on neural networks and the development of models and classifiers that classify medical images into two classes: healthy and diseased, and later capable of indicating existing pathologies based on the obtained image.

The second subject addresses the development of a CRM (Customer Relationship Management) type information system for a hospital institution. Both individually addressed subjects were combined by integrating a medical image classification system directly into the information system, to optimize the workflow for the benefit of both the medical team and the patients, operating in real-time and providing instant information about the presence of a pathology from an acquired image or video.

The structure of the doctoral thesis is divided into 5 chapters as follows:

Chapter 1 presents the scope of the doctoral thesis and the subjects addressed, which involve two different domains and the objectives of the thesis, which integrate these two domains into the construction of a completely new information system.

Chapter 2 presents the current state in the research of medical image analysis systems, automated medical image applications, and automatic diagnosis with the help of artificial intelligence. It reviews current architectures and frameworks for image classification, platforms for building machine learning models, and neural networks specially designed for the detection of ocular and skin pathologies, as well as existing and future databases to be used in the thesis.

Chapter 3 introduces information systems, covering general concepts such as components, benefits, objectives, evaluation criteria, types of information systems found

in institutions, implementation stages of such systems, and objectives of an information system, along with notions about medical information systems, necessary modules, and examples of information systems used in medical institutions.
Chapter 4 presents the proposed solution in the thesis and the general considerations of the solution, presentation of the chosen medical information system, the methodology of implementation for both the medical information system and the automatic pathology detection module (creation of architecture, development of new functionalities, deployment, and integration).
Chapter 5 presents the results obtained during the research and development of the doctoral thesis, the original contributions made, the list of published papers, and future research and development directions for the doctoral thesis topic.
The thesis concludes with the "Bibliography" chapter containing 237 bibliographic references and the annexation of the published works.

The contributions made to the field during the elaboration of the doctoral thesis have been disseminated in the following conferences and specialized journals:

1. [33] El-Khatib, H., **Ștefan, A.-M.,** & Popescu, D. (2023). Performance Improvement of Melanoma Detection Using a Multi-Network System Based on Decision Fusion. Applied Sciences, 13(18), 10536–10536. https://doi.org/10.3390/app131810536
2. [57] **Ștefan, A.-M.;** Rusu, N.-R.; Ovreiu, E.; Ciuc, M. Advancements in Healthcare: Development of a Comprehensive Medical Information System with Automated Classification for Ocular and Skin Pathologies—Structure, Functionalities, and Innovative Development Methods. Appl. Syst. Innov. 2024, 7, 28. https://doi.org/10.3390/asi7020028
3. [58] **Ștefan, A.-M.;** Rusu, N.-R.; Ovreiu, E.; Ciuc, M. Empowering Healthcare: A Comprehensive Guide to Implementing a Robust Medical Information System - Components, Benefits, Objectives, Evaluation Criteria, and Seamless Deployment Strategies. Applied System Innovation. 2024; 7(3):51. https://doi.org/10.3390/asi7030051
4. [59] **Ana-Maria ȘTEFAN,** Elena OVREIU, Mihai CIUC, "Comparative analysis of web-based machine learning models", Romanian Journal of Information Technology and Automatic Control, ISSN 1220-1758, vol. 34(2), pp. 49-63, 2024. https://doi.org/10.33436/v34i2y202404
5. [89] **Ștefan, A.-M**., Paraschiv, E.-A., Ovreiu, S., & Ovreiu, E. (2020). A Review of Glaucoma Detection from Digital Fundus Images using Machine Learning Techniques. 2020 International Conference on E-Health and Bioengineering (EHB). https://doi.org/10.1109/ehb50910.2020.9280218
6. [206] El-khatib, H.**, Ștefan, A.-M**., & Popescu, D., Melanoma Automated Detection System Integrated with an EHR Platform – UPB Scientific Bulletin Series C: Electrical Engineering and Computer Science 2024. Vol. 86, Iss. 1, 2024. ISSN 2286-3540

Chapter 2.

Current status in the research of medical image analysis systems

Medical diagnosis has taken an unexpected turn with the help of artificial intelligence; from robots to automated diagnostic systems, to telemedicine, drug manufacturing, monitoring, intelligent systems (prostheses with brain-machine interface-based systems), and much more. With the help of artificial intelligence, the medical system and workflow will be optimized to the maximum. Diagnosis time will be shortened, diagnosis will be safer and based on a much larger database of examples (which translates to greater evaluator experience), and early diagnosis of pathologies will be possible.

Several hospitals have implemented artificial intelligence algorithms to improve workflow (e.g., the Mayo Clinic in Minnesota). Existing systems for automatic diagnosis are diverse and increasingly numerous; they use images from CT (computer tomography), MRI (magnetic resonance imaging), X-rays, and echography.

Machine Learning (ML) is an area of computer science that deals with analyzing and interpreting patterns, structures in large volumes of data for tasks such as learning and decision support with minimal human expert intervention. ML allows users to input large volumes of data into a computerized algorithm that will analyze them and generate decisions and recommendations. The algorithm also learns from the error function and uses it as input for optimization. Deep learning is a method of machine learning based on the architectural representation of how the biological brain learns.

Artificial Neural Networks progressively learn to perform certain tasks from given examples. For example, in systems for recognizing various instances, the network receives labeled inputs and learns to recognize them by comparing them with known instances. Similar to biological networks, neurons are connected to each other and transmit information from one to another; each of them has a state (which can be 0 or 1) and a weight, which can be modified as the algorithm learns. The role is to create neural network algorithms that learn similarly to the biological brain. The advantage of these artificial networks is that they can access specific abilities that depart from biological capabilities, such as back-propagation and network optimization to obtain the necessary output [15]. Deep neural networks are networks with a number of layers between input and output [16], and each layer calculates the probability of each output. Deep learning models are specialized for a specific task summarized by specific examples depending on the task. A particular model cannot be trained on another dataset (the implementation steps are exemplified in Figure 2.11 and Figure 2.12).

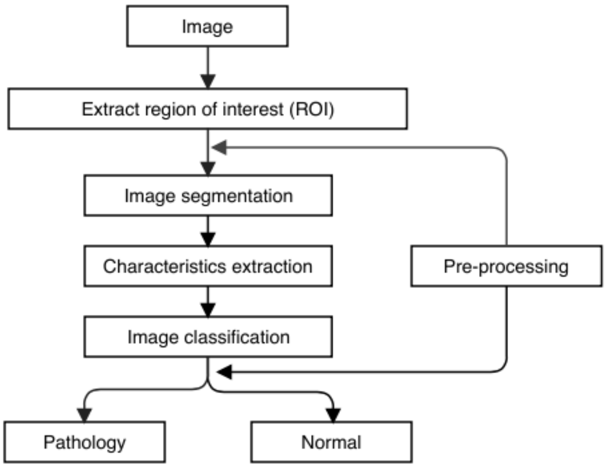

Figure 2.11 *Implementation steps for image classification.*

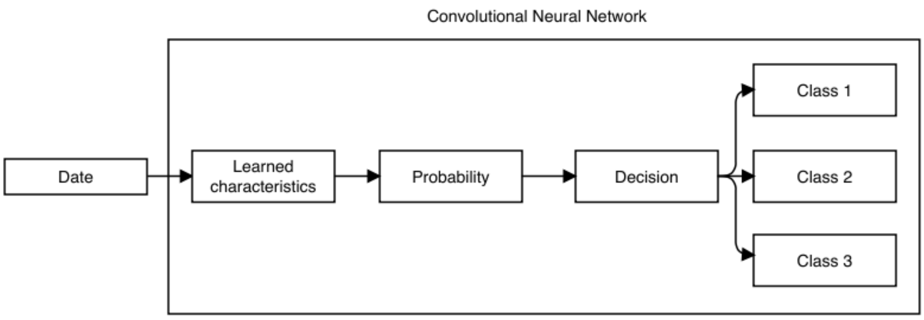

Figure 2.12 *Steps for the training task.*

Training a neural network occurs in two phases. One phase involves passing the input data forward through the network, and the other phase involves updating the gradients and weights. Knowledge transfer can be used to apply the knowledge gained in solving a new problem. This is very useful in cases where there is not enough information for a particular problem, and thus in the classical approach, the network would not have satisfactory accuracy (the steps are described in Figure 2.13 [22-23]).

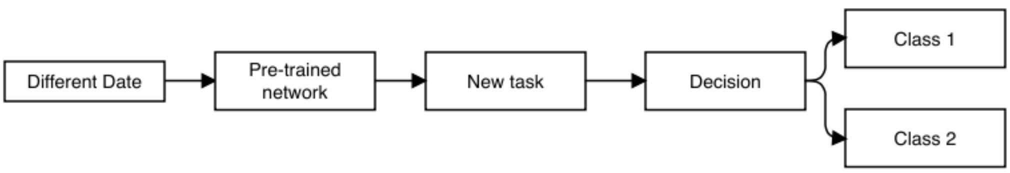

Figure 2.13 *Steps for transfer learning*

The popular frameworks used for image classification are TensorFlow, Keras, Caffe, MATLAB, NVIDIA Caffe, PyTorch, MXNet, and Chainer. Models for machine learning can be created using applications for their rapid construction and exportation of those models for further use on different platforms. In this thesis, Azure Machine Learning, Google Teachable, Google Cloud – Vertex AI, and Salesforce Einstein Vision were investigated.

The focus of this doctoral thesis will be on studying the development of a computer system that integrates automatic image classification applications for two of the leading causes of death globally: Diabetes[1] (103,294 deaths) and Cancer[1](605,213 deaths) [67]. The doctoral thesis domain will focus on intelligent processing of retinal and skin medical images. Various classification systems applied to different image databases will be compared. In the case of skin pathologies, two types of skin cancer (melanoma and basal cell carcinoma) and two other common skin conditions that, if left untreated, can lead to skin cancer (actinic keratosis and seborrheic keratosis) will be studied. For diabetes, complications related to ocular diseases such as diabetic retinopathy, glaucoma, and cataracts will be addressed. Early detection of acute pathologies through such a system can significantly influence patient outcomes, reduce healthcare costs, and promote proactive preventive measures. Furthermore, the use of such an innovative system adds value to the field of digital development for several reasons:

- Rapid Diagnosis via automatic image classification can accelerate the diagnostic process by quickly identifying and categorizing medical images.
- Early Detection becacause the system can assist in the early detection of diseases or conditions, leading to better treatment outcomes.
- Consistency because automated systems provide consistent and standardized analyses of medical images, reducing the likelihood of human errors.
- Time Saving for healthcare professionals that save time through automated image analysis, allowing them to focus more on patient care.
- Rapid Access to labeled images within the system can help doctors make informed decisions quickly.
- Information Collection via automatic image classification provides valuable data for medical research and analysis, aiding in identifying trends, correlations, and treatment effectiveness.
- Population health management by data collection from image classification can be used for population health management and preventive strategies.
- Handling large volumes of data of medical units can efficiently manage and classify a large volume of medical images as the patient and data bases grow.
- Adaptability because the system can be customized for specific medical specialties.
- Process automation of image classification reduces the need for labor-intensive manual image analysis, potentially saving labor costs.
- Workflow pptimization leads to increased efficiency, which can result in cost savings and resource optimization.

- Data security – MIS can ensure secure and controlled access to patient data, including medical images, maintaining their confidentiality.
- Data encryption of sensitive data in images can be encrypted to protect patient information.
- Alerts and notifications for critical findings in medical images, ensuring prompt action.
- Integration for MIS with image classification modules can be integrated with other IT systems in healthcare, creating a cohesive medical ecosystem.
- Remote Access and collaboration for physicians on medical images remotely, improving the accessibility of healthcare services, especially in telemedicine.
- Consistency and quality because automatic image classification ensures a consistent and high-quality standard in image analysis, reducing variations among different healthcare professionals.
- Compliance with regulations because MIS with image classification capabilities can help healthcare organizations comply with regulations and standards regarding the management of medical images.

In conclusion, the importance of early detection of diseases with the help of artificial intelligence lies in the potential to improve treatment outcomes, reduce medical costs, and enhance overall well-being by addressing diseases before they progress to advanced stages. [82] To test algorithms, public databases containing biomedical images were searched for [33], [206], [57-59]:

a. ORIGA (Online Retinal Fundus Image Dataset for Glaucoma Analysis and Research) - is a public database containing 168 color fundus images of the posterior pole of the retina with glaucoma and 482 images without pathologies [138].
b. Messidor - is a database for diabetic retinopathy containing 1200 retinal images (with and without glaucoma) labeled by experts, 800 of which were acquired with dilated pupils [139].
c. EyePacs - is a database available from Kaggle containing 35,126 retinal images of diabetic retinopathy obtained from screening programs [140]
d. HRF (High-Resolution Fundus Image Database) - is a database containing 15 fundus images without pathologies, 15 images with glaucoma, and 15 images with diabetic retinopathy provided by CS5 Laboratory of the Department of Ophthalmology at Friedrich-Alexander University Erlangen-Nuremberg and Brno University of Technology - Department of Biomedical Engineering [141]
e. ISIC (International Skin Imaging Collaboration) - contains 93,083 images of skin lesions and was developed for testing algorithms for melanoma detection [142]
f. Multiple Pathologies Image Database (cataract, glaucoma, and diabetic retinopathy) - available on Kaggle, containing 601 images divided into 5 classes [143]
g. RIM-ONE - a database containing retinal images obtained from ocular fundus images, 200 images with glaucoma and 255 without pathologies [144], [145].

h. DermIS - this database is the most comprehensive database containing dermatological images of almost all types of skin conditions [146].

i. MED-NODE - contains 170 images (70 with melanoma and 100 normal) [147].

j. PH^2 - contains 40 dermatoscopic images representing melanoma and 80 images without pathologies [148]

k. Additionally, a set of images was obtained from Ponderas Academic Hospital, containing 32 images of patients with glaucoma and 34 images of patients without pathologies.

Chaper 3.

Information systems

3.1 General notions about information systems

An information system is a combination of software, hardware, and communication networks designed to collect, process, store, and distribute data for use in workflows and to facilitate decision-making and process optimization.[149] Data becomes information when processed to support decision-making. Information systems can vary considerably in complexity and functionality, from simple systems to complex enterprise systems such as Customer Relationship Management (CRM) software or Enterprise Resource Planning (ERP) systems [57].

3.1.1 The components of a information system

From a socio-technical perspective, information systems are made up of the following components - Human Resources, Equipments, Software, Data, Procedures, Network, Security, Analysis, Control, Environment. All these components work together to create a coherent and functional IS that supports the goals and objectives of an organization. The design and integration of these components are crucial to the system's success in providing accurate, timely information to support relevant decision-making and operational processes.

3.1.2 The benefits and objectives of a computer system

Information systems provide a wide range of benefits for organizations in all fields. These benefits help improve efficiency, decision-making, communication and overall performance. Some key advantages of information systems include:

- Improving the decision-making process
- Increasing efficiency and productivity
- Efficient communication
- Information available in real time
- Improving customer management
- Analysis and reporting
- Strategic planning
- Resource management
- Compliance
- Innovation
- Remote access
- Facilitating documentation management
- Extension [151-159]

Overall, the benefits of information systems are numerous and contribute to the strategic, operational and financial success of organizations. They play a vital role in modern business operations and are critical to maintaining competitiveness in today's digital age.

3.1.3 Evaluation criteria of an information system

Evaluation criteria for an information system are the standards and measures used to evaluate the effectiveness, efficiency, quality and overall performance of the system, such as functionality, ease of use, performance, scalability, reliability, security, interoperability, flexibility, low cost, maintenance and support, integration with business objectives, adaptation to change, reporting and analysis, training, documentation, sustainability [160-163]

3.1.5 Stages of implementing an information system

An information system goes through a series of stages in its implementation to serve the objectives for which it is intended. These stages are part of a methodology developed in 1960 for managing the implementation of complex software projects - SDLC (System Development Life Cycle) which is a structured and systematic approach to the design, development, testing, implementation and maintenance of software applications or systems.

It outlines the various steps and activities involved in creating and managing software, from initial concept to final implementation and ongoing maintenance. The SDLC provides a framework that helps ensure software projects are completed efficiently, on time, and within budget while meeting quality and performance requirements. The main stages or phases of the SDLC typically include Planning, Analysis, Design, Development, Testing, Implementation, Maintenance, Evaluation. There are different methods and models for SDLC implementation, including the Waterfall model, Agile methodologies (such as Scrum and Kanban), and iterative models. Each approach has its own set of principles, practices, and benefits, and organizations can choose the one that best fits their project requirements, team dynamics, and organizational culture. Overall, the SDLC provides a structured framework that guides software development projects through a series of well-defined stages, ensuring that software is developed, tested, and deployed in a systematic and controlled manner [173], [58].

3.2.1 Modules of an information system

A medical information system is typically constructed from several interconnected modules that serve specific functions within a medical organization. These modules facilitate the management of patient information, optimize processes and improve the delivery of healthcare services. While specific modules may vary depending on system design and organization needs, common ones are as follows:

• Electronic Health Record (EHR) is the central component, storing and managing patients' medical records digitally. Includes information such as medical history, diagnoses, treatments, medications, lab test results, and more.
• Patient Registration Module manages patient registration, demographic details and insurance information
• Appointments module allows the management of patients' appointments with healthcare providers, including alerts and notifications.
• Financial module allows the management of the financial-accounting processes of insurance applications and financial transactions related to the services provided.
• Laboratory Information System (LIS) manages laboratory test activities, test results and reporting. It integrates with diagnostic equipment and automates laboratory processes.
• Radiological Information System (RIS) manages radiology and imaging activities, procedures and results, often integrated with picture archiving and communication systems (PACS).
• Pharmaceutical management module manages drug ordering, dispensing and stock control activities. May include checks of drug interactions and their history.
• Telemedicine mode facilitates remote consultations, virtual visits and patient monitoring through telemeters.
• Patient care and nursing mode supports nursing activities, patient care plans and their documentation.

• Intervention management module helps schedule interventions, pre-operative assessments and post-operative care coordination.
• Medical imaging module facilitates the management of medical images, integration with radiology laboratories and facilitates viewing and sharing of images.
• Inventory Module facilitates the management of stocks, expiration dates and the supply chain.
• Human resources and payroll module facilitates the management of employee records, scheduling, payroll processing and workforce management.
• Analysis and reporting module facilitates the provision of data analysis, reporting and performance metrics for informed decision making and evaluation of healthcare processes.
• Patient portal module provides patients with access to medical records, appointment scheduling, communication with healthcare providers and educational resources.
• Administrative and management module supports administrative tasks, resource allocation and general system management.
• Mobile application module allows access to patient information, appointments and communication via mobile devices [189-192].

These modules work together to create a medical information system that improves patient care, optimizes operational efficiency, and supports healthcare providers in delivering high-quality services.

Chapter 4.

Proposed solution

The medical industry is a big business. In the USA in the year 2021 health care expenses reached 4.3 trillion dollars which means 12,914 dollars per person [200] and in the European Union 964 billion dollars. [201]

If the healthcare system is increasingly viewed as a business, healthcare delivery systems should also think in business terms. This objective can be achieved by implementing a customer relationship management system (usually used in the business - sales area). The type of system called CRM technology for managing all of the company's relationships and interactions with customers and potential customers. Using this type of system, the customer's satisfaction increases, the patient has a premium experience and will have access to all the information related to his health at his fingertips (on the phone, tablet or computer) - on the go. The goal is simple: to improve business relationships for business development. A CRM system helps companies stay connected to customers, streamline processes through a shorter turnaround time, and improve profitability through targeted marketing campaigns based on the patient and their condition, attracting patients through a premium healthcare delivery experience. This system can include any number of modules and can be built like a puzzle to be able to interconnect the necessary objects, records and options from each module and can be connected to modules for reporting, analysis and interconnection with other platforms.

The implementation of such a system provides the possibility of implementing a common interface for all existing workflow processes in a medical institution. Teams work together on the same stream and have access to the same kind of real-time information. Ultimately, choosing between these CRM systems depends on your specific needs, budget, and the level of customization and scalability required. In this paper Salesforce was chosen based on its flexibility, mode of operation and developer experience.

The solution proposed in this paper consists in using a information system that integrates an application for automatic classification of medical images. The system chosen for this project was the Salesforce platform, which is a cloud software development company that is the world leader in CRM IT systems (Customer Relationship Management - a 360 view of customers and a global view of work processes and information useful to all workflows). This system has the great advantage of being a SaaS (Software as a Service), which allows users to connect and use cloud-based applications via the Internet in the browser. Salesforce has been around since 1999 and was founded by a former Oracle executive. Salesforce (SF) has a broad user base across industries and its cloud-based structure allows users to access the platform from anywhere, making it a flexible solution for organizations of all sizes. Salesforce has played a significant role in transforming the way businesses manage customer relationships and has become a leader in modern business operations.

Salesforce infrastructure refers to the technology, architecture, and resources that support the operation of the platform that includes the software components, networks, databases, and data centers that work together to ensure the availability, performance, scalability, and security of Salesforce services. Collectively, SF's infrastructure provides the foundation for the delivery of its CRM and cloud services. The combination of advanced technology, scalability, security and redundancy ensures that customers can access and use SF capabilities reliably and securely. The Salesforce platform will be customized as a medical information system. We chose MIS because it has a broad scope that encompasses the entire patient history, including medical and clinical data, but also includes information from other health care providers and sources. MIS are designed to be accessible to a variety of authorized users, including patients themselves.

At the same time, they are designed for interoperability. They enable the exchange of patient data between different organizations and healthcare systems, unlike EMRs for example, which are limited to a single institution and do not allow for easy sharing with third parties. This is crucial for the provision of coordinated care, particularly for patients who see multiple specialists or require care in different settings. In addition, patients have access to their own data through a portal that also allows them to schedule consultations, share information, etc. The proposed solution brings additional benefits from the point of view of having an automatic pathology detection system in a MIS-type system that contains all the workflows for members of the medical teams in the same place:

- Improving patient outcomes
- Prevention of disease progression
- Reduced healthcare costs
- Population health management
- Streamlined workflow
- Data-driven insights
- Research and Development [80-81], [171], [202-204]

Essentially, early detection of pathologies through automated systems integrated into a MIS not only improves individual patient medical care, but also has broader implications for public health, healthcare efficiency, cost reduction, and advances medical research. This is a significant step towards proactive and personalized health management.

4.3 Building of the classification module

Following the research carried out in order to optimize the classification process, some steps are necessary, including pre-processing methods of medical images, such as the extraction of the region of interest, the removal of artifacts (hairs, water drops, etc.), the extraction of the disk optical, image segmentation, feature extraction, noise/artifact removal (such as hair in the case of skin images or bones in the case of radiographs), morphology (edge smoothing), image augmentation, image resizing, binarization, brightness adjustment/correction, enhancement intensity and contrast of images [207-212].
All these methods improve the way data is processed with minimal loss of information. In this chapter we will evaluate tools for creating a model for automatic classification of medical images. To build the models for classifying both skin and eye images, we will use the databases presented in Chapter 2. They were created using both images acquired during fundoscopy and images acquired using a camera, and the one that performed the best, which is the most versatile (performs well on any medical image database) in terms of accuracy, F1 score but also the versatility of further integration into external platforms. For the evaluation of the automatic classification models we used the accuracy and the F1 score because it takes into account both false negative and false positive results. Expressions for performance indicators are exemplified in table 4.2.

Table 4.2 *Expressions for performance indicators*

Performance indicators	Formula
Precision	$\frac{TP}{TP + FP}$
Recall	$\frac{TP}{TP + FN}$
Accuracy	$\frac{TP + TN}{TP + FP + FN + TN}$
F1 score	$2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} = \frac{TP}{TP + \frac{1}{2}(FP+FN)}$

TP = number of true positive results; FP = number of false positives; TN = number of true negative results; FN = number of false negatives.

- Accuracy – performance indicator of a system that measures the performance in the detection of pathologies (proportion of correctly identified classes)
- F1 score – performance indicator that measures the performance of the system in relation to false positive and false negative results
- Sensitivity/Recall – performance indicator that measures the proportion of correct pathology detection (TP). This metric focuses on minimizing false negative results which is very important to not miss cases of incorrectly detected cases.
- Specificity – performance indicator that measures the proportion of correct detection of non-pathology (TN). This metric is important to reduce false alarms and unnecessary further investigations [33], [206], [57-59], [213-218].

Three case studies were conducted to build and evaluate the classification models.

4.3.1 Case Study 1.

In the framework of this study, a web application was built that runs in the browser and allows the uploading of medical images based on which the built model classifies the image into one of two classes: normal image without pathologies or glaucoma. Resources used were Google Collaboratory, Tensor Flow Core, Keras, Streamlit and image database – ACRIMA containing 428 images containing glaucoma and 428 images containing no pathology. The images were divided into three categories: training, validation and testing. The built model had an accuracy and F1 score of 0.8452.

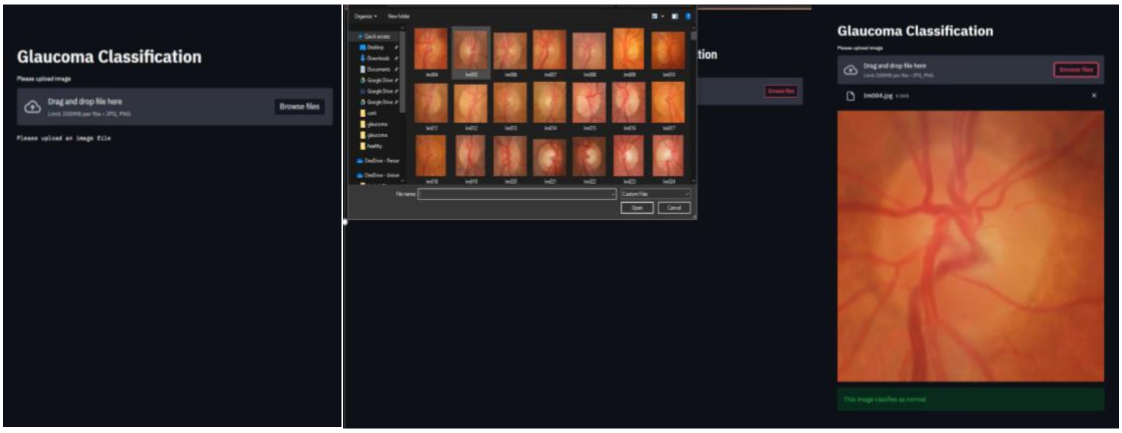

Figure 4.6 *Uploading images and getting the result*

Various ways of integrating this application into an information system were investigated, but given the compatibility of distinct frameworks, security and programming languages, such an integration was not possible, especially since it is not data, but of images.

4.3.2 Case Study 2

In this study [33] we evaluated the performance of four web-based tools (Google Teachlable Machine, Azure Machine Learning, Google Vertex AI and Salesforce Einstein Vision) for creating models for automatic detection of pathologies and a classifier built in MATLAB on decision fusion basis of several convolutional neural networks (DarkNet-53, DenseNet-201, GoogleNet, Inception-V3, InceptionResNet-V2, ResNet-50, ResNet-101 and Xception). The case study was conducted strictly for the detection of melanoma. The comparison was made based on the obtained results of the F1 score because it is more efficient than the accuracy as it takes into account the number of FN and FP. The databases used for testing the adaptability of the systems as well as for their training are ISIC 2020 and DermIS. The model has an accuracy and F1 score of 0.9550 on the DermIS database and 0.9350 on the ISIC database. As a result of the research done, the integration between MATLAB and a computer system proved to be impossible at the time of the research.

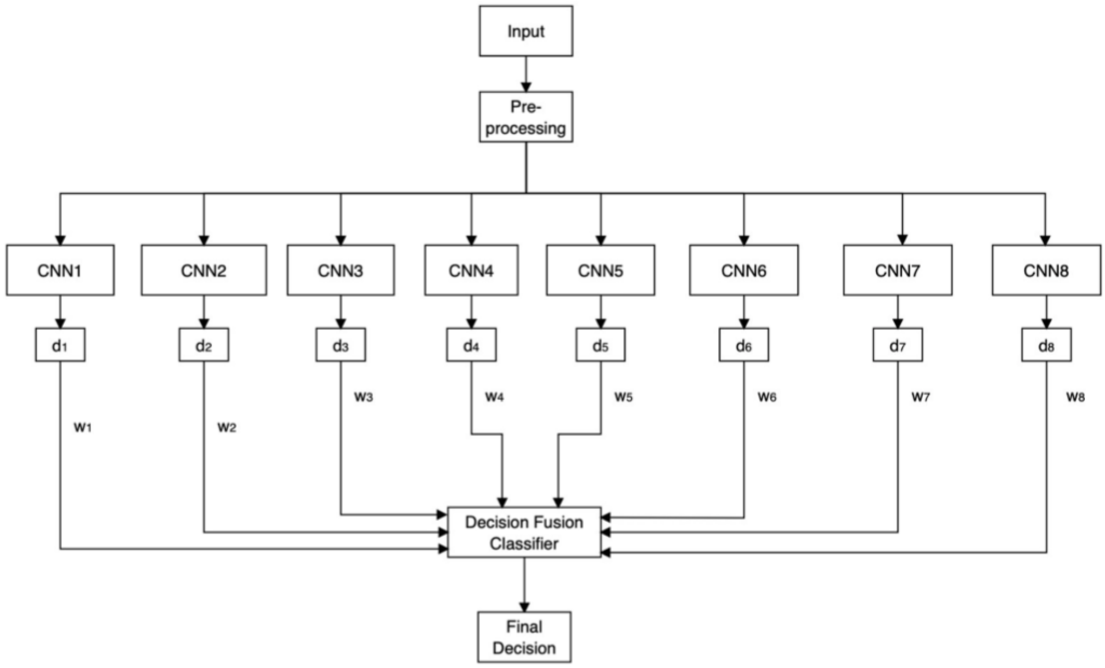

Figure 4.14 *Architecture of the proposed system*

Table 4.5. *Experimental results for the decision fusion classifier vs. other classifiers vs. automated applications.*

Clasification method	Databse	Accuracy	F1 Score
Google Teachable Machine	DermIS	0.7800	0.7786
	ISIC	0.7600	0.7597
Google Vertex AI	DermIS	0.8333	0.8331
	ISIC	0.7400	0.7370
Microsoft Azure Machine Learning	DermIS	0.7967	0.7966
	ISIC	0.8200	0.8200
Salesforce Einstein	DermIS	0.8033	0.8025
	ISIC	0.7933	0.7900
The proposed classifier based on decision fusion	DermIS	0.9550	0.9550
	ISIC	0.9350	0.9350

4.3.3 Case Study 3.

Continuing case study 1, where the performance of a custom classifier based on the fusion of the decisions of several CNN architectures is superior to web-based models, but which, following the extensive study carried out, no reliable option for integration with another platform was found. Therefore, the research will be extended to several databases to verify their adaptability. We performed a case study in which we compared the models with each other, both on databases for ocular pathologies such as ORIGA, Messidor, EyePacs, HRF and Ponderas and on databases containing skin lesions such as ISIC, DermIS, MED-NODE, PH^2, creating both models with two targets (healthy or diseased) but also with three or four targets (healthy or one of the investigated pathologies).

Based on the results obtained in the two studies and the additional research that was carried out, each application has advantages and disadvantages, but considering the requirements of the project, building the model for the automatic detection of pathologies that we will integrate with the Salesforce platform is Google Teachable because the model created must be compatible with SFDC, in addition the model has quite good accuracy on a wide range of images acquired in various ways and is adaptable to many encountered pathologies. To create the architecture of the medical information system with the automatic classification of pathologies, several objects, fields, consoles, layouts, reports, graphs were created for the workflow to be followed. Objects were created to be able to store at the database level all the relevant information about patients, equipment, results, etc. [59].

Table 4.6 *Accuracy of the tested models*

Model	Pathology	Accuracy	F1 score	Average ccuracy
Azure Machine Learning	Glaucoma	0.9631	0.963	94%
	Cataract	0.9024	0.9024	
	Diabetic retinopathy	0.9651	0.9645	
Google Teachable	Glaucoma	0.9713	0.9712	96%
	Cataract	0.9675	0.9674	
	Diabetic retinopathy	0.9476	0.9462	
Google Cloud – Vision AI	Glaucoma	0.9836	0.9836	98%
	Cataract	0.9878	0.9878	
	Diabetic retinopathy	0.9825	0.9822	
Einstein Vision	Glaucoma	0.9188	0.9179	90%
	Cataract	0.8740	0.8735	
	Diabetic retinopathy	0.9083	0.908	
Azure Machine Learning	Melanoma	0.9024	0.9024	93%
	Actinic keratosis	0.9351	0.9352	
	Seborrheic keratosis	0.9579	0.9576	
	Basal cell carcinoma	0.9291	0.929	
Google Teachable	Melanoma	0.9451	0.9451	96%
	Actinic keratosis	0.9457	0.9456	
	Seborrheic keratosis	0.9828	0.9827	
	Basal cell carcinoma	0.9866	0.9865	
Google Cloud – Vision AI	Melanoma	0.9776	0.9776	96%
	Actinic keratosis	0.9524	0.9523	
	Seborrheic keratosis	0.9713	0.9711	
	Basal cell carcinoma	0.9655	0.9654	
Einstein Vision	Melanoma	0.9184	0.9184	91%
	Actinic keratosis	0.9199	0.9198	
	Seborrheic keratosis	0.9195	0.9191	
	Basal cell carcinoma	0.9004	0.9002	

Chapter 5.
Results, Conclusions, Personal Contributions and Future Directions

5.1 Results

5.1.1 Implementation of a medical information system

5.1.1.1 The structure of the developed medical system

In order to achieve the proposed objectives, it started with a CRM-type information system that ensures the tracking of the basic workflow and data processing in an institution that interacts with customers (Figure 5.1). Later the platform was modified and developed both vertically and horizontally, obtaining a new information system that integrates solutions for the automatic classification of medical images using artificial intelligence as well as newly developed applications and functionalities that respond to the pursuit of a complex workflow in a medical institution. The new system obtained "Ana Medical" developed by the author of the thesis that fulfills the proposed objectives and ensures the operation and integration of the application modules that compose it, as well as the newly created functionalities to optimize the workflow or improve the end user experience. The structure of the information system objects are exemplified in Figure 5.2.

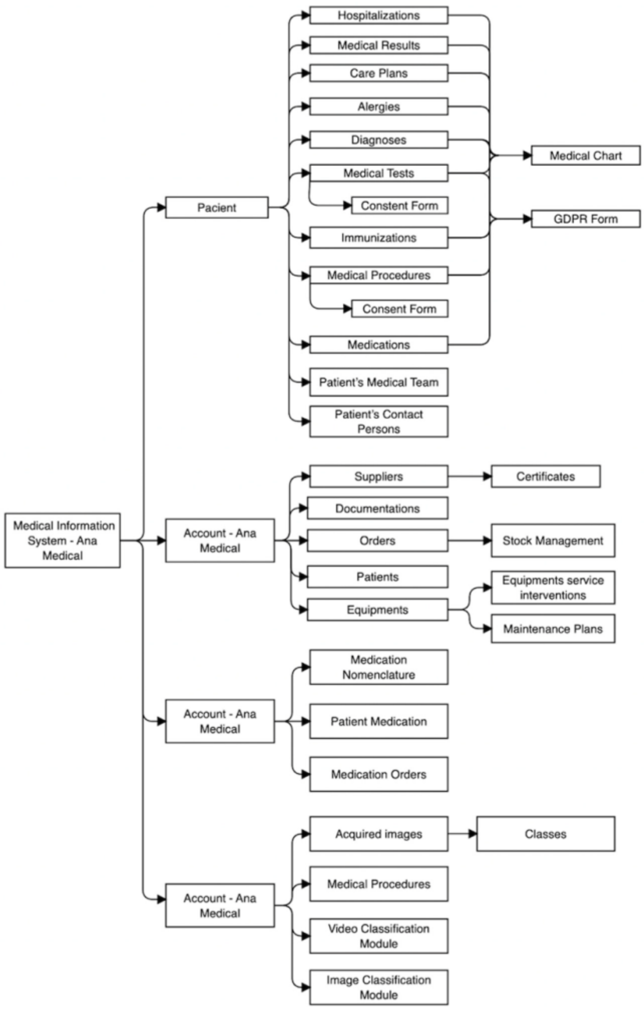

Figure 5.2 *The structure of the Ana Medical system [57]*

The information system is structured in several modules (consoles) (Figure 5.3).

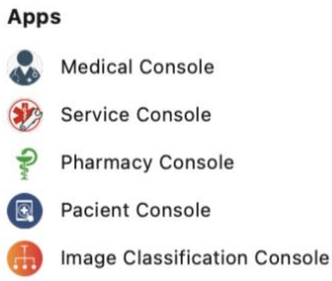

Figura 5.3 *Ana Medical system module*

a. Medical Console - The console is dedicated to the medical team dealing with patient flow and interactions with them from a medical point of view. Dependent on this object are the following patient information:

- Patient Encounters – patient hospitalizations
- Patient Results – results of medical analyzes and procedures
- Patient Care Plans – patient care plans

• Patient Allergies – patients' allergies
• Patient Diagnosis – patient diagnoses
• Patient Medical tests – the patient's medical tests
• Patient Immunizations – the patient's vaccines
• Patient Procedures – patient procedures
• Patient Medication List – the medical treatment that the patient is following
• Patient Teams – the patient's medical team
• Contact Persons – the patient's contact persons

This information about the patient is grouped in Related Lists (which formed a link between the patient and the specific nomenclature). The related list is a component that displays a list of related records based on a specific object. At the level of a patient's record, a series of information about him can be found:

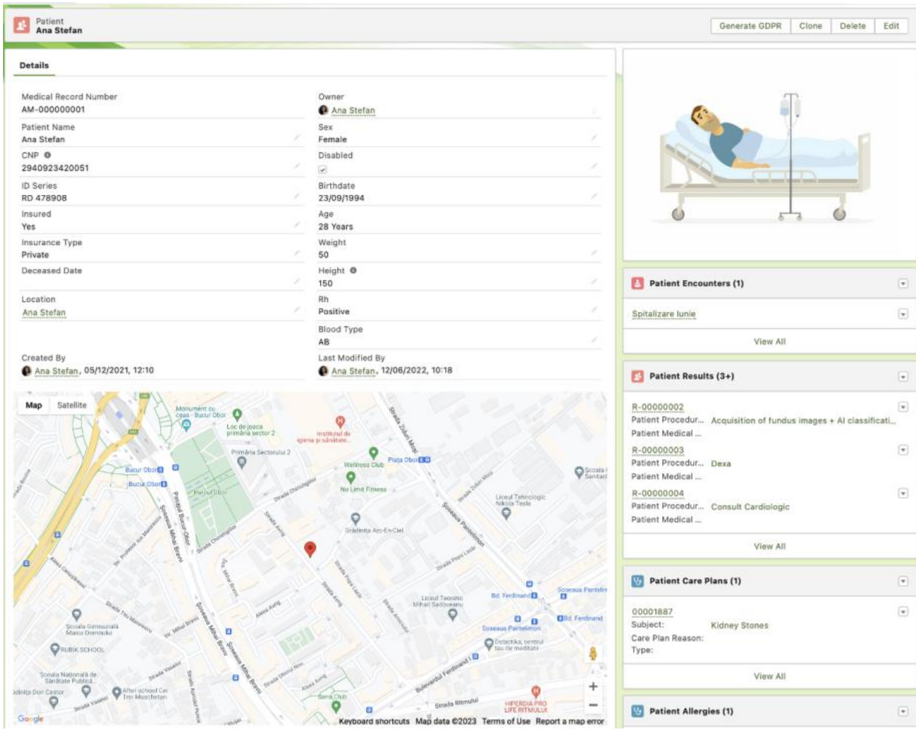

Figure 5.6 *Platform interface at the level of the Patient object*

At the level of this object, several forms are generated, such as the GDPR (General Data Protection Regulation), in order to be able to process personal data in the system, for medical procedures and tests, a patient's informed consent form is generated, but also the patient's medical record with the information generated during his hospitalization in the medical institution (Figure 5.7).

Ana Medical

Consimtamant GDPR

In conformitate cu Regulamentul (UE) 2016/679 al Parlamentului European si al Consiliului din data de 27.04.2016 privind protectia persoanelor fizice in ceea ce priveste prelucrarea datelor cu caracter personal si privind libera circulatie a acestor date si de abrogare.

Imi exprim consimtamantul ca unitatea medico-sanitara sa fie autorizata sa prelucreze datele mele personale prin intermediul platformei CNAS-SIUI (Sistem Informatic Unic Integrat), in scopul realizarii atributiilor si obligatiilor prevazute de actele normative care guverneaza sistemul de asigurari sociale de sanatate si in special celor care decurg din statutul de asigurat al pacientului in sistemul de asigurari sociale de sanatate.

Consimtamantul in ceea ce priveste prelucrarea datelor cu caracter personal, prin intermediul SIUI (Sistem Informatic Unic Integrat), precum si furnizarea datelor rezultate din investigatiile medicale sunt voluntare.

De asemenea, afirm si recunosc ca am fost informat si am inteles pe deplin posibilitatea retragerii consimtamantului, in orice moment, in cazul in care prelucrarea datelor mele personale se efectueaza prin mijloace automatizate si ca acest lucru afecteaza prelucrarea efectuata pe baza consimtamantului inainte de retragerea acestuia, in sensul imposibilitatii inscrierii in sistemul de ocrotire al sanatatii din Romania si de acordare a serviciilor medicale.

Retragerea consimtamantului se poate face prin cerere scrisa datata i semnata la sediul unitatii medico-sanitare la datele de contact mentionate in antet, in atenia reprezentantului legal.

Totodata, stiu ca am dreptul de a depune o plangere in fata ANSPDCP daca am ingrijorari cu privire la modul in care sunt prelucrate datele cu caracter personal.

Prin prezenta declaratie de consimtamant confirm ca sunt de acord ca datele cu caracter personal sa fie prelucrate prin canalul mentionat mai sus si utilizate in limitele astfel specificate.

Nume Ana Stefan
Data 17.03.2023

Semnatura pacient/reprezentant legal______________________

Figure 5.7 *Consent form.*

All of these objects help manage flows in the healthcare facility, such as ordering medications, sending notifications to teams that have different roles in patient care, creating records when a member needs to perform a medical procedure or take certain samples, analyses, they appear scheduled for easy management and the existence of a flow of notifications upon completion of certain activities, receipt of medical results, etc.

b. Service Console - the console is dedicated to the hospital's technical team (medical engineers) who deal with medical equipment - maintenance, calibration, settings, use, etc. The Service Console consists of several objects:

I. Account - Contains the information about the medical institution – “Ana Medical”
¬This object contains Related Lists to other relevant objects:

a. Business Licenses – object that contains information about the institution's fiscal certificates and more

b. Providers – information about the institution's providers (address, website, contact, fiscal information)

c. Patients – patient list of the institution

d. Account Assets – equipment list of insistence

II. Assets – an asset is a piece of equipment owned by the medical institution. It contains information with product images, clinical images, equipment videos, hands-on, etc., equipment documents (product tree, service manuals, equipment declaration of conformity, etc.).

III. Maintenance Plans - at the level of an equipment, maintenance plans are created for the installed equipment base, and at the level of these records, interventions can be automatically created on the respective equipment according to certain criteria.

IV. Work Orders - the medical engineers responsible for the revisions on the equipment, carry out the intervention and generate a report about that intervention:

V. Work Types - the types of interventions that can be performed on medical equipment

c. **Pharmacy Console** - the console is dedicated to the team in charge of the hospital's pharmacy that deals with the drug procurement flow. The Pharmacy Console consists of several objects:

i. Orders – for ordering medicines

ii. Stock – the institution's stock of equipment

iii. Medication List – the nomenclature of medicines

iv. Patient Medication List – linking object between the drug nomenclature and the drugs administered to patients. Objects that help us manage the circulation of medicines.

d. **Patient Console** - the console is dedicated to patients. This is the portal through which patients can view results, treatment plans, medical history, etc. The Patient Console consists of a single object – Patiens, and depending on the patient who logs in, it contains a single record about a single patient with information relevant to him. When scanning the QR code on any medical document generated by the institution, a direct referral is made to this patient console.

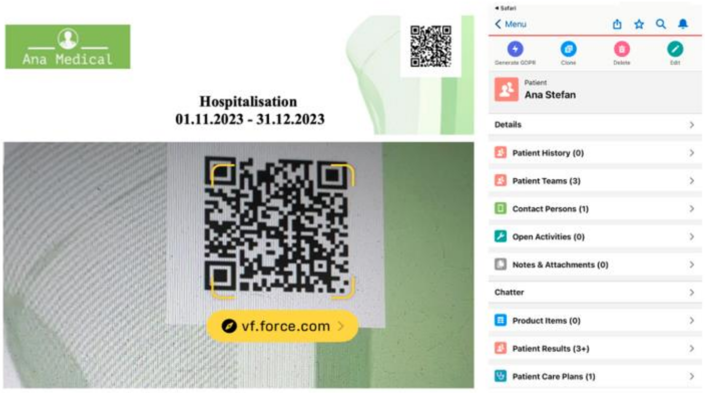

Figure 5.39 *QR code and how it works*

e. Image Classification Console - the console is dedicated to the team of medical engineers specialized in Artificial Intelligence. The console contains the necessary objects for the acquisition and processing of medical images. Medical procedures performed for a patient, images acquired and classes for classification. Each class contains either purchased images or an archive containing purchased images. Both variants can be loaded simultaneously.

Skin pathology detection module - As also exemplified in [33] only that more classes of images were used, namely Melanoma, Actinic keratosis, Seborrheic keratosis, Basal cell carcinoma and class of images without pathologies. Video solution – custom classification application for real-time classification of skin images. The mode of operation of the module is based on the camera of a computer or mobile device on which the platform of the information system is open, it covers the area with the lesion and in real time displays one of the four classes on which the classifier was trained (Figure 4.45).

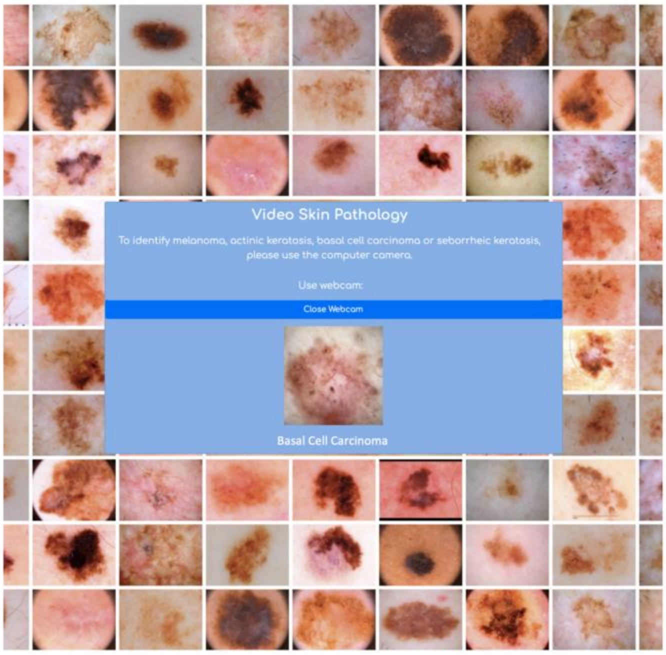

Figure 4.45 *Video classification module for skin lesions [57]*

Eye pathology detection module – in which classes for Glaucoma, Cataract and Diabetic Retinopathy were used to create the model. Image solution - custom classification application for the classification of funduscopic images acquired and stored in the database. The modus operandi of the module is based on uploading an image and receiving the subsequent response according to the 3 pathologies on which the classification model is trained (Figure 4.46).

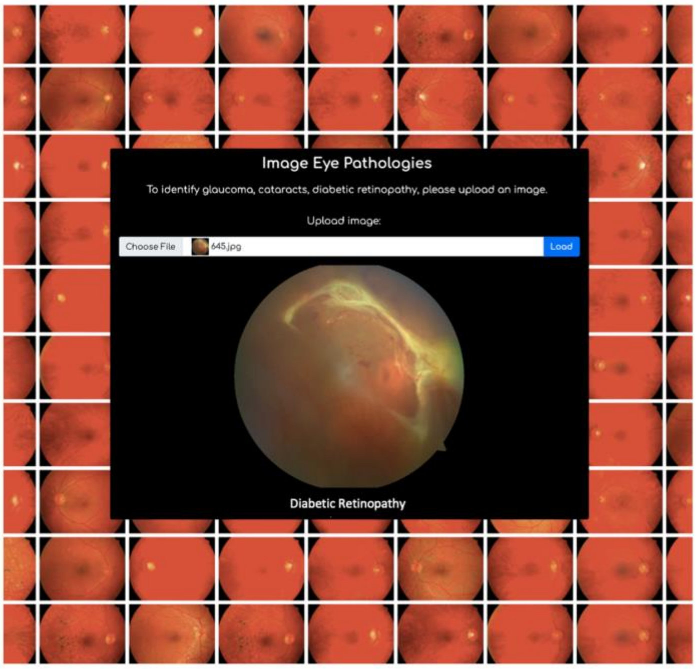

Figure 4.46 *The classification module using images for the classification of eye pathologies [57]*

5.2 Discussions

In this work, two development directions were followed. The first direction consisted in creating an algorithm for classifying medical images that can be successfully integrated into a medical information system. The process consisted of extensive research in the field of machine learning, deep learning and artificial intelligence to achieve automatic diagnosis, neural networks and existing architectures, the differences between them and how they can be used to achieve an automatic classifier (the focus being on the classification retinal fundus images and skin lesions) and databases that can be used for classifier training, research into existing frameworks for implementing such a classifier, web-based platforms for creating classification models that can be later integrated in a classification module within a MIS and the state of the art of automatic medical image classification applications. The best result among the web-based models was Google Vertex AI, but the most versatile and flexible model was Google Teachable Machine and among the custom algorithms the classifier based on the fusion of decisions built in MATLAB, which at the time of the research did not find a viable variant of integration with a computer system. The second direction consisted in the development of a medical information system, a process that consisted in the extensive research of the science of information systems, their types, components, the benefits they bring and the objectives they fulfill, the criteria of evaluation of them, of the planning method and the stages necessary to implement a large-scale project. Subsequently, the MIS and their components were evaluated together with the specific requirements of the medical field. The developed system helps to track most of the workflows in a medical institution that has a patient access platform, separate consoles for the teams of pharmacists, doctors and medical engineers, with the separate mode of processing and clarifying medical images both in real time and and by analyzing a pre-acquired image external to the system. The implemented system is a new one that allows the expansion of multidisciplinary functionalities.

5.4 Personal contributions

Personal contributions are present in all phases of development of both the MIS and the automatic module for the detection of eye and skin pathologies. The development and implementation of a medical informatics system with an integrated module for automatic detection of pathologies involves several crucial steps to ensure a successful implementation of the solution, such as:

a. Identifying needs and planning:

i. Identifying the specific objectives of the information system and the objectives of integrating an automatic module for the detection of pathologies in the MIS

ii. Develop a project plan that describes the scope, timeline, resources, and budget.

iii. Creating a data management strategy, security and integration with existing systems, defining the types of pathologies that need to be detected and the scope of the detection module.

iv. Establish target user groups (healthcare providers, administrators, patients) and their requirements.

b. System design:

i. MIS architecture design, taking into account scalability, databases, data security and its possible interoperability with other healthcare systems.

ii. Establishing the data flow and integration points between the MIS and the automated pathology detection module, the choice of technologies, programming languages and technology frameworks to be used for development.

c. Data integration:

i. Establish data integration mechanisms to transparently transfer patient data, medical images and relevant clinical information between the MIS and the pathology detection module.

d. Development of the information system

i. Development of functionalities and specifications based on system design

ii. Implementing the necessary data structures, algorithms and logic

iii. Designing an intuitive and user-friendly interface for healthcare providers to interact with MIS and access pathology detection results.

iv. Create visualizations and interactive tools to display detected pathologies and relevant clinical data.

e. Developing the machine learning model:

i. Developing and/or selecting suitable machine learning algorithms and models for automatic detection of pathologies based on chosen pathologies and performing case studies confirming the choice made.

ii. Collect and prepare a complete pathology image dataset for training and validation, train and optimize the machine learning model using the dataset to achieve accurate detection results.

f. Integration and testing:

i. Integration of machine learning model and pathology detection module in MIS

ii. Conducting tests to ensure accurate operation of the autodetection module (integration, system and UAT tests) and performing global tests to validate data flow, integration and user interface functionality.

g. Data privacy and security:

i. Implementing data privacy and security measures to protect patient information and medical images.

ii. Ensuring compliance with relevant data protection regulations such as HIPAA or GDPR.

h. Implementation of the system:

i. Implementation of fully integrated MIS with automatic pathology detection module within the healthcare organization.

ii. Configuring databases and infrastructure components, inserting and modifying system data, monitoring system performance.

5.5 Future directions

The next steps to improve the automated pathology detection module solution presented in the thesis include replacing Visualforce pages with LWC, which is a more challenging development solution. However, the advantage of LWC is that the application uses browser resources, not server resources. Another direction to improve the solution consists in the integration of different platforms with the computer system in order to use the computing power present in other platforms built to implement complex solutions, such as integration with MATLAB and expanding the diagnosis of various pathologies using other types of data. than medical images. At the same time to use this solution in the medical industry, another useful tool would be a telemedicine module that can connect patients, medical institutions providing services and emergency services in order to collect data in real time in the computer system. Another useful tool could be integration with native Salesforce Tableau analytics system to analyze collected data to identify trends, patterns and changes in patient condition over time, predictive analytics to forecast potential pathology risks based on historical data and patient profiles combined with the integration of genetic databases to improve pathology detection by considering genetic predispositions and mutations.

At the same time, a direction would be the opportunity to increase the awareness and education of patients and provide them with resources to increase the understanding of the importance of early detection, the potential risks of various pathologies detected late and the actions recommended for their treatment. These ideas can help improve the capabilities and usability of a MIS that integrates automated pathology detection, ultimately improving patient care, diagnostic accuracy, and medical workflows. In order to expand the use of the MIS, it is necessary to develop a financial-accounting module, integration with medical image acquisition equipment or analysis equipment, implementation of bidding modules for medical and health insurance packages, implementation of a data management module about employees and a marketing console to promote the institution and the MIS built. In post-doctoral studies, it is desired to test the system in a beta version, in a medical insistence. The present summary includes in a condensed form the contents of the chapters of the thesis. The numbering of chapters, sub-chapters and sections, figures and tables correspond to those in the thesis. The bibliography is selective with the works present according to the abstract

Bibliography

[33] El-Khatib, H., **Ștefan, A.-M.,** & Popescu, D. (2023). Performance Improvement of Melanoma Detection Using a Multi-Network System Based on Decision Fusion. Applied Sciences, 13(18), 10536–10536. https://doi.org/10.3390/app131810536

[57] **Ștefan, A.-M.;** Rusu, N.-R.; Ovreiu, E.; Ciuc, M. Advancements in Healthcare: Development of a Comprehensive Medical Information System with Automated Classification for Ocular and Skin Pathologies—Structure, Functionalities, and Innovative Development Methods. Appl. Syst. Innov. 2024, 7, 28. https://doi.org/10.3390/asi7020028

[58] **Ștefan, A.-M.;** Rusu, N.-R.; Ovreiu, E.; Ciuc, M. Empowering Healthcare: A Comprehensive Guide to Implementing a Robust Medical Information System - Components, Benefits, Objectives, Evaluation Criteria, and Seamless Deployment Strategies. Applied System Innovation. 2024; 7(3):51. https://doi.org/10.3390/asi7030051

[59] **Ana-Maria ȘTEFAN,** Elena OVREIU, Mihai CIUC, "Comparative analysis of web-based machine learning models", Romanian Journal of Information Technology and Automatic Control, ISSN 1220-1758, vol. 34(2), pp. 49-63, 2024. https://doi.org/10.33436/v34i2y202404

[89] **Ștefan, A.-M**., Paraschiv, E.-A., Ovreiu, S., & Ovreiu, E. (2020). A Review of Glaucoma Detection from Digital Fundus Images using Machine Learning Techniques. 2020 International Conference on E-Health and Bioengineering (EHB). https://doi.org/10.1109/ehb50910.2020.9280218

[136] Ichim, L., & Popescu, D. (2020). Melanoma Detection Using an Objective System Based on Multiple Connected Neural Networks. IEEE Access, 8, 179189–179202. https://doi.org/10.1109/ACCESS.2020.3028248

[147] Giotis, I., Molders, N., Land, S., Biehl, M., Jonkman, M. F., & Petkov, N. (2015). MED-NODE: A computer-assisted melanoma diagnosis system using non-dermoscopic images. Expert Systems with Applications, 42(19), 6578–6585. https://doi.org/10.1016/j.eswa.2015.04.034

[153] Ragowsky, A., Ahituv, N., & Neumann, S. (1996). Identifying the value and importance of an information system application. Information & Management, 31(2), 89–102. https://doi.org/10.1016/s0378-7206(96)01072-5

[206] El-khatib, H.**, Ștefan, A.-M**., & Popescu, D., Melanoma Automated Detection System Integrated with an EHR Platform – UPB Scientific Bulletin Series C: Electrical Engineering and Computer Science 2024. Vol. 86, Iss. 1, 2024. ISSN 2286-3540