Digital Health and Informatics Innovations for Sustainable Health Care Systems
J. Mantas et al. (Eds.)

doi:10.3233/SHTI240558

A Digital Tool Supporting Pathology Practice and Identifying Leucocytes

Dumitru-Cristian APOSTOL[a,1], Antonela-Maria CHIUZBĂIAN[b], Mihaela CRISAN-VIDA[a], Lăcrămioara STOICU-TIVADAR[a] and Marius MARCU[a]
[a]*Faculty of Automation and Computers, University Politehnica Timișoara, Romania*
[b]*Department of Pathology, Emergency County Hospital, Cluj-Napoca, Romania*

Abstract. The aim of this work is to deliver an all-in-one package that contains both the part where the pathologist can manipulate the data as well as predefined models, altogether with the digital pathology interface with a comprehensive component that provides traceability between the identified leucocytes and the underlying possible outcomes of the potential disease. The aim is to directly provide the number of leucocytes and the mass of the cell, only from the image, with minimal intervention from the pathologist, necessary to have a PoC (Proof of Concept) or a prototype. The model was trained on a dataset of around 20,000 models, and the achieved accuracy was approximately 85%. Approximately 82% of the identified areas of interest, as determined by the models, were true positive predictions. The models correctly identified approximately 89% of the actual positive instances - areas of interest - identified by the pathologist. Approximately 6% of the total actual negative instances were incorrectly classified as positive by the models. The tool provides visual scripting, reducing the learning curve for pathology analysis techniques and offers an intuitive interface for healthcare professionals.

Keywords. Pathology, Artificial Intelligence, Intelligent Models.

1. Introduction

In the dynamic landscape of medical diagnostics, the increasing role of artificial intelligence (AI) introduces a paradigm shift. The need for a standardized set of rules and paradigms regarding the use of AI in pathology was very high in the early days of digital pathology [1]. With the evolution and expansion of pathology practices, there arises a pressing demand for tools that not only visualize the intricate details of pathology-related data but also establish a seamless connection between the AI-driven algorithms and the overarching architectural understanding of pathological processes. The need for new paradigms and standards in Digital Pathology surged together with the enhancement of different AI training models [1].

This paper undertakes an exploration of diverse implementations for AI-driven analysis in pathology, akin to the comparative analysis of UML code (generation in the software domain [2-4]). Much like the early days of code modeling, various AI applications strive to bridge the gap between intricate data analytics and a comprehensive understanding of the pathology at hand. Existing AI implementations, ranging from age

[1] Corresponding Author: Dumitru-Cristian APOSTOL; E-mail: dumitru.apostol@student.upt.ro.

recognition to predictive modeling, aim to enhance accuracy, efficiency, and objectivity in pathology diagnoses.

Numerous endeavors, including commercial products like PathAI Translator, DeepMind Health, and others, have ventured into translating pathology-related data into actionable insights through AI. However, as with early code generation tools, these solutions may exhibit limitations in terms of interpretability, completeness, and adaptability to different pathology domains. All of them have shortcomings, especially regarding readability and completeness. First, we will take a brief look at some previous attempts at different AI tools related to the field of pathology, discuss their achievements, elaborate a method that aims to reduce faults of the other existing solutions, and the development of a stand-alone application utilizing the devised method.

A feature that we consider important that is missing in most digital pathology applications are the annotations that provide insight in to possible problems [5]. It is our goal to provide a functional library containing the most prevalent distinguished patterns. There are two relevant groups, the one related to the area of interest, that is constituted of the areas of interest that are highlighted later, at the end of the image processing, the other group is related to the don't care part or the non-interesting areas which usually is something that it doesn't contain possible leucocytes. Related to the class, there is a class for comparing the specific area of interest with the leucocytes that are already in the trained model, as well as the ones that are newly identified.

2. Related work

First In the evolving realm of pathology and Artificial Intelligence (AI), significant strides have been made, particularly in the past decade. However, AI in pathology remains a relatively nascent field within the broader spectrum of healthcare and technology. Despite notable advancements, there persists a need for continuous innovation, particularly in addressing accessibility challenges and enhancing the existing tools [6].

Our analysis of current tools in the AI pathology domain reveals different ramifications. On one side, there are sophisticated solutions catering to industry needs, yet often financially, students or emerging private laboratories, can't afford them. On the other side, free alternatives, while offering desired functionality, come with their own set of limitations, such as design rigidity, usage constraints, and user experience challenges. The focus is on refining the user interface, making it intuitive, and ensuring a seamless interaction for users with varying levels of expertise [7,8]. Improving user experience is pivotal for maximizing the adoption and impact of AI in pathology. Beyond accessibility, the innovation lies in novel problem-solving approaches [9,10]. Our work explores inventive methods, offering solutions that go beyond the conventional and address the unique challenges posed by AI applications in pathology. The cloud serves as a powerhouse, delivering superior computation power and scalable storage crucial for diverse pathology applications. Its attributes, including on-demand scalability, persistent availability, and adaptability, create an ideal environment for handling the complex data and computational demands of pathology analysis. Within pathology, Deep learning (DL) has announced revolutionary capabilities in Natural Language Processing (NLP). This has translated into improved performance in tasks like analyzing pathology reports, identifying key entities, and enhancing the understanding of complex medical language [11].

3. Solution Architecture and Testing Setup

In this paper we present the design and implementation of the architecture that was created in Python with the help of different libraries, like NumPy, OpenCV and TensorFlow.

The application has two main components, the UI part which has a prompter to take a folder with images and analyze the images that are in each specific file, the image is processed through specific layers that are resulted from the segmentation of the images. The result generation component which produces the segmentation of the image, meaning that each image will be analyzed and after that will be spliced into two parts of data one named "important data" and the other named "other elements". The components will be shown in detail in each subsequent part of the paper. The conceptual architecture of the application is represented in Figure 1.

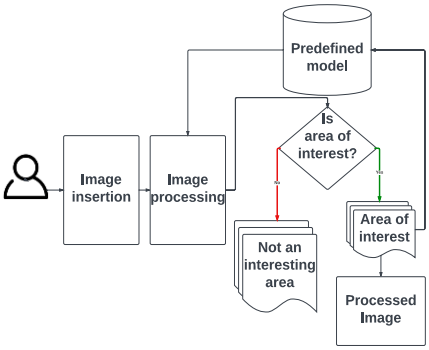

Figure 1. Diagram reflecting the system

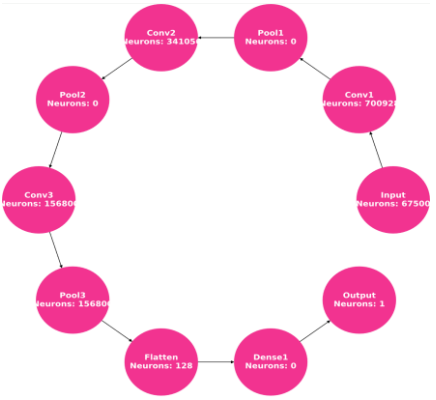

Figure 2. Diagram representing the architecture of the neuronal network.

3.1. Image analysis

The core of the application is image analysis. The receiver, meaning the daemon that is working on the image processing, receives the event from the User Interface (UI), regarding the folder that needs to be analyzed, notifying the component that started an analysis process. The information is segmented and fed to the Leucocyte Analysis, which in turn it will convert the image into the two classes of objects. Once an image that is relevant to the specific class will be added into the specific group. After all the objects that were found in the image will be added to each group, it will circle the objects that are relevant to the "important information" group. Our study involved curating a dataset for pathological analysis by sourcing data from two main repositories: the Raabindata and Pathology Outlines. With approximately 15,000 samples from Raabindata and the other of 5,000, approximately from Pathology Outlines, we aimed for a diverse representation of pathological conditions related to the case study. The solution is based on the following layers: Input Layer, three distinct convolutional layers, three distinct Pooling Layers, Flatten layer, Dense Layer, Dropout Layer as well as the output layer. The total number of neurons that were used is around ~1,422,000, represented in Figure 2.

3.2. Implementation Challenges

An example of a technical challenge faced during the implementation of the model is the assignment of specific threads to specific images. Our approach was to use different file descriptors that retain the data for each specific image. Initially, the process of identifying various cells begins with a focus on a specific region of the body. However, I intend to broaden this approach to encompass the identification of cells in additional anatomical areas.

4. Experimental Results

Figure 3 presents the first experimental results with an image of the tissue before the analysis, and Figure 4 shows the results after the image processing. After training approximately 20,000 models, the achieved accuracy was approximately 85%, with a True Positive Rate (TPR) of around 82% and a Sensitivity (Recall) of approximately 89%. The False Positive Rate (FPR) is around 6%.

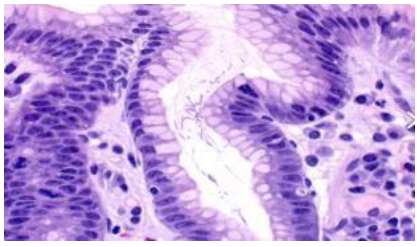

Figure 3.a. A sampling before the implementation of the filtering [12]

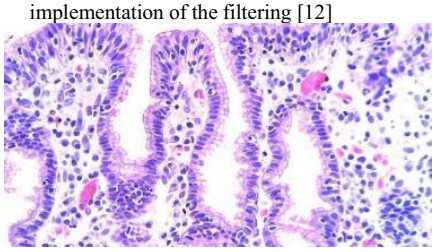

Figure 3.b. A sampling before the implementation of the filtering [13]

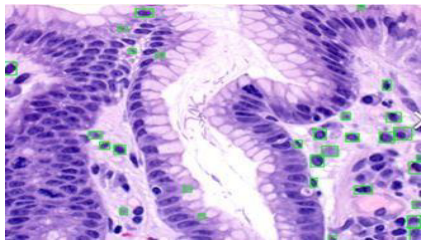

Figure 4.a. The sample after the result was analyzed

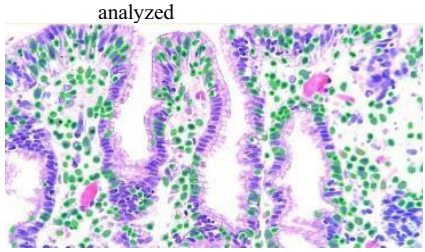

Figure 4.b. The sample after the result was analyzed

Table 1 presents a summary of the results. As seen below in the corresponding figures we can see one that it's easy to recognize the specific areas of interest, but in Figure 4.b, we can see intraepithelial lymphocytosis which is a very important factor in determining the celiac disease.

Table 1. Metrics

Metric	Value
Total models trained	~20,000
Accuracy	~85%
True Positive Rate (TPR)	~82%
Sensitivity (Recall)	~89%
False Positive Rate (FPR)	~6%

5. Conclusions and Future Work

In the realm of pathology diagnostics, there is a need for innovative tools that move beyond conventional approaches. This forward-thinking application is designed to adapt to diverse diagnostic scenarios, prioritizing flexibility, and adaptability. The tool aims to expand language support to accommodate different healthcare professionals and modalities of pathology data. Operational through a specialized processing pipeline, this application leverages advanced algorithms for a comprehensive understanding of pathological conditions. It presents outcomes in a clear, visual format, including benchmarked metrics, aiding in the interpretation of diagnostic results. Looking forward, the application envisions language expansion, exploration of cloud-based solutions, and customizable diagnostic approaches for users, as well as adding more coloration techniques. Future developments also include the incorporation of intelligent correction mechanisms, automating proposed corrections to streamline diagnostic workflows. Future iterations plan to introduce metrics assessing the size and complexity of generated insights, extending beyond traditional parameters. In conclusion, this application aims to redefine the landscape of AI applications in pathology, providing a user-friendly, adaptable tool for healthcare professionals in the dynamic realm of pathology diagnostics.

Acknowledgment

This paper was partially supported by a grant from the Ministry of Research, Innovation and Digitization, CNCS/CCCDI—UEFISCDI, project number ERANET-CHIST-ERA-ECOMOME, within PNCDI III.

References

[1] Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. J Pathol Inform. 2016 Jul 26;7:29. doi: 10.4103/2153-3539.186902.
[2] Tizhoosh HR, Pantanowitz L. Artificial Intelligence and Digital Pathology: Challenges and Opportunities. J Pathol Inform. 2018 Nov 14;9:38. doi: 10.4103/jpi.jpi_53_18.
[3] Wells A, Patel S, Lee JB, Motaparthi K. Artificial intelligence in dermatopathology: Diagnosis, education, and research. J Cutan Pathol. 2021 Aug;48(8):1061-1068. doi: 10.1111/cup.13954.
[4] Nagpal K, et al. Development and Validation of a Deep Learning Algorithm for Gleason Grading of Prostate Cancer From Biopsy Specimens. JAMA Oncol. 2020 Sep 1;6(9):1372-1380.
[5] Wang P, Xiao X, Glissen Brown JR, Berzin TM, et al. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. Nat Biomed Eng. 2018 Oct;2(10):741-748.
[6] You K, et al. Comparison of Core Needle Biopsy and Surgical Specimens in Determining Intrinsic Biological Subtypes of Breast Cancer with Immunohistochemistry. J Breast Cancer. 2017;20:297-303.
[7] Al-Halafi AM. Applications of artificial intelligence-assisted retinal imaging in systemic diseases: A literature review. Saudi J Ophthalmol. 2023 Oct 19;37(3):185-192. doi: 10.4103/sjopt.sjopt_153_23.
[8] Dawson H. Digital pathology - Rising to the challenge. Front Med (Lausanne). 2022 Jul 22;9:888896. doi: 10.3389/fmed.2022.888896. Erratum in: Front Med (Lausanne). 2023 Mar 15;10:1180693.
[9] McGenity C, Clarke EL, Jennings C, Matthews G, et. al. Artificial intelligence in digital pathology: a systematic review and meta-analysis of diagnostic test accuracy. NPJ Digit Med. 2024 May 4;7(1):114.
[10] Zehra T, et al. A suggested way forward for adoption of AI-Enabled digital pathology in low resource organizations in the developing world. Diagn Pathol. 2023 May 18;18(1):68.
[11] Baxi V, Edwards R, Montalto M, Saha S. Digital pathology and artificial intelligence in translational medicine and clinical practice. Mod Pathol. 2022 Jan;35(1):23-32. doi: 10.1038/s41379-021-00919-2.
[12] Helicobacter pylori, https://www.pathologyoutlines.com/topic/stomachhelicobacter.html, Accessed in 10.07.2023.
[13] Celiac disease, https://www.flickr.com/photos/euthman/48467778276, Accessed in 12.07.2023.