Applied Computer Vision on 2-Dimensional Lung X-Ray Images for Assisted Medical Diagnosis of Pneumonia

Lucky V. Aquino [1], Angelo Justin T. Chuakay [2], Vince D. Fernandez [3],
Ralph Joseph S.D. Ligueran [4], Emmanuel H. Valencia [5]

Bachelor of Science in Computer Science, Asian Institute of Computer Studies
AICS Bldg., Commonwealth Ave. cor. Holy Spirit Drive, Brgy. Don Antonio, Quezon City, Philippines

[1] aquino170634@gmail.com
[2] ajtchuakay@gmail.com
[3] vincefernandez02715@gmail.com
[4] ligueran170127@gmail.com
[5] emmanuelvalencia008@gmail.com

Abstract:

The advancements of artificial intelligence in the field of healthcare and biomedical informatics were motivated by two factors: to increase the efficiency in the analysis of medical data and to enhance the accuracy of the analysis to prevent cases of medical misdiagnosis. This research focuses on the application of a specific subfield of artificial intelligence referred to as computer vision in the analysis of 2-dimensional lung x-ray images for the assisted medical diagnosis of pneumonia. A convolutional neural network algorithm was implemented in a Python-coded, Flask-based web application that can analyze x-ray images for the detection of pneumonia. Since convolutional neural network algorithms rely on machine learning for the identification and detection of patterns, a technique referred to as transfer learning was implemented to train the neural network in the identification and detection of patterns within the dataset. Open-source lung x-ray images were used as training data to create a knowledge base that served as the core element of the neural network.

Keyword: *Artificial Intelligence, Biomedical Informatics, Extended Intelligence, Computer Vision, Convolutional Neural Network*

I. INTRODUCTION

The advancements of computer vision technology have led to several significant milestones in the field of artificial intelligence. Computer vision technologies have been applied to various platforms, such as social networking sites, closed-circuit television camera networks, and robotics among others. Recently, due to the growing interest in the development of A.I.-assisted medical diagnostic technologies and healthcare/biomedical informatics, computer vision is slowly but gradually being utilized in the diagnosis of diseases [1]. Medical images, such as x-ray images and CT scans, are highly useful in the diagnosis of certain diseases. Using computer vision, the detection and analysis of these diseases can theoretically yield higher accuracy and reliability in medical diagnosis.

Fig. 1. 2-Dimensional Lung X-Ray Images [2]

There exist various algorithms for the implementation of computer vision technologies. One of them is the Convolutional Neural Network (CNN). A convolutional neural network is a type of artificial neural network. Artificial neural networks are either hardware or software implementations that imitate the structures and operation of natural neural networks that exist in the human brain. A natural neural network is composed of neurons and their corresponding dendrite-axon connections with one another. This complex connection of neurons gave humanity its intelligence and sentience. In computer science, researchers are attempting to imitate this complex network to give machines human-like learning and problem-solving capabilities. This is essential since modern scientific, engineering, financial, and economic problems require vast new approaches, including faster and more efficient computational and problem-solving techniques.

Artificial neural networks are composed primarily of three layers: the input layer, hidden layer, and output layer. Within the different types of artificial neural networks, there exists a specific type known as a deep neural network. Deep neural networks are artificial neural networks that have more than one hidden layer. Deep neural networks also rely on a technique referred to as machine learning to look for patterns within datasets.

A convolutional neural network, therefore, is a deep neural network based on the mathematical operation of convolution that relies on machine learning to detect patterns within datasets. For the convolutional neural network to "learn", training datasets must be fed into the neural network during the machine learning phase [3].

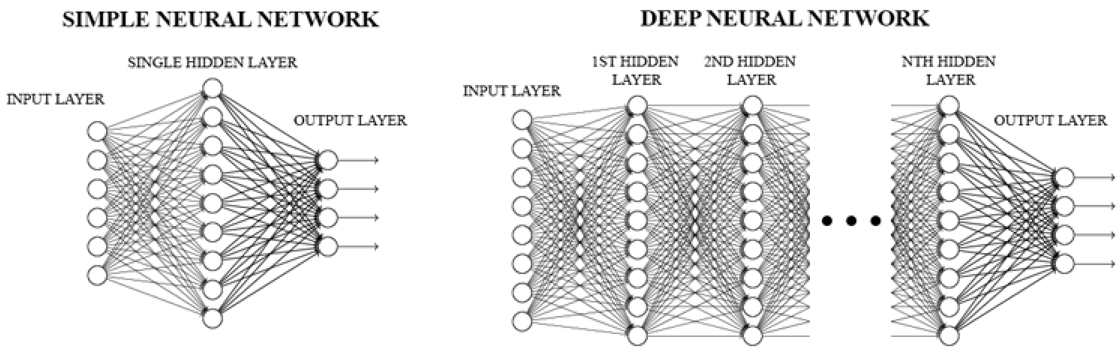

Fig. 2. Comparison of Simple and Deep Neural Networks

Once the machine learning phase is completed using training datasets, the CNN can now use the patterns stored in its knowledge base to analyze and detect certain features within the images. When applied to medical diagnostic technologies, CNN can detect features in medical images that are signs of certain diseases [4]. In this research, the proponents utilized the CNN algorithm to develop a web application that can be used in the diagnosis of pneumonia.

The primary concern in the development of this kind of A.I.-assisted medical diagnostic technology is to prevent cases of medical misdiagnosis. Medical misdiagnosis is cases where medical practitioners unintentionally commit errors in the diagnosis of diseases on patients [5, 6]. False negatives or false positive misdiagnosis can occur due to various factors. Primarily, errors within the medical data are the primary cause. In the case of misdiagnosis of pneumonia through x-ray images, the low contrast ratio in the image can lead to errors upon visual inspection [7]. Human error can also be a significant factor in the causation of medical misdiagnosis. Such cases of medical misdiagnosis of pneumonia can lead to life-threatening conditions [8].

Medical misdiagnosis of pneumonia had occurred in impoverished regions of the world, such as in Sub-Saharan Africa and Southeast Asia. Based on preliminary research, a detailed and comprehensive collection of data regarding the misdiagnosis of pneumonia was conducted in the nations of Nigeria and Malaysia [9, 10]. These studies have revealed the alarming frequency of the misdiagnosis of pneumonia.

Misdiagnosis of pneumonia will lead to mistreatment, or in some cases, overtreatment of the patient. According to the medical article written by Zulfiqar A. Bhutta, 20% of children infected with pneumonia die each year, with misdiagnosis as one of the contributing factors. A case of overtreatment, as explored in the article, is when a patient with viral pneumonia is misdiagnosed with bacterial pneumonia. This misdiagnosis leads to overtreatment, where antibiotic drugs were erroneously administered. Cases like these lead to antimicrobial resistance, which undermines the general effectiveness of antibiotic drugs [11]. Such cases are of great relevance and significance in the field of medical research.

Analysis of medical images is also essential in the discovery of new types of diseases. At the time of the writing of this research paper, the COVID-19 pandemic is still ravaging various nations around the world. Before the implementation of nasal swab testing for the detection of the presence of the SARS-Cov-2 virus in patients, signs of the COVID-19 disease were first detected through lung x-ray images [12]. Due to the novelty of the disease, the x-ray results were misdiagnosed as simply ordinary pneumonia. However, as more COVID-19 x-ray images were gathered and compared with pre-existing x-ray images of pneumonia, patterns were later revealed. This case exemplifies the importance of medical imaging in the diagnosis of diseases. Medical imaging can therefore be improved to produce more accurate results through the help of artificially intelligent tools.

Advancements in artificial intelligence research also gave birth to a new sub-field known as extended intelligence. Extended intelligence is a novel new approach that attempts to combine the strengths of human intelligence with artificial intelligence to solve various problems. Extended intelligence acknowledges the inherent weaknesses of current A.I., such as the lack of true creativity and inherent algorithmic, software, and hardware limitations. Therefore, artificial intelligence should be regarded as just a tool for solving various problems. Human ingenuity combined with the strengths of artificial intelligence could solve the majority of the world's problems [13].

The objectives of the study are the following:

- To develop a Flask-based web application that can detect pneumonia from 2-D lung x-ray images.
- To implement a Convolutional Neural Network algorithm for the web application.
- To use an open-source dataset to train the neural network.

II. LITERATURE REVIEW

This research is divided into six major topics: First, the significance of artificial intelligence in medicine, including its impact in the field of extended intelligence; Second, a study of various cases of medical misdiagnosis of pneumonia; Third, convolutional neural networks and machine learning concepts in computer vision; Fourth, the application of computer vision in the detection of pneumonia on 2D lung x-ray images; Fifth, a study on web-based, medical diagnostic and biomedical informatics technologies; And sixth, a study on web frameworks for the implementation of the CNN algorithm in web applications.

Several textbooks and research papers in high-level mathematics and statistics were also used to define and explore various mathematical concepts to be used, including the statistical metrics for the experimental design to be used for testing.

A. Artificial Intelligence and Extended Intelligence in Medicine

For the topic of the significance of artificial intelligence in medicine, the proponents have selected two research articles. The first research article is "Artificial Intelligence in Medicine", authored by Pavel Hamet and Johanne Tremblay, and published in the journal *Metabolism: Clinical and Experimental* in 2017. This research deals with the development of various implementations of A.I. in medicine, including its societal, cultural, and ethical issues [14].

The second research article is "Extended Intelligence", authored by Joichi Ito, and published in the journal *PubPub* in 2016. This research focuses on the various fields of application of A.I. to further the paradigm of extended intelligence. Examples of medical E.I. were discussed [15].

Information from these research articles is essential in understanding the various socio-cultural and professional impacts of A.I. and E.I. in this research.

B. Study of Various Cases of Medical Misdiagnosis

For the topic of the study of various cases of medical misdiagnosis of pneumonia, the proponents have selected four research articles. The first research article is "Misdiagnosis of Pneumonia, Bronchiolitis and Reactive Airway Disease in Children: A Retrospective Case Review Series in South East, Nigeria", authored by Egbuna Obidike, Tagbo Oguonu, Joy Eze, Ikenna Ndu, Adaeze Ayuk, Mildred Ukoha, Obinna Nduagubam, and Ikechukwu Ogbonna, and published in the journal *Current Pediatric Research* in 2018. This research focuses on the collection and categorization of medical data about pneumonia in children in Nigeria. The research concluded that review and further enhancement of the criteria for the diagnosis of pneumonia are essential to combat misdiagnosis in the region [16].

The second research article is "Misdiagnosis of Community-acquired Pneumonia in Patients Admitted to Respiratory Wards, Penang General Hospital", authored by Ang Choon Seong, Kelvin Beh Khai Meng, Yeang Li Jing, Chin Yuen Quan, Khor Inn Shih, Yoon Chee Kin, and Irfhan Ali Bin Hyder Ali, and published in the journal Medical Journal of Malaysia in 2020. This research focuses on the collection, categorization, analysis, and interpretation of medical data on pneumonia patients in a pediatric hospital in Malaysia. The research concluded the most misdiagnosis was the result of misinterpretation by medical practitioners. A solution to mitigate this problem through enhanced administrative policies was recommended [17].

The third research article is "False Negative Chest X-Rays in Patients Affected by COVID-19 Pneumonia and Corresponding Chest CT Findings", authored by M. Cellina, M. Orsi, T. Toluian, C. Valenti Pittino, and G. Oliva, and published in the journal Radiography in 2020. This research explores four cases of misdiagnosis of COVID-19 pneumonia as regular pneumonia. The research concluded that further research and awareness of false-negative results are essential for this type of emerging disease [18].

The fourth research article is "Childhood Pneumonia in Developing Countries", authored by Zulfiqar A. Bhutta, and published in the journal BMJ in 2006. This medical article focused on the impact of pneumonia in children's health, with research showing that 20% of children suffering from pneumonia die each year, with misdiagnosis as one of the contributing factors. Cases of mistreatment and overtreatment of pneumonia due to misdiagnosis were also explored [19].

These are factual and comprehensive cases of medical misdiagnosis of pneumonia on patients published in credible and respected medical journals. These cases justify the need for a computational tool that can assist medical practitioners in the diagnosis of pneumonia for the purpose of confirmatory analysis.

C. Convolutional Neural Networks and Machine Learning for Computer Vision

For the topic of convolutional neural networks and machine learning concepts in computer vision, the proponents have selected three research articles. The first research article is "Machine Learning in Computer Vision", authored by Asharul Islam Khan and Salim Al-Habsi, and presented at the conference *International Conference on Computational Intelligence and Data Science (ICCIDS 2019)*. This research deals with the concepts of supervised, unsupervised, and semi-supervised machine learning and various new and emerging techniques on how to implement these concepts in computational machines. The research concluded that machine learning technology will advance further as new and more algorithms are developed [20].

The second research article is "Overview of Deep Learning in Medical Imaging", authored by Kenji Suzuki, and published in the journal *Radiological Physics and Technology* in 2017. This research deals with the application of deep learning algorithms in medical imaging. The research concluded that deep learning will be the mainstream technology in medical imaging in the next few decades [21].

The third research article is "Very Deep Convolutional Networks for Large-Scale Image Recognition", authored by Karen Simonyan and Andrew Zisserman, and presented in the conference *3rd International Conference on Learning Representations, ICLR 2015*. This research deals with the development of the VGGNet algorithm for computer vision, and its variant algorithms. The research concluded with accuracy and precision rating results ranging from 70 to ~99% on various testing metrics for the algorithms [22].

These three research articles provided the proponents with comprehensive information about machine learning, with the VGG-16 algorithm, a variant of the VGGNet, as the primary computer vision algorithm that was selected to be implemented for this research.

D. Application of Computer Vision on 2-Dimensional Lung X-Ray Images for the Detection of Pneumonia

For the topic of the application of computer vision on 2-dimensional lung x-ray images for the detection of pneumonia, the proponents have selected seven research articles. The first research article is "Detection of Pneumonia Using Deep Transfer Learning", authored by K. Reddy Madhavi, G. Madhavi, B. Rupa Devi, and Padmavathi Kora, and published in the journal *International Journal of Advanced Trends in Computer Science and Engineering* in 2020. This research deals with the testing for various computer vision algorithms, including the VGG-16 algorithm, for the detection of pneumonia. The research concluded with an accuracy rating result of 99.4% for the VGG-16 algorithm [23].

The second research article is "Pneumonia Detection Using Convolutional Neural Networks", authored by Sammy V. Militante and Brandon G. Sibbaluca, and published in the journal *International Journal of Scientific and Technology Research* in 2020. This research deals with the testing of the VGGNet algorithm, a generalized version of the VGG-16, for the detection of pneumonia. The research concluded with an accuracy rating result of 97% for the VGGNet [24].

The third research article is "Applicability of Various Pre-Trained Deep Convolutional Neural Networks for Pneumonia Classification based on X-Ray Images", authored by Julio Christian Young and Alethea Suryadibrata, and published in the journal *International Journal of Advanced Trends in Computer Science and Engineering* in 2020. This research also deals with the testing of various computer vision algorithms alongside the VGG-16 algorithm. The research concluded with a validation rating result of 80.33% for the VGG-16 algorithm [25].

The fourth research article is "Multitask Learning for Medical Image Classification using VGG Architecture", authored by Yuan Yang, Lin Zhang, Lei Ren, and Yuanjun Lali, and presented in the conference *Proceedings of the 9th International Workshop on Innovative Simulation for Healthcare (IWISH 2020)*. This research tested the VGG-16 algorithm both in pneumonia and ECG datasets. The research concluded with an accuracy rating result of 93% for the VGG-16 algorithm [26].

The fifth research article is "Comparative Analysis of Convolutional Neural Networks Applied in the Detection of Pneumonia Through X-Ray Images of Children", authored by Luan Oliveira Silva, Leandro dos Santos Araújo, Victor Ferreira Souza, Raimundo Matos Barros Neto, and Adam Santos, and published in the journal *Learning and Nonlinear Models* in 2020. This research deals with the testing of five algorithms for the detection of pneumonia, including the VGG-16 algorithm. The research concluded with a precision rating result of 82.80% for the detection of non-pneumonic images, and 91.94% for pneumonic images [27].

The sixth research article is "A Combined Approach Using Image Processing and Deep Learning to Detect Pneumonia from Chest X-Ray Image", authored by Md. Mehedi Hasan, Mir Md. Jahangir Kabir, Md. Rakibul Haque, and Mohiuddin Ahmed, and presented at the conference *2019 3rd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE)*. This research also deals with the testing of various computer vision algorithms alongside the VGG-16 algorithm. The research concluded with an accuracy rating result of 96.2% for the VGG-16 algorithm [28].

The seventh research article is "Artificial Intelligence Empowers Radiologists to Differentiate Pneumonia Induced by COVID-19 versus Influenza Viruses", authored by Houman Sotoudeh, Mohsen Tabatabaei, Baharak Tasorian, Kamran Tavakol, Ehsan Sotoudeh, and Abdol Moini, and published in the journal *Acta Informatica Medica* in 2020. This research deals with the testing of various computer vision algorithms for the detection of COVID-19 and H1N1 pneumonia. The research concluded with an accuracy rating result of 95.37% for the detection of COVID-19 pneumonia, and 83.67% for the detection of H1N1 pneumonia using the VGG-19 algorithm, which is a variant of the VGG-16 algorithm [29].

These seven research articles provided the proponents with detailed and diverse implementations and testing results of the VGG-16 algorithm, including the VGG-19 and VGGNet, for the detection of pneumonia on 2D x-ray images. The testing results also provided the proponents with the knowledge on how to perform the testing for this research.

E. Web-based Medical Diagnostic and Biomedical Informatics Technologies

For the topic of web-based medical diagnostic and biomedical informatics technologies, the proponents have selected three research articles. The first research article is "State of the Art of Semantic Web for Healthcare", authored by Xhemal Zenuni, Bujar Raufi, Florije Ismaili, and Jaumin Ajdari, and published in the journal Procedia - Social and Behavioral Sciences in 2015. This research deals with the application of semantic web in the creation of standardized web-based healthcare technologies with enhanced user accessibility. The research concluded that standardized web technologies will enhance the management and analysis of medical data [30]. This research article motivated the proponents to develop a free-to-use medical diagnostic web application to make medical analytic tools accessible to anyone.

The second research article is "Improving Healthcare with Interactive Visualization", authored by Ben Shneiderman, Catherine Plaisant, and Bradford W. Hesse, and published in the journal Computer in 2013. This research deals with the various techniques and tools on how to visualize medical data. The research concluded with the various pros and cons of various medical data visualization techniques and tools [31]. This research article provided the proponents with information on how to organize medical data for the web application.

The third research article is "Big Data Analytics in Healthcare", authored by M. Ambigavathi and D. Sridharan, and presented at the conference 2018 10th International Conference on Advanced Computing, IcoAC. This research deals with the various tools for healthcare data analytics. The research concluded with the pros and cons of various healthcare data analytic tools [32]. This research article provided the proponents with insight on how to manage the output of the web application.

F. Web Frameworks

For the topic of web frameworks, the proponents have selected four research articles. The first research article is "A Framework for Designing the Architectures of Deep Convolutional Neural Networks", authored by Saleh Albelwi and Ausif Mahmood, and published in the journal Entropy in 2017. This research deals with a web framework implementation for various CNN algorithms. The research concluded that no single web framework can efficiently implement all CNN algorithms, recommending software developers to extensively analyze the specifications and capabilities of their CNN algorithms for efficient implementation [33].

The second research article is "A Low-Effort Analytics Platform for Visualizing Evolving Flask-Based Python Web Services", authored by Patrick Vogel, Thijs Klooster, Vasilios Andrikopoulos, and Mircea Lungu, and presented at the conference 2017 IEEE Working Conference on Software Visualization (VISSOFT). This research deals with the issue of the lack of analytic tools for the evaluation of the performance of various Flask-based web applications. The research concluded with the development of the Flask Dashboard, an analytic tool for the evaluation of the performance of Flask-based web applications [34].

The third research article is "Analysis on Web Frameworks", authored by Dasari Hermitha Curie, Joyce Jaison, Jyoti Yadav, and J. Rex Fiona, and presented at the conference International Conference on Physics and Photonics Processes in Nano Sciences 2019. This research deals with the pros and cons of various web frameworks. The research concluded that web developers should be aware of these pros and cons for the efficient deployment of web applications [35].

The fourth research article is "Implementation of Database using Python Flask Framework", authored by Nidhi Chauhan, Mandeep Singh, Ayushi Verma, Aashwaath Parasher, and Gaurav Budhiraja, and published in the journal International Journal of Engineering and Computer Science in 2019. This research deals with the implementation of a database system for the Flask framework. The research concluded with the feasibility of such implementation [36].
These four research articles gave the proponents the information on the proper selection of the web framework for the efficient development and deployment of the web application.

G. High-Level Mathematics and Statistics

For the topic of high-level mathematics and statistics, the proponents have selected two textbooks and one research article. The first textbook is "Fundamentals of Functional Analysis", authored by S. S. Kutateladze, and published by the Sobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences in 1996 [37]. This textbook is essential for the understanding of the concept of convolution in functional analysis.

The second textbook is "Introduction to Statistics and Data Analysis", authored by R. Peck, C. Olsen, and J. Devore, and published by Thomson Higher Education in 2008 [38]. Information about the experimental design for testing, as well as the formulation of the statistical metrics used, came from this textbook.

For the research paper, the proponents have selected "A Matrix Theory Proof of Discrete Convolution Theorem" authored by B.R. Hunt and published in *IEEE Transactions on Audio and Electroacoustics* in 1971 [39]. The proponents' understanding of the process of convolution in matrices came from this research paper.

These two textbooks and research paper provided enough detail for the understanding of the mathematical foundations of the convolutional neural network, including the statistical tools needed for the experimental design.

III. METHODOLOGY

A. The Mathematics of Convolution – A Functional Analysis Perspective

The definition of convolution varies depending on the field of mathematics. In functional analysis, convolution is defined as the operation between two functions f and g that expresses a third function that indicates how the shape of the graph of one is modified by the other [40].

Definition: Convolution (Functional Analysis)

If f and g are discrete functions, then $f * g$ is the convolution of f and g defined as:

$$(f * g)(t) := \int_{-\infty}^{\infty} f(\tau)g(t-\tau)\,d\tau.$$

Fig. 3. Definition of Convolution according to Functional Analysis

B. The Mathematics of Convolution – A Matrix Theory Perspective

In matrix theory, convolution is defined as the recursive dot product of two matrices. Instead of generalized functions, the operation is done with matrices, A and B, where matrix B is of smaller size than matrix A. Matrix B is referred to as the kernel matrix. The kernel matrix contains the pattern that can be found in matrix A. After a recursive dot product of matrix A to B, the resultant matrix is obtained. If the resultant matrix is the same or similar to the kernel matrix, then it can be concluded that the pattern stored in the kernel matrix exists in the input matrix [41]. This concept is particularly relevant in the implementation of convolutional neural network algorithms.

Definition: Convolution (Matrix Theory)

If $\boldsymbol{A}$ and $\boldsymbol{B}$ are matrices, then $\boldsymbol{A} * \boldsymbol{B}$ is the convolution of $\boldsymbol{A}$ and $\boldsymbol{B}$ defined as:

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} * \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix}$$

$$= \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} x_{(m-i)(n-j)} y_{(1+i)(1+j)}$$

Fig. 4. Definition of Convolution according to Matrix Theory

C. Convolutional Neural Network in Computer Vision

A convolutional neural network, as stated in the introduction, is a type of deep neural network that relies on the operation of convolution for machine learning and analysis. CNN is used primarily on datasets that can be represented as matrices. This means CNN is highly utilized in the field of computer vision, where 2-dimensional images can be represented as matrices. Representation of images into matrices can be achieved using two techniques: either every pixel is assigned a numeric value, converting the entire image into a matrix, or a separate algorithm is used to partition the image into arbitrary sizes, with each grid being assigned a numeric value. Once the image has been represented into a matrix, it will be convoluted into a second, smaller matrix, known as the kernel. The kernel contains the pattern necessary for the analysis of the image. Since the operation of convolution yields another matrix of a smaller size, a pattern is detected if the resulting output matrix matches the kernel matrix, which indicates that the pattern exists within the original matrix. Through this technique, it has become possible to look for patterns within images, as long as such patterns are present within the kernel matrix. Since neural networks contain multiple hidden layers, various kernels containing various patterns can be used within each hidden layer, allowing for the detection and recognition of multiple patterns within the image.

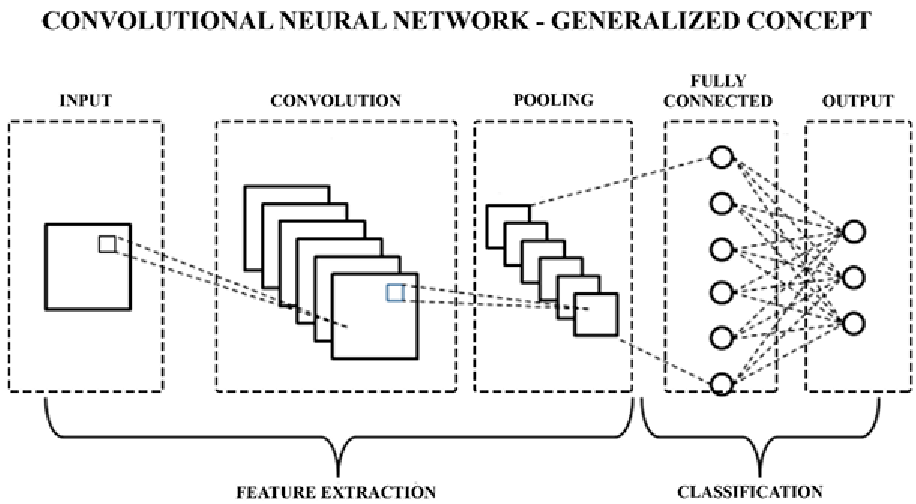

Fig. 5. General Concept of a Convolutional Neural Network Algorithm

The CNN algorithm that is used for this research is the VGG-16, an algorithm developed by Karen Simonyan and Andrew Zisserman from the University of Oxford. VGG stands for Visual Geometry Group, a research laboratory within the Department of Engineering Science of the University of Oxford, in which the two researchers are affiliated.

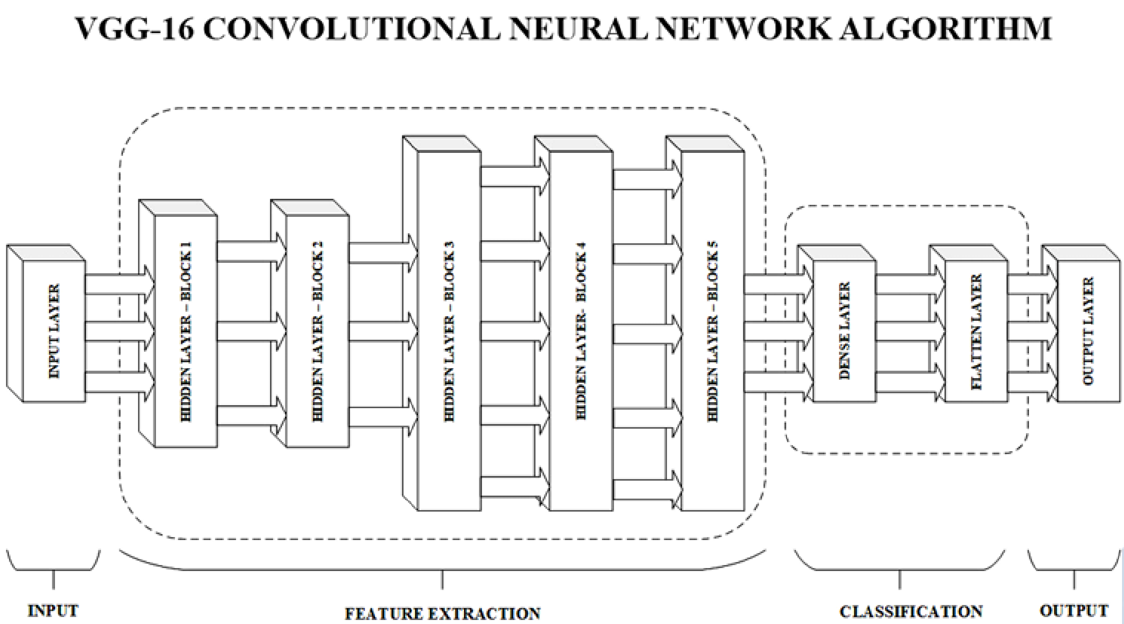

Fig. 6. Block Diagram of the VGG-16 Convolutional Network Algorithm

The VGG-16 algorithm is primarily composed of five blocks of hidden layers, a dense layer, and a flatten layer. Each block of the hidden layer is composed of Cov2D and MaxPooling2D stages. Cov2D is a computational function that can execute a convolutional operation on images. MaxPooling2D is a computational function that minimizes the size of the matrix by getting the highest value within a partition.

For hidden layers of block 1 and 2, the matrices will pass through 2 stages of Cov2D, and a single stage of MaxPooling2D. For block 3,4, and 5, the matrices will pass through three stages of Cov2D and a single stage of MaxPooling2D [42].

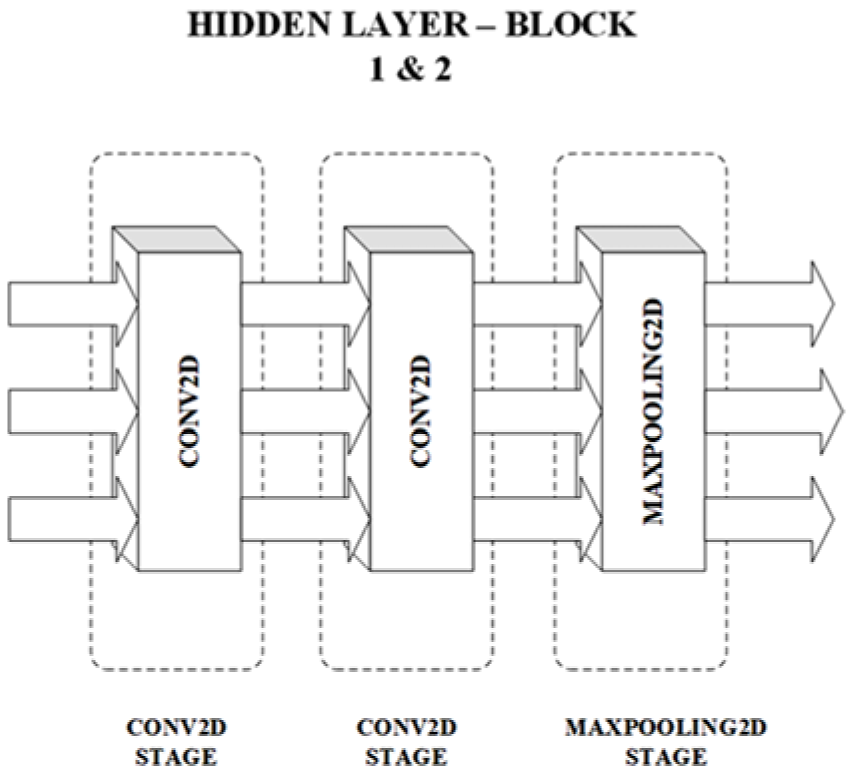

Fig. 7. Conv2D and MaxPooling2D stages for Hidden Layers 1 and 2

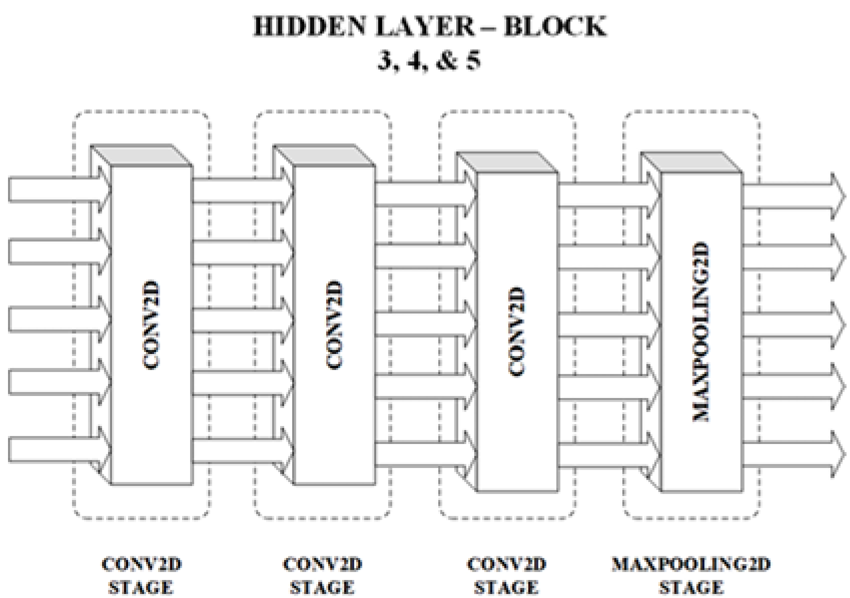

Fig. 8. Conv2D and MaxPooling2D stages for Hidden Layers 3,4, &5

D. Transfer Learning Technique

The kernel matrix can be set by the developer or can be formulated using machine learning. Using machine learning, a knowledge base is formed, containing the kernel matrices necessary for the detection and recognition of patterns present within the training dataset.

In this research, a machine learning model file with a .h5 file extension contains all the data relevant to the analysis of the image. Kernel matrices containing the pattern to be detected are stored in the .h5 file.

Transfer learning is an alternative technique in machine learning that combines supervised and unsupervised machine learning. Supervised machine learning is where the developers feed training data into the neural network during the machine learning phase of software development. Once training data is fed, the neural network detects patterns within the datasets. These patterns are then used to analyze the input data during operational deployment. Unsupervised learning is where the neural network analyzes patterns within datasets, stores those patterns, and uses them to analyze other datasets, all during operational deployment. Technically, a computational machine that uses unsupervised machine learning learns as it operates.

There are various advantages and disadvantages to these two machine learning modes. Supervised learning allows more control over the training dataset since it relies on the developers for the selection and organization of the training dataset, including the training of the neural network itself. However, it cannot accept new knowledge once it is deployed and in operational use. Unsupervised machine learning allows the machine to learn while in operation, however, there is little to no control over what data goes into the machine. Such uncontrolled flow of data may cause errors in the detection of patterns, which will lead to the corruption of the neural network.

Transfer learning attempts to combine the strengths of supervised and unsupervised learning. Training datasets are required during the machine learning phase, storing the patterns it has gathered in its knowledge base. Once deployed and operational, the machine also picks up patterns within data it encounters as it operates, much like in unsupervised learning. This allows for flexibility and further enhancement of its knowledge base. This also allows the machine to yield results with higher accuracy the more data it encounters [43, 44, 45].

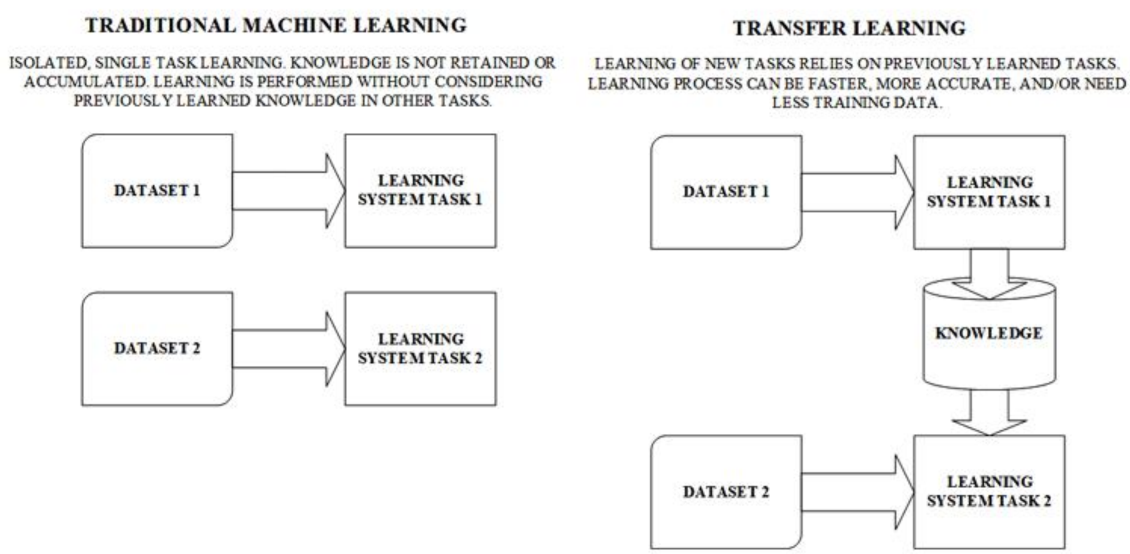

Fig. 9. Comparison of traditional machine learning and transfer learning

E. Training Datasets

Training datasets were obtained from Kaggle, an open-source research community owned by Google LLC, which specializes in data analytics and machine learning research. Kaggle's objective is to provide free and high-quality open-source datasets from various fields of study to further research in data analytics and machine learning.

The dataset used is entitled "Chest X-ray Images for Pneumonia Detection with Deep Learning" published by Tolga Dincer, a data scientist from Connecticut, United States. Before the selection of this dataset, the proponents first considered the use of the dataset entitled "Chest X-Ray Images (Pneumonia)", published by Paul Mooney. The final selection was based on the review of other researchers and the usability rating of the dataset. Since the dataset published by Paul Mooney garnered mixed reviews, with some researchers citing the low quality of the dataset, the proponents opted to use Tolga Dincer's dataset, which is more recent, published by a data scientist and has a usability rating of 8.1, which is higher compared to Paul Mooney's dataset with a usability rating of only 7.5 [46, 47].

Tolga Dincer's dataset is composed of two primary categories, pneumonic and normal. The datasets are also partitioned for training and testing phases. This categorization and classification minimize errors during the training and testing phases. The accompanying documentation of the dataset states the numerical imbalance within the pneumonic and normal categories, where the number of pneumonic x-ray images is almost twice as many compared to the number of normal x-ray images. To accommodate this imbalance, the machine learning program developed for this research utilized data augmentation. Data augmentation is a concept where imbalances within the training dataset are fixed by creating duplicates of the less frequent category. This duplication process includes image manipulations, such as horizontal inversion, zooming-in, and image partitioning. This process creates a more balanced training set, mitigating the numerical imbalance and statistical bias. Data augmentation is essential in yielding accurate and high-quality analysis [48, 49, 50, 51].

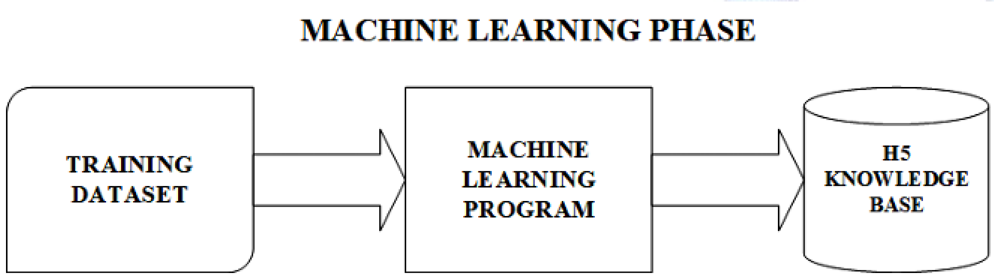

Fig. 10. Machine Learning Phase Diagram for Knowledge Base Generation

F. Implementation Using Flask

To implement the CNN algorithm and utilize the generated .h5 knowledge base file for the detection of pneumonia, a Python application was developed around a Flask framework for web deployment. A Python-compatible cloud server was used to host the web application, including the virtual machine and cloud storage [52, 53, 54].

The web application the proponents have developed follows the standard web application architecture. The front-end is composed of HTML & CSS-coded web pages. The back-end is created using the Python-coded Flask framework, and the storage is where the .h5 file and image storage directory are located [55, 56, 57, 58].

F. Implementation Using Flask

To implement the CNN algorithm and utilize the generated .h5 knowledge base file for the detection of pneumonia, a Python application was developed around a Flask framework for web deployment. A Python-compatible cloud server was used to host the web application, including the virtual machine and cloud storage [52, 53, 54].

The web application the proponents have developed follows the standard web application architecture. The front-end is composed of HTML & CSS-coded web pages. The back-end is created using the Python-coded Flask framework, and the storage is where the .h5 file and image storage directory are located [55, 56, 57, 58].

WEB APPLICATION SYSTEM ARCHITECTURE

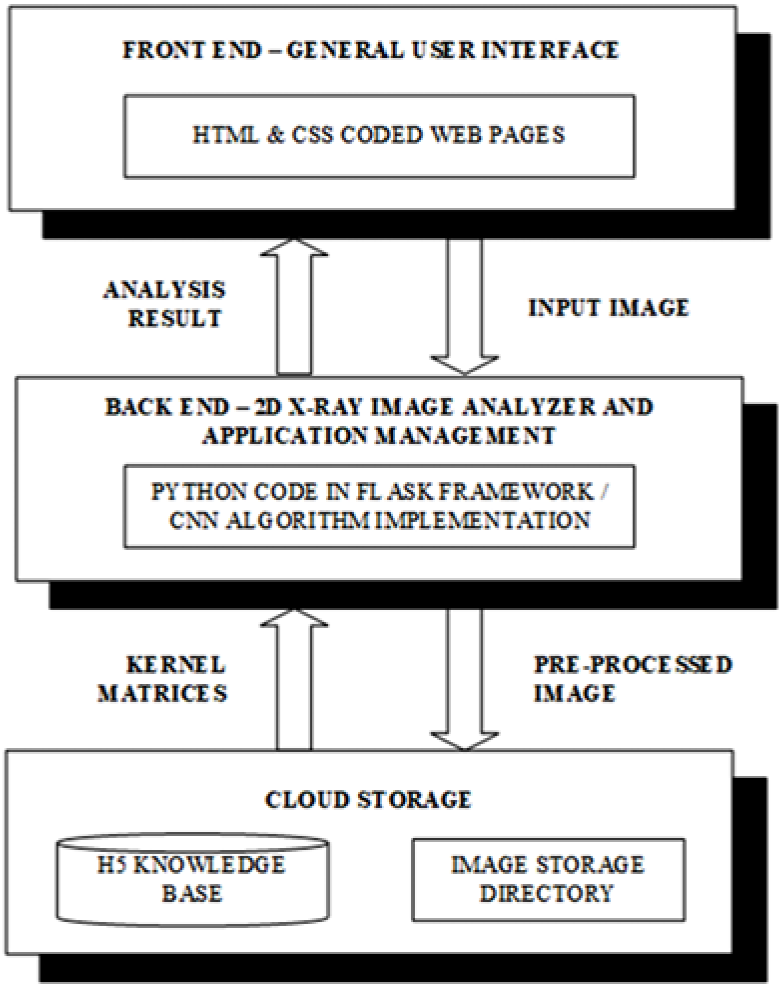

Fig. 11. Web Application System Architecture

The web application is composed of four web pages. The Analyzer page, the Machine Learning page, the Extended Intelligence page, and the Developers page.

The Analyzer page functions both as the home page and the page where the tabulated x-ray results are to be displayed. A short overview of what the web application is about is also displayed. The Machine Learning page contains the details about how machine learning is conducted for the web application, as well as the statistical results. The Extended Intelligence page details the relevance of extended intelligence in the field of biomedical informatics. The Developers page has the details of the development team, the research adviser, as well as the abstract of this research paper.

Fig. 12. Web Application User Interface

IV. RESULTS AND DISCUSSION

IV. RESULTS AND DISCUSSION

A. Quantitative Approach, Experimental Design and Statistical Methods

To test the precision and effectiveness of the developed web application, a confirmatory test was conducted using a quantitative approach and statistical methods.

The experimental design was a 5-trial confirmatory test. In this approach, a confirmatory test was conducted five times using five random and independently selected test datasets. For each trial, a sample of one-hundred 2-dimensional lung x-ray images was gathered from the open-source dataset. The selection process was conducted using a statistical approach known as Simple Random Sampling without Replacement (SRSWOR). In this approach, a sample was selected from the population randomly with no duplicates. This sample served as the test dataset. Fifty images that were labeled as pneumonic were selected using SRSWOR, while another fifty that were labeled as non-pneumonic were also selected using SRSWOR. This fifty pneumonic and non-pneumonic 2-D lung x-ray images composed the test dataset of size 100 per trial. The 1:1 ratio of pneumonic and non-pneumonic images ensured less statistical bias. These test datasets were then analyzed using the web application.

To verify the accuracy of the web application, the resulting preliminary diagnosis should match the label of each x-ray image. To do this, three statistical measures were formulated for this specific experimental design, the Diagnosis Precision Percentage per Trial (DPP-Ti), General Diagnostic Precision Percentage (GDPP), and the General Diagnostic Error Percentage (GDEP). The Diagnostic Precision Percentage per Trial is defined as the arithmetic mean of the matched diagnosis in a specific trial, multiplied by 100%. To compute this, the number of matched diagnosis was tallied, and the number is divided by the sample size, which is 100. This arithmetic mean is then multiplied to 100%. The General Diagnostic Precision Percentage is defined as the arithmetic mean of all five Diagnostic Precision Percentage per Trial. This is computed by summing all five DPP-Ti, and dividing it by 5. The General Diagnosis Error Percentage is defined as the error rating and is computed by subtracting the GDPP from 100 % [59].

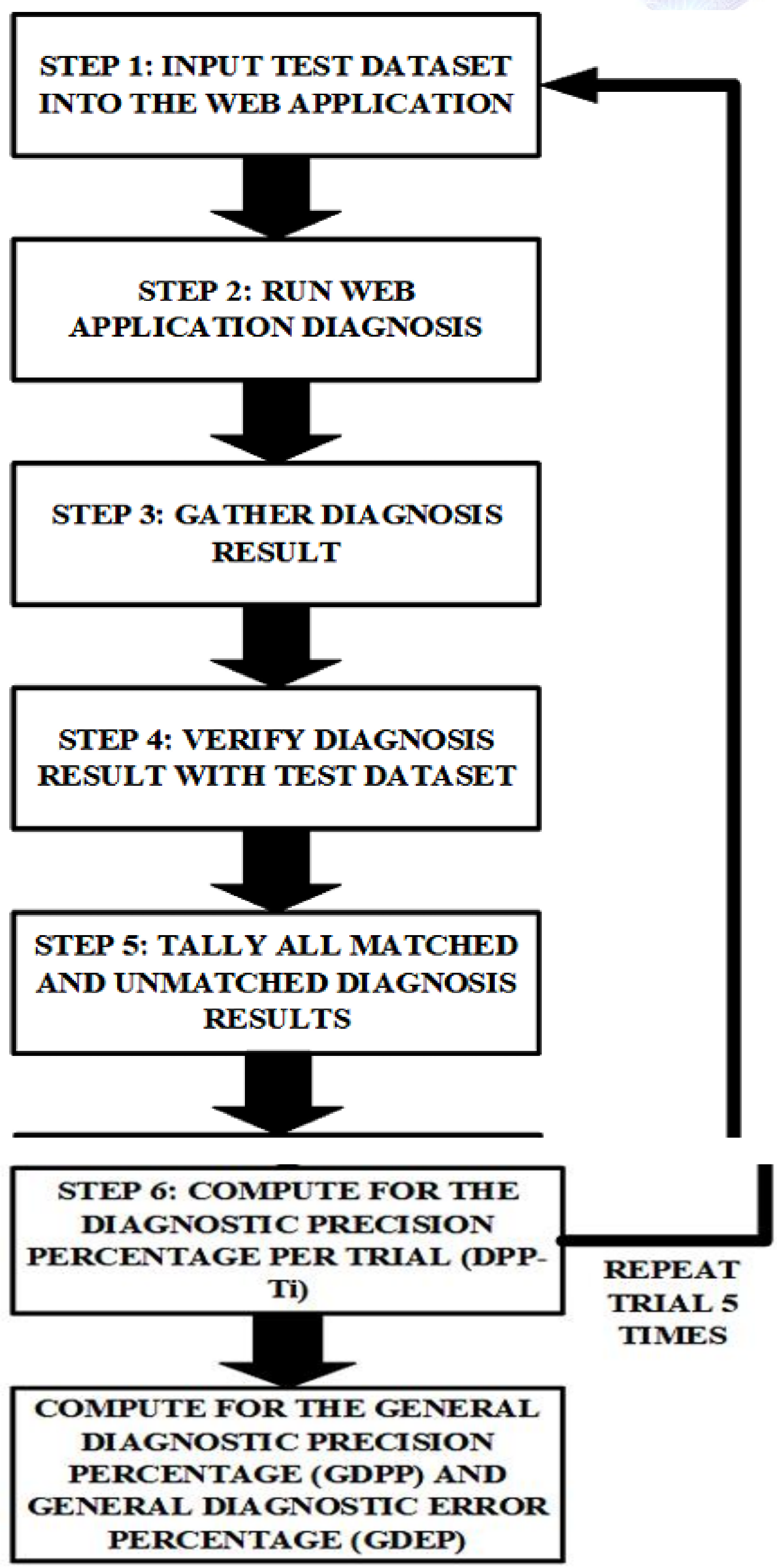

Fig. 13. Experimental Design for Validation of Precision of Web Application

Table 1.
Statistical Formulas for Precision Validation

Statistical Formulas for Precision Validation
Arithmetic Mean ($\bar{x}$) with sample size n: $$\bar{x} = \frac{\sum_{i=1}^{n} xi}{n}$$
Diagnostic Precision Percentage per Trial (DPP-Ti): $$\text{DPP-Ti} = \frac{\text{TOTAL NUMBER OF MATCHED DIAGNOSIS}}{100} \times 100\%$$
General Diagnostic Precision Percentage (GDPP): $$\text{GDPP} = \frac{\sum_{i=1}^{5} \text{DPP}-\text{Ti}}{5}$$
General Diagnostic Error Percentage (GDEP): $$\text{GDEP} = 100\% - \text{GDPP}$$

B. Experimental Result

Here is the tabulated result of the 5-Trial Confirmatory Test, with the calculated Diagnostic Precision Percentage per Trial, General Diagnostic Precision Percentage, and General Diagnostic Error Percentage.

Table 2.
Result of the 5-Trial Confirmatory Test (Tabulated)

Result of the 5-Trial Confirmatory Test	
Trial 1:	
Matched Diagnosis	Tally Result: 92
Unmatched Diagnosis	Tally Result: 8
Diagnosis Precision Percentage of Trial 1:	92 %
Trial 2:	
Matched Diagnosis	Tally Result: 93
Unmatched Diagnosis	Tally Result: 7
Diagnosis Precision Percentage of Trial 2:	93 %
Trial 3:	
Matched Diagnosis	Tally Result: 90
Unmatched Diagnosis	Tally Result: 10
Diagnosis Precision Percentage of Trial 3:	90 %
Trial 4:	
Matched Diagnosis	Tally Result: 91
Unmatched Diagnosis	Tally Result: 9
Diagnosis Precision Percentage of Trial 4:	91 %
Trial 5:	
Matched Diagnosis	Tally Result: 90
Unmatched Diagnosis	Tally Result: 10
Diagnosis Precision Percentage of Trial 5:	90 %
Final Result	
General Diagnosis Precision Percentage:	91.2 %
General Diagnosis Error Percentage:	8.8 %

Here is the result of the 5-Trial Confirmatory Test in Confusion Matrix Format. The confusion matrix shows the relationship between the label and the corresponding diagnosis result of the web application on each test images.

Table 3.
Result of the 5-Trial Confirmatory Test (Confusion Matrix Format)

Result of the 5-Trial Confirmatory Test in Confusion Matrix Format		
	Labelled As "Pneumonic"	Labelled As "Not Pneumonic"
Diagnosed as "Pneumonic"	*TRUE POSITIVE* Tally 1:50 Tally 2: 50 Tally 3: 50 Tally 4: 50 Tally 5: 50	*FALSE POSITIVE* Tally 1: 8 Tally 2: 7 Tally 3: 10 Tally 4: 9 Tally 5: 10
Diagnosed as "Not Pneumonic"	*FALSE NEGATIVE* Tally 1: 0 Tally 2: 0 Tally 3: 0 Tally 4: 0 Tally 5: 0	*TRUE NEGATIVE* Tally 1: 42 Tally 2: 43 Tally 3: 40 Tally 4: 41 Tally 5: 40

C. Statistical Interpretation

A General Diagnostic Precision Percentage of 91.2% and a General Diagnosis Error Percentage of 8.8 % were computed from the 5-Trial Confirmatory Test. This means that the developed web application is precise with its diagnosis 91.2% of the time, and can commit errors in its diagnosis for 8.8% of the time. This result is consistent with the referenced research papers concerning the application and implementation of the VGG-16 algorithm for the detection of pneumonia on x-ray images, where precision ratings range from 90 to 98%.

The confusion matrix illustrates that the web application commits no errors in the diagnosis of pneumonia on pneumonia-labeled images, but commits errors when diagnosing non-pneumonic images. The error is that certain non-pneumonic images were diagnosed as pneumonia, which is a false positive error. There are no false-negative errors found. This means for 8.8 % of the time, errors in the diagnosis will be false positive.

V. CONCLUSIONS AND RECOMMENDATIONS

Since the GDPP has a value of 91.2%, the experimental result concluded that the developed web application is precise for 91.2% of the time. A GDEP value of 8.8% tells that an error in the diagnosis can happen 8.8% of the time. The confusion matrix of the experimental result illustrates that only false positive errors are possible with no false-negative errors found.

This error is reasonable since false positive errors (non-pneumonic but diagnosed as pneumonic) have less serious consequences than false-negative errors (pneumonic but diagnosed as non-pneumonic).

The error can be attributed to the inherent limitation of the VGG-16 algorithm, as well as the quality of the training and test datasets. Since transfer learning relies heavily on the training datasets, the quality of the analysis is linked directly to the quality of the dataset. High-quality datasets in large quantities are essential in yielding high-precision diagnoses. Small errors within the test dataset can manifest in the generated knowledge base and analysis result. Various iterations of the CNN algorithm will yield various precision ratings.

The successful generation of the .h5 knowledge base proved that the CNN algorithm implementation for machine learning can generate patterns based on analyzing open-source datasets. The successful implementation and deployment of the web application to the cloud server proved that such a system can be feasibly deployed in such a platform. The experimental data proved the high precision of the analysis, proving that it can be used in the diagnosis of pneumonia under the supervision of medical practitioners.

ACKNOWLEDGEMENT

The proponents would like to give their sincerest gratitude to the researchers in the field of artificial intelligence, extended intelligence, biomedical informatics, computer vision, and convolutional neural networks for their valuable contributions in their specific fields. Their contribution to the creation of new knowledge not only led to the advancement of their fields of study but the advancement of humanity in general.

Lastly, the proponents would like to give their sincerest gratitude to their thesis and research adviser, Prof. Manuel Luis C. Santos. His guidance, criticisms, suggestions, and motivations became instrumental for the success of this research.

REFERENCE

[1][14] Pavel Hamet and Johanne Tremblay. 2017. Artificial Intelligence in Medicine. *Metabolism: Clinical and Experimental* 69 (Jan. 2017), 36-40. DOI: https://doi.org/10.1016/j.metabol.2017.01.011

[2][46] Tolga Dincer. 2020.Chest X-Ray Images for Pneumonia Detection with Deep Learning, ver. 1.0. (Aug. 2020). Retrieved Feb. 17, 2021 from https://www.kaggle.com/tolgadincer/labeled-chest-xray-images

[3][20] Asharul Islam Khan and Salim Al-Habsi. 2019. Machine Learning in Computer Vision.In *International Conference on Computational Intelligence and Data Science (ICCIDS 2019),* September 6 – 7, 2019, Gurugram, India. Elsevier Procedia, Amsterdam, Netherlands, 1444–1451. DOI: https://doi.org/10.1016/j.procs.2020.03.355

[4][21] Kenji Suzuki. 2017. Overview of Deep Learning in Medical Imaging. *Radiological Physics and Technology* 10, 3 (Sept. 2017), 257-273. DOI: https://doi.org/10.1007/s12194-017-0406-5

[5][9][16] Egbuna Obidike, Tagbo Oguonu, Joy Eze, Ikenna Ndu, Adaeze Ayuk, Mildred Ukoha, Obinna Nduagubam, and Ikechukwu Ogbonna. 2018. Misdiagnosis of Pneumonia, Bronchiolitis and Reactive Airway Disease in Children: A Retrospective Case Review Series in South East, Nigeria. *Current Pediatric Research* 22, 2 (June 2018), 166-171.

[6][10][17] Ang Choon Seong, Kelvin Beh Khai Meng, Yeang Li Jing, Chin Yuen Quan, Khor Inn Shih, Yoon Chee Kin, and Irfhan Ali Bin Hyder Ali. 2020. Misdiagnosis of Community-Acquired Pneumonia in Patients Admitted to Respiratory Wards, Penang General Hospital. *Medical Journal of Malaysia* 75, 4 (July 2020), 385-390.

[7][12][18] M. Cellina, M. Orsi, T. Toluian, C. Valenti Pittino, and G. Oliva. 2020. False Negative Chest X-Rays in Patients Affected by COVID-19 Pneumonia and Corresponding Chest CT Findings. *Radiography* 26 (2020), 189-194. DOI: https://doi.org/10.1016/j.radi.2020.04.017

[8][11][19] Zulfiqar A. Bhutta. 2006. Childhood Pneumonia in Developing Countries. *BMJ* 333, 7569 (Sept. 2006), 612–613. DOI: https://doi.org/10.1136/bmj.38975.602836.be

[13][15] Joichi Ito. 2016. Extended Intelligence. *PubPub* (Feb. 2016), 7 pages. DOI: https://doi.org/10.21428/10224

[22][42] Karen Simonyan and Andrew Zisserman. 2015. Very Deep Convolutional Networks for Large-Scale Image Recognition. In *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings,* May 7 - 9, 2015, San Diego, California. arXiv, Ithaca, NY, 14 pages.

[23][43] K. Reddy Madhavi, G. Madhavi, B. Rupa Devi, and Padmavathi Kora. 2020. Detection of Pneumonia Using Deep Transfer Learning. *International Journal of Advanced Trends in Computer Science and Engineering* 9, 5 (Sept.-Oct. 2020), 8934–8937. DOI: https://doi.org/10.30534/ijatcse/2020/292952020

[24][48] Sammy V. Militante and Brandon G. Sibbaluca. 2020. Pneumonia Detection Using Convolutional Neural Networks. *International Journal of Scientific and Technology Research* 9, 4 (April 2020), 1332-1337.

[25][49] Julio Christian Young and Alethea Suryadibrata. 2020. Applicability of Various Pre-Trained Deep Convolutional Neural Networks for Pneumonia Classification based on X-Ray Images. *International Journal of Advanced Trends in Computer Science and Engineering* 9, 3 (May-June 2020), 2649–2654. DOI: https://doi.org/10.30534/ijatcse/2020/24932020

[26][44] Yuan Yang, Lin Zhang, Lei Ren, and Yuanjun Lali. 2020. Multitask Learning for Medical Image Classification using VGG Architecture. In *Proceedings of the 9th International Workshop on Innovative Simulation for Healthcare (IWISH 2020),* September 16 - 18, 2020, Online. CAL-TEK SRL, Cosenza, Italy. 20-26. https://doi.org/10.46354/i3m.2020.iwish.004

[27][50] Luan Oliveira Silva, Leandro dos Santos Araújo, Victor Ferreira Souza, Raimundo Matos Barros Neto, and Adam Santos. 2020. Comparative Analysis of Convolutional Neural Networks Applied in the Detection of Pneumonia Through X-Ray Images of Children. *Learning and Nonlinear Models* 18, 2 (2020), 4–15. DOI: https://doi.org/10.21528/lnlm-vol18-no2-art1

[28][45] Md. Mehedi Hasan, Mir Md. Jahangir Kabir, Md. Rakibul Haque, and Mohiuddin Ahmed. 2019. A Combined Approach Using Image Processing and Deep Learning to Detect Pneumonia from Chest X-Ray Image. In *2019 3rd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE),* December 26 - 28, 2019, Rajshahi, Bangladesh. IEEE, Piscataway, NJ, 71 pages. https://doi.org/10.1109/icecte48615.2019.9303543

[29][51] Houman Sotoudeh, Mohsen Tabatabaei, Baharak Tasorian, Kamran Tavakol, Ehsan Sotoudeh, and Abdol Moini. 2020. Artificial Intelligence Empowers Radiologists to Differentiate Pneumonia Induced by COVID-19 versus Influenza Viruses. *Acta Informatica Medica* 28, 3 (Sept. 2020), 190-195. DOI: https://doi.org/10.5455/aim.2020.28.190-195

[30][52] Xhemal Zenuni, Bujar Raufi, Florije Ismaili, and Jaumin Ajdari. 2015. State of the Art of Semantic Web for Healthcare. *Procedia - Social and Behavioral Sciences* 195 (2015), 1990–1998. DOI: https://doi.org/10.1016/j.sbspro.2015.06.213

[31][53] Ben Shneiderman, Catherine Plaisant, and Bradford W. Hesse. 2013. Improving Healthcare with Interactive Visualization. *Computer* 46, 5 (May 2013), 26–34. DOI: https://doi.org/10.1109/mc.2013.38

[32][54] M. Ambigavathi and D. Sridharan. 2018. Big Data Analytics in Healthcare.In *2018 Tenth International Conference on Advanced Computing (ICoAC),* December 13 - 15, 2018, Chennai, India. IEEE, Piscataway, NJ, 269-276. DOI: https://doi.org/10.1109/icoac44903.2018.8939061

[33][55] Saleh Albelwi and Ausif Mahmood. 2017. A Framework for Designing the Architectures of Deep Convolutional Neural Networks. *Entropy* 19, 6 (June 2017), 21 pages. DOI: https://doi.org/10.3390/e19060242

[34][56] Patrick Vogel, Thijs Klooster, Vasilios Andrikopoulos, and Mircea Lungu. 2017. A Low-Effort Analytics Platform for Visualizing Evolving Flask-Based Python Web Services. In *2017 IEEE Working Conference on Software Visualization (VISSOFT),* September 18 – 19, 2017, Shanghai, China. IEEE, Piscataway, NJ, 5 pages. https://doi.org/10.1109/vissoft.2017.13

[35][57] Dasari Hermitha Curie, Joyce Jaison, Jyoti Yadav, and J. Rex Fiona. 2019. Analysis on Web Frameworks. In *International Conference on Physics and Photonics Processes in Nano Sciences 2019,* June 20 – 22, 2019, Eluru, India. IOP Publishing, Bristol, United Kingdom, 7 pages. https://doi.org/10.1088/1742-6596/1362/1/012114

[36][58] Nidhi Chauhan, Mandeep Singh, Ayushi Verma, Aashwaath Parasher, and Gaurav Budhiraja. 2019. Implementation of Database using Python Flask Framework. *International Journal of Engineering and Computer Science* 8, 12 (Dec. 2019), 24894–24899. DOI: https://doi.org/10.18535/ijecs/v8i12.4390

[37][40] S. S. Kutateladze. 1996.Kluwer Texts in the Mathematical Sciences, Vol. 12: Fundamentals of Functional Analysis. Sobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia.

[38][59] Roxy Peck, Chris Olsen, and Jay Devore. 2008. *Introduction to Statistics and Data Analysis* (3rd ed.). Thomson Higher Education, Belmont, California.

[39][41] B.R. Hunt. 1971. A Matrix Theory Proof of Discrete Convolution Theorem. *IEEE Transactions on Audio and Electroacoustics* 19, 4 (Dec. 1971), 285-288.

[47] Paul Mooney. 2018.Chest X-Ray Images (Pneumonia), ver. 2.0. (March 2018). Retrieved February 10, 2021 from https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia/metadata