DEEP LEARNING FOR MEDICAL IMAGE ANALYSIS

Robi Dany Riupassa

01 – Overview of Medical Image

02 – AI (ANN and DL)

03 – Medical Image Analysis using Deep Learning Algorithms

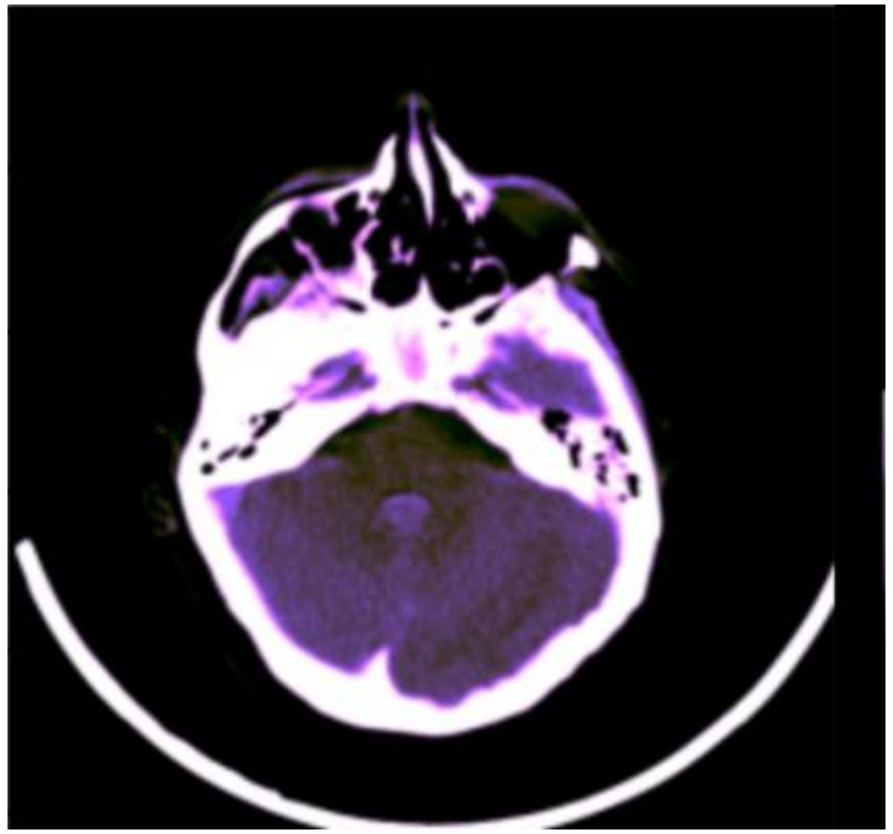
03 – Challenges and Limitations

04 – Future Prospects

MEDICAL IMAGE

In the health care system, there has been a dramatic increase in demand for medical image services, e.g. Radiography, endoscopy, Computed Tomography (CT), Mammography Images (MG), Ultrasound images, Magnetic Resonance Imaging (MRI), Magnetic Resonance Angiography (MRA), Nuclear medicine imaging, Positron Emission Tomography (PET) and pathological tests. Besides, medical images can often be challenging to analyze and time-consuming process due to the shortage of radiologists.

Computed Tomography (CT) of the Brain

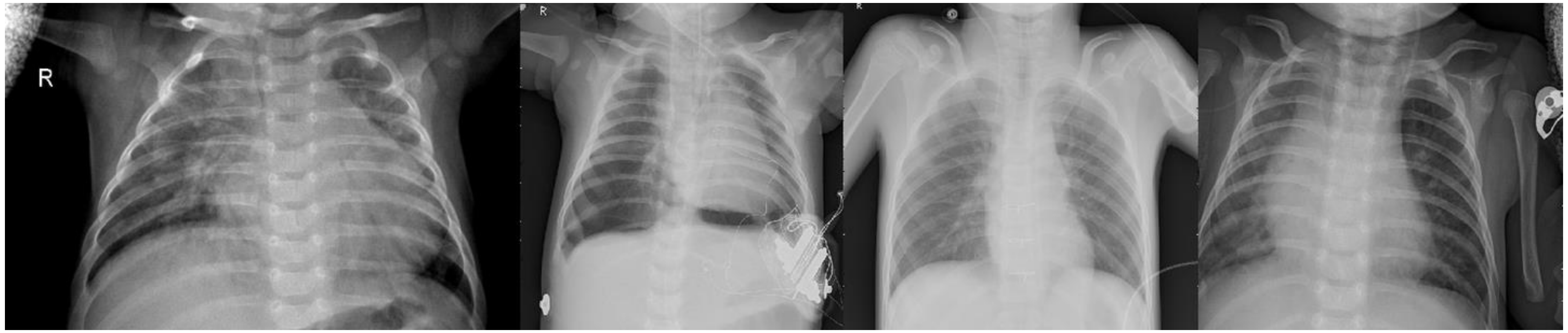


Medical image analysis based on deep learning approach

Muralikrishna Puttagunta[1] • S. Ravi[1]

MEDICAL IMAGE

In the health care system, there has been a dramatic increase in demand for medical image services, e.g. Radiography, endoscopy, Computed Tomography (CT), Mammography Images (MG), Ultrasound images, Magnetic Resonance Imaging (MRI), Magnetic Resonance Angiography (MRA), Nuclear medicine imaging, Positron Emission Tomography (PET) and pathological tests. Besides, medical images can often be challenging to analyze and time-consuming process due to the shortage of radiologists.



Pediatric Pneumonia Chest X-ray

Medical image analysis based on deep learning approach

Muralikrishna Puttagunta[1] · S. Ravi[1]

AI (ANN and DL)

Artificial Intelligence (AI) can address these problems. Machine Learning (ML) is an application of AI that can be able to function without being specifically programmed, that learn from data and make predictions or decisions based on past data. ML uses three learning approaches, namely, supervised learning, unsupervised learning, and semi-supervised learning. The ML techniques include the extraction of features and the selection of suitable features for a specific problem requires a domain expert. Deep learning (DL) techniques solve the problem of feature selection. DL is one part of ML, and DL can automatically extract essential features from raw input data

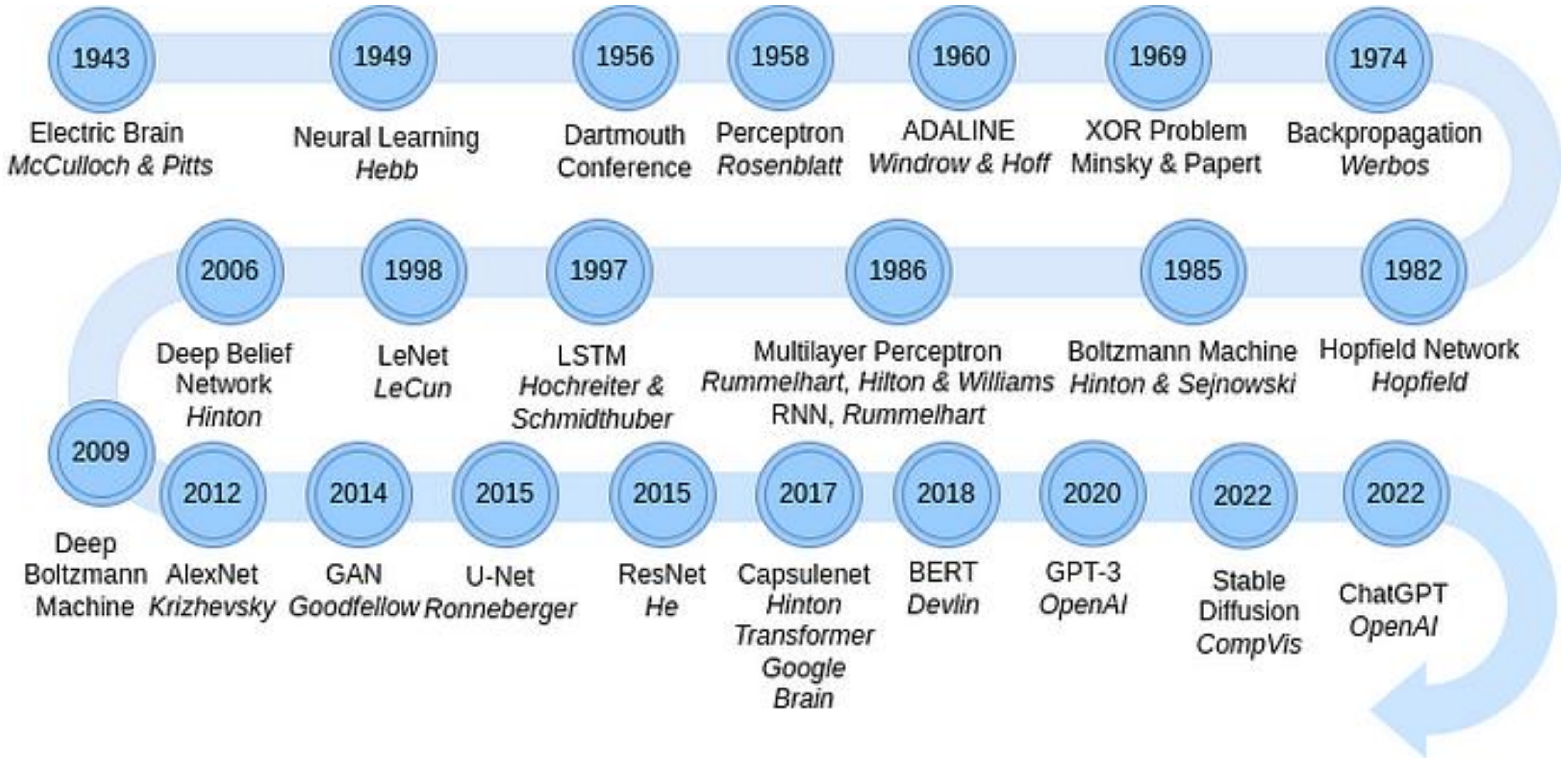
Medical image analysis based on deep learning approach

Muralikrishna Puttagunta[1] • S. Ravi[1]

Artificial Neural Networks (ANN)

Artificial Neural Networks (ANN) form the basis for most of the DLA. ANN is a computational model structure that has some performance characteristics similar to biological neural networks. ANN comprises simple processing units called neurons or nodes that are interconnected by weighted links. A biological neuron can be described mathematically in Eq. (1). Figure 3 shows the simplest artificial neural model known as the perceptron.

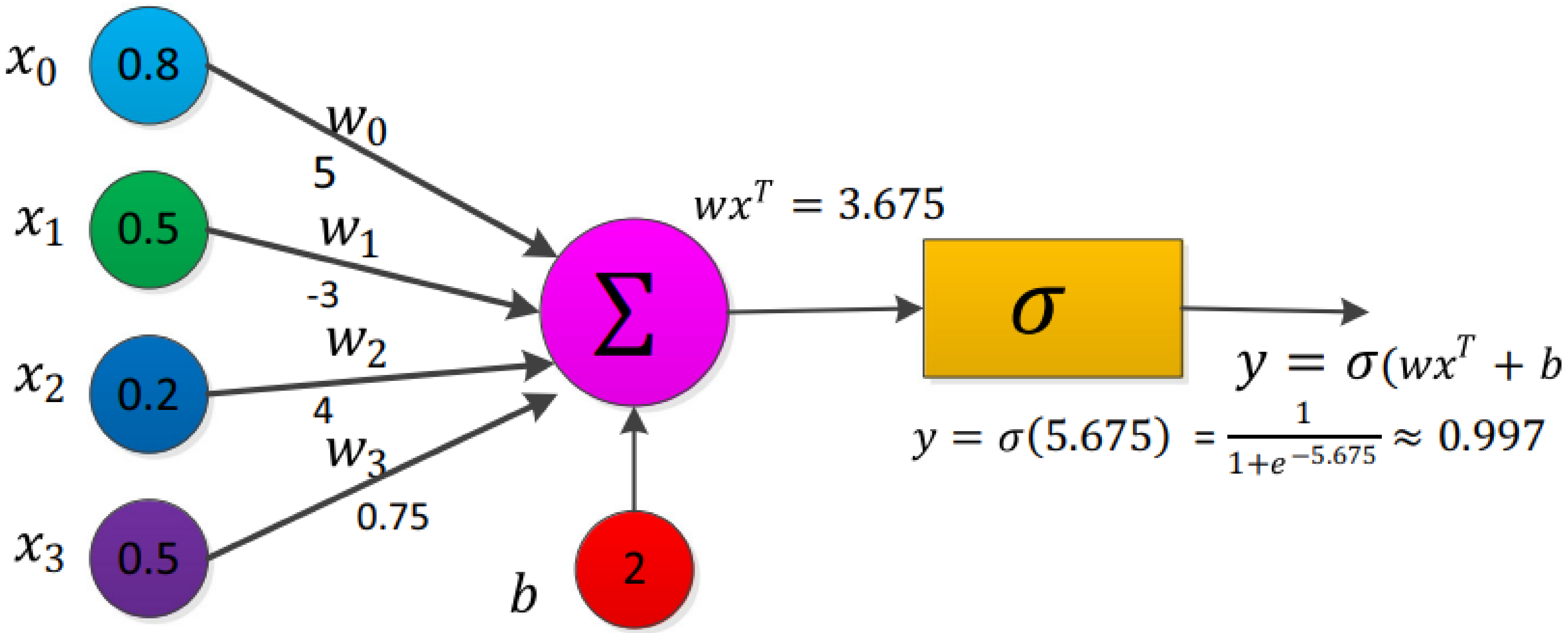
$$y = f\left(wx^{T} + b\right) \tag{1}$$



[Unraveling the Tapestry: The Evolution of Neural Networks | by Rushikesh Kahat | Medium]

Artificial Neural Networks (ANN)

In the neural networks, the learning process is modeled as an iterative process of optimization of the weights to minimize a loss function. Based on network performance, the weights are modified on a set of examples belonging to the training set. The necessary steps of the training procedure contain forward and backward phases. For Neural Network training, any of the activation functions in forwarding propagation is selected and BP training is used for changing weights.



Medical image analysis based on deep learning approach

Muralikrishna Puttagunta[1] • S. Ravi[1]

Artificial Neural Networks (ANN)

Table 1 Activation functions

Function name	Function equation	Function derivate
Sigmoid [86]	$f(x) = \frac{1}{1+e^{-x}}$	$f'(x) = f(x)(1 - f(x))$
Hyperbolic tangent [87]	$f(x) = \tanh(x) = \frac{2}{1+e^{-2x}} - 1$	$f'(x) = 1 - f(x)^2$
Soft sign activation	$f(x) = \frac{x}{1+\lvert x\rvert}$	$f'(x) = \frac{1}{(1+\lvert x\rvert)^2}$
Rectified Linear Unit [68, 104] (ReLU)	$f(x) = \begin{cases} 0 & x < 0 \\ x & x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$
Leaky Rectified Linear Unit [94] (leaky ReLU)	$f(x) = \begin{cases} \alpha x & x < 0 \\ x & x \geq 0 \end{cases}$	$f'(x) = \begin{cases} \alpha & x < 0 \\ 1 & x \geq 0 \end{cases}$
Parameterized Rectified Linear Unit(PReLU) [47]	PReLU is the same as leaky ReLU. The difference is $\propto$ can be learned from training data via backpropagation	
Randomized Leaky Rectified Linear Unit [180]	$f(x) = \begin{cases} \alpha x & x < 0 \\ x & x \geq 0 \end{cases}$	$f'(x) = \begin{cases} \alpha & x < 0 \\ 1 & x \geq 0 \end{cases}$
Soft plus [32]	$f(x) = \ln(1 + e^x)$	$f'(x) = \frac{1}{1+e^{-x}}$
Exponential Linear Unit (ELU) [24, 137]	$f(x) = \begin{cases} \propto(e^x - 1) & x < 0 \\ x & x \geq 0 \end{cases}$	$f'(x) = \begin{cases} f(x) + \propto & x < 0 \\ 1 & x \geq 0 \end{cases}$
Scaled exponential Linear Unit (SELU) [67]	$f(x) = \lambda \begin{cases} \propto(e^x - 1) & x < 0 \\ x & x \geq 0 \end{cases}$	$f'(x) = \begin{cases} f(x) + \lambda\propto & x < 0 \\ \lambda & x \geq 0 \end{cases}$

Medical image analysis based on deep learning approach

Muralikrishna Puttagunta[1] · S. Ravi[1]

Deep learning

A. RESTRICTED BOLTZMANN MACHINE

Restricted Boltzmann machine (RBM) is a kind of probabilistic graphical models that can be interpreted as stochastic neural networks [28]. A typical two-layer RBM includes a visible layer that contains the input we know and a hidden layer that contains the latent variables, as described in Fig. 2(a). RBMs are organized as a bipartite graph, where each visible neuron is connected to all hidden neurons and vice versa, but any two units are not connected in the same layer. RBMs have seen successful applications in many fields, such as collaborative filtering [29] and network anomaly detection [30]. Multiple stacked RBM layers can form a deep belief network (DBN), which consists of a visible layer and multiple hidden layers. The training of a DBN follows a layer-by-layer method, where each layer is treated as an RBM trained on top of the previously trained layer [31]. Many applications can benefit from the structure of DBNs, such as fault detection classification in industrial environments, threat identification in security alert systems, and emotional feature extraction out of images [17].

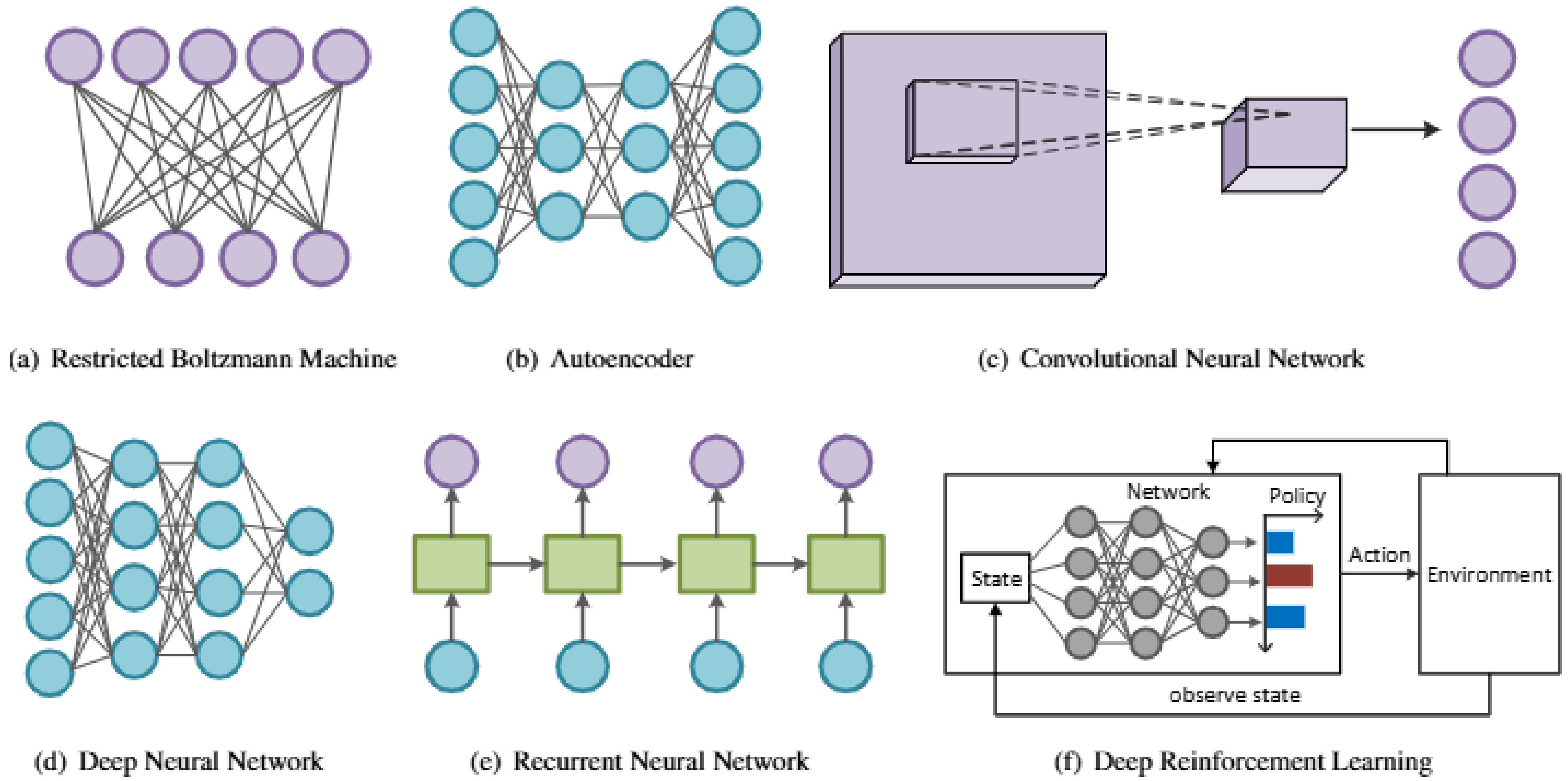


FIGURE 2: The structures of different deep learning models.

Deep Learning for Edge Computing Applications: A State-of-the-art Survey

FANGXIN WANG[1] (Student Member, IEEE), MIAO ZHANG[1], XIANGXIANG WANG[1], XIAOQIANG MA[2], AND JIANGCHUAN LIU[1] (Fellow, IEEE)
[1]School of Computing Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada
[2]School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China
Corresponding Author: Xiaoqiang Ma (mxqhust@gmail.com).

Deep learning

B. AUTOENCODER

An autoencoder includes an input layer and an output layer that are connected by one or multiple hidden layers [32], as illustrated in Fig. 2(b). The shape of the input layer and the output layer are the same. The AE can be divided into two parts, i.e., an encoder and a decoder. The encoder learns the representative characteristics of the input and transforms it into other latent features (usually in a compressing way). And the decoder receives the latent features of the encoder and aims to reconstruct the original form of the input data, minimizing the reconstruction error. Similarly, an AE can be formed as a deep architecture by stacking multiple layers into the hidden layer. There are several variants and extensions of AEs, such as sparse AE [33], denoising AE [34], and variational AE [35].

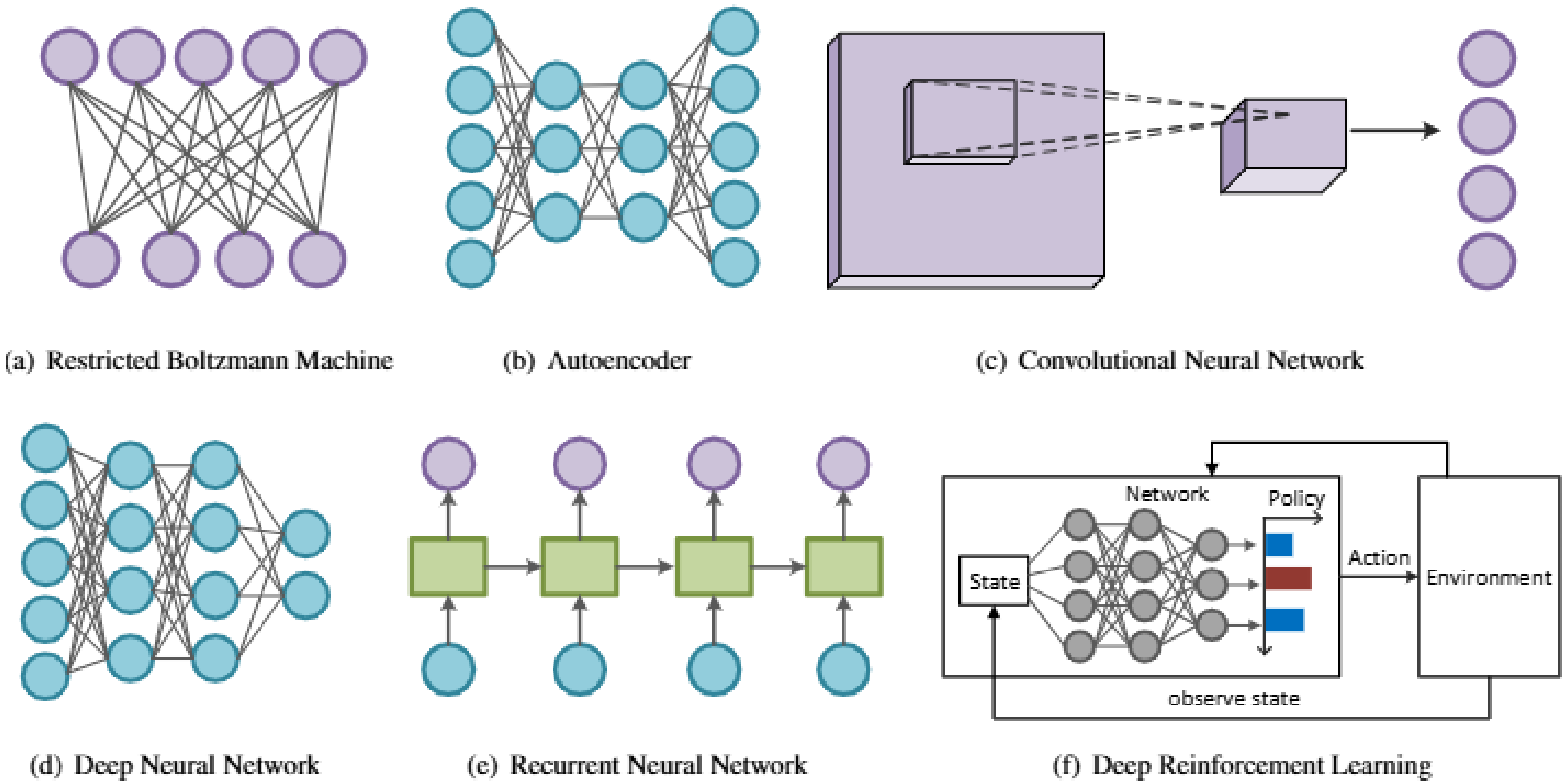


FIGURE 2: The structures of different deep learning models.

Deep Learning for Edge Computing Applications: A State-of-the-art Survey

FANGXIN WANG[1] (Student Member, IEEE), MIAO ZHANG[1], XIANGXIANG WANG[1], XIAOQIANG MA[2], AND JIANGCHUAN LIU[1] (Fellow, IEEE)
[1]School of Computing Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada
[2]School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China
Corresponding Author: Xiaoqiang Ma (mxqhust@gmail.com).

Deep learning

C. DEEP NEURAL NETWORKS

Compared to the traditional artificial neural network (ANN) that has shallow structure, deep neural network (DNN) (or deep fully connected neural network) usually has a deeper layer structure for more complicated learning tasks [32]. A DNN consists of an input layer, several hidden layers, and an output layer, where the output of each layer is fed to the next layer with activation functions. At the last layer, the final output representing the model prediction is produced. Optimization algorithms such as Stochastic Gradient Decent (SGD) [36] and backpropagation [37] are mostly used in the training process. DNNs are widely used in feature extraction, classification and function approximation.

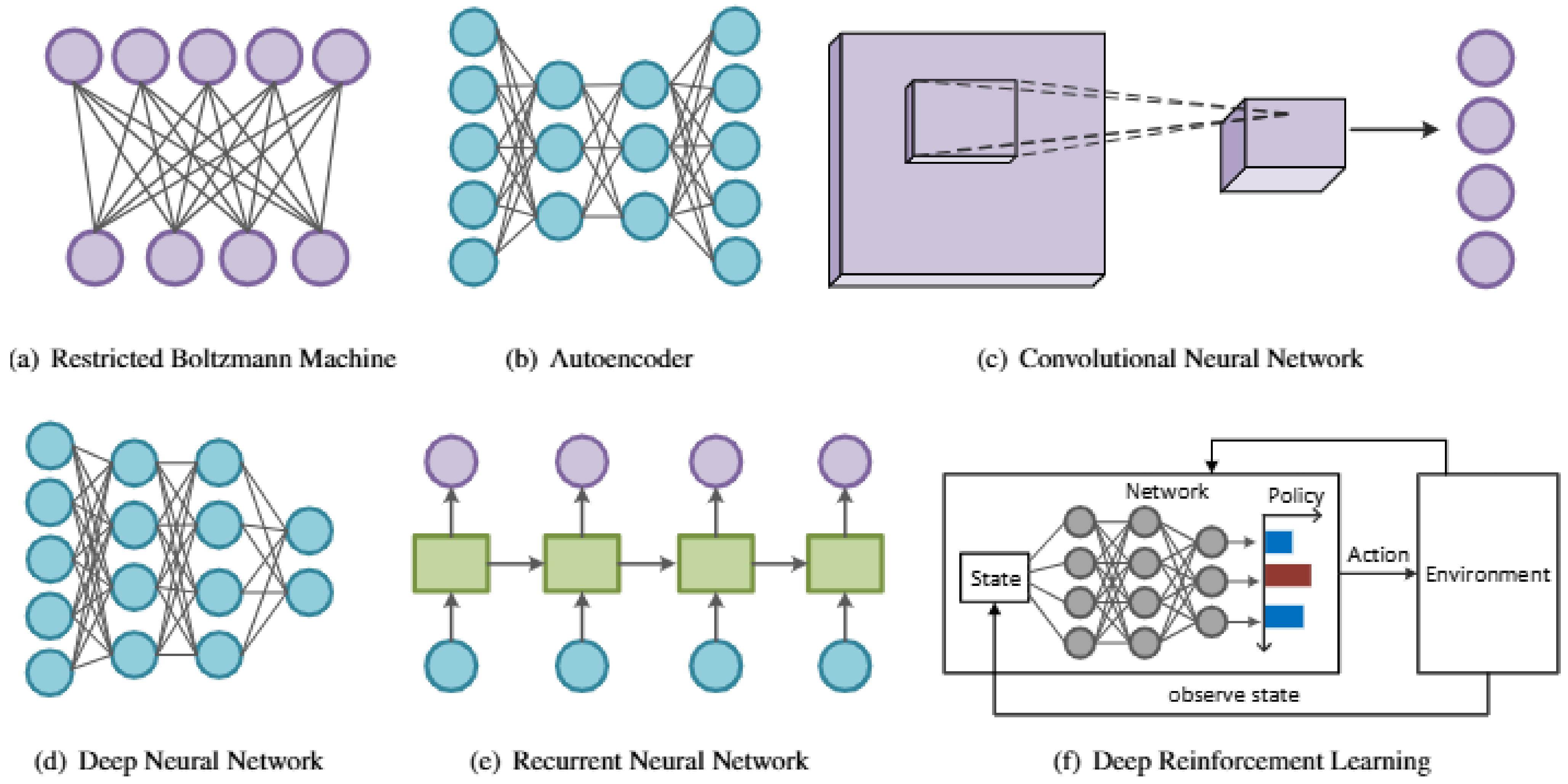


FIGURE 2: The structures of different deep learning models.

Deep Learning for Edge Computing Applications: A State-of-the-art Survey

FANGXIN WANG[1] (Student Member, IEEE), MIAO ZHANG[1], XIANGXIANG WANG[1], XIAOQIANG MA[2], AND JIANGCHUAN LIU[1] (Fellow, IEEE)
[1]School of Computing Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada
[2]School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China
Corresponding Author: Xiaoqiang Ma (mxqhust@gmail.com).

Deep learning

D. CONVOLUTIONAL NEURAL NETWORKS

Convolutional neural networks (CNNs) are designed to process data that comes in the form of multiple arrays, for example, a color image composed of three 2D arrays containing pixel intensities in the three color channels [27]. A CNN receives 2D data structures and extracts high-level features through convolutional layers as described in Fig. 2(c), which is the core of CNN architecture. By going through the 2D data with a set of moving filters and the pooling functions, CNN extracts the spatial correlations between adjacent data by calculating the inner product of the input and the filter. After that, a pooling block is operated over the output to reduce the spatial dimensions and generate a high-level abstraction. Compared to traditional fully connected deep networks, CNN can effectively decrease the parameter numbers of network and extract the spatial correlations in the raw data, mitigating the risk of overfitting [38]. The above advantages make CNN achieve significant results in many applications, such as object detection [39] and health monitoring [40].

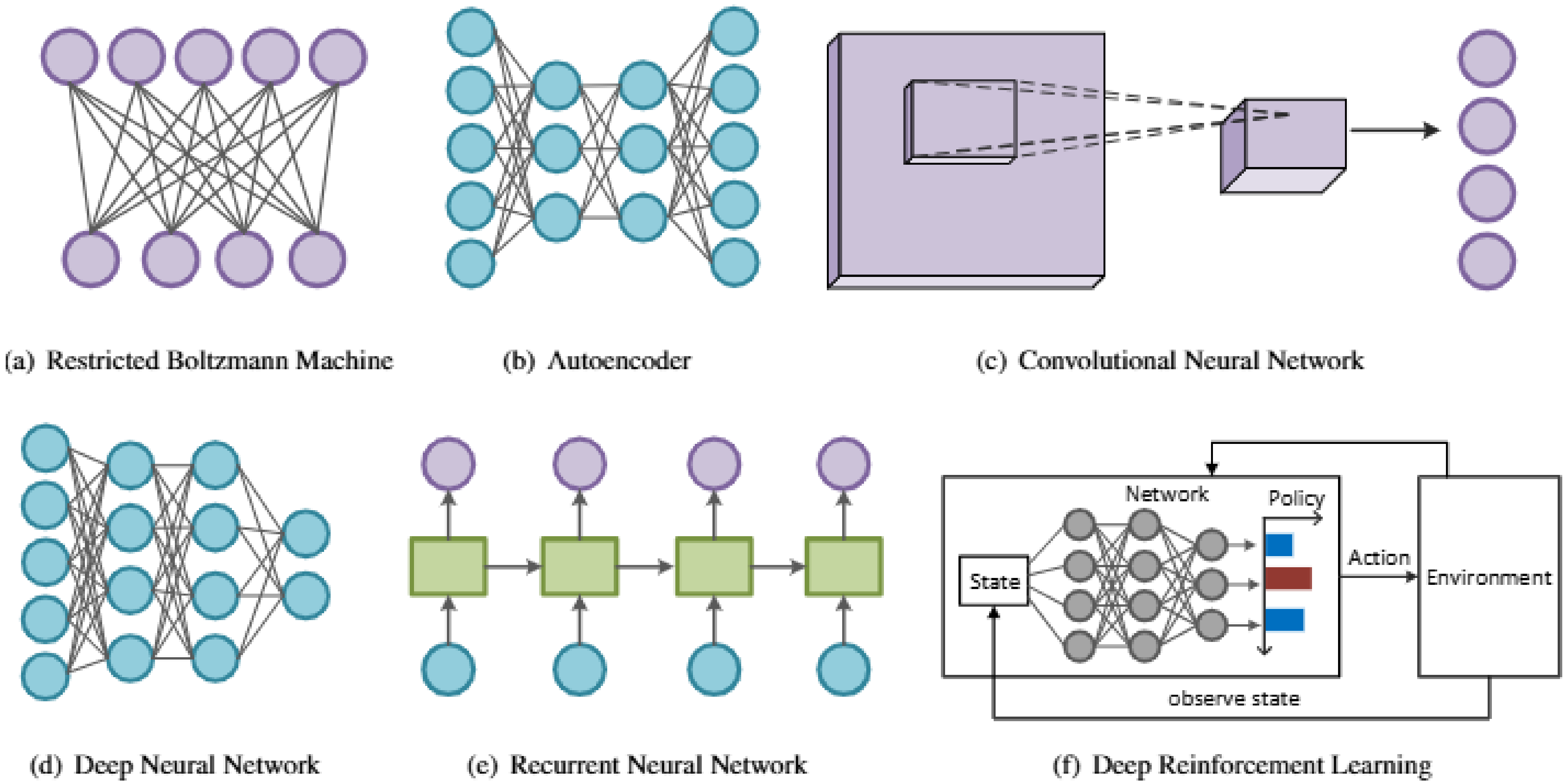


FIGURE 2: The structures of different deep learning models.

Deep Learning for Edge Computing Applications: A State-of-the-art Survey

FANGXIN WANG[1] (Student Member, IEEE), MIAO ZHANG[1], XIANGXIANG WANG[1], XIAOQIANG MA[2], AND JIANGCHUAN LIU[1] (Fellow, IEEE)
[1]School of Computing Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada
[2]School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China
Corresponding Author: Xiaoqiang Ma (mxqhust@gmail.com).

Deep learning

E. RECURRENT NEURAL NETWORKS

Different from CNNs that are good at abstracting spatial features, recurrent neural networks (RNNs) are designed for processing sequential or time-series data. The input to an RNN includes both the current sample and the previously observed samples. Specifically, each neuron of an RNN layer not only receives the output of its previous layer but also receives the stored state of from previous time steps, as depicted in Fig. 2(e). With this special architecture, RNN is able to remember previous information for integrated processing with the current information. However, RNNs can only look back for a few steps due to the gradient explosion and long-term dependencies. To solve this problem, Long Short-Term Memory (LSTM) network [41] is proposed to control the flow of information. In LSTM model, the forget gate is utilized to control the cell state and decide what to keep in the memory. Through the learning process, stored computations in the memory cells are not distorted over time, which particularly achieves better performance when data is characterized in long dependency [42]. RNN and LSTM are widely used in various sequential scenarios, such as language processing [43] and activity recognition [44].

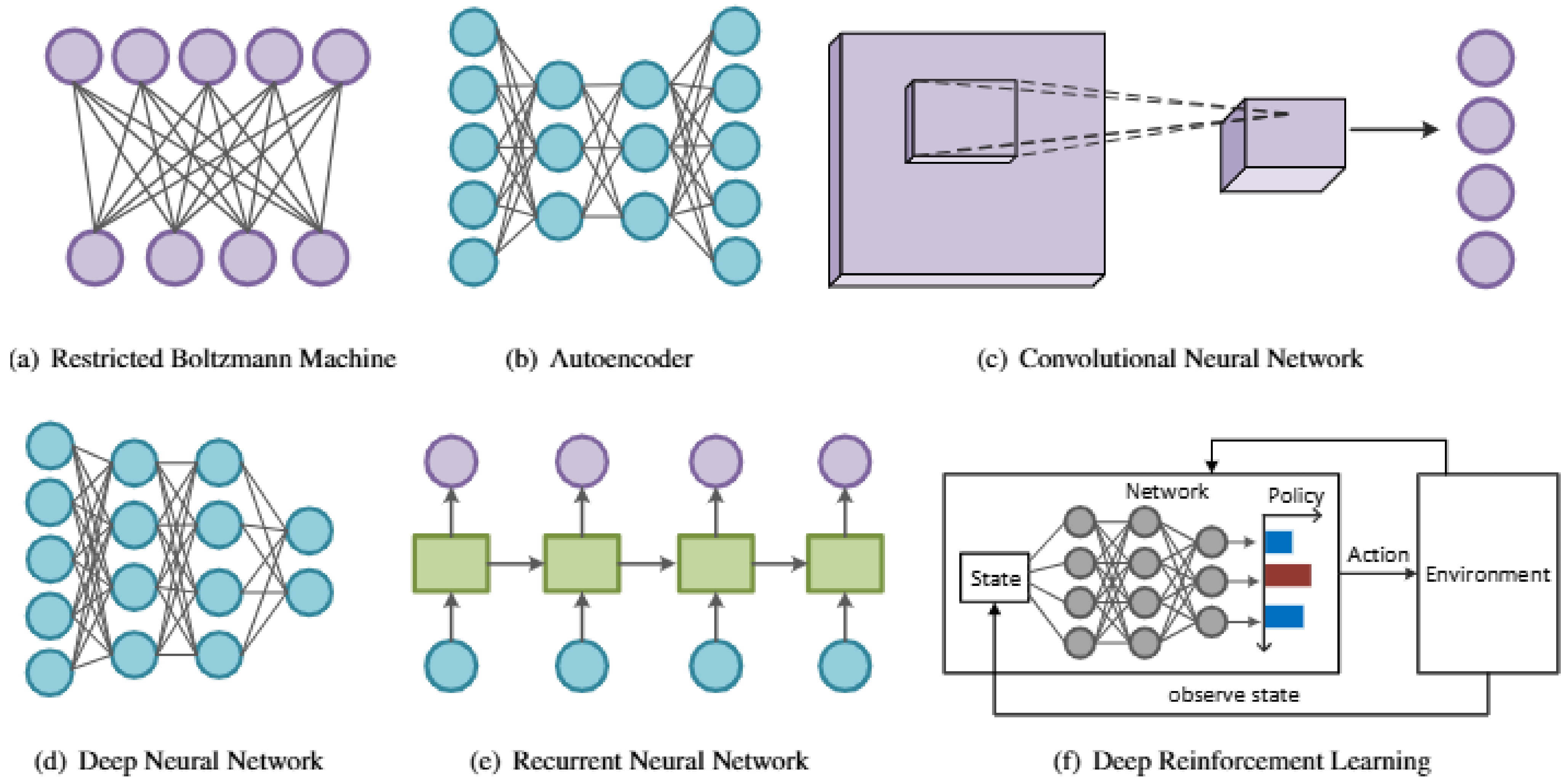


FIGURE 2: The structures of different deep learning models.

Deep Learning for Edge Computing Applications: A State-of-the-art Survey

FANGXIN WANG[1] (Student Member, IEEE), MIAO ZHANG[1], XIANGXIANG WANG[1], XIAOQIANG MA[2], AND JIANGCHUAN LIU[1] (Fellow, IEEE)
[1]School of Computing Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada
[2]School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China
Corresponding Author: Xiaoqiang Ma (mxqhust@gmail.com).

Deep learning

F. DEEP REINFORCEMENT LEARNING

Deep reinforcement learning (DRL) [7] is a combination of deep learning (DL) and reinforcement learning (RL) [45]. It aims to build an agent that is able to learn the best action choices over a set of states through the interaction with the environment, so as to maximize the long-term accumulated rewards. Different from traditional RL, DRL utilizes a deep neural network to represent the policy given its strong representation ability to approximate the value function or the direct strategy. DRL can be categorized into value-based models, such as Deep Q-Learning (DQL), Double DQL [46] and Duel DQL [47], and policy-gradient-based models, such as deep deterministic policy gradient (DDPG) [48] and asynchronous advantage actor-critic (A3C) [49]. The DRL has been successfully applied in many fields, such as computer gaming [7], chess gaming [8] and rate adaptation [50].

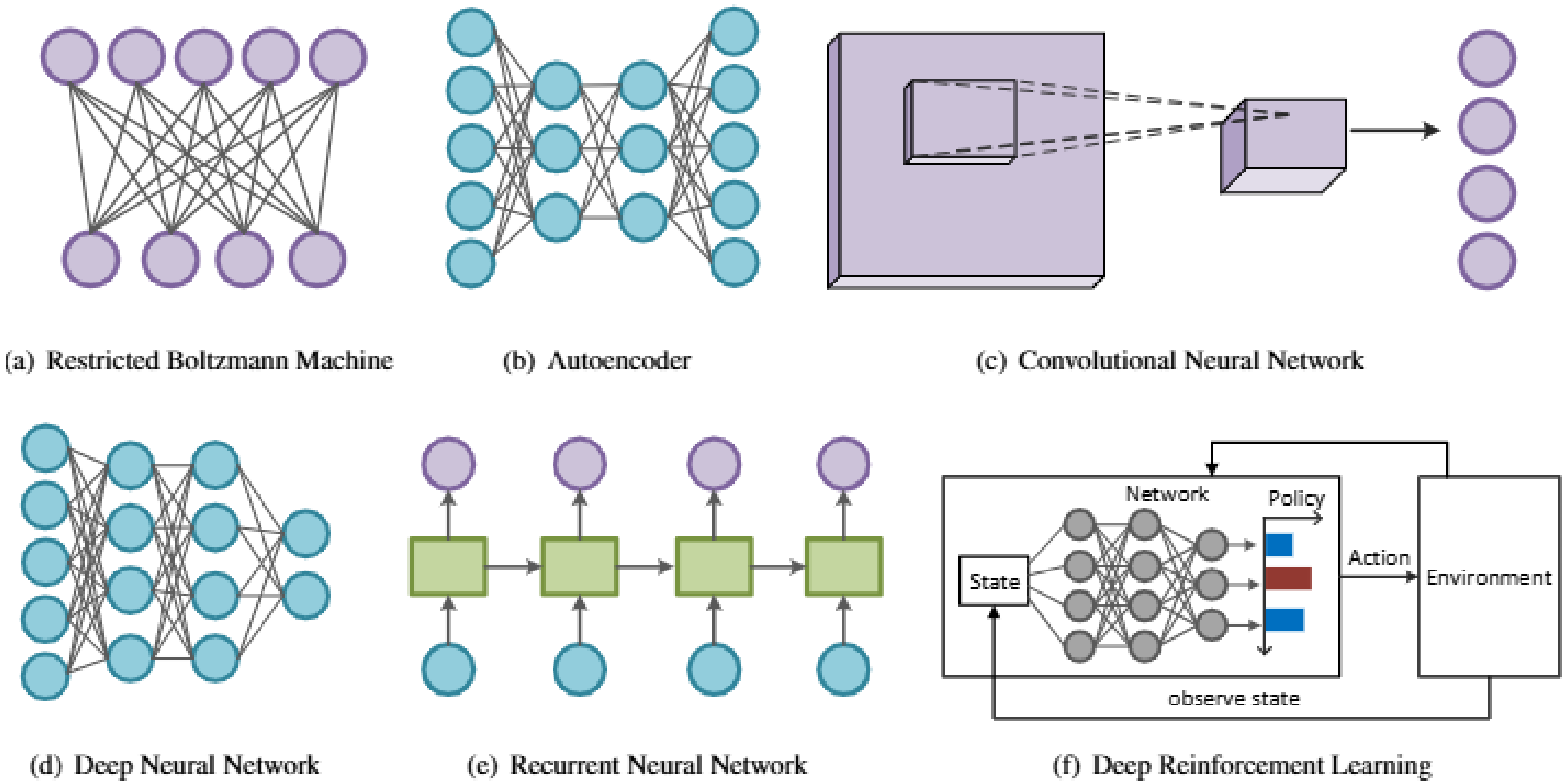


FIGURE 2: The structures of different deep learning models.

Deep Learning for Edge Computing Applications: A State-of-the-art Survey

FANGXIN WANG[1] (Student Member, IEEE), MIAO ZHANG[1], XIANGXIANG WANG[1], XIAOQIANG MA[2], AND JIANGCHUAN LIU[1] (Fellow, IEEE)
[1]School of Computing Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada
[2]School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China
Corresponding Author: Xiaoqiang Ma (mxqhust@gmail.com).

Use of deep learning in medical imaging

Table 3 An overview of the DLA for the study of X-ray images

Reference	Dataset	Method	Application	Metrics
Lo et al.,1995 [89]	–	CNN	Two-layer CNN, each with 12 5 × five filters for lung nodule detection.	ROC
S.Hwang et al. 2016 [57]	KIT, MC, and Shenzhen	Deep CNN	The first deep CNN-based Tuberculosis screening system with transfer learning technique	AUC
Rajpurkar et al. 2017 [122]	ChestX-ray14	CNN	Detects Pneumonia using CheXNet is a 121-layer CNN from a chest X-ray image.	F1 score
Lopes & Valiati 2017 [91]	Shenzhen and Montgomery	CNN	Comparative analysis of Pre-trained CNN as feature ex-tractors for tuberculosis detec-tion	Accuracy, ROC
Mittal et al. 2018 [99]	JSRT	LF-SegNet	Segmentation of lung field from CXR images using Fully convolutional encoder-decoder network	Accuracy
E.J.Hwang et al. 2019 [58]	57,481 CXR images	CNN	Deep learning-based automatic detection (DLAD) algorithm for tuberculosis detection on CXR	ROC
Souza et al. 2019 [148]	Montgomery	CNN	Segmentation of lungs in CXR for detection and diagnosis of pulmonary diseases using two CNN architecture	Dice coefficient
Hooda et al. [53]	Shenzhen, Montgomery, Belarus, JSRT	CNN	An ensemble of three pre-trained architectures ResNet, AlexNet, and GoogleNet for TB detection	Accuracy, ROC
Xu et al. 2019 [181]	chest X-ray14	CNN, CXNet-m1	Design a hierarchical CNN structure for a new network CXNet-m1 to detect anomaly of chest X-ray images	Accuracy, F1-score, and AUC
Murphy et al. 2019 [103]	5565 CXR images		Deep learning-based CAD4TB software evaluation	ROC

Use of deep learning in medical imaging

Table 3 (continued)

Reference	Dataset	Method	Application	Metrics
Rajaraman and Antani 2020 [119]	RSNA, Pediatric pneumonia, and Indiana,	CNN	An ensemble of modality-specific deep learning models for Tuberculosis (TB) detection from CXR	Accuracy, AUC, CI
Capizzi et al. 2020 [15]	Open data set from radiologykey.com	PNN	The fuzzy system, combined with a neural network, can detect low-contrast nodules.	Accuracy
Abbas et al. 2020 [2]	196 X-ray images	CNN	Classification of COVID-19 CXR images using Decompose, Transfer, and Compose (DeTraC)	Accuracy, SN, SP
Basu et al. 2020 [7]	225 COVID-19 CXR images	CNN	DETL (Domain Extension Transfer Learning) method for the screening of COVID-19 from CXR images	Accuracy
Wang & Wong 2020 [165]	13,975 X-ray images	CNN	A deep convolutional neural network COVID-Net design for the detection of COVID-19 cases	Accuracy, SN, PPV.
Ozturk et al. 2020 [110]	127 X-ray images	CNN	Deep learning-based DarkCovid net model to detect and classify COVID-19 cases from X-ray images	Accuracy.
Loey et al. 2020 [90]	306 X-ray images	AlexNet google Resnet18	A GAN with deep transfer learning for COVID-19 detection in limited CXR images.	Accuracy,
Apostolopoulos & Mpesiana 2020 [3]	1427 X-ray images	CNN	Transfer Learning-based CNN architectures to the detection of the Covid-19.	Accuracy, SN, SP

Challenges of the DL Applications in Medical Image Analysis

The lack of high-quality annotated data is one of the greatest problems with deep learning (DL) algorithms used for medical image analysis. For DL models to perform well and generalize, they need a lot of labeled data. But getting high-quality annotations for medical photos is challenging for a number of reasons: restricted accessibility: Because it is expensive and time-consuming to capture and annotate medical pictures, the amount of data from annotated images is constrained (76). Additionally, the process of annotating calls for medical professionals with particular training and understanding, who are not always available. Due to changes in patient anatomy, imaging modality, and disease pathology, medical pictures are complicated and extremely varied. Annotating medical images requires a high degree of accuracy and consistency, which can be challenging for complex and heterogeneous medical conditions. Privacy and ethical issues: The annotation process has the potential to make medical photographs containing sensitive patient data vulnerable to abuse or unauthorized access. Medical image analysis has a significant difficulty in protecting patient privacy and confidentiality while preserving the caliber of annotated data. Annotating medical pictures requires making subjective assessments, which might result in bias and variability in the annotations. These variables may have an impact on the effectiveness and generalizability of DL models, especially when the annotations are inconsistent among datasets or annotators (77). To address the challenge of limited availability of high-quality annotated data, several approaches have been proposed, including:

- Transfer learning: To enhance the performance of DL models on smaller datasets, transfer learning uses pre-trained models that have been learned on big datasets. By using this method, the volume of annotated data needed to train DL models may be decreased, and the generalizability of the models can be increased.
- Data augmentation: By applying modifications to already-existing, annotated data, data augmentation includes creating synthetic data. The diversity and quantity of annotated data available for DL model training may be increased using this method, and it can also raise the models' resistance to fluctuations in medical pictures.
- Active learning: Active learning involves selecting the most informative and uncertain samples for annotation, rather than annotating all the data. This approach can reduce the annotation workload and improve the efficiency of DL model training.
- Collaborative annotation: Collaborative annotation involves engaging medical experts, patients, and other stakeholders in the annotation process to ensure the accuracy, consistency, and relevance of annotations to clinical needs and values.

Medical image analysis using deep learning algorithms

Mengfang Li[1†], Yuanyuan Jiang[2†], Yanzhou Zhang[2]* and Haisheng Zhu[3]

Challenges of the DL Applications in Medical Image Analysis

Overall, addressing the challenge of limited availability of high-quality annotated data in medical image analysis requires a combination of technical, ethical, and social solutions that can improve the quality, quantity, and diversity of annotated data while ensuring patient privacy and ethical standards.

Deep learning algorithms for medical image analysis have a significant problem in terms of data quality. The model's performance may be considerably impacted by the caliber of the data utilized to train the deep learning algorithms (78). Obtaining medical pictures may be difficult, and their quality can vary based on a number of variables, such as the image capture equipment used, the image resolution, noise, artifacts, and the imaging technique. Furthermore, the annotations or labels used for training can also impact the quality of the data. Annotations may not always be accurate, and they may suffer from inter-and intra-observer variability, which can lead to biased models or models with poor generalization performance. To overcome the challenge of data quality, researchers need to establish robust quality control procedures for both image acquisition and annotation. Additionally, they need to develop algorithms that can handle noisy or low-quality data and improve the accuracy of annotations. Finally, they need to develop methods to evaluate the quality of the data used to train the deep learning models (79).

Medical image analysis using deep learning algorithms

Mengfang Li[1†], Yuanyuan Jiang[2†], Yanzhou Zhang[2*] and Haisheng Zhu[3]

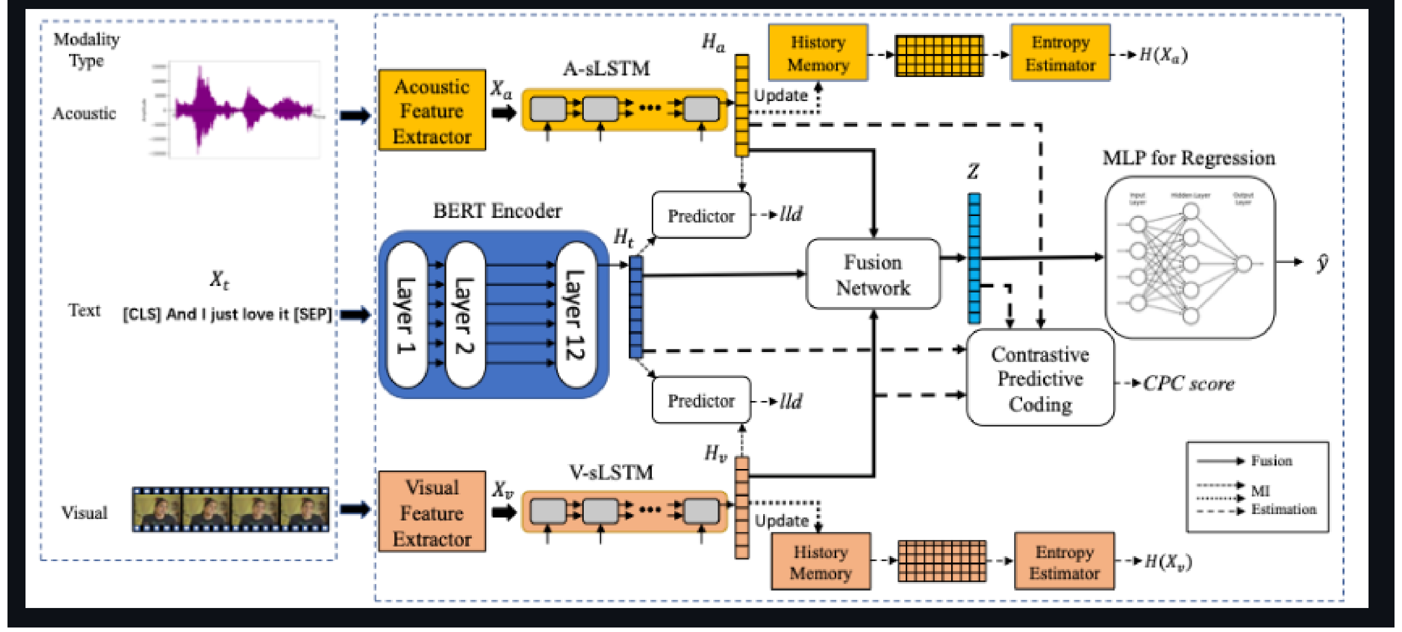
Future works

Future research in the fast-developing field of medical image analysis utilizing deep learning algorithms has a lot of potential to increase the precision and effectiveness of medical diagnosis and therapy. Some of these areas include:

7.8.1. Multi-modal image analysis

Future research in medical image analysis utilizing deep learning algorithms will focus on multi-modal picture analysis. Utilizing a variety of imaging modalities, including MRI, CT, PET, ultrasound, and optical imaging, allows for a more thorough understanding of a patient's anatomy and disease (101). This strategy can aid in enhancing diagnostic precision and lowering the possibility of missing or incorrect diagnoses. Multi-modal picture data may be used to train deep learning algorithms for a range of tasks, including segmentation, registration, classification, and prediction. An algorithm built on MRI and PET data, for instance, might be used to identify areas of the brain afflicted by Alzheimer's disease. Similarly, a deep learning algorithm could be trained on ultrasound and CT data to identify tumors in the liver. Multi-modal image analysis poses several challenges for deep learning algorithms. For example, different imaging modalities have different resolution, noise, and contrast characteristics, which can affect the performance of the algorithm. Additionally, multi-modal data can be more complex and difficult to interpret than single-modality data, requiring more advanced algorithms and computational resources (102). To address these challenges, researchers are developing new deep learning models and algorithms that can integrate and analyze data from multiple modalities. For example, multi-modal fusion networks can be used to combine information from different imaging modalities, while attention mechanisms can be used to focus the algorithm's attention on relevant features in each modality. Overall, multi-modal image analysis holds promise for improving the accuracy and efficiency of medical diagnosis and treatment using deep learning algorithms. As these technologies continue to evolve, it will be important to ensure that they are being used safely, ethically, and in accordance with relevant laws and regulations.

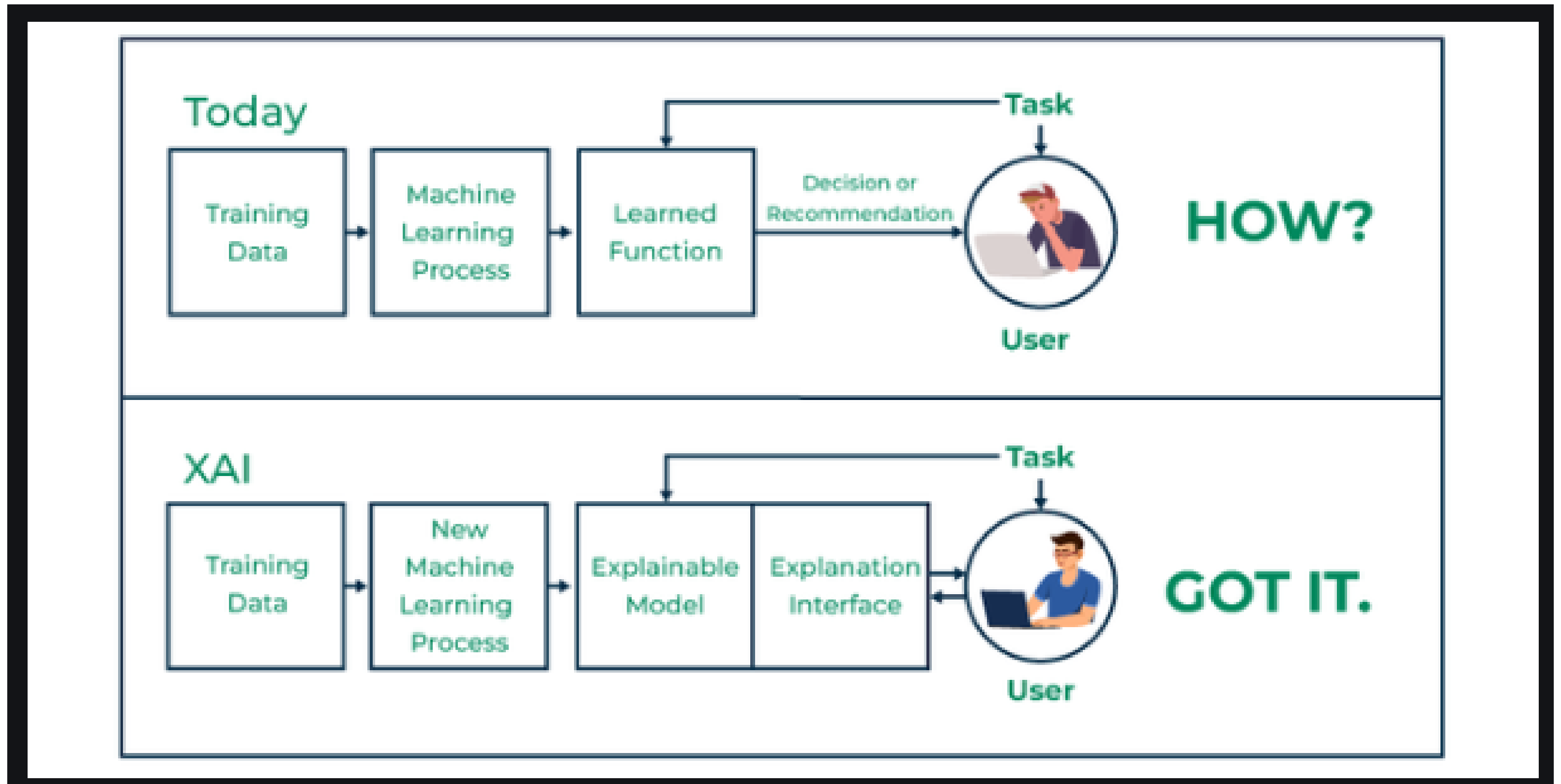
Multimodal-informax (MMIM) synthesizes fusion results from multi-modality input through a two-level mutual information (MI) maximization. We use BA (Barber-Agakov) lower bound and contrastive predictive coding as the target function to be maximized. To facilitate the computation, we design an entropy estimation module with associated history data memory to facilitate the computation of BA lower bound and the training process.



Future works

7.8.2. Explainable AI

Future research in deep learning algorithms for medical image analysis will focus on explainable AI (XAI). XAI is the capacity of an AI system to explain its decision-making process in a way that is intelligible to a human (103). XAI can assist to increase confidence in deep learning algorithms when employed in the context of medical image analysis, guarantee that they are utilized safely and morally, and allow clinicians to base their judgments more intelligently on the results of these algorithms. XAI in medical image analysis involves developing algorithms that can not only make accurate predictions or segmentations but also provide clear and interpretable reasons for their decisions. This can be particularly important in cases where the AI system's output contradicts or differs from the clinician's assessment or prior knowledge. One approach to XAI in medical image analysis is to develop visual explanations or heatmaps that highlight the regions of an image that were most important in the algorithm's decision-making process. These explanations can help to identify regions of interest, detect subtle abnormalities, and provide insight into the algorithm's thought process (104). Another approach to XAI in medical image analysis is to incorporate external knowledge or prior information into the algorithm's decision-making process. For example, an algorithm that analyzes brain MRIs could be designed to incorporate known patterns of disease progression or anatomical landmarks. Overall, XAI holds promise for improving the transparency, interpretability, and trustworthiness of deep learning algorithms in medical image analysis. As these technologies continue to evolve, it will be important to ensure that they are being used safely, ethically, and in accordance with relevant laws and regulations (105).



Explainable Artificial Intelligence Concept

Explainable Artificial Intelligence(XAI) - GeeksforGeeks

Future works

7.8.3. Transfer learning

Future research in the field of deep learning-based medical image processing will focus on transfer learning. Transfer learning is the process of using previously trained deep learning models to enhance a model's performance on a new task or dataset. Transfer learning can be particularly helpful in the interpretation of medical images as it can eliminate the requirement for significant volumes of labeled data, which can be challenging and time-consuming to gather. Researchers can use pre-trained models that have already been trained on huge datasets to increase the precision and effectiveness of their own models by taking advantage of the information and representations acquired by these models. Since transfer learning can do away with the need for large amounts of labeled data, which can be difficult and time-consuming to collect, it can be very useful in the interpretation of medical pictures. By utilizing the knowledge and representations amassed by pre-trained models that have previously been trained on massive datasets, researchers may utilize them to improve the accuracy and efficacy of their own models (106). The pre-trained model could be a useful place to start for the medical image analysis problem since it enables the model to learn from less data and might lessen the possibility of overfitting. Additionally, transfer learning may increase the generalizability of deep learning models used for medical picture interpretation. Medical image analysis models may be able to develop more reliable and generalizable representations of medical pictures that are relevant to a wider range of tasks and datasets by making use of pre-trained models that have learnt representations of real images. Transfer learning has the potential to enhance the effectiveness, precision, and generalizability of deep learning models used for medical image interpretation. As these technologies continue to evolve, it will be important to ensure that they are being used safely, ethically, and in accordance with relevant laws and regulations.

Transfer Learning / Pre-trained Model

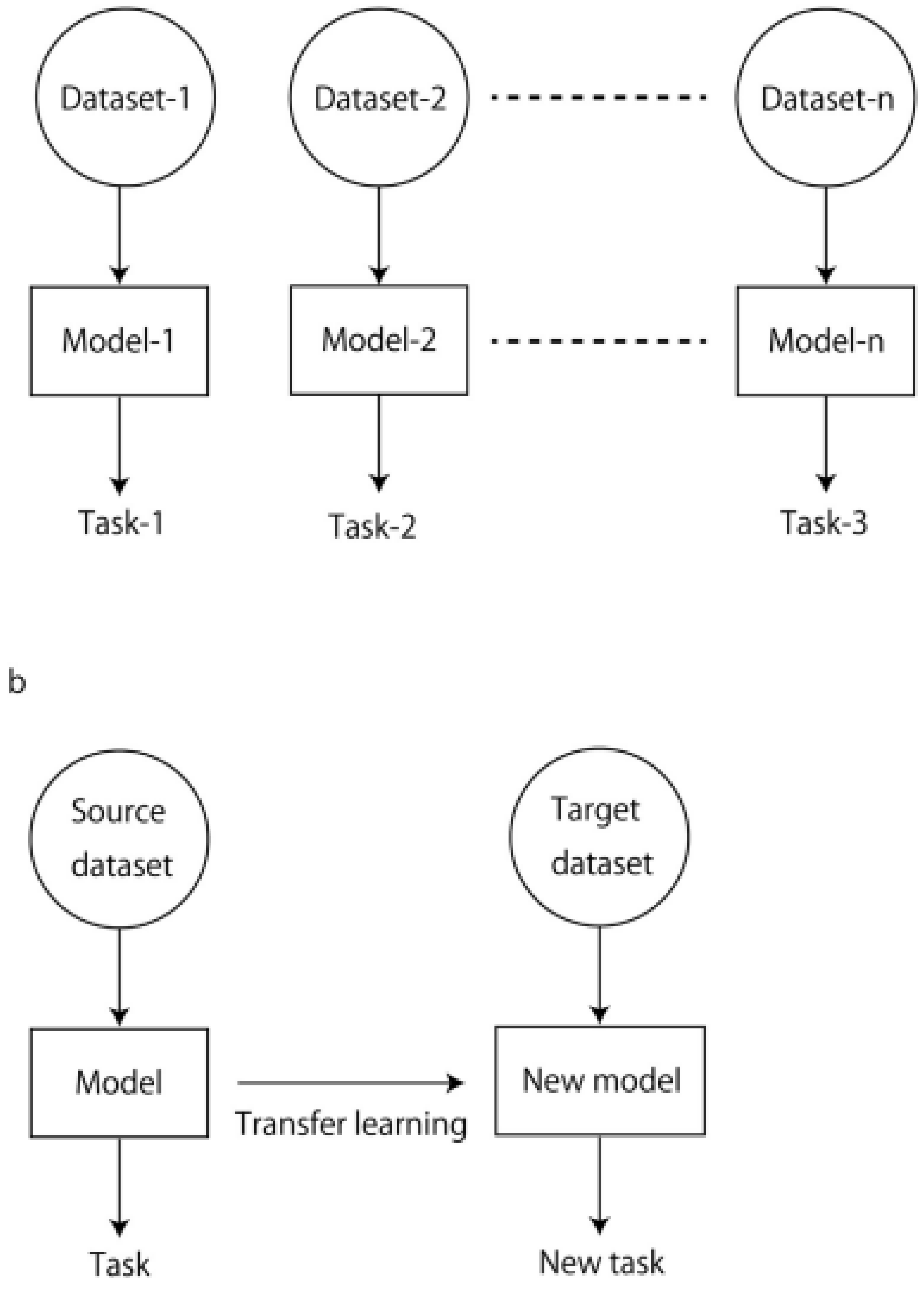


Fig. 1. Schematic diagrams for conventional machine learning approach and transfer learning process. (a): An illustration of conventional machine learning. (b): An illustration of transfer learning.

Transfer Learning / Pre-trained Model

NVIDIA

Home AI Data Center Driving Gaming Pro Graphics Robotics Healthcare Startups AI Podcast NVIDIA Life

What Is a Pretrained AI Model?

A pretrained AI model is a deep learning model that's trained on large datasets to accomplish a specific task, and it can be used as is or customized to suit application requirements across multiple industries.

Why Are Pretrained AI Models Used?

Instead of building an AI model from scratch, developers can use pretrained models and customize them to meet their requirements.

To build an AI application, developers first need an AI model that can accomplish a particular task, whether that's identifying a mythical horse, detecting a safety hazard for an autonomous vehicle or diagnosing a cancer based on medical imaging. That model needs a lot of representative data to learn from.

This learning process entails going through several layers of incoming data and emphasizing goals-relevant characteristics at each layer.

To create a model that can recognize a unicorn, for example, one might first feed it images of unicorns, horses, cats, tigers and other animals. This is the incoming data.

Then, layers of representative data traits are constructed, beginning with the simple — like lines and colors — and advancing to complex structural features. These characteristics are assigned varying degrees of relevance by calculating probabilities.

As opposed to a cat or tiger, for example, the more like a horse a creature appears, the greater the likelihood that it is a unicorn. Such probabilistic values are stored at each neural network layer in the AI model, and as layers are added, its understanding of the representation improves.

To create such a model from scratch, developers require enormous datasets, often with billions of rows of data. These can be pricey and challenging to obtain, but compromising on data can lead to poor performance of the model.

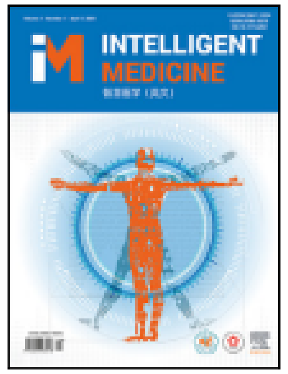
Transformers

Intelligent Medicine 3 (2023) 59–78

Contents lists available at ScienceDirect

Intelligent Medicine

journal homepage: www.elsevier.com/locate/imed



Review

Transformers in medical image analysis

Kelei He [1,2,#], Chen Gan [2,#], Zhuoyuan Li [1,2,#], Islem Rekik [3,4,#], Zihao Yin [2], Wen Ji [2], Yang Gao [2,5], Qian Wang [6,*], Junfeng Zhang [1,2,*], Dinggang Shen [6,7,8,*]

[1] Medical School of Nanjing University, Nanjing, Jiangsu 210093, China
[2] National Institute of Healthcare Data Science at Nanjing University, Nanjing, Jiangsu 210093, China
[3] BASIRA Laboratory, Faculty of Computer and Informatics Engineering, Istanbul Technical University, Istanbul, Turkey
[4] School of Science and Engineering, Computing, University of Dundee, UK
[5] State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, Jiangsu 210093, China
[6] School of Biomedical Engineering, ShanghaiTech University, Shanghai 201210, China
[7] Department of Research and Development, Shanghai United Imaging Intelligence Co., Ltd., Shanghai 200030, China
[8] Shanghai Clinical Research and Trial Center, Shanghai 201703, China

Transformers

K. He, C. Gan, Z. Li et al. Intelligent Medicine 3 (2023) 59–78

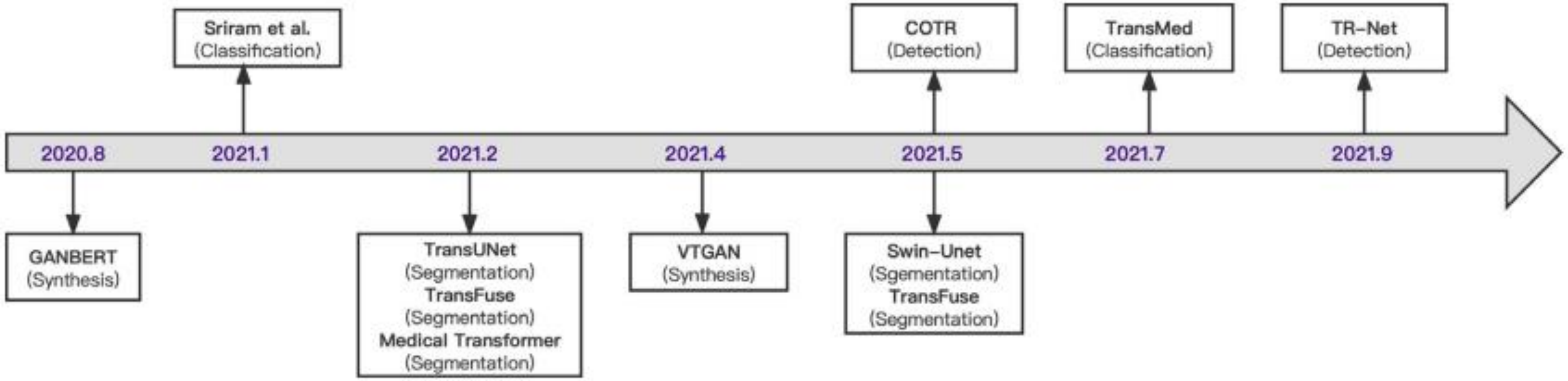


Figure 1. The development of transformers in medical image analysis. Selected methods are displayed relating to classification, detection, segmentation, and synthesis applications.

Transformers

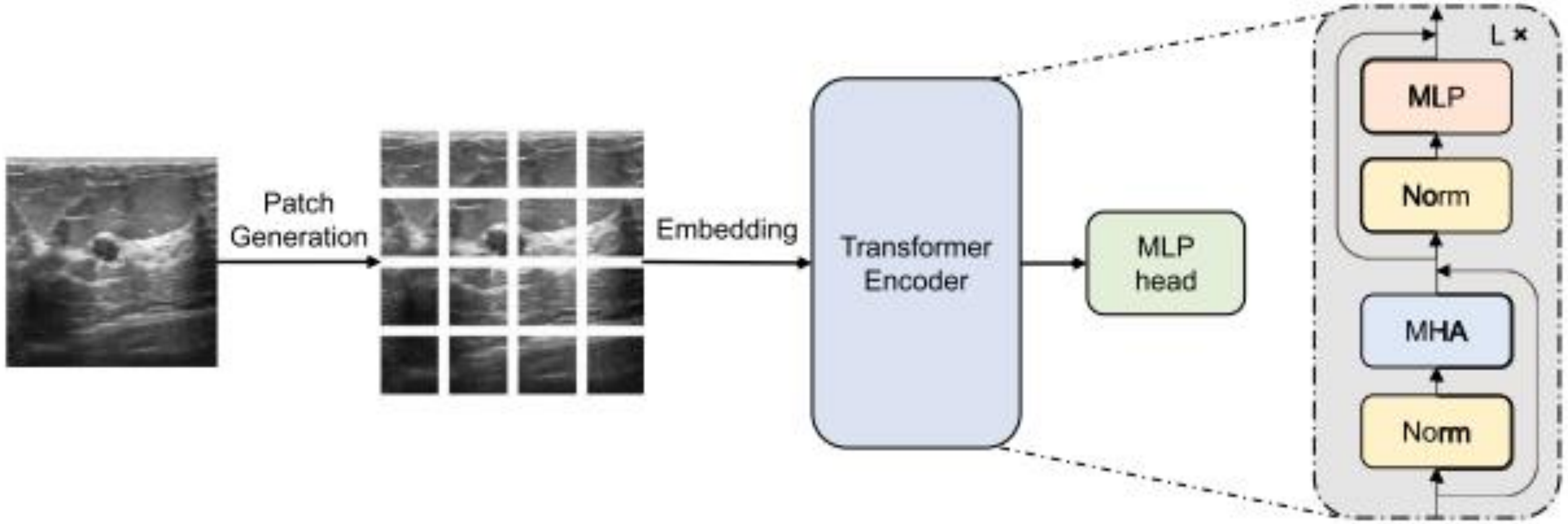


Fig. 6. Detailed structure of the transformer used in the MIA field. MHA refers to the multi-head attention module. Norm represents layer normalization and MLP illustrates the multilayer perception module.

SegFormer

Transformers
Search documentation Ctrl+K
V4.46.3 EN 135,569
Installation
Adding a new model to `transformers`
TUTORIALS
Run inference with pipelines
Write portable code with AutoClass
Preprocess data
Fine-tune a pretrained model
Train with a script
Set up distributed training with 🤗 Accelerate
Load and train adapters with 🤗 PEFT
Share your model

SegFormer

Overview

The SegFormer model was proposed in [SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers]() by Enze Xie, Wenhai Wang, Zhiding Yu, Anima Anandkumar, Jose M. Alvarez, Ping Luo. The model consists of a hierarchical Transformer encoder and a lightweight all-MLP decode head to achieve great results on image segmentation benchmarks such as ADE20K and Cityscapes.

The abstract from the paper is the following:

We present SegFormer, a simple, efficient yet powerful semantic segmentation framework which unifies Transformers with lightweight multilayer perception (MLP) decoders. SegFormer has two appealing features: 1) SegFormer comprises a novel hierarchically structured Transformer encoder which outputs multiscale features. It does not need positional encoding, thereby avoiding the interpolation of positional codes which leads to decreased performance when the testing resolution differs from

SegFormer
Overview
Usage tips
Resources
SegformerConfig
SegformerFeatureExtractor
SegformerImageProcessor
SegformerModel
SegformerDecodeHead
SegformerForImageClassification
SegformerForSemanticSegmentation
TFSegformerDecodeHead
TFSegformerModel
TFSegformerForImageClassification

SegFormer

https://huggingface.co/nvidia/mit-b0

Hugging Face | Search models, datasets, users... | Models | Datasets | Spaces | Posts | Docs | Enterprise | Pricing

nvidia/mit-b0 | like 30 | Follow NVIDIA 4,577

Image Classification | Transformers | PyTorch | TensorFlow | imagenet_1k | segformer | vision | Inference Endpoints | arxiv:2105.15203 | License: other

Model card | Files and versions | Community 5 | Train | Deploy | Use this model

Edit model card

SegFormer (b0-sized) encoder pre-trained-only

SegFormer encoder fine-tuned on Imagenet-1k. It was introduced in the paper SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers by Xie et al. and first released in this repository.

Disclaimer: The team releasing SegFormer did not write a model card for this model so this model card has been written by the Hugging Face team.

Model description

SegFormer consists of a hierarchical Transformer encoder and a lightweight all-MLP decode head to achieve great results on semantic segmentation benchmarks such as

Downloads last month
43,045

Inference Examples

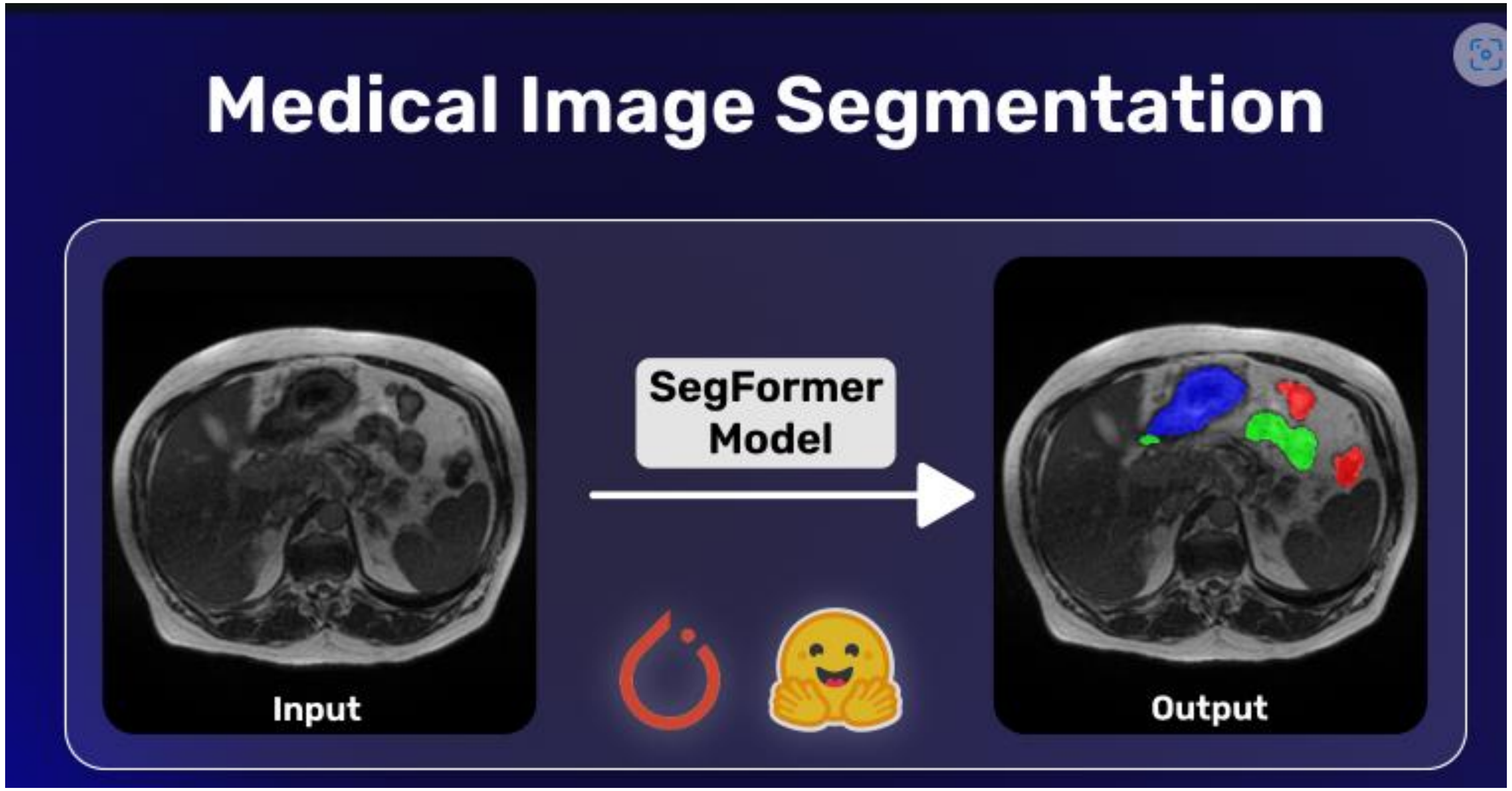
Image Classification

This model does not have enough activity to be deployed to Inference API (serverless) yet. Increase its social visibility and check back later, or deploy to Inference Endpoints (dedicated) instead.

Model tree for nvidia/mit-b0

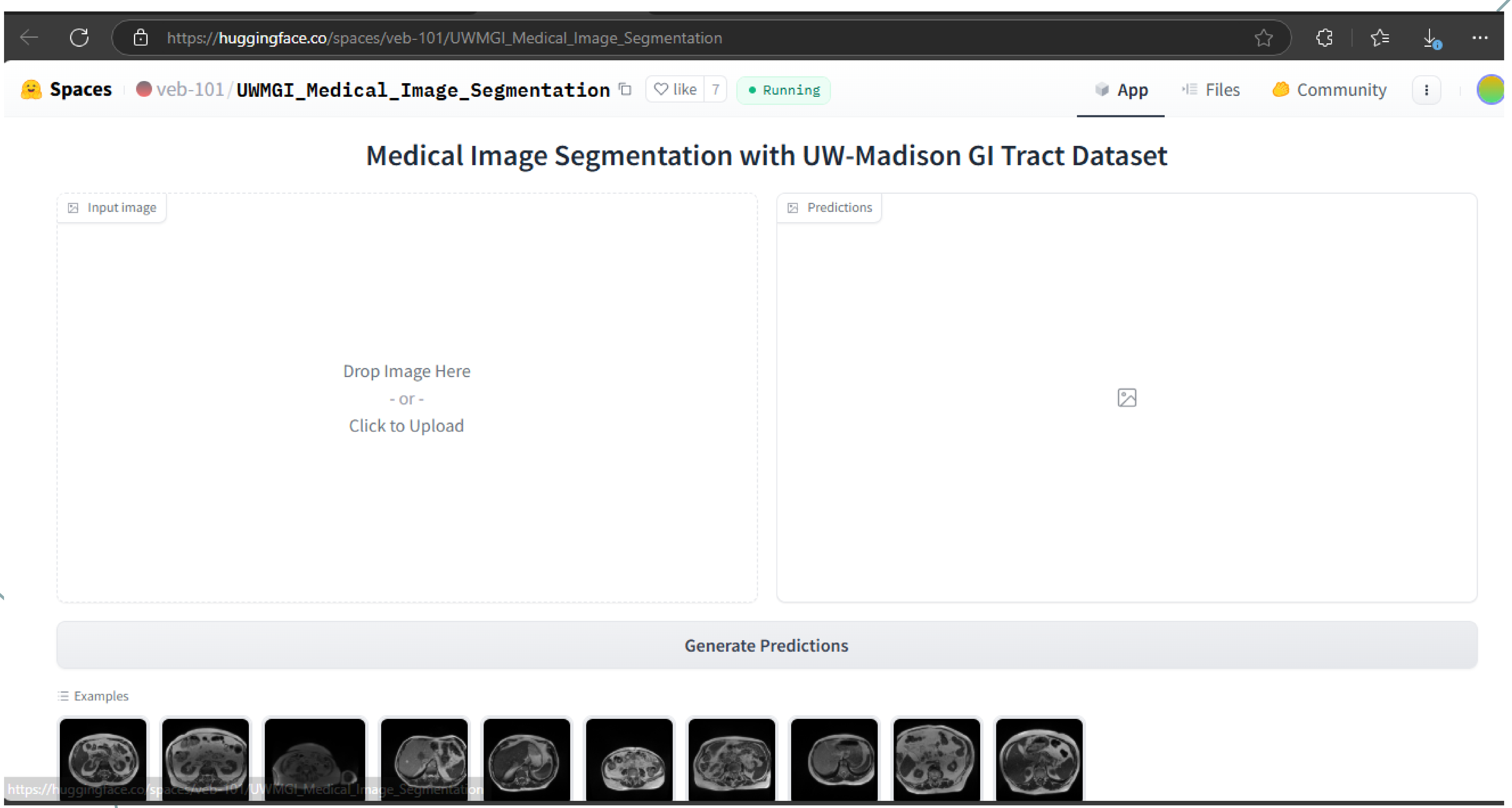
Adapters	2 models
Finetunes	319 models
Quantizations	1 model

SegFormer



[learnopencv/Medical-Image-Segmentation-Using-HuggingFace-&-PyTorch/README.md at master · spmallick/learnopencv](#)

SegFormer


https://huggingface.co/spaces/veb-101/UWMGI_Medical_Image_Segmentation
Spaces
veb-101/UWMGI_Medical_Image_Segmentation
like 7
Running
App
Files
Community
Medical Image Segmentation with UW-Madison GI Tract Dataset
Input image
Drop Image Here
- or -
Click to Upload
Predictions
Generate Predictions
Examples

End to End DL Pipeline

Informatics in Medicine Unlocked 50 (2024) 101582

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

journal homepage: www.elsevier.com/locate/imu

Deep learning-based mobile application for the enhancement of pneumonia medical imaging analysis: A case-study of West-Meru Hospital

Japheth Mumo Kimeu *, Michael Kisangiri, Hope Mbelwa, Judith Leo

School of Computational and Communication Science and Engineering, Nelson Mandela African Institution of Science and Technology (NM-AIST), P.O. Box 447, Arusha, Tanzania

A B S T R A C T

Pneumonia remains a significant global health challenge, demanding innovative solutions. This study presents a novel approach to pneumonia diagnosis and medical imaging analysis, leveraging advanced technologies. The study used a Literature Review Methodology to study various scientific articles and involved healthcare staff, including Doctors, Nurses, Radiologists and the community, in sharing their requirements for the study. The findings led to the proposal for the integration of Deep Learning techniques, including Convolutional Neural Network (CNN), as well as tools like YOLOv8, Roboflow, and Ultralytics, to revolutionize pneumonia detection and classification. The EfficientDet-Lite2 model architecture was subsequently used to generate a TensorFlow Lite Model, deployable in both Android and iOS mobile applications. The study's outcomes reveal a substantial improvement in the precision and recall metrics. These results signify a promising step forward in empowering healthcare professionals with timely and reliable results for optimal patient management.

End to End DL Pipeline

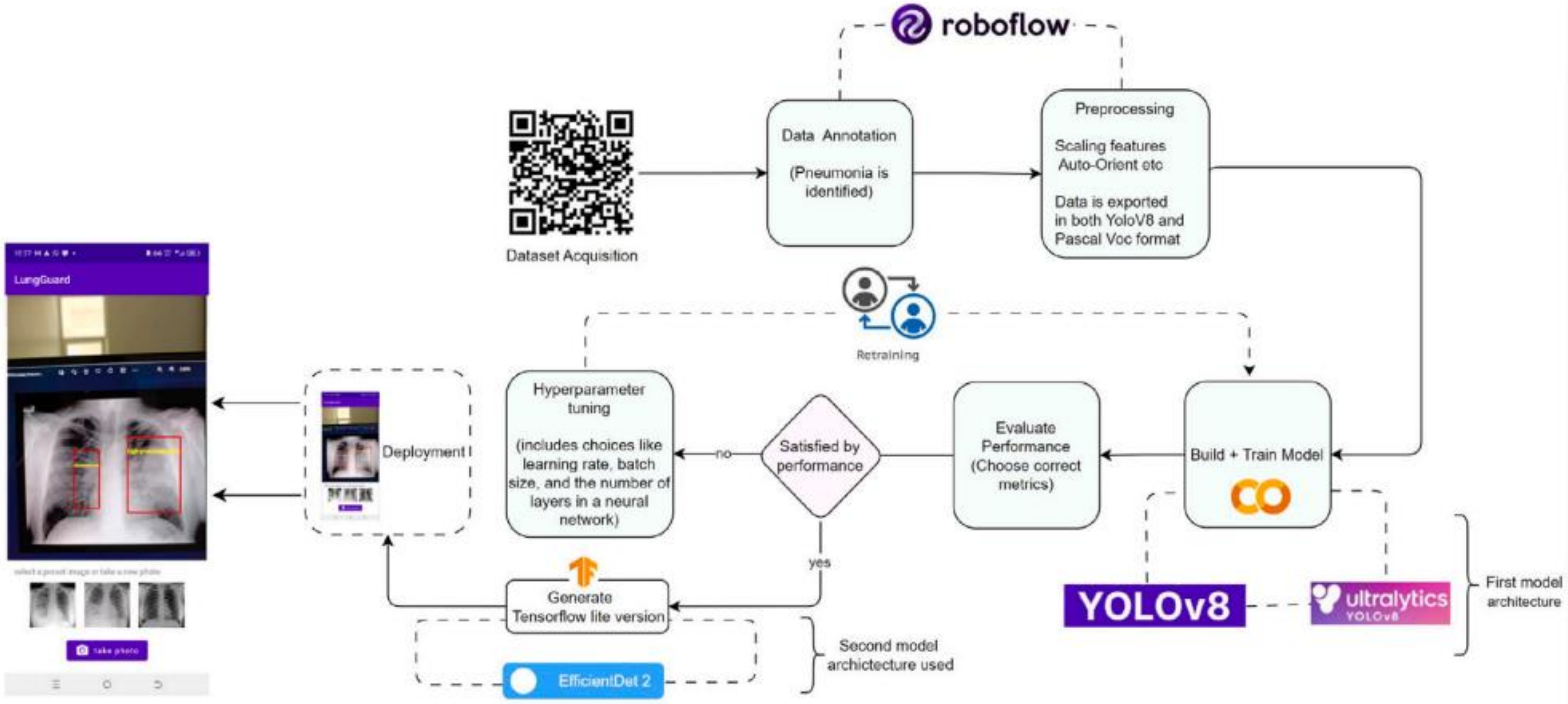


Fig. 5. Block diagram of the system.

End to End DL Pipeline

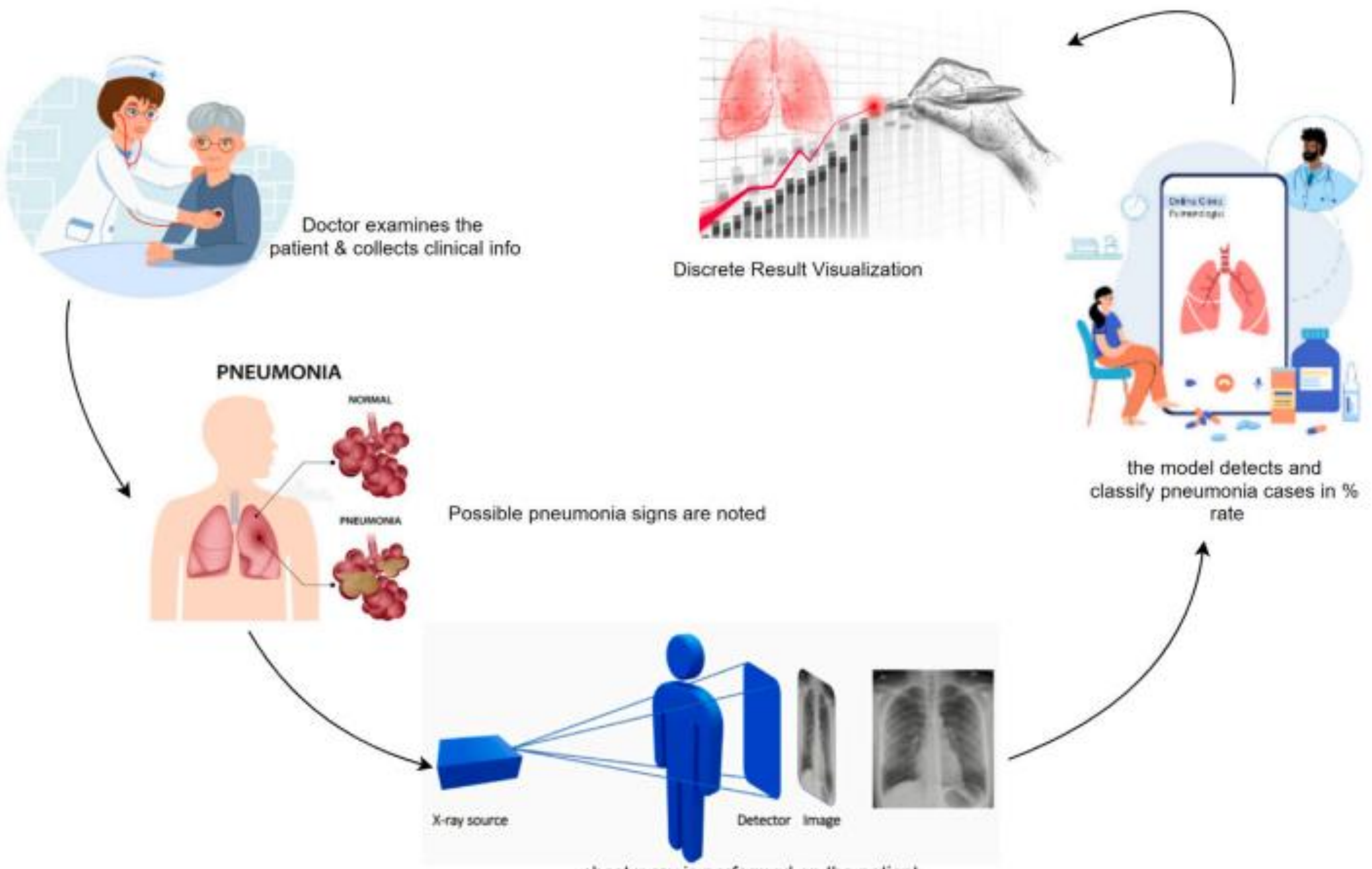
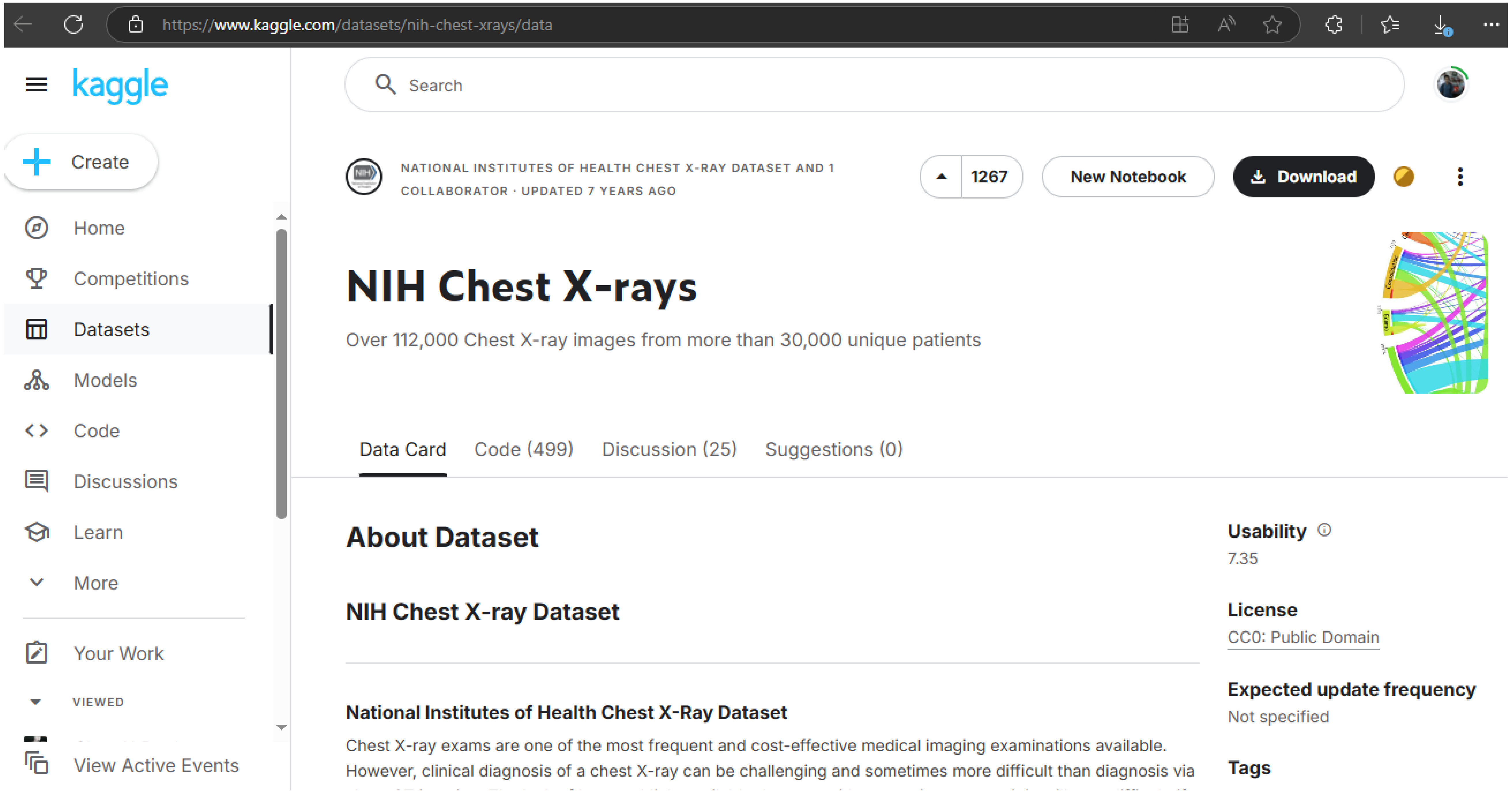
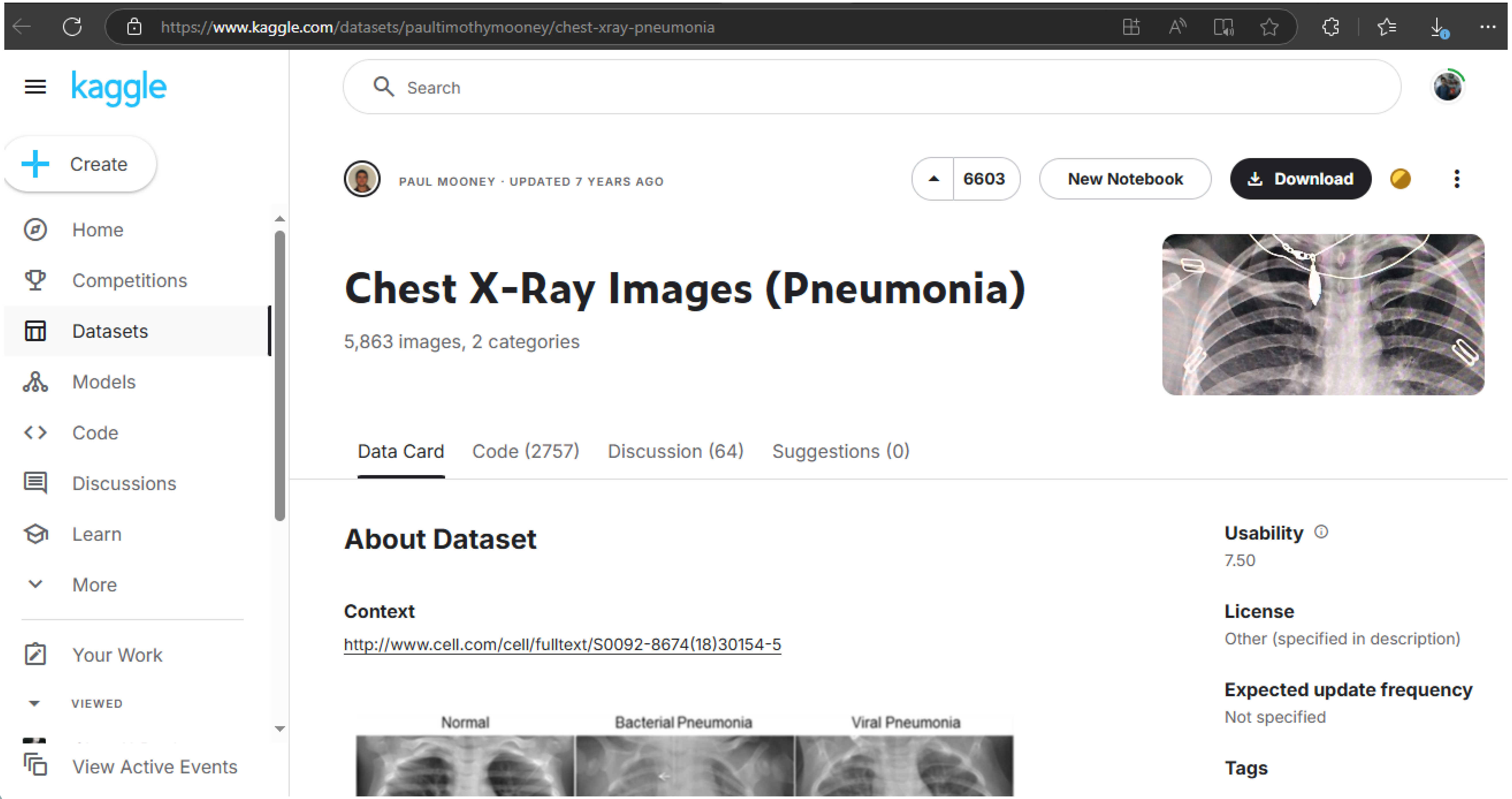


Fig. 7. The Proposed Deep Learning Pneumonia Predictive System.

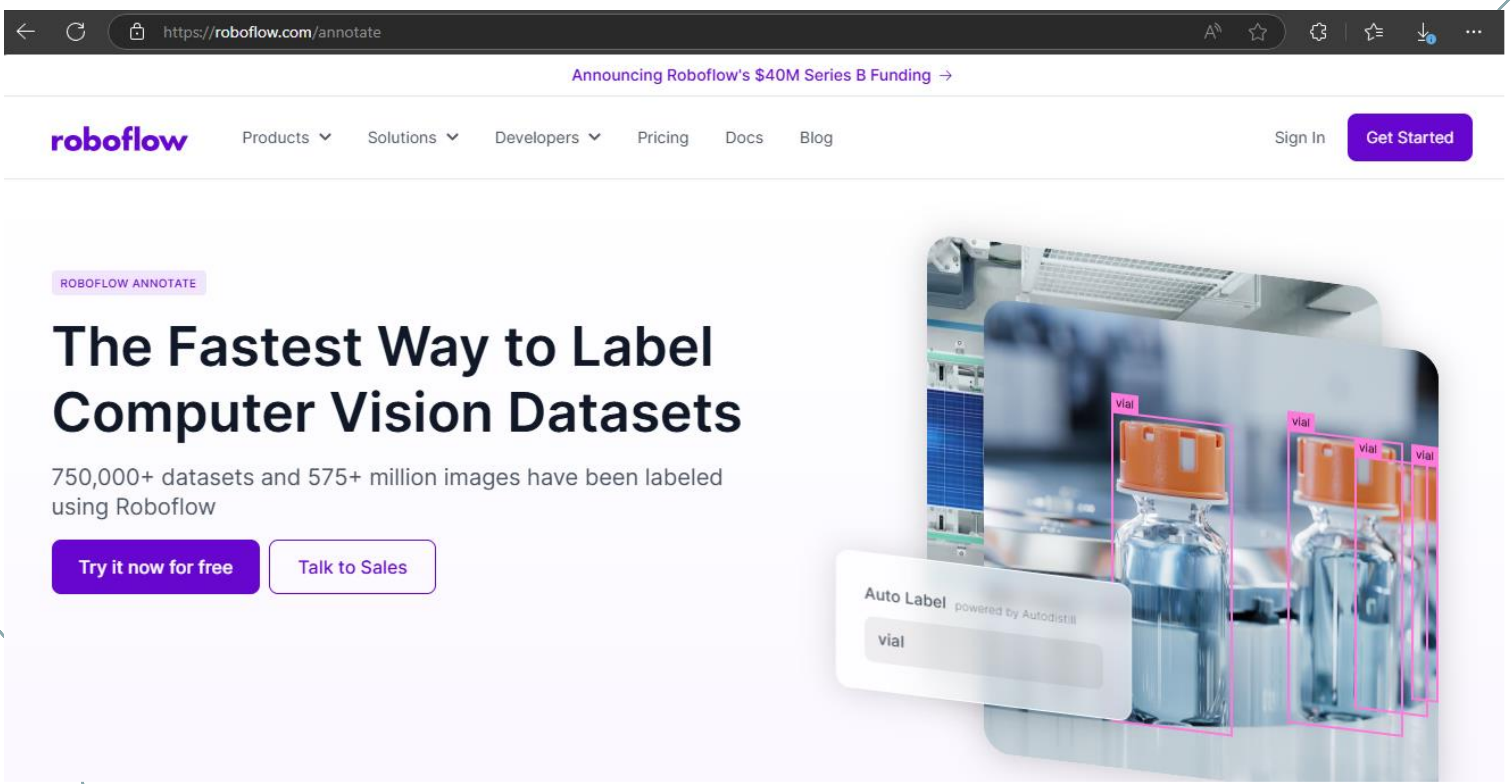
End to End DL Pipeline



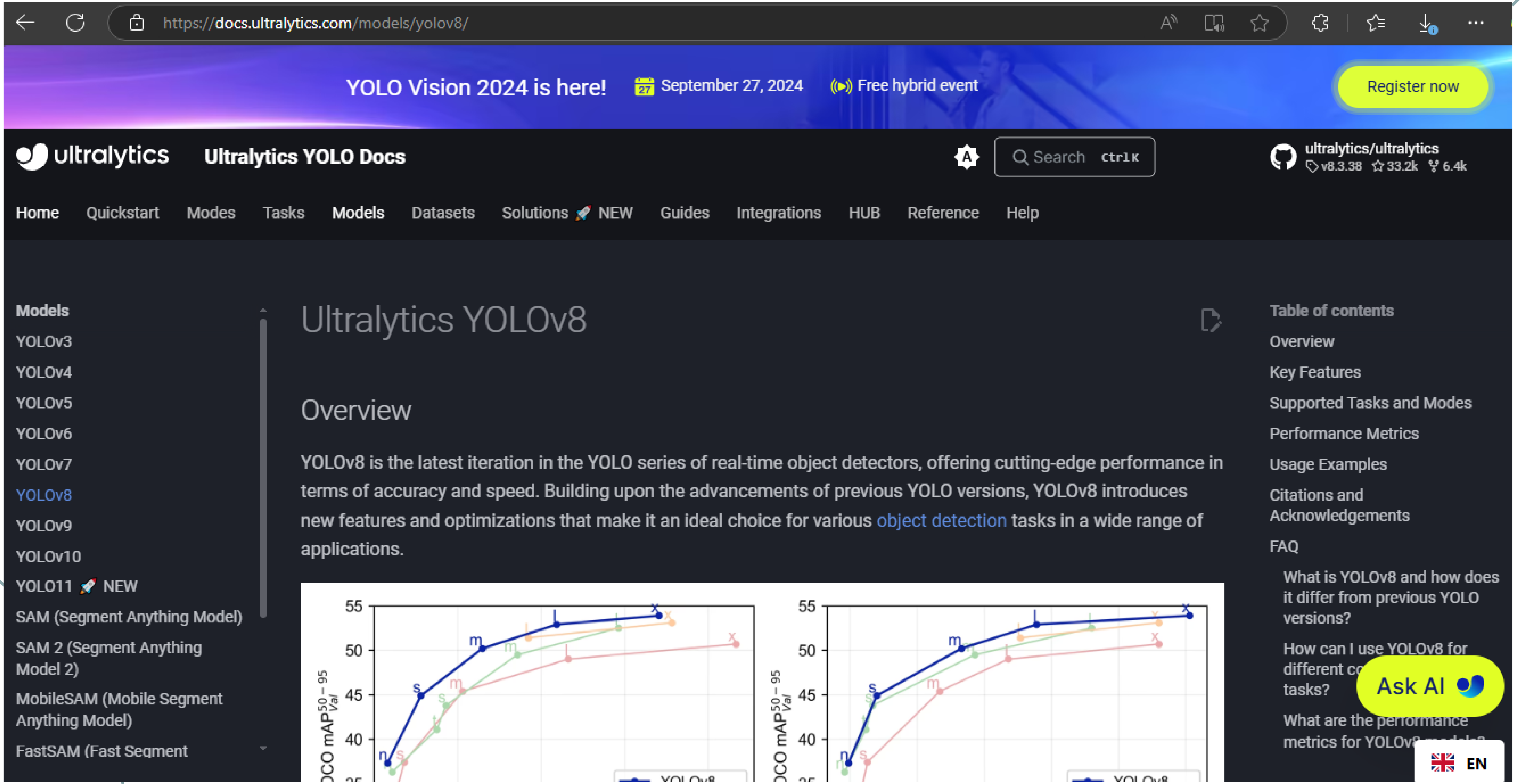
End to End DL Pipeline



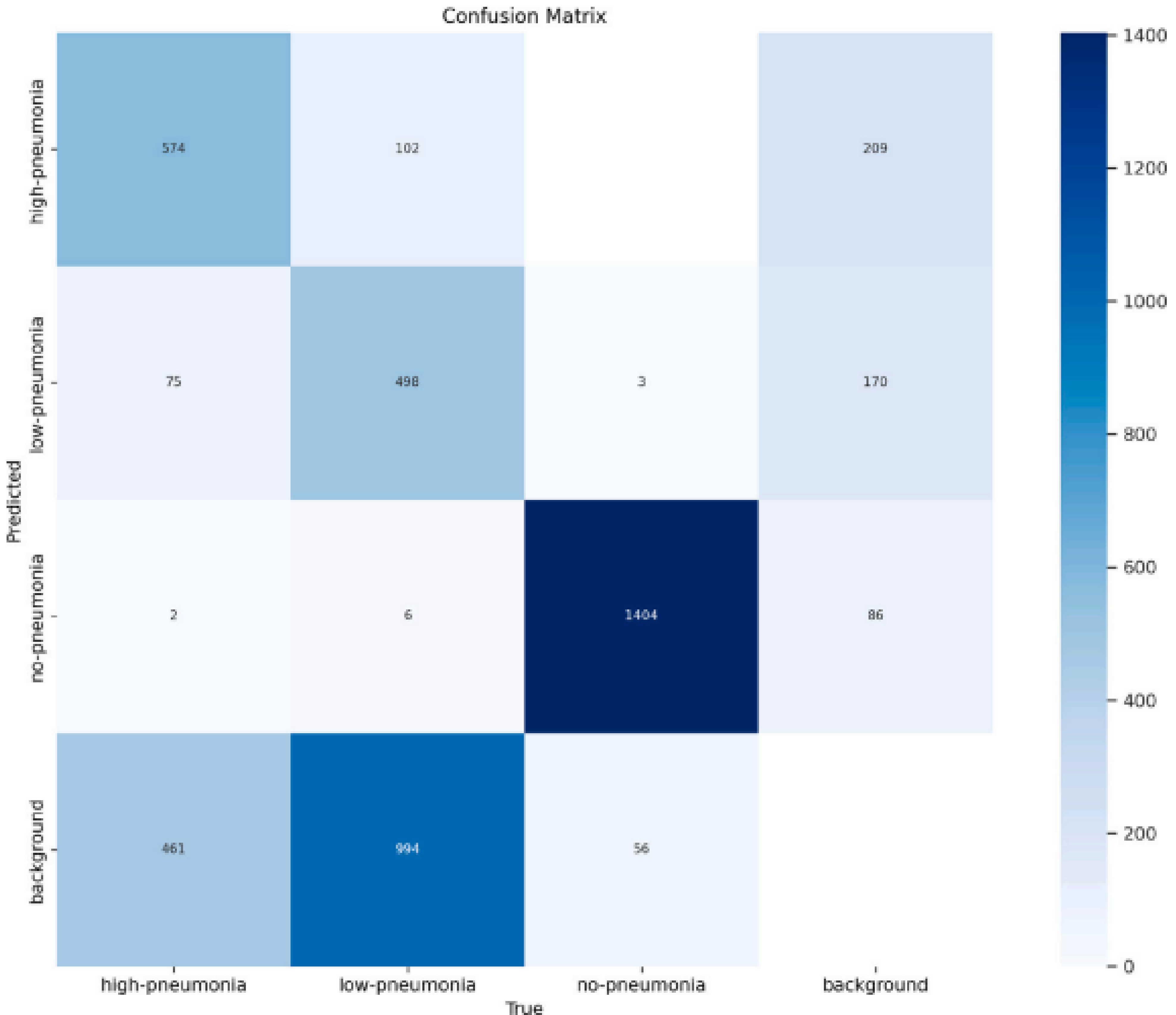
End to End DL Pipeline



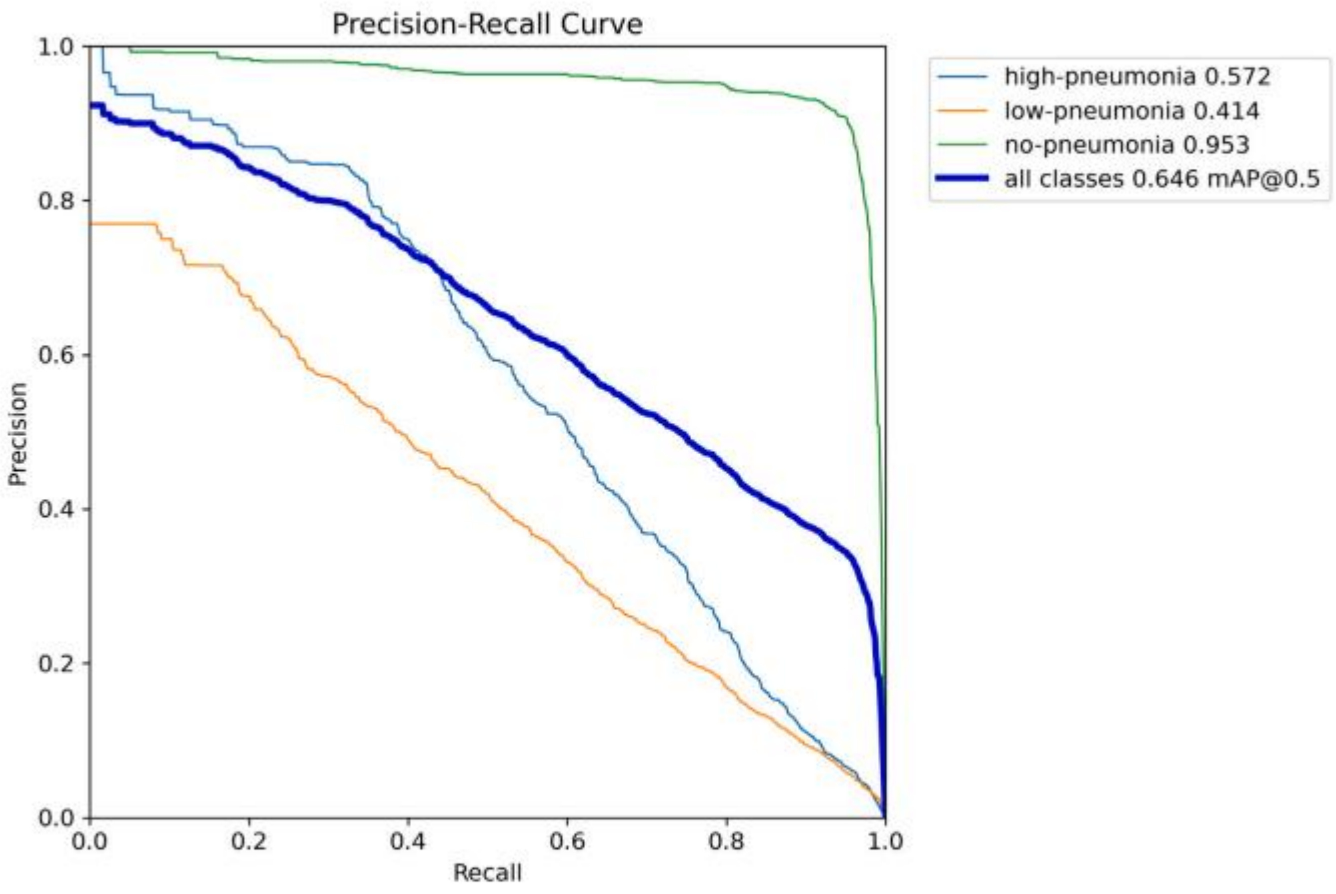
End to End DL Pipeline



End to End DL Pipeline



End to End DL Pipeline



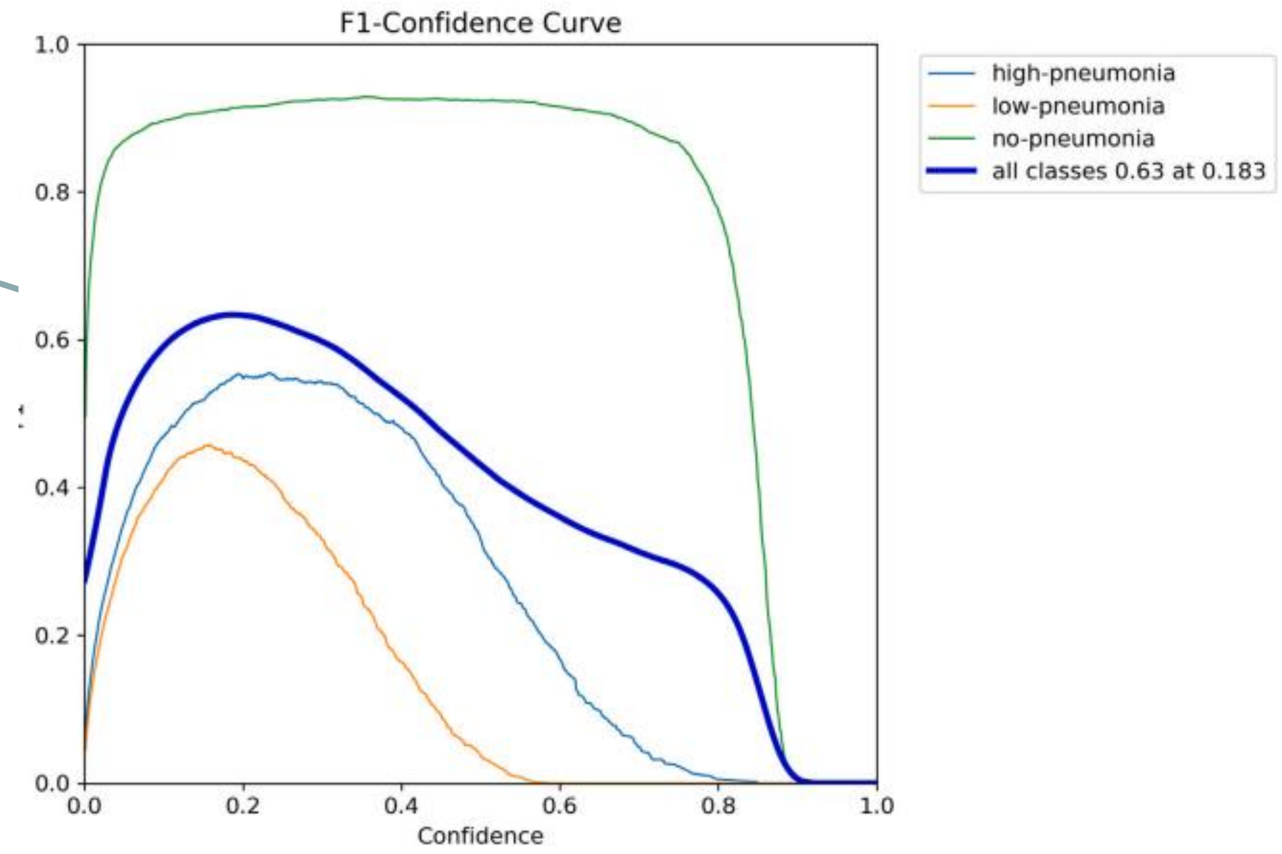


Fig. 13. F1-Confidence Curve for different classes.

We acknowledge our pneumonia detection model's initial performance, which currently stands at a commendable precision of 62.4% and a recall of 65.6%. With a commitment to refining our approach, we anticipate significant performance improvements. To augment our capabilities, we have also integrated a pre-trained YOLOv8n model, celebrated for its real-time processing and accuracy in object detection tasks. The trained YoloV8 Model is then passed through the EfficientDet2 Model Architecture to generate the final version of the deployable TensorFlow Lite model. This versatile addition extends the scope of our solution, enabling it to efficiently identify pneumonia-related anomalies in radiological images. As we progress, our focus remains on optimizing this combined model to further elevate its precision and recall, reinforcing its potential impact in the realm of medical diagnostics.

End to End DL Pipeline

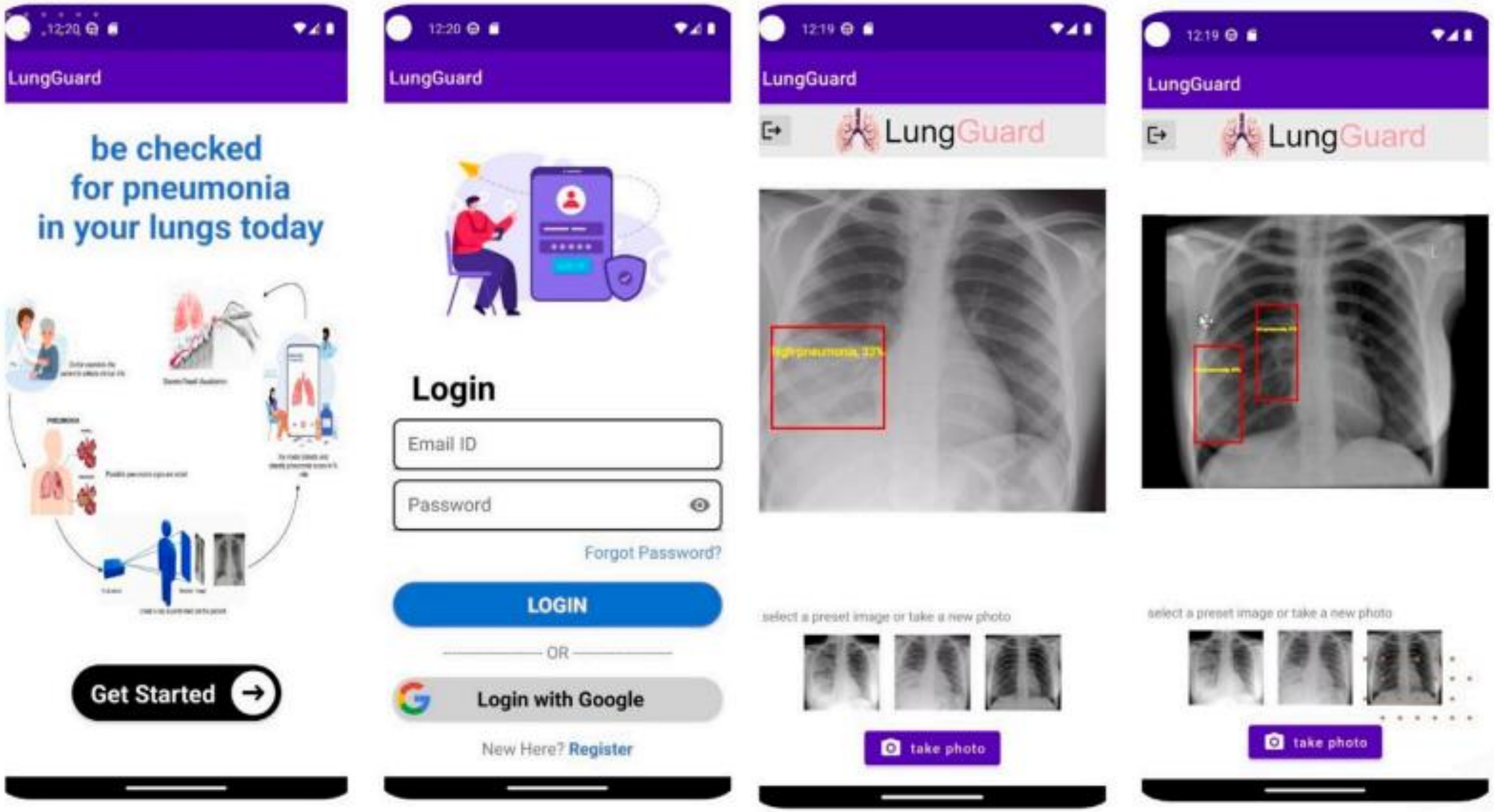


Fig. 14. Results of the Deployed TensorFlow Lite Model.

End to End DL Pipeline

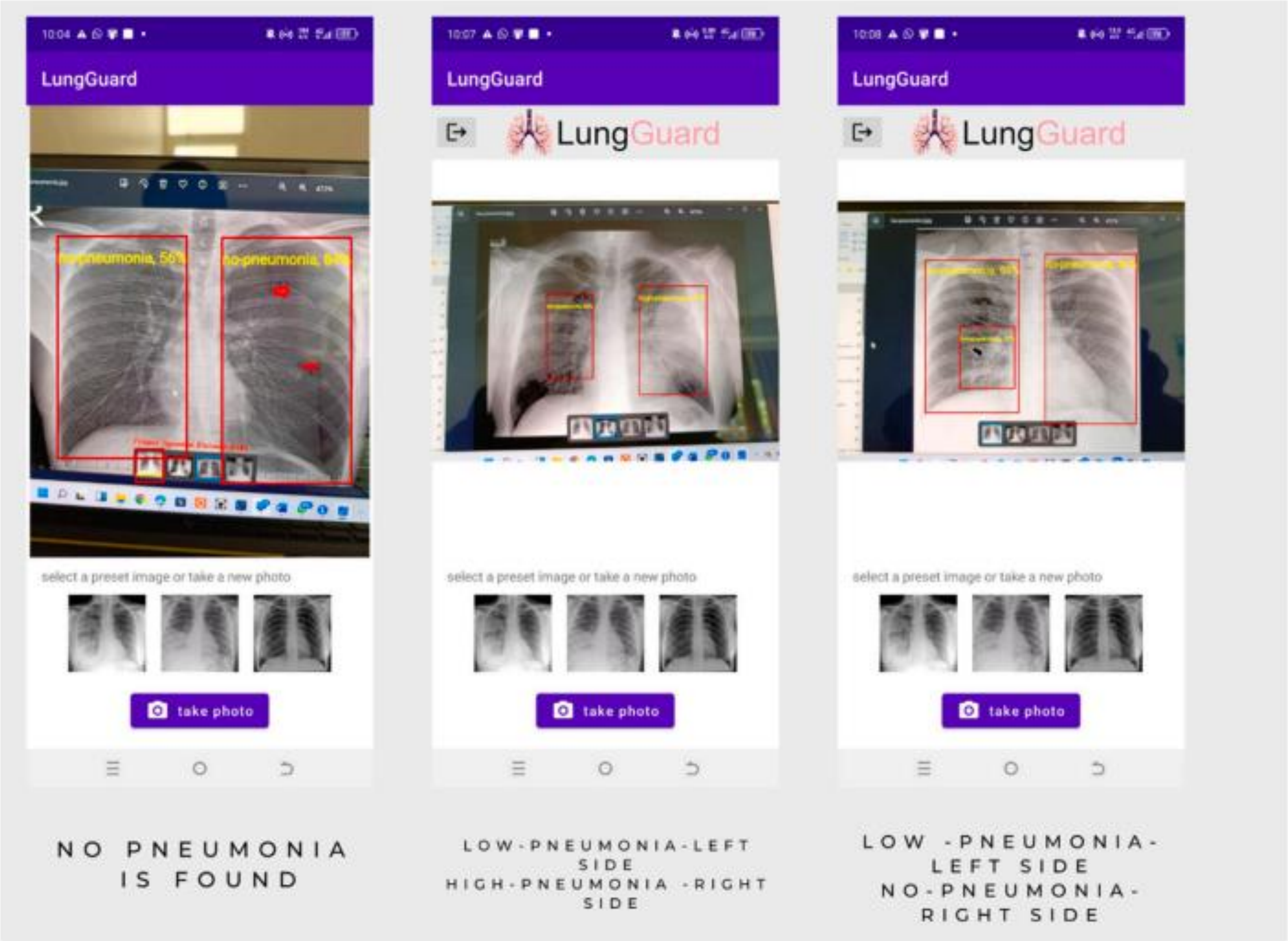


Fig. 15. Predictive Analysis 01.

End to End DL Pipeline

Challenges in model development and deployment

The model shows strong performance in detecting the "no-pneumonia" class, indicating its reliability in identifying cases without pneumonia. However, improvements are needed for the "high-pneumonia" and "low-pneumonia" classes. Addressing class imbalance through re-weighting or oversampling, incorporating advanced data augmentation, and exploring more sophisticated architectures could enhance recall and reduce false positives. Calibration techniques and fine-tuning hyperparameters may improve training convergence and performance. Collecting more labeled data, especially for underrepresented classes, will further improve generalization and detection capabilities. These enhancements will bolster the model's overall effectiveness in pneumonia detection and classification, supporting its potential for practical applications and further research. Training a large dataset was a challenge due to limited Google Colab resources. The annotation process as well required significant time as each image needed to be examined. Furthermore, maintaining TFLite code required ongoing adjustments due to changing dependencies. Keeping up with updates and ensuring compatibility became a recurring challenge, which also applied to Flutter packages and Kotlin language changes.

Perkembangan Topik ini di Indonesia

HEALTHSENSE : JOURNAL OF PUBLIC HEALTH PERSPECTIVE

Vol. 01 No. 01, Oktober 2024
journal.ininnawaparaedu.com

Studi Literatur Penerapan Deep Learning dalam Analisis Citra Medis di Indonesia

Minhajul Yusri Khairi[1], Eliyah Acantha Manapa Sampetoding[2*], Yulita Sirinti Pongtambing[3]

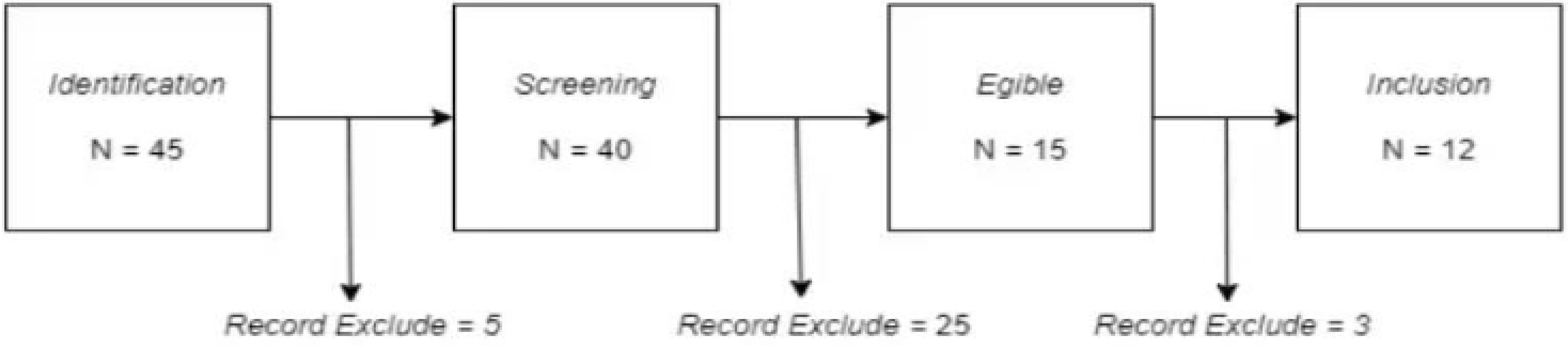
[1,2*] Sistem Informasi, Universitas Hasanuddin
[3] Administrasi Kesehatan, Universitas Negeri Makassar

[1]minhajulkhairi@gmail.com, [2]eliyahacantha@unhas.ac.id, [3]yulita.sirinti@unm.ac.id*

ABSTRAK

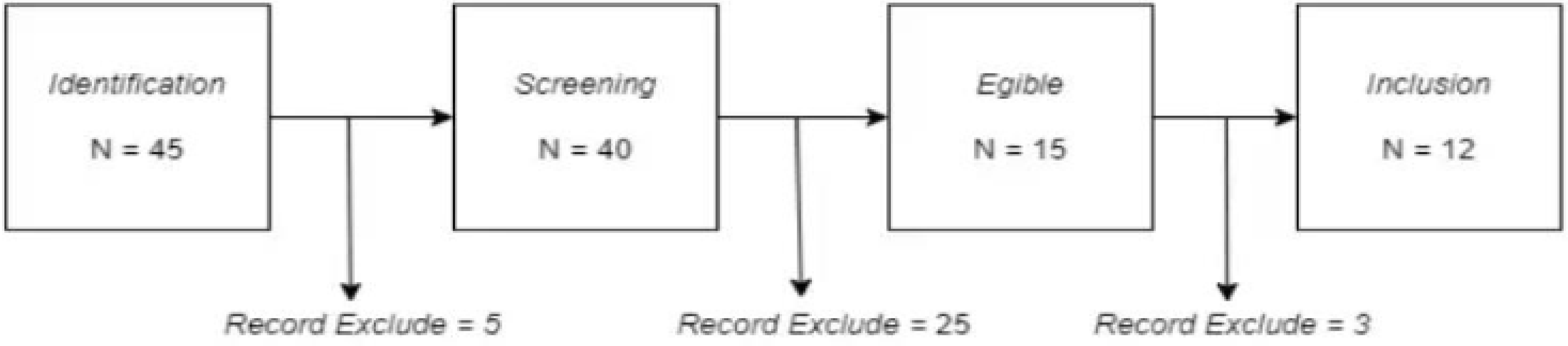
Artificial Intelligence (AI) telah berkembang pesat, termasuk di bidang kesehatan, khususnya dalam analisis citra medis. Salah satu cabangnya, Deep Learning (DL), menunjukkan kemampuan luar biasa dalam mengidentifikasi pola dan mengolah data citra medis. Penelitian ini bertujuan untuk melakukan tinjauan pustaka sistematis mengenai penerapan deep learning dalam analisis citra medis di Indonesia, menggunakan metode Systematic Literature Review (SLR) berbasis model Kitchenham. Dari 45 artikel yang diidentifikasi, 12 artikel dipilih berdasarkan kriteria inklusi, eksklusi, dan relevansi untuk dianalisis lebih lanjut. Hasil studi menunjukkan bahwa metode deep learning seperti Fourier Adaptive Recognition System (FARS) dan Residual Neural Network (ResNet50) telah berhasil meningkatkan akurasi diagnosa penyakit. Namun, tantangan yang dihadapi meliputi keterbatasan infrastruktur medis, kurangnya data berkualitas, serta perlunya penerapan yang lebih luas di fasilitas kesehatan. Temuan ini menunjukkan potensi besar deep learning untuk meningkatkan pelayanan kesehatan di Indonesia, asalkan tantangan-tantangan tersebut dapat diatasi.

Perkembangan Topik ini di Indonesia



Gambar 1. Proses Pemilihan Artikel

Perkembangan Topik ini di Indonesia



Gambar 1. Proses Pemilihan Artikel

Perkembangan Topik ini di Indonesia

Judul	**Teknologi**	**Klasifikasi**	**Hasil**
Enhancing medical image analysis with unsupervised domain adaptation approach across microscopes and magnifications	Fourier Adaptive Recognition System (FARS)	Digital	Penelitian ini memperkenalkan FARS sebagai model yang dirancang untuk pengenalan adaptif parasit malaria, namun juga memiliki potensi dalam diagnostik tumor dan kanker. Dengan memanfaatkan pelabelan segmentasi semantik, pelatihan adversarial, dan Color Domain Aware Fourier Domain Adaptation (F2DA), FARS meningkatkan ekstraksi fitur lintas domain dan magnifikasi. Hasilnya menunjukkan peningkatan kinerja signifikan pada adaptasi lintas domain dan magnifikasi, menjadikannya langkah maju dalam pengenalan parasit dan citra medis lainnya.

Perkembangan Topik ini di Indonesia

Simply Fine-Tuned Deep Learning-Based Classification for Breast Cancer with Mammograms	Convolutional Neural Network (ResNet50)	Digital	Penelitian ini mengimplementasikan model Residual Neural Network 50 (ResNet50) bersama dengan algoritma gradient adaptif, adaptive moment estimation, stochastic gradient descent, serta teknik augmentasi data dan fine-tuning untuk mengklasifikasikan massa payudara dari gambar mammogram. Hasil evaluasi menunjukkan bahwa model ini mampu mengklasifikasikan kanker payudara ke dalam kategori jinak dan ganas dengan hasil yang memuaskan dalam hal akurasi, nilai-p, AUC, sensitivitas, presisi, F1-score, spesifisitas, dan kappa. Model ini menunjukkan kelayakan dalam membantu diagnosis kanker payudara dari gambar mammogram.

di Universitas Matana?

Pengaruh Variasi Optimizer, Batch Size dan Jumlah EPOCH Terhadap Deteksi dan Klasifikasi Kanker Darah Menggunakan MobileNet-V2

Grace Cara Eirene Kakunsi

THANK YOU