ҚАЗАҚСТАН РЕСПУБЛИКАСЫНЫҢ ҒЫЛЫМ ЖӘНЕ ЖОҒАРЫ БІЛІМ МИНИСТРЛІГІ
МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РЕСПУБЛИКИ КАЗАХСТАН
MINISTRY OF SCIENCE AND HIGHER EDUCATION OF THE REPUBLIC OF KAZAKHSTAN

ХАЛЫҚАРАЛЫҚ УНИВЕРСИТЕТІ
МЕЖДУНАРОДНЫЙ УНИВЕРСИТЕТ
INTERNATIONAL UNIVERSITY

ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ

МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES

2022 (3) 4

Қазан-желтоқсан

ISSN 2708–2032 (print)
ISSN 2708–2040 (online)

Халықаралық ақпараттық және коммуникациялық технологиялар журналы
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)

Меншіктенуші: «Халықаралық ақпараттық технологиялар университеті» АҚ (Алматы қ.).

Қазақстан Республикасы Ақпарат және әлеуметтік даму министрлігінің Ақпарат комитетінде – **20.02.2020** жылы берілген.
№ KZ82VPY00020475 мерзімдік басылым тіркеуіне қойылу туралы куәлік.

Тақырыптық бағыты: ақпараттық технологиялар, әлеуметтік-экономикалық жүйелерді дамытудағы цифрлық технологиялар, ақпараттық қауіпсіздік және коммуникациялық технологияларға арналған.

Мерзімділігі: жылына 4 рет.

Тиражы: 100 дана

Редакцияның мекенжайы: 050040, Алматы қ-сы, Манас к-сі, 34/1, 709-кабинет, тел: +7 (727) 244-51-09).

E-mail: ijict@iitu.edu.kz Журнал сайты: https://journal.iitu.edu.kz

---

Международный журнал информационных и коммуникационных технологий
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)

Собственник: АО «Международный университет информационных технологий» (г. Алматы).

Свидетельство о постановке на учет периодического печатного издания в Министерство информации и общественного развития Республики Казахстан **№ KZ82VPY00020475,** выданное от **20.02.2020 г.**

Тематическая направленность: информационные технологии, информационная безопасность и коммуникационные технологии, цифровые технологии в развитии социо-экономических систем**.**

Периодичность: 4 раза в год.

Тираж: 100 экземпляров.

Адрес редакции: 050040 г. Алматы, ул. Манаса 34/1, каб. 709, тел: +7 (727) 244-51-09).

E-mail: ijict@iitu.edu.kz Сайт журнала: https://journal.iitu.edu.kz

«International Journal of Information and Communication Technologies»
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Owner: International Information Technology University JSC (Almaty).
The certificate of registration of a periodical printed publication in the Ministry of Information and Social Development of the Republic of Kazakhstan, Information Committee No. KZ82VPY00020475, issued on 20.02.2020.
Thematic focus: information technology, digital technologies in the development of socio-economic systems, information security and communication technologies
Periodicity: 4 times a year.
Circulation: 100 copies.
Editorial address: 050040. Manas st. 34/1, Almaty. +7 (727) 244-51-09). E-mail: ijict@iitu.edu.kz
Journal website: https://journal.iitu.edu.kz

МАЗМҰНЫ

СОДЕРЖАНИЕ

CONTENTS

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Vol. 3. Is. 4. Number 12 (2022). Pp. 70–83
Journal homepage: https://journal.iitu.edu.kz
https://doi.org/*10.54309/IJICT.2022.12.4.006*

COMPARATIVE ANALYSIS OF DEEP LEARNING METHODS FOR PNEUMONIA DETECTION ON X-RAY IMAGES

K.A. Gamri[1*], B.S. Omarov[2], Bohdan Haidabrus[3]

Gamri Khaled Adnan — Master's student, Department of Computer Engineering, International University of Information Technology
Omarov Batyrkhan Sultanovich — PhD, Assistant Professor of the Department of MKM, International University of Information Technology. Associate Professor of the Department "Information Systems", Kazakh National University. al-Farabi
Bohdan Haidabrus — PhD, Cand.of Tech.Science Riga Technical University, Riga, Latvia
https://orcid.org/0000-0002-9040-9058

Abstract. Pneumonia is a potentially fatal bacterial illness that affects one or both lungs in humans and is frequently caused by the bacterium Streptococcus pneumoniae. According to the World Health Organization, pneumonia accounts for one in every three fatalities in India (WHO). Expert radiotherapists must evaluate chest X-rays used to diagnose pneumonia. Thus, establishing an autonomous method for identifying pneumonia would be advantageous for treating the condition as soon as possible, especially in distant places. Convolutional Neural Networks (CNNs) have received a lot of interest for illness categorization due to the effectiveness of deep learning algorithms in evaluating medical imagery. Furthermore, features gained by pre-trained CNN models on large-scale datasets of X-ray pictures are extremely effective in image classification tasks. Several Convolutional Neural Networks were seen to categorize x-ray pictures into two groups, pneumonia and non-pneumonia, using various parameters, hyperparameters, and number of convolutional layers modified by the authors. The study analyzes six different models. The first and second models each include two and three convolutional layers. VGG16, VGG19, ResNet50 and Inception-v3 are the other four pre-trained models.

Keywords: Convolutional Neural Networks, Pneumonia detection, medical imaging VGG Net and ResNet

For citation: K.A. Gamri, B.S. Omarov, Bohdan Haidabrus. Comparative analysis of deep learning methods for pneumonia detection on x-ray images // INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES. 2022. Vol. 3. Is. 4. Number 12. Pp. 70–83 (In Russ.). DOI: *10.54309/IJICT.2022.12.4.006.*

РЕНТГЕНДІК СУРЕТТЕ ПНЕВМОНИЯНЫ АНЫҚТАУДЫҢ ТЕРЕҢ ОҚУ ӘДІСТЕРІН САЛЫСТЫРМАЛЫ ТАЛДАУ

Х.А. Гамри[1]*, Б.С. Омаров[2] , Bohdan Haidabrus[3]

Гамри Халед Аднан — "Компьютерлік инженерия" кафедрасының магистранты, Халықаралық ақпараттық технологиялар университеті
Омаров Батырхан Сұлтанұлы — PhD, Халықаралық ақпараттық технологиялар университеті, Математикалық және компьютерлік модельдеу кафедрасының ассистент профессоры. әл-Фараби Қазақ ұлттық университеті "Ақпараттық жүйелер" кафедрасының доценті.
Bohdan Haidabrus — PhD, Канд.Технология туралы.Ғылым Рига Техникалық Университеті, Рига, Латвия

Аннотация. Пневмония-бұл адамның бір немесе екі өкпесіне әсер ететін және көбінесе стрептококк бактериясынан туындаған өлімге әкелетін бактериялық ауру. Дүниежүзілік денсаулық сақтау ұйымының мәліметі бойынша (ДДҰ), әрбір үшін ші өлім себепшісі - пневмония. Тәжірибелі радиотерапевттер пневмонияны диагностикалау үшін қолданылатын кеуде қуысының рентгенографиясын баға-лауы керек. Осылайша, пневмонияны анықтаудың автономды әдісін құру бұл ауруды, әсіресе шалғай жерлерде ерте емдеу үшін тиімді болар еді. Конволю-ция лық нейрондық желілер (CNN) медициналық кескіндерді бағалау кезінде терең оқыту алгоритмдерінің тиімділігіне байланысты ауруларды санаттарға бөлуге үлкен қызығушылық тудырды. Сонымен қатар, ауқымды рентгендік дерек тер жиынында алдын ала дайындалған CNN үлгілері арқылы алынған сипатта малар кескіндерді жіктеу тапсырмаларында өте тиімді. Бірнеше конволю циялық нейрондық желі әртүрлі параметрлерді, гиперпараметрлерді және авторлар ұсынған конволюциялық қабаттардың санын қолдана отырып, рентген сәулелерін пневмония бар және пневмония жоқ деп екі топқа жіктейтіні байқалды. Ұсынылып отырған зерттеу пневмонияны анықтау мәселесі үшін алты түрлі терең оқыту моделін талдайды. Бірінші және екінші модельдер әрқайсысы екі және үш конволюциялық қабаттарды қамтиды. Сонымен қатар, алдын-ала машықтандырылған VGG16, VGG19, ResNet50 және Inception-v3 модельдері қарасырылады.

Түйін сөздер: конволюционды нейрондық желілер, пневмонияны анықтау, медициналық бейнелеу, VGG Net және ResNet

Дәйексөз үшін: Х.А. Гамри, Б.С. Омаров, Bohdan Haidabrus. Рентгендік суретте пневмонияны анықтаудың терең оқу әдістерін салыстырмалы талдау // ХАЛЫҚАРАЛЫҚАҚПАРАТТЫҚ-КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ. 2022. Том. 3. Is. 4. Нөмірі 12. 70–83 бет (орыс тілінде). DOI: 10.54309/IJICT.2022.12.4.006.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ ВЫЯВЛЕНИЯ ПНЕВМОНИИ НА РЕНТГЕНОВСКИХ ИЗОБРАЖЕНИЯХ

Х.А. Гамри[1]*, Б.С. Омаров[2] , Bohdan Haidabrus[3]

Гамри Халед Аднан — магистрант кафедры «Компьютерная инженерия», Международный университет информационных технологий
Омаров Батырхан Султанович — PhD, Ассистент профессор кафедры МКМ, Международного университета информационных технологии. Доцент кафедры «Информационные системы», Казахский национальный университет им. аль-Фараби
Bohdan Haidabrus — Кандидат технических наук, Рижский технический университет, Рига, Латвия

Аннотация. Пневмония является потенциально смертельным бактериальным заболеванием, которое поражает одно или оба легких человека и часто вызывается бактерией Streptococcus pneumoniae. По данным Всемирной организации здраво охранения, на пневмонию приходится каждый третий смертельный исход в Индии (ВОЗ). Опытные радиотерапевты должны оценивать рентген груд ной клетки, используемый для диагностики пневмонии. Таким образом, создание автономного метода выявления пневмонии было бы выгодно для скорейшего лечения заболевания, особенно в отдаленных районах. Сверточные нейронные сети (CNN) вызвали большой интерес для категоризации болезней из-за эффективности алгоритмов глубокого обучения при оценке медицинских изображений. Кроме того, функции, полученные предварительно обученными моделями CNN на крупномасштабных наборах данных рентгеновских снимков, чрезвычайно эффективны в задачах классификации изображений. Было замечено, что несколько сверточных нейронных сетей классифицируют рентгеновские снимки на две группы, пневмонию и не пневмонию, используя различные пара-метры, гиперпараметры и количество сверточных слоев, модифицированных авторами. В исследовании анализируются шесть различных моделей. Каждая из первой и второй моделей включает в себя два и три сверточных слоя. VGG16, VGG19, ResNet50 и Inception-v3 — это четыре другие предварительно обученные модели.

Ключевые слова: сверточные нейронные сети, обнаружение пневмонии, медицинская визуализация, VGG Net и ResNet

Для цитирования: Х.А. Гамри, Б.С. Омаров, Bohdan Haidabrus. Сравнительный анализ методов глубокого обучения для выявления пневмонии на рентгеновских изображениях // МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ. 2022. Том. 3. Is. 4. Номер 12. Стр. 70–83 (на русском языке). DOI: 10.54309/IJICT.2022.12.4.006.

Introduction

1.In recent years, Computer Aided Designs (CAD) have emerged as the dominant study topic in machine learning. Existing CAD systems have previously been shown to help the medical field, especially in the identification of breast cancer, mammography, lung nodules, and so on. Significant characteristics are of the utmost relevance when using Machine Learning (ML) algorithms to medical pictures. As a result, most earlier methods employed hand-crafted features for constructing CAD systems based on image analysis (Dev Kumar Das, 2013). However, the handcrafted features with limits that varied depending on the job were incapable of providing many useful features. Deep Learning (DL) models, notably Convolutional Neural Networks (CNNs), have been used to extract valuable features in picture classification applications (Ali Sharif Razavian, 2014). This feature-extraction procedure necessitates transfer learning approaches, in which pre-trained CNN models learn generic features on large-scale datasets like ImageNet, which are then transferred to the needed job. The availability of pre-trained CNN models such as AlexNet, VGGNet, Inception, ResNet, and DenseNet (Alex Krizhevsky, 2012; Karen Simonyan, 2014; Chollet, 2016; Kaiming, 2016; Gao Huang, 2017) greatly facilitates the method of important feature extraction. Furthermore, classification using high-rich extracted features improves picture classification performance (Heba Mohsen, 2017).

For performance comparison, VGG16, VGG19, ResNet50, and Inception-v3 were used. One of the most difficult challenges in constructing deep networks is vanishing gradient. During back propagation in the vanishing gradient issue, the gradients become infinitesimally tiny, resulting in the loss of integral information. As a result, the network's accuracy reaches a plateau and subsequently begins to deteriorate. The models utilized in this article used several strategies to tackle the vanishing gradient problem. Deep network training has several constraints, including the need for a big dataset, the usage of a significant number of computational resources to achieve high performance, and the tedious process of fine-tuning each parameter and hyper-parameter to produce the best results.

The following sections comprise the research paper:

Section 1 presents the topic and discusses its significance. Section 2 describes the work's approach, including the architecture of a basic CNN model and the specific models provided in this research. This section also goes over the dataset that was used to train and evaluate the six models. Section 3 displays the outcomes of each model, and Section 4 ends the investigation. Section 5 contains a list of references.

Methodology

The article includes several phases, beginning with the import of the dataset from Kaggle. The dataset had been preprocessed. Following that, the dataset was partitioned into train and test sets, each of 5216 and 624 pictures. The six models, each with a distinct architecture, were trained using the training dataset (Szegedy et al., 2017). Each model was trained for 20 epochs, using 32 and one training and testing batch sizes, respectively. The validation accuracy of models 1, 2, VGG16, VGG19, ResNet50, and Inception-v3 was calculated after training and testing. In Figure 1 a pneumonia deep learning-based screening structure is introduced, where the program uses a deep learning algorithm to predict whether the images of the lung of the suspected patient are normal or affected with pneumonia.

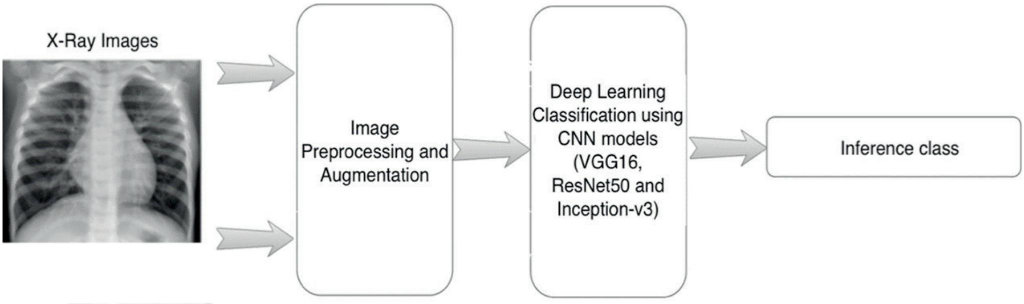

Figure 1 – “Medical image prediction framework”

Dataset

The dataset was obtained from Kaggle and is titled as "Chest X-Ray Images (Pneumonia)". This 1.16 GB dataset comprises 5216 training photos and 624 testing images. The images in this collection are grayscale and 64 X 64 in size. The dataset contains three types of images: normal, bacterial, and viral pneumonia (Mooney, 2018) and samples of it are depicted on Figure 2:

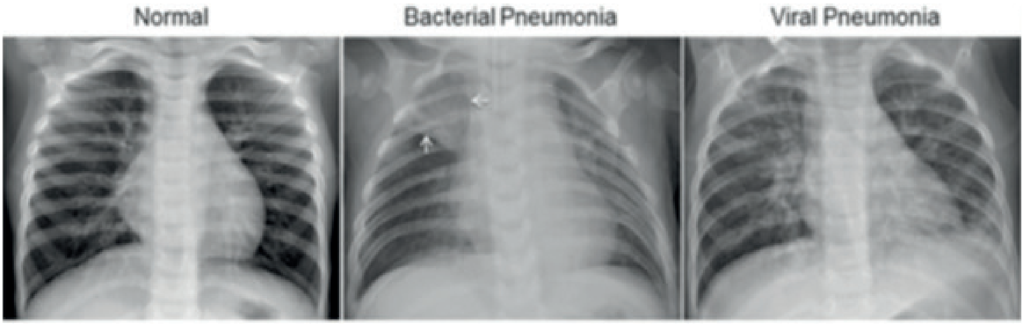

Figure 2 – “Displays the "Chest X-Ray dataset (pneumonia)," which includes three types of images: normal, bacterial, and viral”

Overview of deep learning models for pneumonia detection

CNN Architecture

Figure 3 depicts CNN, a feed-forward neural network. It has four processing layers: the convolutional layer, the pooling layer, the flattening layer, and the fully connected layer (Albarqouni et al., 2016). The subheadings below provide a full discussion of each layer in the CNN architecture.

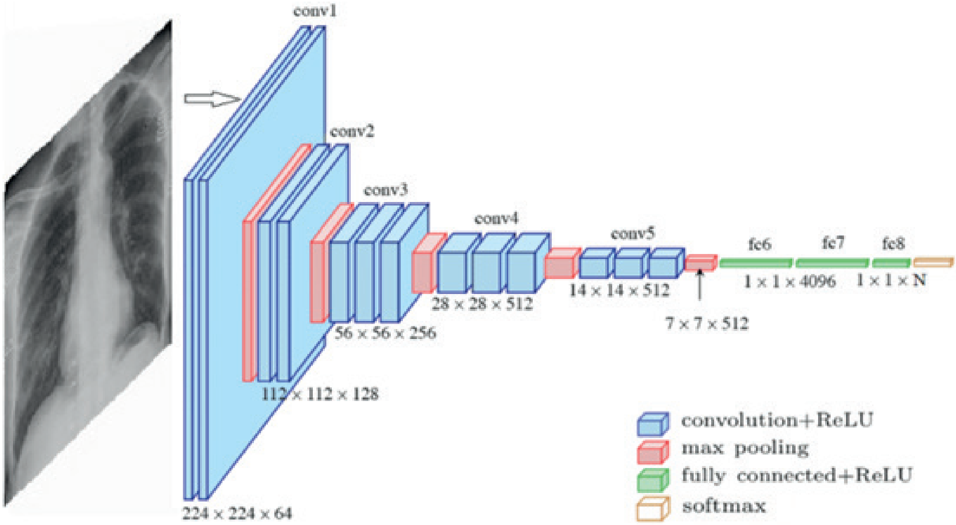

Figure 3 – "CNN architecture"

Convolutional layer

The incoming picture is transformed into a matrix. The convolution process is performed between the input matrix and a 3X3 feature detector/filter/kernel, yielding a feature map (Xu et al., 2015). This procedure decreases the image's size, making it easier to process. This also results in information loss, but the feature detector retains the essential components of the image (Rubin et al., 2018). Multiple feature detectors are used to the input matrix to create our first convolutional layer, which is a layer of feature maps. This layer is subjected to further pooling and flattening before being input into the fully-connected layer.

Activation functions

All six models employed two separate activation functions. These are the activation functions for ReLU and softmax. The most common activation function is the ReLU function. The rectified linear function is a linear function that is applied to the feature map-based convolutional layer. If the input is positive, the ReLU function returns one. If not, the output is zero (Alex Krizhevsky, 2012). Because the ReLU function avoids and corrects the vanishing gradient problem, neural network models that employ it are easier to train and perform better than models that use other activation functions such as sigmoid or hyperbolic tangent activation functions. The ReLU function (Kingma, 2014) is denoted by f(x), as shown in Equation 1.

$$f(x)=max(0,x) \tag{1}$$

Softmax is another extensively used activation function. The Softmax function converts inputs or logits to a probability distribution. The sum of the distribution's output probabilities equals one. Logits are the outputs of the network's logit layer or final layer.

These are unprocessed forecast values ranging from minus infinity to infinity. The cost function that is most employed with softmax is categorical cross entropy. In all six models, the Softmax activation function was applied.

Pooling layer

The pooling layer's aim is to further downsample the input picture. To put it another way, to lower the size of the input image (Karen Simonyan, 2014). The number of picture parameters is minimized, lowering computing complexity. Maxpooling and average-pooling are the sub-sampling techniques utilized in the models. Max-pooling is a discretization method that uses samples. The 2X2 pooling layer operates over each feature map, scaling its dimensionality with the 'MAX' function. Maxpooling chooses the highest pixel value from the image window currently covered by the feature detector (Xu et al., 2015). Max pooling assists models in recognizing the image's key elements.

Another subsampling approach is average pooling, which computes the average value from the window of the picture now covered by the feature detector. Max-pooling is beneficial for recognizing picture salient characteristics, but average pooling allows the neural network to identify the entire image. In comparison to max-pooling, the average pooling strategy preserves more information.

Fully connected layers

The combined feature map is smoothed out into a column before being sent into the neural network (Cireşan et al., 2011). This allows the neural network to simply handle the produced feature maps. The input picture is sent into the fully connected layer after passing through the convolutional and pooling layers and the flattening layer. Input forward propagates as weights are calculated. A prediction is made by the network. We generate a cost function based on the prediction, which in this case is categorical crossentropy. The cost function indicates how effectively a network performs. Back propagation, weight tweaking, and feature mapping are used to improve the network once the cost function is calculated.

This forward and back propagation process continues until the network is fully optimized. The Adam optimizer was utilized in all six models (Shin et al., 2016). Adam is a search engine optimization method. The Adam optimizer is used to repeatedly adjust the network weights depending on the training data. The Adam optimizer is excellent for networks, training on huge datasets or parameters, and it is simple to construct, computationally efficient, and requires little memory.

Reducing overfitting

Dropout was used to decrease overfitting in models 2, VGG16, VGG19, and ResNet. With a probability of 0.5, the dropout mechanism resets the output of each hidden neuron to zero. Neurons with 0 initialization do not participate in forward or backward propagation (Baldi et al.). As a result, complicated coadaptations of neurons are reduced since each neuron must accomplish something meaningful without relying on other neurons in the same layer. As a result, neurons are pushed to acquire several prominent traits that are helpful when combined. Another method for reducing overfitting is data augmentation. To avoid overfitting, the learning rate of the models was also adjusted.

The learning rate is a hyper-parameter that governs the amount to which the network's weights are adjusted in relation to the loss gradient. This hyper-parameter influences the network's ability to converge to some local minima.

Model architecture

Model 1

The first trained model consists of two convolutional layers: the first has 32 feature maps that use the ReLU function, and the second has 64 feature maps that use the ReLU function. Following each convolutional layer, max-pooling layers of 2X2 dimensions are employed. Behind these layers is a flattening layer. There are two dense layers used: one with 256 output perceptrons and the other with two output perceptrons with the softmax function. The learning rate has been decreased to 0.001. The Adam optimizer was employed as the cost function, with categorical cross-entropy.

Model 2

The second model is made up of three convolutional layers: the first convolutional layer has 32 feature maps that use ReLU, the second convolutional layer has 64 feature maps that use ReLU, and the third convolutional layer has 128 feature maps that use ReLU. Following each convolutional layer, max pooling layers of 2X2 dimensions are employed. The first dense layer contains 256 output perceptrons using ReLU and the second dense layer has two output perceptrons using softmax function. A dropout layer is also included. The model's learning rate is decreased to 0.0001. The Adam optimizer was employed as the cost function, with categorical cross-entropy.

VGG16 and VGG19

VGG16 is a CNN model created by K. Simonyan and A. Zisserman. It was one of the most prominent models entered in the 2014 ILSVRC competition. On the ImageNet dataset, this model achieves 92.7 % top-5 test accuracy. In all, the network contains 16 layers. VGG16 added numerous 3X3 kernel-sized filters one after the other, replacing the previous models' enormous kernel-sized filters. The depth of the neural network increases as the number of kernel layers increases. This allows the neural network to identify and grasp more complicated features and patterns. Vgg16 is made up of 3x3 convolutional layers, 2x2 average-pooling layers, and fully linked layers. The neural network's starting width is 64. After each pooling layer, the neural network's diameter doubles. Each of the first two completely linked levels contains 256 channels, while the third layer has two channels. The first two hidden layers use the ReLU activation function, whereas the last layer uses the softmax activation function. After each 256 channel dense layer, dropout was applied. The network's learning rate is 0.0001. The Adam optimizer was employed as the cost function, with categorical cross-entropy. The representational depth of VGG16 is advantageous for classification accuracy.

Vgg19, a variation of VGG16, is a 19-layer convolutional neural network that is mostly used for image categorization. Its basic architecture is identical to VGG16's. The sole variation in VGG19 is the use of two thick layers with 256 channels and a learning rate of 0.00001.

ResNet50

ResNet is an image classification algorithm that stands for residual network. ResNet from Microsoft obtained a top 5 error rate of 3.57 % on the ImageNet dataset and won the ILSVRC classification challenge in 2015 (He et al., 2016). Convolutional layers in the network contain 3X3 filters, and downsampling is done directly by convolutional layers with a stride of 2. The network's final layer is a fully linked layer with 256 and two channels that use ReLU and softmax activation functions, respectively. The network's learning rate is 0.000001. The Adam optimizer was employed as the cost function, with categorical cross-entropy. ResNet employs shortcut connections to address the issues of degraded accuracy and vanishing gradient that arise in deep neural networks. These connections allow the network to bypass levels that it deems unnecessary for training. This decreases training error and allows the network to converge faster than other networks. Figure 4 demonstrates how shortcut connections function in the ResNet50 model.

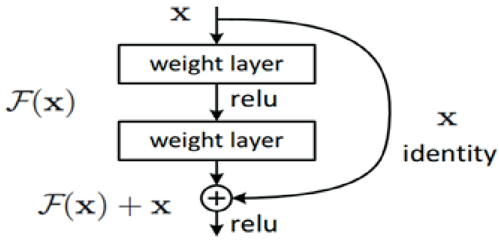

Figure 4 – "Shortcut connection"

Inception-v3

Figure 5 depicts Inception v3, a convolutional neural network used for image categorization. Inception v3 is a 42-layer CNN. It has several variations, including inceptionv1/google net, inceptionv2, and inceptionv4. Inception v1 finished first in the ILSVRC 2015 competition. GoogleNet/inception v1 was released in 2015, and with each subsequent version, more features were added. Inception v1 introduced auxiliary classifiers. Auxiliary classifiers were implemented to avoid or prevent each layer's activation from descending to zero. Inception v2 introduced batch normalization. By decreasing the internal covariate shift, this approach corrects the problem of vanishing gradients and zero activations.

Additional factorization was originally utilized in Inception v3 to minimize the amount of network connections/parameters while maintaining network performance. The network's learning rate is 0.000001. The Adam optimizer was employed as the cost function, with categorical cross-entropy. The fundamental architecture of the inception-v3 network is seen in Figure 5.

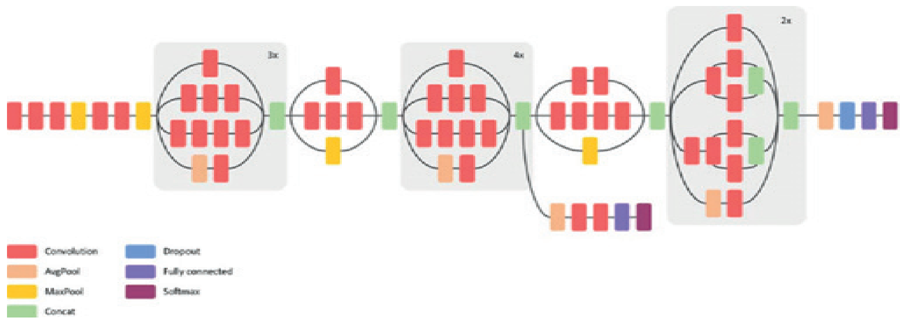

Figure 5 – "Inception-v3 architecture"

Results

Evaluation Criteria

Six models were trained and evaluated using the Chest X-Ray Images (Pneumonia) dataset, which included 5216 images for training and 624 images for testing. The identical data preprocessing approach was utilized for all six models. Accuracy, recall, and F1 are performance indicators used to examine and select the top performing models. Choosing an acceptable performance metric for a classification assignment is a significant difficulty. As assessment criteria Accuracy, Recall, and F1 score were used. The accuracy measure is the model's validation or classification accuracy.

The recall is employed as a performance evaluation metric in the identification of patients infected with bacterial pneumonia and viral pneumonia. If an actual positive patient is anticipated to be negative, the outcome might be disastrous for the patient's health. Whereas precision is an excellent metric to analyze scenarios with a high false positive cost. A false positive implies that the model misidentified Chest X-Ray Images that did not show bacterial or viral pneumonia infection as having bacterial or viral pneumonia. Precision refers to how many observations are truly positive and how accurate the model is. If the suggested model's accuracy is low, we may get the erroneous diagnosis. The F1 score performance metric is superior than accuracy and recall because it balances them for the unequal Normal, Bacterial Pneumonia, and Viral Pneumonia class distributions with a high percentage of real negatives.

The Accuracy (Hemanth et al., 2014) is given by Equation 2 as follows:

$$\text{Accuracy} = \frac{\text{t\_pos} + \text{t\_neg}}{\text{t\_pos} + \text{t\_pos} + \text{f\_pos} + \text{f\_neg}} \tag{2}$$

Recall and F1 score are given by Equation 3 and Equation 4:

$$\text{Recall} = \frac{\text{t\_pos}}{\text{t\_pos} + \text{f\_neg}} \tag{3}$$

$$\text{F1} = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}} \tag{4}$$

The terms t pos, t neg, f pos, and f neg in the above formula stand for True Positive, True Negative, False Positive, and False Negative, respectively. The recall is a measurement of the total number of genuine, relevant results returned. When the cost of false negatives is significant, model recall is critical. Sensitivity is another term for recall. The F1 Score is, in general, the harmonic mean of accuracy and recall. If a model's F1 Score is high, it suggests it has less false positives and false negatives. It is calculated as a weighted average of recall and accuracy.

Analysis of CNN Models

For the purpose of simplifying the experimental result assessments, the three classes of normal patients, bacterial pneumonia, and viral pneumonia have been united into one class called infected. As a result, the findings were classified as pneumonia expected and normal. The confusion matrix offers information about the classifier's mistake. It is used to describe the classification model's performance on test pictures when true values are known. It is a summary of the production outcomes. The following are CNN model confusion matrices:

Table 1 – "Confusion matrix of Model 1"

True Label	Predicted label	
	165	69
	23	367

Table 2 – "Confusion matrix of Model 2"

True Label	Predicted label	
	192	42
	6	384

The aforementioned confusion matrices are used to compute the recall and F1 Score of CNN models. Based on the findings obtained while training and testing on the dataset, a comparative analysis of performance metrics of two CNN models is shown below.

Table 3 – "Model 1 and 2 performance comparison"

	Accuracy	Recall	F1 Score
Model 1	85.26 %	94 %	89 %
Model 2	92.31 %	98 %	84 %

Model 1 exhibited 92.52 % training accuracy and 19.33 % training loss, respectively. Model 1's validation accuracy is 85.26 %, but its validation loss is 38.36 %. Similarly, the training accuracy and training losses for Model 2 are 96.30 % and 9.98 %, respectively. Model 2 achieved validation accuracy and validation loss of 92.31 % and 25.23 %, respectively. As a result, Model 2 has outperformed Model 1 since it has a better value for each performance metric. Model 2 is not just a higher performing model; it is also a consistent and efficient model, scoring above 90 % in all three performance categories and having an extremely high recall of 98 %. Model 1 is more prone to overfitting than Model 2.

Confusion matrices of pre trained models are given below:

Table 4 – "Confusion matrix of VGG16 model"

True Label	Predicted label	
	168	66
	14	376

Table 5 – "Confusion matrix of VGG19 model"

True Label	Predicted label	
	182	52
	20	370

Table 6 - “Confusion matrix of Resnet50 model”

True Label	Predicted label	
	104	130
	10	380

Table 7 - “Confusion matrix of Inception-v3 model”

True Label	Predicted label	
	116	118
	63	327

Confusion matrices illustrate the error produced by the classifier models, and it is noticed that Model 2 has a 6.7 % error (the lowest among all), while the error observed in all other models is greater than 10 %, with ResNet50 having the highest error rate of 21 %. Based on the findings obtained during training and testing on the dataset, a comparative study of performance metrics of four pre-trained models (VGG16, VGG19, ResNet50, and Inception-v3) is shown below.

Table 8 - “Performance comparison of pre trained models”

	Accuracy	Recall	F1-score
VGG16	87.18 %	96 %	90 %
VGG19	88.46 %	95 %	91 %
ResNet50	77.56 %	97 %	84 %
Inception-v3	70.99 %	84 %	78 %

Table 9 - “Values of accuracy and loss achieved by each model”

	Training accuracy	Training loss	Validation accuracy	Validation loss
VGG16	95.61 %	12.03 %	87.17 %	37.94 %
VGG19	92.85 %	18.01 %	88.46 %	34.29 %
ResNet50	94.29 %	14.32 %	77.56 %	68.36 %
Inception-v3	88.96 %	28.20 %	70.99 %	97.56 %

ResNet50 and Inception-v3 exhibit significant overfitting due to the big disparity in training and validation accuracy. These two models have a high validation loss and a low validation accuracy or classification accuracy. As a result, these two models perform poorly.

VGG16 and VGG19, on the other hand, exhibit less overfitting. Their validation accuracy is likewise excellent. The above comparison study shows that VGG19 surpasses all other models, having earned the greatest values for classification accuracy and F1 Score. Its recall is lower than VGG16's, but it performs better overall. These four models are deep neural networks with several layers. Given the reduced amount of the dataset used for training and testing, their validation accuracy is lower than that of the CNN models (shallow networks) mentioned above. Deep neural networks are likely to outperform CNN algorithms when greater datasets are utilized (shallow networks).

Conclusion

This study contrasts six high-performance neural networks for real-time applications.

In this study, recall is a significant performance assessor since it is required to reduce the frequency of false negatives in the case of medical imaging. Model 2 recall is as high as 98 %, while VGG19 recall is equally high at 95 %. The f1 scores for the Model 2 and VGG19 networks were 94 % and 91 %, respectively.

Because of their outstanding performance across all performance criteria, the Model 2 and VGG19 models may be efficiently employed by medical officials for early identification of pneumonia in both children and adults. A huge number of x-ray pictures may be analyzed fast to offer extremely precise diagnostic results, allowing healthcare systems to deliver more efficient patient care and lower death rates.

To improve the classification accuracy of all models, it is planned to fine-tune each parameter and

hyper-parameter in the future (Rajpurkar et al., 2017) proposed the ChexNet model, a fast and accurate model suitable for real-time applications. The models given in this study may be expanded to classify various illnesses with great accuracy, like CheXNet did. The models' overall performance can be enhanced by using more datasets.

REFERENCES

Ali Sharif Razavian, Hossein Azizpour, Josephine Sullivan and Stefan Carlsson (2014). "CNN features off-the-shelf: an astounding baseline for recognition", Proceedings of the IEEE conference on computer vision and pattern recognition workshops. — Pp. 806813. — 2014.

Alex Krizhevsky, Ilya Sutskever and Geoffrey E Hinton (2012). "Imagenet classification with deep convolutional neural networks", Advances in neural information processing systems. — Pp. 10971105. — 2012.

Albarqouni S., Baur C., Achilles F., Belagiannis V., Demirci S. & Navab N. (2016). Aggnet: deep learning from crowds for mitosis detection in breast cancer histology images. IEEE transactions on medical imaging, 35(5). — Pp. 1313–1321.

Chollet F. (2016). "Xception: deep learning with separable convolutions".

Cireşan D.C., Meier U., Masci J., Gambardella L.M. & Schmidhuber J. (2011). High-performance neural networks for visual object classification. arXiv preprint arXiv:1102.0183.

Dev Kumar Das, Madhumala Ghosh, Mallika Pal, Asok K Maiti and Chandan Chakraborty (2013). "Machine learning approach for automated screening of malaria parasite using light microscopic images", Micron. — Vol. 45. — 2013.

Heba Mohsen, El-Sayed A El-Dahshan, El-Sayed M El-Horbaty and AbdelBadeeh M Salem, "Classification using deep learning neural networks for brain tumors", Future Computing and Informatics Journal, 2017Gao Huang, Zhuang Liu, Laurens Van Der Maaten and Kilian Q Weinberger, "Densely Connected Convolutional Networks", CVPR. — Vol. 1. — Pp. 3. — 2017.

He K., Zhang X., Ren S. & Sun J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition. — Pp. 770–778.

Hemanth D.J., Vijila C.K.S., Selvakumar A.I. & Anitha J. (2014). Performance improved iterationfree artificial neural networks for abnormal magnetic resonance brain image classification. Neurocomputing, — 130. — Pp. 98–107.

Karen Simonyan and Andrew Zisserman (20). 1 "Very deep convolutional networks for large-scale image recognition". — 2014.

Kingma D.P. & Ba J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun (2016). "Deep residual learning for image recognition", Proceedings of the IEEE conference on computer vision and pattern recognition. — Pp. 770778. — 2016.

Mooney P.T. (2018). Chest X-Ray Images (Pneumonia). Kaggle. Retrieved from *https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia*

Szegedy C., Ioffe S., Vanhoucke V. & Alemi A.A. (2017, February). Inception-v4, inception-resnet and the impact of residual connections on learning. In Thirty-First AAAI Conference on Artificial Intelligence.

Shin H.C., Roberts K., Lu L., Demner-Fushman D., Yao J. & Summers R.M. (2016). Learning to read chest x-rays: Recurrent neural cascade model for automated image annotation. In Proceedings of the IEEE conference on computer vision and pattern recognition. — Pp. 2497–2506.

Rubin J., Sanghavi D., Zhao C., Lee K., Qadir A. & Xu-Wilson M. (2018). Large scale automated reading of frontal and lateral chest x-rays using dual convolutional neural networks. arXiv preprint arXiv:1804.07839.

Rajpurkar P., Irvin J., Zhu K., Yang B., Mehta H., Duan T. & Lungren M.P. (2017). Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225.

Xu Y., Jia Z., Ai Y., Zhang F., Lai M., Eric I. & Chang C. (2015, April). Deep convolutional activation features for large scale brain tumor histopathology image classification and segmentation. In 2015 international conference on acoustics, speech and signal processing (ICASSP). — Pp. 947–951.

ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ

МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES

Правила оформления статьи для публикации в журнале на сайте:
https://journal.iitu.edu.kz
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Собственник: АО «Международный университет информационных технологий» (Казахстан, Алматы)

ОТВЕТСТВЕННЫЙ РЕДАКТОР
Ералы Диана Русланқызы

КОМПЬЮТЕРНАЯ ВЕРСТКА
Жадыранова Гульнур Даутбековна

Подписано в печать 15.12.2022.
Формат 60х881/8. Бумага офсетная. Печать - ризограф.7,0 п.л. Тираж 100
050040 г. Алматы, ул. Манаса 34/1, каб. 709, тел: +7 (727) 244-51-09.

Издание Международного университета информационных технологий
Издательский центр КБТУ, Алматы, ул. Толе би, 59