Int J Syst Assur Eng Manag
https://doi.org/10.1007/s13198-022-01788-x

ORIGINAL ARTICLE

A deep neural architecture for SOTA pneumonia detection from chest X-rays

Sravani Nalluri[1] · **R. Sasikala**[1]

Received: 21 May 2022 / Revised: 24 September 2022 / Accepted: 30 September 2022

Abstract Pneumonia among children is a leading cause of death in India, and it gains a lot of researchers' attention to develop early detection tools. Due to a lack of the number of radiologists, especially in rural India, the traditional method of diagnosing pneumonia does not address the real-time issues related to early stages. This paper presents a deep learning model, NASNet (Neural Architecture Search Network), pre-trained on ImageNet to predict pneumonia very early stage through chest x-rays of patients. With 2.6 million trainable parameters, the proposed model can run even on a mobile phone with good precision, recall, and an F1 score to detect pneumonia. This approach thus proves to be significantly better than the current state-of-the-art models. It can also help trained radiologists to get a second opinion/ validation of pneumonia diagnosis.

Keywords Deep learning · Pneumonia · Chest X-Ray · NASNet

1 Introduction

Pneumonia is an intense respiratory infection that causes irritation or liquid in the lungs. It is caused by various infectious agents, including viruses, fungi, and bacteria. Two of the most widely recognized reasons for severe bacterial pneumonia are Haemophilus influenzae type b and Streptococcus pneumonia. Pneumococcus and Hib are the primary sources of severe bacterial pneumonia in youngsters, together representing roughly 60% of pneumonia deaths in kids younger than five years. In India, it is assessed that in 2013 pneumococcus caused more than 60,000 pneumonia deaths, and Hib caused more than 35,000 pneumonia deaths. In addition to this high mortality burden, these two bacteria are responsible for a large number of instances of pneumonia and a considerable number of cases of crippling meningitis in kids each year (Walker et al. 2013; O'brien, K. L., Wolfson, L. J., Watt, J. P., Henkle, E., Deloria-Knoll, M., McCall, N. & Cherian, T. 2009; Watt et al. 2009). Arising microbes like Burkholderiapseudomallei are progressively perceived as a significant reason for Community-Acquired Pneumonia (CAP) in Southeast Asian countries (Currie et al. 2001).

Pneumonia continued to be the primary source of fatality in children worldwide, with India representing 20% of deaths and a higher possibility of childhood pneumonia as compared to other countries (McAllister et al. 2019; Simonyan and Zisserman 2014). India adds to roughly 23% of worldwide and 36 percent of World Health Organisation's pneumonia patients under five years (Sharma et al. 2013). Reliable evaluations of the illness burden are not accessible, especially for the adult populace. The meagre information for grown-ups comes from tertiary consideration showing medical clinics utilizing cross-sectional studies (Guan et al. 2018). A study from Mumbai discovered that Streptococcus pneumonia and Gram-negative bacteria were more common in serious pneumonia, and that 19% of all patients had extreme CAP (SCAP) (Dagaonkar et al. 2012). A new study highlights the significance of pneumococci in the obtrusive pneumococcal infections in India.

Hib antibodies are currently being utilized in more than 185 nations worldwide (Loharikar et al. 2016). The illness has been practically disposed of in industrialized nations, where immunization has been for more than 20 years (Toğaçar et al. 2020; Akter, et al. 2021). In non-industrial

✉ R. Sasikala
sasikala.ra@vit.ac.in

[1] VIT University School of Computer Science and Engineering, Vellore, Tamil Nadu, India

Published online: 03 November 2022

countries that have brought the antibody into their public inoculation programs, Hib pneumonia and meningitis have been radically reduced (Watt et al. 2009). Pneumococcal antibodies are being utilized in more than 115 countries (Record 2006). There has been a practically complete end of pneumococcal infection in nations where pneumococcal immunizations have been presented. Immunization has also been displayed to diminish the sickness among unvaccinated children and adults in a populace – an impact of "herd immunity" (Feikin, D. R., Kagucia, E. W., Loo, J. D., Link-Gelles, R., Puhan, M. A., Cherian, T., … & Serotype Replacement Study Group 2013; Davis et al. 2013). In 2011, the Government of India presented the Hib-containing pentavalent antibody in a staged way. The pentavalent immunization gives assurance against five infections: diphtheria, lockjaw, pertussis, hepatitis B, and Hib. Somewhere in the range of 2011 and 2013, the antibody was presented in 14 states and association regions. In 2013, the National Technical Advisory Group on Immunization suggested the public scale-up of the vaccine.

Community-Acquired Pneumonia, CAP, is the primary source of mortality and dreariness with considerable clinical and financial effects. Albeit a few pathogens are believed to cause the infection, information on the pathogen circulation is not consistently addressed across the nations. A few factors like topographical locale, age, and study period impact the frequency of CAP in adults. Nonetheless, solid and reliable information over a long period is accessible from a few nations. Reports suggest that almost 2.4 million deaths happen among all ages because of Lower Respiratory Tract Infections (LRTIs) (Walker et al. 2013). Sub-Saharan Africa, Southeast Asia, and South Asia have recorded higher casualties. In 2016, 197.05 million cases of pneumococcal pneumonia were accounted for worldwide, resulting in the primary source of LRTI related morbidity and mortality.

Universally, mortality because of LRTI remained unchanged from 2005 to 2015, even though age-normalized death rates fell by 19.5 percent (Walker et al. 2013). Consistent expansion in the hospitalization rates, including Intensive Care Units (ICU) because of Community-Acquired Pneumonia, particularly in aged people, is observed (http:, , www.who.int, maternal_child_adolescent, epidemiology, gappd-monitoring, en, . xxxx; Nalluri and Sasikala 2020). The case casualty rate goes from 2 to 20 percent, up to 50 percent in patients admitted to ICUs, and differs between medical care settings, geological area, patient classes, and age (Rajpurkar et al. 2017). This account survey centers around the bacterial CAP in immune-competent adults with particular importance on existing modalities and holes in diagnostics, ideal usage of testing systems, and individualized treatment choices with attention to Indian situations.

1.1 Background

Researchers have been working on predicting a patient's pneumonia using machine learning and deep learning techniques on CT scans, X-Rays, and Vocal data. Some studies have approached the solution with a machine learning model, whereas others have trained deep learning models on massive datasets. As a result of this extensive research, significant progress has been made in predicting pneumonia. In this paper, we continue the research to predict pneumonia more efficiently and reliably.

M. Jamshidi et al. (Xu et al. 2015) present AI-based techniques to detect COVID-19 using various approaches such as Extreme Machine Learning (ELM), Long/Short Term Memory (LSTM), including Generative Adversarial Networks (GANs). Research has been done to compare various approaches to predict COVID-19 infection (Anthimopoulos et al. 2016). The models presented by Alakus, T. B., & Turkoglu (Anthimopoulos et al. 2016) achieve a precision of 86.75% and an accuracy of 86.66% to predict the likelihood of COVID-19 in a patient. The models were tested on 18 laboratory findings from 600 patients.

Yujin Oh, Sangjoon Park, and Jong Chul Ye (He et al. 2016) presents an interesting approach to predict COVID-19 with CXR using limited training datasets. Their proposed method is influenced by analyzing potential imaging biological markers of CXR radiographs. (He et al. 2016) It uses segmented mask for data pre-processing with the pre-trained model on ResNet-18, achieving an of 93.2%. Daniel G. Pankratz et al. (Pankratz et al. 2017) use a Logistic Regression model to identify Usual Interstitial Pneumonia by training the model with parameters such as age, smoking history, etc. Exome-enriched RNA sequencing was performed on 233 TBBs. The model is trained to classify a given record as UIP or Non-UIP (Pankratz et al. 2017). This study was done on only 84 subjects and tested on 31 test subjects and achieved an AUC of 0.86 with a specificity of 86% and sensitivity of 63%. Alakus and Turkoglu 2020 utilized scan of the multinational dataset to predict COVID-19 pneumonia using 3D models and hybrid 3D models on 1387 scans achieving an accuracy of 90% with a specificity of 84%. Devulapalli, et al. 2021; Oh et al. Aug. 2020 presents a custom deep learning model to differentiate between CAP and other lung ailments. The model, COVNet, trained on 4352 CT scans from 3322 patients and achieves a specificity of 90%.

SARS-Net, a new model with convolutional layers and inception blocks with one of its variations utilizing a convolutional graph network, has been introduced in Panwar, et al. (2020). SARS-Net aims to detect COVID-19 and is trained on the COVIDx dataset consisting of only 13,975 CXR images. They utilized 90% of the dataset for training

and the rest for testing, achieving an accuracy of 97.6% and sensitivity of 92.9%.

Pranav Rajpurkar et al. designed a convolutional neural network that can detect pneumonia and 14 other lungs related diseases called CheXNet. It is a 121-layer convolutional neural network trained on the ChestX-ray14 dataset. This dataset contains little more than 100,000 images of Chest X-Ray frontal-view with 14 different types of lung diseases, four Stanford radiologists annotated a test set with which the performance of CheXNet was evaluated, and this exceeded the average radiologist performance of the F1-score metric. This model performed better than the models proposed respectively by Wahl et al. (2020); Zhao et al. 2020).

A group of researchers from India (Loharikar et al. 2016) used various combinations of CNNs for feature extraction and machine learning models for classification on a set of 112,000 Chest X-Rays. The images were resized to 224×224, and the combination of DenseNet169 with the Support Vector Machine for classification turned out to be the best within the scope of their research. They achieved an AUC of 80%.

Reviewing many research papers on pneumonia detection methods (Record 2006) with artificial intelligence revealed that Chest X-Rays are subjected to highly complex processing using edge detection, Gaussian filters, and morphological operations to detect meaningful features. Such features are used in various classification models, such as Support Vector Machines, ResNet, and Random Forest. Many have tried modern deep learning models like DenseNet to achieve decent results. But due to the high complexity of models because of the large number of parameters, training such models on massive datasets may be impractical. Performance metrics such as precision and recall are critical compared to the model's accuracy due to the dire consequences of false negatives (http:, , www.who.int, maternal_child_adolescent, epidemiology, gappd-monitoring, en, . xxxx).

With a risk factor-based modeling approach, The Lancet Child and Adolescent Health (Wahl et al. 2020) introduces a detailed assessment of state-explicit pneumonia rates in kids in India. The authors assessed the change in pneumonia over the long run by figuring the impact of fleeting changes in the pervasiveness of notable pneumonia hazard factors like hunger, deficient vaccination, and openness to indoor air contamination on the rate.

A model to detect thorax diseases with lateral and frontal chest X-rays was modelled by Rubin et al. (Rubin et al. 2018). For the recognition of images on a large scale, MIMIC-CXR data is used. The dataset was split into training, testing, and validation sets as 70%, 20%, and 10%, respectively. To improve the overall performance of model, data augmentation was used, which made the model robust. Their DualNetConvNet model achieved average AUC (Area Under Curve) of 0.721 and 0.6882 for PA and AP, respectively.

To classify pulmonary tuberculosis, a deep CNN model was developed by Lakhani et al. (Kaushik et al. 2020). To classify chest X-rays, they used transfer learning models of AlexNet and GoogleNet. The dataset was split into training, testing, and validation sets as 70%, 15%, and 15%, respectively. With pre-processing and data Augmentation, they achieved an AUC of 0.99. The model achieved a precision of 100% and a recall of 97.3%.

Xu et al. (Devulapalli, et al. 2021) developed a deep learning CNN model for classifying and dividing cerebrum tumor MRIs. They employed various approaches such as feature selection and data augmentation and achieved a classification accuracy of 89.7%—Guan et al. (Guan et al. 2018) designed an AG-CNN model to recognize thorax ailments. The chest X-ray14 dataset was used to detect thorax disease from chest X-ray images. Attention-guided global and local branch CNN was used for classification purposes. An AUC of 0.868 was achieved by them and was better than other models.

Anthimopoulos et al. (Devulapalli and Krishnan 2021) developed a convolution neural network with five convolution layers to detect lung disease patterns in a dataset with 15,000 images belonging to 7 different types of diseases. Contracting many models, they used average pooling, leaky ReLU, and three fully connected layers. This model accomplished a classification-accuracy of nearly 85.1%. Another approach has been presented by Panwar, Harsh, et al. (Zhao et al. 2020) which detects COVID-19 using nConvNet(Krizhevsky et al. 2012). They use the data augmentation techniques such as rotation range and vertical flip along with RGB re-ordering to generate augmented data. With this approach, they achieved an error rate of 21.43% and an overall accuracy of 78.57%.

The above Related work discussed various deep learning architectures on different datasets. However, the best model with specific domain skills can be done by training optimized architecture search strategies. This paper proposes a novel approach to detect pneumonia using various deep learning models. We present that NASNet achieves state-of-the-art results in terms of precision and recall and, at the same time, is also very efficient regarding computing power.

2 Methodology

Pneumonia is a complex problem to solve because many other diseases resemble pneumonia. A machine learning model needs a well-processed and quality dataset to learn and understand very minute details of data. To properly distinguish between pneumonia and non-pneumonia, this paper uses a publicly available dataset released by Wang

et al. (Nalluri and Sasikala 2020), which includes images of 14 different diseases similar to pneumonia. As mentioned in the pre-processing data section, multiple pre-processing steps are performed before using it to train the model.

Figure 6 describes that the input image will be fed to the NASNET model for matching with the trained image data available in the search space. Further, the search strategy will be selected based on the precision of the deep architecture utilized in the training process. The best result will be selected as the predicted outcome of the query image. Various models trained for NASNET have been discussed below.

2.1 Dataset description

This dataset contains 112,200 Chest X Rays of 30,800 patients. Each record in this dataset is labeled with 14 types of thoracic disease labels. For our specific task of identifying pneumonia, we've labeled all pneumonia-positive images as one and the rest as 0.

We split the dataset between training, testing, and validation data so that each contains an equal number of positive and negative pneumonia samples. This ensured that the model didn't output wrong predictions because of skewed data. Since the dataset is skewed, we also used weights to output while training the model.

The training dataset contains data from 27,833 patients with 97,748 images, whereas validation data includes 7442 images of 1763 patients and the test dataset contains 531 images from 400 patients. We made sure that there was no overlap in training, testing, and validation datasets.

2.2 Data pre-processing

We re-scaled the images (Figs. 1, 2) to have a dimension of 224 × 224x3. The data is normalized with min–max normalization, as shown in Eq. (1)Mean and standard deviation

Fig. 1 Normal chest X-ray

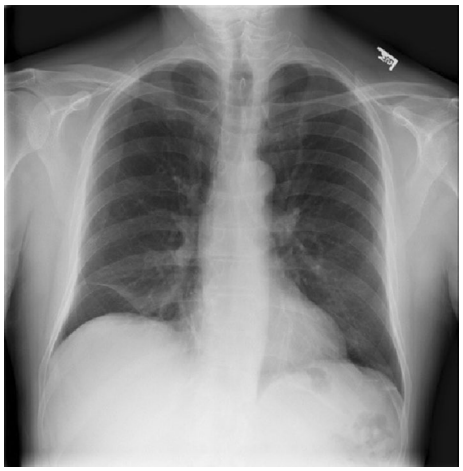


Fig. 2 Pneumonia chest X-Ray

data in ImageNet are used to standardize the data as shown in Eq. (2). To increase the data size of pneumonia cases, data augmentation, such as horizontal and vertical flip wear, was applied. Augmentation doubled the true positives of data and thus decreased the skewness of data.

$$x = \frac{x - min}{max - min} \tag{1}$$

$$x = \frac{(x - \mu)}{\sigma} \tag{2}$$

where.

min: 0, the minimum value of pixel in a specific dimension of image in the entire dataset.

max: 255, the max value of a pixel in a specific dimension of image in the entire dataset.

μ: mean of ImageNet.

σ: standard deviation of ImageNet.

μ: Means[R,G,B]: [0.485, 0.456, 0.406].

σ: standard-deviations[R,G,B]: [0.229, 0.224, 0.225].

Figure 1 is the normal chest X-ray image and Fig. 2 is the Pneumonia chest X-ray. The Fig. 2 has more fluid in the image around the lungs.

3 CNN architecture

Convolutional Neural Networks are feedforward neural networks with multiple convolutional, activation, pooling, flattening, and fully-connected layers. In CNN, the input is read as a matrix or an array. A Convolutional filter is applied on the matrix, which slides over the input, performs element-wise multiplication, and stores the sum in another array called feature-map. Usually, A 3xN filter is applied over the

input of shape AxBxN, resulting in a 2D feature map. N represents the number of channels in the input; for example, RGB image has three channels. Similarly, multiple feature maps are created by applying different filters. All of them are stacked together, resulting in a K-dimensional feature map where K is the number of filters involved.

3.1 Convolution layer

The first ever convolutional neural network was used by Yann LeCunn, the inventor of CNNs, on images to detect zip codes. In 2012, in the ImageNet competition ILSVRC, Krizhevsky et al. (Toğaçar et al. 2020) used a five-layered CNN model followed by max-pooling and fully connected layers. This architecture had 60 million parameters and used dropout to prevent overfitting and make the model robust. This model achieved a top-five error rate of 17% in 2014. This paper also uses convolutional layers with multiple filter sizes to detect minute data details. A highly accurate model with minimal filter size acquiring accuracy of 92.6% was developed by Simoyan et al. (Akter, et al. 2021).

3.2 4.2 Activation function

Activation functions are used to make the model learn non-linear patterns in the data and hence become very important in deep learning. In this paper, we make the use of ReLU (Rectified.

Linear Unit (Rubin et al. 2018) and softmax activation function. Mathematically, ReLU function is stated as below Eq. (3).

$$f(x) = max(0, x) \tag{3}$$

3.3 Rectified liner unit (ReLU)

In the last dense layer of the models, activation function "sigmoid" Eq. (4) is used as presented in this paper. A probability distribution function is created by normalizing this activation function. Sigmoid is used in two-class classification models, in our case Pneumonia and Non-Pneumonia.

3.4 Sigmoid

$$Sigmoidfunctionequation\sigma(x) = \frac{1}{1 + e^{-x}} \tag{4}$$

3.5 Pooling layer

In much famous modern deep learning architecture, a convolutional layer is followed by a pooling layer. This layer ensures that the model can remember the essential

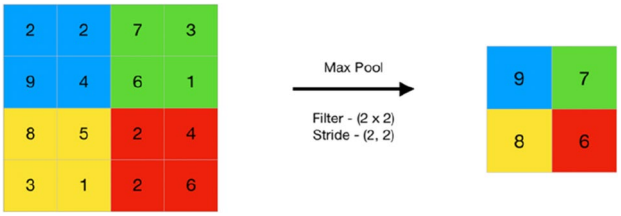


Fig. 3 Max pooling

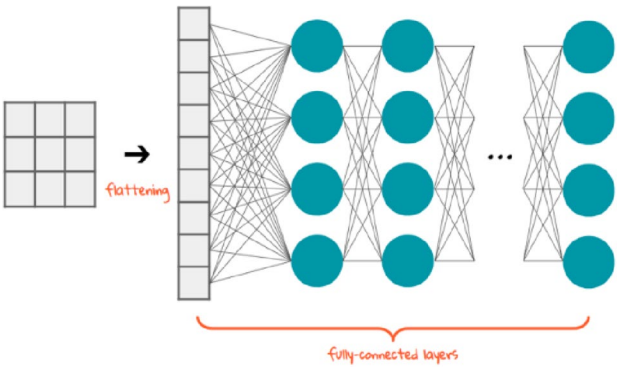


Fig. 4 Flattening to fully connected layer

information from the input. It does that by picking up the pixel with the highest value in a subset of the matrix, usually 2×2. This layer also helps reduce the model's complexity and allows the model to train faster. The models presented in this paper use max-pooling (Fig. 3) techniques it helps recognize essential features of the input. We have also used max pooling to prevent the model from over-fitting.

3.6 Flattening layer and fully connected layers

This layer, shown in (Fig. 4) usually comes at the end of convolution and pooling layers and transforms the input of size AxBxC into a column vector of size A*B*C. Each fully connected layer is followed by either another fully connected layer with fewer parameters or an activation function such as sigmoid or softmax. Each node of a fully connected layer is connected to all the nodes of the next layer. Hence most of the convolutional neural network parameters are found in fully connected layers. Feature extraction is done by layers which means that they extract useful information from data, and based on that data, a prediction is made by the network (Devulapalli, et al. 2021; Devulapalli and Krishnan 2021). After forward propagation, a cost function is calculated, which essentially estimates how well one model performs. After assessing the cost, backpropagation is initiated, which corrects the weights and biases introduced in the model based on the cost. This interaction is repeated until the network accomplishes optimum performance. An

optimizer such as SGD (Stochastic Gradient Descent) and ADAM uses the backpropagation process.

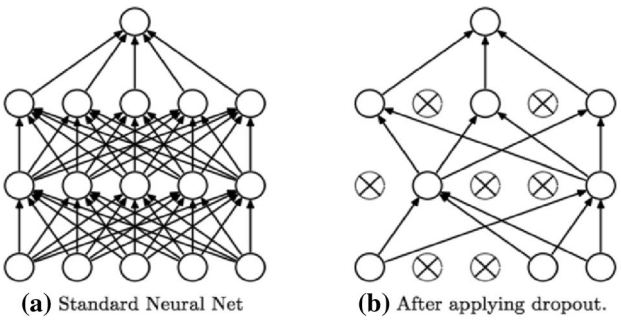
3.7 Reducing overfitting

Due to skewed data, that is, having many records with no pneumonia, the machine learning model may become better at just the training data and may not become robust to any other data. To prevent this possibility and make the model robust and prevent the vanishing gradient problem, We used the dropout layer in our model with a value of 0.3, randomly turning on and off a few neurons in the network. The value of 0.3 signifies that 30% of any random set of neurons will be off during a forward pass. This forces the model to learn a different path to reach the output and thereby increases the changes of the model to predict the outcome correctly. Data Augmentation techniques such as horizontal and vertical flip also helped us reduce the overfitting and generalizing better over test data. Visual representation of dropout as shown in (Fig. 5).

4 Transfer learning and fine-tuning

As the name suggests, transfer learning is transferring learning from one model to another. With transfer learning, it is possible to take the learning from a model trained on an extensive dataset and then precisely tune it to another dataset. This approach is instrumental in training the model faster, and at the same time, it is also computationally less expensive. Transfer learning helps to train and analyze multiple models quickly.

In this paper, we use transfer learning to analyseVGG-Net, DenseNet121, ResNet50 and NASNet. We specifically used the models trained on ImageNet and marked the last layers of these pre-trained models as trainable. The final few layers of the model contain dataset-specific information; hence, re-training them on another dataset makes the models compatible with new data. For NASNet, we add another fully connected layer with 256 parameters and another layer with one parameter stating pneumonia or not. This was followed by a sigmoid function which provided the likelihood of pneumonia. As pre-trained models are fast to train, we use multiple pre-trained models with different hyperparameters to determine which combination of hyperparameters and models works best for Pneumonia Detection. For example, We experiment with pre-trained DenseNet121 with various combinations of batch size, activation function, loss function, etc., similarly for NASNet and ResNet50.



(a) Standard Neural Net **(b)** After applying dropout.

Fig. 5 Dropout visual representation

5 Architecture

We used 2 models, ResNet50 (Sudheer et al. 2019) and DenseNet121 (Huang et al. 2017) with various hyper parameter configurations. This model aims to find a model with fewer parameters and high accuracy. In this section, we compare and contrast Simple CNN, ResNet50, and DenseNet121. (Fig. 6).

5.1 ResNet

ResNetsshown in (Fig. 7) extended the behavior of a simple CNN by including the skip connection with this Eq. (5) all the information learned by the n-1th convolution layer is present with the nth convolution layer; hence, the information propagates better and stays longer.

$$x_l = H_l(x_l - 1) + x_{l-1} \tag{5}$$

5.2 DenseNet

DenseNets (Huang et al. 2017) make the first difference from ResNets (Sudheer et al. 2019) in the way information is passed. DenseNets concatenates feature maps of the output layer with the incoming feature maps instead of adding them. Hence, the equation turns out to be: Eq. (6)

$$x_l = H_l([x_0 + x_1 + .. + x_{l-1}]) \tag{6}$$

A similar issue we faced in our work on ResNets, (Fig. 8), the combining activity of feature maps is impossible when they are of different sizes regardless of the combining activity being an addition or a concatenation of feature maps. Therefore, and the same way we used for ResNets (Sudheer et al. 2019), Dense Nets (Huang et al. 2017) (Fig. 8) are divided into dense blocks. Within a partnership, the dimensions of the feature maps remain the same, but the number of filters changes between them. These layers between the dense blocks are called Transition Layers. They are used for the down-sampling by applying batch normalization, a 1×1 convolution, and 2×2 pooling layers. The channel

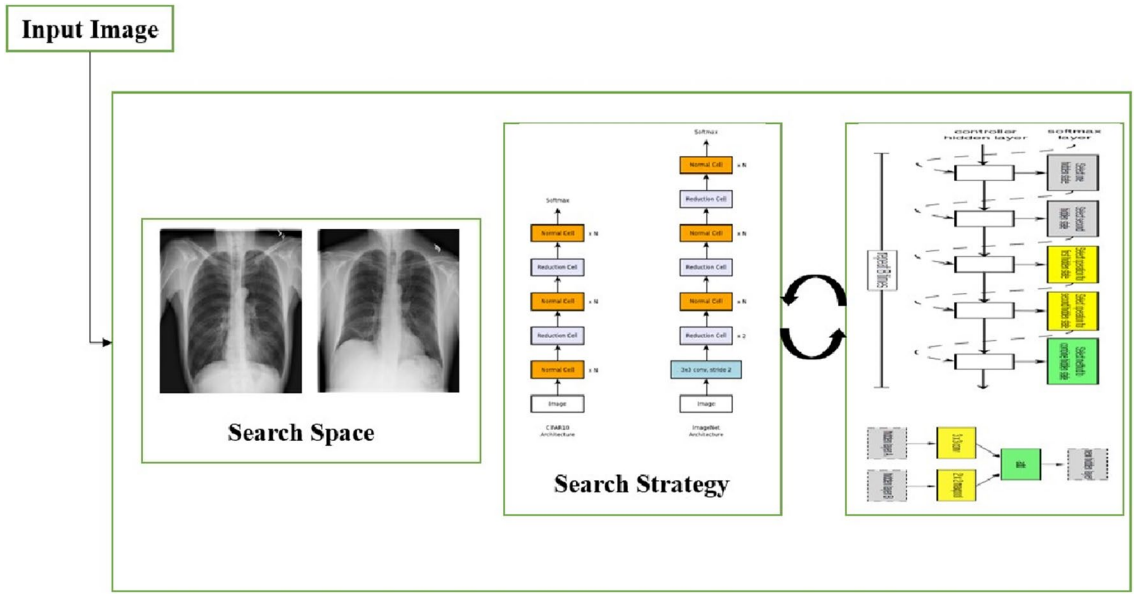


Fig.6 The architecture of the proposed method

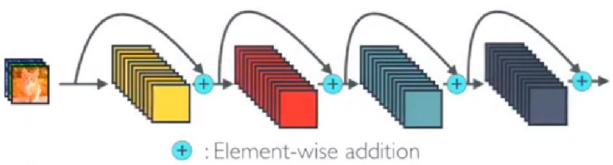


Fig.7 ResNet

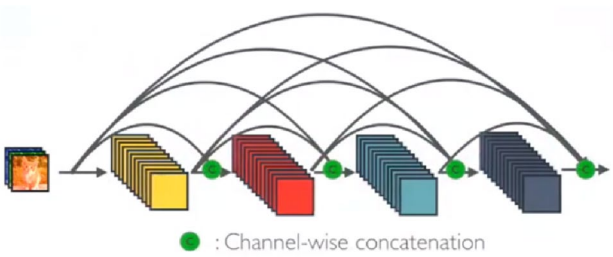


Fig. 8 Dense block

dimensions increase at every layer because of the concatenation of feature maps. 'k' is the growth rate hyperparameter that maintains the amount of information stored in each layer of the network. If H_1 produces k feature maps, then the generalized equation for determining the number of feature maps by lthlayer is as shown in Eq. (6).

Feature maps act as the information or learning data of the network. Every layer has access to feature maps of previous layers, and hence they have collective knowledge. Each layer adds new information or feature map to this collective knowledge, in concrete k feature maps of data. Because of this, DenseNet can save and use the information from previous layers better than ResNets and conventional ConvNet.

5.3 Propose method: NASNET

Google Brain creates NASNet. Authors of NASNet (Jamshidi et al. 2020) presented to search for an architectural building block on a tiny data and afterward move the block to a bigger data. Mainly, the team searched for an excellent convolutional layer on Canadian Institute for Advanced Research-10 data first and then applied it to ImageNet data by adding similar cells one upon the other. A regularization method called Scheduled Drop Path is also presented, which essentially increases the generalization of NASNet. Finally, the NASNet model accomplishes cutting-edge results with smaller model sizes and lower intricacy (FLOPs).

In this paper, we use pre-trained NASNet model and evaluate it on various hyper-parameters such as activation function, optimiser, batch size etc. which helped us achieve the highest accuracy among the current solutions with least computational cost.

5.4 NASNET architecture

- As presented by NASNet, (Fig. 9), authors, the best blocks are searched by the model using reinforcement learning.
- N is the number of motif repetitions which is a free parameter along with convolutional filters at starting step, which are used for scaling.

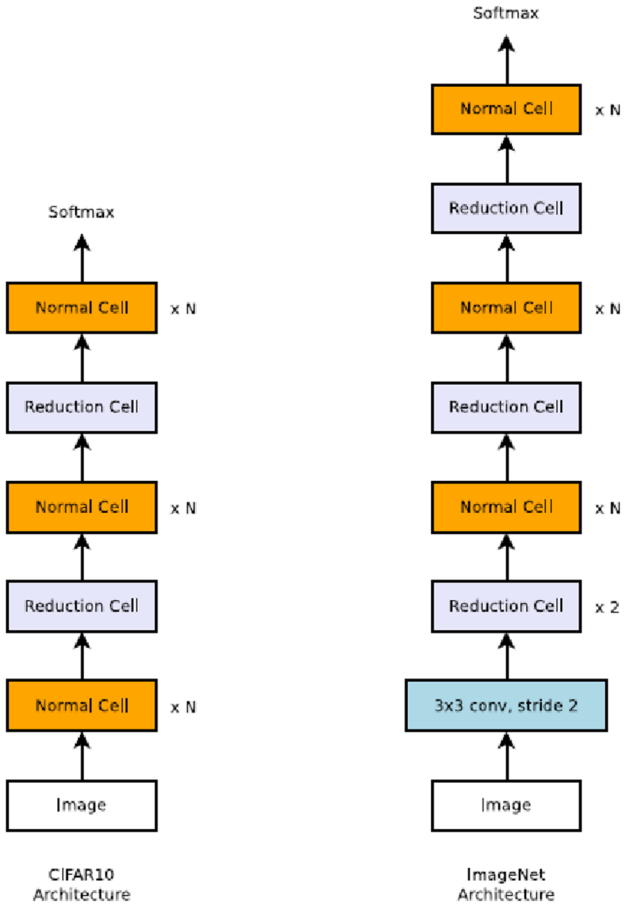


Fig. 9 CIFAR 10 & imagenet architecture

- These cells are known as normal cells and reduction cell. A normal cell performs convolution operation and returns a feature map of same size as input. A reduction cellbasically performs an operation similar to max pooling and returns a feature map which is half the size of input in all dimensions.
- The structures of normal and reduction cell are identified through a controller unit employing RNNs.

5.5 Controller model architecture

- The controller, (Fig. 10) is an RNN which predicts the remaining structure of normal and reduction cell given the first two hidden states as below.
- As a first step A hidden state(hj)is selected from current(hi), and previous (hi-1) hidden statues or from hidden states created in previous blocks.
- Second step includes selecting another hidden state (hj + 1) from the same options as in previous step.
- In step 3, we have to select an operation to be applied on hidden state (hj)
- In step 3, we have to select an operation to be applied on hidden state (hj + 1)
- Final step includes selecting a method to combine (hj + 1) and (hj).

There are certain only operations which are applied by the controller model to search for the convolution block.

Operations allowed in controller RNN.

– Identity 1 × 1 convolution
– 1 × 7 convolution followed by a 7 × 1 convolution filter.
– 1 × 3 convolution followed by a 3 × 1 convolution filter.
– 3 × 3 dilated convolution
– 3 × 3 average pooling
– Max pooling of 3 × 3, 5 × 5, 7 × 7 filter size.
– Depth wise-separable convolution of 3 × 3, 5 × 5. 7 × 7 filter size

We trained the multiple machine learning models of NASNet, ResNet, DenseNet and evaluated the outcome of these models on various parameters. We used logistic loss function as mentioned in Eq. (12, 13)to evaluate the loss and used various optimisers such as ADAM, AdaGrad, SGD to optimise the model. Multiple models of NASNet A, NASNet B and NASNet C are shown below in (Figs. 11, 12, 13) respectively. All NASNets mentioned below uses same fundamental concept of convolution, pooling and activations but differs in the way their normal and reduction cells are designed.

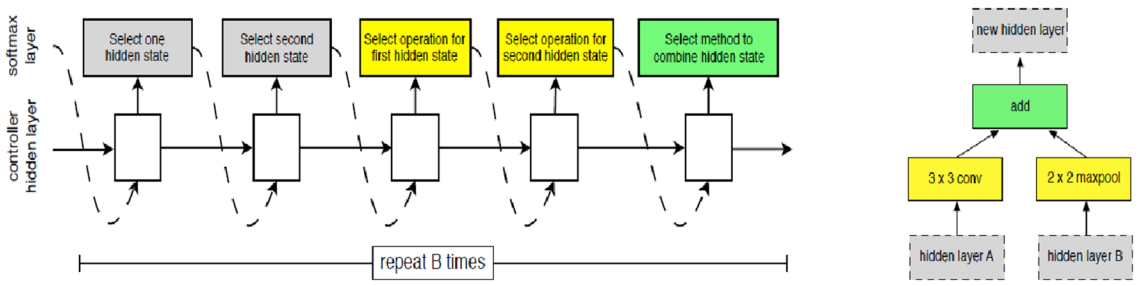


Fig. 10 Controller RNN

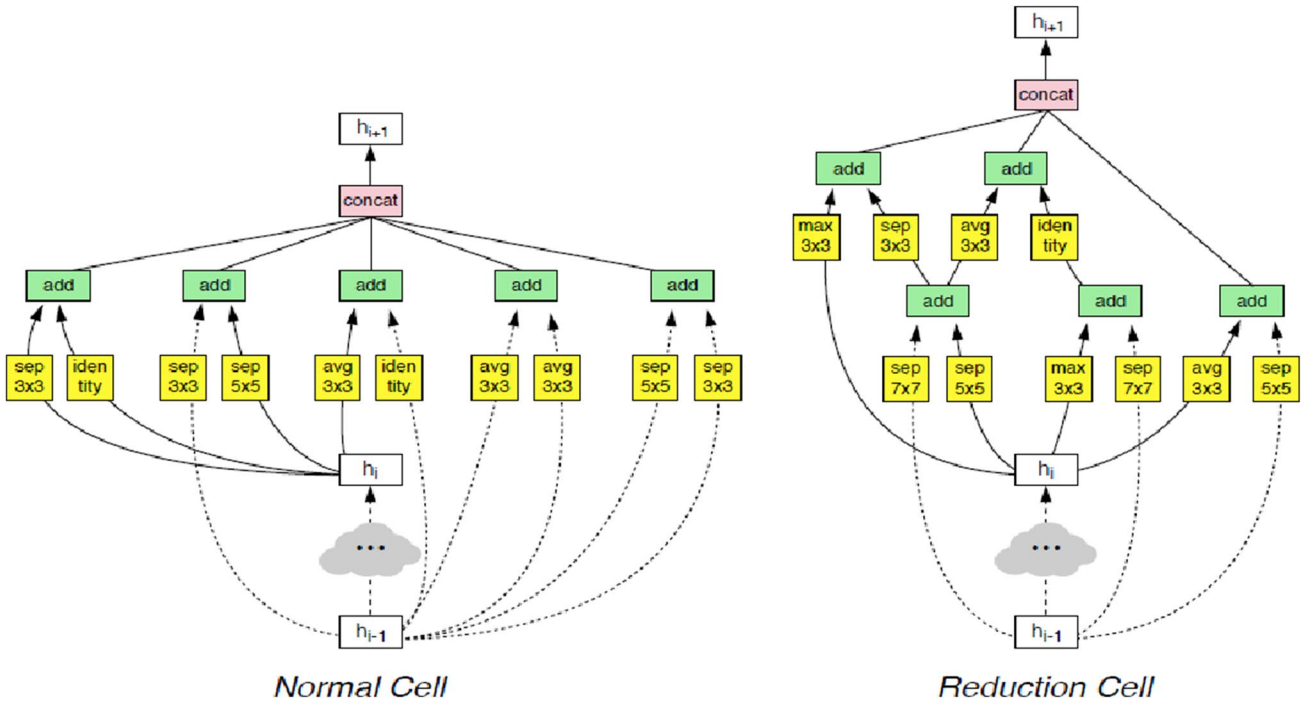


Fig. 11 NASNet A

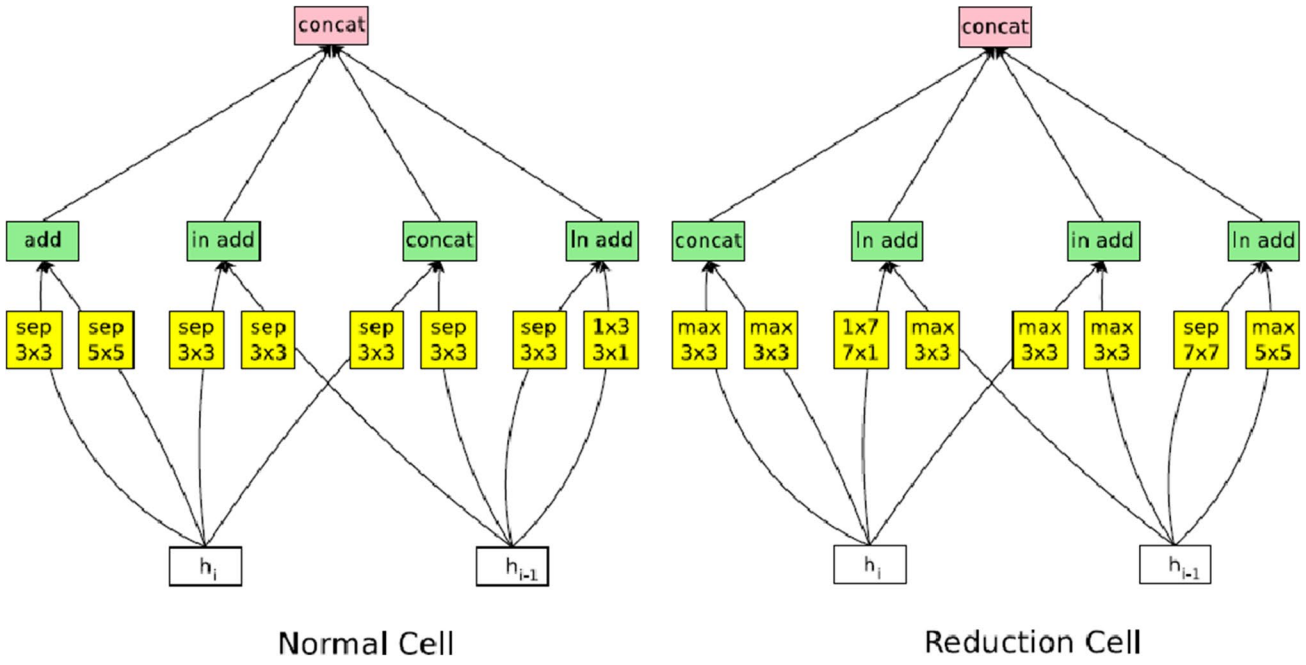


Fig.12 NASNet B

6 Model training

We trained our deep learning model on NVIDIA Super Computer Lab for Deep learning in Bennett University, Noida. (Fig. 14) shows the flow diagram of model.

NVIDIA DGX-1 V100 has NVIDIA Pascal, powered by Tesla V100 accelerators. As we had a massive dataset of 42 gigabytes, we employed this supercomputer, experimented

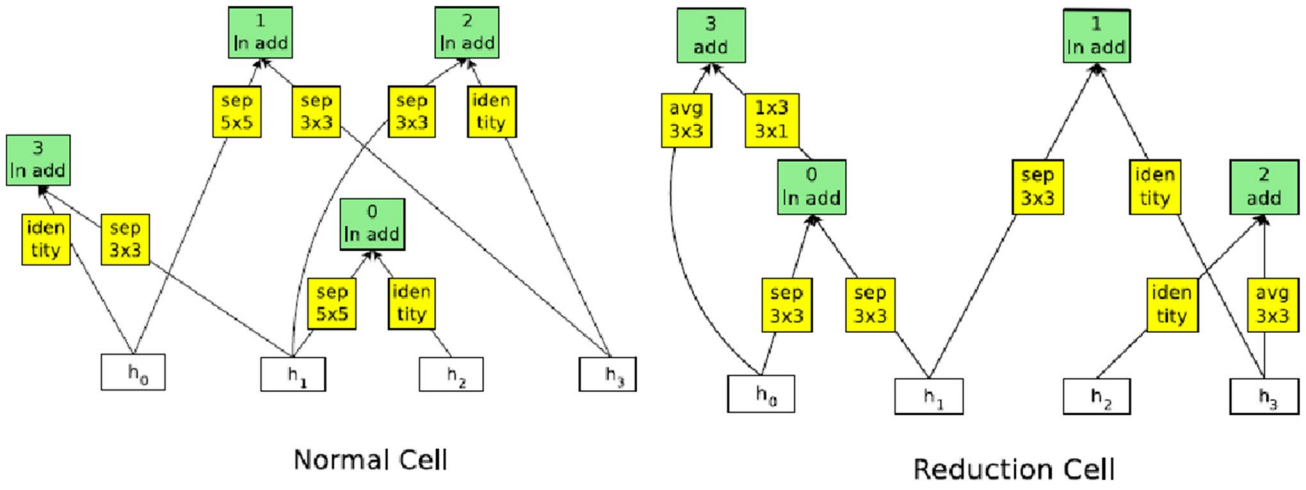


Fig. 13 NASNet C

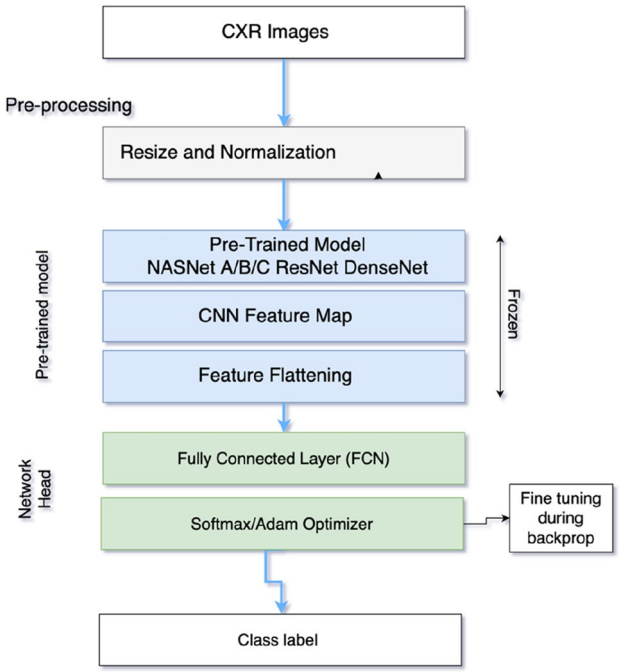


Fig. 14 Flow diagram

with, and trained on various hyperparameters. We trained the model parallelly on 6GPUs and used TensorFlow to maintain the consistency and connection among GPUs training model parallelly on different parts of training data. We used fourfold cross-validation to get the validation dataset while training.

Due to the massive size of the dataset, we employed Keras data generators, where the model read data from the SSD. We experimented with various optimizers such as Stochastic Gradient Descent (Figs. 15, 16), Adam,

model	depth	# params	error rate (%)
DenseNet ($L = 40, k = 12$) [26]	40	1.0M	5.24
DenseNet($L = 100, k = 12$) [26]	100	7.0M	4.10
DenseNet ($L = 100, k = 24$) [26]	100	27.2M	3.74
DenseNet-BC ($L = 100, k = 40$) [26]	190	25.6M	3.46
Shake-Shake 26 2x32d [18]	26	2.9M	3.55
Shake-Shake 26 2x96d [18]	26	26.2M	2.86
Shake-Shake 26 2x96d + cutout [12]	26	26.2M	2.56
NAS v3 [71]	39	7.1M	4.47
NAS v3 [71]	39	37.4M	3.65
NASNet-A (6 @ 768)	-	3.3M	3.41
NASNet-A (6 @ 768) + cutout	-	3.3M	2.65
NASNet-A (7 @ 2304)	-	27.6M	2.97
NASNet-A (7 @ 2304) + cutout	-	27.6M	2.40
NASNet-B (4 @ 1152)	-	2.6M	3.73
NASNet-C (4 @ 640)	-	3.1M	3.59

Fig. 15 Comparison of various Models w.r.t error rate and parameters

AdaGrad, and SGD with momentum. Stochastic Gradient Descent (SGD) persists only one learning rate, which does not change while training, whereas Adam, adjusts the learning rate dynamically during the training. To find the minima of the loss function, ResNet50 also implements Adam optimizer. Adam calculates the moving average of the gradient *mt* Eq. (7)/squared gradients ***vt*** Eq. (8) and the parameters to alter the learning rate during the training.

$$m_t = \beta_1 m_t - 1 + \left(1 - \beta_1\right) g_t \tag{7}$$

$$v_t = \beta_2 v_t - 1 + \left(1 - \beta_2\right) g_t^2 \tag{8}$$

$\boldsymbol{m_t}$is the first moment of gradient.
$\boldsymbol{v_t}$is the second moment of gradient.

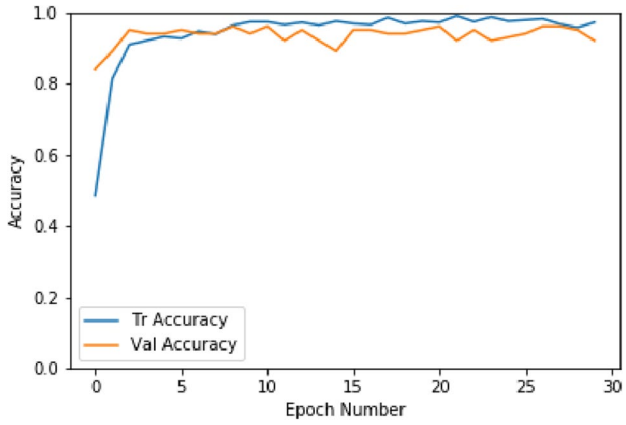


Fig. 16 NASNET A accuracy graph

In the above Eq. (8) on the right, *vt* are set to zero's. Biases are changed to correct value by calculating 1st and 2nd moment estimates in Eq. (9) (Sudheer et al. 2019):

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t}\hat{v}_t = \frac{v_t}{1-\beta_2^t} \tag{9}$$

Following Adam update rule Eq. (10) is used to update the weights:

$$(\theta) : \theta_{t+1} = \theta_t - \eta\frac{\hat{m}_t}{\sqrt{\hat{v}_t}+\varepsilon} \tag{10}$$

***η*:** Eta is the learning rate hyperparameter. It is the rate at which model learns from the input. Very high learning rate could make the training unstable and very low learning rate could make the training process very slow.

β_1 : isexponential decay rate for the first moment estimates.

β_2 : isexponential decay rate for the second-moment estimates (e.g. 0.999).

ε: It'sa small number to prevent division by zero exception (e.g. 10E-8).

We trained two variants of each model with different input sizes and two different optimizers: Adam and SGD, for 28 epochs. Alongside, as we observed that DenseNet121 with SGD as optimizer performs better than others, we also trained a variation of DenseNet121 with SGD and batched size of 32.

Activation functions such as sigmoid and hyperbolic tangent experience saturation near their input's middle point. Using these activation functions can result in inadequate training and problems such as vanishing gradient (Huang et al. 2017) and exploding gradient. To overcome these problems, we use the ReLU (Rectified Linear Unit) activation function, Eq. (11)

$$g(x) = max(0, x) \tag{11}$$

7 Evaluation metrics and analysis

7.1 Loss function: log loss

$$L(x, y) = -(y * log(x) + (1-y) * (log(1-x)) \tag{12}$$

$$Wherex = p(Y = 1|X) \tag{13}$$

With above experiments as shown in Table 1, has described detailed comparison of the proposed and existing deep learning models accuracy by varying with number of parameters, batch size. It also shows the precision, accuracy and F1 score. (Table 2) We achieved F1-score significantly better than existing systems with 1/10th of the parameters as shown in (Fig. 15).

F1 score: CheXNet 0.435 (0.387, 0.481).

F1 score: NASNet 0.9749 (0.97, 0.98).

8 Precisionvs. computational demand

9 Conclusion and future work

In India, the number of lives lost due to Pneumonia has increased drastically. We proposed a novel approach using deep learning to detect Pneumonia in rural and urban areas. We experimented with various deep learning models and presented that NASNet not just achieves state-of-the-art results in terms of accuracy and loss. It is also efficient for computing power. It has minor trainable parameters, with 2.6 M, with the highest precision and recall. Further, the proposed methodology can be used for prediction with as low compute power as in the mobile app. This will not only help in rural India, where there is a lack of radiologists but will also help radiologists have a quick and reliable second opinion on pneumonia detection (Figs. 16, 17, 18 19).

Funding None.

Declarations

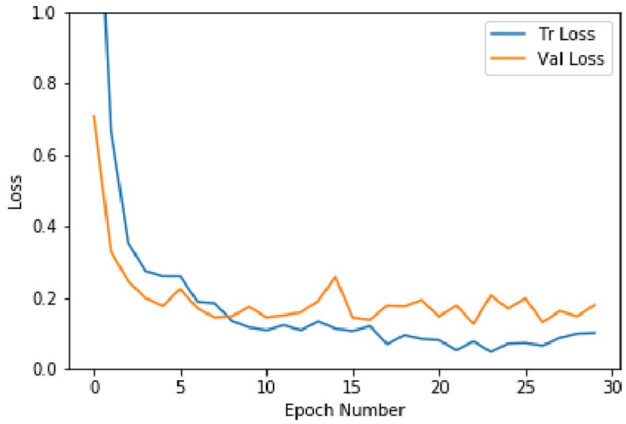


Fig.17 NASNet A loss graph

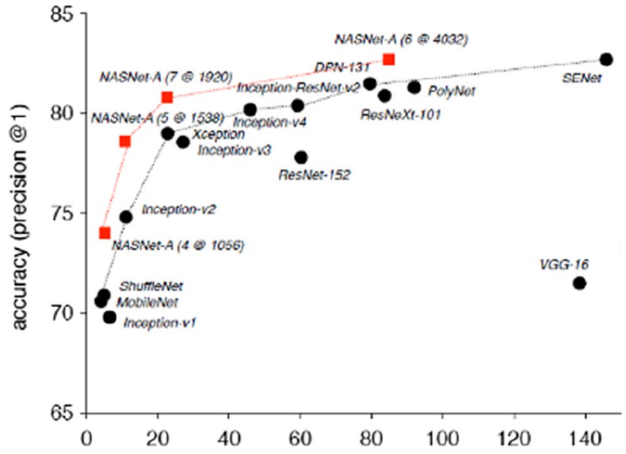


Fig. 18 parameters millions

Table 2 Comparison of execution time with proposed and existing models

S. no	Model	Execution time in (S)
1	NASNet-A (6@768)	720
2	NASNet B (4@1152)	715
3	NASNet C	708
4	CheXNet	702
5	ResNet50	690

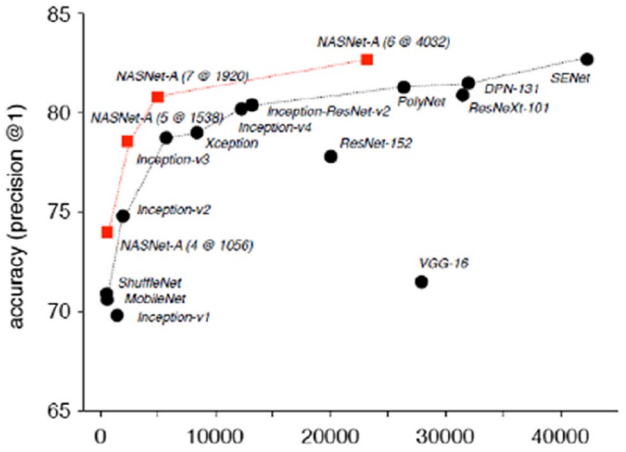


Fig. 19 Add and multiplication operations

Table 1 Evaluation metrics

Model	Precision(*P*)	Recall(R)	F1 Score	Accuracy(%)	Batch Size	(%) Error	# params
NASNet-A (6@768)	**0.97**	**0.98**	**0.9749**	**97.35**	32	**2.65**	3.41 M
NASNet B (4@1152)	0.94	0.93	0.934	94.35	32	3.73	**2.65 M**
NASNet C	0.91	0.87	0.8895	92.34	64	3.71	3.1 M
CheXNet	0.387	0.481	0.435	92.63	32	3.74	27.2 M
ResNet50	0.31	0.33	0.319	91.85	64	5.23%	29.8 M

NASNet-A has higher Precision, Recall, F1 Score. Accuracy, Batch size and low error rate when compaer to other models whereas NASNet B required less number of parametrs when compared to other models

Conflicts of interest The authors declare that they have no conflicts of interest.

Human participants or animals resources This article does not contain any studies involving animals performed by any of the authors.

Informed consent This article does not contain any studies involving human participants performed by any of the authors.

References

Akter S et al (2021) COVID-19 detection using deep learning algorithm on chest X-ray images. Biology 10(11):1174

Alakus TB, Turkoglu I (2020) Comparison of deep learning approaches to predict COVID-19 infection. Chaos, Solitons Fractals 140:110120

Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S (2016) Lung pattern classification for interstitial lung

diseases using a deep convolutional neural network. IEEE Trans Med Imaging 35(5):1207–1216
Currie BJ, Mayo M, Anstey NM, Donohoe P, Haase A, Kemp DJ (2001) A cluster of melioidosis cases from an endemic region is clonal and is linked to the water supply using molecular typing of Burkholderiapseudomallei isolates. Am J Trop Med Hyg 65(3):177–179
Dagaonkar R S, Udwadia Z F, Sen T, Nene A, Joshi J, Rastogi S A & Pandey K (2012) Severe community acquired pneumonia in Mumbai, India: Etiology and predictive value of the modified British thoracic society rule. In D61. respiratory clinical epidemiology (pp. A6060-A6060). American Thoracic Society.
Davis SM, Deloria-Knoll M, Kassa HT, O'Brien KL (2013) Impact of pneumococcal conjugate vaccines on nasopharyngeal carriage and invasive disease among unvaccinated people: review of evidence on indirect effects. Vaccine 32(1):133–145
Devulapalli S, Krishnan R (2021) Remote sensing image retrieval by integrating automated deep feature extraction and handcrafted features using curvelet transform. J Appl Remote Sens 15(1):016504
Devulapalli, Sudheer, et al. (2021) "Experimental evaluation of unsupervised image retrieval application using hybrid feature extraction by integrating deep learning and handcrafted techniques." Materials Today: Proceedings.
Feikin, D. R., Kagucia, E. W., Loo, J. D., Link-Gelles, R., Puhan, M. A., Cherian, T., … & Serotype Replacement Study Group (2013) Serotype-specific changes in invasive pneumococcal disease after pneumococcal conjugate vaccine introduction: a pooled analysis of multiple surveillance sites. PLoSMedicine 10(9):e1001517
Guan Q, Huang Y, Zhong Z, Zheng Z, Zheng L, & Yang Y (2018) Diagnose like a radiologist: Attention guided convolutional neural network for thorax disease classification. arXiv preprint arXiv:1801.09927
Harmon SA, Sanford TH, Xu S et al (2020) Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. Nat Commun 11:4080. https://doi.org/10.1038/s41467-020-17971-2
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).
http://www.who.int/maternal_child_adolescent/epidemiology/gappd-monitoring/en/
Huang G, Liu Z, Van Der Maaten L, & Weinberger K Q (2017) Densely connected convolution networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp 4700–4708).
Jamshidi M et al (2020) Artificial intelligence and COVID-19: deep learning approaches for diagnosis and treatment. IEEE Access 8:109581–109595. https://doi.org/10.1109/ACCESS.2020.3001973
Kaushik V S, Nayyar A, Kataria G, & Jain R (2020). Pneumonia detection using convolutional neural networks (CNNs). In: Proceedings of first international conference on computing communications, and cyber-security (IC4S 2019) (pp. 471–483).
Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. Adv Neural Inf Process Syst 25:1097–1105
Kumar A, Tripathi AR, Satapathy SC, Zhang YD (2022) SARS-net: COVID-19 detection from chest x-rays by combining graph convolutional network and convolutional neural network. Pattern Recogn 122:108255
Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, Xia J (2020) Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: evaluation of the diagnostic accuracy. Radiology 296(2):E65–E71
Loharikar A, Dumolard L, Chu S, Hyde T, Goodman T, Mantel C (2016) Status of new vaccine introduction—worldwide, september 2016. Morb Mortal Wkly Rep 65(41):1136–1140
McAllister DA, Liu L, Shi T, Chu Y, Reed C, Burrows J, Nair H (2019) Global, regional, and national estimates of pneumonia morbidity and mortality in children younger than 5 years between 2000 and 2015: a systematic analysis. Lancet Glob Health 7(1):e47–e57
Nalluri S, Sasikala R (2020) An insight into application of big data analytics in healthcare. Int J Data Mining, Model Manage 12(1):87–117
O'brien, K. L., Wolfson, L. J., Watt, J. P., Henkle, E., Deloria-Knoll, M., McCall, N. & Cherian, T. (2009) Burden of disease caused by streptococcus pneumoniae in children younger than 5 years: global estimates. The Lancet 374(9693):893–902
Oh Y, Park S, Ye JC (2020) Deep learning COVID-19 features on CXR using limited training data sets. IEEE Trans Med Imaging 39(8):2688–2700. https://doi.org/10.1109/TMI.2020.2993291
Pankratz DG, Choi Y, Imtiaz U, Fedorowicz GM, Anderson JD, Colby TV, Martinez FJ (2017) Usual interstitial pneumonia can be detected in transbronchial biopsies using machine learning. Ann Am Thorac Soc 14(11):1646–1654
Panwar H et al (2020) Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet. Chaos, Solitons Fractals 138:109944
Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T & Ng A Y (2017) Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225.
Record WE (2006) WHO Position Paper on Haemophilus influenzae type B conjugate vaccines. Week Epidemiol Record 24(47):1
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
Sharma R, Deoskar R, Bargaje M, Kumar P, & Agarwal Y (2013) A study of etiological and clinical profile of community acquired pneumonia in a tertiary care hospital in Western India
Simonyan K, & Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
Sudheer D, R SethuMadhavi, and P Balakrishnan (2019) "Edge and texture feature extraction using canny and haralick textures on SPARK cluster." proceedings of the 2nd international conference on data engineering and communication technology. Springer, Singapore, 2019.
Toğaçar M, Ergen B, Cömert Z (2020) COVID-19 detection using deep learning models to exploit social mimic optimization and structured chest X-ray images using fuzzy color and stacking approaches. Comput Biol Med 121:103805
Wahl B, Knoll MD, Shet A, Gupta M, Kumar R, Liu L, McAllister DA (2020) National, regional, and state-level pneumonia and severe pneumonia morbidity in children in India: modelled estimates for 2000 and 2015. The Lancet Child Adolescent Health 4(9):678–687
Walker CLF, Rudan I, Liu L, Nair H, Theodoratou E, Bhutta ZA, Black RE (2013) Global burden of childhood pneumonia and diarrhoea. The Lancet 381(9875):1405–1416
Wang, Xiaosong, Peng, Yifan, Lu, Le, Lu, Zhiyong, Bagheri, Mohammadhadi, and Summers, Ronald M (2007) Chestx-ray8: hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. arXiv preprint arXiv:1705.02315.
Watt JP, Wolfson LJ, O'Brien KL, Henkle E, Deloria-Knoll M, McCall N, Cherian T (2009) Burden of disease caused by Haemophilus influenzae type b in children younger than 5 years: global estimates. The Lancet 374(9693):903–911

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

Zhao D, Yao F, Wang L, Zheng L, Gao Y, Ye J, Gao R (2020) A comparative study on the clinical features of coronavirus 2019 (COVID-19) pneumonia with other pneumonias. Clin Infect Dis 71(15):756–761

Zoph B, Vasudevan V, Shlens J, & Le Q V (2018) Learning transferable architectures for scalable image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp 8697–8710).