greatlearning
Learning for Life

Capstone Project

Pneumonia Detection Challenge

Problem Statement

- **Here's the backstory and why solving this problem matters.**
- Pneumonia accounts for over 15% of all deaths of children under 5 years old internationally. In 2017, 920,000 children under the age of 5 died from the disease.
- It requires review of a chest radiograph (CXR) by highly trained specialists and confirmation through clinical history, vital signs and laboratory exams. Pneumonia usually manifests as an area or areas of increased opacity on CXR. However, the diagnosis of pneumonia on CXR is **complicated because of a number of other conditions in the lungs such as fluid overload (pulmonary edema), bleeding, volume loss (atelectasis or collapse), lung cancer, or post-radiation or surgical changes**. Outside of the lungs, fluid in the pleural space (pleural effusion) also appears as increased opacity on CXR. When available, comparison of CXRs of the patient taken at different time points and correlation with clinical symptoms and history are helpful in making the diagnosis.
- CXRs are the most commonly performed diagnostic imaging study. A number of factors such as positioning of the patient and depth of inspiration can alter the appearance of the CXR, complicating interpretation further. In addition, **clinicians are faced with reading high volumes of images every shift.**

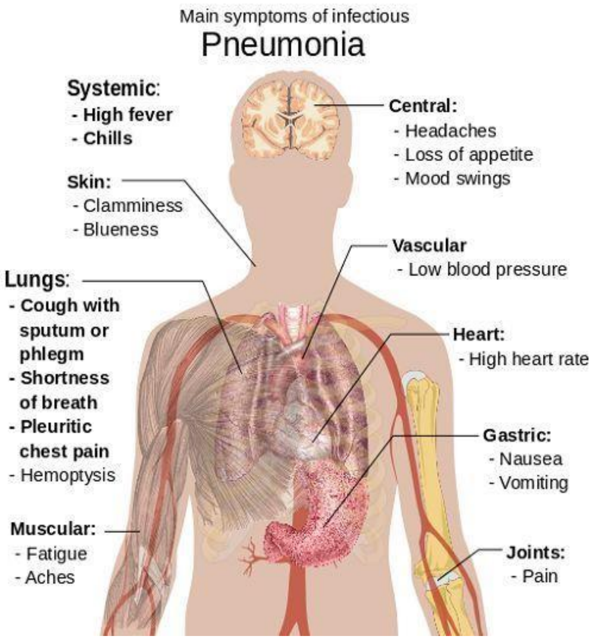
About Pneumonia

Symptoms to detect Pneumonia

- The Diagnosis of pneumonia on CXR(is complicated because of a number of other c (pulmonary edema), bleeding, volume loss (atelectasis or collapse), lung cancer, or p

Objective/Business case

- Is to detect the presence of pneumonia using binary classification and create a boundary box for the same.
- These objectives have been tried by different research groups. However here in this project we are trying improve the mapping identification and efficiency.
  Further, it can be enhanced with image captioning and report generation using statistical NLP.



Scope

- Identify pneumonia is there or not, binary class of identification
- Create boundary box for the identified region on the Image

Business Value

Business Value

- Automating Pneumonia screening in chest radiographs, providing affected area details through bounding box
- Assist physicians to make better clinical decisions or even replace human judgement in certain functional areas of healthcare (eg, radiology).
- Guided by relevant clinical questions, powerful AI techniques can unlock clinically relevant information hidden in the massive amount of data, which in turn can assist clinical decision making.

Dataset

About Dataset

We are using the dataset provided by RSNA on Kaggle with the below features:

- stage_2_train.csv - the training set. Contains patientIds and bounding box / target information.

  for each image.

  patientId _- A patientId. Each patientId corresponds to a unique image. x_ - the upper-left x coordinate of the bounding box.

  y_ - the upper-left y coordinate of the bounding box. width_ - the width of the bounding box.

  height_ - the height of the bounding box.

  Target_ - the binary Target, indicating whether this sample has evidence of pneumonia.
- stage_2_train_images - set of training images
- stage_2_test_images - set of test images

Dicom original images

Medical images are stored in a special format called as DICOM files (*.dcm). They contain a combination of header metadata as well as underlying raw image arrays for pixel data

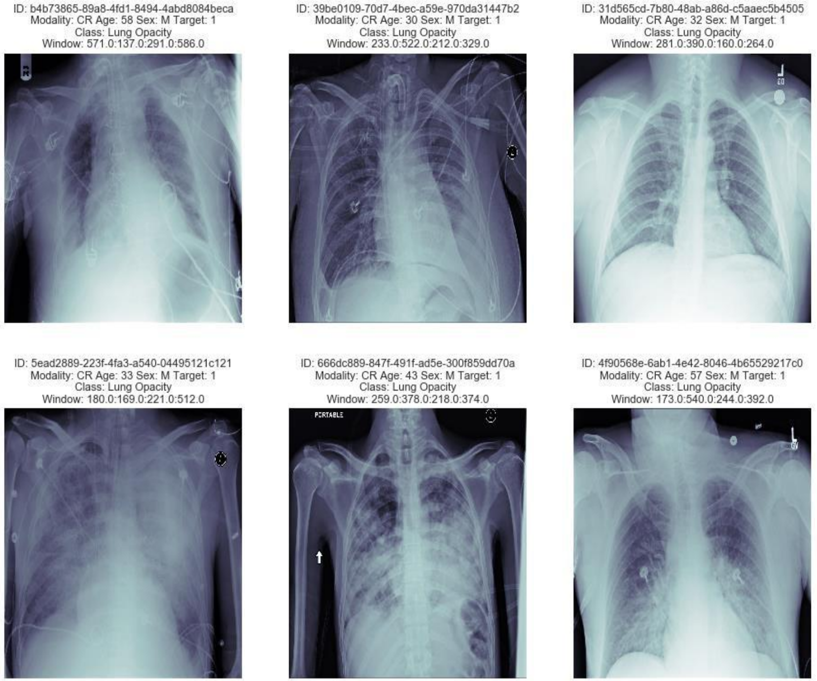
greatlearning
Learning for Life

Exploratory Data Analysis

greatlearning
Learning for Life

Exploratory Visualization(with lung opacity)

- Tissues with sparse material, such as lungs which are full of air, do not absorb the X-rays and appear black in the image. Dense tissues such as bones absorb the X-rays and appear white in the image.
- While we are theoretically detecting “lung opacities”, there are lung opacities that are not pneumonia related.
- In the data, some of these are labeled “Not Normal No Lung Opacity”. This extra third class indicates that while pneumonia was determined not to be present, there was nonetheless some type of abnormality on the image and oftentimes this finding may mimic the appearance of true pneumonia. These non pneumonia “Not Normal” detections end up being a primary source of frustration in building models.
- The images below show a few of the examples in the “Not Normal” class.



We would like to represent the images with the overlay boxes superposed. For this, we will need first to parse the whole dataset with Target = 1 and gather all coordinates of the windows showing a Lung Opacity on the same image.

greatlearning
Learning for Life

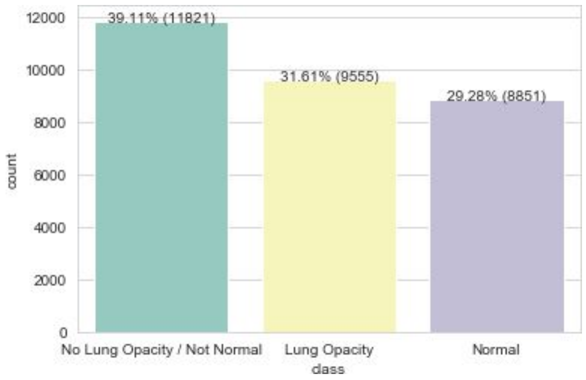
Exploratory Data Analysis

Preliminary EDA insights

class_info dataset **30227** rows and **2** columns

26684 unique patients

Class Distribution



In [52]:
```
# Let's merge now the two datasets, using Patient ID as the merge criteria.
train_class_df = train_labels_df.merge(class_info_df, left_on='patientId', right_on='patientId', how='inner')
train_class_df.sample(5)
```

Out[52]:

	patientId	x	y	width	height	Target	class
6061	3dcc0fa9-312e-427f-9a03-132bcb72b3e0	265.0	148.0	189.0	433.0	1	Lung Opacity
33134	ed20cdc6-e8f7-4551-a8d5-6b360b3f9c71	245.0	315.0	145.0	290.0	1	Lung Opacity
17385	8aee90c4-2018-4326-83b9-a7080655597f	NaN	NaN	NaN	NaN	0	No Lung Opacity / Not Normal
23284	afef1dff-d4a2-47fa-bfbc-ce4f4e16e9eb	NaN	NaN	NaN	NaN	0	No Lung Opacity / Not Normal
13610	7281385b-4397-44f1-9176-6e7d72773eb3	NaN	NaN	NaN	NaN	0	No Lung Opacity / Not Normal

Preliminary EDA insights –training images

Training has **26684** unique images

- Images are dcm format
- Every image has the following details
    - -Patient sex;
    - -Patient age;
    - -Modality;
    - -Body part examined;
    - -View position;
    - -Rows & Columns;
    - -Pixel Spacing.

Preliminary EDA insights -train labels dataset

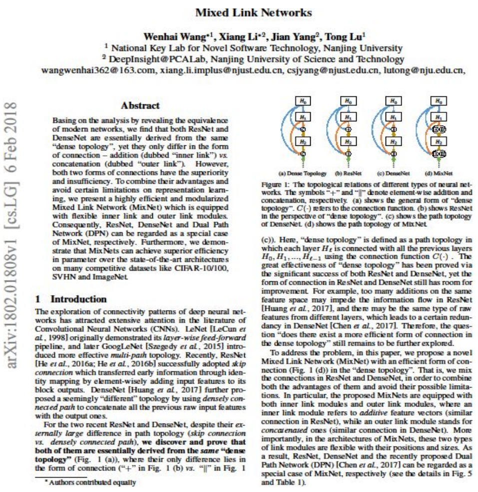
- **30227** rows and **6** columns
- -patientId and x, y coordinates, width, height of the bounding boxes and target indicating Pneumonic case or not
- Missing values –nil
- **26684** unique patients

	patientId	x	y	width	height	Target
0	0004cfab-14fd-4e49-80ba-63a80b6bddd6	NaN	NaN	NaN	NaN	0
1	00313ee0-9eaa-42f4-b0ab-c148ed3241cd	NaN	NaN	NaN	NaN	0
2	00322d4d-1c29-4943-afc9-b6754be640eb	NaN	NaN	NaN	NaN	0
3	003d8fa0-6bf1-40ed-b54c-ac657f8495c5	NaN	NaN	NaN	NaN	0
4	00436515-870c-4b36-a041-de91049b9ab4	264.0	152.0	213.0	379.0	1

Model building

- Mixed Link Networks with mix of DenseNet & ResNet (using Adam Optimizer)

Reference: Below 2 papers are referred for the model building.



Mixed Link Networks

Wenhai Wang*[1], Xiang Li*[2], Jian Yang[2], Tong Lu[1]
[1] National Key Lab for Novel Software Technology, Nanjing University
[2] DeepInsight@PCALab, Nanjing University of Science and Technology
wangwenhai362@163.com, xiang.li.implus@njust.edu.cn, csjyang@njust.edu.cn, lutong@nju.edu.cn,

arXiv:1802.01808v1 [cs.LG] 6 Feb 2018

Abstract

Basing on the analysis by revealing the equivalence of modern networks, we find that both ResNet and DenseNet are essentially derived from the same "dense topology", yet they only differ in the form of connection – addition (dubbed "inner link") *vs.* concatenation (dubbed "outer link"). However, both two forms of connections have the superiority and insufficiency. To combine their advantages and avoid certain limitations on representation learning, we present a highly efficient and modularized Mixed Link Network (MixNet) which is equipped with flexible inner link and outer link modules. Consequently, ResNet, DenseNet and Dual Path Network (DPN) can be regarded as a special case of MixNet, respectively. Furthermore, we demonstrate that MixNets can achieve superior efficiency in parameter over the state-of-the-art architectures on many competitive datasets like CIFAR-10/100, SVHN and ImageNet.

1 Introduction

The exploration of connectivity patterns of deep neural networks has attracted extensive attention in the literature of Convolutional Neural Networks (CNNs). LeNet [LeCun *et al.*, 1998] originally demonstrated its *layer-wise feed-forward* pipeline, and later GoogLeNet [Szegedy *et al.*, 2015] introduced more effective *multi-path* topology. Recently, ResNet [He *et al.*, 2016a; He *et al.*, 2016b] successfully adopted *skip connection* which transferred early information through identity mapping by element-wisely adding input features to its block outputs. DenseNet [Huang *et al.*, 2017] further proposed a seemingly "different" topology by using *densely connected path* to concatenate all the previous raw input features with the output ones.

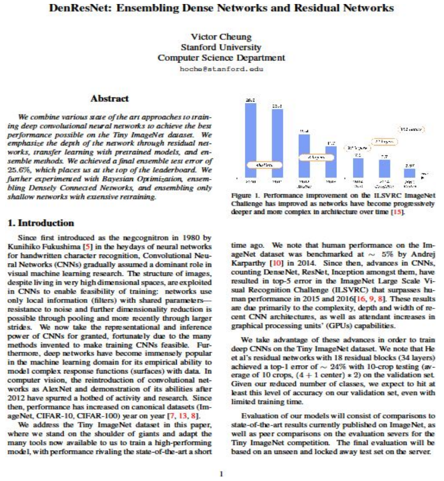
For the two recent ResNet and DenseNet, despite their *externally large* difference in path topology (*skip connection vs. densely connected path*), **we discover and prove that both of them are essentially derived from the *same* "dense topology"** (Fig. 1 (a)), where their only difference lies in the form of connection ("+" in Fig. 1 (b) *vs.* "‖" in Fig. 1 (c)). Here, "dense topology" is defined as a path topology in which each layer H_ℓ is connected with all the previous layers $H_0, H_1, ..., H_{\ell-1}$ using the connection function $C(\cdot)$. The great effectiveness of "dense topology" has been proved via the significant success of both ResNet and DenseNet, yet the form of connection in ResNet and DenseNet still has room for improvement. For example, too many additions on the same feature space may impede the information flow in ResNet [Huang *et al.*, 2017], and there may be the same type of raw features from different layers, which leads to a certain redundancy in DenseNet [Chen *et al.*, 2017]. Therefore, the question "does there exist a more efficient form of connection in the dense topology" still remains to be further explored.

To address the problem, in this paper, we propose a novel Mixed Link Network (MixNet) with an efficient form of connection (Fig. 1 (d)) in the "dense topology". That is, we mix the connections in ResNet and DenseNet, in order to combine both the advantages of them and avoid their possible limitations. In particular, the proposed MixNets are equipped with both inner link modules and outer link modules, where an inner link module refers to *additive* feature vectors (similar connection in ResNet), while an outer link module stands for *concatenated* ones (similar connection in DenseNet). More importantly, in the architectures of MixNets, these two types of link modules are flexible with their positions and sizes. As a result, ResNet, DenseNet and the recently proposed Dual Path Network (DPN) [Chen *et al.*, 2017] can be regarded as a special case of MixNet, respectively (see the details in Fig. 5 and Table 1).

(a) Dense Topology (b) ResNet (c) DenseNet (d) MixNet

Figure 1: The topological relations of different types of neural networks. The symbols "+" and "‖" denote element-wise addition and concatenation, respectively. (a) shows the general form of "dense topology". $C(\cdot)$ refers to the connection function. (b) shows ResNet in the perspective of "dense topology". (c) shows the path topology of DenseNet. (d) shows the path topology of MixNet.

* Authors contributed equally



DenResNet: Ensembling Dense Networks and Residual Networks

Victor Cheung
Stanford University
Computer Science Department
hoche@stanford.edu

Abstract

We combine various state of the art approaches to training deep convolutional neural networks to achieve the best performance possible on the Tiny ImageNet dataset. We emphasize the depth of the network through residual networks, transfer learning with pretrained models, and ensemble methods. We achieved a final ensemble test error of 25.6%, which places us at the top of the leaderboard. We further experimented with Bayesian Optimization, ensembling Densely Connected Networks, and ensembling only shallow networks with extensive retraining.

Figure 1. Performance improvement on the ILSVRC ImageNet Challenge has improved as networks have become progressively deeper and more complex in architecture over time [15].

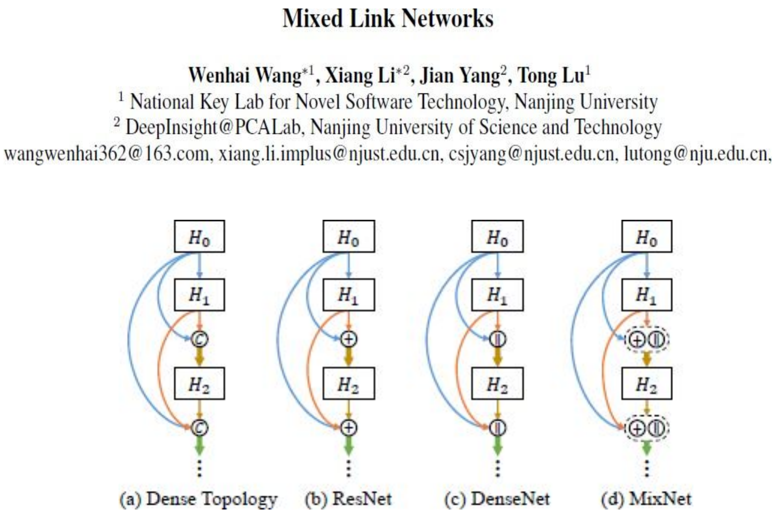
1. Introduction

Since first introduced as the negcognitron in 1980 by Kunihiko Fukushima [5] in the heydays of neural networks for handwritten character recognition, Convolutional Neural Networks (CNNs) gradually assumed a dominant role in visual machine learning research. The structure of images, despite living in very high dimensional spaces, are exploited in CNNs to enable feasibility of training: networks use only local information (filters) with shared parameters—resistance to noise and further dimensionality reduction is possible through pooling and more recently through larger strides. We now take the representational and inference power of CNNs for granted, fortunately due to the many methods invented to make training CNNs feasible. Furthermore, deep networks have become immensely popular in the machine learning domain for its empirical ability to model complex response functions (surfaces) with data. In computer vision, the reintroduction of convolutional networks as AlexNet and demonstration of its abilities after 2012 have spurred a hotbed of activity and research. Since then, performance has increased on canonical datasets (ImageNet, CIFAR-10, CIFAR-100) year on year [7, 13, 8].

We address the Tiny ImageNet dataset in this paper, where we stand on the shoulder of giants and adapt the many tools now available to us to train a high-performing model, with performance rivaling the state-of-the-art a short time ago. We note that human performance on the ImageNet dataset was benchmarked at $\sim$ 5% by Andrej Karparthy [10] in 2014. Since then, advances in CNNs, counting DenseNet, ResNet, Inception amongst them, have resulted in top-5 error in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) that surpasses human performance in 2015 and 2016[16, 9, 8]. These results are due primarily to the complexity, depth and width of recent CNN architectures, as well as attendant increases in graphical processing units' (GPUs) capabilities.

We take advantage of these advances in order to train deep CNNs on the Tiny ImageNet dataset. We note that He et al's residual networks with 18 residual blocks (34 layers) achieved a top-1 error of $\sim$ 24% with 10-crop testing (average of 10 crops, $(4 + 1 \text{ center}) * 2$) on the validation set. Given our reduced number of classes, we expect to hit at least this level of accuracy on our validation set, even with limited training time.

Evaluation of our models will consist of comparisons to state-of-the-art results currently published on ImageNet, as well as peer comparisons on the evaluation severs for the Tiny ImageNet competition. The final evaluation will be based on an unseen and locked away test set on the server.



Challenges

- Training of 26k+ images on a local laptop for hours using different CNN Models (infrastructure)
- Understanding of CNN based model development papers
- Tuning of the models and right utilization of the parameters.

greatlearning
Learning for Life

Thank you!