Disease Detecting in NIH Chest X-ray Images

Author Names Go Here Using the Author Name Style (Required)

Affiliations and Addresses Go Here Using the Affiliation and Address Style (Required)  
publications21@aaai.org (email is text roman font using Affiliation and Address Style (Required)

[*At least 3 carriage returns must separate the author affiliation block from the beginning of the two-column text block*]

Abstract

Chest X-ray exams are one of the most frequent and cost-effective medical imaging examinations available.

However, clinical diagnosis of a chest X-ray can be challenging and sometimes more difficult than diagnosis via chest CT imaging. By using machine learning and deep learning techniques, we aim to ease the detection and classification of various lung disease. This paper proposes and compares the performance between ML model with Manual Feature extraction, ML Models with CNN features and deep learning Neural network models. This paper concludes with the best performing model from through experimentation.

Introduction*[[1]](#footnote-1)*

Large-scale multimedia data has accelerated the development of tasks in computer vision, such as image retrieval and classification, video analysis. In radiological studies, Chest X-ray (CXR) is one of commonly used screening techniques in thorax disease diagnosis, such as Pneumonia, nodules, cardiomegaly, effusion etc. Thousands of CXR images are captured in hospitals, making the computer-aided diagnosis (CAD) very important but challenging. Therefore, automatic analysis of CXR images would effectively assist clinical diagnosis and pathology finding. However, chest X-ray image analysis is a challenging task which suffers from the intrinsically complex relations of different pathologies. In this paper, we conduct experiments with deep learning techniques and multiclass and Multi label machine learning techniques.

The dataset used in this project is from the National Institute of Health (NIH). This NIH Chest X-ray Dataset is comprised of 112,120 X-ray images with disease labels from 30,805 unique patients. To create these labels, the authors used Natural Language Processing to text-mine disease classifications from the associated radiological reports. The labels are expected to be >90% accurate and suitable for weakly-supervised learning.

There are 15 classes (14 diseases, and one for "No findings"). Images can be classified as "No findings" or one or more disease classes.

Atelectasis, Consolidation, Infiltration, Pneumothorax, Edema, Emphysema, Fibrosis, Effusion, Pneumonia, Pleural thickening, Cardiomegaly, Nodule Mass, Hernia.

Commonly, CXR images are labeled with one or more pathologies, which makes the CXR image classification a multi-label problem. In the NIH Chest X-ray dataset, each image is annotated with multiple lung-related pathologies. In this paper we propose and experiment with multi-label and multiclass Machine Learning models.

Deep learning has made noticeable progress in field of medical image analysis, such as classification, segmentation or detection, image registration. In this paper we propose and have designed purpose built Convolutional Neural Network and experiment with multiple Pretrained Deep learning Models such as Resnet 161, Resnet 152, Resnet121, VGG16.

Summarized Contributions

* CNN Pretrained models – experimenting with hyper parameter configuration of pretrained ML models
* Traditional ml models with CNN features – extracting features generated by our CNN models and using them to train traditional ML models.
* Traditional ml models with Manual features. – using Manual feature extraction methods like GLCM, LBP and discrete Wavelet transforms to extract features from CXR images and training traditional ML models with these features.

Graphical user interface, application

Description automatically generated

Timeline

Description automatically generated with low confidence

Graphical user interface

Description automatically generated with medium confidence

Background

Manual Feature Extraction

Local Binary Pattern

LBP is an effective texture pattern descriptor introduced by Ojala et al. [23] to describe the local texture patterns of an image. It is widely used in the applications based on [image processing](https://www.sciencedirect.com/topics/engineering/image-processing). The LBP works in a block size of 3×3, in which the center pixel is used as a threshold for the [neighboring pixel](https://www.sciencedirect.com/topics/engineering/neighboring-pixel), and the LBP code of a center pixel is generated by encoding the computed threshold value into a decimal value.

Text, letter

Description automatically generated

Grey Level Cooccurrence Matrix

In statistical texture analysis, texture features are computed from the statistical distribution of observed combinations of intensities at specified positions relative to each other in the image. According to the number of intensity points (pixels) in each combination, statistics are classified into first-order, second- order and higher-order statistics. The Gray Level Cooccurrence Matrix (GLCM) method is a way of extracting second order statistical texture features.

Given an image composed of pixels each with an intensity (a specific gray level), the GLCM is a tabulation of how often different combinations of gray levels co-occur in an image or image section. Texture feature calculations use the contents of the GLCM to give a measure of the variation in intensity (a.k.a. image texture) at the pixel of interest.

Discrete Wavelet Transforms

The fundamental idea behind wavelets is to analyze signal according to scale. It has gained a lot of interest in the area of signal processing, numerical analysis and mathematics during recent years. Generally, the wavelet transform is an advanced technique of signal and image analysis. It was developed as an alternative to the short time Fourier to overcome problems related to its frequency and time resolution properties. The basic idea of DWT is to provide the time-frequency representation. The2D-DWT represents an image in terms of a set of shifted and dilated wavelet functions

Text, letter

Description automatically generated

{LH, UL, HH}, and Nj =N/2. In this paper LH, UL and UVI are called wavelet or DWT sub-bands.

Diagram, schematic

Description automatically generated

Machine Learning Models

Multinomial Logistic regression

In statistics, multinomial logistic regression is a classificationmethod that generalizes logistic regression to multiclass problems, i.e., with more than two possible discrete outcomes. That is, it is a model that is used to predict the probabilities of the different possible outcomes of a categorically distributed dependent variable, given a set of independent variables (which may be real-valued, binary-valued, categorical-valued, etc.).

Multinomial logistic regression is used when the dependent variable in question is nominal (equivalently *categorical*, meaning that it falls into any one of a set of categories that cannot be ordered in any meaningful way) and for which there are more than two categories.

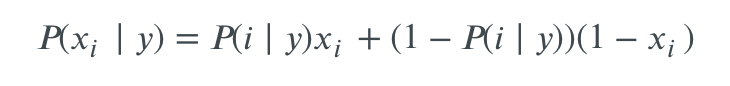
Bernoulli NB

Bernoulli Naive Bayes is a variant of Naive Bayes. This is used for discrete data, and it works on Bernoulli distribution. The main feature of Bernoulli Naive Bayes is that it accepts features only as binary values like true or false, yes or no, success or failure, 0 or 1 and so on. So, when the feature values are binary, we know that we must use Bernoulli Naive Bayes classifier.

Text

Description automatically generated

The decision rule for Bernoulli Naïve Bayes:



Gaussian NB

[Gaussian Naive Bayes](https://iq.opengenus.org/gaussian-naive-bayes/) is a variant of [Naive Bayes](https://iq.opengenus.org/text-classification-naive-bayes/) that follows Gaussian normal distribution and supports continuous data.

When working with continuous data, an assumption often taken is that the continuous values associated with each class are distributed according to a normal (or Gaussian) distribution. The likelihood of the features is assumed to be-(sd – standard deviation)

**P(x, mean, sd) = (1 / (sqrt(2 \* PI) \* sd)) \* exp(-((x-mean^2)/(2\*sd^2)))**

Gaussian Naive Bayes supports continuous valued features and models each as conforming to a Gaussian (normal) distribution. An approach to create a simple model is to assume that the data is described by a Gaussian distribution with no co-variance (independent dimensions) between dimensions. This model can be fit by simply finding the mean and standard deviation of the points within each label, which is all what is needed to define such a distribution.

Decision Trees

Decision tree learning or induction of decision trees is one of the predictive modelling approaches used in statistics, data mining and machine learning. It uses a decision tree (as a predictive model) to go from observations about an item (represented in the branches) to conclusions about the item's target value (represented in the leaves). Tree models where the target variable can take a discrete set of values are called classification [trees](https://en.wikipedia.org/wiki/Decision_tree); in these tree structures, [leaves](https://en.wikipedia.org/wiki/Leaf_node) represent class labels and branches represent [conjunctions](https://en.wikipedia.org/wiki/Logical_conjunction) of features that lead to those class labels. Decision trees where the target variable can take continuous values (typically real numbers) are called regression [trees](https://en.wikipedia.org/wiki/Decision_tree). Decision trees are among the most popular machine learning algorithms given their intelligibility and simplicity.

Extra Tree Classifier

Extremely RandomizedTrees Classifier (Extra Trees Classifier) is a type of ensemble learning technique which aggregates the results of multiple de-correlated decision trees collected in a “forest” to output its classification result. In concept, it is very similar to a Random Forest Classifier and only differs from it in the manner of construction of the decision trees in the forest.

Each Decision Tree in the Extra Trees Forest is constructed from the original training sample. Then, at each test node, each tree is provided with a random sample of k features from the feature-set from which each decision tree must select the best feature to split the data based on some mathematical criteria (typically the Gini Index). This random sample of features leads to the creation of multiple de-correlated decision trees.

Deep Learning Models

Convolutional Neural Networks

A Convolutional Neural Network, also known as CNN or ConvNet, is a class of neural networks that specializes in processing data that has a grid-like topology, such as an image. A digital image is a binary representation of visual data. It contains a series of pixels arranged in a grid-like fashion that contains pixel values to denote how bright and what color each pixel should be.

The human brain processes a huge amount of information the second we see an image. Each neuron works in its own receptive field and is connected to other neurons in a way that they cover the entire visual field. Just as each neuron responds to stimuli only in the restricted region of the visual field called the receptive field in the biological vision system, each neuron in a CNN processes data only in its receptive field as well. The layers are arranged in such a way so that they detect simpler patterns first (lines, curves, etc.) and more complex patterns (faces, objects, etc.) further along. By using a CNN, one can [enable sight to computers](https://www.datascience.com/blog/computer-vision-in-artificial-intelligence).

Diagram, engineering drawing

Description automatically generated

**Convolutional Layer**

The convolution layer is the core building block of the CNN. It carries the main portion of the network’s computational load.

This layer performs a dot product between two matrices, where one matrix is the set of learnable parameters otherwise known as a kernel, and the other matrix is the restricted portion of the receptive field. The kernel is spatially smaller than an image but is more in-depth. This means that, if the image is composed of three (RGB) channels, the kernel height and width will be spatially small, but the depth extends up to all three channels.

A picture containing text, electronics, screenshot

Description automatically generated

Related Work

This Is a Subsubsection Heading. Use of This Style Sheet Is Required for all subsubsections.

Project Description

Experiments

Conclusion

All AAAI papers must use author year citations in the text (Engelmore and Morgan 1986). Numbered references may not be used. The references must be formatted using the exact references style shown the Author

Citation Style and Reference Style

All AAAI papers must use author year citations in the text (Engelmore and Morgan 1986). Numbered references may not be used. The references must be formatted using the exact references style shown the Author Formatting Instructions for Authors Using Word. Citations must be placed in alphabetical order. Lines must be justified. References may not be numbered. No indent is allowed. The style sheet will automatically add the proper amount of space between entries; no additional space should be added.

References

Clancey, W. J. 1984. Classification Problem Solving. In *Proceedings of the Fourth National Conference on Artifi­cial Intelligence,* 49-54. Menlo Park, Calif.: AAAI Press.

Engelmore, R., and Morgan, A. eds. 1986. *Blackboard Sys­tems.* Reading, Mass.: Addison-Wesley.

Hasling, D. W.; Clancey, W. J.; and Rennels, G. R. 1983. Strategic Explanations in Consultation. *The International Journal of Man-Machine Studies* 20(1): 3–19.

Rice, J. 1986. Poligon: A System for Parallel Problem Solving, Technical Report, KSL-86-19, Department of Computer Science, Stanford University, Stanford, CA.

Robinson, A. L. 1980a. New Ways to Make Microcircuits Smaller. *Science* 208:1019-1026.

1. Copyright © 2021, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved. [↑](#footnote-ref-1)