

X-RAY PNEUMONIA IMAGE CLASSIFICATION WITH DEEP LEARNING

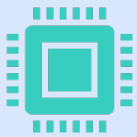


The Problem



- Globally, a child dies of pneumonia every 39 seconds.
- Pneumonia is the leading cause of morbidity and mortality in children younger than the age of 5, killing more children than HIV/AIDS, malaria, and measles combined.
- Chest X-rays are primarily used for the diagnosis of this disease. However, even for a trained radiologist, it is a challenging task to examine chest X-rays.

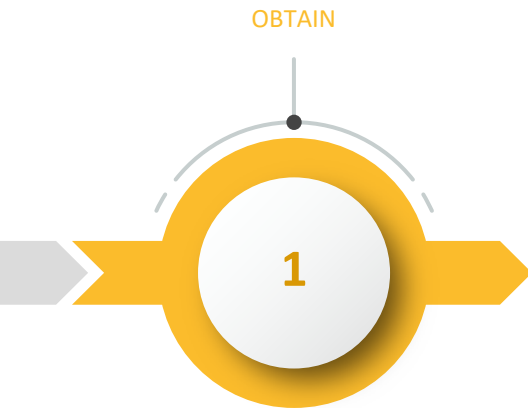
The Solution



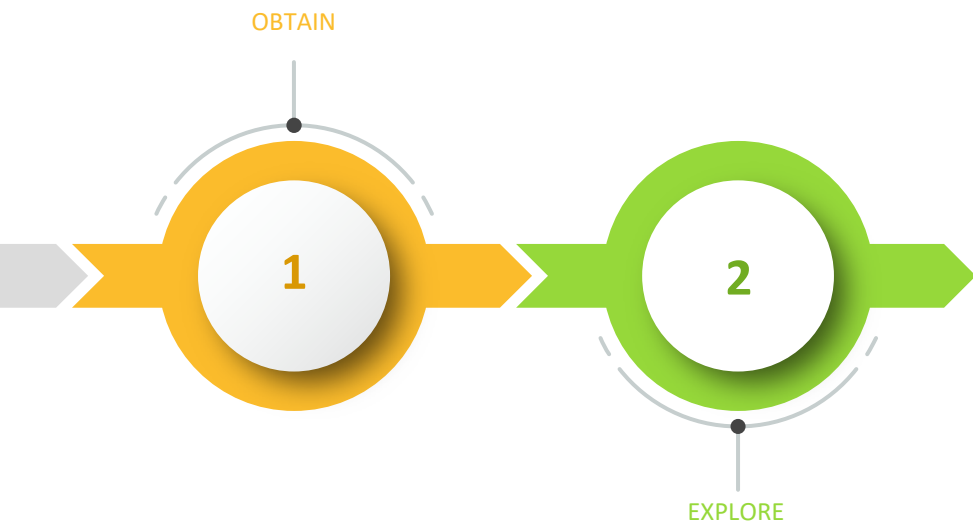
To solve this, deep learning (DL), a branch of machine learning (ML), is developed to detect hidden features in images which are not apparent or cannot be detected even by medical experts.

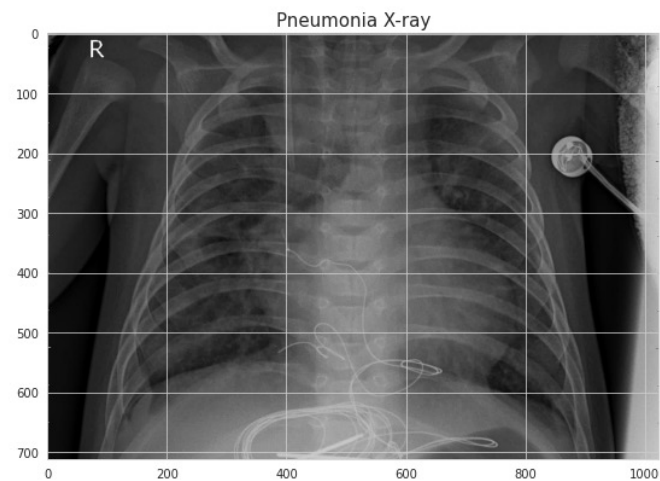
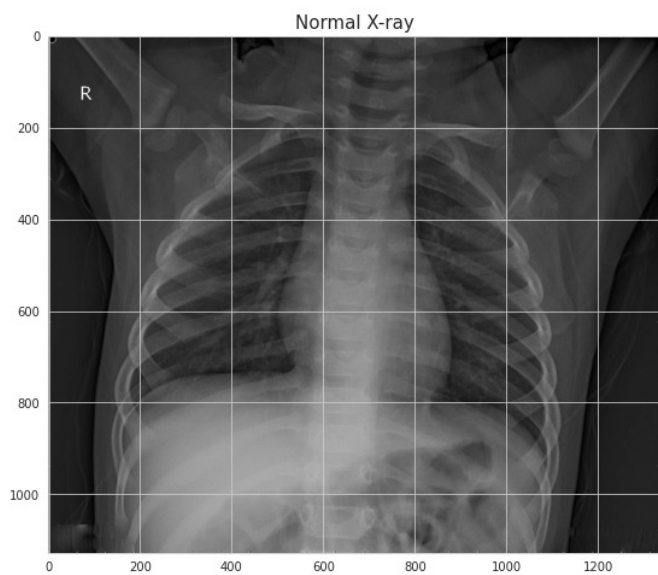


With AI system aiding medical experts in expediting the diagnosis, earlier treatment can be prescribed, resulting in improved clinical outcomes.



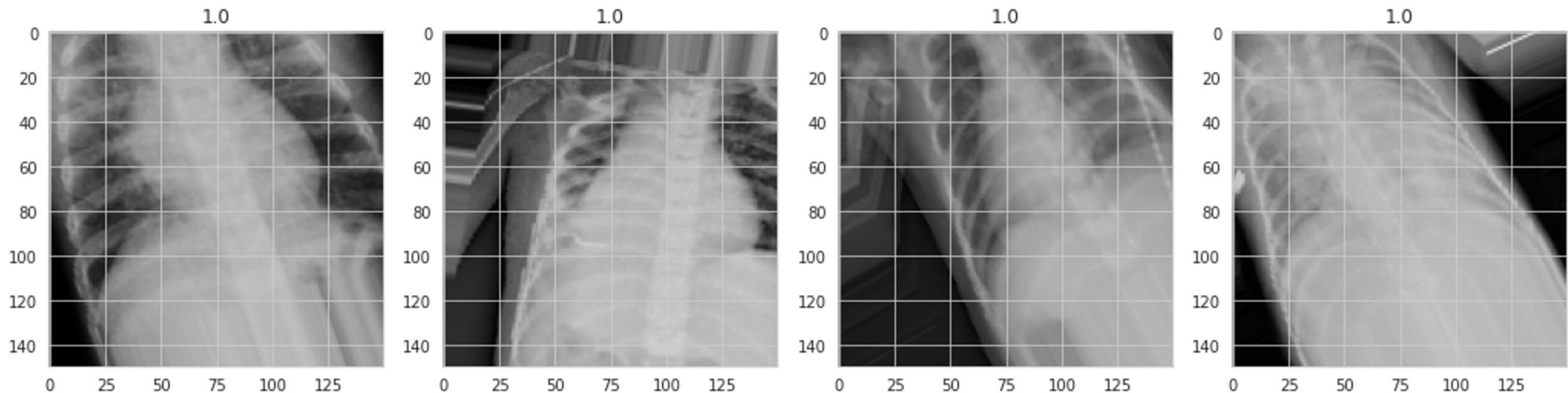
- Chest X-ray images (anterior-posterior) were selected from pediatric patients of one to five years old from Guangzhou Women and Children's Medical Center.
- There are 5,863 X-Ray images (JPEG)
- 2 categories: Normal & Pneumonia

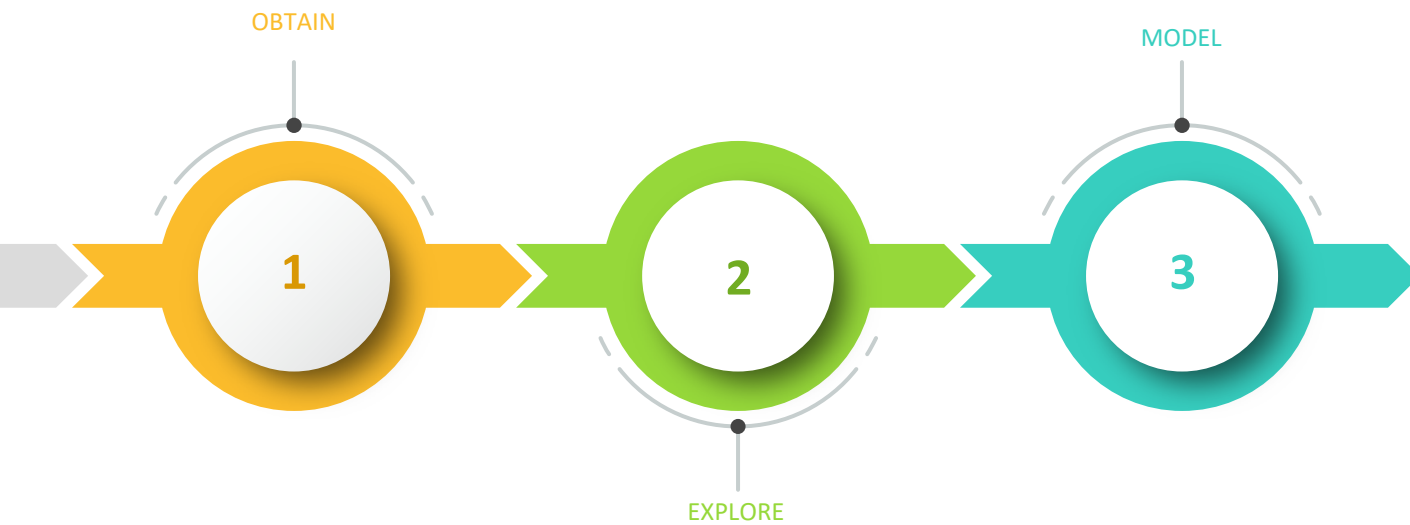


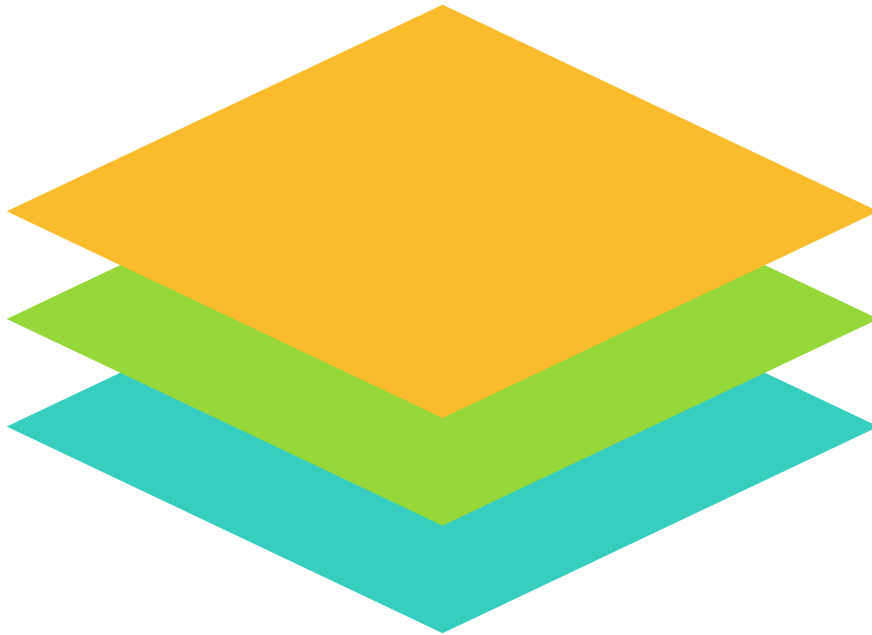





Data Augmentation

- To build a powerful image classifier using very little training data, image augmentation is usually required to boost the performance of deep networks.
- Image augmentation artificially creates training images through different ways of processing images, such as random rotation, shifts, shear and flips, etc. of each training instances

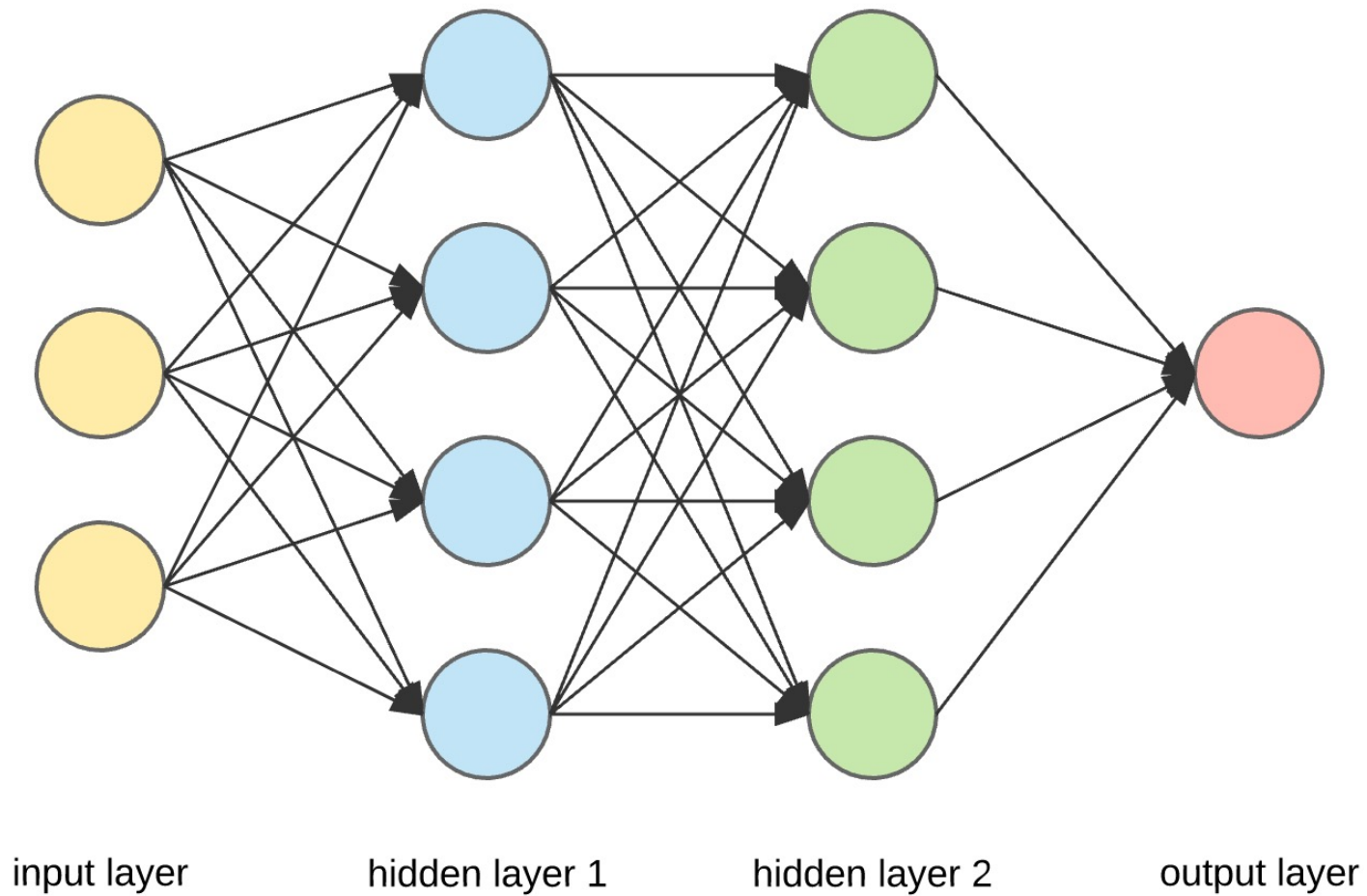






-  MULTILAYER PERCEPTRON MODEL (MLP)
-  CONVOLUTIONAL NEURAL NETWORK MODEL (CNN)
-  TRANSFER LEARNING with VGG16 CNN MODEL

Multilayer Perceptron



Multilayer Perceptron

input

dense_input: InputLayer	input:	[(None, 4096)]
	output:	[(None, 4096)]

hidden layer 1

dense: Dense	input:	(None, 4096)
	output:	(None, 32)

hidden layer 2

dense_1: Dense	input:	(None, 32)
	output:	(None, 32)

hidden layer 3

dense_2: Dense	input:	(None, 32)
	output:	(None, 64)

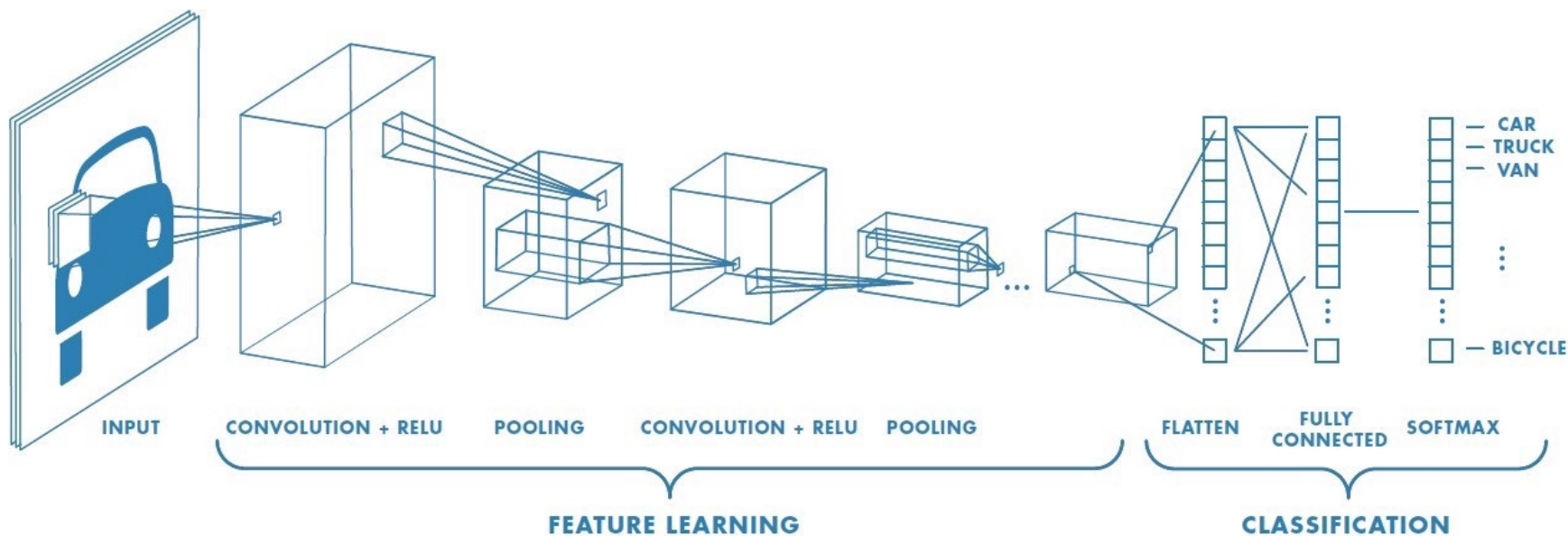
hidden layer 4

dense_3: Dense	input:	(None, 64)
	output:	(None, 128)

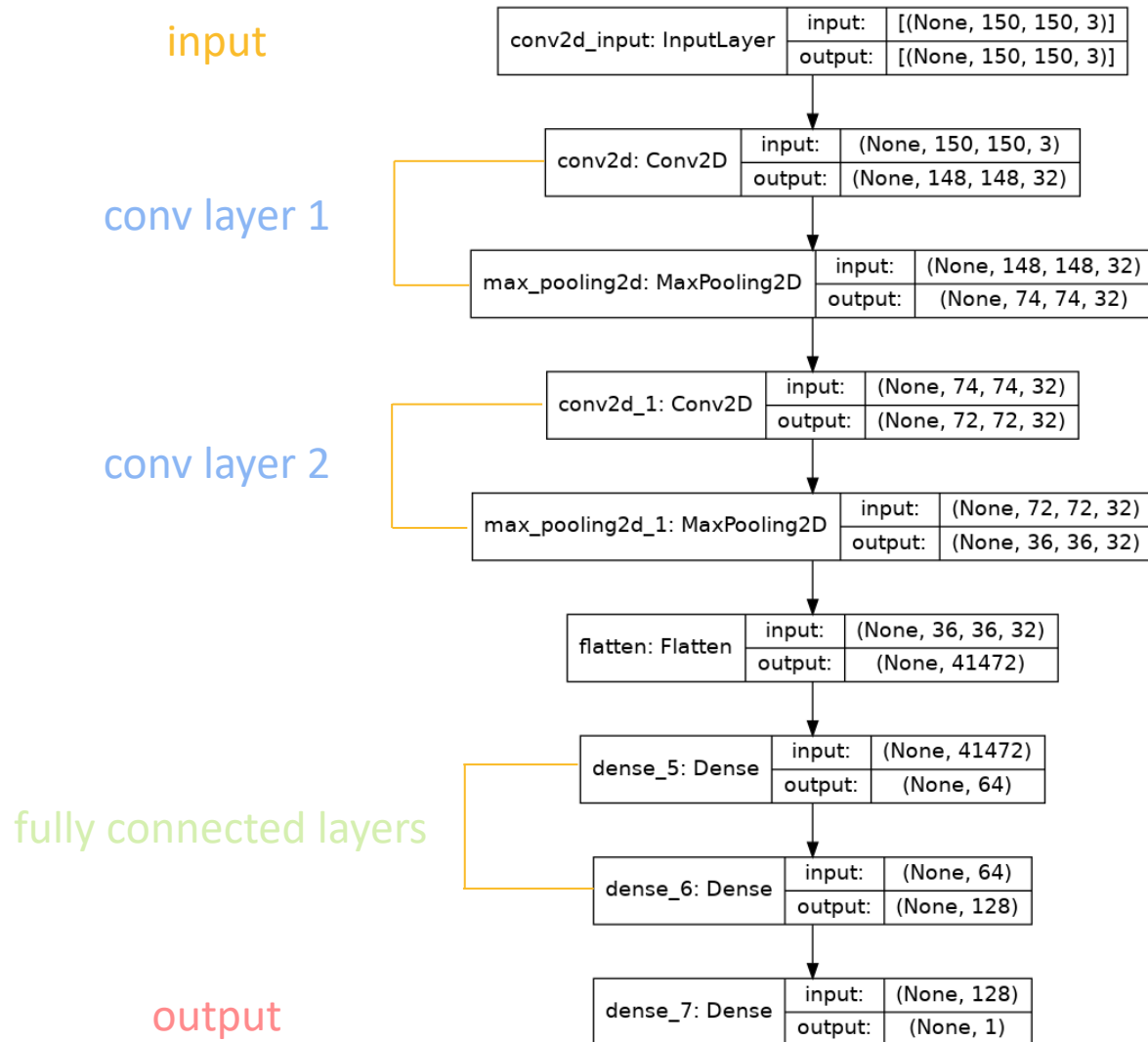
output

dense_4: Dense	input:	(None, 128)
	output:	(None, 1)

Convolutional Neural Network

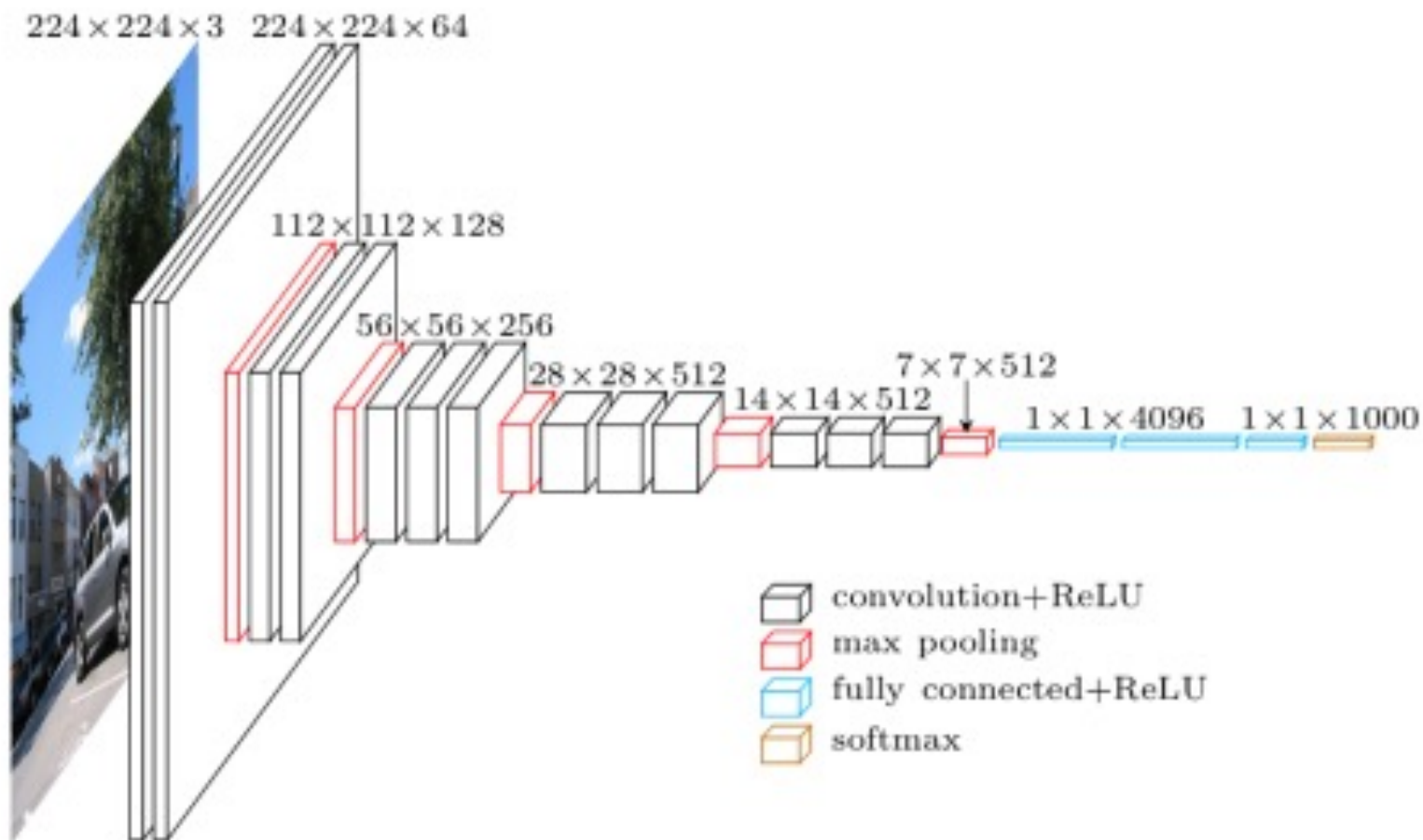


Convolutional Neural Network



Transfer Learning with VGG16

IMAGENET



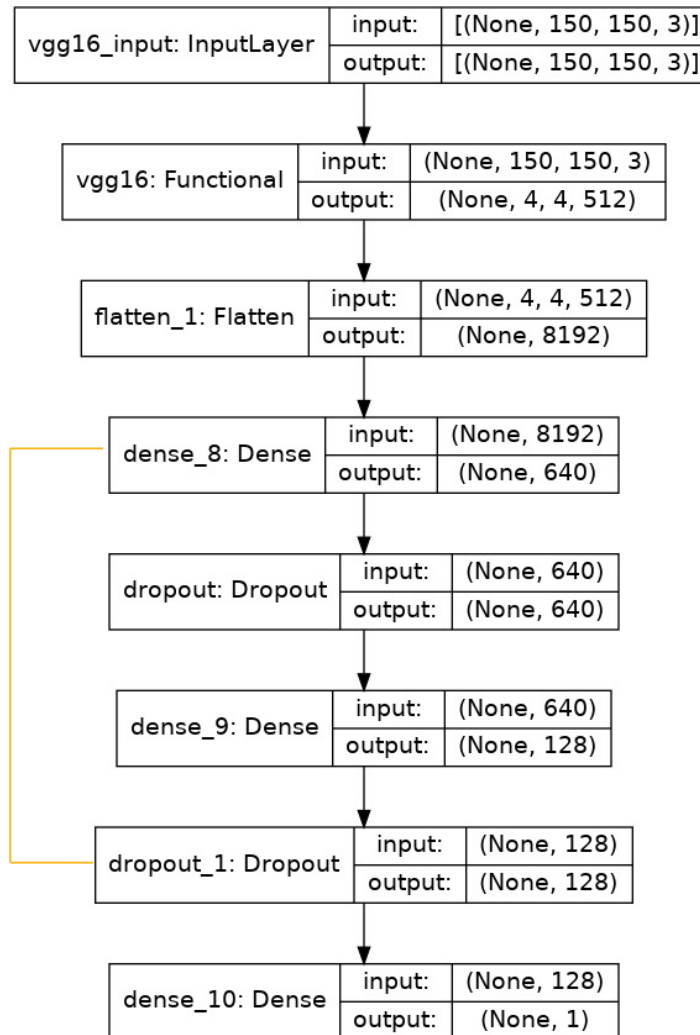
Transfer Learning with VGG16

input

VGG16

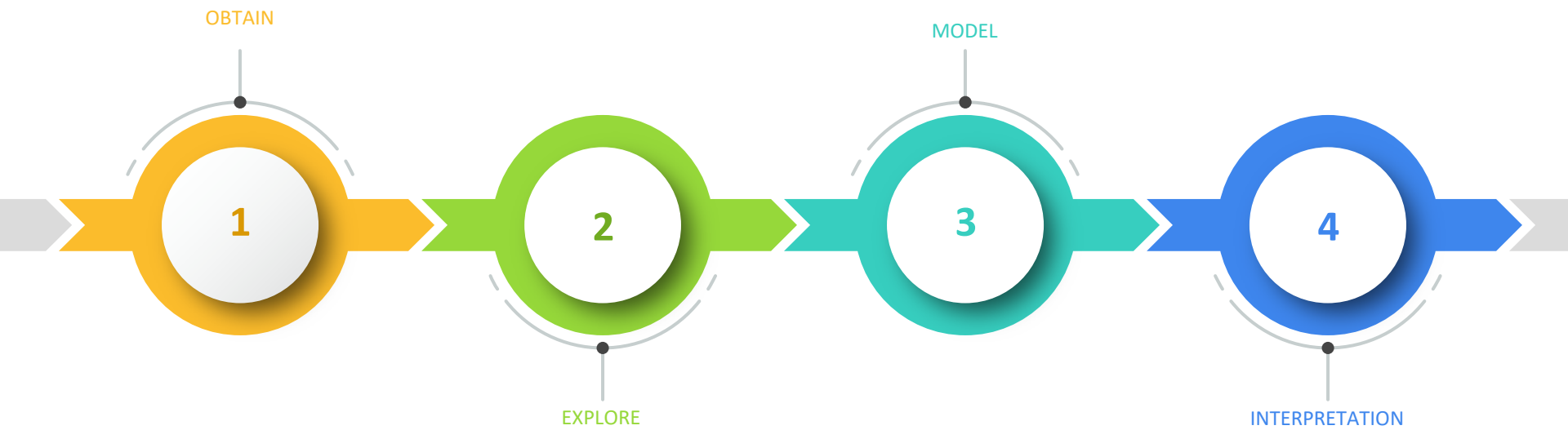
fully connected layers

output



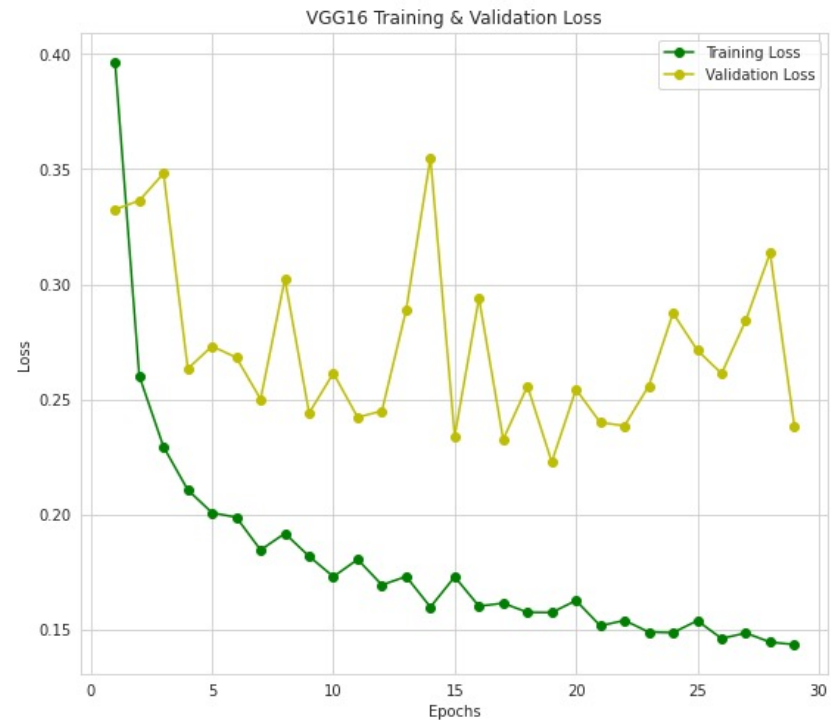
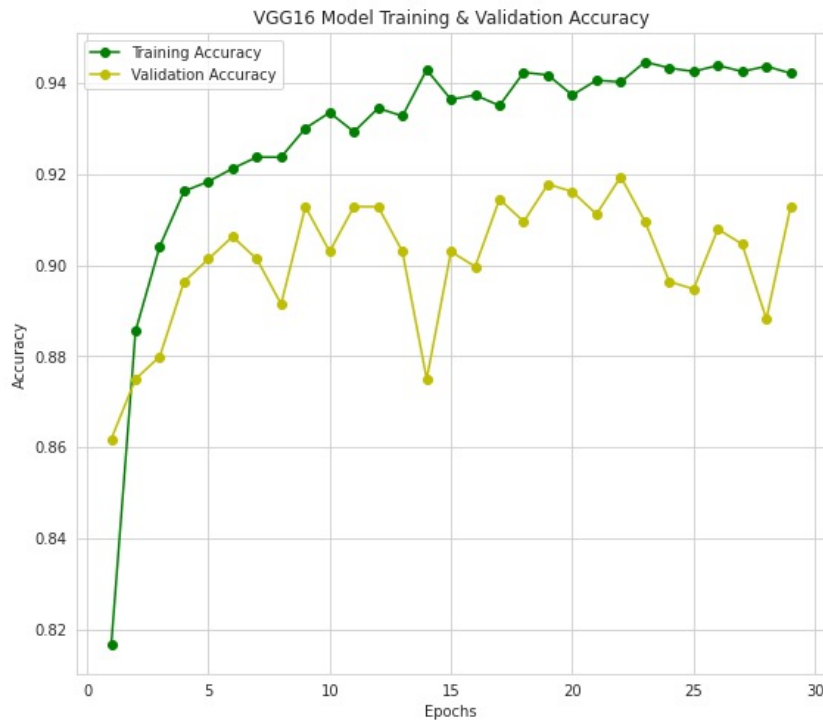
Summary of Key Findings

	Model	Accuracy	Precision	Recall	F1 Score	AUC
0	Multilayer Perceptron Model	0.79	0.84	0.74	0.75	0.74
1	Convolutional Neural Network Model	0.88	0.89	0.85	0.86	0.85
2	Transfer Learning: VGG16 CNN Model	0.92	0.91	0.90	0.91	0.90



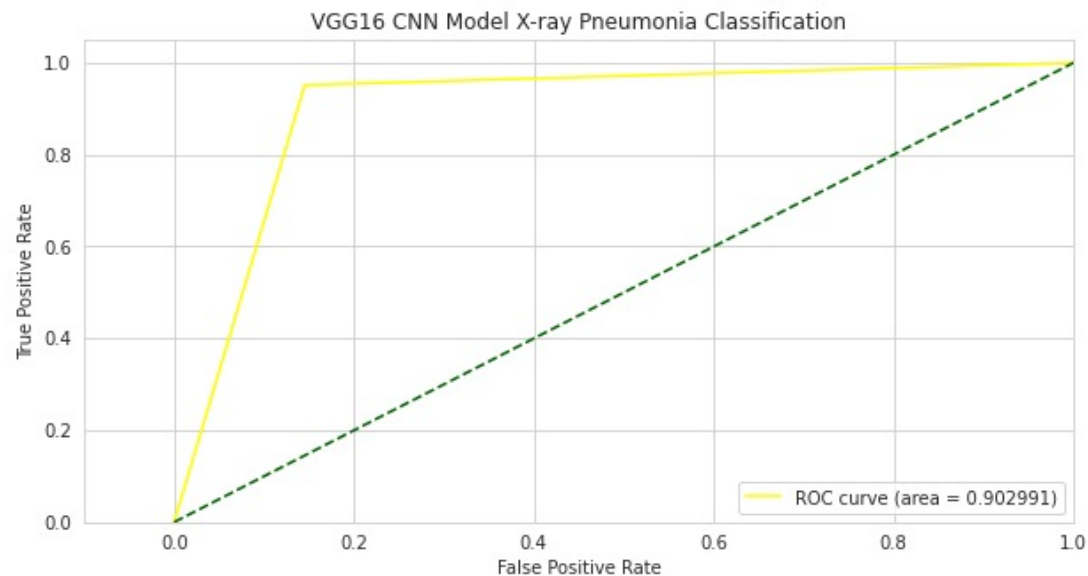
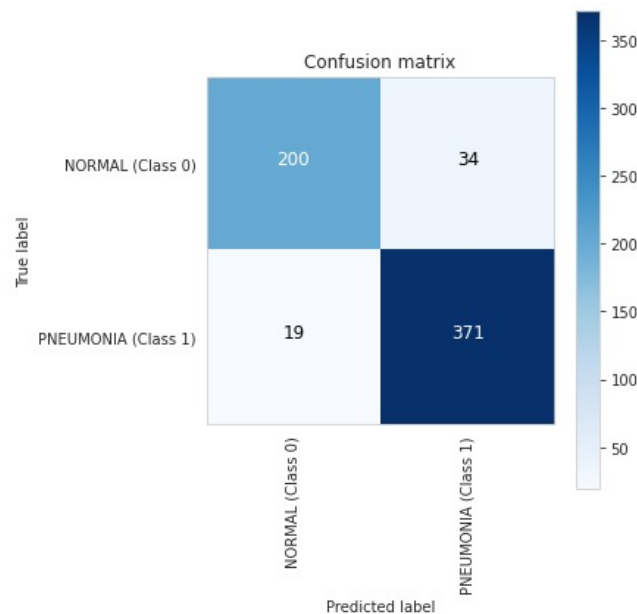
Best Model

Transfer Learning with VGG16



Train accuracy = 95%

Validation accuracy = 92%



Accuracy of 92%

Recall/sensitivity of 90%

Precision/specificity of 91%

FN < FP

The area under the ROC curve of 90%

Future Work

1. Build a multi-class classification model to distinguish between Normal, Viral Pneumonia, and Bacterial Pneumonia
2. Combine CNN models with other classifiers such as Support Vector Machine (SVM)
3. Tune parameters such as learning rate, batch size, optimizer, number of layers, types of layer, number of neurons per layer, and the type of activation functions for each layer. GridSearchCV or RandomizedSearchSV can be used to achieve this.

Recommendations

Pneumonia are largely preventable in this age group and failure in long-term management or preventing it will result in increases of the risk of developing chronic pulmonary disorders in later adult life. Chang (2013) suggests:

1. Focus on solving problems that are associated with increased risk of pneumonia such as overcrowding, access to clean water, malnutrition, anemia, young maternal age, low birth weight, and exposure to tobacco smoke and other environmental pollutants
2. Invest in resources to collect data systematically, especially in poor countries
3. Develop a universally agreed diagnostic gold standard for childhood pneumonia, especially one that can also differentiate between bacterial and non-bacterial pneumonia, which is currently a major limitation in clinical research in this area

The background features abstract, organic shapes in light blue and purple. A large, light blue circular shape is on the left, and a large, purple, wavy shape is on the right. The text "THANK YOU" is centered in a bold, blue, sans-serif font.

THANK YOU

The background features abstract organic shapes in shades of blue and purple. A large, light blue shape is on the left, and a large, dark blue to purple shape is on the right. A smaller, light blue shape is at the bottom left.

APPENDIX

Reference

Chang, A. B., Ooi, M. H., Perera, D., & Grimwood, K. (2013). Improving the Diagnosis, Management, and Outcomes of Children with Pneumonia: Where are the Gaps?. *Frontiers in pediatrics*, 1, 29. <https://doi.org/10.3389/fped.2013.00029>

Geron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras and TensorFlow: concepts, tools, and techniques to build intelligent systems* (2nd ed.). O'Reilly.

Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C., Liang, H., Baxter, S. L., McKeown, A., Yang, G., Wu, X., Yan, F., Dong, J., Prasadha, M. K., Pei, J., Ting, M., Zhu, J., Li, C., Hewett, S., Dong, J., Ziyar, I., Shi, A., ... Zhang, K. (2018). Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell*, 172(5), 1122–1131.e9. <https://doi.org/10.1016/j.cell.2018.02.010>

Unicef. (2021, April 07). Pneumonia in children statistics. Retrieved April 13, 2021, from <https://data.unicef.org/topic/child-health/pneumonia/#:~:text=A%20child%20dies%20of%20pneumonia,of%20the se%20deaths%20are%20preventable.>

