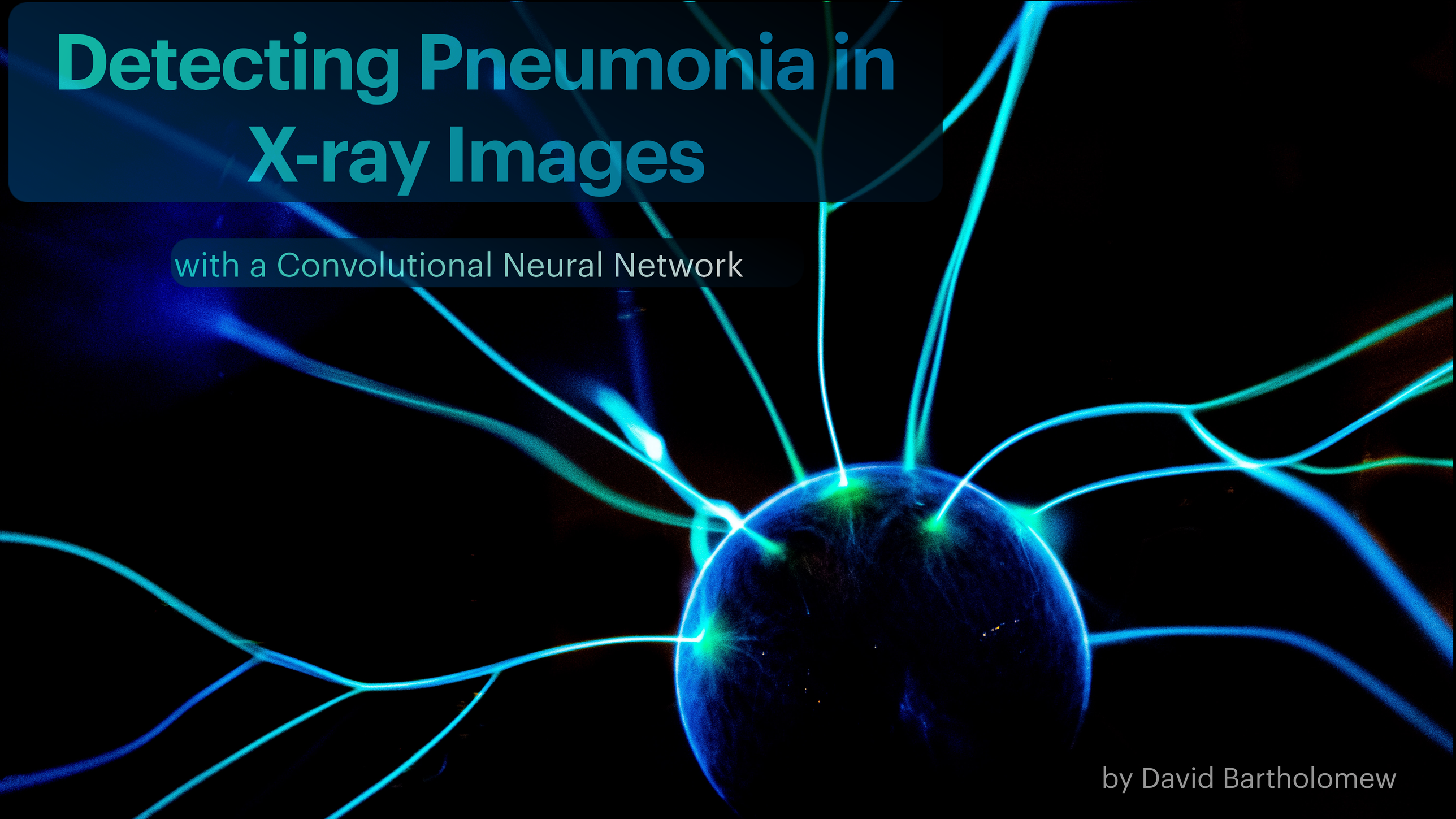
Detecting Pneumonia in X-ray Images

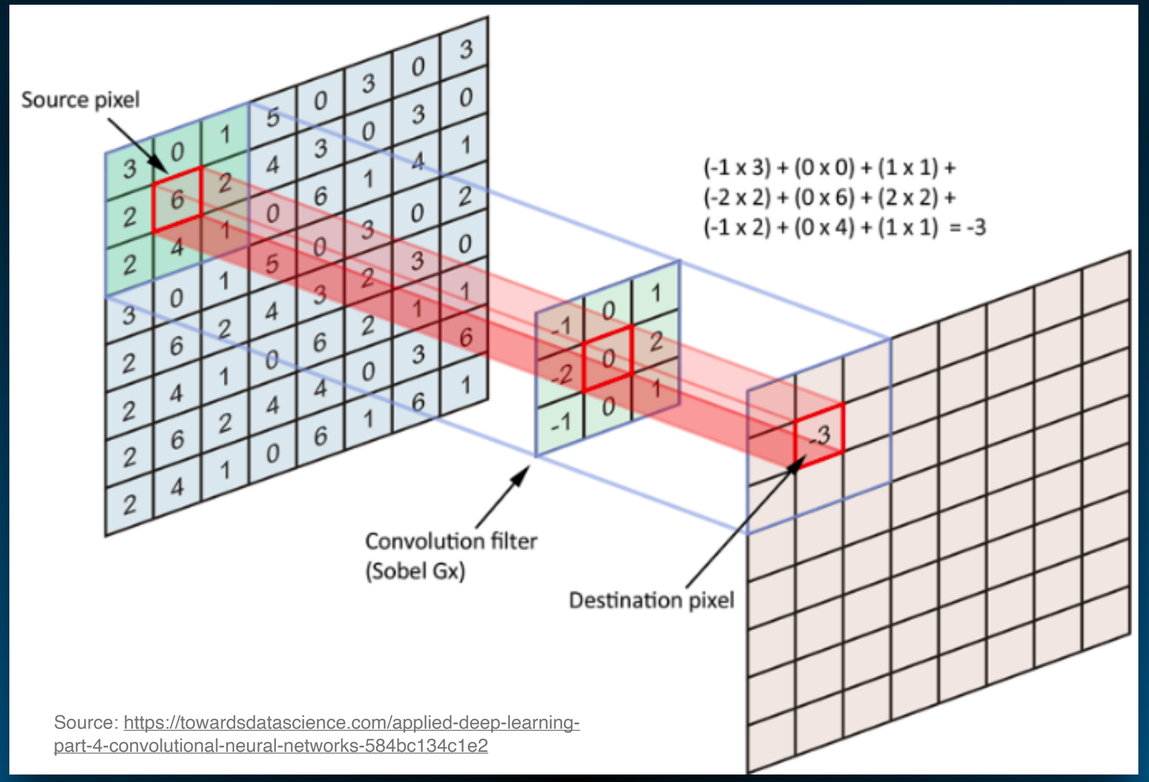
with a Convolutional Neural Network



by David Bartholomew

What is a convolutional neural network (CNN)?

- A CNN takes in the raw pixel values of an image and extracts features from the image such as edges or shapes by performing convolutional operations with specified filters. The CNN then builds feature maps, highlighting important features in an image.
- More plainly, it's a model commonly used in visual imagery to detect image features



Source: https://towardsdatascience.com/applied-deep-learning-part-4-convolutional-neural-networks-584bc134c1e2

Why use a CNN to detect pneumonia?

- Across the world, there is a shortage of radiologists
- In certain countries, due to staffing shortages, medical images are sent to be analyzed offsite to determine results
- Without a radiologist, relying on medical professionals without expertise in radiology could lead to misinterpreted results
- Getting quick, accurate results can be a difference-maker for certain patients

The Importance of Detecting Pneumonia

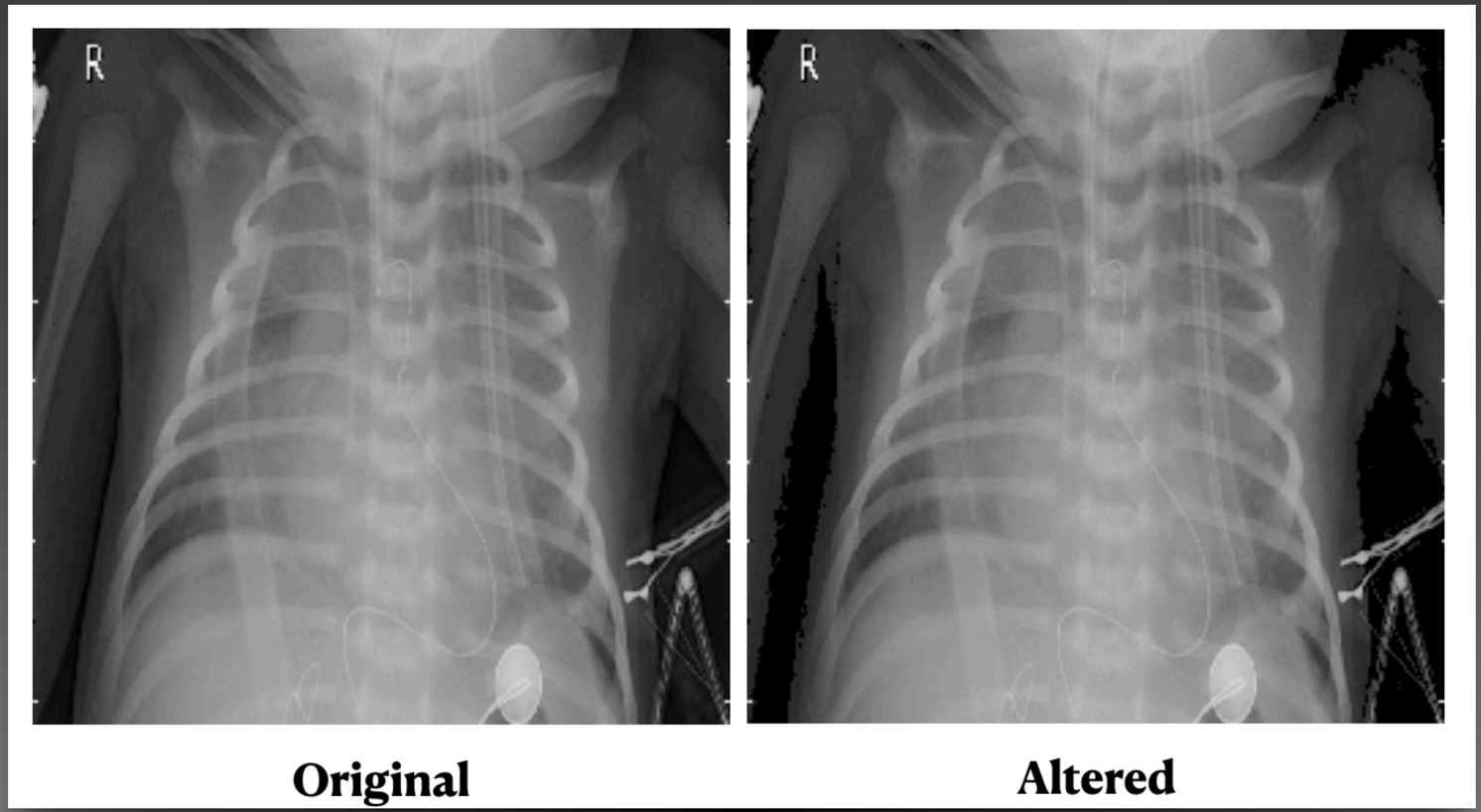
- The images in this model are of pediatric patients under 5 years old
- According to WHO, Pneumonia accounts for 15% of deaths for children under 5
- Pneumonia caused by bacteria can be treated with antibiotics, but only one third of children receive them
- Streamlining the process of accurately detecting pneumonia in children is a necessity, and it truly could save lives

The Data

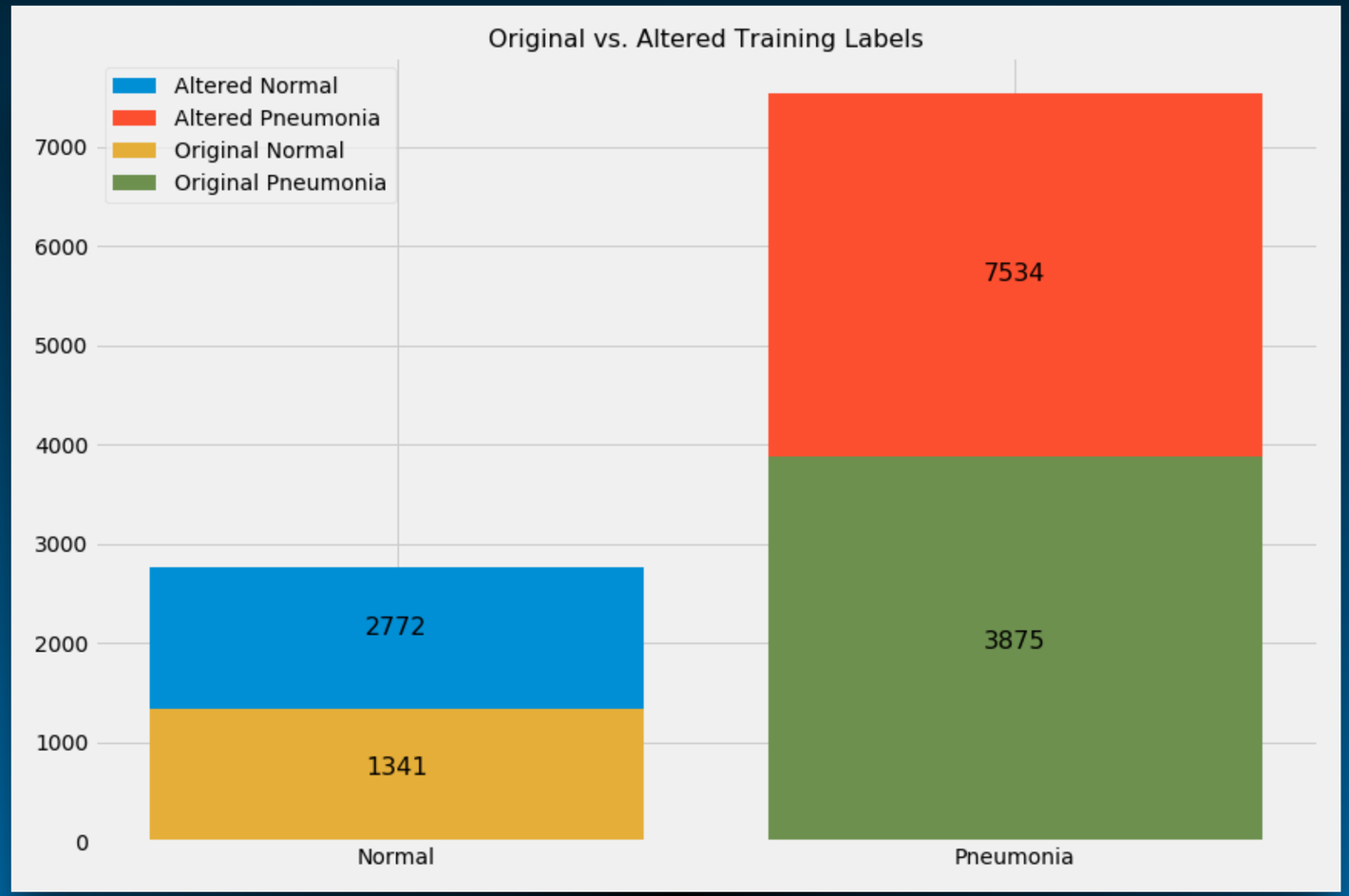
- The data for this project was downloaded directly from https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia
- It contains 5,856 x-ray images of pediatric patients with and without pneumonia
- All images were converted to true color images (RGB) and resized to 75x75x3 (pixels), so that the model could be run on a local machine

Data Augmentation

- Artificially increasing the size of training data to give the model more diversity in the data and generalize well to predicting new images. In the altered images, pixel values of under 40 were changed to 0. Essentially, this converted darker gray areas to black, highlighting more important features in the image and returning higher model accuracy. (Differences are subtle, yet effective).

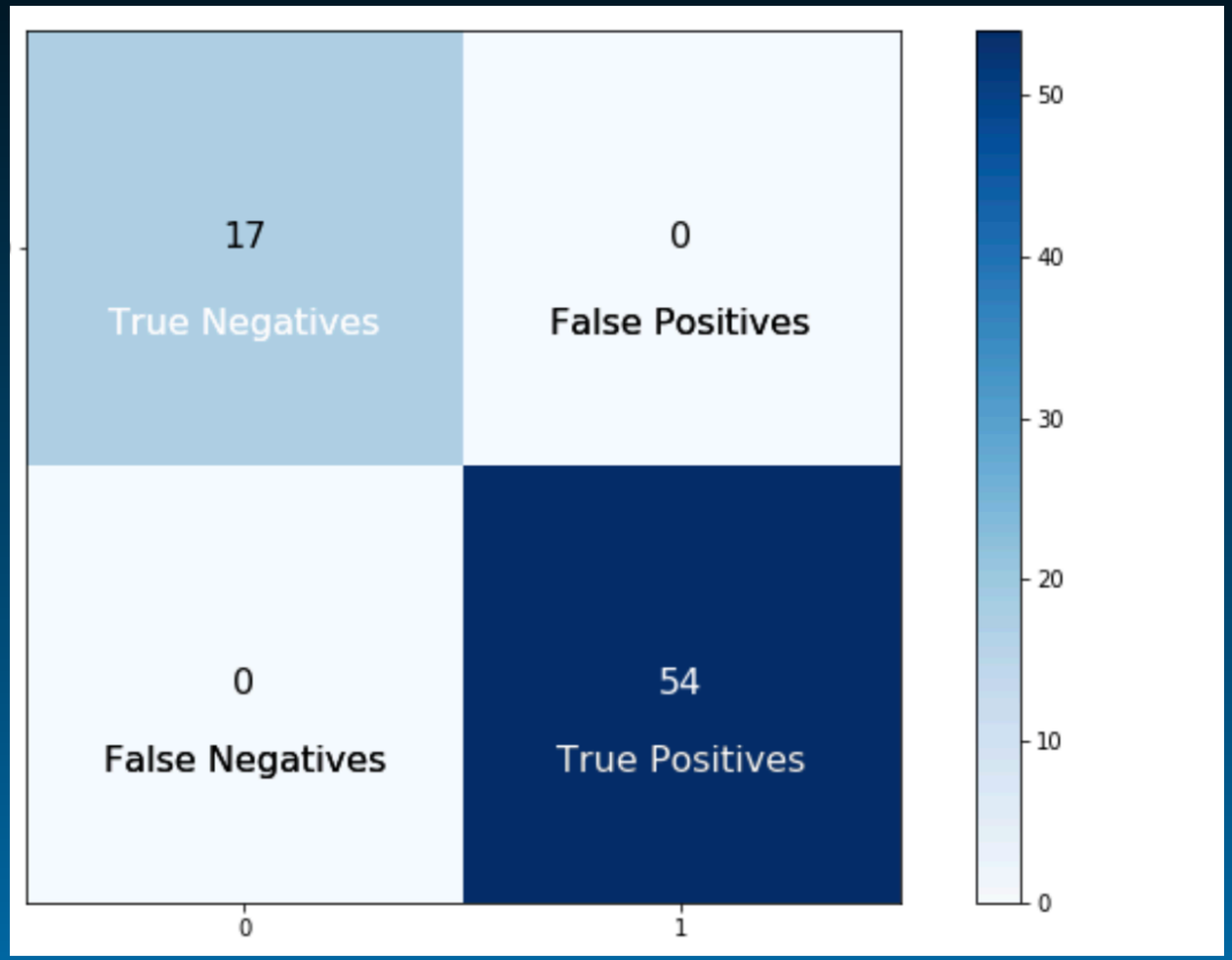


- Here we can see the original training data vs. the augmented training set with a breakdown of “normal” vs. “pneumonia”
- Additionally, the original data was combined and re-split to increase the size of the test set and get a better picture of overall accuracy



Final Model Results

- To achieve higher accuracy, transfer learning was used. A prebuilt model (Inception-ResNet-v2) was retrained and used as the starting point for the final model.
- Out of 71 images in the unseen test set, the final model resulted in 100% accuracy, classifying 17 true negatives and 54 true positives.



Conclusion

- A convolutional neural network can have a wide variety of applications
- In this case, we saw the need for a CNN model due to a shortage of radiologists and the need for getting quick, accurate medical image results
- To emphasize this need, we saw that 15% of children under 5 die with pneumonia
- This particular CNN model was built with x-ray images of pediatric patients. Using techniques like data augmentation and transfer learning resulted in a final model with 100% accuracy on the unseen test images.
- Additional analysis could include further classifying bacterial or viral pneumonia.

Project Link

- Full details of this project are available at this link:

https://github.com/dbarth411/dsc-mod-4-project-v2-1-online-ds-sp-000

References

- Marlee Long, “The Radiologist Shortage and The Potential of AI”, April 1, 2020, https://www.aidoc.com/blog/is-radiologist-shortage-real/
- Vincent Tatan, “Understanding CNN (Convolutional Neural Network)”, December 23, 2019, https://towardsdatascience.com/understanding-cnn-convolutional-neural-network-69fd626ee7d4
- World Health Organization, August 2, 2019, https://www.who.int/news-room/fact-sheets/detail/pneumonia

Thank you!