Memorandum

Date: February 21, 2025
To: The Gates Foundation
From: Zwiller Strategic Solutions (Thomas Zwiller)
Subject: Augmenting Pneumonia Diagnosis in Malawi

1. Executive Summary

The data science team at Zwiller Strategic Solutions (ZSS) has developed a Nuanced Convolutional Neural Network computer vision model with an overall accuracy of 94.43%. It can detect the top 30% of pneumonia cases with 100% accuracy. If deployed, the model could improve pneumonia detection in Sub-Saharan Africa by augmenting medical professionals. To properly deploy the model to Malawi, we request $2,000,000 USD to serve as a starter fund for a pilot program.

2. Background

Per the United Nations Children's Fund, a child dies of pneumonia every 43 seconds,[1] and it "kills more children than any other infectious disease, claiming the lives of over 700,000 children under five (5) every year, or around 2,000 per day."[1] The disease also disproportionately impacts those in Sub-Saharan Africa. In 2021, the three countries with the highest estimated annual death rate of children under five from pneumonia were South Sudan (292.7), Nigeria (270.3), and the Central African Republic (264.5)[2]. Part of the disproportionate effect can be attributed to malnutrition; children whose immune systems have been weakened by malnutrition or undernourishment are much more susceptible to contracting pneumonia.[3] According to the Economist, "some 19.1% of the African population are malnourished,"[4] resulting in a population that is generally more susceptible to contracting pneumonia. Another key risk factor is the presence of any pre-existing illnesses,[3] which can lead to an increased susceptibility to pneumonia. This risk factor is also prevalent in Africa; in 2022, 1.3 million children (ages 0-14) were estimated to be living with human immunodeficiency virus (HIV).[5] The World Health Organization (WHO) believes that the mother-to-child rate of transmission is anywhere from 15% to 45% and can occur as early as during pregnancy and as late as breastfeeding.[5]

Despite the increased prevalence of pneumonia in Africa, the disease is manageable. Diagnostically speaking, the disease can be detected in numerous ways; physicians can utilize chest X-rays, polymerase chain reaction tests (PCR tests), blood cultures, or bronchoscopies.[6] Once diagnosed, physicians can prescribe antibiotics or antiviral medicines.[7] These diagnostic tools are not a panacea, however. In 2022, researchers found that although 80% of the African population lives in rural areas, over 90% of African radiologists "work in urban or preurban settings."[8] While there are radiology facilities that can serve the rural population, the machines in these facilities are "often old and not adequately maintained or serviced."[8] As a result, diagnosis can be made harder by poor imaging, if it is even available.

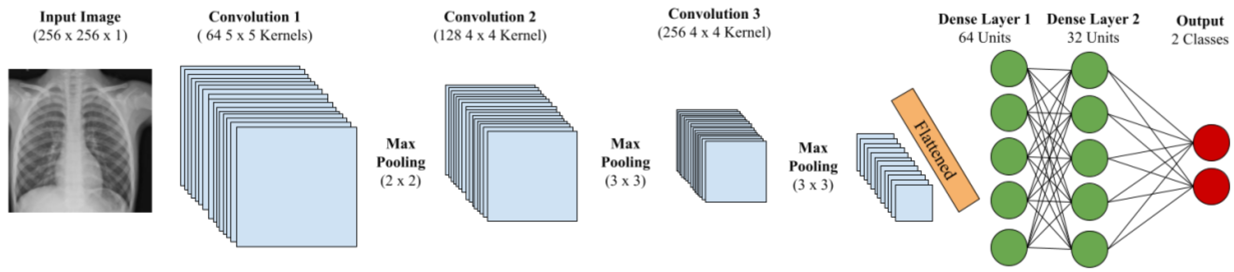
3. Analysis

Our team at ZSS began developing a deep-learning computer vision model by experimenting with a parameterized Convolutional Neural Network (pCNN). Initially, the team limited the number of epochs while exploring hyperparameters, such as the total number of filters, kernel size, and dense layers. Once the initial hyperparameter grid search was completed, a second dense layer was added, and the total number of epochs was increased to 20 to ensure the model had been adequately exposed to the data. After optimizing the model's hyperparameters, it achieved an area under the curve (AUC) of 98.14 and an accuracy of 94.96%.

Once the initial benchmark of 98.14 had been established, the team implemented data augmentation, which expanded the training dataset by altering the existing images. This involved shifting the images' height and width, flipping them horizontally and vertically, and rotating them. These enhanced images will help prepare the model for any poor-quality imaging it will encounter when deployed, contributing to an increased AUC score of 98.53 and an accuracy of 94.43%.

While the model's overall accuracy slightly dipped, the improved AUC increased its accuracy when dealing with the top 5 percent of healthy patients.

After the team developed a robust pCNN, we created a Nuanced Convolutional Neural Network (nCNN), which enabled us to customize each of the model's dense layers. The architecture of the final model is shown below:

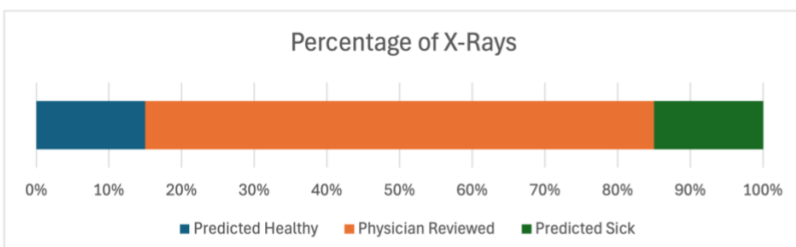


During training, the x-ray image[15] is broken down by the initial kernels before undergoing max pooling, which helps with processing. This image is repeated twice, with more kernels in each iteration. The resulting information is then flattened into a vector before being passed to the dense layers. These layers mimic the human brain's thinking process before activating one of the two output neurons.

The finalized version of the nCNN model achieved an AUC of 98.47 and an accuracy of 95.13% after undergoing 20 epochs of training. The version of the nCNN that utilized data augmentation saw the AUC improve to 98.66 while achieving an accuracy of 94.43%.

4. Model Cost-Benefit Analysis and Recommendation

Because none of the models achieved an AUC score of 99+, the team recommends that the model be used to augment the diagnosis process by making predictions for the top 30% of all sick diagnoses and the top 30% of all healthy diagnoses while leaving the remaining 70% to professionals. This model will speed up the diagnosis process for the attending physicians and allow them to predict cases that are less clear to the model.



Using this existing framework, the team ran a cost-benefit analysis on the four developed models and the option of allowing the attending physicians to continue making 100% of the diagnoses. The accuracy rates for the physicians were taken from a 2018 study, while the model's accuracy rates are from testing data.

Model	Cost per 1000 Cases	TN Accuracy	FP Accuracy	TP Accuracy	FN Accuracy
Non-Augmented pCNN	$54.569.14	99.42%	0.58%	98.52%	1.48%
Augmented pCNN	$54,347.38	99.42%	0.58%	99.42%	0.58%
Non-Augmented nCNN	**$54,190.52**	100%	0%	100%	0%
Augmented nCNN	$54,204.47	99.42%	0.58%	100%	0%
Naïve (Current Physician)[14]	$73,980.50	90.60%	9.40%	20.60%	79.40%

The cost assumptions and a more detailed pricing table are included in the appendix under the Pricing Breakdown and Pricing Table sections.

Based on the cost-benefit analysis, the team recommends using the non-augmented Nuanced Convolutional Neural Network Model, which achieved the best combined accuracy and AUC score. Combining the nCNN model with the augmentation strategy model will save approximately $19,789.98 per 1,000 cases diagnosed.

Appendix

Pricing Assumptions

The team's primary assumption for the cost-benefit analysis is that the X-rays are a sunk cost and, thus, not factored into any of the calculations. Since an X-ray is needed to make a prediction, all four categories require one, and any associated costs (the time with the doctor, the time spent taking the X-ray, and the model's cost prediction) would apply to all four categories.

The secondary assumption is that the cost-benefit analysis only pertains directly to the cost of pneumonia treatment. For example, a patient correctly diagnosed as not having pneumonia will likely have another illness, just not pneumonia. Because it is impossible to know what they have, the cost is assumed to be $0 because they do not have pneumonia.

The third and final assumption is that any costs will only relate to the top 30% of the predictions made by the model. Because the final iteration of the model did not reach an AUC of 99+, it is not advisable to make it a fully automated process as it does not correctly capture enough of the false positives in the training set. As a result, the process does not apply to all the cases but to the top 30% of healthy and those with pneumonia. All remaining cases will need to be examined by the prescribing doctor. Therefore, 70% of the model's cost is the cost for the physician.

Pricing Breakdown

True Positive: In Malawi, the most readily available form of penicillin is amoxicillin,[11] and current WHO guidelines suggest that the best treatment course for pneumonia is a five-day schedule, which would require two doses[9] of amoxicillin for five (5) days,[10] resulting in 10 pills. The current cost of one 250mg tab is $1.60 USD, resulting in a price of **$16.00 USD**.[11]

True Negative: As the x-ray cost is not accounted for, a correct negative diagnosis costs **$0.00 USD**.

False Positive: A false positive costs **$16.00 USD**, resulting in the same treatment plan as a true positive. Any costs associated with treating the actual condition are not considered, as they are not pneumonia-related.

False Negative: In the case of an incorrect negative diagnosis, the team assumes that the patient would need to be admitted to the hospital ($137)[12] shortly after being diagnosed and would likely need an extended form of the amoxicillin treatment ($32.00), plus an additional x-ray ($10.98)[13] to confirm that they do have pneumonia. The resulting total cost is **$179.98 USD**.

Pricing Table

		Accuracy Rate (Limited to top 30 % for Models)				Healthy		Sick		Healthy		Sick		Proportions Based on Test Population	
	Cost per 1000	TN ACC	FP ACC	TP ACC	FN ACC	TN Cost	FP Cost	TP Cost	FN Cost	TN Result	FP Result	TP Result	FN Result	Sick	Healthy
Non-Augmented pCNN	$54,569.14	0.9942	0.0058	0.9852	0.0148	$0.00	$16.00	$16.00	$179.98	149.39	0.87	148.04	2.22	500.87	499.13
Augmented pCNN	$54,347.38	0.9942	0.0058	0.9942	0.0058	$0.00	$16.00	$16.00	$179.98	149.39	0.87	149.39	0.87	500.87	499.13
Non-Augmented nCNN	**$54,190.52**	1	0	1	0	$0.00	$16.00	$16.00	$179.98	150.26	0.00	150.26	0.00	500.87	499.13
Augmented nCNN	$54,204.47	0.9942	0.0058	1	0	$0.00	$16.00	$16.00	$179.98	149.39	0.87	150.26	0.00	500.87	499.13
Naive (Current Doctor)	$73,980.50	0.906	0.094	0.206	0.794	$0.00	$16.00	$16.00	$179.98	453.79	47.08	103.18	397.69	500.87	499.13
	Models Costs use top 30% of healthy and sick and 70% of Physician Cost														

The pricing table utilizes the 2018 physician diagnosis accuracy as the base naïve model. The model rates all came from their respective confusion matrix diagrams and have been scaled to a test population of 100 patients. All model costs include 70% of the naïve model, as the process is augmentation, not automation.

Works Cited

1. *Pneumonia in Children Statistics - UNICEF Data*. UNICEF. (2024, November). https://data.unicef.org/topic/child-health/pneumonia/

2. Dadonaite, B., & Roser, M. (2024, January). *Pneumonia*. Our World in Data. https://ourworldindata.org/pneumonia

3. World Health Organization. (2022, November 11). *Pneumonia in Children*. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/pneumonia#:~:text=Pneumonia%20is%20the%20single%20largest,low%2Dtech%20medication%20and%20care.

4. Roche, C. (2023). *Breaking the Cycle of Chronic Child Malnutrition in Sub-Saharan Africa*. Economist Impact. https://impact.economist.com/perspectives/sites/default/files/economist_impact_breaking_the_cycle_of_chronic_malnutrition_in_ssa_feb_23.pdf

5. World Health Organization. (2022a). *HIV/AIDS*. World Health Organization. https://www.afro.who.int/health-topics/hivaids#:~:text=In%20the%20African%20region%2C%20an,live%20to%20their%20fifth%20birthday.

6. U.S. Department of Health and Human Services. (2022, March 24). *Pneumonia Diagnosis*. National Heart Lung and Blood Institute. https://www.nhlbi.nih.gov/health/pneumonia/diagnosis

7. U.S. Department of Health and Human Services. (2022b, March 24). *Pneumonia Treatment*. National Heart Lung and Blood Institute. https://www.nhlbi.nih.gov/health/pneumonia/treatment#:~:text=Antibiotics%20may%20be%20prescribed%20for,your%20pneumonia%20may%20come%20back.

8. Kawooya, M. G., Kisembo, H. N., Remedios, D., Malumba, R., Del Rosario Perez, M., Ige, T., Hasford, F., Brown, J. K., Lette, M. M., Mansouri, B., Salama, D. H., Peer, F., & Nyabanda, R. (2022, March 26). *An Africa Point of View on Quality and Safety in Imaging*. Insights into Imaging. https://pmc.ncbi.nlm.nih.gov/articles/PMC8959275/

9. World Health Organization. (2014). *Revised WHO Classification and Treatment of Childhood Pneumonia at Health Facilities*. Institutional Repository for Information Sharing. https://iris.who.int/bitstream/handle/10665/137319/9789241507813_eng.pdf?sequence=1

10. Ginsburg, A.-S., Mvalo, T., Nkwopara, E., McCollum, E. D., Phiri, M., Schmicker, R., Hwang, J., Ndamala, C. B., Phiri, A., Lufesi, N., & May, S. (2020, July 2). *Amoxicillin for 3 or 5 days for chest-indrawing pneumonia in Malawian children*. The New England Journal of Medicine. https://pmc.ncbi.nlm.nih.gov/articles/PMC7233470/

11. Khuluza, F., & Haefele-Abah, C. (2019, February 12). *The Availability, Prices, and Affordability of Essential Medicines in Malawi: A Cross-Sectional Study*. PloS one. https://pmc.ncbi.nlm.nih.gov/articles/PMC6372227/

12. Shin, H., Ngwira, L. G., Tucker, A., Chaisson, R. E., Corbett, E. L., & Dowdy, D. W. (2020, February 28). *Patient-Incurred Cost of Inpatient Treatment for Tuberculosis in Rural Malawi*. Tropical Medicine & International Health : TM & IH. https://pmc.ncbi.nlm.nih.gov/articles/PMC7658961/#:~:text=Despite%20free%20government%20services%2C%20inpatients,concentration%20index:%20%E2%88%920.279

13. MacPherson, P., Webb, E. L., Kamchedzera, W., Joekes, E., Mjoli, G., Lalloo, D. G., Divala, T. H., Choko, A. T., Burke, R. M., Maheswaran, H., Pai, M., Squire, S. B., Nliwasa, M., & Corbett, E. L. (2021, September 9). *Computer-Aided X-ray Screening for Tuberculosis and HIV Testing Among Adults with Cough in Malawi (The Prospect Study): A Randomised Trial and Cost-Effectiveness Analysis*. PLoS medicine. https://pmc.ncbi.nlm.nih.gov/articles/PMC8459969/

14. Uwemedimo, O. T., Lewis, T. P., Essien, E. A., Chan, G. J., Nsona, H., Kruk, M. E., & Leslie, H. H. (2018, April 9). *Distribution and Determinants of Pneumonia Diagnosis Using Integrated Management of Childhood Illness Guidelines: a Nationally Representative Study in Malawi*. BMJ Journals. https://gh.bmj.com/content/bmjgh/3/2/e000506.full.pdf

15. Generating Chest X-Ray Progression of Pneumonia Using Conditional Cycle Generative Adversarial Networks - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Example-images-from-the-Chest-X-Ray-Image-dataset_fig5_373207773

Link to final model