Predicting Child Mortality Rate from Satellite Imagery Using CNNs

Raaisa Moktader[1], Shayana Venukanthan[1], Tim Wu[1]

[1]Department of Computer Science, Stanford University

Background

- Progress towards the **United Nations' Sustainable Development Goals** (SDGs) has been hindered by lack of data on key indicators
- Recent advances in machine learning make it possible to use abundant data from satellites to **provide insight on progress toward SDGs**
- We hope to facilitate gauging progress towards SDGs **especially in remote, less accessible locations**

Figure 1: UN Sustainable Development Goals

Problem Statement

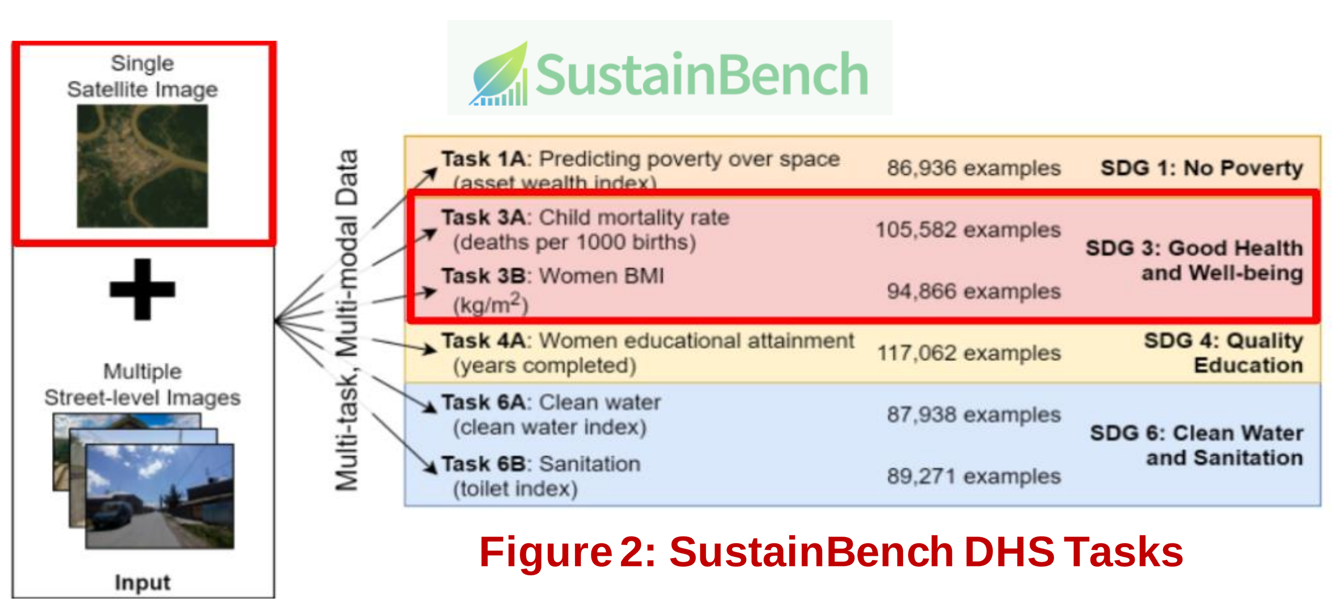

Figure 2: SustainBench DHS Tasks

- **Input:** Single 255 x 255 x 3px Satellite Image
  Output: Predicted child mortality rate (deaths per 1000 births)
- Apply a **pre-trained CNN model** to produce scores for range of values
- Take the **expectation** across all classes for final prediction
- Evaluate using the **Pearson's r^2 coefficient** of determination

Dataset

- **SustainBench dataset**
  - 56 countries
  - 105, 582 datapoints
  - Labelled with child mortality rates (5-166)
- **Raw Data:**
  - 255 x 255 x 8 px images
  - Bands: blue, green, red, SWIR 1, SWIR 2, thermal, NIR, nightlights

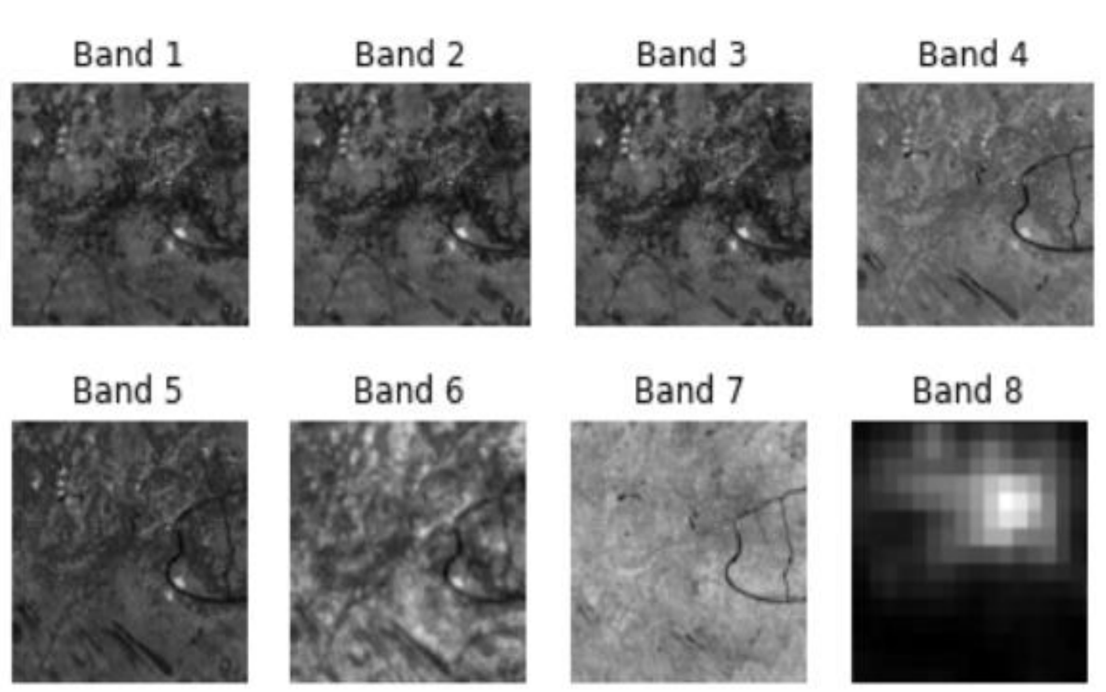

Figure 3: Grayscale images of 8 LandSat 5/7/8 bands

Methods

Landsat-8 Band Combinations

- Landsat-8 satellite image inputs have 8 bands, but we want 3-band inputs.
- Candidates for 3-band subsets were chosen based on what intuitively would affect child mortality rate. Four selected candidates shown below.
- Ran the model on all four candidates during hyperparameter tuning

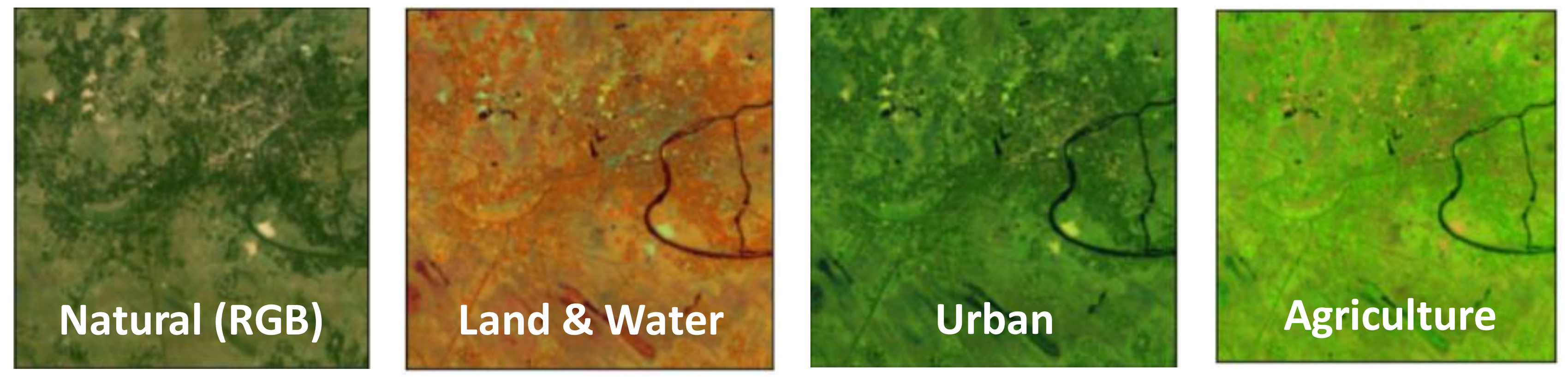

Figure 4: RGB representations of our chosen band combinations

Modified ResNet Model

- Transfer learning with ResNet-18 pretrained on ImageNet
- Add **Fully Connected (FC) layer** to output 167 scores
- Take expectation for final prediction
- **Loss function:** **Mean Absolute Error (MAE)**
- **Evaluation Metric:** **Pearson's r^2 coefficient** of determination

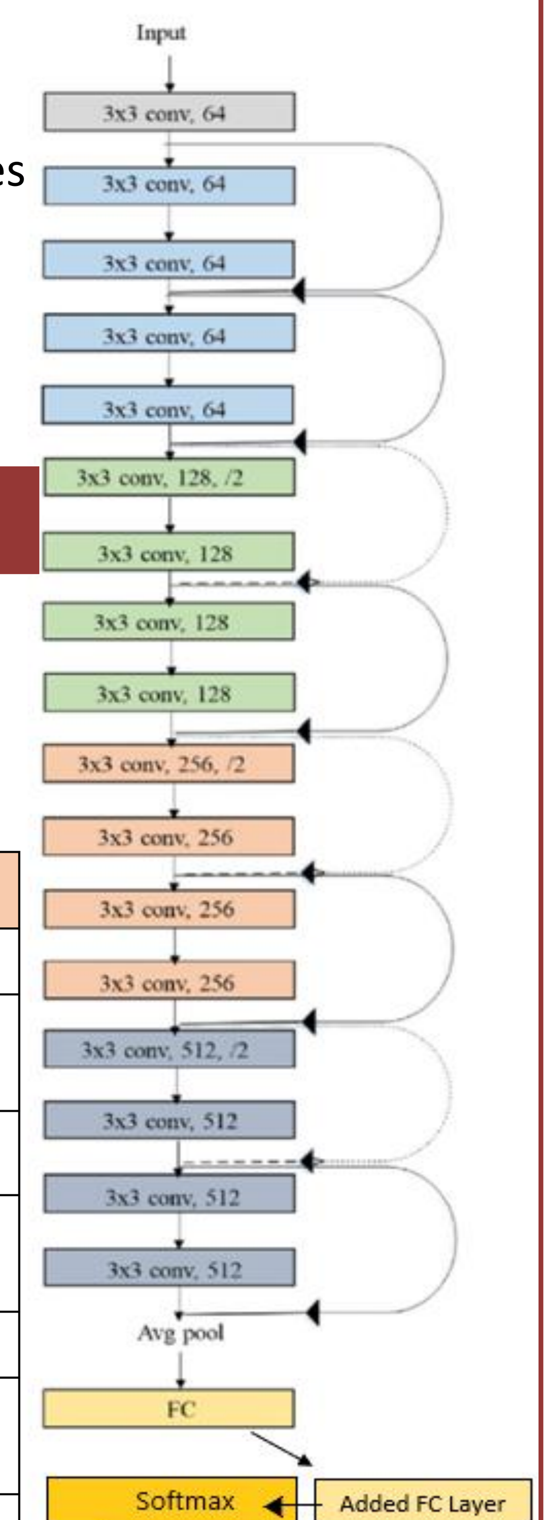

Figure 5: Our ResNet-18 + FC-167 Model

Experiment & Analysis

- Hyperparameters tuned, in order: learning rate, L2 regularization, ResNet model type, # of frozen ResNet layers, Landsat-8 satellite bands used.

Key Hyperparameter Tuning Results Summary	
Hyperparameters	**Val r^2**
Untuned ResNet-18 model w/ 0 frozen layers on RGB channels	0.000
Best *learning rate* (lr=1e-3)	0.1490
lr=1e-3, best *L^2 regularization* (weight decay=1e-3) on ResNet-18	0.1771
lr=1e-3, wd=1e-3 on *ResNet-34*	0.1091
lr=1e-3, wd=1e-3 on ResNet-18 w/ *6 frozen layers* (RGB channels) - Best Model	**0.1790**
lr=1e-3, wd=1e-3 on ResNet-18 (6 frozen layers), *Land & Water channels*	0.1636

- Each row above represents the model performance with the optimal value or an interesting value for each hyperparameter (italicized) we tuned

Final Results

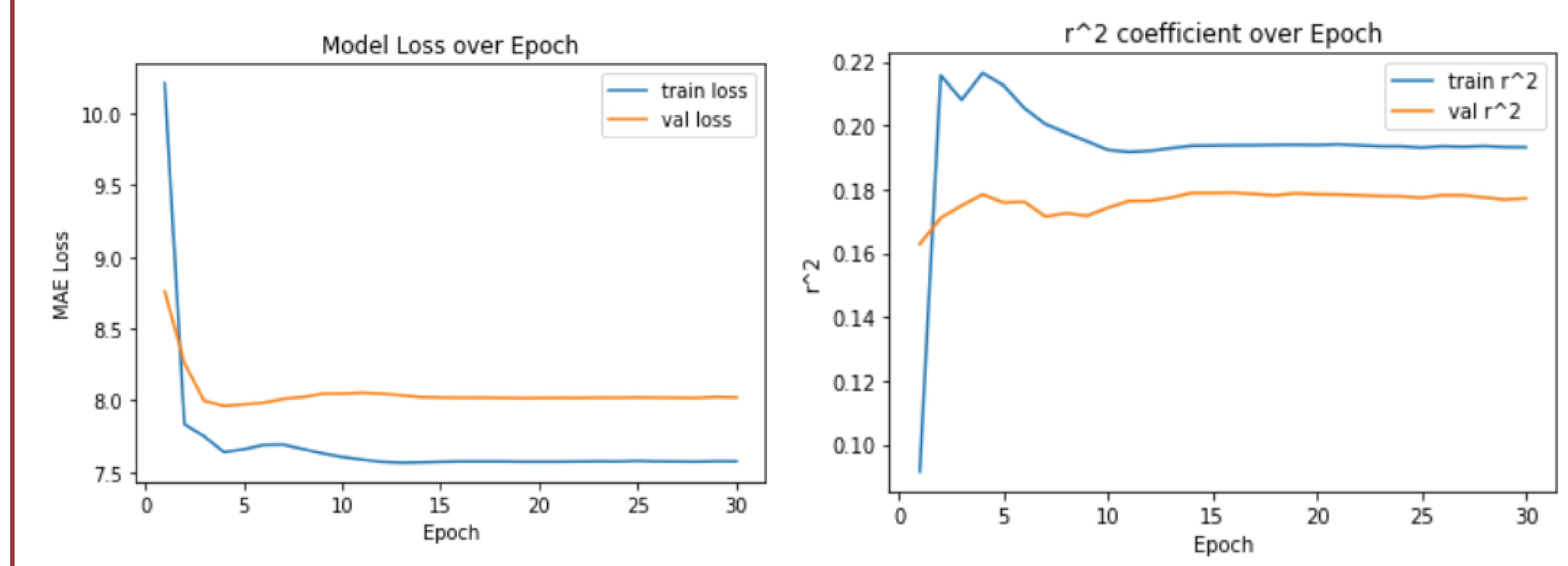

Figure 6: Results for our final modified ResNet-18 + FC-167 model

Models Comparison

Model Comparison		
Model	**Val r^2**	**Test r^2**
SustainBench KNN (Yeh et. al.)	0.0395	0.0700
CNN Baseline (Milestone)	0.0109	0.0052
ResNet-18 + FC-167 (Final Model)	**0.1790**	**0.0922**

Saliency Maps

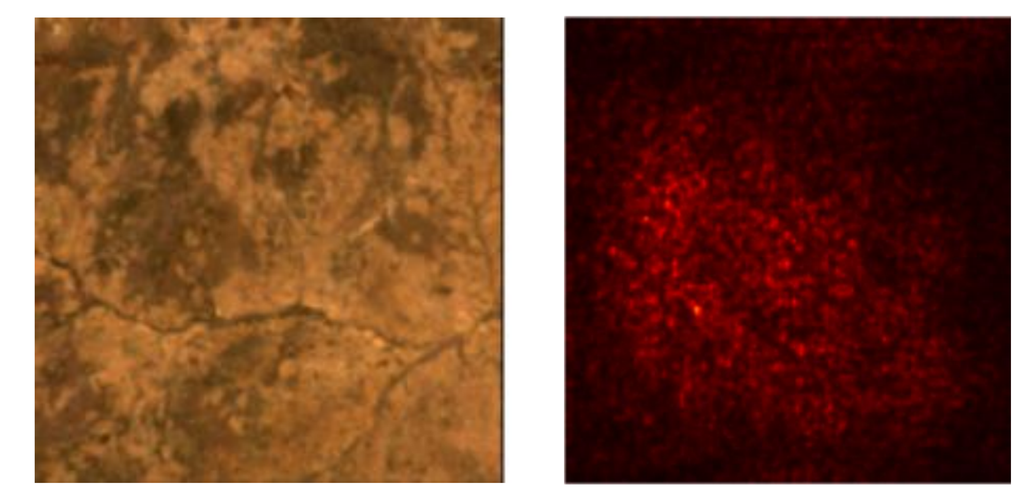

Figure 7: RGB image (left) and saliency map (right)

- Fails to truly capture the land's features such as rivers or vegetation
- Large room for improvement

Conclusions & Future Work

- Our modified ResNet-18 architecture is **state-of-the-art**, compared to SustainBench's KNN baseline model.
- Despite beating baseline, large drop between train, val r^2 vs. test r^2
- Future work:
  - More extensive search of optimal band combinations
  - Ensemble models: linear combination (with learnable weight parameters) of several ResNets across different band combinations.
  - Deeper neural networks: more attempts at ResNet-34, ResNet-50

References

Yeh et. al. , Sustainable development goals with machine learning. *In Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*, 2021