Predicting Child Mortality Rate from Satellite Imagery Using CNNs

Raaisa Moktader³, Shayana Venukanthan³, Tim Wu³

³Department 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.

Problem Statement

- 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 r² 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
- 8 Band Combinations

Experiment & Analysis

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

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

Evaluated using the Mean Absolute Error (MAE) for each hyperparameter (italicized) we tuned

Key Hyperparameter Tuning Results Summary

<table>
<thead>
<tr>
<th>Hyperparameters</th>
<th>Val r²</th>
<th>Test r²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Untuned ResNet-18 model w/ 0 frozen layers on RGB channels</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Best learning rate (lr=1e-3)</td>
<td>0.1490</td>
<td>0.1490</td>
</tr>
<tr>
<td>lr=1e-3, best L2 regularization (weight decay=1e-3) on ResNet-18</td>
<td>0.1171</td>
<td>0.1171</td>
</tr>
<tr>
<td>lr=1e-3, wd=1e-3 on ResNet-34</td>
<td>0.1091</td>
<td>0.1091</td>
</tr>
<tr>
<td>lr=1e-3, wd=1e-3 on ResNet-18 w/ 6 frozen layers (RGB channels) - Best Model</td>
<td>0.1790</td>
<td>0.1790</td>
</tr>
<tr>
<td>lr=1e-3, wd=1e-3 on ResNet-18 (6 frozen layers), Land &amp; Water channels</td>
<td>0.1636</td>
<td>0.1636</td>
</tr>
</tbody>
</table>

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

Final Results

- Model Comparison
  - 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

- 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 r² vs. test r²
- 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