Using satellite imagery, transfer learning, and [redacted] survey data to predict poverty in Southwest Asia

An expansion of the World Bank's *Predicting Poverty from the Sky* (2021)

Laniah Lewis, Taymour Siddiqui, Michael Chan, Rus Adamovics-Davtian

September 3, 2021

California Polytechnic State University

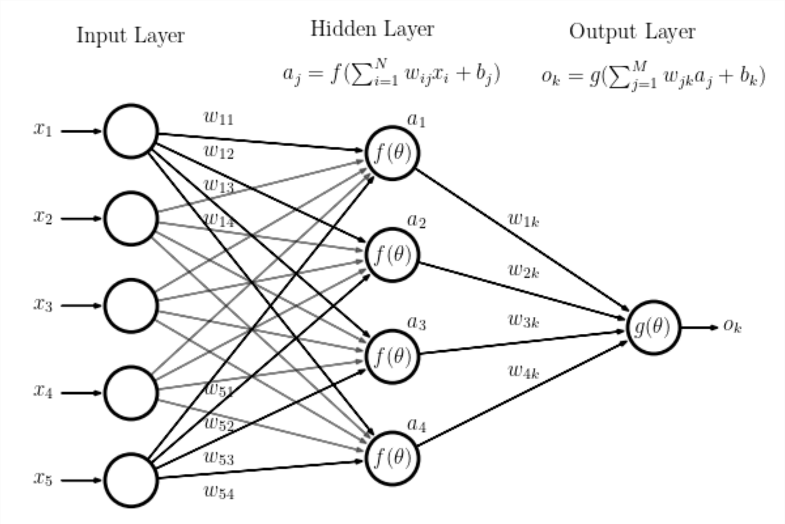
Table of contents

Introduction

Motivation

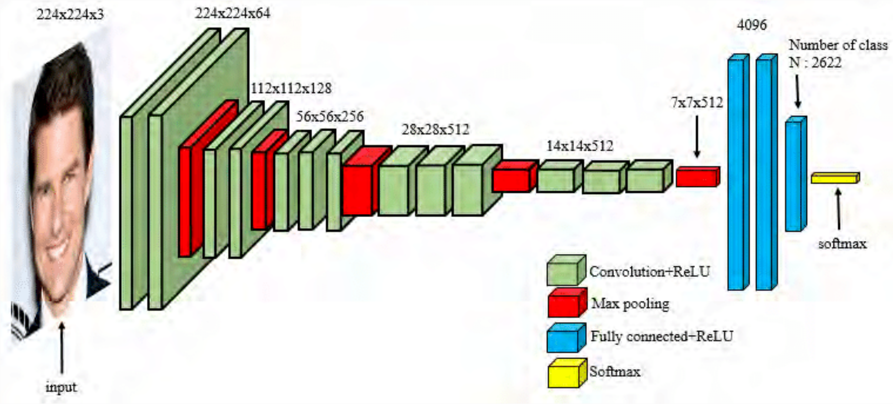
- More than 9.2 percent of the world lives in extreme poverty on less than 1.9 dollars per day.
- Measuring poverty allows us to target resources where they are needed most and most effectively reduce poverty.
- Much new research has focused on how satellite imagery can be used as a cheap way to achieve updated measures of poverty, as opposed to, traditional methods like surveys.
- Our paper focuses on how we can make these models more accurate and seeing how generalizable the methods are to a new geographic area.

What is a Neural Network Model?



https://www.astroml.org/book_figures/chapter9/fig_neural_network.html

What is a Convolutional Neural Network Model (CNN)?



https://www.cs.toronto.edu/~lczhang/aps360_20191/lec/w04/arch.html

Basics of our approach: Transfer Learning

- Transfer Learning- a model trained on one task is re-purposed on a second related task.
- Transfer learning allows us to do advanced deep learning techniques without having millions of images, a lot of time, or deep expertise.
- We make use of a pre-trained CNN that is able to learn from the satellite data that we have.
- The CNN takes in the satellite image data that we have and selects features that are helpful in predicting nighttime lights.
- The features are then used to predict poverty using machine learning.

Research Question

The 3 main questions we hoped to answer with our research were:

1. Can our improvements to the model achieve higher accuracy in predicting poverty on a binary level? The improvements being incorporating multispectral imagery, expanding our gridsearch, and tuning the parameters of the model(more on this later)
2. Can our model also be used to accurately predict poverty on a continuous level?
3. Can our machine learning model be used to predict other economic variables related to well-being besides poverty?

Literature Review: Origins of the method

- Cash transfers are one of the most common methods used to fight poverty in developing countries.
- However, these cash transfer programs have been shown to be most effective when impoverished groups are identified correctly.
- It is also difficult for organizations to measure the long term effects without this method.
- (Henderson2012) described how economic activity can be measured from space and introduced US Air Force Weather Service satellite night-lights data as a useful proxy for economic activity, providing a new way to measure poverty

Literature Review: Development of the method

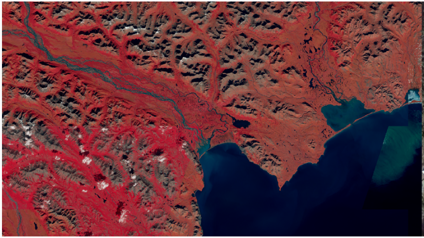
- (Jean2016) and (Yeh2020) have shown the validity of this method in predicting poverty in African Countries.
- (Helber2017) tested the accuracy and effectiveness of incorporating multi-spectral imagery into pre-trained neural networks and found that it maintained fairly high classification accuracy.
- Recent research discussed by (Burke2021) focused on creating powerful machine learning models that use satellite images as inputs to predict wealth indexes.

Data Summary

Daytime light intensity

LANDSAT 8 Satellite Image Data

- 50,583 48x48 Daytime Images
- Red, Green, and Blue (RGB) Bands for standard images
- 4 different non-visible spectral bands also available




Nighttime light intensity

- Visible Infrared Imaging Radiometer Suite (VIIRS) images
- Used k-means clustering to discretize night time radiance into 3 bins
- Over 1.8 million averaged night time light radiance values in 2014
- Down sampled the two majority classes to size of minority class
- Each bin contains 16,861 nighttime light values for a total of 50,583 values

[redacted] Survey

- Panel data with anonymized household ID
- Detailed demographic information available. 186 columns of information.
- Geospatial data for exact home locations of [redacted] participants.
- Poverty index and asset variables. For example: Does this household own a television, fan, refrigerator, etc.

Summary of [redacted] Data

Table 1: [redacted] Survey Data, Mean Statistics

Age Statistics Per HH	*Oldest Person*	*Youngest Person*
	55.72	3.6
Localities	*Urban*	*Rural*
	25.43%	74.57%
Unconditional Cash Grant	*Proportion of Recipients*	*Proportion of Non-Recipie*
	[redacted]%	[redacted]%

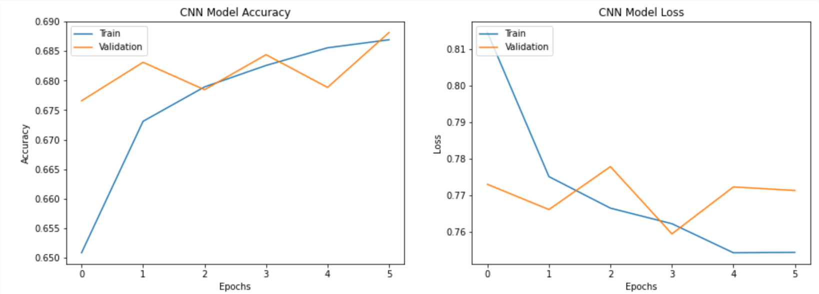
HH = household.

Problem Description

Original Model (Pre-Trained VGG 16)

Layer	Feature Map	Size	Activation	Type
Input Image	1	48 X 48 X 3	Relu	VGG 16
2 X Convolution	64	48 X 48 X 64	Relu	VGG 16
Max Pooling	64	24 X 24 X 64	Relu	VGG 16
2 X Convolution	128	24 X 24 X 128	Relu	VGG 16
Max Pooling	128	12 X 12 X 128	Relu	VGG 16
3 X Convolution	256	12 X 12 X 256	Relu	VGG 16
Max Pooling	256	6 X 6 X 256	Relu	VGG 16
3 X Convolution	512	6 X 6 X 512	Relu	VGG 16
Max Pooling	512	3 X 3 X 512	Relu	VGG 16
3 X Convolution	512	3 X 3 X 512	Relu	VGG 16
Max Pooling	512	1 X 1 X 512	Relu	VGG 16
Global Max Pooling	512	512	Relu	VGG 16
FC Flatten	512	512	Relu	Custom
FC Dense	100	100	Relu	Custom
Dropout	100	100	Relu	Custom
Output	3	3	Softmax	Custom

Original Model Performance



Incorporation of more spectral bands

- Spectral bands are captured image data that represent a specific wavelength range(RGB,Infrared,etc).
- Incorporating all spectral bands that we have available, capture all of the important features that we have available in our data.
- To incorporate all bands, we estimate a total of 5 CNN models, one for the RGB values and 4 models for each of the rest of the 4 bands being plugged into the CNN individually.
- We then combine all the features from these 5 CNN models and use machine learning to predict poverty using these features.

Hyperparameter adjustments of pre-trained CNN model predicting night-time lights

- Froze Pre-Trained CNN Layers
- Experimented with number of nodes, number of hidden layers, drop out rates, batch normalization, optimization function, type of pre-trained model, batch size, metric to optimize
- Original model architecture worked best

Grid Search Optimization with Machine Learning Models

- Expanded binary grid search
  parameters, ranges of model parameters
- Created grid search for continuous measures of poverty
  regression models, model parameters, ranges of model parameters

Results

NTL predictions

Inputs	Accuracy	Precision	Recall	F1-Score
Band 1	66.5 %	68.3 %	66.9 %	66.3 %
Bands 2,3,4 (RGB)	68.8 %	69.0 %	69.2 %	68.8 %
Band 5	60.9 %	61.3 %	60.6 %	61.0 %
Band 6	61.6%	65.4%	61.8%	61.1%
Band 7	64.5 %	64.4 %	64.4 %	64.4 %

* Trained on ~40,000 DTL images to predict ~10,000 NTL

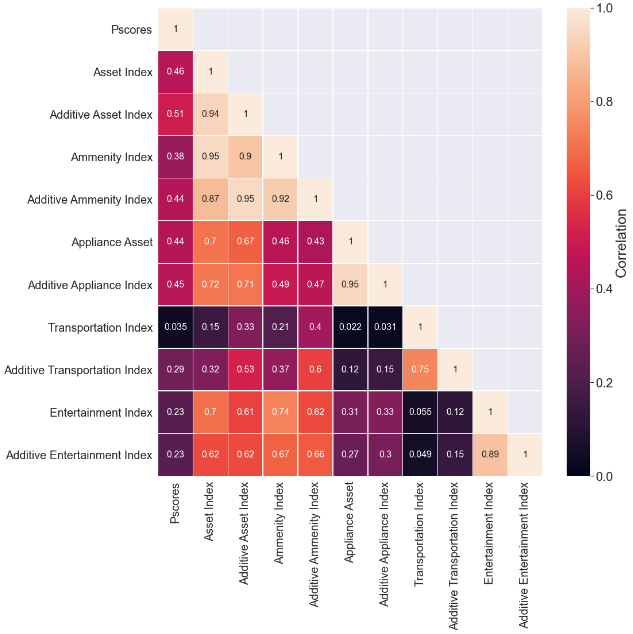
Poverty estimation

Table 2: Machine Learning Model Performance Metrics Prior To & Following Contributions

Binary	*Prior*	*Following*
Accuracy	.449	.487
Precision	.452	.465
Recall	.964	.8922
F1	.616	.6121
Continuous		
R2	N/A	.0319
MSE	N/A	134.7657
Correlation	N/A	.18174

Machine learning model predicts poverty on a binary and continuous level

Correlation matrix between poverty score and asset indices



Asset index measures

Table 3: Asset Indices Performance Metrics

	R2	*MSE*	*Correlation*
Main	.0774	.412	.2785
Additive Main	.0589	3.5188	.2511
Additive Appliance	.144	.4369	.395

Poverty Estimation using individual assets

Table 4: Performance Metrics of Fan Asset

	Accuracy	*Precision*	*Recall*	*F1*
Fan Asset	0.9126	0.9184	0.9918	0.9537

- 10.95% of households did not have fans (10.94%)

Robustness

- For our model to be robust, we want consistent outputs when making changes.
- Hyperparameter of batch size value was changed to test for stable results.
- 1000, 64, 125, 256, and 300 were used and the CNN model did not seem to be sensitive to these changes.
- Which is good! Everything depends on this first-stage model.

Conclusion

Main takeways

- COVID-19 halted the progress in decreasing poverty levels
- Extended current research using 2-stage modeling approach to predict poverty
- Achieved a 3.8% overall validated accuracy improvement (44.9% to 48.7%) in predicting binary measure of poverty
- Modified framework to accommodate predicting a continuous poverty measure (R^2 = .0319)

Shortcomings

- No significant improvements to CNN models compared to pre-trained base model
- Limited image training data ($\approx$ 40,000) due to sub-sampling and keeping balanced NTL bins
- Extremely computationally expensive

Future Research

- Continue to use data science approaches in predicting poverty levels and not strictly traditional methods
- Implement similar two-stage modeling technique on additional, richer data sources
- Have access to higher computing power
- Other measurements of well-being

Policy Implications

- [redacted]
- Other conditional and unconditional support programs
- Use when surveying is unfeasible.

Acknowledgments

A huge thank you to:

- Professor Katya Vasilaky
- World Bank team member Rob Marty
- World Bank team member Alice Duhaut
- California Polytechnic DxHub

Questions?