# Semi-Supervised Multitask Learning on Multispectral Images using W-GANs for Predicting Poverty

Anthony Perez, Swetava Ganguli
CS 231N Convolutional Neural Networks | Stanford University

## Introduction & Problem Statement

**Target Application:**

Obtaining reliable data describing economic livelihoods at a granularity that is informative to policy-makers requires expensive and logistically difficult surveys, particularly in the developing world. We would like to use multi-spectral satellite images from LandSat7 to predict poverty related metrics such as an Asset Wealth Index (AWI), nightlights, population density, distance to nearest road, land cover type, etc.

**Research Question:**

Does Multi-task learning that incorporates a WGAN-GP loss allow us to overcome the scarcity of labels to perform semi-supervised learning?

**Previous Approaches:**

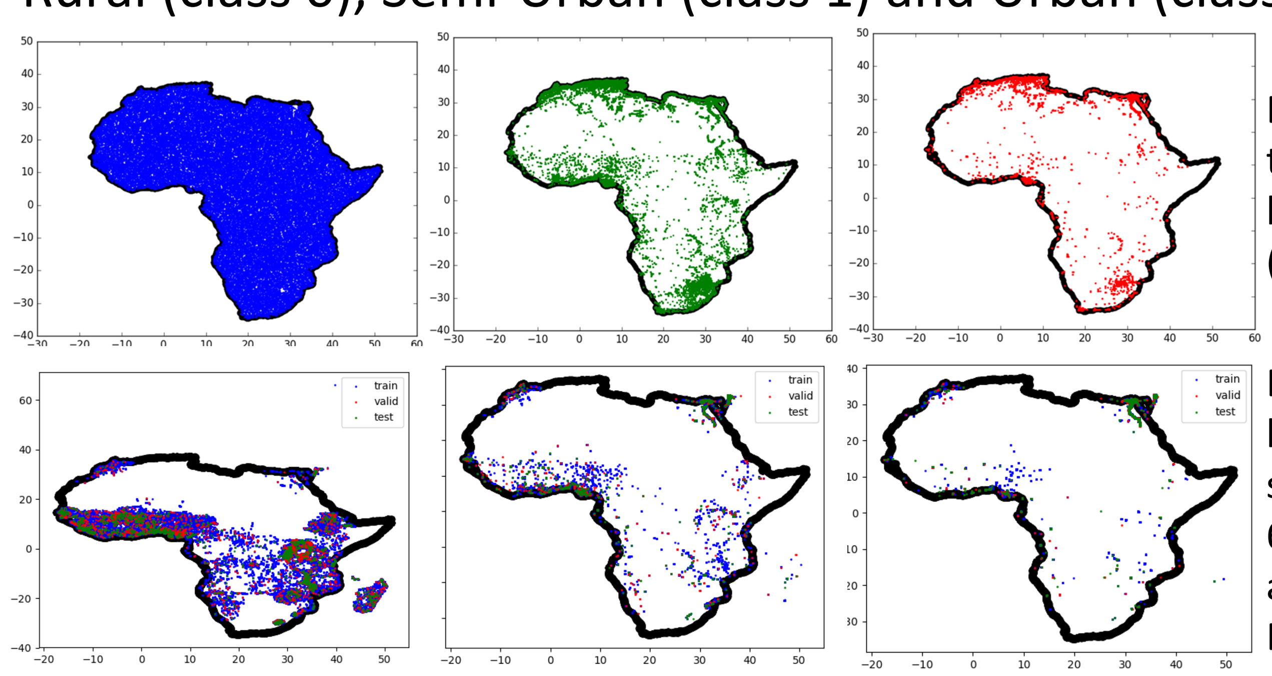
1. [1] use transfer learning to predict nightlight intensity
2. [2] propose the gradient penalty method for stable training of WGANs
3. [3] propose semi-supervised GAN training

**Problem Statement:**

Given a small number (5%) of labeled satellite images and large number (95%) of unlabeled satellite images, we use the semi-supervised loss in [3] across multiple tasks simultaneously and the WGAN-GP loss proposed in [2] as an additional task for stable training of the generator. Tasks are weighted proportional to their importance in predicting the asset wealth score.

## Datasets & Processing

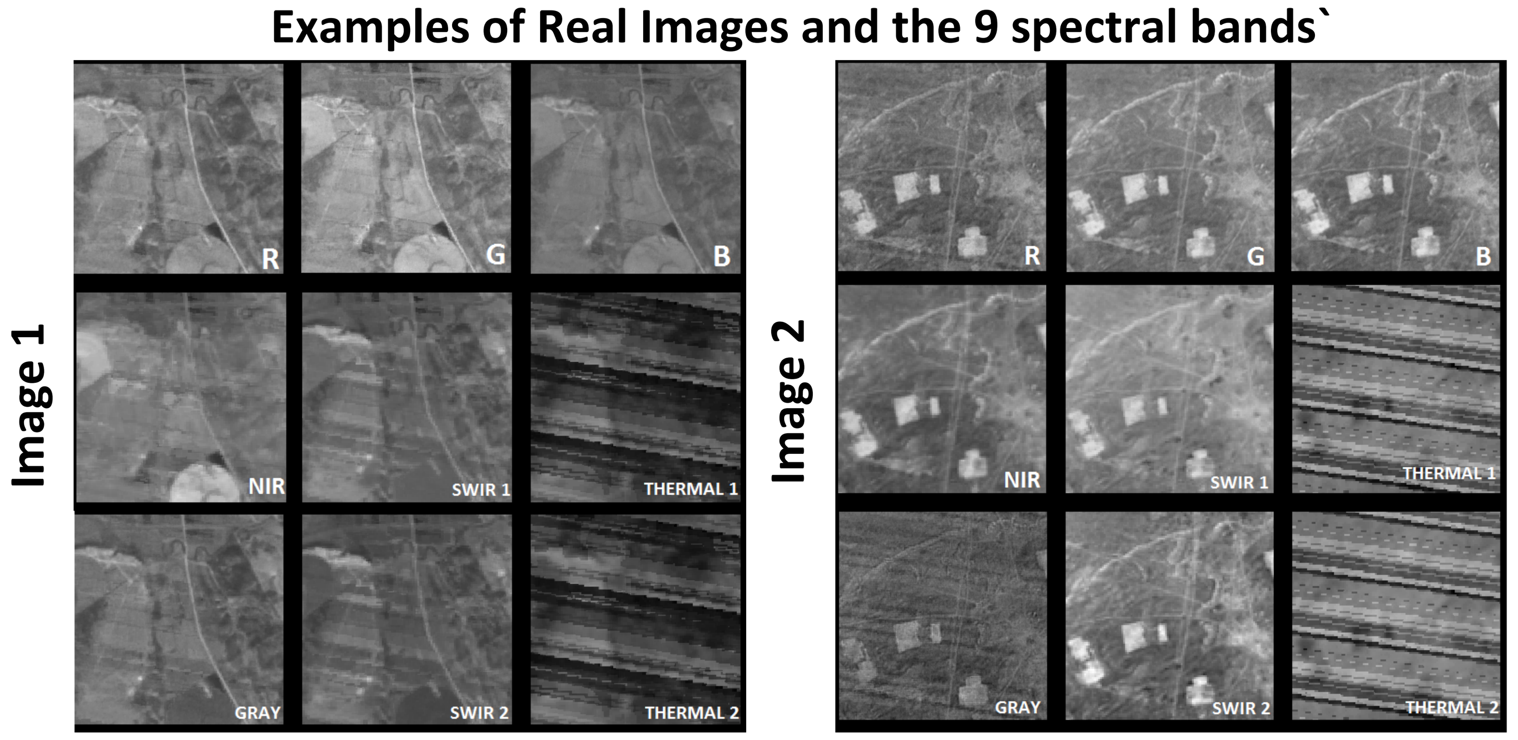
Data points are labeled based on nightlight classes of which there are 3: Rural (class 0), Semi-Urban (class 1) and Urban (class 2).



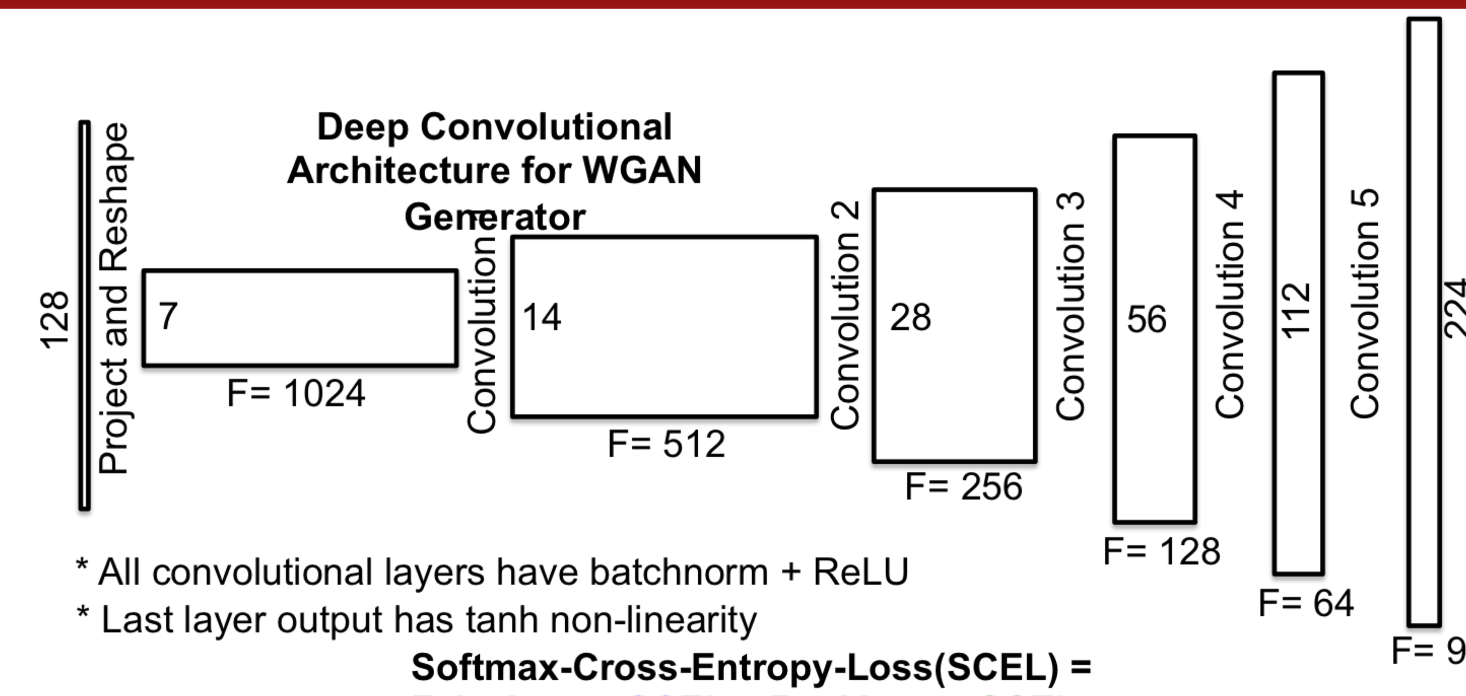
Data samples from all over the African continent balanced by nightlight classes (All Africa Dataset)

Data samples from near the locations where DHS data surveys are available with 65% rural, 20% semi-urban and 15% urban (Around DHS Dataset)

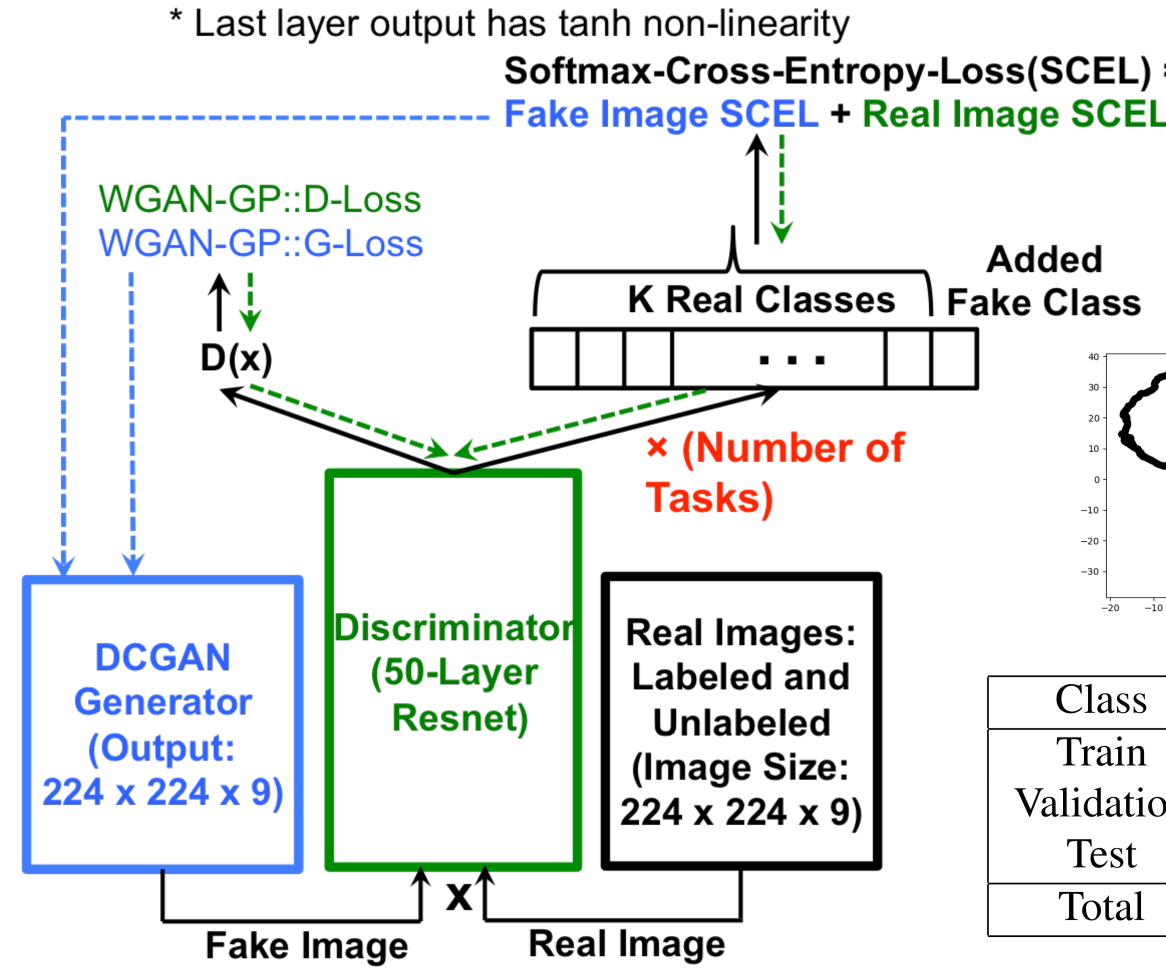
Examples of Real Images and the 9 spectral bands`

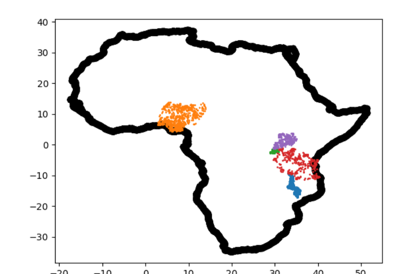



## Model and Training Algorithm




* All convolutional layers have batchnorm + ReLU
* Last layer output has tanh non-linearity



Locations where DHS data surveys are available

**Dataset Sizes**

| Class | 0 | 1 | 2 |
|---|---|---|---|
| Train | 59896 | 18486 | 13140 |
| Validation | 2805 | 883 | 788 |
| Test | 1330 | 681 | 711 |
| Total | 64031 | 20050 | 14639 |

## Loss Functions

### Loss Function Definitions from [2]

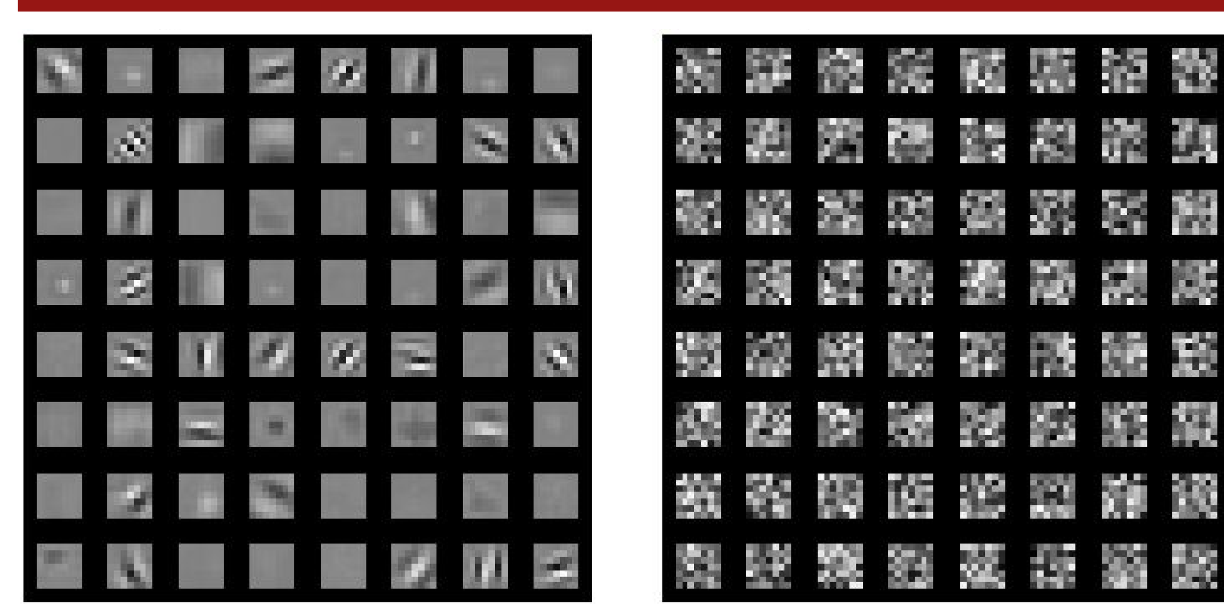
$$
\begin{aligned}
L &= -\mathbb{E}_{\boldsymbol{x},y\sim p_{\text{data}}(\boldsymbol{x},y)}[\log p_{\text{model}}(y|\boldsymbol{x})] - \mathbb{E}_{\boldsymbol{x}\sim G}[\log p_{\text{model}}(y=K+1|\boldsymbol{x})] \\
&= L_{\text{supervised}} + L_{\text{unsupervised}}, \text{ where} \\
L_{\text{supervised}} &= -\mathbb{E}_{\boldsymbol{x},y\sim p_{\text{data}}(\boldsymbol{x},y)} \log p_{\text{model}}(y|\boldsymbol{x}, y<K+1) \\
L_{\text{unsupervised}} &= -\{\mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})} \log[1-p_{\text{model}}(y=K+1|\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{x}\sim G} \log[p_{\text{model}}(y=K+1|\boldsymbol{x})]\}
\end{aligned}
$$

### WGAN-GP Loss for WGAN Training from [3]

**D-Loss:** $D_w(\tilde{\boldsymbol{x}}) - D_w(\boldsymbol{x}) + \lambda(\|\nabla_{\hat{\boldsymbol{x}}} D_w(\hat{\boldsymbol{x}})\|_2 - 1)^2$

**G-Loss:** $-D_w(G_\theta(\boldsymbol{z}))$

## Discriminator Network Visualization



On the left are shown the 64 conv first layer NIR filters with "same init" at convergence. On the right are the 64 conv first layer NIR filters with "random init"

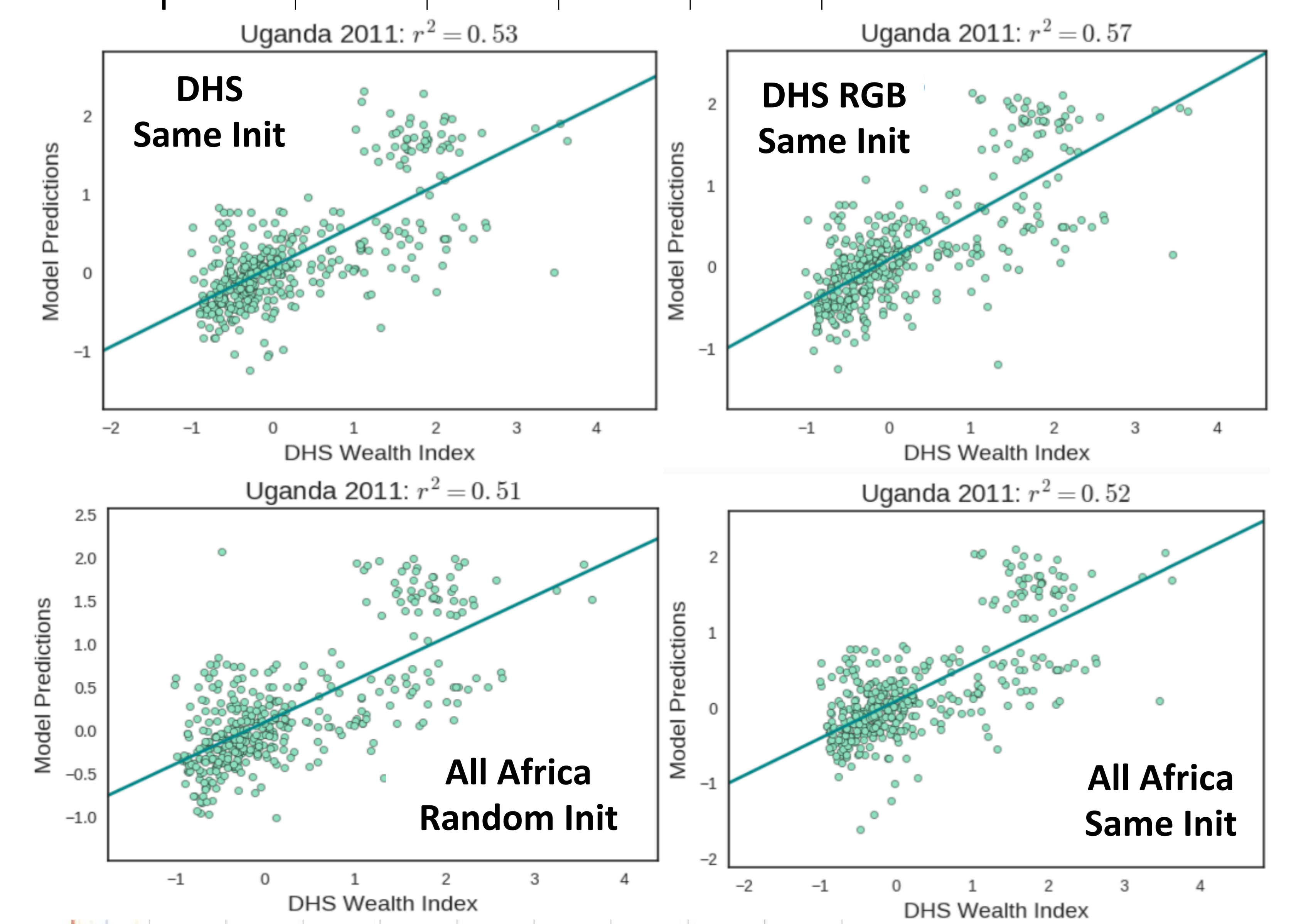
RGB Bands always initialized with ResNet Pretrained Weights. For the hyperspectral bands:
Random Init = Truncated Normal with Mean and Standard Deviation from Resnet Pretrained Weights
Same Init = Mean of the RGB channels of the ResNet Pretrained weights

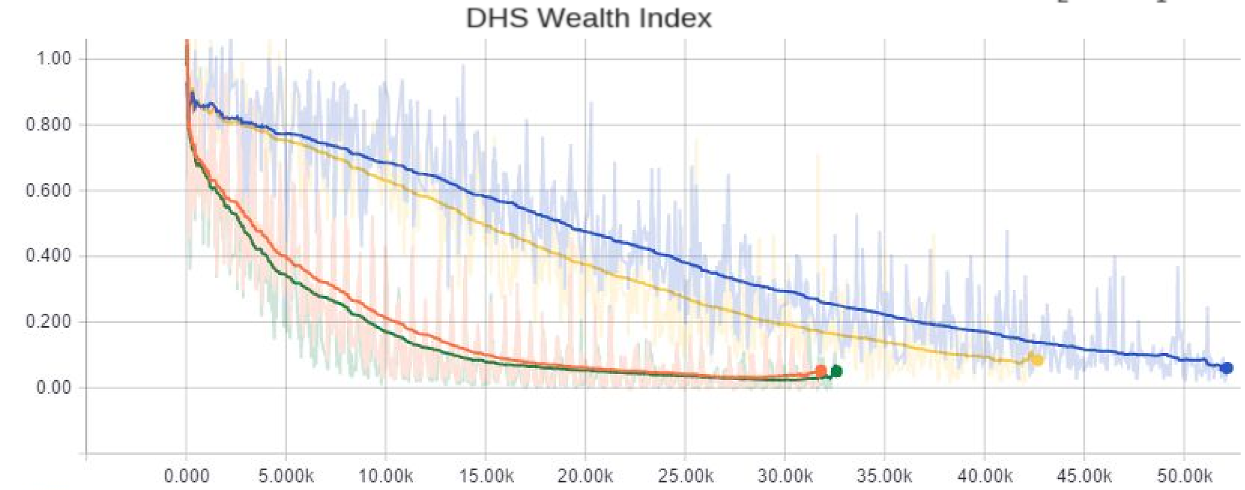
## Results

| Name | DHS Dataset | Same Init | RGB Only | Asserts $r^2$ | Training Accuracy | Validation Accuracy |
|---|---|---|---|---|---|---|
| All Africa Random Init | X | X | X | 0.51 | 96% | 66% |
| All Africa Same Init | X | ✔ | X | 0.52 | 98% | 70% |
| DHS Same Init | ✔ | ✔ | X | 0.53 | 98% | 65% |
| DHS Same Init RGB Only | ✔ | ✔ | ✔ | 0.57 | 97% | 68% |
| Nightlights | N/A | N/A | N/A | 0.484 | 100% | 100% |
| [1] | N/A | N/A | N/A | 0.66 | N/A | N/A |

**AWI is predicted using a linear model over features extracted from the network. The metrics used to evaluate a model are the accuracy of prediction and the Pearson Correlation Coefficient**

$$\rho_{X,Y} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y}$$



**Convergence behavior of loss:**
**(Orange) DHS Same Init**
**(Green) DHS RGB Same Init**
**(Blue) All Africa Random Init**
**(Yellow) All Africa Same Init**

## Conclusions & Future Work

1. (C) Model is capable of over-fitting to data showing sufficient capacity
2. (C) Asset Wealth Index prediction better than simply using nightlights
3. (FW) Adding more tasks will help reduce over-fitting and increase feature generalization. Train WGAN Generator to convergence and use it for the semi-supervised learning tasks

## References & Acknowledgements

1. Jean, et al. "Combining satellite imagery and ML to predict poverty", Science, 2016.
2. Gulrajani, et al. "Improved Training of Wasserstein GANs", arXiv:1704:00028, 2017
3. Salimans, et al. "Improved Techniques for Training GANs", arXiv:1606:03498, 2016

4. Prof. Stefano Ermon, Neal Jean, Volodymyr Kuleshov and the Sustainability and AI Lab at Stanford University