

Estimating Global Subnational HDI using Satellite Imagery and Convolutional Neural Networks

Dhruv Jain

Received August 24, 2023

Accepted November 04, 2023

Electronic access December 15, 2023

The Human Development Index (HDI) has increasingly become a key metric in informing governments and organizations across the world about the state of a country or region's progress and development. It is widely used by governments, corporations, and media. Yet, despite being so instrumental, collecting all the elements required to calculate HDI is a very expensive and highly error prone task. This especially applies to poorer developing countries where the insights from HDI are most needed. On the other hand, NASA provides very detailed nighttime satellite imagery and gridded population estimates for every corner of the world. My paper seeks to create a framework for estimating HDI using this data. I process the data by merging all the different data sources to ultimately create a train and test dataset of thousands of 200-by-200-pixel images, each representing some corner of the Earth. I design a multimodal convolutional neural network with two equally weighted input streams, a numerical stream for population, and a convolutional stream for the satellite image. This model can accurately predict HDI with mean absolute error of 0.0945 points, based on comparisons between predicted and actual values in the test dataset. There is room for improvement in the accuracy and structure of the model. While the specifics will depend on the need of the organization or government using the model, a more acceptable error range would be closer to around 0.01, similar to existing models and errors in data collection. Still, a framework like this can be used by various governments and organizations all over the world to estimate HDI.

Introduction

As the world rapidly develops, it is important to keep track of the socioeconomic variables that can quantify this progress. These may include indicators like Human Development Index (HDI), GDP Per Capita, or percent of people living below the poverty line. HDI is a composite indicator that strives to quantify key dimensions in human development and serve as a more well-rounded indicator of human development than relying solely on economic measures like income¹. The goal is to reflect the quality of life of a country's average citizen. It is calculated by the United Nations Development Programme (UNDP) and was introduced in 1990. It measures three aspects found important to human development: a long and healthy life, access to education, and a decent standard of living². It uses four key metrics to calculate a score between zero and one that is given to each country. These metrics include life expectancy at birth, expected years of schooling, mean years of schooling, and gross national income per capita (GNI per capita). Life expectancy is the number of years a person can expect to live. Essentially, for our purposes, it is the average number of years a newborn in a given country can expect to live if mortality trends at birth do not change throughout the person's lifetime³. It makes sense that life expectancy is a metric. If more people are living longer, than those people are enjoying a higher quality of life, with good healthcare and

nutrition. Mean years of schooling refers to the average number of years of schooling received by adults aged 25 and older. Expected years of schooling calculates the years of schooling a child can expect to receive if the current enrollment trends in the country do not change. Education has historically been a massive impetus for change and can lead to significant economic growth. Thus, it plays a big part in the HDI index as well². GNI per capita is income data calculated by the World Bank using various sources. It is adjusted for price changes over time and price differences between countries as well and is expressed in international-\$ at 2017 prices². More income inherently means easier access to food, education, health, and all the creature comforts associated with a higher standard of living. It is still important to note that GNI per capita doesn't consider things like income inequality. But it still makes sense why it is part of the HDI index.

Each of these metrics are then normalized between 0 and 1, based on minimum and maximum value "goalposts" set by the UNDP. Basically, through this process, three dimensional indices are calculated: health, education, and standard of living. The health index is found through normalizing life expectancy. The education index is found through taking the arithmetic mean of the normalized mean years of schooling and expected years of schooling. The standard of living index is found through normalizing GNI per capita. Then, these indices are combined through finding their geometric mean

to ultimately produce a score between 0 and 1². This score is then given to each country with some minor exceptions where enough data isn't available. In Subnational HDI, a score is given to each subregion of the world. Subregions are administrative regions within countries such as provinces and states. In large countries like China for example, there can be quite a difference between the levels of development of each region⁴. All of this helps to create a multidimensional indicator that can quantify aspects of human development more thoroughly and holistically than just a single indicator like income ever could. Through all these steps, there is now a useful baseline for measurement and comparison everyone can use. HDI is widely used by people around the world. It is used by governments, media, corporations, academia, and policy makers⁵. It has become a system through which countries measure their progress. It plays a big role in determining where governments and NGOs decide to allocate their foreign development funds⁶ or how corporations such as pharmaceutical companies set prices⁷. The importance of HDI cannot be overstated.

Despite all of this, HDI is not a perfect indicator. For example, one common criticism is that it values certain arbitrary tradeoffs that may not be in line with what one might consider as "human development". For example, a paper published in 2018 showed that if Senegal were to increase its mean years of schooling by one year, it could lose a quarter of its GNI, but still retain the same HDI⁸. How meaningful are these tradeoffs? Who decided them? What defines a "higher" quality of life can be a very subjective question, nodding to our deeply held individual philosophical beliefs. In recent years, a plethora of alternative indices have been developed to address some of these limitations⁹. These include but are not limited to the World Happiness Report or Social Progress Index. In addition to these concerns, finding the data needed to calculate HDI, especially in the developing countries where progress in human development needs to be tracked the most, is extremely difficult. HDI is calculated using many different data sources, and the accuracy depends on how each country collects their data, but the trend is clear. The more developed a country, the more precise is their data⁶. Some elements of the index are harder to calculate than others. Life expectancy is relatively straightforward and accurate. GNI per capita, on the other hand, is not that simple. It requires collecting the comparable prices for up to 200 different markets at the same time, and the methodology for doing that is complex and uses subjective assumptions⁸. This process is very costly, and thus is not done annually. For years when GNI per capita is not explicitly calculated, it is extrapolated from existing years. Overall, with all these considerations, HDI statistics can be incorrect from between 0.03 and 0.11 standard deviations, significant on the 0 to 1 scale⁶. HDI, a metric important to the modern world, is a costly and error prone index.

Despite all of its limitations, HDI still remains by far the

most widely accepted and ubiquitously used index⁵. This begs the question, can there be a way to estimate HDI using data that is more easily available? This estimate could be an alternate approach, not used to replace HDI, but to complement it as an estimate that is regularly updated. One that could help governments, corporations, and NGOs make important decisions.

There are thousands of working satellites orbiting around the earth every single day. There is a plethora of satellite imagery available online. Using machine learning techniques, perhaps we can take this easily available data and predict a hard to calculate socioeconomic variable such as HDI. The approach I describe here uses nighttime satellite imagery and gridded population data to estimate Subnational HDI on a global scale. The basic idea is that when an area is economically developed, there will be more artificial light present in the corresponding satellite image¹⁰. This is because richer countries tend to have universal electrification and a lot more economic activity. This economic activity could include things like factories, oil rigs, highways, and other light emitting sources. These richer countries will be more urbanized as well. They will have more light producing urban centers with streetlights, billboards, and skyscrapers. These rich countries will also generally have higher life expectancy and levels of education. Therefore, more artificial light means higher HDI. Research has shown that nighttime lights can be a good proxy for development, even at the local level¹¹. There is an important caveat though. The absence of light doesn't necessarily mean a lack of development. Some of the world's most developed countries have some very noticeably "dark" areas, such as the Western United States or Australian Outback. While there is no easy way to quantify something as broad and abstract as darkness, satellite imagery shows both these areas with little to no artificial light¹². Yet, these places are highly developed⁴. What explains this is a lack of population. Both Australia and the United States are highly urbanized countries, with rates of 86 and 83 percent respectively¹³. This means that most of the population is concentrated in a few relatively small "bright" urban areas. A small population in rural areas with low amounts of artificial light means high development. Compare this with Sub-Saharan Africa where urbanization rates are generally lower than 50 percent¹³. The "dark" rural areas in this region have a massive population as relatively fewer people live in the "bright" urbanized regions. A high population with low amounts of artificial light means low levels of development. All of this is to say that when using nighttime imagery as a proxy for human development, the population of the region in question must be considered as well.

It seems that we can get a sense of a region's HDI by comparing artificial light with the population living in that region. Figure 1 shows two regions at night from space: the contiguous United States and Sub-Saharan Africa. The US has less

than a third of the population of Sub-Saharan Africa¹⁴, yet it emits way more artificial light¹². Correspondingly, the US has a much higher HDI than Sub-Saharan Africa⁴.

Predicting socioeconomic indicators through satellite imagery is an idea that has been implemented many times before, though the specifics may be different¹⁰. For example, in research published in *Science* in 2016, the authors mention the same problem with collecting reliable socioeconomic data in developing countries and use robust machine learning techniques with publicly available high resolution satellite data to predict up to 75 percent of the economic variation between two regions¹⁵. They specifically focus on Africa. They also mention how many people have tried using nighttime imagery, and that while showing promise in discerning larger country wide trends, it isn't as effective in finding the differences between two relatively poor, economically similar regions. Researchers at IIT Ropar used satellite images of Indian villages and cities, not to predict a single outcome like HDI or poverty, but rather multiple related indicators, such as access to electricity, roof material, or access to tap water¹⁶. These indicators were collected in the Indian census and the model was trained on that data in tandem with satellite images. This multiple variable approach was found to be reliable, and a Convolutional Neural Network (CNN) was used for this. Current estimation methods for socioeconomic variables like the ones referenced above generally tend to focus on a specific region or specific aspect of development. In contrast, I seek to create a broad framework that can estimate HDI globally in any corner of the world, no matter the level of development. There has been a very recent attempt at estimating HDI through machine learning and satellite imagery, but it is very complex and uses multiple features and extensive amounts of data¹⁷. While this is very useful and does help with model accuracy, it also defeats the purpose of using data that is easy to collect and process. My goal is to create a simple framework that can accurately predict HDI while solely relying on two easily available pieces of data. Given the previous analysis done, it seems that using population data in tandem with nighttime satellite imagery is a good strategy for generating accurate HDI estimates. By combining nighttime satellite imagery with available gridded population data, is it possible to accurately predict HDI using machine learning?

Methodology

This project was carried out in Python. The first thing I did was download and format Subnational HDI data which would be used as the target variable for model training and testing¹⁸. I used the 2016 Subnational HDI data from the Global Data Lab⁴. This was a .xlsx table containing thousands of subregions across the world with their calculated HDI in 2016. Each one of these subregions had a certain GDL code. This code

was just a way for Global Data Lab to refer to the various subregions without causing any confusion. I also downloaded a .shp file which contained a table with all the subregion names, the coordinates of their geographical boundaries, and the same GDL code. I was able to merge the two datasets on that GDL code using Pandas. The second piece of data was nighttime satellite imagery from NASA's Black Marble Project¹². The highest resolution (500m) available allowed me to download a flat map of the world at night in 2016 in color with a resolution of 86400 by 43200 pixels. This was split up into 8 separately downloadable regional tiles, each being 21600 by 21600 pixels. The whole flat map was in the equirectangular projection. Lastly, for the gridded population data of the world, I used NASA SEDAC's population estimation service¹⁴. The highest resolution available corresponded to 1 kilometer. This dataset also used 8 separate ASCII text files, each representing data from the same eight tiles as the nighttime satellite imagery data. Figure 2 displays what regions of the world each tile corresponds to.

The goal was to have a dataset with thousands of 200 by 200-pixel nighttime satellite images, with a population number and HDI score associated with each image. This format would allow me to train a neural network on the image and population data with a target variable of HDI. To achieve this, a lot of data processing was needed.

I first focused on processing one tile before generalizing that strategy to the rest of them. I used NumPy to load the ASCII file corresponding to that tile. This loaded a two-dimensional array with a shape of 10800 by 10800. Each element in this array represented the population of a certain geographical area. That area was the corresponding square in the 10800 by 10800 grid that spanned the whole geographical area of the tile. Most of the elements in the array, representing a small $1km^2$ square of the earth, will have a population of zero. This makes sense as most of the earth is ocean or very sparsely populated land. These elements were represented with the number -9999 in the array. To fix averaging errors, I just replaced those values with zero. Next, I reshaped the NumPy array to have a shape of 108 by 108. Each element in this new array just represented a sum of all the values in the corresponding 10 by 10 subarray of the old array. Essentially, I made the squares of my grid significantly larger, $100km^2$, instead of just $1km^2$. Each element in this new array was the population of the corresponding 200 by 200-pixel square of the large nighttime satellite image tile. For processing the nighttime satellite imagery, I loaded the 21600 by 21600-pixel .jpeg tile into Python using a library called Pillow. I then cut up the large tile into smaller 200 by 200-pixel squares. This gave me 108^2 images, or the same number of elements in the new array representing population.

I now had 11664 square 200 by 200-pixel images with the population for each image. But most of these images were

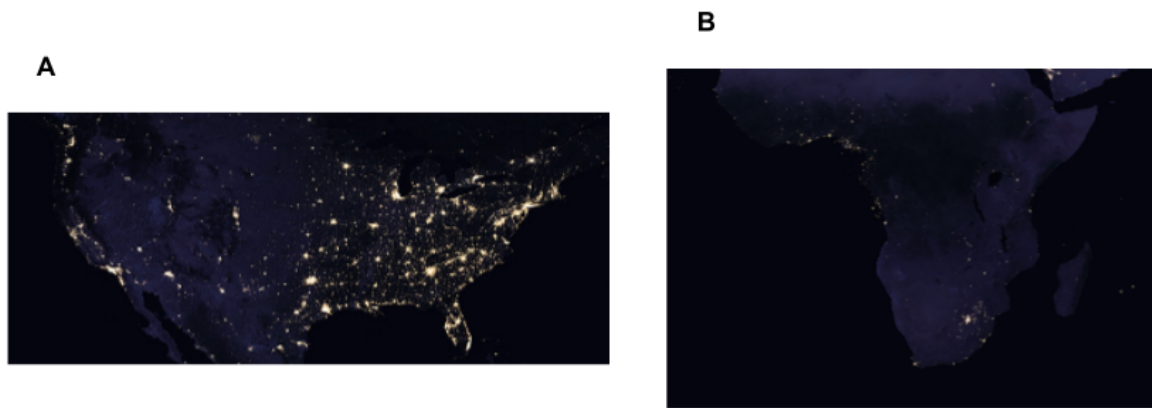


Fig. 1 Comparison of United States and Sub-Saharan Africa. As seen in figure 1A, the United States emits a greater amount of artificial light than Sub-Saharan Africa seen in figure 1B¹².

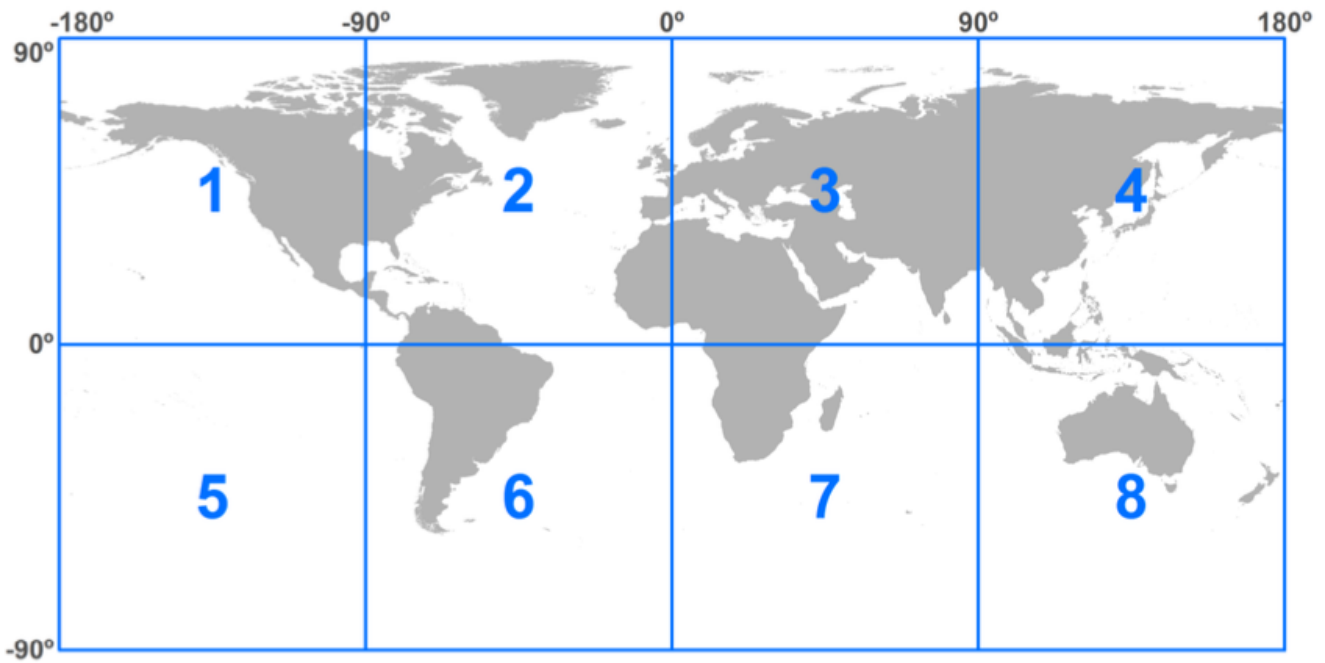


Fig. 2 The location of all eight tiles on a world map¹⁴.

completely dark with no artificial light at all. This was not very good data for the neural network to train on, given that my theory relied on measuring artificial light. To alleviate this problem, I took out all the images where the population was less than 5000. This is a bit of an arbitrary number, but through hyperparameter tuning, I found that most images above this population did have at least some sort of artificial light. This hyperparameter tuning process involved me testing four different population thresholds: 10000, 7000, 5000, and 3000. I took out all the images with a population lower than the threshold. For the first two thresholds, almost every single image left in the dataset had a decent amount of artificial light. When the threshold was 3000, a significant number of images had no artificial light. Since there is no easy way to exactly quantify how much artificial light is present in each image, I had to physically look at all of them and get a sense of what the data was looking like. In this process, 5000 came out to be the best number. This allowed me to keep most of the images. While there were some “dark” images, the majority of images had some level of artificial light.

Finding the human development index for each one of these square images was the hardest part. The flat world map of nighttime satellite imagery I used was in the equirectangular projection, so distances and shapes were warped, especially near the poles. This means that in order to find the coordinates of each 200 by 200-pixel square image, I needed to use the reverse transformation formula for the equirectangular projection. Equation 1 represents the reverse transformation formulas to find both the latitude and longitude in an equirectangular projection.

Equation 1: Reverse transformations for both latitude and longitude.

$$\text{longitude} = 360 \left(\frac{\text{pixel}_x}{\text{image}_{\text{width}}} \right) - 180$$

$$\text{latitude} = 90 - 180 \left(\frac{\text{pixel}_y}{\text{image}_{\text{height}}} \right)$$

I used the Equation 1 to find the geographic coordinates of the center of each square image. Once I obtained the coordinates for each image, I used them to find the corresponding HDI. Since I had already merged the two Global Data Lab tables to create a dataset with the subregion names, HDI score, and the coordinates of their shapes, this task was simple. For each square image, I first checked if its center geographic coordinate was already inside a subregion. If this was true, which in most cases it was, the image would just get assigned the HDI of that subregion. If this wasn't the case, a common theme in coastal images, I found the subregion whose boundary was closest. I then gave the image the HDI of that subregion.

All of this was just for Tile 1. I repeated this whole process for the other seven tiles that made up the world. After doing this, I had created a dataset with 13499 images, each a 200 by 200-pixel square of nighttime satellite imagery representing some corner of the earth. Each image came with the geographic coordinate of its center, the population within that image, and the HDI score of that image. With all this ready, a model could now be trained on this data. Figure 3 and Table 1 represent some samples of the processed data.

I also plotted the HDI values of each image on a flat map. With the goal of exploring larger trends in the data and visualizing the geographic spread of HDI values. Figure 4 represents the HDI data plotted for each image at its coordinate on a world map.

While multiple data sources were aggregated, there are still some limitations with this data. With the gridded population data, the resolution of the estimate depends on the granularity of the administration areas used to collect the data in each country¹⁴. This means that there will be some amount of error based on this variance in size. While the satellite imagery is high resolution, its nighttime nature doesn't capture a lot of the nuance. It treats all light the same, and doesn't show other aspects of development, such as park area, lot size, and the presence of slums. It also captures oil rigs and wildfires as artificial. Even the target variable, HDI, has a significant amount of error due to the factors already discussed in the introduction. Given these limitations, it would be interesting to see how accurate of a prediction this machine learning model could give us.

In terms of machine learning techniques, I used a multi-modal convolutional neural network architecture that took in both an array representing the square image and numerical population data as an input. This model had two input streams. The first was the CNN stream that took in the image as an input. The input shape was (200, 200, 3) as each pixel in the 200 by 200 image was represented by a three-element array of RGB values. This image array was then preprocessed by scaling all the values between negative one and one. Then this preprocessed array was inputted into the MobileNetV2 pretrained model available in TensorFlow. This lightweight model is very good for image classification and object detection tasks¹⁹. While that isn't exactly I did, the complexity of this model was useful for the image regression task at hand. After MobileNetV2, there is a Flatten layer. This essentially flattens the output from MobileNet. After this, there are a couple of Dense fully connected layers with an output Dense layer made up of 10 neurons. The ReLU activation function was used. The idea behind this architecture was to continually decrease the number of neurons in each subsequent layer while maintaining a decent amount of complexity. The hope was that, with the complexity, the model would also pick up information about things like urbanization in addition to just the

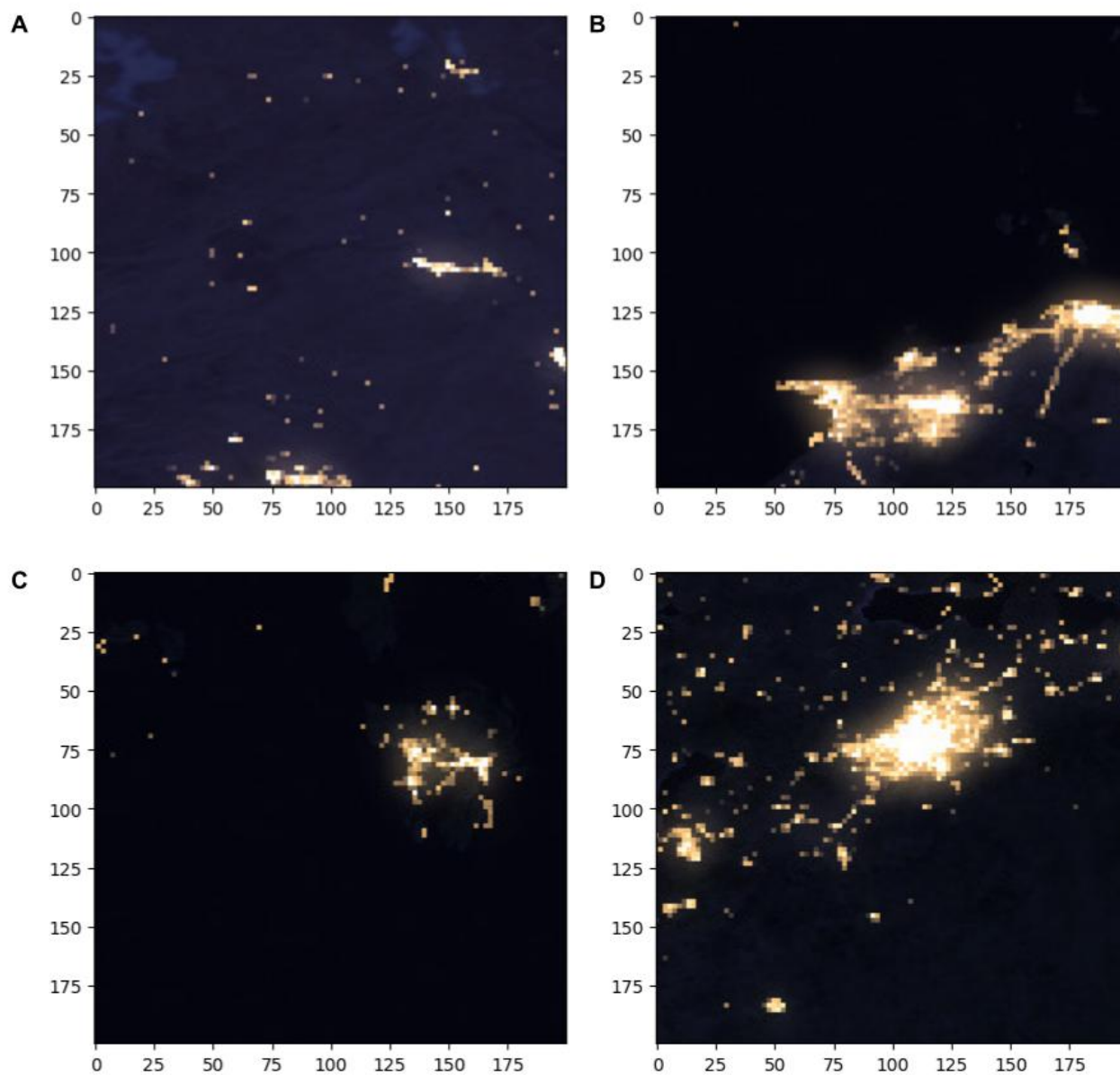


Fig. 3 Four example images. Each one is a 200 by 200-pixel square with varying levels of artificial light. The different natural landscapes are evident as well.

amount of artificial light in each image. All of this defined the CNN stream. The numerical stream dealing with population took in one numerical input and normalized it to values between negative one and one. I then used three Dense layers to follow it, with 64, 32, and 10 neurons. They all had the ReLU

activation function. In summary, the numerical stream used a basic feedforward neural network.

Then, to combine both the numerical and CNN streams, I took their 10 neuron outputs and just concatenated or stacked them on top of each other. The idea was that the same number

	Figure 3A	Figure 3B	Figure 3C	Figure 3D
HDI	0.803	0.747	0.800	0.738
Coordinates	(38.75, 37.08)	(37.92, 120.42)	(6.25, 99.58)	(19.58, -101.25)
Population	136310	1032561	137960	1378143

Table 1 Numerical data for each image

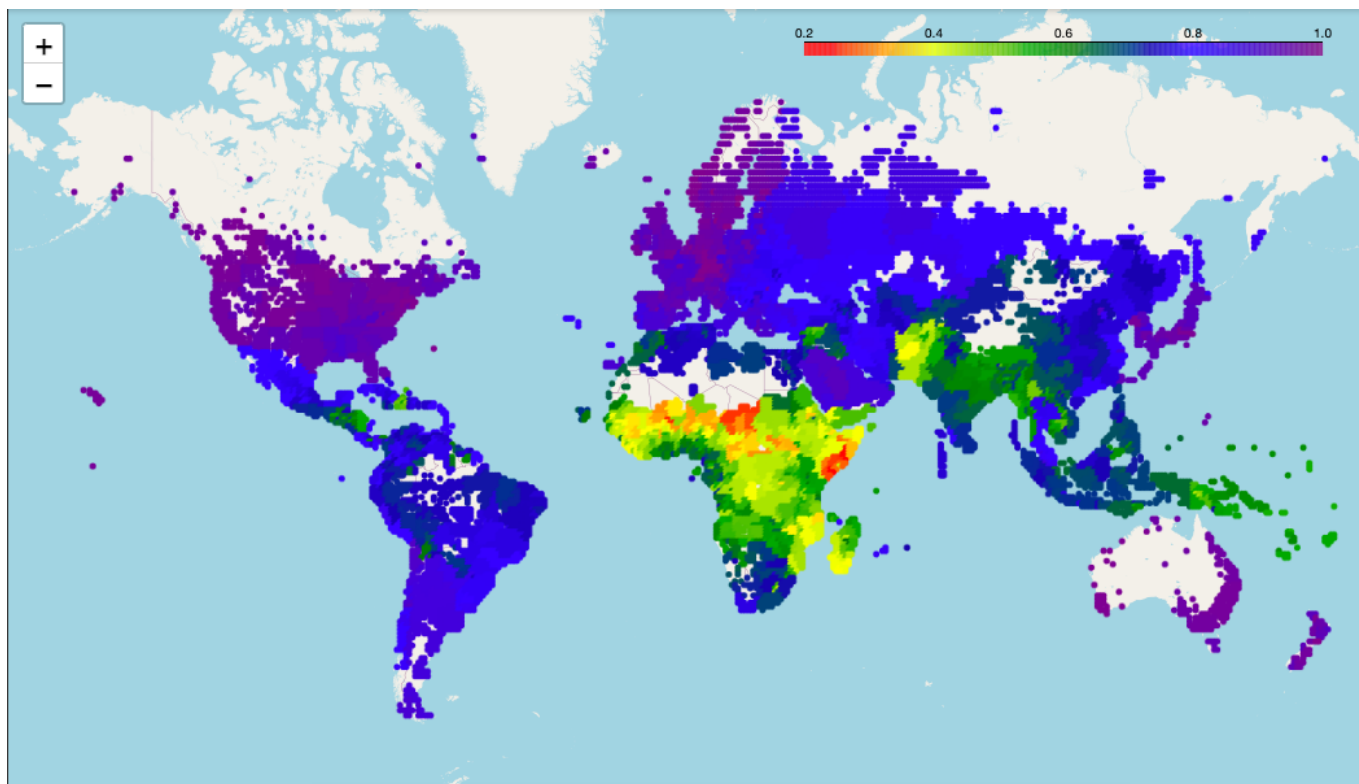


Fig. 4 HDI values plotted on world map. As can be seen, a large amount of the world was not included in the images due to low population density and not meeting the minimum population threshold. There is a wide range of HDI values across the world and even within some large countries.

of neurons for each stream represented giving equal importance to each type of data. Intuitively, it made sense to give equal weightage to both population and nighttime satellite imagery as both were important in figuring out the HDI. Finally, I followed this with four more Dense layers, with 16, 10, 4, and the final one neuron output layer representing the HDI prediction. Figure 5 represents a flowchart of the model architecture.

This architecture was chosen to give equal weight to both the satellite image data and numerical population data. Experiments using other model architectures were tried as well. Originally, the MobileNetV2 wasn't incorporated in the model. Instead, it was replaced with a simple convolution neural network, with only a couple of Conv2D layers, and a Dropout and Flatten layer. This model did not work well. It had a mean absolute error of 0.14, which seemed decent at

first. But all its predictions were essentially the same value, very close to the mean HDI. It seemed that the model defaulted to predict the mean, rather than picking up on the nuances between images.

Due to this, I decided to increase the complexity of my model, in the hope that something with more trainable parameters will actually pick up on the trends we want it to. This is how I ultimately decided on the specific model architecture above.

In terms of fine-tuning, I experimented with training 10 epochs. The accuracy started getting worse after 8 epochs due to overfitting though. I also first designed the model with fewer Dense layers. I only had two dense layers after concatenating instead of 4. But the model performed better with the increased complexity of 4 layers.

Multimodal CNN Architecture

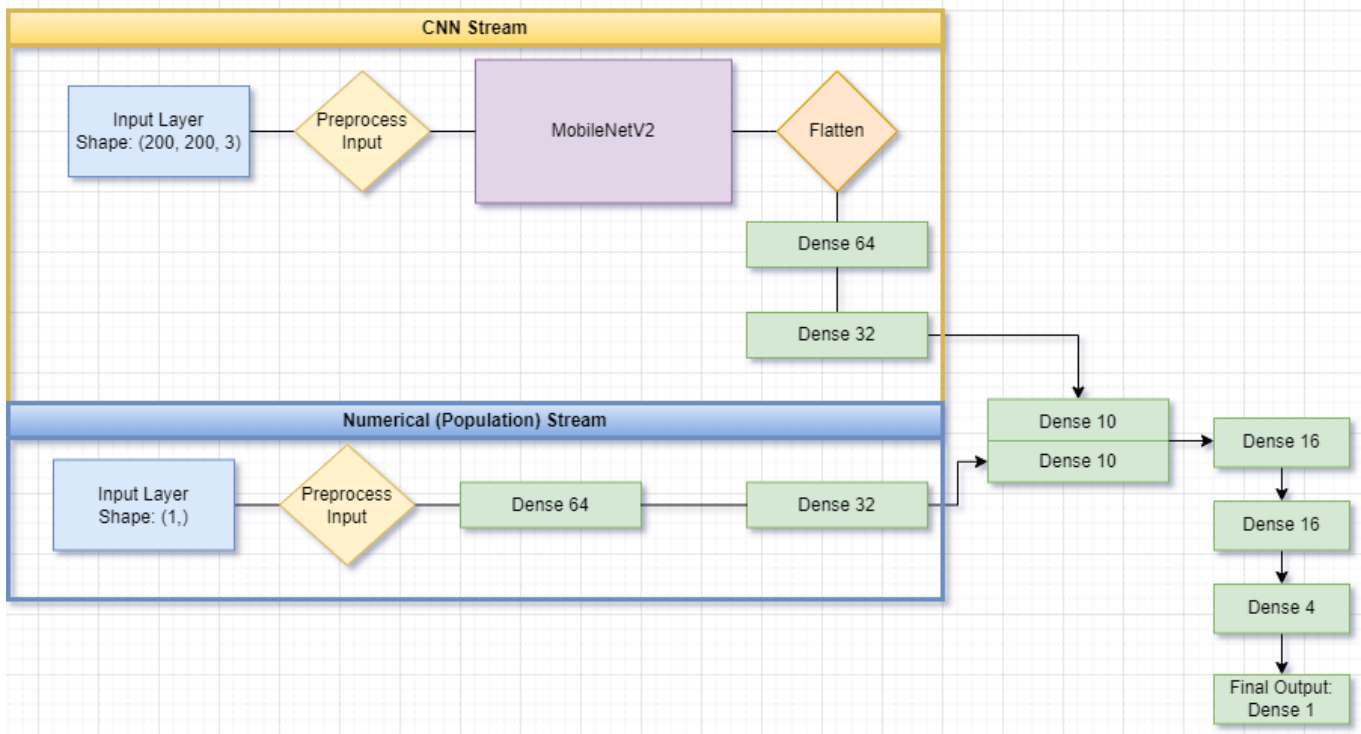


Fig. 5 Flowchart of multimodal CNN architecture. As seen above, the data from the two streams is concatenated and passed through a final feed-forward network to output one neuron with the numerical HDI estimate.

Ultimately, for the training process, I used an 80-20 train split on the 13499 sample image data. This gave 10799 images in total to train on. The loss function used was mean squared error, the optimizer used was Adam, and 8 epochs were used with a batch size of 32 images. To test the accuracy, I used mean absolute error as my metric, and tested it on the 2700 samples that made up the test dataset.

Results

This trained model had a mean absolute error of 0.0945 points on the test dataset. This number was calculated by first finding the absolute value of difference between the predicted and actual value of HDI for each entry in the test dataset. Then the mean of all of these differences is taken to produce this number. Figure 6 displays a comparison of the histograms of the model's predictions and the actual HDI values in the test dataset. Plotting and comparing both the histograms, it is clear that the model can predict a wide range of HDI values, but still has trouble predicting the highest of values. We will analyze

this more thoroughly in the Discussion section.

The standard deviation of the HDI in the test dataset is also much higher than the mean absolute error, at about 0.16. Figure 7 shows some examples comparing the model's predicted and actual values.

Figure 8 shows the predicted HDI values for all the images plotted on a world map.

Discussion

The model does a pretty good job of predicting a wide variety of HDI values to a high degree of accuracy. Obviously, "high" is a subjective term. A mean absolute error of 0.0945 seems small, but that number represents potentially decades of progress and development. That error is the difference in HDI between the United States and a country like Turkey. Also, other existing models have been able to predict HDI with a much lower error closer to 0.01²⁰, albeit using methods that require a lot more data collection. Simply put, this model is not good at distinguishing between nuances in countries and

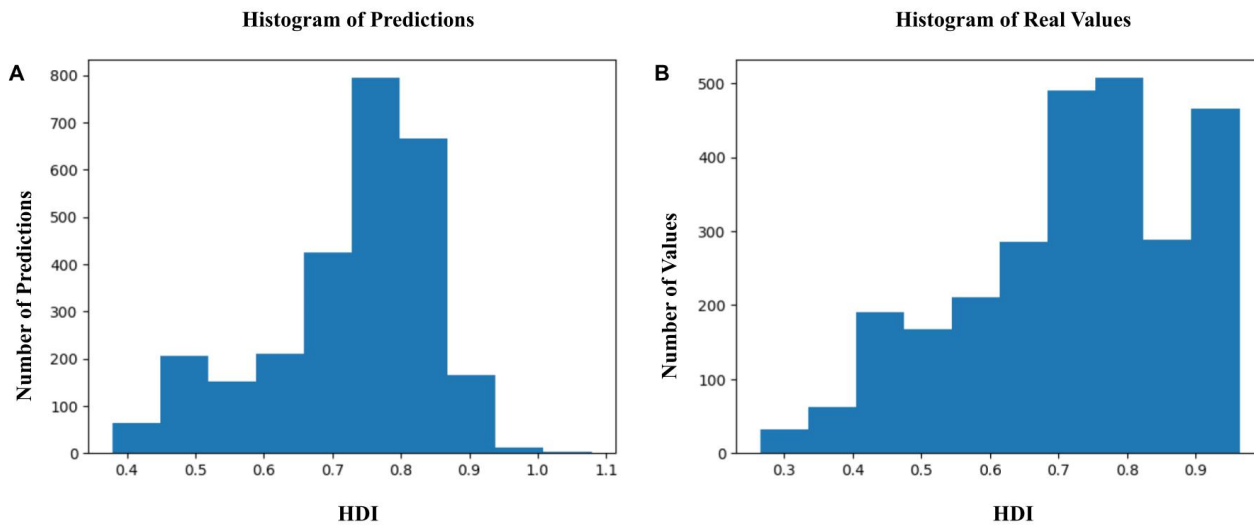


Fig. 6 A comparison of two histograms representing the model’s predictions and actual HDI values. As seen above, the model can predict a wide range of HDI values but has trouble predicting the highest of values greater than 0.9. The horizontal axis represents the HDI number, while the vertical axis represents the number of samples within that HDI range.

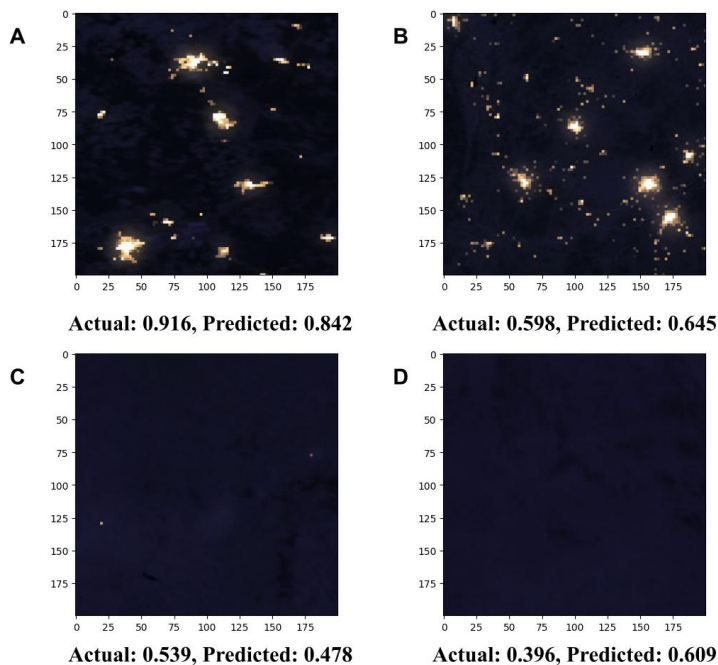


Fig. 7 Examples of the model’s predictions. The model both underestimates and overestimates the HDI depending on the given image. Figures 7A and 7B shows a cluster of heavily populated urbanized areas while Figures 7C and 7D show rural areas with relatively little artificial light.

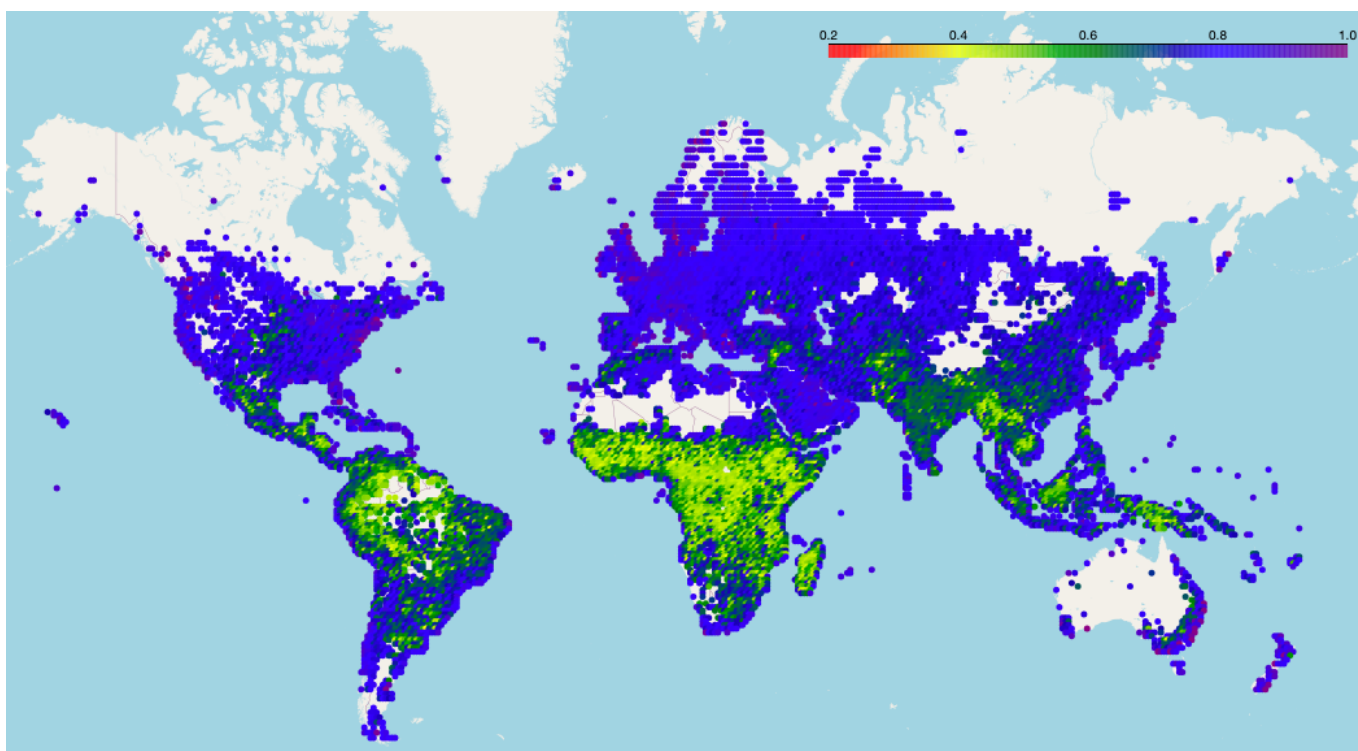


Fig. 8 Predicted HDI values plotted on world map.

regions with similar levels of development. To be fair, that is a lot to be asking from a model that is only using population and satellite imagery to predict an index that uses completely different parameters. This model is still broadly able to distinguish between higher and lower levels of development through data that is much easier to collect than the data that makes up the actual HDI. The biggest issue with the model is distinguishing between areas that are highly developed on a global scale, yet still have disparities between them. For example, the model can't distinguish between development in the former Soviet bloc and Western Europe. This makes sense as both areas have achieved nearly universal electrification²¹, making it hard to distinguish development using nighttime satellite imagery. On a global scale, both regions are considered to be relatively highly developed. This also explains why in the histograms in Figure 6, the model was struggling to predict the highest of values. It seems that whenever two areas have universal electrification and a relatively high level of development, such as the former Soviet Bloc and Western Europe, the model just defaults to predicting an HDI of around 0.8. It can't distinguish that extra 0.1 difference in development Western Europe has. In terms of improving this issue, experimenting with the model's structure and fine-tuning it even more would help. This could indeed make a significant difference in limiting error. We could also use higher resolution data, that could

capture the real-time nuances between regions regarding artificial light. But ultimately, it would make sense to incorporate other data, such as daytime satellite imagery, as other models have done. The model also has a strong preference for categorizing coastal areas as highly developed, such as the Eastern Seaboard of the US. This also makes sense as coastal areas generally are more developed and contain a lot of urbanized areas with a high HDI. The model is most strongly able to distinguish areas with a very low amount of light compared to population. It clearly gives Sub-Saharan Africa and parts of Asia relatively low scores compared to the rest of the world. It is even able to pick out countries such as Afghanistan and Syria from their surroundings, two unfortunately war-ravaged places, and gives them a lower HDI score. There were some areas where the model did underperform. For areas in the Amazon Rainforest, scores were significantly lower than what they should've been. In Figure 8, images in the inland North-Central part of South America, where the Amazon rainforest is located, have a predicted HDI in the 0.5-0.6 range. In Figure 4, where the actual HDI value is presented, images in that same region have an HDI value close to 0.7. This works out to be an error as great as 0.2 points. This is significantly higher than the mean absolute error of 0.0945. I am not sure why this is. One possible reason is that rainforest areas tend to show up darker in satellite images due to the tree canopy. Other

physical features, such as glaciers and deserts generally look less dark in nighttime satellite images. This may be confusing the model into thinking that those areas have more artificial light than they really do. For forested areas, this could lead to the model measuring less artificial light than there actually is, thus contributing to the lower HDI scores. Overall, inland areas tended to perform worse no matter the location. Comparing the world maps shown in Figure 8 and Figure 4, the model can accurately pick out the general trends in HDI on a global scale.

It is very interesting that something as complicated and multidimensional as HDI can be predicted with such a high degree of accuracy through parameters that are completely different. It opens the possibilities to wonder if there may be other ways to measure human well-being. Having a model like this has massive implications for the world. Governments, corporations, and nonprofits can use this model with their own data to accurately estimate the HDI for their target region and scale. They can then make informed decisions which serve their interests the best. They are not at the mercy of statistics from a certain year, country, or tied to a certain subregion anymore. Satellite data is available very easily nowadays, and population data is still much easier to collect than something like income. Using this model with the appropriate data, anyone can estimate the HDI accurately in any corner of the Earth. That's a very useful application of these machine learning techniques. This paper also corroborates and confirms what other researchers in this area have found: satellite imagery can be utilized effectively to measure socioeconomic conditions¹⁰. While a lot of research involves predicting socioeconomic conditions in a narrower scope, looking at certain cities or countries, I expand upon that research by creating a successful framework that is more broadly applicable.

Conclusion

We can now see that it is possible to estimate a broad socioeconomic indicator such as HDI on a global scale using satellite imagery and population. This framework and model will be useful for a wide variety of people. Ultimately though, there is still a lot more that can be done. The accuracy of the model is still not good enough to estimate small scale changes in development. The next step would be to fine tune this model and get it to be the most accurate it can. It would also be useful to use some more real-time, higher resolution data to train the model. This gives us more data to work with, and potentially opens the doors to understanding small scale changes and differences in development. Ultimately, this is an emerging field, and models like the one I created will become instrumental in measuring socioeconomic development in the coming decades as the world gets richer and tries to bring a higher quality of life to those who need it the most.

Acknowledgements

Thank you for the guidance of Mark Lisi from Yale University in the development of this research paper.

References

- 1 *Human development index (HDI)*, <https://hdr.undp.org/data-center/human-development-index#/indicies/HDI>, Retrieved August 13, 2023, from.
- 2 M. Roser, *Human development index (HDI)*. *Our World in Data*, <https://ourworldindata.org/human-development-index>, Retrieved August 13, 2023, from.
- 3 M. Roser, E. Ortiz-Ospine and H. Ritchie, *Our World in Data*.
- 4 S. H. database, *Global Data Lab*, <https://globaldatalab.org/shdi/>, Retrieved August 13, 2023, from.
- 5 K. Dervis and J. Klugman, *Revue D'économie Politique*, **121**, 73–92.
- 6 H. Wolff, H. Chong and M. Auffhammer, *National Bureau of Economic Research*.
- 7 R. Bate and K. Boateng, *SSRN Electronic Journal*.
- 8 S. Ghislandi, W. Sanderson and S. Scherbov, *Population and Development Review*, **45**, 219–233.
- 9 F. Comim, *Beyond the HDI? Assessing alternative measures of human development from a capability perspective*, <https://hdr.undp.org/content/beyond-hdi-assessing-alternative-measures-human-development-ca> UNDP (United Nations Development Programme).
- 10 T. Ghosh, S. Anderson, C. Elvidge and P. Sutton, *Sustainability*, **5**, 4988–5019.
- 11 A. Bruederle and R. Hodler, *PLOS ONE*, **13**, 0202231.
- 12 *Earth at night: Flat maps*, <https://earthobservatory.nasa.gov/features/NightLights/page3.php>, Retrieved August 13, 2023, from.
- 13 H. Ritchie and M. Roser, *Urbanization*. *Our World in Data*, <https://ourworldindata.org/urbanization>, Retrieved September 17, 2023, from.
- 14 *Gridded population of the world (GPW), v4 — SEDAC*, <https://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-count-rev11/data-download>, Retrieved August 13, 2023, from.
- 15 N. Jean, M. Burke, M. Xie, W. Davis, D. Lobell and S. Ermon, *Science*, **353**, 790–794.
- 16 S. Pandey, T. Agarwal, C. Krishnan and N., *Proceedings of the AAAI Conference on Artificial Intelligence*, **32**, year.
- 17 L. Sherman, J. Proctor, H. Druckenmiller, H. Tapia and S. Hsiang, *National Bureau of Economic Research*.
- 18 J. Smits and I. Permanyer, *Scientific Data*, **6**, year.
- 19 M. Sandler and A. Howard, *MobileNetV2: The next generation of on-device computer vision networks*, <https://ai.googleblog.com/2018/04/mobilenetv2-next-generation-of-on.html>.
- 20 P. Akin and T. Koc, *Adiyaman University Journal of Science*.
- 21 H. Ritchie, M. Roser and P. Rosado, *Our World in Data*.