Using Satellite Images to Track Relative Socioeconomic Development in India

Aatif Nisar Dar, **Badrinath Padmanabhan**, Mrunal Atul Kadane, Aaditeshwar Seth
Indian Institute of Technology, Delhi

July 8 – July 11, 2024

Introduction

Challenges:

- Traditional census data is valuable but outdated and slow to update.
- Satellite data offers promise, but using daytime imagery or nightlights alone has limitations[1].
- Models trained on daytime satellite images lack temporal robustness i.e. models trained on one year may not produce estimates for other years. Models trained on nightlights data are not suitable for village-level tracking because of low variability for rural areas.

Proposed Solution:

- Combine daytime satellite images with nightlight data to create a satisfactory temporally-robust model.
- Use 2011 Indian census data as a ground truth labels as done in Gulgulia et. al.[2] to train a CNN model on satellite imagery for producing initial development estimates.
- Train a regression model on a set of features; outputs from CNN model, nightlights-based features, and population features from census 2011 to estimate Relative Wealth Index (RWI)[3] of the target village.
- Ensure Temporal Robustness by evaluating the model, trained on census data from 2011, on census data from 2001 on those indicators that are common for these census years.
- Creating a standardized relative wealth index at a village level by following the methodology provided by Demographics and Health Surveys[4].
- Produce Village level development estimates over a period of two decades i.e. for the years 2003, 2011, and 2019.
- Generate development estimates for 14 states covering over 77.8% of India's population.

Contributions:

- Provides standardized development estimates for villages.
- Enables studying village development dynamics and trends.
- Informs policy decisions for targeted interventions.

Data

- Datasets are utilized in this study:
  1) Census of India
  2) Village shapefiles
  3) Landsat 7 satellite data
  4) DMSP and VIIRS nightlights data.
- The Census of India provides valuable socio-economic indicators at the village level, serving as labels for machine Learning models.
- We break down census data into categories, as shown in the Table, and then use them to categorize the satellite data.

Indicator	Type	Resources
Bathroom facility	Rudimentary	No Latrine Facility
	Intermediate	Pit Latrine
	Advanced	Pipe Sewer/Septic Tank
Fuel for cooking	Rudimentary	Firewood
	Intermediate	Cow Dung/Kerosene
	Advanced	LPG/PNG/Biogas
Main source of water	Rudimentary	Well/Spring/River
	Intermediate	Tube Well/Hand Pump
	Advanced	Tap Water/Treated Water
Literacy	Literate population percent	
Asset ownership	TV	
	Telephone	
	Two wheeler (2w)	
	Four wheeler (4w)	

Table 1. Categorization of villages by grouping census variables

Methodology and Results

1) We start with a CNN based architecture where the model takes a village satellite image as an input and predicts the development class (Rudimentary, Intermediate, or Advanced). Five models were trained, one for each indicator (IND).

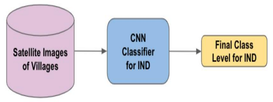


Fig 1. Architecture 1

Indicator	Train Accuracy (%)	Validation Accuracy (%)
MSW	75.34	74.10
ASSET	77.06	78.93
FC	72.09	79.47
BF	77.09	79.47
LIT	66.09	79.47

Table 2. Architecture 1 results

2) We take the softmax outputs from Architecture 1 as features and combine them with the same features obtained from neighboring villages, along with population features and nightlights-based features (Shown in Table 3).

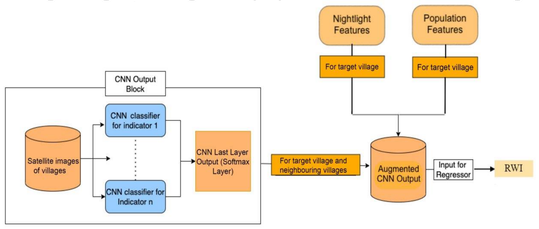


Fig 2. Architecture 2

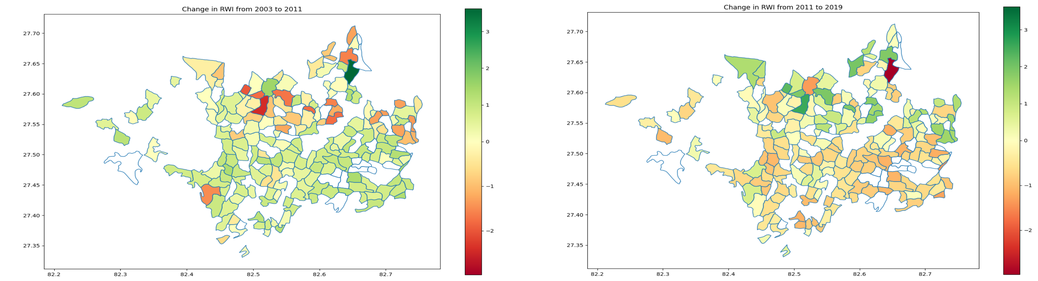
SI No.	Feature	Description (%)
1	$\frac{1}{D}$	Nearest economic hub
2	$\frac{2}{D}$	Nearest and largest economic hub
3	$\frac{1}{D}$	Nearest and most intense economic hub
4	$\frac{S*I}{D}$	Nearest, largest, and most intense economic hub

Table 3. List of Nightlight features. Here D is the distance between the villages and the economic hub

- The outcome of Architecture 2 is a RWI value of a particular village. We calculate an aggregated RWI value for each state by combining the RWI values of all the villages within that state using the following formula:

$$Aggregated\ RWI\ of\ a\ State = \frac{\sum\ (RWI\ of\ a\ village * Population\ of\ that\ village)}{Total\ Population\ of\ the\ state}$$

- Below figures represents the change in RWI values over years (2003-11 and 2011-19) for villages in Tulsipur block in UP.



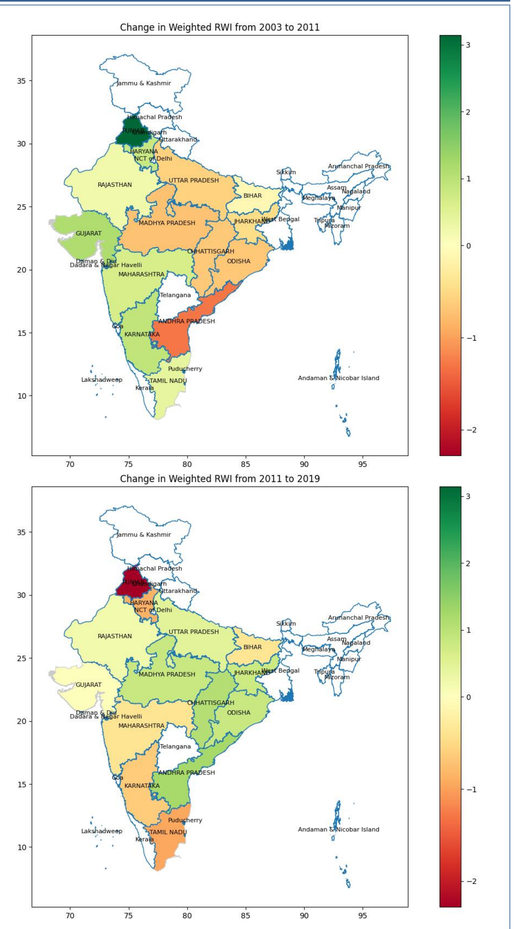


Fig 3. Change in RWI over the years (2003-11 and 2011-19) for 14 states in India

Future Work

- Use of Transformer based model instead of convolution based model for training.
- Use of models pretrained on labelled satellite data instead of models pretrained on ImageNet or other big datasets.
- Study impact of welfare expenditure on socio-economic development.
- Expand the set of indicators to include more meaningful features from census data.

References

1) Ni, Ye, et al. "An investigation on deep learning approaches to combining nighttime and daytime satellite imagery for poverty prediction." *IEEE Geoscience and Remote Sensing Letters* 18.9 (2020): 1545-1549.
2) Gulgulia, Anant, et al. "Tracking Socio-Economic Development in Rural India over Two Decades Using Satellite Imagery." *ACM Journal on Computing and Sustainable Societies* 1.2 (2023): 1-31.
3) Chi, Guanghua, et al. "Microestimates of wealth for all low-and middle-income countries." *Proceedings of the National Academy of Sciences* 119.3 (2022): e2113658119.
4) https://dhsprogram.com/topics/wealth-index/

Contact Emails

Aatif Nisar Dar: Aatif.dar11@gmail.com
Badrinath Padmanaban: badrinath95@gmail.com
Mrunal Atul Kadhane: Mrunal.Atul.Kadhane.cs121@cse.iitd.ac.in
Aaditeshwar Seth: aseth@cse.iitd.ac.in