Economic Activity Estimation Using Satellite Imagery and Machine Learning

By

ERRAJI, Brahim

Supervised By: ***Pr .Younes Jabrane*** and ***M. Oussama Abboudi***

End-of-studies project report

AL KHWARIZMI

UNIVERSITY MOHAMMED IV POLYTECHNIC

&

SADID Program

P-CURIOSITY LAB

August 13, 2021

Preface

As part of the Data Science Undergraduate Program at UM6P, students are required to do an end-of-studies internship at the end of the sixth semester for three months. The aim of which is to practice acquired knowledge, tools, and techniques in a real-world project/problem. My project of choice is entitled "Economic activity estimation in Morocco using satellite imagery and machine learning". The goal behind this project is to fit a model capable of predicting the economic activity level in a given area simply by providing a satellite image of that area. Having such a model will help track economic activity developments in real-time helping authorities measure the effectiveness of different sustainable development strategies. I chose this project because it combines all the pillars of data science: Acquiring the data (Satellite imagery and economic activity indicators), cleaning the data and leaving the relevant information only, and fitting a model to predict the economic activity level using different ML algorithms and comparing the results and finally delivering the ready-to-use model to the authorities to help them make informed decisions.

Acknowledgement

I would like to express my gratitude to my two supervisors Professor Younes Jabrane and M. Oussama Abboudi for their guidance and help throughout the period of making this project. I also would like to thank everyone at University Mohamed VI Polytechnic especially Al Khwarizmi department for the quality of life and education that I benefited from during the last three years.

Abstract

In the context of my end-of-studies internship I worked on a society-oriented problem, due to the lack of reliable methods, Morocco has been struggling to track economic developments in its regions. In this project, we propose an inexpensive approach to estimating economic activity using satellite imagery, nightlight intensities, and survey datasets. After both training our model directly to estimate economic activity and using a proxy (Nightlights), we found that the model's performance is highly dependant on the resolution of the images and on the correct matching of images with their respective labels (economic activity level).

Keywords— Artificial Intelligence, Machine Learning, Computer Vision, Satellite Imagery

Contents

Introduction

Until now Morocco used mainly manual methods to estimate economic activity in its regions. namely the general census in which the state hires people to go door-to-door and ask people various socio-economic questions. It is not hard to hard to see the flaws of this method: expensive because it involves additional employees and fees, unreliable because it counts on the citizens' integrity to answer the questions correctly, but most importantly it is slow. the general census is done once every ten years in Morocco making it impossible to track economic developments in real-time.

Figure 1: Moroccan General Census 2014

Having a method that is fast, cheap, and doesn't require much human intervention would allow obtaining time series of economic activity in the regions of Morocco. This will help when making a multitude of crucial decisions that influence directly the human population, such as wealth distribution, tax on property value determination...

Below is the distribution of poverty following the Moroccan general census of 2014. Distribution is subject to continuous change especially in times of pandemic and economic crisis. This emphasizes the need for a method for continuous tracking of poverty and economic activity.

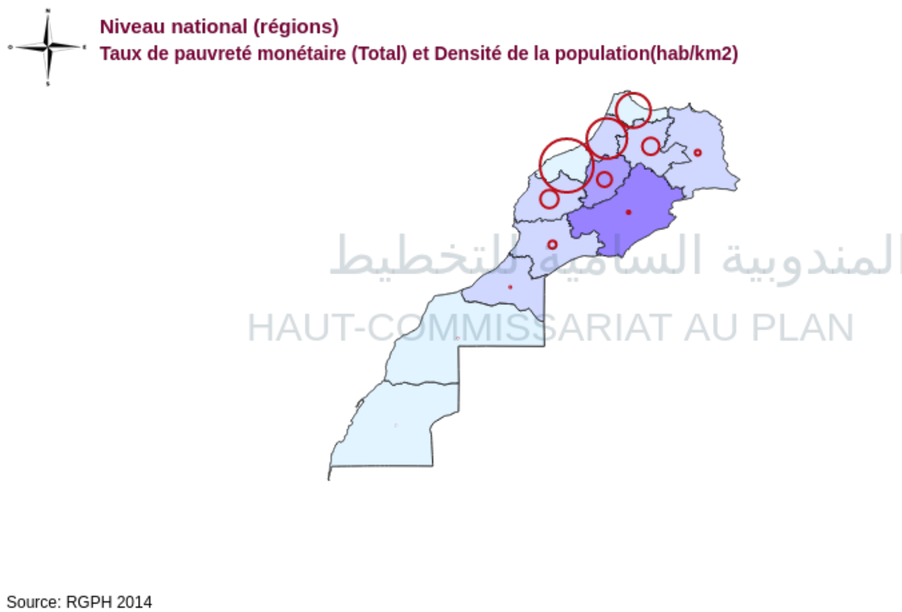


Figure 2: The distribution of poverty in Morocco

About P-Curiosity Lab

P-curiosity lab is an innovation LAB committed to ensure a community-led transformation for rural communities across Africa, and this by providing inclusive, sustainable and innovative services. P-Curiosity LAB (PCL) is hustles with innovative people their world-shaping ideas working with experts and mentors. Located at University Mohammed VI Polytechnic*(UM6P), The P Curiosity LAB use biomimetic as inspiration process. The Lab members can explore, prototype and test innovative services with the aim to serve farmer sustainably and finding a new opportunities to improve services business. In P-Curiosity LAB students, researchers, experts and entrepreneurs feed each other. This exchange between the world of science and that of the company allows the realization of new ideas and the commercialization of the services that come from it. In addition, the proximity between the different actors accelerates the development process.

Satellite imagery

Satellite imagery has an immense amount of information and features, including those that can explain the variance in economic activity indicators. such as household expenditure and poverty rates...
Therefore getting a computer to understand what a poor and a rich region look like would help in estimating the economic activity in those regions. It helps that satellite imagery is now ubiquitous, with

over 2,666 satellites orbiting the earth, many of which are used purely for earth observation purposes. *Sentinel-2* is a widely known earth observation mission that has launched in 2017. The mission's main goal was provide free high spatial resolution (10 m to 60 m) satellite imagery of the entire earth for research and education. In this project we used images of 480 different regions of Morocco, provided by sentitel-2. These regions are equally distributed over economic activity levels to avoid over-fitting whilst training the model. it is worth mentioning that sentinel-2 isn't the only satellite imagery provider. In fact there are many free and paid options with a resolution that can go up to 0.5m per pixel. namely the *Mohammed VI A and B* satelittes operated by Morocco, due to financial complications we couldn't use their images in the making of this project.

Figure 3: Sentinel-2 satellite

We aim to exploit the advancements in computer vision and machine learning to extract features that influence the variance in economic activity.

Project report structure

In this project report, we will follow a logical sequence of chapters. we will start by exploring the state-of-the-art in economic activity estimation: what has been done and what are the results and performance. Then we will move on to investigate the recent findings in machine learning and computer vision, which are two important components in the making of this project. finally, we will expose our dataset and model along with a discussion of the obtained results.

Chapter 1

State of the art : Economic activity estimation

In this chapter we will explore the efforts that has been done to digitize the economic activity estimation process.

To date, many approaches have been investigated to estimate economic activity efficiently and cost-effectively. One of the most active research entities in this regard is the Word Bank. The first approach the World Bank considered to attack was exploiting the ubiquity of cell phones to determine economic activity levels [4]. the assumption was Call Detail Records (CDRs) obtained from cellular phones provide highly granular real-time data that can be used to assess socio-economic behavior including consumption, mobility, and social patterns.

Another attempt by The world to use data for economic activity estimation was by using satellite imagery of the country of Siri Lanka [2] and linking it with local survey results to fit a model capable of doing the mapping poverty throughout the whole country only by providing high-resolution images of the country. The results were impressive, confirming the intuition that satellite imagery hold indicative information about social-economic activities:

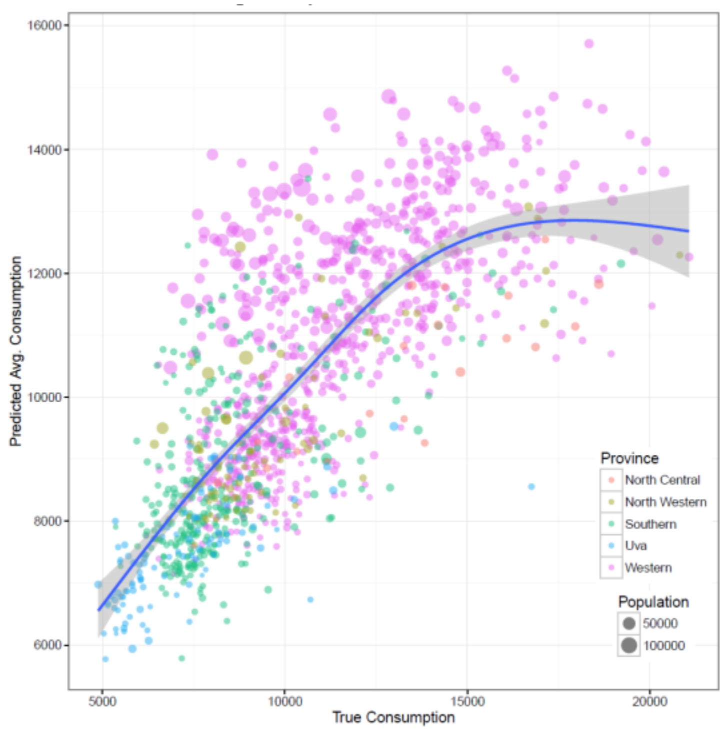


Figure 1.1: Model diagnostic plot of predicted against true average consumption

Using their model The World Bank confidently mapped poverty in the country of Siri Lanka :

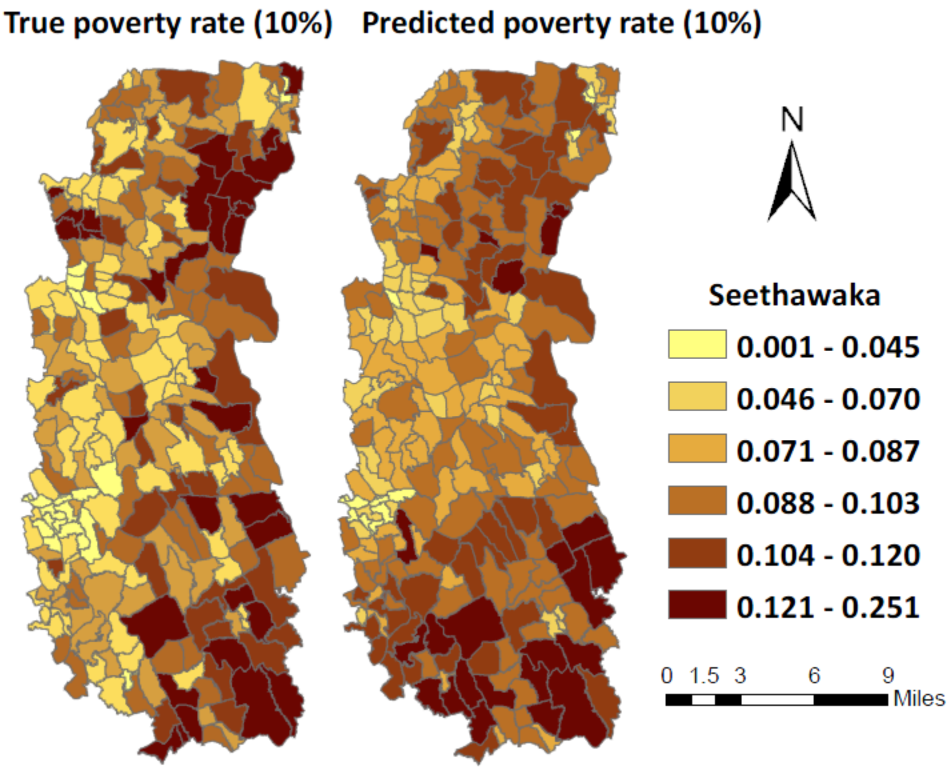


Figure 1.2: Distribution of predicted vs true poverty in Siri Lanka

One more import paper in this conxtext, is the one published by a grad student from Stanford [6].The paper implemented a unique and innovative method to take advantage of satellite imagery, nightlight intensities, and survey datasets to map poverty in poor African countries. Using sing survey and satellite data from five African countries Nigeria, Tanzania, Uganda, Malawi, and Rwanda, the paper show how a convolutional neural network can be trained to identify image features that can explain up to 75% of the variation in local-level economic outcomes. The proposed methods method, which requires only publicly available data, could transform efforts to track and target poverty in developing countries.

The results of this method were satisfying given they only used publicly available data:

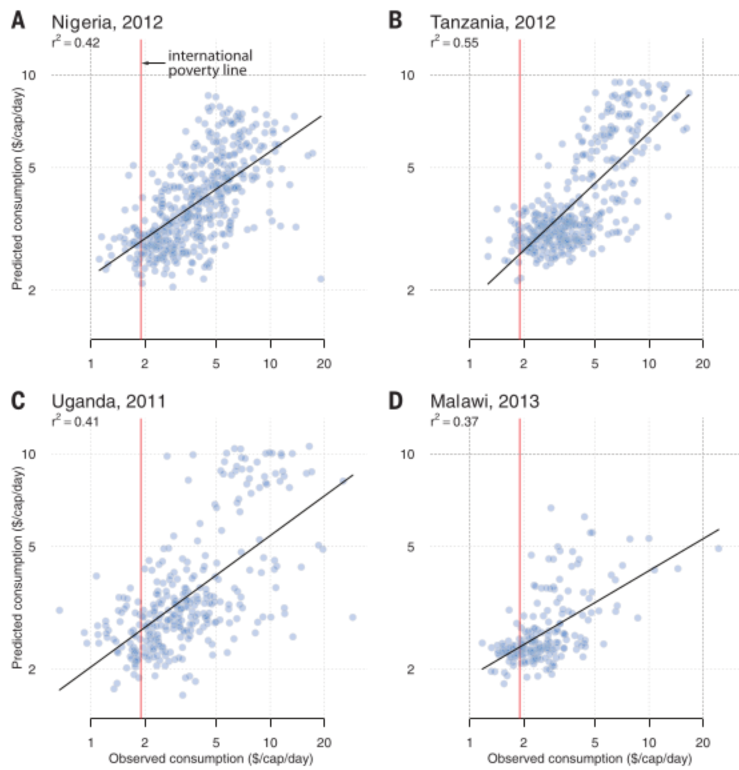


Figure 1.3: Predicted cluster-level consumption from transfer learning approach (y axis) compared to survey-measured consumption (x axis).

As we discovered in this chapter, many entities are interested in solving this lasting and difficult problem. The next chapter will the advances in computer vision and machine learning which will help immensely in the realisation of this project.

Chapter 2

State of the art : Machine Learning And Computer Vision

In this chapter we will explore the different advances in computer vision and machine learning, namely CNNs (Convolutional Neural Networks).

Machine learning

Machine learning has come a long way in the past few decades, with just a simple idea behind, ML can nowadays solve problems that until only recently were thought to be impossible for a machine to process, namely driving a car, talking and generating phrases like a human and much more.

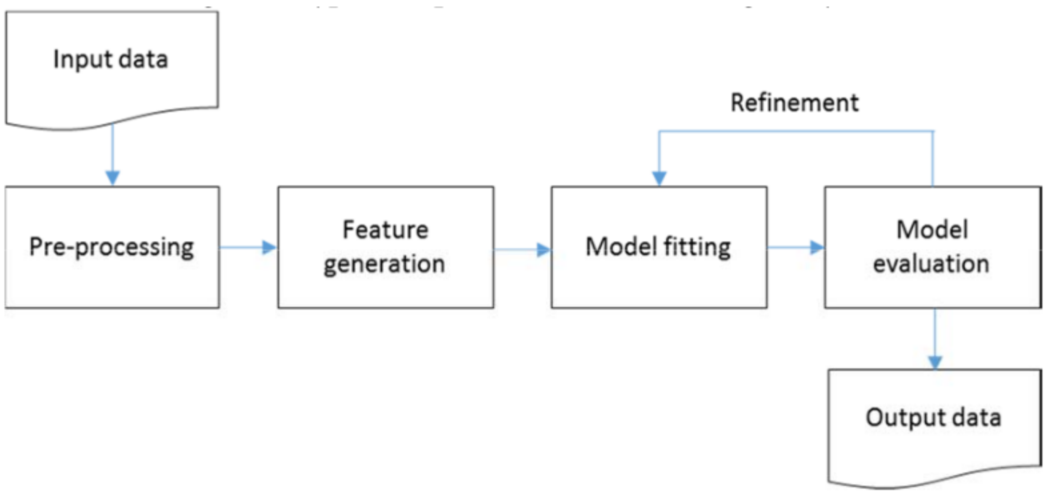


Figure 2.1: Typical Steps in a Machine-Learning Analysis

Convolutional Neural Networks

Convolutional Neural Networks are the most important artificial neural network architecture today for almost any computer vision and image processing-related AI tasks. CNNs are regularized versions of multilayer perceptrons. Multilayer perceptrons usually mean fully connected networks, that is, each neuron in one layer is connected to all neurons in the next layer. The "full connectivity" of these networks make them prone to overfitting data. Typical ways of regularization, or preventing overfitting, include: penalizing parameters during training (such as weight decay) or trimming connectivity (skipped connections, dropout, etc.) CNNs take a different approach towards regularization: they take advantage of the hierarchical pattern in data and assemble patterns of increasing complexity using smaller and simpler patterns embossed in their filters. Therefore, on a scale of connectivity and complexity, CNNs are on the lower extreme.

ResNets

CNNs architectures are released regularly giving better and better results every time. The most known architecture family are the ResNets: Resnet [3] is a deep convolutional neural network architectural design. it is considered as the most popular CNN architecures with over 20,000 citations.

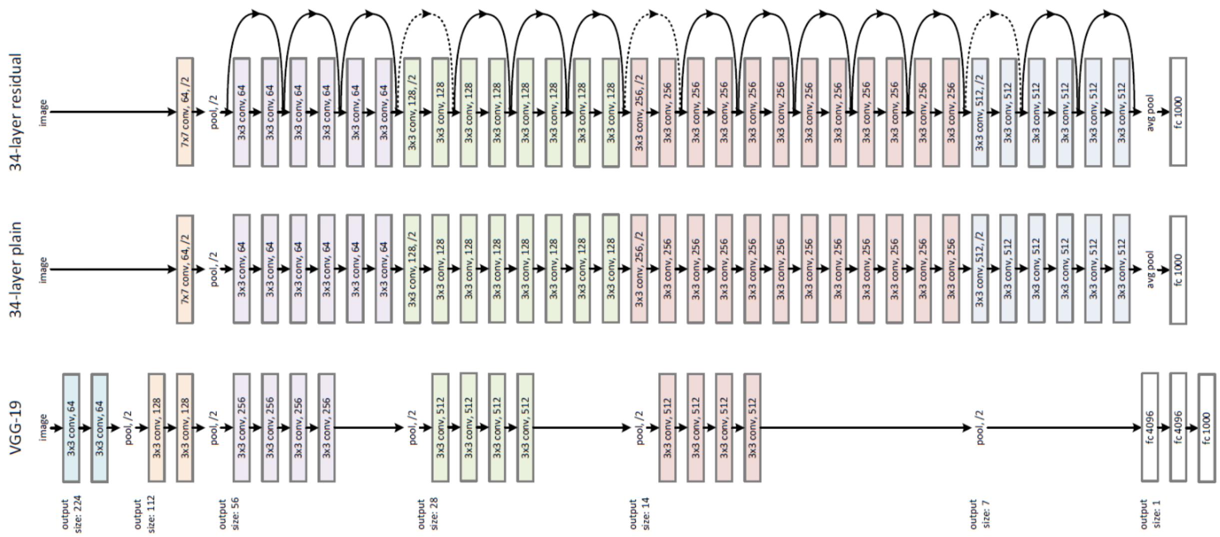


Figure 2.2: The ResNet32 CNN architecture

AlexNets

AlexNet architecture won the Imagenet large-scale visual recognition challenge in 2012. The model was proposed in 2012 in the research paper named Imagenet Classification with Deep Convolution Neural Network by Alex Krizhevsky and his colleagues.

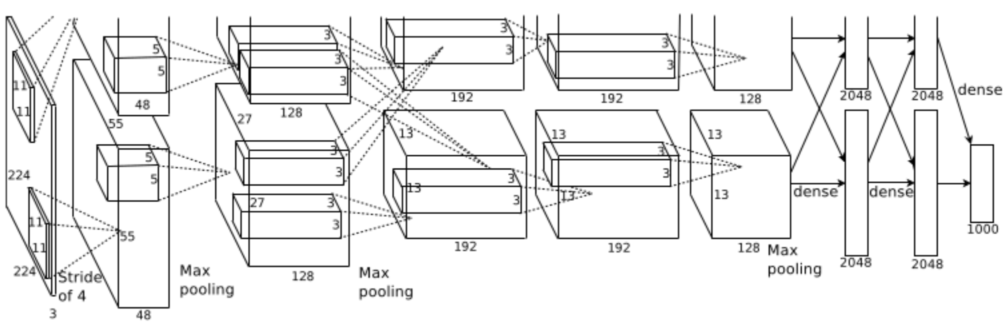


Figure 2.3: AlexNet Architecture

Transfer learning

Transfer learning is a machine learning method where a model developed for a task is reused as the starting point for a model on a second task.

It is a popular approach in deep learning where pre-trained models are used as the starting point on computer vision and natural language processing tasks given the vast compute and time resources required to develop neural network models on these problems and from the huge jumps in skill that they provide on related problems.

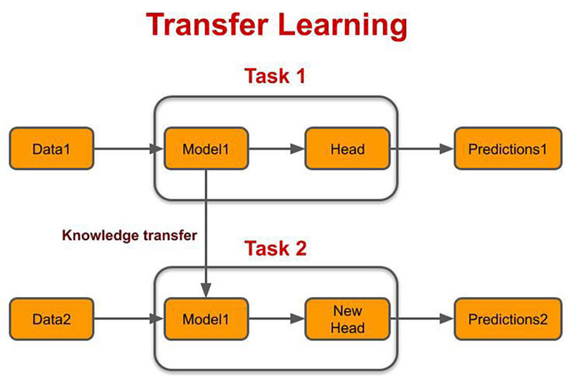


Figure 2.4: Illustration of Transfer Learning

Most deep learning practitioners nowadays use pretrained models on large datasets such as ImageNet [1]. and adjust the neural network's weight to their particular problem.

Computer vision

Computer vision has been a field of study for decades now, giving computers the ability to see is a long standing goal for researches. and the recent advances are very promising. Computers are now able to describe with confidence the content of an image, give the correct decision to a car to drive straight turn left or right and much more. And the advances in computer vision can be split into multiple categories:

Object Detection

Object Detection consists of identifying the objects presents in a given image from a list of predefined objects.

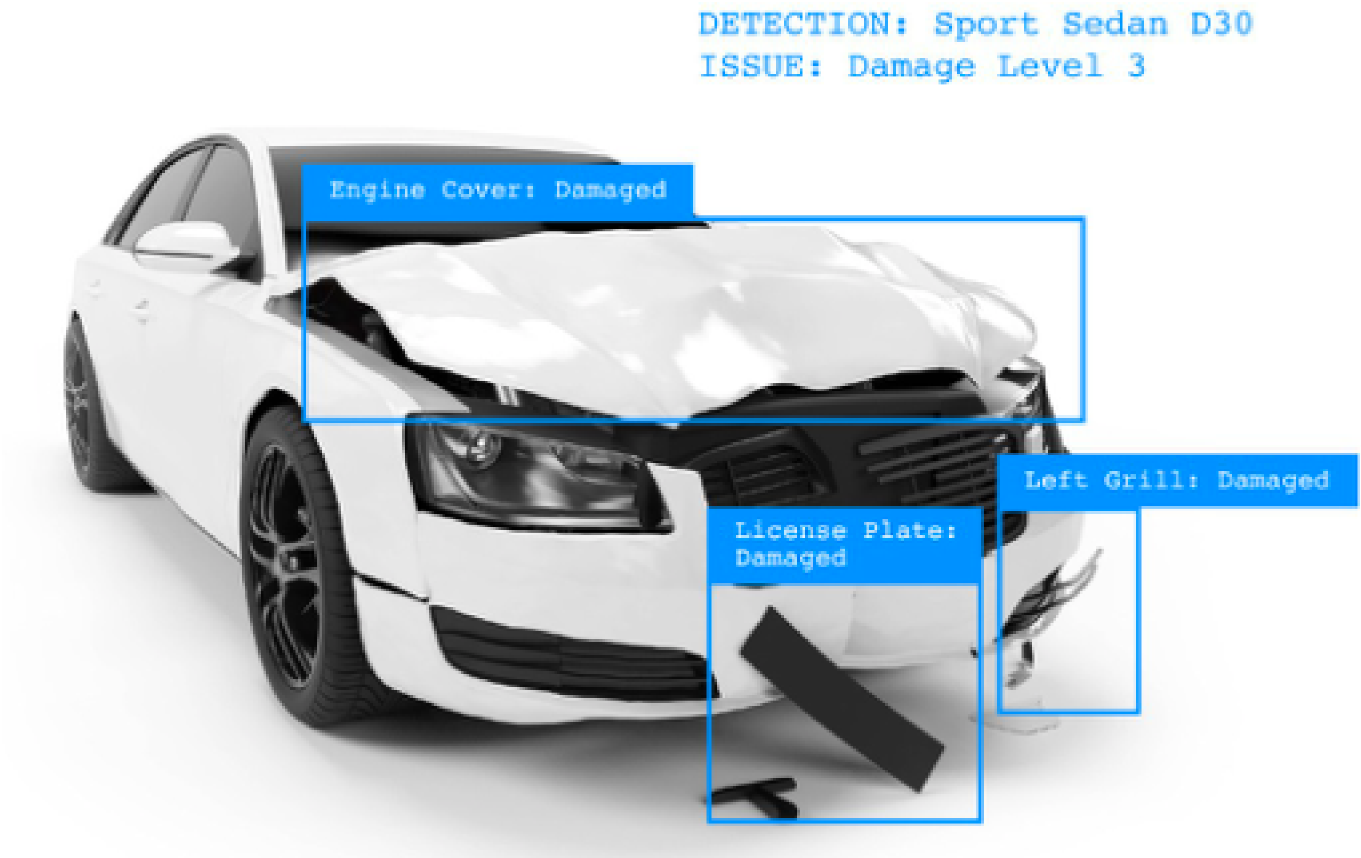


Figure 2.5: Object detection in computer

Image segmentation

Image segmentation consists of grouping pixels into meaningful or perceptually similar regions.

Figure 2.6: Image segmentation in computer vision

Deep Tracking

Locating an object in successive frames of a video is called tracking.

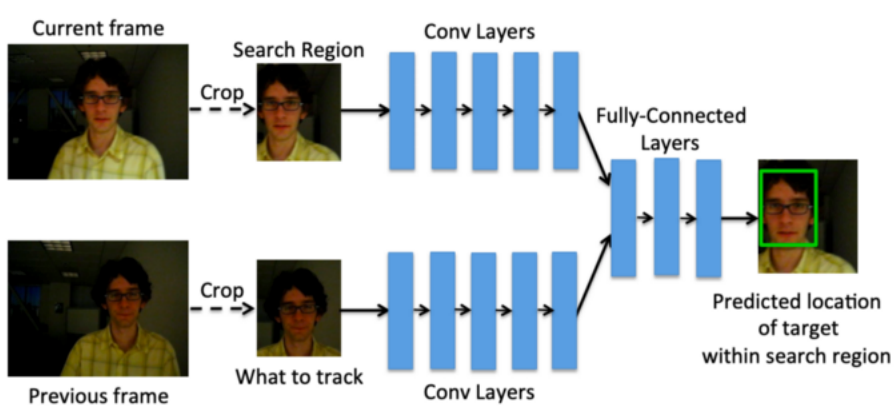


Figure 2.7: Tracking a face in a video

Generative Adversarial Networks

Generative modeling is an unsupervised learning task in machine learning that involves automatically discovering and learning the regularities or patterns in input data in such a way that the model can be used to generate or output new examples that plausibly could have been drawn from the original dataset.

Figure 2.8: Images generated using GANs

Chapter 3

Contribution

In this chapter, we will discuss our contribution and the methods we used to combine satellite imagery and machine learning to estimate economic activity in the regions of Morocco, as well as the results of our trained model. It is worth mentioning that is our dataset contains only point from the north part of Morocco leaving the south part as an upcoming area of study.

3.1 Dataset description

In order to get a labeled dataset of satellite images, we had to find survey data that included economic activity indicators and provided gps location of each observation (interrogated household). The DHS (Demographic and Health Surveys) Program corresponded perfectly to our needs. The Demographic and Health Surveys (DHS) Program is responsible for collecting and disseminating accurate, nationally representative data on health and population in developing countries. The project is implemented by ICF International and is funded by the United States Agency for International Development (USAID) with contributions from other donors such as UNICEF, UNFPA, WHO, and UNAIDS. Only incovience was that the latest Moroccan dataset in DHS dates back to 2004.

The DHS methodology consists of going door-to-door and prompting people to fill a survey containing several questions about different topics: violence, employment, famine...
And they pinpoint every household they question using GPS providing at the end the center point of each cluster of households with an added error of 2km on average to keep people's privacy. A cluster of households is defined as a varying number of households that share the same characteristics, the same economic activity

in our case.

The 2004 Moroccan DHS contains the information of 11513 households that belong to 480 clusters. the distribution of which is shown below :

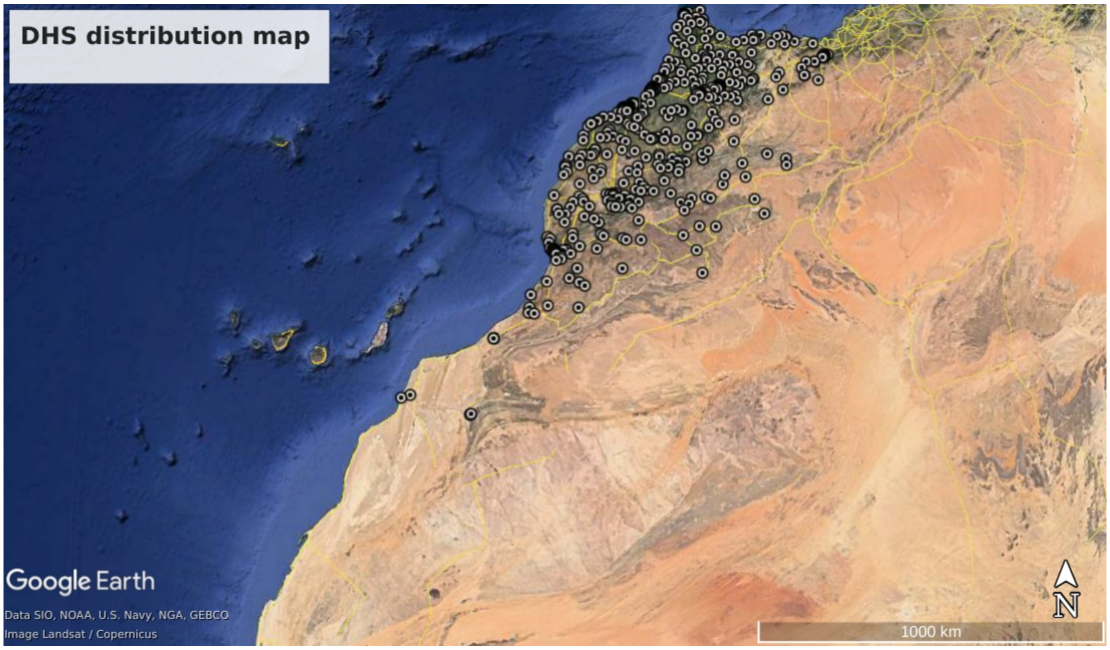


Figure 3.1: DHS Morocco 2004 survey GPS data

We can see that the of the households are condensed in the big cities.

The dataset has many variables for each household but we will leave only the economic activity level, which is a class from 1 to 5.

HV001	Latitude	Longitude	HV270
1	27.090887	-13.417127	3
2	27.157745	-13.189701	5
3	26.742351	-11.681967	4
4	26.762514	-11.652784	3
5	29.170336	-9.718877	4

Table 3.1: Samples from the DHS dataset

Let's look at the distribution per class:

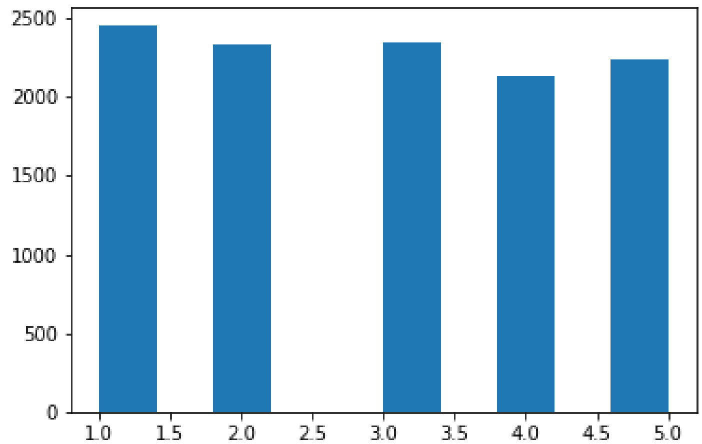


Figure 3.2: DHS Morocco 2004 survey GPS data

We see that the households are distributed equally over the 5 classes which is crucial to avoid over-fitting.

Now we have to get images of every point the DHS dataset. We used *Google Earth Engine* API to extract Sentinel-2 images of the given area. The images are 400x400 pixels and have a resolution of 10m per pixel which is a fairly high quality that allows for the details to be seen.

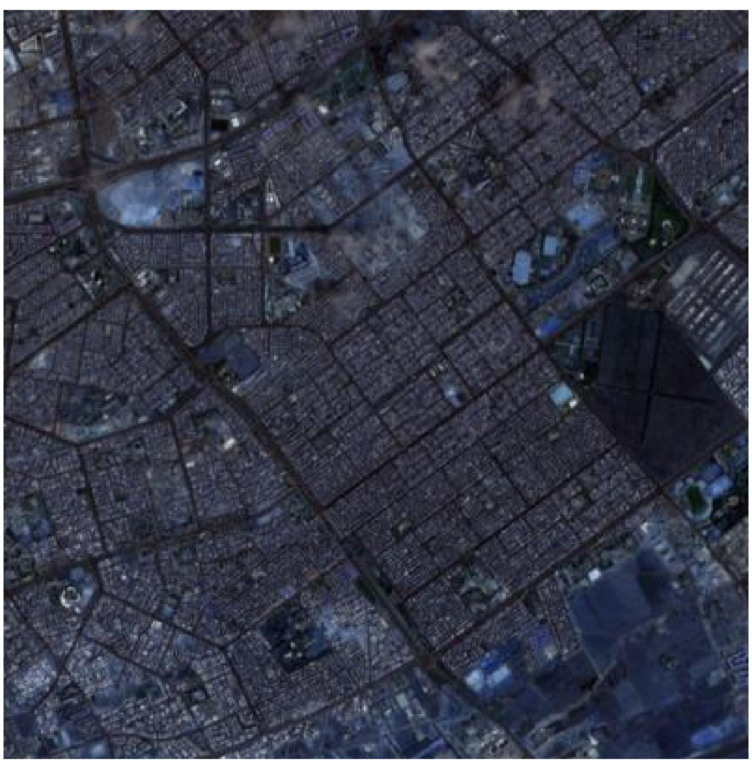
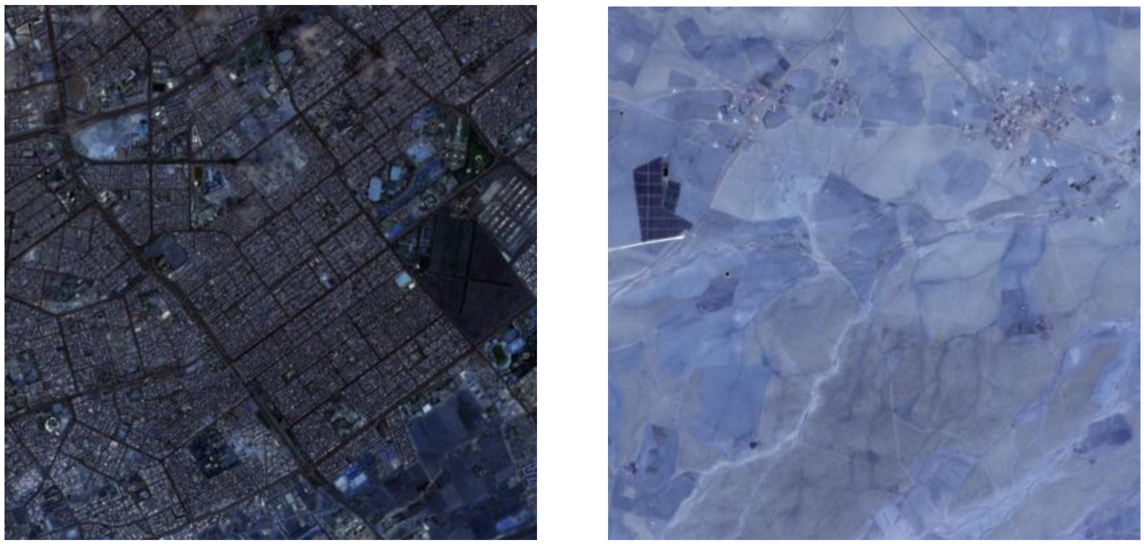


Figure 3.3: Satellite Image of an area of economic activity class 5

The images dates varies from 2015 to 2017 which is far from 2004 the release date of the DHS dataset. This can prove to be problematic in the training phase because of the mismatch of images and labels.



(a) Class 5 (b) Class 1

Figure 3.4: 2 satellite images from classes 1 and 5

By looking at the figure above, it is possible to visually distinguish between the images of different economic activity classes.

Satellite Imagery Format

There are five types of resolution when discussing satellite imagery in remote sensing: spatial, spectral, temporal, radiometric and geometric. For reasons of simplicity when stripped the images from all the bands except for the colors bands: RED, GREEN, BLUE. Compressed in a file format called GeoTIFF. GeoTIFF allows storing additional information about the image that just the pixels value as with PNG and JPG formats. additional information implies additional storage size making it tricky to store and manipulates this type of files.

Nightlight intensity data

To extract nightlight intensity data we used the VIIRS (The Visible Infrared Imaging Radiometer Suite) dataset, that is a public dataset that gets released by NASA on a monthly/yearly basis and contains nightlight intensities of the entire globe.

Figure 3.5: Visualization of VIIRS dataset over Europe

We limited our study to north Morocco and linked each point from the DHS survey dataset to its associated nighttime light intensity.The nighttime light intensity is calculated by averaging the pixels values of the 10km x 10 km surrounding the centroid.

HV001	cluster_lat	cluster_lon	HV270	nightlights	nl_class
1.0	27.090887	-13.417127	3.0	0.917380	Medium
2.0	27.157745	-13.189701	5.0	4.769447	Medium
3.0	26.742351	-11.681967	4.0	1.443238	Medium
4.0	26.762514	-11.652784	3.0	1.443238	High
5.0	29.170336	-9.718877	4.0	0.258637	Low

Table 3.2: Sample from the complete dataset (with nighlight classes)

The intuition behind using nightlight intensity for economic activity is that, the more active an area is the more luminous it will be at night. Let's test that hypothesis:

```
##Enconding the classes
all_data_with_nl.loc[all_data_with_nl['nl_class'] == 'High','nl_class_enc'] = 3
all_data_with_nl.loc[all_data_with_nl['nl_class'] == 'Medium','nl_class_enc'] = 2
all_data_with_nl.loc[all_data_with_nl['nl_class'] == 'Low','nl_class_enc'] = 1
corr=all_data_with_nl['HV270'].corr(all_data_with_nl['nl_class_enc'])
print("Correlation between nightlight classes and economic activity levels is : " ,corr)
```

```
Correlation between nightlight classes and economic activity levels is :  0.5740534331706977
```

Figure 3.6: Correlation between NL class and EA class

0.57 is understandable given that the nighttime light intensities and economic activity labels are from different years.

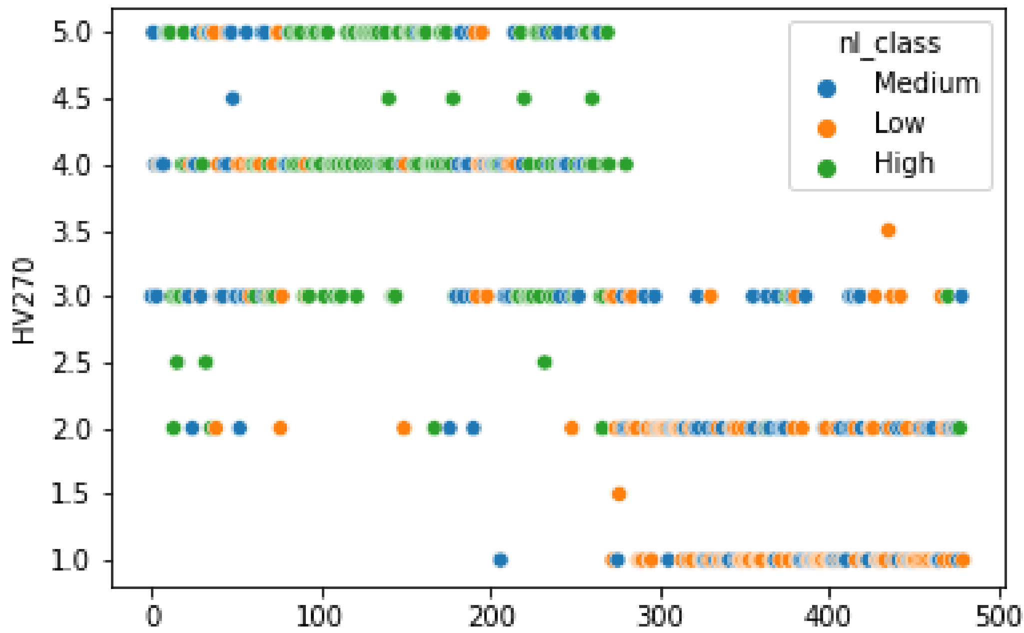


Figure 3.7: Visualisation of NL class versus EA class

Let’s look at the distribution of images per class:

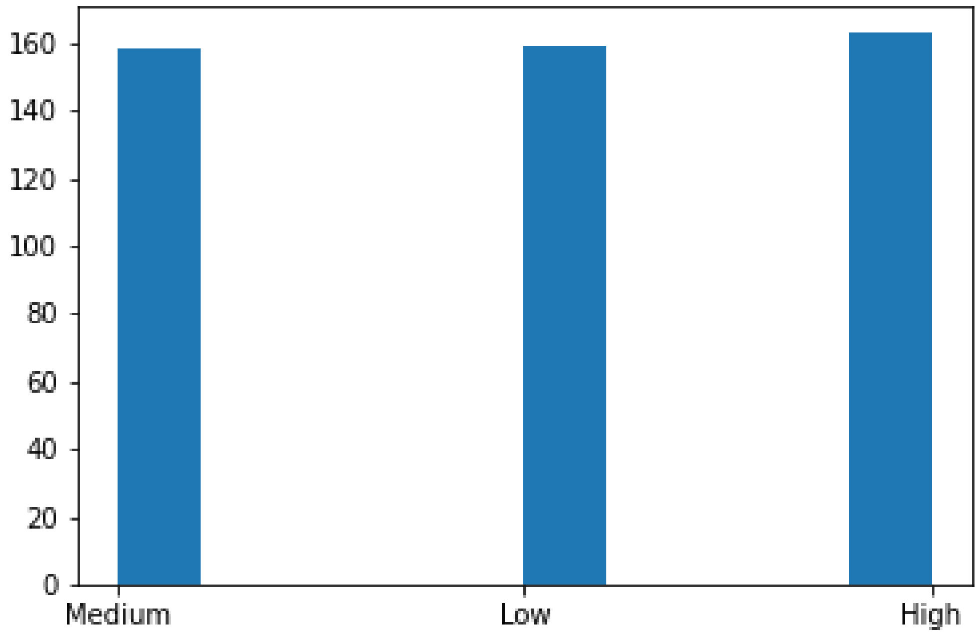


Figure 3.8: Distribution of images per class

We see that the images are equally distributed, this was done on purpose to avoid overfitting.

3.2 Training the model

Using direct economic activity indicator

In this part, we will train our model on the gathered data using the different mentioned architectures.
the training was done using the Fastai [5] python library. which is a deep learning library which provides practitioners with high-level components that can quickly and easily provide state-of-the-art results in standard deep learning domains. Because the images' resolution is big, working with it directly is not possible due to the computional cost. Therefore all of the images are resized to 224x224 to make the training faster. After the data preparation and loading, we look for the best learning rate value:

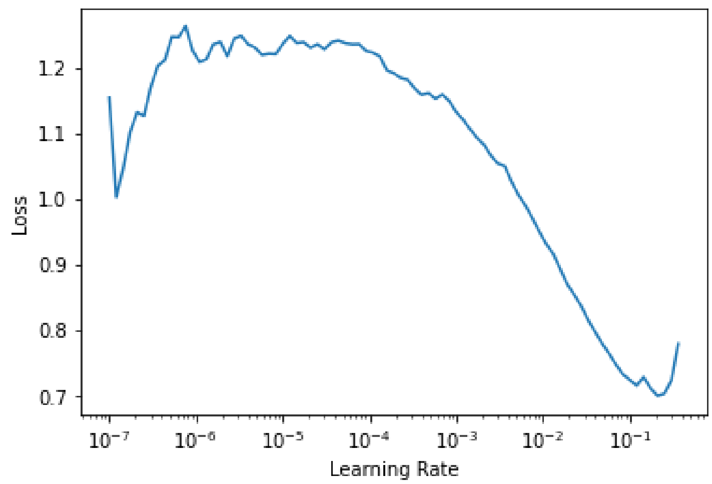
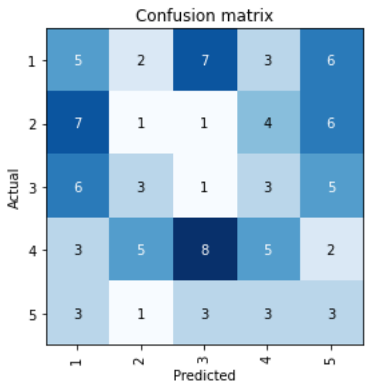


Figure 3.9: Different values of learning rate and the associated loss

We reserve 20% of the data for validation. The obtained confusion matrix is as follows:



(a) Confusion matrix

epoch	train_loss	valid_loss	accuracy	time
0	2.587138	2.041600	0.187500	01:50
1	2.670662	1.973483	0.197917	00:03
2	2.667472	1.935776	0.281250	00:03
3	2.623023	1.962516	0.229167	00:03
4	2.616668	1.967270	0.229167	00:03

(b) Epochs

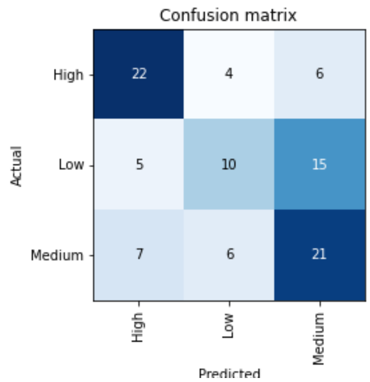
Figure 3.10: Training metrics

Using nightlight intensities as a proxy

In this part we will train our model using the gathered nighttime light intensity data. The classification problem is formatted as follows, each cluster is given a nightlight class based on its nighttime light intensity, either Low, Medium or High. The assumption here is that we are capable of mapping the nighttime light classes to actual economic activity classes.

The training was done using the same python library Fastai [5], with the ResNet32 architecture and with a model that has been pretrained on ImageNet [1].

the results are as follows:



(a) Confusion matrix

epoch	train_loss	valid_loss	accuracy	time
0	1.850294	1.382092	0.395833	00:03
1	1.848534	1.345437	0.427083	00:03
2	1.838774	1.289580	0.458333	00:03
3	1.822479	1.252413	0.500000	00:03
4	1.804737	1.222162	0.489583	00:03
5	1.781038	1.191822	0.500000	00:03
6	1.769217	1.176457	0.531250	00:03
7	1.716184	1.165709	0.531250	00:03
8	1.715602	1.164354	0.500000	00:03
9	1.707043	1.159600	0.531250	00:03

(b) Epochs

Figure 3.11: Training metrics

We notice that the accuracy gets signaficatly better that with the previous problem.

3.3 Discussions

Data Is Inaccurate. linking images from 2015 to label from 2004 deemed the results of the model to be low. a solution for this might be to find images from 2004 which is not very helpful for current authorities. Another solution is to find an alternative survey data source that is more up to date.

Images are corrupted an important chunk of the collected images are either cloudy or corrupted due to the fact that they are on the edge of the full image. the solution for this is taking those corrupted images out and replacing them with clean ones.

Images are of low quality. because these are publicly available, they aren't the best in terms of quality. at best the quality of the images are 10m per pixel barely shows any details that explain the variance in economic activity.

In this chapter we saw the limitations of our method as it is highly dependant on the quality of the provided datasets and images.

Conclusion and perspectives

Conclusion

This project's main goal is to find an alternative to the old methods used for estimating economic activity in Morocco, namely census survey data which is expensive and thus doesn't get updated that often. In the case of Morocco, the general census is done every ten years. Which makes it difficult to evaluate economic developments in real-time. Many other countries experience the same scarcity in Methods that serve this purpose, and consequently, financial and social entities and researchers have focused their attention the find a creative data-driven approach that estimates correctly the economic activity in a given area. The fruits of such efforts are several scientific papers that investigate many paths ways to this goal, including satellite imagery, cellphone data, and much more. Many of these methods are replicable in Morocco due to the similarity in socio-economic features Morocco has with the subject countries of these papers. In this project, we decided to combine several approaches, namely from the Standford University paper [6] and from the World Bank research team [2].

The process of defining the problem and identifying the needs and resources took the biggest chunk of time because we understood that with only clean data that we can have results. We started by looking for survey data from the last general census, but we had to abandon that since no geo referencing is provided of each household, which is crucial for our purposes since we want to get satellite images of that pinpoint. One alternative solution was using the DHS (Demographic and Health Survey) program Moroccan dataset. The only drawback was that the most recent Moroccan dataset dates back to 2004.

The second challenge was to find high-quality satellite imagery of the points provided in the DHS dataset, many paid and free options were considered. namely the google earth engine API and the Mohammed VI satellite, which are free and paid respectively. Mohammed VI satellite images are x20 times higher quality than the free alternatives. But because of the high cost of the images, we had to go for the google earth engine's sentinel-2 dataset.

After successfully assigning the correct label to each image, the environment was prepared for machine learning. The results were disappointing because of numerous reasons discussed in the contribution chapter.

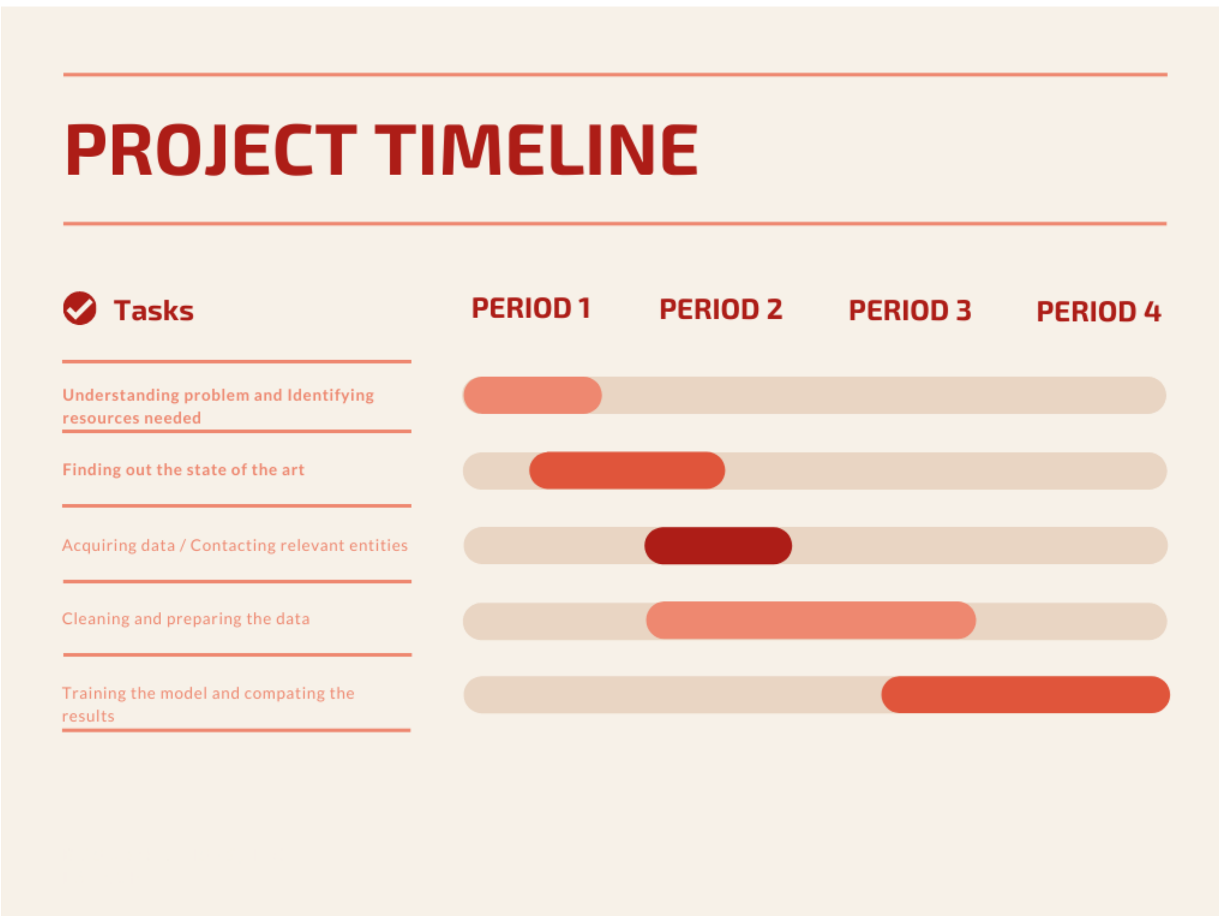
Perspectives

In light of the disappointing training results, we realize that the model's performance is highly dependant on the accuracy of the images and labels. Furthermore, it depends on the resolution of the images. Therefore a way to overcome this is by limiting the points of study to smaller regions and buying the associated images, which will make the costs reasonable.

Another improvement point is experimenting with economic activity proxies instead of working directly with the indicators provided in the DHS dataset. To name a few proxies: property value, nightlight intensity, cellphone usage...

Appendix A

Appendix



Bibliography

[1] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, 2009.

[2] Newhouse David Locke Engstrom Ryan, Hersh Jonathan Samuel. Poverty from space : using high-resolution satellite imagery for estimating economic well-being (english). *The World Bank*, 548(7665):43–51, 2017.

[3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015.

[4] Frias-Martinez Vanessa Whitby Andrew Frias-Martinez Enrique Hernandez Marco, Hong Lingzi. Estimating poverty using cell phone data evidence from guatemala. *The World Bank*, 548(7665):43–51, 2018.

[5] Jeremy Howard and Sylvain Gugger. Fastai: A layered api for deep learning. *Information*, 11(2):108, Feb 2020.

[6] Neal Jean, Marshall Burke, Michael Xie, W Matthew Davis, David B Lobell, and Stefano Ermon. Combining satellite imagery and machine learning to predict poverty. *Science*, 353(6301):790–794, 2016.