Google Earth Predicts School Success

Luke Denoncourt [ldenonc@iu.edu], Mustafa Keskin [mkeskin@iu.edu], Ravi Sharma, [ravishar@iu.edu]

Engr E533, Deep Learning Systems, Indiana University

Introduction

Artificial Intelligence and Machine Learning with the surrounding worry of whether they are helpful or hurtful tools [1-2] is always up for debate. Besides ethical, legal, and social concerns, Deep/Convolutional Neural Networks (DNN) capabilities are limitless and get the most attention from the public [3]. There is no doubt about DNN's unprecedented rise and its dominance in electronics, scientific research, software, and hardware. With the help of developers, custom tools and libraries, such as ONNX, Caffe2, PyTorch, MXNet, Keras, and Tensorflow, are developed and upgraded constantly and rigorously, hence, widely used by the public [4-6].

Lately, with the computational power of high-tech companies [7], DNN applications in computer vision and image datasets for classification, image generation, and prediction have become a click of a button on our smartphones, tablets, or laptops. Use cases vary greatly from road, vehicle, and hazard monitoring for self-autonomous driving to fake scene generation and image correction in social media posts [8-9]. These are only a very few application examples of what DNN can do nowadays.

Separately, satellite images may offer a powerful tool for monitoring the most valuable resource, that humans have, Earth. It allows us to sense and track air, land, water, ice, and plantations. These images are used by governmental and private institutions for a variety of applications such as meteorology, oceanography, agriculture, forestry, geology, mapping, area assessment, intelligence, warfare, and education [10-11]. Recent studies are researching if satellite images of a particular building, or a block of buildings carry any interesting information buried deep inside [12-15].

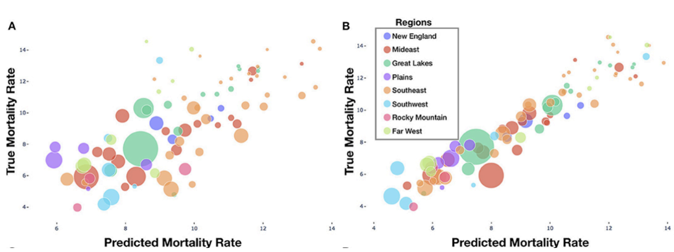

Figure 1: Exemplary results from Ref. [13]

Ref. [12] examined if DNN can predict the success of students based on satellite images of schools in Brazil and the Philippines. One can argue that there may be no relation between scores of students to types of buildings, or roofs of buildings seen by satellite. However, there might be indirect relations between the quality of the buildings, the isolation of campuses, the number of buildings to the status, or living conditions of the community. This hidden relationship can be explored by the computational complexity of DNNs.

In a similar timeframe, Ref [13] studied if there is any relationship between the mortality rate in the community and satellite images of school surroundings in US communities. Exemplary results of Ref. [13] are shown in Figure 1. Authors in both papers have done a significant amount of preprocessing of data both image-wise also blending with different text or results like one hot vector approach to improve the robustness of the analysis [12-13].

1. Project Description

In this paper, we investigated if DNN can predict school success by using satellite images of school surroundings in the State of California. This study is built from ground zero including collecting data, cleaning, processing, and implementing DNN as shown in the framework of Figure 2.

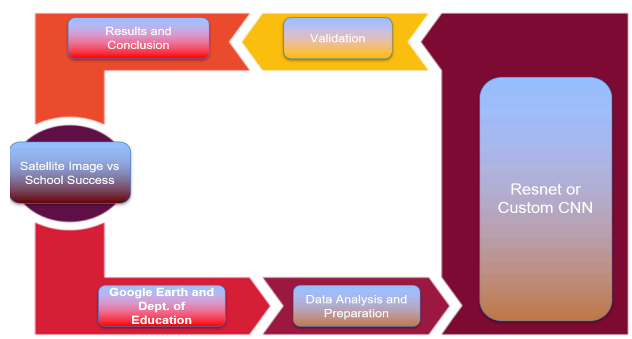

Figure 2 Methodology and Framework of the project.

During the initial phase of the project, all team members studied references and familiarised with the research question. In the second phase of the project, one of the team members focused on extracting and downloading school images from the Google API. The second team member worked on finding school success data from the Department of Education of the USA as well as from the California Department of Education. The third team member studied projects done in similar research areas and worked on getting ResNet ready for the final phase of the project. In the third phase of the project, we still partitioned the workload as capturing new satellite images for two sets (Zoomed-in/out), coding and implementing ResNet models and writing the paper. For the last 3-4 weeks, we maintained constant communication by few emails per day, a Slack channel, and Zoom meetings (on average one per 3-4 days and almost daily Zoom before the submission).

In summary, the architecture of the proposed project is captured in Figure 3, where satellite images of schools (three San Diego schools for illustration only) enter the ResNet and are processed through this DNN. Please note that the initial weights of the ResNet are downloaded, then, all layers are made trainable. Eventually, the plan is to estimate school success at the output layer as a percentage.

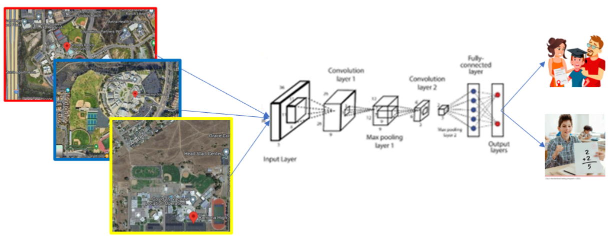

Figure 3: Project Architecture

2. Data Collection and Analysis

There are two different types of data collection in this project. The first one is related to the success of schools and the second one is related to satellite images.

2.1 School Success Data Collection and Analysis

Data for School academic success was collected from Ref [16] for the school year of 2014-2015 in terms of scores of maths achievement on standardized tests as well as the total number of students taking the exams. For this study, public schools were selected from California to compare students who theoretically all took the same math exam, as different states could administer different exams, while also still including the largest sample size available. There are 9 columns in this data including, the name of the state, district ID, district name, school ID, the name of the school, date of the exam, the total number of students who took the proficiency exam, and the percentage of the students who passed the proficiency exam.

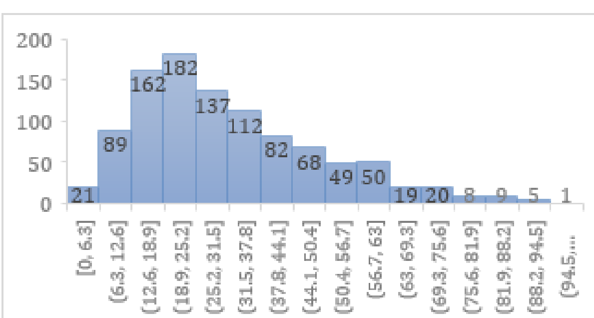

Figure 4: Percentage of Students above the proficiency

A histogram of the percentage of students passing the proficiency is shown in Figure 4 and it is observed as a Gaussian distribution with a low skew on the upper scale.

Since addresses and zip codes are missing in the success data, locations of the CA schools were gathered from the school directory website from the California Department of Education website [17]. The final dataset was also a subset of schools having at least 100 test takers. Due to schools having as low as one test taker and the output of successful passing of the test was in terms of percentages, the schools with a lower number of test takers would have greater variability in the data due to random chance, thus reducing model accuracy when predicting off the satellite images. A histogram of the number of students taking the exam is shown in Figure 5.

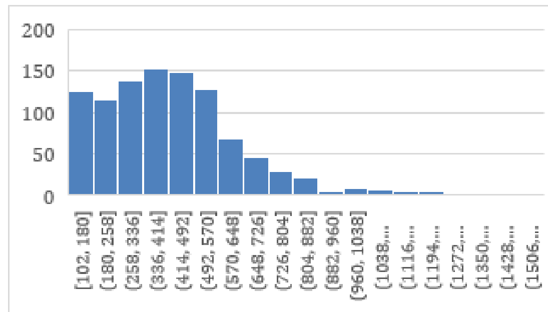

Figure 5: Number of students taking the exam per school

2.2 Satellite Imagery Data Collection and Analysis

The desired satellite image data would include differing levels of zoomed-in images containing no cloud coverage and clear resolution of schools across our dataset. The source of data fitting those criteria would come from the Google Earth Engine. Originally, Landsat images such as the Sentinel Landsat dataset were considered, but the image resolution was not refined enough for the intended zoom level of this study [18]. The Google Map API was also refined for Java developers and therefore not an option to gather data. The last remaining option was using the package Geemap, "a python package for interactive geospatial analysis and visualization with Google Earth Engine", in an unconventional manner.

Geemap is an interactive Google Earth Engine map interface that contains numerous types of base maps and the ability to superimpose Landsat images onto the map [19]. The map of interest contained the ESRI World Imagery satellite map layer, which contains, at a maximum, 0.5-meter resolution images for the United States within the last three to five years [20]. GPS coordinates of schools were plotted in the Geemap using the ESRI World Imagery layer and retrieved via screenshots after downloading the corresponding Geemap HTML file at two differing image scales.

Figure 6: A sample satellite view of the school image

The first images were taken with a total area of 6,820 km^2, originally 1890*755, and then scaled to 244*244. The second more zoomed-in images have a total area of 856 km^2, with 1890*755 size scaled to 244*244, as shown in Figure 6, i.e. the ideal dimensions for the ResNet model [21-22].

As suggested by Prof. Kim, there are also two versions of the school surroundings captured. These two versions are zoomed-in and zoomed-out versions as shown in Figure 7 (please note that they are rescaled below to fit to single figure). In total, close to 2,200 images were captured including both types of aerial views. However, only ~900 for each were kept for DNN due to the availability of test scores.

Figure 7: Zoomed-in and out versions of a sample

3. Resnet Implementation and Results

ResNet50 is used for this specific study as it can capture diverse patterns and variations in the data and is a well-known and widely used architecture [12-15]. ResNet-50 has, as the name suggests, 50 layers. These layers consist of 5 blocks, and each block has a group of residual structures. These so-called residual blocks provide a hysteresis or memory for saving information from earlier layers. Eventually, this will help the network to learn better representations of the input data. In addition, there is also 1x1 convolution, which allows faster training bypassing some portion of the internal layers. Eventually, this resolved gradient vanishing issues and improved the training overall.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 resnet50 (Functional)       (None, 7, 7, 2048)        23587712

 global_average_pooling2d (  (None, 2048)              0
 GlobalAveragePooling2D)

 dense (Dense)               (None, 256)               524544

 dropout (Dropout)           (None, 256)               0

 dense_1 (Dense)             (None, 1)                 257

=================================================================
Total params: 24112513 (91.98 MB)
Trainable params: 24059393 (91.78 MB)
Non-trainable params: 53120 (207.50 KB)
```

Figure 8: The first sDNN model

Before directly inputting the data into the ResNet model, image augmentation was applied using the imgaug library in Python. The augmentation methods employed encompassed horizontal and vertical flips, image rotation of -10 to +10 degrees, as well as adjustments to brightness and contrast. We implemented a total of six ResNet50 models, at the same time, the model architecture for all six models remains almost identical with ResNet50 serving as the base model in Figure 8 along with a Global Average Pooling layer, a dense layer with 256 neurons with Relu activation, a Drop Out layer with 0.2 dropout rate, and the final layer with one neuron and linear/sigmoid activation depending upon target values. The variations in the models include input data, loss function, target prediction values, and the use of image augmentation on the input data.

Details of the models are as follows:

1) Model 1
   a) Input: Zoomed-in Images of size 244*244
   b) Target: The percentage of students that scored above proficiency
   c) Validation Dataset: 20%
   d) Loss: Mean Squared Error
   e) **Result:** Got a Validation MAE of 15.29 which indicates the predicted values lie ±15% of the actual values
2) Model 2
   a) Input: Zoomed-in Images of size 244*244
   b) Target: Values above or below the mean
   c) Validation Dataset: 20%
   d) Loss: Mean Squared Error
   e) **Result:** Got a Validation MAE of 0.18
3) Model 3
   a) Input: Zoomed-in Images of size 244*244
   b) Target: 0 if the score is below average and 1 if it is above average
   c) Validation Dataset: 20%
   d) Loss: Binary Cross entropy
   e) **Result:** An accuracy of around 53.85% was achieved which is not too good for a Binary Classification problem, but for this specific Model since the data didn't have a normal distribution, and the low accuracy of our model combined didn't produce the desired outputs
4) Model 4
   a) Input: Both Zoomed in and Zoomed out Images of size 244*244
   b) Target: The percentage of students that scored above passing
   c) Validation Dataset: 20%
   d) Loss: Mean Squared Error
   e) **Result:** A Validation MAE of around 14.60 was achieved, better than just the zoomed-in images. This suggests that the; surrounding areas of the school somehow play a role in influencing the number of students who pass.
5) Model 5
   a) Input: Zoomed out images of size 224*224
   b) Target: Binary classification of above or below mean passing percentage of students
   c) Validation Dataset: 20%
   d) Loss: Binary_crossentropy
   e) **Result**: Validation accuracy of 57.63%
6) Model 6
   a) Input: Combined zoomed-in and zoomed-out images of size 224*224
   b) Target: Binary classification of above or below mean passing percentage of students
   c) Validation Dataset: 20%
   d) Loss: Binary_crossentropy
   e) **Result**: Validation accuracy of 62.09%

The prediction results are shown throughout the code and one of them is presented in Figure 9.

```
In [13]:   1 test = np.expand_dims(test, axis=0)

           Predicted score for School No.1 as 63% where the actuall score was 61%

In [14]:   1 model.predict(test)
           1/1 [==============================] - 1s 790ms/step
Out[14]:   array([[63.08008]], dtype=float32)
```

Figure 9: Prediction vs actual result for model 1

Conclusion

In this study, the relationship between aerial views of the schools and the success of students using those buildings is examined. This research question was impossible to examine perhaps 20 years ago, however, with the computational power of GPUs and advanced DNN architectures, it became much easier to investigate.

One of the most difficult parts of this project is bringing the pieces together. Collecting Big data is still a huge problem, and it requires tedious, time-consuming, and laborious work to make it reliable for DNN to process. Extracting satellite images or information from web pages is becoming more difficult rather than easier even with certain Python libraries. Some of the issues we faced, were inconsistency of data between different websites (or data in itself), data protection-related obstacles, limited privileges for developers, or a wide variety of data types. In addition, every member spent a great amount of time just googling to find a trustworthy website as well as a reasonable amount of data. We chose a dataset pre-covid since COVID-19 by itself introduced a lot of unknowns to education and student success, unfortunately.

As presented in the Python code, success percentages of schools predicted by ResNet using satellite images only vary in very low single digits such as 63% to 61%. Hence, this suggests a good prediction mechanism for schools' success by using aerial images of their surroundings. Since we used both zoomed-in and zoomed-out images, we found that the surrounding areas of the school do play a role in the success of the students of that school.

Due to time limitations, we were not able to add additional data sources such as crime rates per zip code or feature activation maps into this study to evaluate the model or prediction further. Even though, we searched other data sources on the internet, most of the data is not open to the public on this particular topic or only available with payment.

Regarding the continuation of this project, the Authors suggest expanding it to all USA schools, categorizing middle, elementary, and high schools, comparing private and public schools, or implementing a hybrid Machine Learning model using different types of data such as crime rate, poverty rate, ethnicity, income rate per zip code, etc.

Acknowledgments

The authors thank Prof. Minje Kim for teaching a top-notch course covering the ins and outs of Deep Neural Networks, in addition to his time for reviewing our presentation and providing insightful suggestions about where to focus on our project and encouragement on our work. In addition, the authors are also grateful to each other for working harmoniously until the end of the project. As can be seen from Ref. [12-13], similar work had taken about 3-4 years for a complete research team to complete and eventually was published internationally in prestigious journals.

Appendix: Summary of Extra/secondary Project

As presented in the review meeting with Prof. Minje Kim, we briefly worked on a fallback project in the background. Basically, during the search for available datasets for this study, we found a Kaggle dataset that has house images from California. Few people worked on this dataset for classifying images based on prices or square footage using NN structures on the Kaggle web page [23]. Their percentage was quite low around 40% accuracy. It was observed that there are not only exterior pictures of houses but also some fake images like a map, grass field, or storefront. There are also some interior images of houses. This is identified as a main problem by us for the low percentages of earlier work not to accurately guess the price of the house. For that reason, a two-stage NN approach is proposed as shown below. This approach is implemented using 2 CNN stages. The accuracy for classification of the real exterior was around 94% in the first stage and these real images were later applied to the second CNN for predicting house price. By implementing a two-stage CNN network and resolving the issues in the consistency of the dataset, the previously reported accuracy of 40% is improved to 74%. Codes related to this extra project are two files with names starting with "EXTRA."

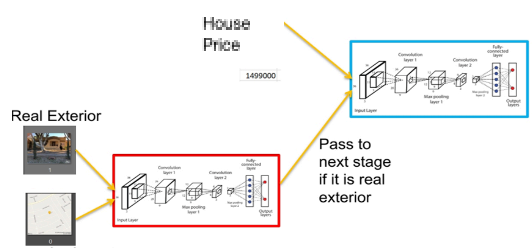

Figure 10: Two-Stage CNN Architecture

References

[1] S. Spielberg, "A.I. Artificial intelligence", Warner Bros, June 29, 2001. https://en.wikipedia.org/wiki/A.I._Artificial_Intelligence
[2] A. Proyas, "I, Robot, Davis Entertainment," July 2004, https://en.wikipedia.org/wiki/I,_Robot_(film)
[3] J. Chadwick, "The Game is Over! Google's DeepMind says it is close to achieving 'human-level' artificial intelligence," Daily Mail, UK, May 2018, https://www.dailymail.co.uk/sciencetech/article-10828641/Googles-DeepMind-says-close-achieving-human-level-artificial-intelligence.html
[4] W. G. Hatcher et. al., "A Survey of Deep Learning: Platforms, Applications, and Emerging Research Trends," in *IEEE Access*, vol. 6, pp. 24411-24432, 2018.
[5] L. Jiao et al., "A Survey of Deep Learning-Based Object Detection," *in IEEE Access*, vol. 7, pp. 128837-128868, 2019.
[6] V. Borisov et al., "Deep Neural Networks and Tabular Data: A Survey," in *IEEE Transactions on Neural Networks and Learning Systems*, p. 1-21, Dec. 2022.
[7] S. K. Moore, "Google, Intel, Nvidia Battle in Generative AI Training," *IEEE Spectrum*, 12 Nov 2023.

[8] S. Kang et al., "7.4 GANPU: A 135TFLOPS/W Multi-DNN Training Processor for GANs with Speculative Dual-Sparsity Exploitation," IEEE International Solid-State Circuits Conference, pp. 140-142, Feb. 2020.
[9] J. Liu, "Survey of the Image Recognition Based on Deep Learning Network for Autonomous Driving Car," *2020 5th International Conference on Information Science,* Shenyang, China, 2020, pp. 1-6.
[10] C. Polykretis, et. al., “Digital Educational Geoinformatics Methodologies for Monitoring Landscape – GeoLand,” March 2022, Erasmus project. https://www.geolandproject.eu/wp-content/uploads/2022/05/GEOLAND-Handbook-EN.pdf
[11] S. Ackerman, et. al., “Satellite Applications for Geoscience Education,” University of Wisconsin Madison. http://cimss.ssec.wisc.edu/sage/about/outline.html
[12] D. Runfola et. al., "Using satellite data and deep learning to estimate educational outcomes in data-sparse environments," pages 87-97, 12 Nov 2021, https://doi.org/10.1080/2150704X.2021.1987575
[13] I. Maduako et.al., Automated School Location Mapping at Scale from Satellite Imagery-Based on Deep Learning. *Remote Sens.* 2022, *14*, 897. https://doi.org/10.3390/rs14040897.
[14] J. J. Levy et. al, "Using Satellite Images and Deep Learning to Identify Associations Between County-Level Mortality and Residential Neighborhood Features Proximal to Schools: A Cross-Sectional Study," Front Public Health, 2021 Nov 5, https://doi.org/10.3389/fpubh.2021.766707.
[15] Special Issue "Land Use Classification with GIS and Remote Sensing Data Based on AI Technology.
[16] Ed-Facts Data Files, The Department of Education, USA, https://www2.ed.gov/about/inits/ed/edfacts/data-files/index.html
[17] School Directories, The California Department of Education, https://www.cde.ca.gov/ds/si/ds/pubschls.asp
[18] Google Earth API Developer, https://developers.google.com/earth-engine/datasets/catalog
[19] A Python package for geospatial analysis, https://geemap.org/
[20] World Imagery, https://www.arcgis.com/home/item.html?id=10df2279f9684e4a9f6a7f08febac2a9\
[21] Geemap, Common Module, https://geemap.org/common/?h=zoom_level#geemap.common.zonal_stats_by_group
[22] Map Control, Zoom Levels, https://blogs.bing.com/maps/2006/02/25/map-control-zoom-levels-gt-resolution
[23] House Prices and Images – SoCal, https://www.kaggle.com/datasets/ted8080/house-prices-and-images-socal