DEPRIVATION INDEX PREDICTION USING DEEP LEARNING

Coursework 3

Student ID: 251238730

FM9528 - Banking Analytics

Word Count: 2197

December 17, 2021

TABLE OF CONTENTS

1. Executive Summary

Light detection and ranging (LIDAR) is a remote-sensing technology that uses light in the form of pulsed laser to measure distances to the Earth[1]. This allows the exploration of natural and manmade environments with accuracy and flexibility [1]. It can collect large scale data at a low cost which could be used for the analysis of urban inequality[2].

Deprivation refers to the lack of material benefits considered as basic necessities. In this coursework, deep learning technology is used to develop a model that could be used to predict the deprivation index linked to crime using the LIDAR images, to understand the effectiveness of LIDAR in analyzing the multidimensional deprivation.

VGG and a RESNET based models were created which gave an MSE of 0.3851 and 0.3767 respectively.

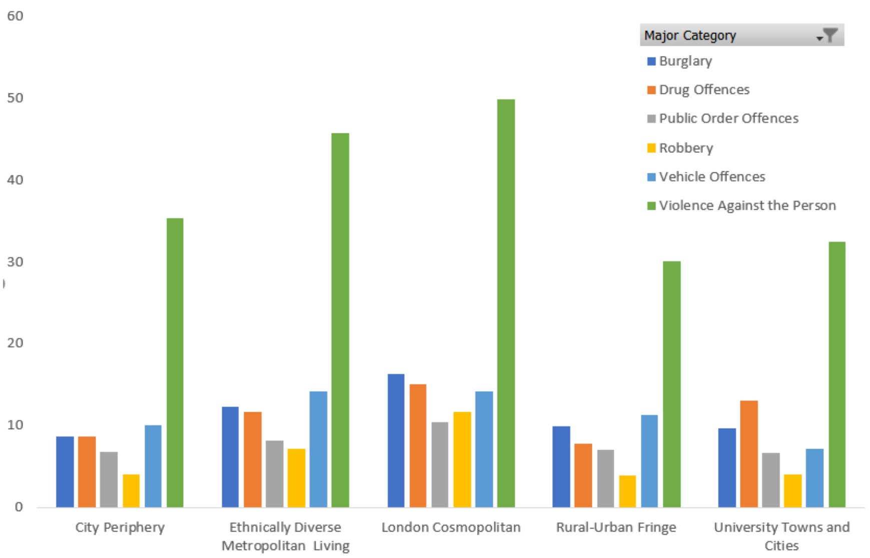
2. Effectiveness of LIDAR in Crime Index Prediction

Crime Prediction has been on highlight in the first two decades of the 21st Century. Crime prediction models can be developed through both qualitative and quantitative approaches. Qualitative approaches include environmental scanning and Delphi method which is useful in understanding the nature of the criminal activities. Quantitative methods focus on prediction of the future crime through time series analysis [3].

The coursework focuses the prediction of the deprivation index linked to crime based on the LIDAR images of various locations in London, UK. It is expected that due to the wide ranging nature of criminal activities, the prediction of the crime index based on just the LIDAR images may not be very effective. However, the natural and the manmade environments present in the LIDAR images may be good indicators of certain types of crimes.

It is also validated from the analysis in latter sections that, although the predictive model showcases reasonably good performance, a key aspect that has not been taken into account is the nature of the crime. It is observed that there is not a declining trend in the validation loss although there is a constant decline in the training loss. This is an indication that LIDAR images cannot be effectively used to predict overall crime index.

The nature of the criminal activities is diverse as listed in the London Police website [4].The geographical distribution is varied as well.



It can be noticed that drug offences are high in University towns, and vehicle offences are high in rural areas.

Whereas, burglary is prevalent in cosmopolitan areas

Distribution of the average daily crime activities across various localities
(Data obtained from https://data.london.gov.uk/dataset/recorded_crime_summary)

If the Deprivation index related to crime was split based on the crime categories, we can expect the model to find more relevant features that could be used to predict the crime index from LIDAR images. For instance, (Breetzke, 2012) discusses effect of altitude and slope on patterning of burglary[9].

3. Deep learning based models for Crime Index Prediction

In this section, a deep learning model is developed for crime index prediction based on the LIDAR images.

3.1. Data Overview

Two major sources of data were used for model development:

- 47360 LIDAR images of locations in London, UK.
- Embedding data with geographic information of the and deprivation index values.

- Image ID was used to connect the embeddings data with the LIDAR images.
- There was only 36723 images with embeddings.
- A few set of images were **monochromatic** (only one color) and was not useful in training the model. These have been removed.

```
from IPython.display import Image
Image(filename='LIDAR/LIDAR_72255.png')
```

Example of a monochromatic image

MonochromaticImages
LIDAR/LIDAR_72255.png
LIDAR/LIDAR_32013.png
LIDAR/LIDAR_63873.png
LIDAR/LIDAR_63872.png
LIDAR/LIDAR_51810.png
LIDAR/LIDAR_72012.png
LIDAR/LIDAR_53043.png

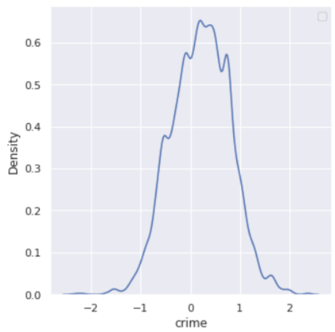
List of Monochromatic Images

- The final set of images available for training and validation was 36716
- The crime index also seemed to be standardized as it is nearly normally distributed.

	crime
count	36723.000000
mean	0.211408
std	0.588681
min	-2.354000
25%	-0.199000
50%	0.221000
75%	0.627000
max	2.377000

```
sns.displot(data = embeddings, x = "crime", kind='kde')
plt.legend(loc='upper right')
```

```
No handles with labels found to put in legend.
<matplotlib.legend.Legend at 0x7ff9f75f3390>
```



Summary Statistics and Distribution of the Index Crime.

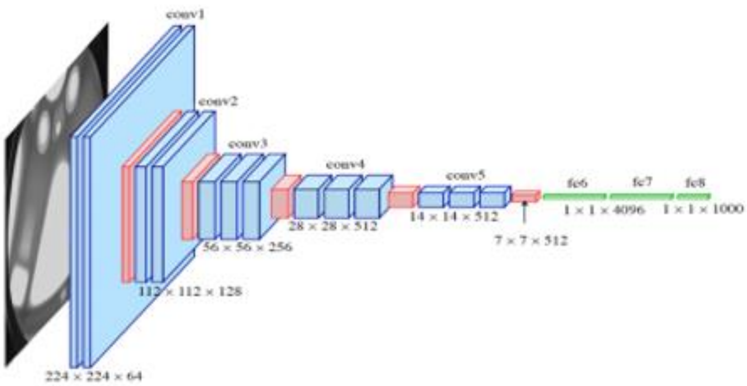
- The data was split into training and testing set (80-20).

```
from sklearn.model_selection import train_test_split

train_df,test_df = train_test_split(embeddings[['image_location','crime']],
                                    test_size=0.2, random_state=251238730)
```

3.2. Creating VGG Model for crime index prediction

- The VGG-16 model was created to solve the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) with images from about 1000 categories [4].



VGG-16 Network Architecture [5]

- Since the VGG-16 has been trained on a large set of images, it is trained to identify a lot of features that helps in classification.
- At the initial layers, the trained parameters identifies simple features like shapes and edges, and latter layers will have complex features such as faces.
- While preserving the weights of the initial layers, the few final layers in the network can be trained for custom applications. This is called Transfer learning.

- **Steps for creating the VGG model for crime index prediction**

1) The weights of the VGG-16 model trained on the imagenet data is loaded from keras.

```
from tensorflow.keras.applications.vgg16 import VGG16, preprocess_input
model = VGG16(weights = 'imagenet',      # The weights from the ImageNet competition
              include_top = False,       # Do not include the top layer, which classifies.
              input_shape= (224, 224, 3) # Input shape. Three channels, and BGR (NOT RGB!!!)
             )
```

2) Create a new model and copy the weights. Only the last two convolutional layers are set as trainable .

3) Add a fully connected block with three dense layers having a dropout of 0.6. The dropout randomly delete connections to some neurons and reduces overfitting.
4) For all layers except the last layer set the activation function as ReLU. Using ReLU ensures that gradients are large enough to avoid vanishing gradients problem.
5) Since crime index is continuous and have negative values, the activation function in the last layer is linear.

```
# Set layer as trainable.
# CBModel.layers[14].trainable = True
CBModel.layers[15].trainable = True
CBModel.layers[16].trainable = True
```

```
# We now add the new layers for prediction - since the index can be negative we use a linear activation
CBModel.add(Flatten(input_shape=model.output_shape[1:]))
CBModel.add(Dense(128, activation = 'relu'))
CBModel.add(Dropout(0.6))
CBModel.add(Dense(128, activation = 'relu'))
CBModel.add(Dropout(0.6))
CBModel.add(Dense(64, activation = 'relu'))
CBModel.add(Dropout(0.6))
CBModel.add(Dense(1, activation = 'linear'))
```

6) Compile the model.
    a. The loss function is set as Mean Squared Error as it is regression.
    b. The optimizer used is RMSProp as it performs well in regression and time series models.
    c. The learning rate is set at 1e-6 as high learning rate would not allow convergence at the global optima.

7) ImageDataGenerator module is used to pass images as batches for training. Here, data augmentation techniques such as zoom and shear is set at 0.2 and flipping of images is allowed. This will help in building additional robustness.

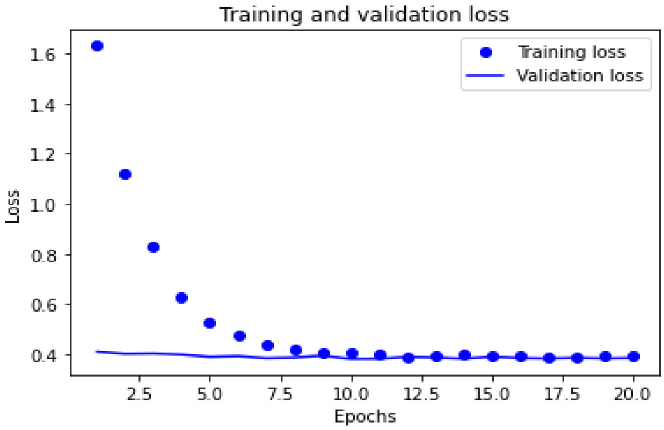
```
train_datagen = ImageDataGenerator(
                                    rescale=1./255,
                                    shear_range=0.2,
                                    zoom_range=0.2,
                                    horizontal_flip=True,
                                    vertical_flip=True,
                                    preprocessing_function=preprocess_input,
                                    validation_split = 0.2
                                    )
```

8) In the data loader for test set, no augmentation is applied.
9) Model is trained for 20 epochs using the batch size for training as 64 and validation as 32.

```
epochs = 20

# Train!
CBModel.fit(
            train_generator,
            epochs=epochs,
            validation_data=validation_generator,
            steps_per_epoch = 64, # Usually cases / batch_size = 176. Reduced to 32 so it runs faster.
            validation_steps = 32 # Number of validation steps. Again cases / batch_size = 44.
          )
```

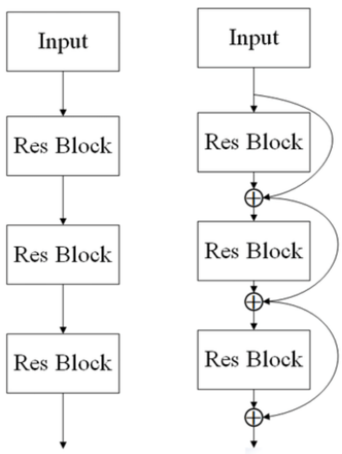
- After training the model, it was observed that the training loss is decreasing till epoch 10 and then it converges. However, there is not much variation in the validation loss.



It can be inferred that just using LIDAR for crime prediction might not be very effective.

3.3. <u>Creating Resnet Model for crime index prediction</u>

- Resnet50 model (Residual Networks) is also convolutional network which won the Imagenet challenge in 2015[6].
- It introduced the concept of skip connection. In the skip connection, along with the output of the current convolution, output of the previous layers are added as well.



By including skip connections the features learned by the model in previous layers are transferred forward using skip connections.

Resnet skip connections [7]

- This has helped in improving the performance of the model by a great margin.

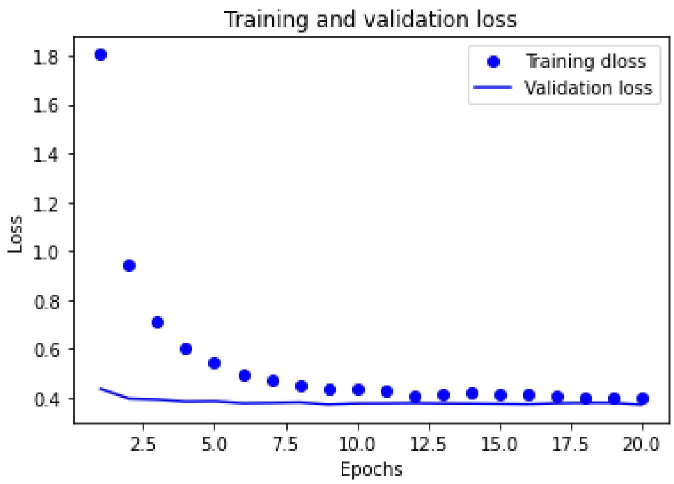
- **Steps for training the Resnet Model for crime index prediction**

  1) The weights of the Resnet trained on the Imagenet data is loaded using keras

```
# Import model with input layer
base_model = ResNet50V2(weights = 'imagenet',
                       include_top = False,
                       input_shape= (224, 224, 3)
                       )
```

2) The input layer for the resnet model is designed and the fully connected layers are added at the end. In the final layer we use the linear activation function like in VGG.
3) The learning rate for the model is set at 1e-6 and mean squared error loss is used while training the model.
4) The data generation is similar to that of the VGG model. However while training the resnet model, we initially warmup the model learning rate using a small value for a few epochs before starting the training.
5) This avoids the model from not converging as the learning rate is low initially when the weights are randomly initialized.
6) During the model training phase we specify callbacks which helps to save the model when the training loss converges.

- After training the model we observe the training and the validation loss. The final validation loss for the model is 0.3767



It can be observed that while the training loss is constantly reducing, there is a small reduction in the validation loss in the first few epochs. It can be inferred that just using LIDAR for crime prediction might not be very effective.

- The MSE value on the validation set for the VGG16 is 0.3851 whereas the MSE value for the Resnet50 is 0.3767.
- There is only minor variation in the MSE, hence we would prefer using the VGG model for the analysis as it is sequential and less complex.

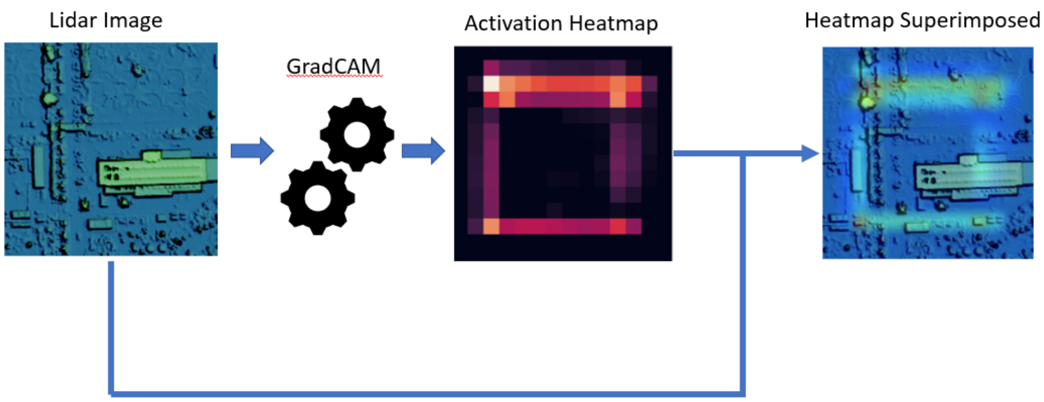
3.4. Visualization of the network using GradCAM

- Gradient-weighted Class Activation Mapping (Grad-CAM), uses the gradients flowing into the final convolutional layer to highlight important regions in the image[8] .
- We use the Grad-CAM technique to check the regions of the LIDAR image that the model considered for prediction.

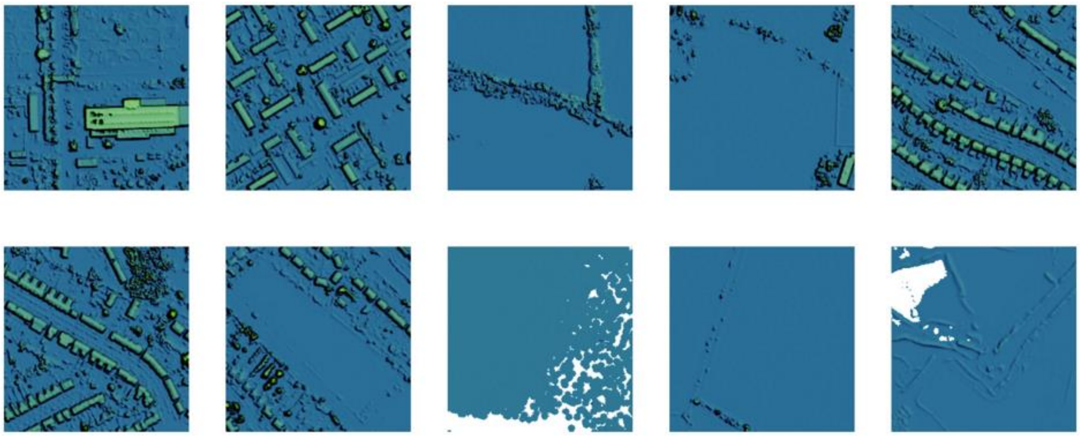
- This technique improves explainability and can act as signal for further training, network redesign or hyperparameter tuning if the model is not learning the right insights.

Steps to perform Grad-CAM

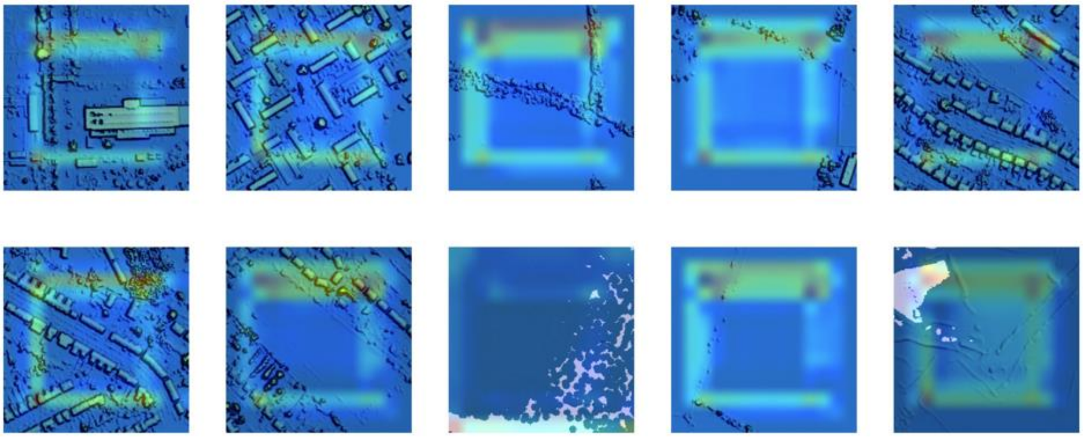
1) Create a model that maps the input image to the activations of the last conv layer
2) Create a model that maps the activations of the last convolutional layer to the final class predictions
3) Compute the gradient of the top predicted class for our input image with respect to the activations of the last conv layer
4) Multiply each channel in the feature map by gradients to get the heatmap.
5) The heat map generated is superimposed on the image to visualize the areas that are getting activated for model prediction.



Images



Grad-CAM Activations



- From the above images, it can be observed that the top and the bottom edges are getting activated which means that the model uses features in these parts to predict the crime index.
- It can be noticed that the model is looking for areas with cluster of small buildings or barren lands without much population as areas prone to crime.
- However, the learning from the model is limited due to the points discussed in Section 2. Training using additional features such as nature of crime may help in improving the model performance.

4. Using LIDAR for multidimensional deprivation index

The urban areas is ever expanding and absorbing most of the population growth. As a result, there is challenges related to access to education, healthcare etc. which results in increasing inequality. Reducing these inequalities is key to ensure sustainable development[2]. Deep learning can be used effectively to predict the deprivation indexes of an area using the satellite images such as LIDAR. This can help in identifying specific areas of concern and targeting policies for development.

Since LIDAR can clearly give the topography of the area along with the manmade and natural environments existing in the area, it can be very useful indicators of some deprivation indexes such as health, income and living environment. (Suel et al., 2021)[2] has a multimodal approach for the prediction of the quality of the living environment. (Andersson et al., 2019)[10] discusses the use of street level imagery for the predicting incidence of Dengue cases. But for prediction of certain indices additional information is requires.

Due to the lack of the well-formed regulations in the remote sensing technology space, the use of LIDAR presents a few ethical and privacy issues. Although the LIDAR images are quite low resolution, when combined with additional sources of information could cause national security concerns. Additionally, since the technology spans across the international arena where certain countries may have lenient rules related to remote sensing, there should be an international policy[11].

The use of LIDAR also has some impact on the personal privacy concerns. Remote sensing technology at a metre level resolution that relate to the physical space of an individual and around his home poses a privacy issue. Specifically for LIDAR, when the remote sensing data is analyzed with data from multiple sources such as social media data has the potential to deliver personal information like relationships, property and time utilization [12]. With the advances in the image processing technologies, criminals could also get a near real time image of an individual’s home to determine occupancy, visibility, value of property etc. which may lead to organized burglary[12].

Standardization and harmonization of data can help to reduce the liability fear. It can help to categorize the data based on the need which enables scientific research and enhances data security by access control. Global organizations should work together to develop international standards and regulations for use and sharing of remote sensing data[11]. It is highly essential to enforce rules and regulations in this fast growing space as it as wide ranging applications.

Based on the analysis, it is evident that LIDAR data can be very effective in the analysis of deprivation and can be used to target areas of concern. However since this technology is growing at a fast pace it is essential to focus on creation of standards and protocols for the use of LIDAR, so that it can be used for research safely without any ethical or privacy concerns.

5. Appendix

- The link to the colab Notebook for the assignment is: https://colab.research.google.com/drive/1iXUaC40tRjwJP1B01fulgtzhPtEN8ppc?usp=sharing

- The link to the Excel File for crime analysis is: https://docs.google.com/spreadsheets/d/1YZ9ZCzGlvLVSEA6hNv9je5rM6xwMw2HG/edit?usp=sharing&ouid=102855196681769508439&rtpof=true&sd=true

- Word Count: 2197

6. References

[1] US Department of Commerce, N. O. & A. A. (2012, October 1). *What is Lidar*. NOAA's National Ocean Service. Retrieved December 15, 2021, from https://oceanservice.noaa.gov/facts/lidar.html

[2] Suel, E., Bhatt, S., Brauer, M., Flaxman, S., & Ezzati, M. (2021). Multimodal deep learning from satellite and street-level imagery for measuring income, overcrowding, and environmental deprivation in urban areas. Remote Sensing of Environment, 257, 112339. https://doi.org/10.1016/j.rse.2021.112339

[3] Department of Justice Canada, Schneider, S. (2002). PREDICTING CRIME: A REVIEW OF THE RESEARCH (No. rr2002-7e). Research and Statistics Division Policy, Integration and Coordination Section. https://www.justice.gc.ca/eng/rp-pr/csj-sjc/jsp-sjp/rr02_7/rr02_7.pdf

[4] Hassan, M. U. (2021, February 24). *VGG16 – Convolutional Network for Classification and Detection*. Neurohive. https://neurohive.io/en/popular-networks/vgg16/

[5] Ferguson, Max & ak, Ronay & Lee, Yung-Tsun & Law, Kincho. (2017). Automatic localization of casting defects with convolutional neural networks. 1726-1735. 10.1109/BigData.2017.8258115

[6] Dwivedi, P. (2021, December 7). Understanding and Coding a ResNet in Keras - Towards Data Science. Medium. https://towardsdatascience.com/understanding-and-coding-a-resnet-in-keras-446d7ff84d33

[7] Wang, Z., Yi, P., Jiang, K., Jiang, J., Han, Z., Lu, T., & Ma, J. (2019). Multi-Memory Convolutional Neural Network for Video Super-Resolution. IEEE Transactions on Image Processing, 28(5), 2530–2544. https://doi.org/10.1109/tip.2018.2887017

[8] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2019). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. International Journal of Computer Vision, 128(2), 336–359. https://doi.org/10.1007/s11263-019-01228-7

[9] Breetzke, G. D. (2012). The effect of altitude and slope on the spatial patterning of burglary. Applied Geography, 34, 66–75. https://doi.org/10.1016/j.apgeog.2011.10.017

[10] Andersson, V. O., Cechinel, C., & Araujo, R. M. (2019). Combining street-level and aerial images for dengue incidence rate estimation. 2019 International Joint Conference on Neural Networks (IJCNN). https://doi.org/10.1109/ijcnn.2019.8851749

[11] Schoenmaker, A. (2011, April). Community remote sensing legal issues. Retrieved December 17, 2021, from https://swfound.org/media/62081/schoenmaker_paper_community_remote_sensing_legal_issues_final.pdf

[12] Slonecker, T., Shaw, D., & Lillesand , T. (1998, June). Emerging legal and ethical issues in advanced remote ... Emerging Legal and Ethical Issues in Advanced Remote Sensing Technology. Retrieved December 17, 2021, from https://ibis.geog.ubc.ca/courses/geog373/lectures/Handouts/PERS_Remote_Sensing_Ethics.pdf