

Using Satellite Imagery and Convolutional Neural Networks for Building Identification in Points of Interest from GPS Data

Since the year 2000, GPS technology has seen rapid advancements in computer software applications. GPS receivers are now extensively used to track mobile entities equipped with GPS antennas or in-vehicle systems, including both pedestrians and vehicles. It is crucial to recognize that the performance of GPS receivers varies significantly, and the software used for GPS monitoring possesses inherent limitations. For example, satellite signals can be reflected or diffracted by urban infrastructure such as buildings, leading to signal delays that disrupt the accuracy of real-time data transmission. These issues can cause errors such as the disappearance of tracked entities or disordered movement trajectories. Hashemi (2017) highlighted a scenario where GPS signal drift led to the detection of implausibly short travel distances when identifying user points of interest—locations where individuals live, work, or spend significant time. Even on a smaller scale, challenges persist; for instance, building structures can obstruct GPS signals when trying to determine a person's home address, misplacing coordinates to nearby roads rather than the actual intended locations. This introductory discussion underscores the need for enhanced methods to address the inaccuracies introduced by urban settings in GPS data.

However, satellite imagery offers a solution to the challenges of poor indoor signal reception by enabling the direct extraction of geographic images to distinguish between buildings and non-building areas. Alsabhan, Alotaiby, and Dudin (2022) have effectively employed the U-Net semantic segmentation model to identify buildings from satellite imagery, facilitating the accurate delineation of building polygons in a single image. This approach primarily captures underlying geographic information rather than merely classifying images as building or not. Building on this, the current paper proposes a novel research direction that uses satellite imagery for binary classification of points of interest identified from GPS data, determining if they are actual buildings or mistakenly identified locations appearing on streets. The primary aim of this study is to correct indoor positioning biases evident in GPS data. The foundational data for this study is sourced from a publicly available Amazon dataset, which documents the GPS usage of over 30,000 users in the Elephant and Castle area in South London, as depicted in Figure 1. This area shows heightened GPS activity near the Thames River in January 2018. Furthermore, this research utilizes recent satellite imagery updates from Google Maps, at a resolution of 640x640 pixels. The computational model employed is the VGG-16 architecture, a robust convolutional neural network pre-trained on the ImageNet dataset.

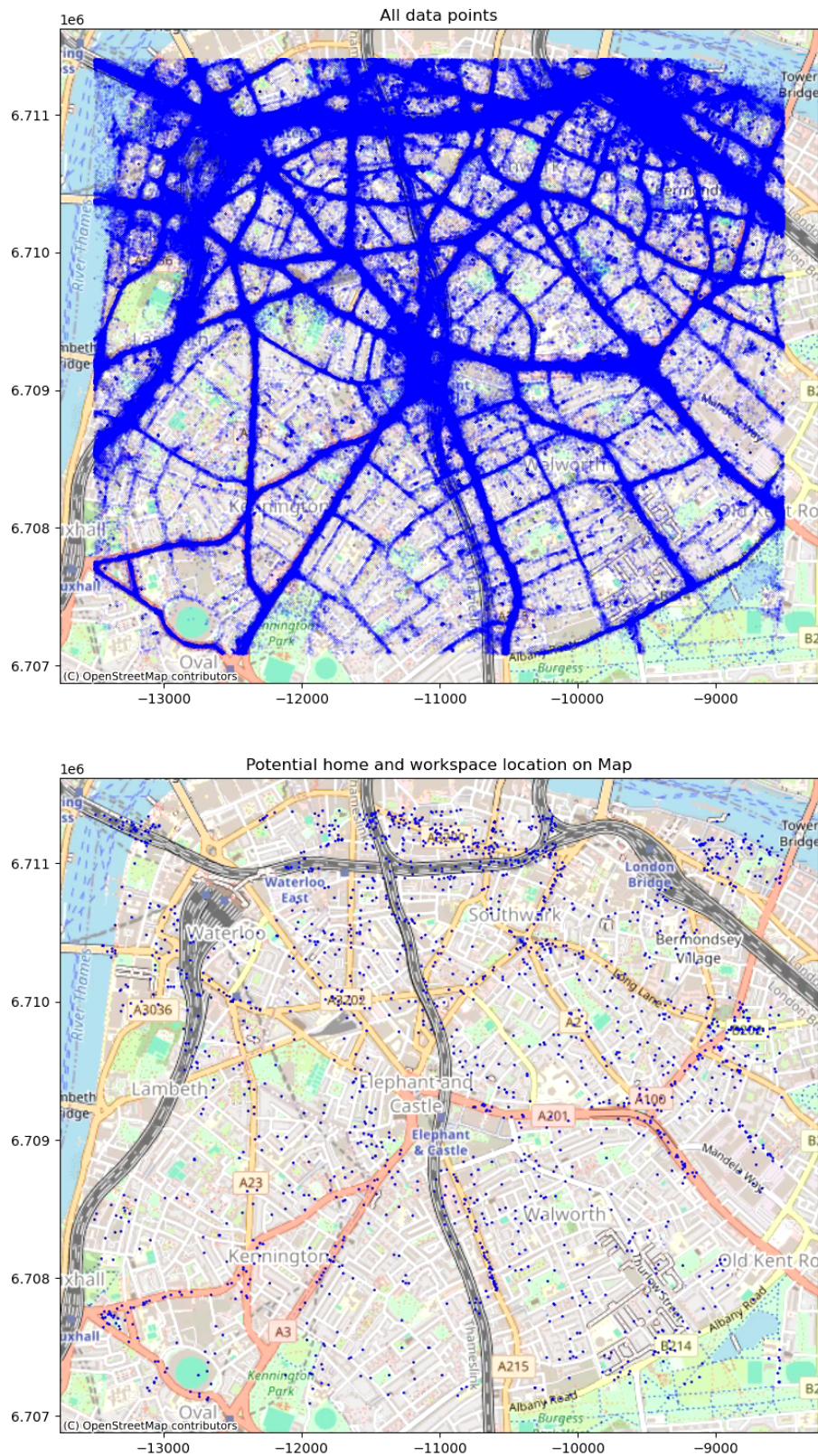


Figure 1

In processing the extensive GPS data, I applied the stop point detection method proposed by Li et al. (2008), which uses fixed time intervals and distances to identify all potential stopping points. This method was selected for its effectiveness in distinguishing brief stops from continuous travel. By setting specific thresholds for time and distance, I conducted an exhaustive search to pinpoint these locations. Subsequent filtering by time and grouping by location helped to refine and average the preliminary data. Given the constraints imposed by map boundaries, many points were clustered around perimeter areas and near stations. To address this, I implemented buffer zones and a density-based clustering algorithm, enhancing the precision of the categorized data. The refined set of points, totaling 1,848, are depicted in Figure 1 (Lower).

These identified points likely represent the home and workplace locations of the dataset's users. In the absence of signal interference, these should correspond to buildings. I merged these points with the road network and geographic building data from OpenStreetMap for an initial categorization. Using the Google Maps API, I then extracted satellite images for each point. To ensure the accuracy of the data used in modeling, I performed a manual secondary classification, resulting in two image categories: buildings and streets. I also discarded any irrelevant or unclear images to maintain model efficiency. Figure 2 illustrates examples of this classification, showing the key characteristics of the two image types. The final dataset comprised 504 images labeled as 'buildings' and 564 as 'streets.' The slight numerical disparity between these categories is unlikely to introduce significant bias into the model, according to Model and Shamir (2015).

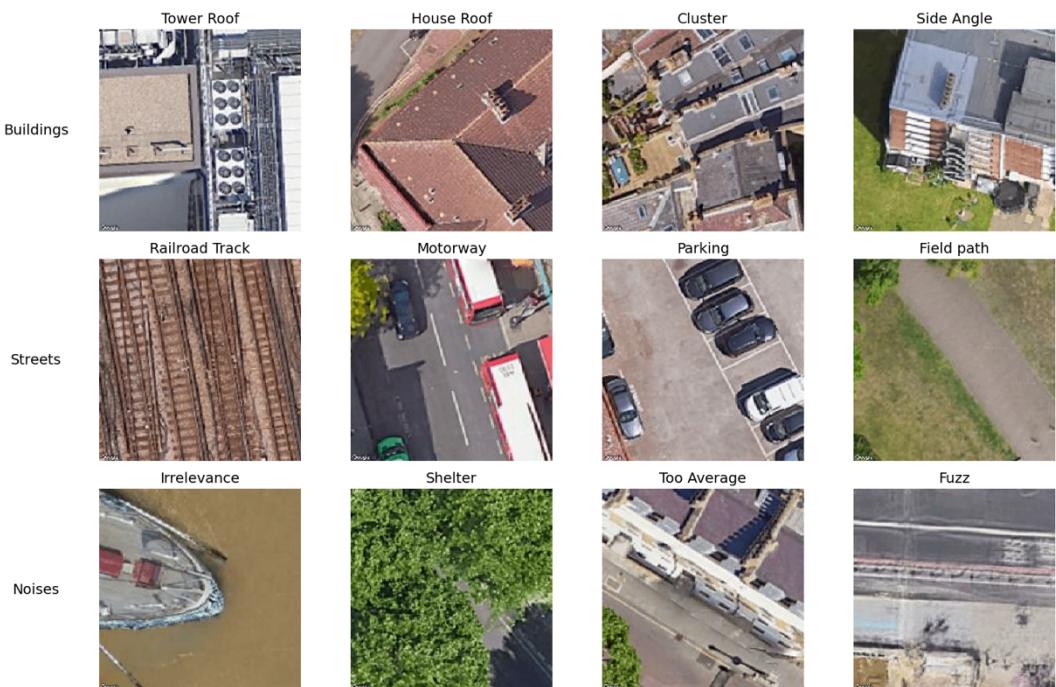


Figure 2

The research pipeline, outlined in Figure 3, begins with data processing followed by the construction of the machine learning model. I employed the VGG-16 architecture as the base model for feature extraction due to its proven effectiveness in similar applications, as evidenced by Tammina (2019) and Ye et al. (2021), who achieved notable performance in their experiments. Additionally, Li and Wang (2021) successfully used VGG-16 to analyze post-hurricane building images, underscoring its utility in satellite remote sensing.

Before model training, I conducted a series of preprocessing steps to standardize pixel values and augment the dataset, thereby helping to prevent overfitting. These steps included rescaling pixel values, applying random rotations, adjusting image dimensions, horizontal flipping, and zooming (Brownlee, 2019). I then divided the dataset into an 80% training set and a 20% validation set, resizing the images to 224x224 pixels to meet the input specifications of the VGG-16 model. I set a conservative learning rate and chose binary cross-entropy as the loss function to optimize for classification accuracy. The model was configured to train over 30 epochs with a batch size of 32, incorporating an early stopping callback to cease training if the validation accuracy failed to improve for five consecutive epochs.

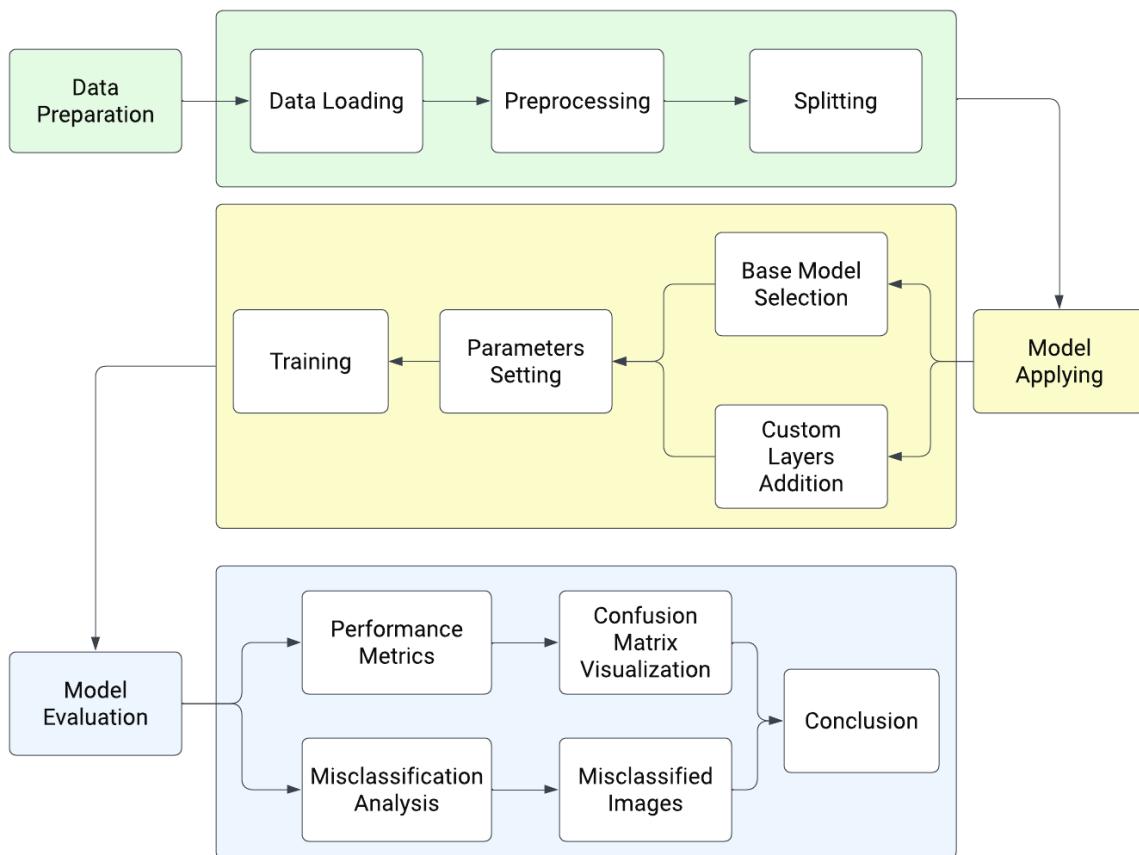


Figure 3

After completing the iterations, I generated line plots for the training and validation losses and accuracies, as shown in Figure 4. From these plots, it can be observed that with increasing epochs, both the training and validation losses exhibit a decreasing trend, while their accuracies correspondingly increase. This is a positive indication, suggesting that the model is effective for both the training and validation sets, and there is no evidence of overfitting apparent from the plots.

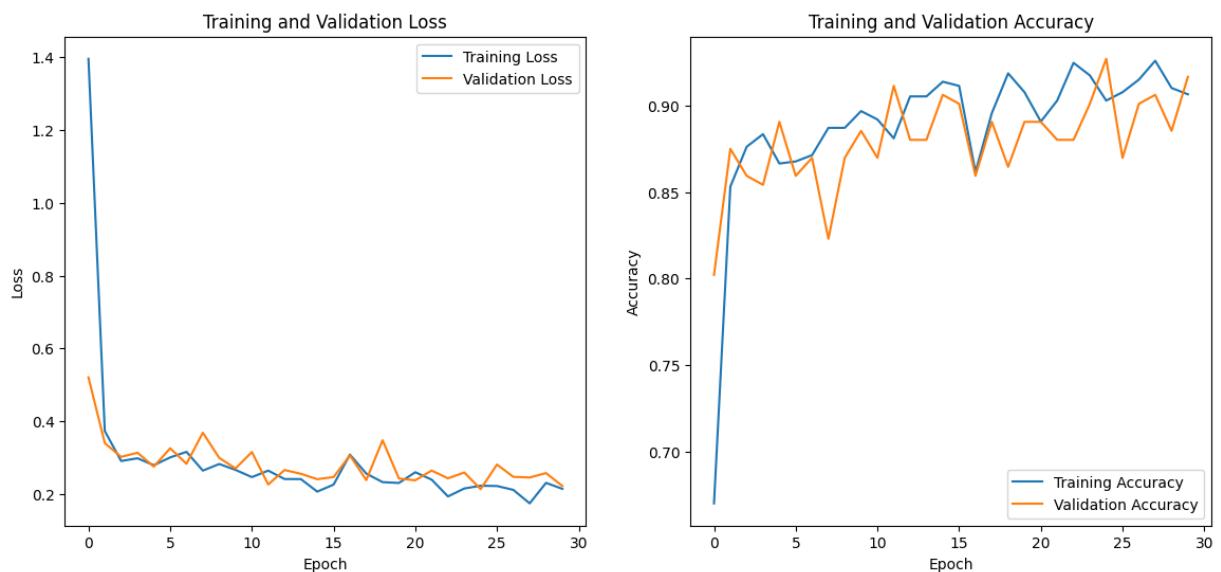


Figure 4

Next, from Figure 5, the confusion matrix, it can be observed that the model's recognition of streets (TP) is slightly better than that of buildings (TN), indicating a potentially stronger capability of capturing street features by the model. To delve deeper into the underlying reasons, I generated Figure 6, the Misclassification table. This figure displays examples of False Negatives (FN) and False Positives (FP). For FN errors, one possible reason could be due to the large scaling of the images, as in example 2, where only the rooftops are visible in the entire image, which may be confused with street images. Similarly, for FP errors, all examples are combinations of street and building images. Therefore, these errors in identification can also be considered as having plausible reasons.

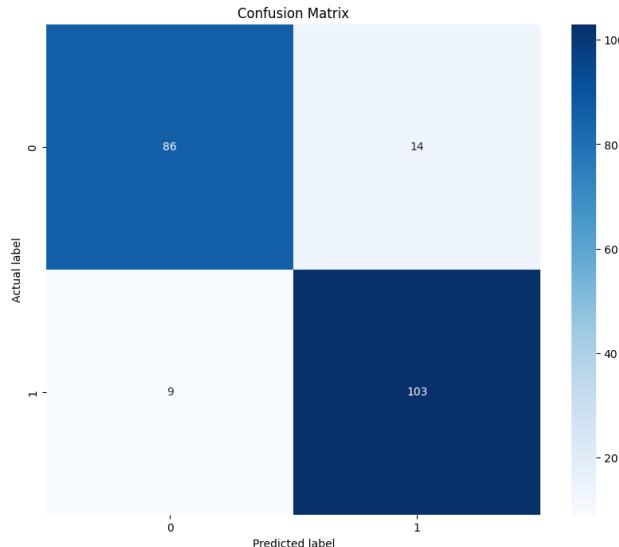


Figure 5

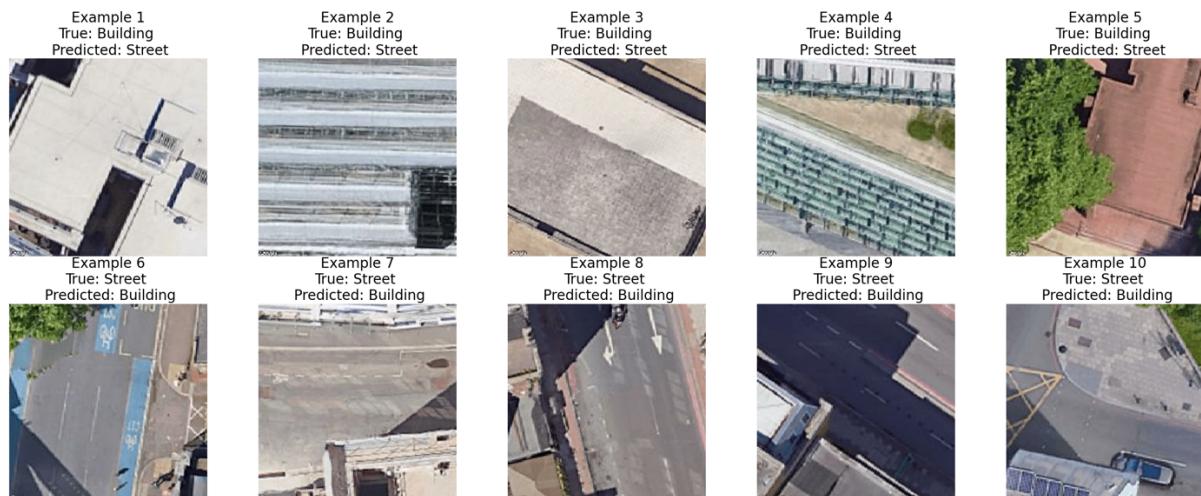


Figure 5

	precision	recall	f1-score	support
0	0.93	0.82	0.87	100
1	0.85	0.95	0.90	112
accuracy			0.89	212
macro avg	0.89	0.88	0.89	212
weighted avg	0.89	0.89	0.89	212

Table 1

Finally, I produced my performance analysis. The high accuracy rate indicates that the model has excellent overall recognition capabilities for these two classes of images. Additionally, for both the building and street categories, the precision and recall are high, indicating that the model can not only correctly identify most positive samples (high recall) but also is usually correct when it claims an image belongs to a certain class (high precision). The F1 score is 0.89, which considers both the classifier's precision and robustness. My model has high F1 scores for both classes, indicating a good balance between accuracy and consistency.

Overall, this experiment was successful, especially considering the model's good performance with only 1000 freely available data samples. The application of the model can assist both governmental and commercial sectors in better researching and determining location data, which is significant for policy-making and business decisions. However, it is important to acknowledge the limitations of this experiment. Firstly, the experiment only serves as a binary study, classifying images into two categories: buildings and streets. However, the real situation is more complex, as there may be more data not belonging to these two categories, similar to the noise I excluded. A more accurate and powerful model should consider more critical factors. Secondly, another issue to consider is the temporal validity of the data. The GPS data is from 2018, while Google Maps street view images have been updated in the past two years. Therefore, there may be changes in the geographical locations, building facilities, or road infrastructure at different time points, leading to errors.

Reference

Alsabhan, W., Alotaiby, T. and Dudin, B. (2022) 'Detecting Buildings and Nonbuildings from Satellite Images Using U-Net', *Computational Intelligence and Neuroscience*. Edited by A. Bhardwaj, 2022, pp. 1–13. Available at: <https://doi.org/10.1155/2022/4831223>.

Brownlee J. (2019) 'How to Configure Image Data Augmentation in Keras', *MachineLearningMastery.com*, 11 April. Available at: <https://machinelearningmastery.com/how-to-configure-image-data-augmentation-when-training-deep-learning-neural-networks/> (Accessed: 22 April 2024).

Hashemi, M. (2017) 'Intelligent GPS trace management for human mobility pattern detection', *Cogent Engineering*. Edited by M. Chadli, 4(1), p. 1390813. Available at: <https://doi.org/10.1080/23311916.2017.1390813>.

Huang, G., Yang, B. and Jiang, K. (2011) 'New Algorithms on the Solution to Drifting Problem of GPS Positioning', *Journal of Computers*, 6(1), pp. 83–89. Available at: <https://doi.org/10.4304/jcp.6.1.83-89>.

Kjærgaard, M.B. et al. (2010) 'Indoor Positioning Using GPS Revisited', in P. Floréen, A. Krüger, and M. Spasojevic (eds) *Pervasive Computing*. Berlin, Heidelberg: Springer, pp. 38–56. Available at: https://doi.org/10.1007/978-3-642-12654-3_3.

Li, H. and Wang, X. (2021) 'Robustness Analysis for VGG-16 Model in Image Classification of Post-Hurricane Buildings', in *2021 2nd International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE)*. *2021 2nd International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE)*, pp. 401–407. Available at: <https://doi.org/10.1109/ICBASE53849.2021.00081>.

Li, Q. et al. (2008) 'Mining user similarity based on location history', in *Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems*. New York, NY, USA: Association for Computing Machinery (GIS '08), pp. 1–10. Available at: <https://doi.org/10.1145/1463434.1463477>.

Percival, D.B. and Walden, A.T. (2000) *Wavelet Methods for Time Series Analysis*. Cambridge University Press.

Tammina, S. (2019) *Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images*, *International Journal of Scientific and Research Publications (IJSRP)*. Available at: <https://doi.org/10.29322/IJSRP.9.10.2019.p9420>.

Verma, R. et al. (2024) 'Comparison of home detection algorithms using smartphone GPS data', *EPJ Data Science*, 13(1), pp. 1–22. Available at: <https://doi.org/10.1140/epjds/s13688-023-00447-w>.

Yang, D.K. et al. (2006) 'Design and Realization of Delay Mapping Receiver Based on GPS for Sea Surface Wind Measurement', in *2006 1ST IEEE Conference on Industrial Electronics and Applications. 2006 1ST IEEE Conference on Industrial Electronics and Applications*, pp. 1–4. Available at: <https://doi.org/10.1109/ICIEA.2006.257106>.

Ye, M. et al. (2021) 'A Lightweight Model of VGG-16 for Remote Sensing Image Classification', *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, pp. 6916–6922. Available at: <https://doi.org/10.1109/JSTARS.2021.3090085>.