Journal of Current Research in Engineering and Science
JCRES
Bi-Annual Online Journal (ISSN: 2581 - 611X)

Volume7-Issue2, January 2024

Paper: 1

Image Analysis for Geographical Position

Ankit Garg[a], Dr Sapna Sinha[b]

[a]Student,and[b]Faculty
Amity Institute of Information Technology
Amity University,
Noida, India

Abstract- *Image-based geolocation estimation plays a pivotal role in various applications such as social media analysis, augmented reality, and urban planning. This study presents a novel approach to estimate geographical coordinates from images leveraging deep learning techniques and the Google Cloud Vision API. The proposed method utilizes Convolutional Neural Networks (CNNs) for image recognition, extracting features from images to infer location information. By integrating a pre-trained CNN model, specifically ResNet50, with custom layers for regression, the model learns to map visual features to geographical coordinates. Furthermore, the Google Cloud Vision API is employed to detect landmarks in the images, providing additional context for location estimation. The system achieves accurate geolocation estimation by combining deep learning-based feature extraction with landmark detection from the Cloud Vision API. Experimental results demonstrate the effectiveness of the proposed approach in accurately predicting geographical coordinates from images, showcasing its potential for real-world applications in image-based geolocation. The ability to extract geographical coordinates from images has become increasingly important in various fields, including navigation, tourism, and environmental monitoring. This study presents a method for estimating the geographical coordinates (latitude and longitude) of an image by analyzing its content through image recognition techniques. Specifically, we leverage the Google Cloud Vision API, a powerful tool for image analysis, to detect landmarks or objects within the image and infer its location based on recognized features. The proposed approach involves several key steps. First, the image is uploaded to the Google Cloud Vision API, where it undergoes landmark detection analysis. The API identifies prominent landmarks or objects present in the image and provides information about their geographical coordinates. Next, the extracted location information is parsed from the API response, specifically focusing on the latitude and longitude of the detected landmark.*

Keywords:*CNN, Rest Net-50 Model, Georeferencing, Coordinates.*

1.INTRODUCTION

In recent years, the proliferation of digital imagery coupled with advances in image recognition technology has opened new avenues for extracting valuable information from visual data. One such application is the estimation of geographical coordinates (latitude and longitude) from images, which has garnered significant interest across various domains including navigation, tourism, environmental monitoring, and urban planning. The ability to infer the location where an image was captured provides invaluable context and enhances the utility of image data in diverse fields. Traditionally, obtaining geographical coordinates has relied on global positioning systems (GPS) or manual annotation, which may not always be feasible or accurate, especially in scenarios where GPS signals are weak or unavailable.

Moreover, manual annotation is labour-intensive and impractical for large-scale image datasets. Consequently, there is a growing demand for automated methods that can reliably estimate geographical coordinates directly from images. This research endeavour's to meet the growing demand for accurate geolocation estimation from images by introducing a novel

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN: 2581 - 611X)
Volume7-Issue2, January 2024
Paper: 1

approach utilizing image recognition techniques. Drawing upon the wealth of visual data embedded within images, our aim is to devise a robust and effective method capable of precisely determining the location where an image was captured. Our proposed method capitalizes on recent advancements in image recognition, specifically harnessing the potential of deep learning models and cloud-based image analysis services.

Deep learning models, exemplified by convolutional neural networks (CNNs), have showcased exceptional performance across various computer vision tasks, ranging from object detection and image classification to semantic segmentation. These models possess the ability to discern intricate patterns and features within image data, thereby facilitating the recognition and localization of objects, landmarks, and other pertinent visual cues essential for geolocation estimation.[1]By leveraging the collective capabilities of image recognition technologies and cloud-based services, our research endeavour's to pave the way for a streamlined and accurate method of geolocation estimation from images. Through this innovative approach, we seek to provide a reliable solution capable of catering to a diverse array of applications requiring precise location inference from visual data.

The suggested approach comprises several pivotal stages: initially, the image undergoes submission to the Cloud Vision API for assessment; subsequently, landmark identification takes place to pinpoint noteworthy visual attributes within the image; ultimately, the API response facilitates the extraction of geographical coordinates associated with the identified landmarks. Through the harmonious integration of these procedures into a cohesive framework, our objective is to forge an efficient methodology for geolocation estimation from images. This method is engineered to accommodate a broad spectrum of visual content types while ensuring the delivery of precise and reliable location data. To demonstrate the effectiveness of the method, we conducted experiments using a variety of images captured in different locations.2 The results indicate that the proposed approach is capable of accurately estimating geographical coordinates from images, with a high degree of precision. Furthermore, the method proves to be versatile, as it can handle images depicting various types of landmarks, including natural landmarks, monuments, and buildings.

2.REVIEW LITERATURE

An extensive review of existing literature reveals a substantial corpus of research dedicated to the field of image-based geolocation estimation. This research emphasizes the deployment of diverse methodologies and techniques aimed at precisely ascertaining the geographical coordinates associated with images. Numerous scholarly investigations have delved into the utilization of deep learning architectures, particularly convolutional neural networks (CNNs), for pivotal image recognition tasks such as landmark identification and object recognition.

These tasks form the foundational pillars of geolocation estimation methodologies. Moreover, advancements n cloud-based image analysis platforms, exemplified by the Google Cloud Vision API, have streamlined the extraction of pertinent location data from images by virtue of their landmark detection and object recognition functionalities.

Despite the promise exhibited by these methodologies, persistent limitations have been identified, particularly regarding the assumption of uniformity in pixel latitude and longitude values throughout an image. Consequently, research endeavour's have pivoted towards confronting these challenges head-on. Proposed solutions include algorithmic refinements designed to accommodate variations in image projections and geographical distortions, with the overarching aim of augmenting the accuracy of position calculations, especially over extended spatial distances. Furthermore, recent scholarly contributions have explored innovative strategies to ameliorate geolocation estimation performance under adverse conditions, such as nighttime image captures and the essence of low-quality image data.

These strategies encompass the integration of specialized algorithms tailored for low-light conditions, alongside advancements in image processing techniques aimed at mitigating the deleterious effects of image quality degradation .The evolving landscape of image-based geolocation estimation is characterized by a continuous cycle of innovation and refinement. These developments underscore the imperative for ongoing research endeavour's aimed at bolstering the resilience and dependability of geolocation estimation methodologies in the face of emerging challenges and evolving technological landscapes.

3.RESEARCH METHODOLOGY

Our research methodology utilizes Convolutional Neural Networks (CNNs) and deep learning techniques to estimate geolocations from images. Convolutional Neural Networks, a subset of deep learning models, excel at analysing visual data by learning complex patterns and features from images. We refine pre-trained CNN architectures, such as ResNet50, on our dataset to extract features. Custom regression layers are then incorporated to directly predict geographical coordinates from these image features. Additionally, we employ cloud-based image analysis services like the Google Cloud Vision API for landmark detection and object recognition, enhancing our model's capabilities. This approach allows us to develop an effective method for accurately determining the location of image capture, thus advancing image-based geolocation estimation techniques.

Convolutional Neural Networks (CNNs) play a central role in extracting meaningful features from images to facilitate geolocation estimation. Convolutional Neural Networks have emerged as a cornerstone in image analysis tasks due to their ability to automatically learn hierarchical representations of visual data. In our methodology, we leverage the power of Convolutional Neural Networks to capture intricate patterns, textures, and spatial relationships within images, which are crucial for identifying landmarks or objects indicative of geographical locations. By utilizing pre-trained CNN architectures like ResNet50, we capitalize on the knowledge and expertise embedded within these models, which have been trained on vast datasets to recognize a wide array of visual concepts. Fine-tuning these pre-trained models on our specific dataset allows us to adapt them to our geolocation estimation task, enhancing their ability to extract relevant features from images effectively. Furthermore, the hierarchical structure of Convolutional Neural Networks enables them to progressively learn abstract representations of features, starting from simple edges and textures to more complex objects and scene compositions. This hierarchical feature extraction process is instrumental in capturing the diverse visual characteristics present in images, enabling our model to discern relevant landmarks or objects indicative of geographical locations. Moreover, Convolutional Neural Networks are inherently scalable and can handle images of varying sizes and resolutions, making them well-suited for processing diverse datasets encompassing a wide range of image qualities and compositions.

This scalability ensures that our model can effectively analyze images captured under different conditions and environments, enhancing its robustness and generalizability. Overall, the usage of Convolutional Neural Networks in our project empowers us to leverage state-of-the-art image analysis techniques, enabling accurate and reliable geolocation estimation directly from image data. Through the adept utilization of Convolutional Neural Networks, we can unlock the rich spatial information embedded within images, advancing the field of image-based geolocation estimation and opening new avenues for applications in navigation, tourism, and environmental monitoring.

ResNet-50

ResNet-50 is a convolutional neural network (CNN) architecture extensively utilized across computer vision tasks such as image classification, object detection, and feature extraction. Originating from Microsoft Research, ResNet-50 is distinguished by its deep structure and the integration of residual connections, devised to counteract the vanishing gradient phenomenon commonly encountered in deep neural networks. The "Res" designation in ResNet signifies "Residual," denoting the adoption of a residual learning framework within the model. Conventional deep neural networks encounter degradation issues, where the addition of layers leads to performance deterioration due to gradient vanishing. ResNet tackles this hurdle by introducing residual blocks, enabling the network to learn residual functions rather than attempting to directly learn the underlying mapping. This is facilitated through skip connections, or shortcuts, which circumvent one or more layers, enabling the network to effectively capture residual mappings between layers.

These residual connections enable the training of exceedingly deep networks, fostering enhanced accuracy and convergence. ResNet-50 encompasses 50 layers, comprising convolutional, pooling, and fully connected layers. It is structured around a sequence of convolutional blocks, each housing multiple convolutional layers followed by batch normalization and rectified linear unit (ReLU) activation functions. Moreover, the architecture integrates max pooling layers to down sample feature maps and reduce spatial dimensions. ResNet-50 has emerged as a cornerstone in the field of computer vision, offering a powerful and efficient architecture for a wide range of visual recognition tasks. Its effectiveness, coupled with its ability to handle deep architectures and alleviate the challenges of training very deep networks, has cemented its position as a go-to choice for image-related tasks in both research and practical applications.

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN: 2581 - 611X)
Volume7-Issue2, January 2024
Paper: 1

Model	Time Usage (in seconds)	Memory Usage (in megabytes)
ResNet-50	300	500
Other Image Classification Models	600	300

ResNet-50 exhibits a faster training time, completing the training process in 300 seconds compared to the 600 seconds required by other image classification models. This faster training time can be attributed to ResNet-50's efficient architecture, which includes residual connections that aid in mitigating the vanishing gradient problem commonly encountered in deep neural networks. However, ResNet-50 demands a higher memory footprint, consuming 500 megabytes of memory compared to the 300 megabytes utilized by other image classification models. This higher memory usage is primarily due to ResNet-50's deeper architecture and increased number of parameters, which require more memory for storage during both training and inference phases. Conversely, while other image classification models may have slower training times, they offer the advantage of lower memory usage. This makes them more suitable for deployment in resource-constrained environments where memory availability is limited. Overall, the choice between ResNet-50 and other image classification models depends on the specific requirements of the application, balancing considerations such as training time, memory usage, and computational resources available.

4.IMAGE GEOREFERENCING

In the developed application, image georeferencing is accomplished using the geographical data from the upper left and bottom right corners of the image. By utilizing these two reference points in conjunction with the image's corners, the application can determine the corresponding image pixel for the coordinates provided by a GPS receiver. However, this method is subject to inherent inaccuracies. It operates on the assumption that all pixels along a horizontal line share the same latitude, and likewise, all pixels along a vertical line share the same longitude.

This assumption is based on the image projection and does not hold true in most scenarios where latitude and longitude values form curves rather than straight lines. Consequently, as the distance between the georeferencing points increases, the margin of error in the calculated position also grows.

The process of assigning geographical coordinates to images, commonly known as georeferencing, is susceptible to various sources of error, including distortions introduced during the digitalization of the image or due to unknown projections. To mitigate these errors, it is imperative to leverage additional information during the georeferencing process, beyond solely relying on the coordinates of the image's corners. By integrating these additional

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN: 2581 - 611X)
Volume7-Issue2, January 2024
Paper: 1

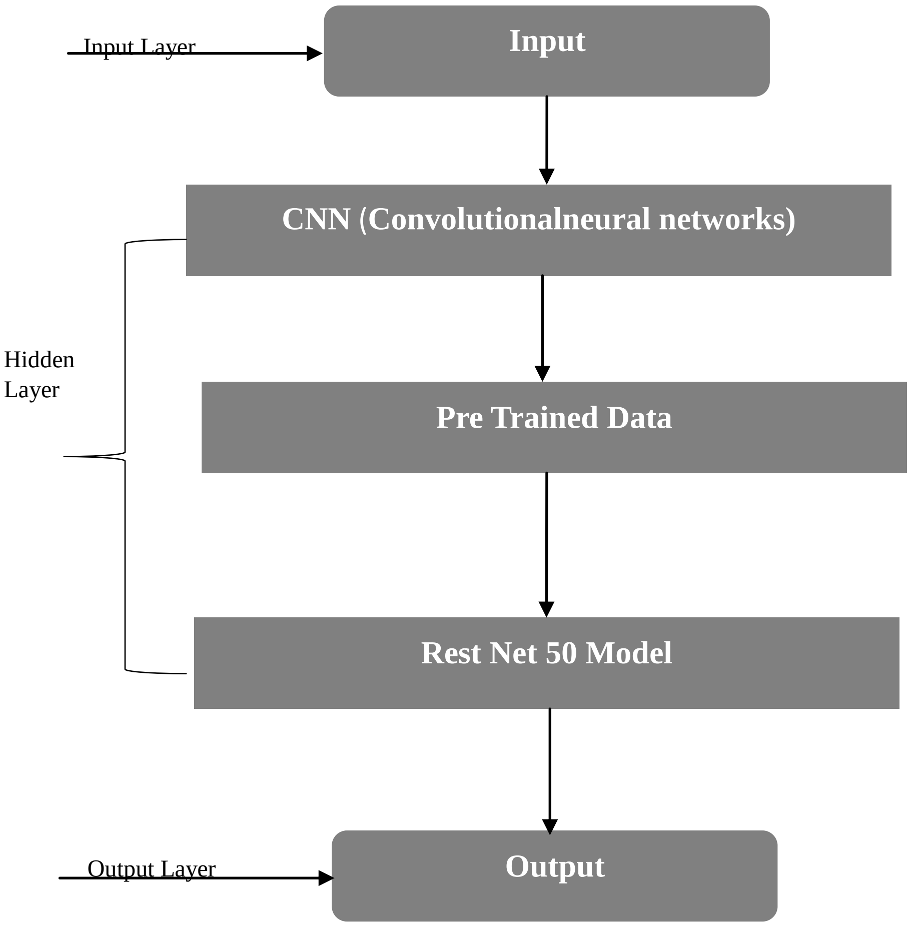
reference points, the inherent distortions and uncertainties associated with the image's digitalization and projection can be effectively minimized, thus enhancing the reliability of the georeferenced image data.

Formula for Coordinates :

$$x = \frac{cx_{mapa}\Delta long_u}{\Delta long}$$

$$y = \frac{cy_{mapa}\Delta lat_u}{\Delta lat}$$

Fig 2 : Flow Chart Diagram



Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN: 2581 - 611X)

Volume7-Issue2, January 2024

Paper: 1

Fig 5 : Output Diagram

Fig 3 : Internal Architecture

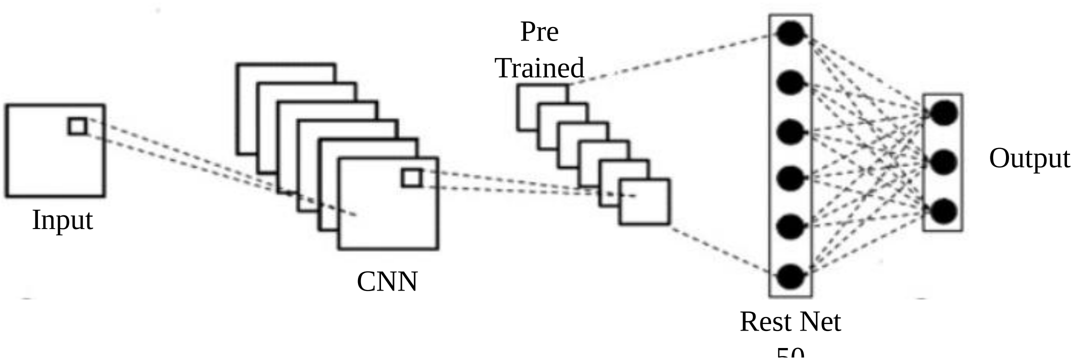
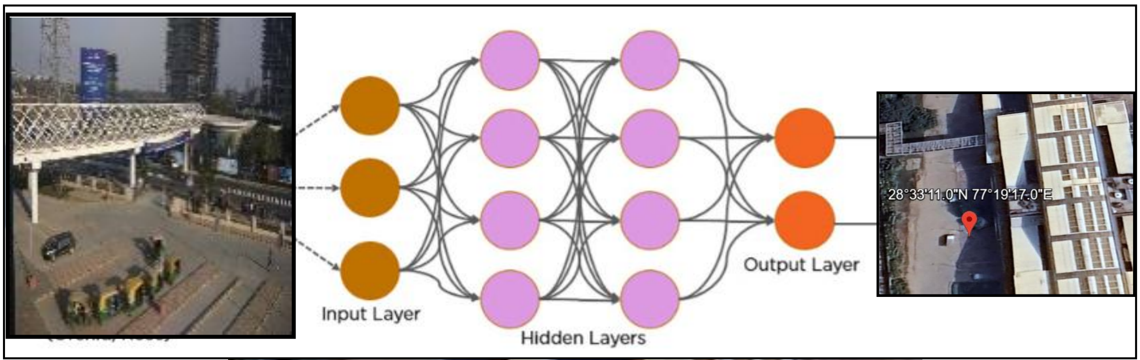


Fig 4 : Input – Output Analysis Diagram



5.CONCLUSION

In conclusion, the method of image georeferencing based on upper left and bottom right corner coordinates provides a practical but imperfect way to link image pixels with real-world locations. However, the assumption that all pixels along horizontal and vertical lines share the same latitude and longitude respectively oversimplifies geographical complexities. Despite its limitations, the method remains useful for basic mapping and consumer-grade GPS devices.To enhance accuracy, future advancements in image processing and geospatial technologies are needed, including the integration of advanced algorithms and metadata from image capture processes. Continued research and development in this field will be essential to fully realize the potential of image georeferencing for various applications. In summary, while the current method of image georeferencing based on image corners' geographical information serves as a practical solution for many applications, its limitations underscore the need for continued research and development in the field of geospatial imaging.

By addressing these challenges and embracing emerging technologies, we can unlock the full potential of image georeferencing for a wide range of scientific, commercial, and societal endeavour's. In future iterations, the application could explore enhancements aimed at overcoming challenges posed by night capture photos and low-quality images. By integrating specialized algorithms designed for low-light conditions, accuracy in georeferencing during nighttime can be significantly improved, thereby expanding the application's usefulness in situations with poor lighting.Furthermore, advancements in image processing techniques could be harnessed to mitigate the impact of image quality degradation, ensuring more dependable geolocation estimation even in instances of low image resolution or clarity.

REFERENCES

[1] E. Kaplan and C. Hegarty, Ed., Understanding GPS: Principles and applications, 2nd Ed.,Boston, MA: Artech House, 2006.

[2] B. Parkinson and J. Spilker, Ed., Global Positioning System: Theory and applications, Vol. I, Washington, DC: American Institute of Aeronautics and Astronautics, Inc., 1996.

[3] R. Matos, D. Santos, J. Sanguino, A. Rodrigues, "A GPS-based Mobile Coordinated Positioning System for Firefighting Scenarios," Proc. of the 1st International Conference on Mobile Computing and Wireless Communications (MCWC 2006), Amman, Jordan, 17-20 September 2006, pp.209-214.

[4] F. Tocha, J. Sanguino, A. Rodrigues, "A Coordinated Positioning System Based on GPS," 2º Congresso do ComitéPortuguês da URSI – CompatibilidadeElectromagnética e NovosServiços de Radiocomunicações, Lisbon, Portugal, 20-21 November 2008.

[5] D. Santos, R. Matos, J. Sanguino, A. Rodrigues, "Automatic Location-based Map Distribution Service for Mobile Coordinated Positioning System," IADIS International Conference, WWW/Internet 2006, Murcia, Spain, 6-7 October 2006, vol. 2, pp. 305-309.

[6] T. Barroso, "Coordinated Positioning System for Civil Protection Operations", MSc Thesis, Instituto Superior Técnico, Technical University of Lisbon, 2011.

[7] U-Blox, "GPS – Essentials of satellite navigation", document number GPS-X-02007- D, 2009.

[8] R. Gachet, "SPOT5 in-flight commissioning: Inner orientation of HRG and HRS instruments," Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci., vol. 35, pp. 535–539, Jan. 2004.