INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
SHORT ABSTRACT OF THESIS

Name of the Student	:	Avinash Kumar Chouhan
Roll Number	:	176101103
Programme of Study	:	Ph.D.

Thesis Title: **DeepGeoVision: Deep Learning Techniques for Various Remote Sensing Vision Tasks**

Name of Thesis Supervisor(s)	:	Prof. Arijit Sur
Thesis Submitted to the Department/ Center	:	Department of Computer Science and Engineering
Date of completion of Thesis Viva-Voce Exam	:	11.12.2024
Keywords for description of Thesis Work	:	Remote Sensing Vision Tasks, Multitask Learning, Semantic Segmentation, Change Detection, Image Super resolution, Unsupervised Domain Adaptation, Nighttime Aerial Tracking, Image Translation

---

SHORT ABSTRACT

The development of intelligent automated vision applications is a need of the hour in remote sensing due to the exponential growth in data generated by aerial and satellite sensors. This thesis proposes eight deep learning models for four important Remote Sensing Vision Tasks: Semantic Segmentation (SS), Change Detection (CD), Image Translation (IT), and Unsupervised Domain Adaptation (UDA). SS can be modeled as an image-to-image mapping where pixel-level classification is required. Pixel-level classification is challenging for high-resolution aerial images due to tiny objects in low frequency, and more information details for such small objects are needed for dense semantic labeling. Two deep models are proposed for effective SS through capturing multi-scale and multi-context features and long-range dependency mapping. CD in satellite images is an important research area as it has a wide range of applications in natural resource monitoring, geo-hazard detections, urban planning, etc. Identifying physical changes on the ground and avoiding spurious changes due to other reasons like co-registration issues, changes in illumination conditions, sun angle, and presence of cloud and fog is a challenging task. We proposed two deep models to resolve issues for the recent existing CD methods related to better feature representation and far pixel relation extraction. IT objective is to convert input to output, which may be from another domain or resolution. Two models are proposed for IT, one for SAR to optical translation and the other for multi-frame super-resolution, which utilizes multi-scale attentive discrimination and handles spatial and temporal pixel dependencies using multiple approaches of global average pooling, multiple size kernels, and self-attention. UDA is an important task that transfers learned knowledge from the source domain to the unseen target domain. Domain shift is the major challenge UDA methods face, caused by differences in appearance, distribution, decision boundary, sensor platforms, capturing conditions, etc. We propose two domain adaptation deep models for domain adaptive segmentation and nighttime aerial tracking using a masked domain dual adaptation approach, joint adversarial alignment, consistency enforcement, feature dissimilarity-based alignment, and reconstruction-based adaptation. Experimentation results reveal that the proposed eight models achieved state-of-the-art results on eighteen benchmark datasets. The effectiveness of the proposed models is validated through extensive quantitative and qualitative analysis.