roshini chelimela

Abstract

**Self-Supervised Pretraining of Transformers for Satellite Image Time Series Classification – 2021**

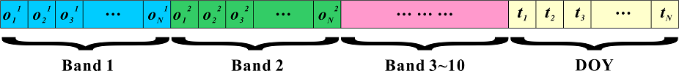
Satellite image time series (SITS) classification is important in remote sensing with applications across various fields. Deep learning methods, though effective, often overfit when labelled data is limited. To address this, a novel self-supervised pre-training scheme for a Transformer-based network is proposed. This approach leverages large-scale unlabelled data by training the model to predict randomly contaminated observations from an entire time series of a pixel. The goal is to learn general-purpose spectral-temporal representations related to land cover semantics using the inherent temporal structure of SITS. After pre-training, the network can be fine-tuned on small-scale labelled datasets for specific tasks, transferring the general knowledge to improve performance and reduce overfitting. Experiments on three benchmark datasets confirm the efficacy of this method.

Crop Classification USING Satellite TIME SERIES DATA

**Data Characteristics:**

To address the issue of insufficient labeled data, self-supervised learning has been proposed. This paradigm trains models on unlabeled data by creating supervised tasks from the data itself. The general pipeline involves:

1. Pretraining: Learning representations by solving a pretext task on large-scale unlabeled data.



1. Fine-tuning: Adapting these representations to a downstream task with limited labeled data.



Ex: [0.3,0.4,0.35,0.32,0.5,0.6,...],[15,30,45,60,75,90,...],[wheat]``

 **Observation Data**: This is the primary data used by your machine learning models. It provides the temporal and spectral information necessary for identifying different crops.

 **DOY Data**: This helps the model understand the time aspect of the observations, which is crucial for analysing seasonal changes in crop growth.

 **Class Label**: When available, it is used for supervised learning to train the model to recognize specific crops. In self-supervised learning, the model might use this data for validation after pretraining.

**Literature Papers Abstract:**

1. **HSI-BERT: Hyperspectral Image Classification Using the Bidirectional Encoder Representation from Transformers:** **2020**

introduces a novel deep learning framework for classifying hyperspectral images. The proposed model, HSI-BERT, leverages BERT's multi head self-attention (MHSA) mechanism to address the limitations of existing CNN-based methods, such as limited receptive fields and inflexibility. HSI-BERT's global receptive field captures pixel dependencies regardless of spatial distance, enabling dynamic and flexible input regions. The model's robust generalization ability allows it to adapt to different region shapes without retraining, making it a powerful tool for accurate pixel-level classification. The paper demonstrates that HSI-BERT outperforms traditional CNN models in both accuracy and computational efficiency on multiple hyperspectral image datasets.

**2.Spatio-Temporal Transfer Learning for Mapping Irrigated Areas using Sentinel-1 - 2020**

The study presents an innovative deep learning approach for mapping irrigated areas using satellite imagery from Sentinel-1, focusing on the challenge of transferring knowledge between different geographical regions. Initially, a convolutional neural network (CNN), termed the "Teacher Model," is trained on a large dataset from Catalonia, Spain, characterized by extensive labelled samples of irrigated and non-irrigated areas. To adapt this model to the West Occitanie region in France, which has a limited number of labelled samples, the researchers employ a "distill before refine" strategy. This involves distilling the teacher model into a more compact "student model" and then refining it with the available data from the target area. The results show that this transfer learning method outperforms other techniques, including random forests and direct application of the CNN, demonstrating its potential for enhancing large-scale, accurate irrigation mapping and supporting better water resource management.

**Model Characteristics:**

An embedding function is a mathematical transformation that maps high-dimensional or categorical data into a lower-dimensional continuous vector space. This function is particularly useful for representing complex data structures in a way that makes them more suitable for machine learning models.

**Attention Mechanism in Deep Learning**

The attention mechanism is a fundamental concept in deep learning, particularly in models dealing with sequential data such as text and time series. It allows the model to focus on different parts of the input data when producing an output.

Attention Formula:

The attention mechanism can be defined as:

Attention (Q, K, V) = SoftMax (QK^T/√dk) V

Explanation of Terms:

Q (Query): The set of queries, representing the current processing element's state that looks for relevant information in the input data.

K (Key): The set of keys, representing all possible elements in the input data.

V (Value): The set of values, representing the actual content that needs to be processed or attended to.

dk: The dimensionality of the keys.

**The key contributions of base paper are threefold:**

1) For the first time, author proposed a self-supervised pretraining

scheme to cope with the problem of insufficient labeled

samples for the SITS analysis.

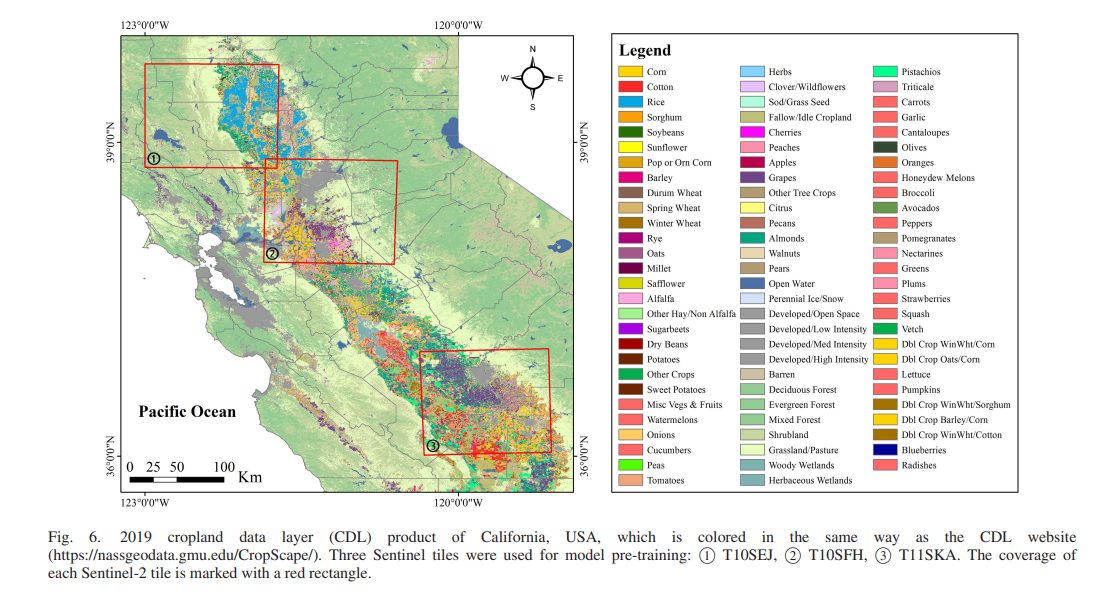
2) Author introduces an end-to-end deep learning architecture

as an effective alternative to convolutional and recurrent

neural networks for the SITS classification.

3) Author conducted comprehensive experiments on three largescale datasets to validate the effectiveness of the proposed method.

**Results:**



**Ideas for Enhancement:**

The evaluation on three benchmark datasets have revealed that the proposed pretraining scheme is generally effective for several deep learning models, i.e., CNN, Bi-LSTM, and transformer.