Contents

1   LEARNING TO DETECT ROADS

1.1   Unsupervised Pretraining

Traditionally neural networks have been initialized with small random weights. However, it has recently been shown that using an unsupervised learning procedure to initialize the weights can significantly improve the performance of neural networks. Using such an initialization procedure has been referred to as pretraining.

A set of inclusion and exclusion criteria was ascertained as competency factors to identify previous studies and subjects based on the purpose of this work. The exclusion factors were as follow.

- The full text of the papers was not provided by publishers;
- Remote sensing images were not used in the he papers;
- Peer-reviewed papers, such as conferences and journals;
- Articles written in English;

Although the methods utilized for road extraction used different data, this study can provide the following important outcomes

(1) The capabilities of deep learning methods for road extraction are more effective than those of regular approaches
(2) The low efficiency of the proposed methods in terms of data quality, training dataset, and model hyperparameters is presented Table(1)
(3) Occlusions, such as shadows, cars, and buildings, are similar to road features, such as colors, reflectance, and patterns. Road extraction remains challenging owing to such issues Figure(1)
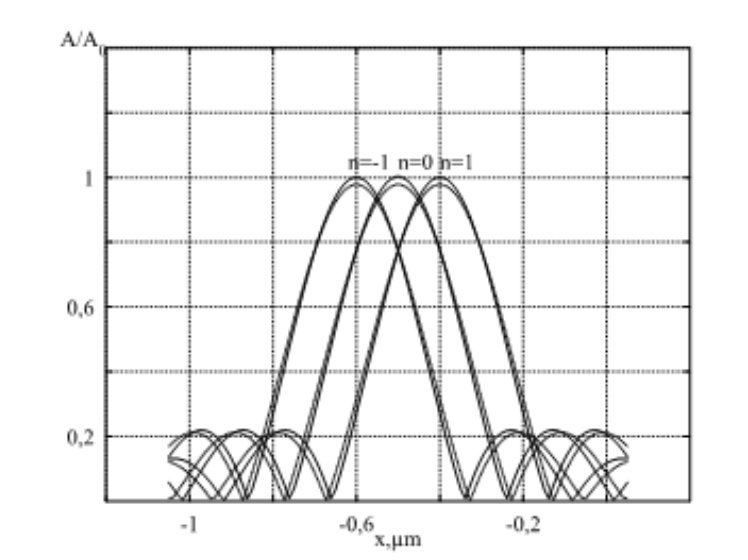

Fig. 1. The pixel resolution of most

Table 1. The confusion matrix is a detailed

Year	Channel	Resolution
2017	N,R,G,B	15cm/px
2018	N,R,G,B	10cm/px

where Z is a normalizing constant and the energy $E(v, h)$ formula(1) is:

$$E(V, h) = \sum_{i=1} V_i^2 - \left(\sum_{k=1} b_k h_k + \sum_{i=1} w_i V_i h_k \right). \tag{1}$$

In this work, we use Mean Squared Error (MSE) formula(2) as the loss function:

$$^N \frac{1}{N} \sum V_i^2 ||Net(I_i; W) - S_i||^2, _{,i=1} \text{ where: } N - \text{ is the number} \tag{2}$$

of the training samples, W – stochastic gradient descent.

1.2 Adding Rotations

When training the neural network f we found that it is useful to rotate each training case by a random angle each time it is processed. Since many cities have large areas where the road network forms a grid, training on data without rotations will result in a model that is better at detecting roads at certain orientations. By randomly rotating the training cases the resulting models do not favor roads in any particular orientation.

We propose a large-scale learning approach to road detection that addresses all three problems as follows

- We use synthetic road/non-road labels that we generate from readily available vector road maps. This allows us to generate much larger labelled datasets than the ones that have been used in the past;
- By using neural networks implemented on a graphics processor as our predictors we are able to efficiently learn a large number of features and use a large context for making predictions;

One major study of our work is the effectiveness of training multi-spectral data in image recognition.

(1) We simultaneously investigate the impact of using models with different complexities Figure(2);
(2) Aerial farmland images contain annotations with vastly different sizes;
(3) In order to justify our choice of using 512 × 512 windows to construct the Agriculture-Vision dataset, we additionally generate two versions of the dataset with different window sizes;

The dropout method operates on the fully connected layers to avoid overfitting because a fully connected layer usually contains a large number of parameters formula(3)
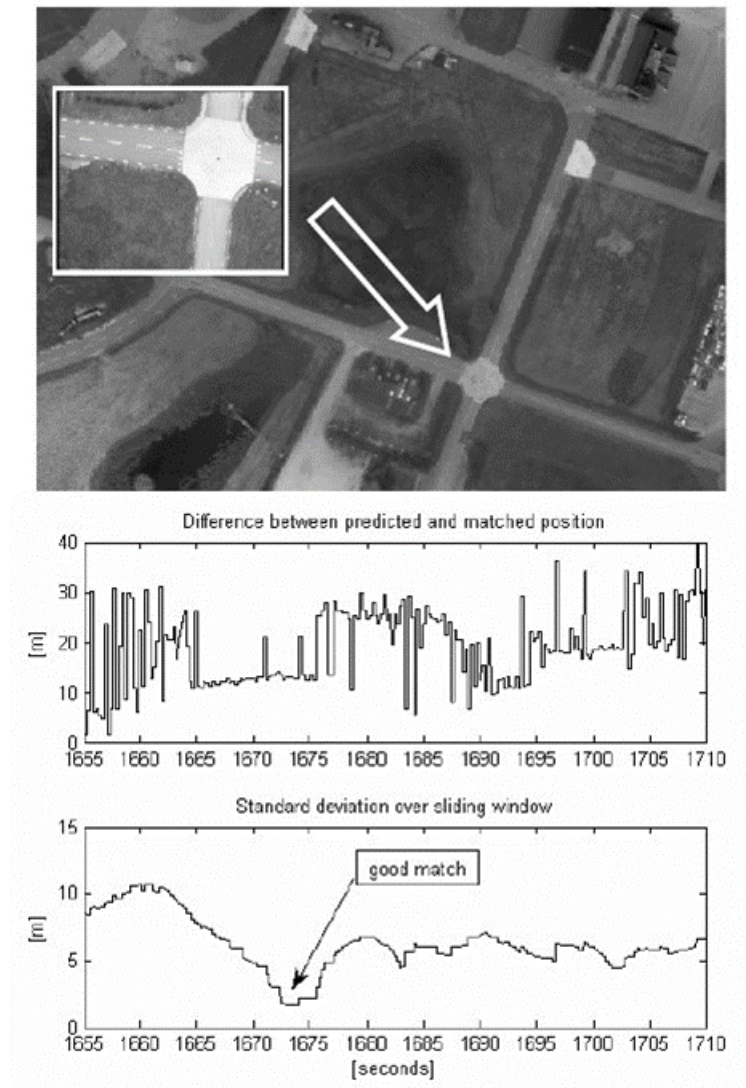
$$P(y_i) = \frac{e_i^Y}{\sum_{k=1}^{c} e_i^y} \tag{3}$$

Fig. 2. The pixel resolution