Alexander A. DOUDKIN, Alexander V. INYUTIN, Albert I. PETROVSKY, Maxim E. VATKIN
United Institute of Informatics Problem, National Academy of Science of Belarus
Surganov st., 6, 220012, Minsk, Belarus,
{doudkin, itekan}@newman.bas-net.by, {avin, vatkin}@lsi.bas-net.by, http://lsi.bas-net.by
Bolesław SOBKOWIAK
Industrial Institute of Agricultural Engineering
Starolecka st., 31, 60963, Poznan, Poland, http://www.pimr.poznan.pl

THREE-LEVEL NEURAL NETWORK FOR DATA CLUSTERIZATION ON IMAGES OF INFECTED CROP FIELD

Summary

The objective of this research was to use neural network approach for segmentation problem of agricultural landed-fields in remote sensing data. A neural network clusterization algorithm for segmentation of the color images of crop field infected by diseases that change usual color of agricultural plants is proposed. It can be applied for cartography of fields infected by plant diseases to reduce the use of plant protection products.

Introduction

Machine vision systems are widely used for inspection of growing plants to recognize their diseases using trichromatic features of leaves [1]. A goal is to sort data into some groups according to given parameters, i.e. to solve segmentation problem.

One approach for segmenting agricultural landed-fields in digital aerial images is using a generalization of region growing techniques combined with deformable models [2]. This mixed approach is called Region Competition. The goal of this approach is to alleviate the tasks of digitizing the region contours, to obtain the vector representation of the features that appear in an aerial photo. Deformable models (snakes) are defined as elastic curves that dynamically adapt a vector contour to a region of interest by applying energy minimization techniques. At the same time, given the problem of agricultural-land segmentation, we need region-growing approach to divide the raster image into homogeneous parcels. Region Competition combines the best features of Snakes/Balloon models and Region Growing techniques. In operation time, these techniques are applied to the case of having only two regions: the parcel to be segmented and its complementary.

The ARTMAP neural network can also be used in remote sensing [3]. ARTMAP belongs to the family of adaptive resonance theory (ART) networks, which are characterized by their ability to carry out fast, stable learning, recognition, and prediction, with a training procedure that requires only one pass through the data. These features differentiate ARTMAP from the family of feed forward multilayer perceptions (MLPs). The ARTMAP neural network mapping method presented here automatically produces vegetation maps from spectral and terrain data. As a supervised learning system, ARTMAP is trained by example. The approximately accuracy of ARTMAP mapping is about 80%.

The methods based on mathematical morphology are also used for the agricultural field image segmentation. We are proposing the image segmentation algorithm by the grayscale pseudo-skeleton operation [4]. The grayscale pseudo-skeleton is a modification of classical binary skeleton in the mathematical morphology. In this approach the image processing algorithm consist from 4 stage - morphology gradient, grayscale pseudo-skeleton operation, watershed transformation and region merging. There is an operation of a watershed in mathematical morphology, which allows to segment the map with a high accuracy and which does not use the threshold operation. As a result, the application of the watershed immediately to the gradient of the initial map gives too many areas. It will be explained by the presence of noise on the map and of the insignificant variations of brightness invisible to an eye. As a result, a common time of the map processing is sharply increased owing to the stages of boundaries check and areas merging. The applying of grayscale pseudo-skeleton operation is allowed to reduce effect of oversegmentation.

In [5] it is proposed to use modular neural networks for data processing for detect and identify vegetation (areas of vegetation) infected or polluted by bio-agents. The reduction of influence of distorting factors is done, and classification of spectral curves of chemical components is provided using this approach. Vegetation state estimation (i.e. the detection of chemical components) was carried out on the base of spectral curves shape obtained from leaf reflectance. The modular neural networks can provide better results than traditional neural network paradigms.

The results of [6] have shown the feasibility of image capture/processing and fuzzy logic control in the development of a precision farming herbicide application system. Weeds can be located by the greenness method and a fuzzy logic controller automatically manages herbicide applications to obtain effective weed control, reduce costs, and minimize soil and water pollution. The fuzzy logic membership functions are very easy to modify and control instructions can be obtained quickly. These manageable properties are essential to precision farming systems.

In [7] a remote sensing technology for automatically detection rice field is proposed. The principal of this technology is applying a region based classification by means of integrating geographical data and domain knowledge with multitemporal Image. Based on the principal, there methods of investigating the temporal Normalized Difference of Vegetation Index profile to detect rice fields were implemented. They are profile Matching, peak Detection, and difference classification.

The Advanced Spaceborne Thermal Emission and Reflectance Radiometer (ASTER) offers improved spatial, spectral and radiometric resolutions for various applications. Hence, in [8] purpose was to evaluate the utility of multispectral ASTER imagery in the discrimination and mapping of agricultural crops, soil and related land cover types. Four agricultural land cover attributes were specifically considered for spectral separability assessment and mapping: crop type, crop growth stages, soil colour and soil texture. The supervised classification in this study utilised the minimum distance to means algorithm.

In [9] the segmentation algorithm was proposed that bases on clusterisation. The results of image processing are shown in Fig. 1. Feature extraction for this algorithm is discussed in [10].

In this paper we propose neural network realization of clusterization that allows increasing accuracy of disease recognition.

Three- level neural network for clusterusation

Since color of plant on a RGB-picture is represented by color of leaves, time of day and overcast, we have allocated as input parameters for approximation of disease plant functions the following parameters:

1. Averages RGB components of plant leaves.
2. Height of the sun above horizon.
3. Average spectral capacity of daylight at moment of photographing.

As input training data for system are represented in the form of a set of discrete values, it is necessary to develop an approximator which could interpolate disease level values depending on any values of observable parameters.

It is proposed to use as such approximator a three-layer neural network with back-propagation training. The first layer has 5 neurons according to quantity of the parameters, the second layer has 3 neurons (it is obtained empirically). The output layer has one neuron, which activity gives a level of plant infection.

Training algorithm

The training algorithm of the proposed neural network is based on back-propagation algorithm. For each type of diseases we train an individual neural network. Thus, for unknown input data we receive N various diseases s_i, where $i = 1 \ldots N$. Further, we define a membership function to the disease as following: $\mu_i = \dfrac{s_i}{\sum_i s_i}$.

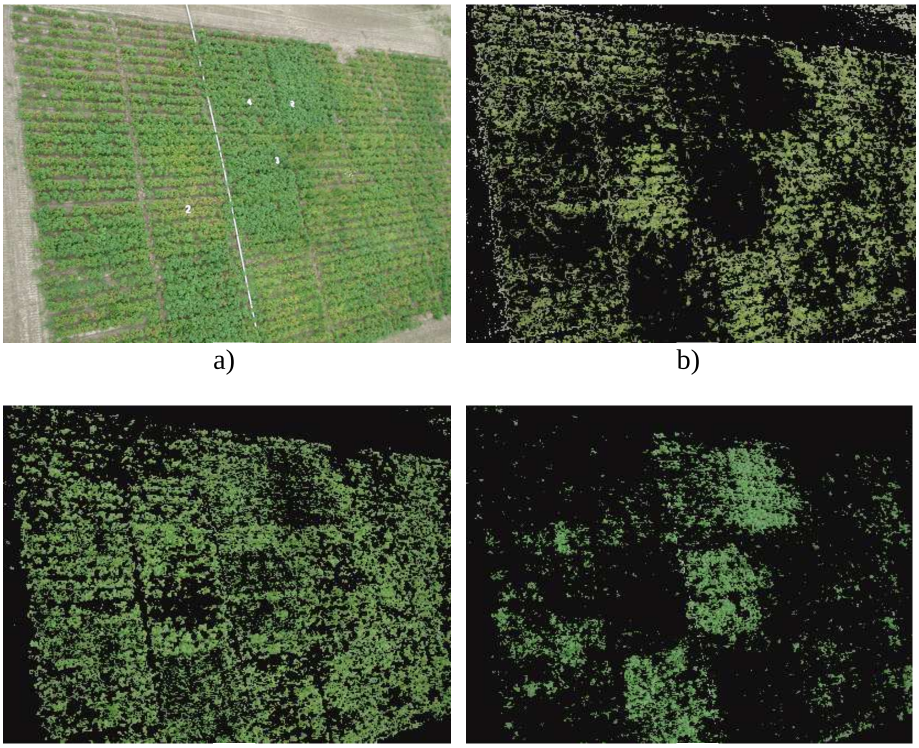
Back-propagation algorithm of return distribution is described by following relations:

$$E^p = \frac{1}{2}\sum_{j=1}^{N_j} \beta_j^2 = \frac{1}{2}\sum_{j=1}^{N_j} (d_j^p - y_j^p)^2$$

- mean-square deviation of a network for p-th training image, d_j - desirable target activity of j-th output neuron y_j - real activity of j-th output neuron. $y_j^p = F(s_j^p)$ - target activity of j-th neuron;

$s_j^p = \sum_i w_{i,j} y_i^p + \theta_j$ – the weighed sum on an input j- th neuron, where $F(\cdot)$ – function of neuron activation, $w_{i,j}$ – weight factor between i-th and j- th neurons, θ_j – threshold of j- th neuron;

$\Delta_p w_{i,j} = -\alpha \dfrac{\partial E^p}{\partial w_{i,j}}$ – amount of change between i-th



a) b) c) d)

Fig. 1. a) Input image, b) Segments of severe disease, c) Segments of medium disease, d) Sound segments

and j- th neurons, where α - an adaptive step;

$$\frac{\partial E^p}{\partial w_{i,j}} = \frac{\partial E^p}{\partial y_j^p} \cdot \frac{\partial y_j^p}{\partial s_j^p} \cdot \frac{\partial s_j^p}{\partial w_{i,j}}$$ – a partial weight factor derivative of a mean-square deviation;

$$\frac{\partial E^p}{\partial y_j^p} = \beta_j = d_j^p - y_j^p$$ – a deviation of j- th output neuron;

$$\frac{\partial y_j^p}{\partial s_j^p} = F'(s_j^p)$$ – a weighed sum derivative of activation function;

$$\frac{\partial s_j^p}{\partial w_{i,j}} = y_i^p$$ – a partial weight factor derivative of the weighed sum;

$$w_{i,j}(t+1) = w_{i,j}(t) + \Delta_p w_{i,j}(t) =$$
$$= w_{i,j}(t) - \alpha \cdot \beta_j \cdot F'(s_j^p) \cdot y_i^p$$ – the final formula of weight factors changes between the hidden and output neuron layers.

Similar relations are true for i– th neuron of hidden layer.

As a result we receive the final formula for correction of weight factors between the input and hidden neuron layers:

$$w_{k,i}(t+1) = w_{k,i}(t) + \Delta_p w_{k,i}(t) =$$
$$= w_{k,i}(t) - \alpha \cdot \beta_i \cdot F'(s_i^p) \cdot y_k^p \quad .$$

For each type of diseases we train an individual neural network.

Thus, for unknown input data we receive N various diseases s_i, where $i = 1 \ldots N$.

Method clusterization of fuzzy data

Let's realize clusterization according to possible diseases of plants: we consider that the plant belongs to j- th cluster according to the following criterion:

$\max_i(s_i) > S_t$,where S_t - the threshold defining, if the plant is ill.

We define a grade of membership of a plant to disease μ by the function $\mu_i = \frac{s_i}{\sum_i s_i}$. We make clusterization according to possible plant diseases: we shall consider, that the plant belongs to j- th cluster according to the following criterion:

$\mu_j = \max_i(\mu_i)$, $\max_i(s_i) > S_t$, where S_t - the threshold, defining if the plant is ill.

References

[1] Boleslaw Sobkowiak, Tadeusz Pawlowsky. Analiza obrazow kolorowych zaifekowanych pól uprawnych. Część 1. Preprowadzenie testu I ocena wyników badań // Wybrane zagadnienia ekologiczne we współczesnym rolnictwie. – PIMR, Poznań, 2005.- P. 113-117.

[2] Margarita Torre, Petia Radeva. "Agricultural-Field Extraction on Aerial Images by Region Competition Algorithm," icpr, vol. 01, no. 1, p. 1313, 15th 2000.

[3] G. A. Carpenter, S. Gopal, and C. E. Woodcock, "A neural network method for efficient vegetation mapping," Remote Sensing Environment, vol. 70, no. 3, pp. 326--338, Dec. 1999.

[4] A.V. Inyutin. The algorithm of image segmentation by grayscale pseudo-skeleton. Proceedings of the III International Conferences on Neural Networks and Artificial Intelligence (ICNNAI 2003), November 12-14, Minsk. Belarus, p.263-265.

[5] Natalia Kussul, Michael Kussul, Anatoly Sachenko, George Markowsky, Sergiy Ganzha. Remote Sensing of Vegetation Using Modular Neural Networks. Proceedings of the III International Conferences on Neural Networks and Artificial Intelligence (ICNNAI 2003), November 12-14, Minsk. Belarus, p.232-234.

[6] Yang, C.C.; Prasher, S.O. ; Landry, J.A. ; Perrett, J. , Ramaswamy, H.S. 2000. Recognition of Weeds with Image Processing and their use with Fuzzy Logic for Precision Farming. Canadian Agricultural Engineering. 42(4) : 195-200

[7] Tseng, Y.H., P.H. Hsu, and Y.H. Chen, 1998. Automatic Detecting Rice Fields by Using Multispectral Satellite Images, Land-parcel Data and Domain Knowledge, Proceedings of the 19th Asian Conference on Remote Sensing, Manila, Philippines, 16-20 November, pp. R-1-1~R-1-7.

[8] Apan, Armando and Kelly, Rob and Jensen, Troy and Butler, David and Strong, Wayne and Basnet, Badri (2002) Spectral Discrimination And Separability Analysis Of Agricultural Crops And Soil Attributes Using Aster Imagery. In 11th Australasian Remote Sensing and Photogrammetry Conference, 2-6 September, pages 396-411, Brisbane, Queensland.

[9] Analysis of colour images of infected crop field. Part 2. Two-stage algorithm of segmentation and improvement of color images of infected crop field / Alexander A. Doudkin et.// Wybrane zagadnienia ekologiczne we wspolczesnym rolnictwie – PIRM, Poznan, 2005.- P. 118-122.

[10] Ваткин М.Е., Ганченко В.В., Дудкин А.А., Петровский А.И., Собковяк Б. Выделение информативных признаков болезни картофеля по цветовым характеристикам листьев

The research is partially supported by the Belarusian Republican Foundation of Fundamental Research, grant T05MC-060.