# How Can a Massive Training Artificial Neural Network (MTANN) Be Trained With a Small Number of Cases in the Distinction Between Nodules and Vessels in Thoracic CT?[1]

Kenji Suzuki and Kunio Doi

**Rationale and Objectives.** To demonstrate that a massive training artificial neural network (MTANN) can be adequately trained with a small number of cases in the distinction between nodules and vessels (non-nodules) in thoracic computed tomography (CT) images.

**Materials and Methods.** An MTANN is a trainable, highly nonlinear filter consisting of a linear-output multilayer artificial neural network model. For enhancement of nodules and suppression of vessels, we used 10 nodules and 10 non-nodule images as training cases for MTANNs. The MTANN is trained with a large number of input subregions selected from the training cases and the corresponding pixels in teaching images that contain Gaussian distributions for nodules and zero for non-nodules. We trained three MTANNs with different numbers (1, 9, and 361) of training samples (pairs of the subregion and the teaching pixel) selected from the training cases. In order to investigate the basic characteristics of the trained MTANNs, we applied the MTANNs to simulated CT images containing various-sized model nodules (spheres) with different contrasts and various-sized model vessels (cylinders) with different orientations. In addition, we applied the trained MTANNs to nontraining actual clinical cases with 59 nodules and 1,726 non-nodules.

**Results.** In the output images for the simulated CT images by use of the MTANNs trained with small numbers (one and nine) of subregions, model vessels were clearly visible and were not removed; thus, the MTANNs were not trained properly. However, in the output image of the MTANN trained with a large number of subregions, various-sized model nodules with different contrasts were represented by light nodular distributions, whereas various-sized model vessels with different orientations were dark and thus were almost removed. This result indicates that the MTANN was able to learn, from a very small number of actual nodule and non-nodule cases, the distinction between nodules (spherelike objects) and vessels (cylinder-like objects). In nontraining clinical cases, the MTANN was able to distinguish actual nodules from actual vessels in CT images. For 59 actual nodules and 1,726 non-nodules, the performance of the MTANN decreased as the number of training samples (subregions) in each case decreased.

**Conclusions.** The MTANN can be trained with a very small number of training cases (10 nodules and 10 non-nodules) in the distinction between nodules and non-nodules (vessels) in CT images. Massive training by scanning of training cases to produce a large number of training samples (input subregions and teaching pixels) would contributed to a high generalization ability of the MTANN.

**Key Words.** Computer-Aided Diagnosis (CAD); Lung Nodule; Cancer; Thoracic CT; Artificial Neural Network.

© AUR, 2005

**Acad Radiol 2005;** 12:1333–1341

[1] From the Kurt Rossmann Laboratories for Radiologic Image Research, Department of Radiology, The University of Chicago, 5841 South Maryland Avenue, Chicago, IL 60637. Supported by USPHS Grants Nos. CA62625 and CA98119. Received January 10, 2005; revision received and accepted June 16. **Address correspondence to:** KS. e-mail: suzuki@uchicago.edu

© AUR, 2005
doi:10.1016/j.acra.2005.06.017

Lung cancer is the leading cause of cancer deaths among Americans (1). Low-dose helical computed tomography (LDCT) has been used for early detection of lung cancer (2–6). Radiologists, however, may fail to detect lung nodules in CT images that are visible in retrospect (7,8). Therefore, a computer-aided diagnostic (CAD) scheme for detecting lung nodules in CT images (9–15) has been investigated as a tool for improving radiologists' detection accuracy. A major problem with current CAD schemes is a relatively large number of false positives, which is likely to lower radiologists' efficiency in using a CAD scheme. Therefore, it is important to reduce the number of false positives as much as possible while a high sensitivity is maintained. It is difficult, however, to eliminate false positives without removal of any true-positive nodules, because variations in patterns of nodules and non-nodules are large (eg, there are various-sized nodules with different contrasts and various-sized lung vessels with different orientations in CT images; actually, the major source of false positives are lung vessels) (16).

Artificial neural networks (ANNs) have been applied for distinction between lesions and nonlesions (false positives) (17,18) and for distinction between malignant and benign lesions (19–22) in CAD schemes, and they have been shown to be useful for various CAD schemes (17–24). For achieving a high and reliable performance for nontraining cases, a large number of training cases (eg, 400–800 cases) are commonly required (25,26). If an ANN is trained with only a small number of cases, the generalization ability (performance for nontraining cases) will be lower (ie, the ANN may fit only the training cases); this is known as "overtraining" (or "overfitting") (27). Because diagnostic radiology is progressing rapidly as technology advances, the timely development of CAD schemes is important. However, it is very difficult to collect a large number of abnormal cases for training, particularly for a CAD scheme with a new modality, such as lung cancer screening with multidetector-row CT (MDCT).

Massive training ANNs (MTANNs) have been developed for reducing the number of false positives in CAD schemes for LDCT images (16) and chest radiographs (28). With an MTANN, 54% of 1,726 false positives were removed without eliminating any of 58 true-positive nodules in a database of 63 LDCT scans containing 63 primary lung cancers (16). The MTANN was trained with only 10 nodules and 10 non-nodules (29), whereas other ANNs usually require training with a large number of cases because ANNs generally have a large number of parameters to be determined. However, it was not clear how and why the MTANN can be trained with a small number of cases and can provide a high performance even for nontraining cases.

Our purpose in this study was to demonstrate and to verify that an MTANN can be trained with a small number of cases in the distinction between nodules and vessels (non-nodules) in a CAD scheme for detecting nodules in thoracic CT images.

## MATERIALS AND METHODS

### Database

Our database in this study consisted of 68 LDCT scans acquired from 68 patients who participated voluntarily in a lung cancer screening program between 1996 and 1999 in Nagano, Japan (2). The CT examinations were performed on a mobile CT scanner (CT-W950SR; Hitachi Medical, Tokyo, Japan). The CT scans were acquired with a low-dose protocol of 120 kVp, 25 mA or 50 mA, 10-mm collimation, and a 10-mm reconstruction interval at a helical pitch of 2. The pixel size was 0.586 mm or 0.684 mm. Each reconstructed CT slice had an image matrix size of 512 × 512 pixels, and the number of gray levels was 4,096. The number of CT slices per scan was 31 or 33. The 68 scans included 71 lung cancers that were determined by biopsy or surgery. The size (effective diameter) of the 71 cancers ranged from 6 mm to 24 mm, with a mean of 14 mm. These cancer cases included nodules in three different categories—pure ground glass opacity (GGO or non-solid) nodules (24% of nodules), mixed GGO (or part-solid) nodules (30%), and solid nodules (46%). A training set for the MTANNs used in this study included 10 LDCT scans containing 10 nodules obtained from our "missed" cancer database (8), in which 38 cancers were not reported or misreported during the initial clinical interpretation and were identified retrospectively.

### Our CAD scheme

Our CAD scheme for detecting lung nodules in CT (30) consisted of a difference-image technique, a multiple gray-level–thresholding technique, extraction of image features, and a rule-based scheme. To summarize the methodology, lung segmentation was performed by use of thresholding. Nodules in the segmented lungs were enhanced by use of the difference-image technique. Nodule candidates were identified by application of the multiple

Download English Version:

https://daneshyari.com/en/article/9387333

Download Persian Version:

https://daneshyari.com/article/9387333

Daneshyari.com