Deep Sensing: Cooperative Spectrum Sensing Based on Convolutional Neural Networks

Woongsup Lee, Minhoe Kim, Dong-Ho Cho, and Robert Schober

Abstract

In this paper, we investigate cooperative spectrum sensing (CSS) in a cognitive radio network (CRN) where multiple secondary users (SUs) cooperate in order to detect a primary user (PU) which possibly occupies multiple bands simultaneously. Deep sensing, which constitutes the first CSS framework based on a convolutional neural network (CNN), is proposed. In deep sensing, instead of the explicit mathematical modeling of CSS, the optimal strategy for combining the individual sensing results of the SUs is obtained with a CNN based on training sensing samples. Accordingly, an environment-specific CSS is found in an adaptive manner regardless of whether the individual sensing results are quantized or not. Through simulation, we show that the performance of CSS can be significantly improved by the proposed deep sensing scheme, especially in the low signal-to-noise ratio (SNR) regime, even when the number of training samples is moderate.

Index Terms

Cognitive radio network, cooperative spectrum sensing, deep learning, convolutional neural network, spatial correlation.

I. INTRODUCTION

In cognitive radio networks (CRNs), secondary users (SUs) dynamically utilize the unused channels which are owned by primary user (PU) [1]. Given that the top priority of SUs is not to interrupt the operation of the PU, it is of utmost importance to determine whether the PU is present or not. Consequently, spectrum sensing to detect the presence of the PU is one of the

W. Lee is with the Department of Information and Communication Engineering, Gyeongsang National University, Republic of Korea. M. Kim and D.H. Cho are with the School of Electrical Engineering, Korea Advanced Institute of Science and Technology. R. Schober is with the Institute for Digital Communication of the Friedrich-Alexander-University Erlangen-Nürnberg, Erlangen, Germany.
most important research topic in CRNs [2]–[5]. Furthermore, the tolerable interference margin of the PU is typically small, which makes spectrum sensing in CRNs challenging.

Given that the sensing results of individual SUs are susceptible to errors due to the fluctuating channel conditions, cooperative spectrum sensing (CSS), where the individual sensing results from multiple SUs are combined to determine the presence of the PU, has been proposed. However, deriving the optimal CSS strategy is difficult in general as it depends on the network environment in which each SU is situated [3]. For example, SUs which are close to the PU (e.g., SU A and B in Fig. 1) are likely to detect the PU more reliably than SUs which are far away from the PU (e.g., SU C in Fig. 1). Moreover, because of the spatial correlation of wireless channels, SUs which are close to each other (e.g., SU A and B in Fig. 1) are likely to report similar sensing results. The problem becomes more complicated if the PU can utilize multiple spectrum bands simultaneously and the SUs do not know in advance which bands are used by the PU. In addition, the leakage of transmit power to adjacent bands also needs to be taken into account for CSS. In fact, deriving the optimal CSS mathematically is extremely difficult. Hence, simple yet efficient CSS strategies such as the $K$-out-of-$N$ scheme, which decides that a PU is present if $K$ SUs out of $N$ SUs detect the PU, are widely considered in the literature [3], [6]. Although the heterogeneity of the individual sensing results of the SUs is considered in [3], [6], a simplified system model is assumed for CSS which entails a performance degradation.

Recently, deep learning has gained considerable attention in many fields of computer science, especially in image processing, speech recognition, and natural language processing, due to the significant performance gains it can achieve compared to conventional schemes [7]–[9]. A deep neural network (DNN) is a multi-layer network which emulates the neurons of the human brain. In a DNN, the optimal strategy to classify data, e.g., speech and images, is learned from a large amount of sample data via the back-propagation algorithm, without the need for developing
complicated hand-crafted mathematical models for the data. By building a sufficiently large network, the DNN based model can emulate the behavior of highly nonlinear and complicated systems, which is one of the benefits of deep learning.

Although DNNs have been mainly applied for image processing and speech recognition so far, they can also be applied in wireless communication systems, especially for classifying communication signals [7], [10]–[12]. In [7], spectrum sensing for a single SU was considered and DNNs were shown to yield a significant performance gain over conventional schemes. Moreover, the authors of [11] showed that the data traffic type can be determined accurately with Long Short Term Memory (LSTM), which is a special type of a recurrent neural network (RNN). Furthermore, in [12], a deep belief network was employed for the classification of PUs.

In this paper, we propose deep sensing which is a new DNN-based CSS scheme for CRNs. The contributions of this paper can be summarized as follows.

1) We propose deep sensing, which is a CSS scheme for CRNs employing convolutional neural networks (CNNs), a particular type of DNN. In the proposed scheme, the optimal strategy for combining individual sensing results is learned from training sensing samples, independent of the type of sensing decisions at individual SUs, i.e., both hard decision (HD) and soft decision (SD) sensing are included [4].

2) Deep sensing exploits the spatial and spectral correlations of the channels such that the CSS combining strategy is optimized according to the location of the SUs and the characteristics of the PU in an adaptive manner. To the best of the authors’ knowledge, this is the first work that applies deep learning for CSS.

3) We evaluate the performance of the proposed scheme based on computer simulations. Our results confirm that the proposed scheme can achieve higher sensing accuracy compared to conventional approaches, especially when the power of the noise is large. Moreover, we show that the proposed scheme achieves a high performance even when the number of training samples is moderate, which facilitates the adoption of the proposed scheme in practice.

The remainder of this paper is organized as follows. In Section II, we describe the system model. The proposed deep sensing scheme is introduced in Section III. Simulation results are provided in Section IV, and Section V concludes the paper.
II. System Model

We assume that $N_{SU}$ SUs and a single PU are randomly distributed in a given area and move with velocity $v$ such that the locations of the SUs and the PU change over time. Moreover, we assume a multi-channel system with $N_B$ bands whose bandwidth is $W$. The PU can simultaneously utilize $N_{Bp}$ consecutive bands, while the SUs do not know which bands are used by the PU. Furthermore, we assume that the transmit power of the PU in a given band is fixed to $P$ and can be also leaked to adjacent bands. We assume that the amount of power leakage to the adjacent bands is $\eta P$, where $\eta$ is the proportion of power leakage. In addition, we assume that the noise power spectral density is $N_0$.

A simplified path-loss model with path-loss exponent $\alpha$ and path-loss constant $\beta$ is adopted for modeling the channel between the PU and the SUs. Let $d_i$ be the distance between the PU and SU$_i$, then the path-loss is $\beta d_i^\alpha$. The effect of multi-path fading, $g_{i,j}$, is also considered where $i$ is the index of the SU and $j$ is the index of the band. In this work, $g_{i,j}$ is modeled as an independent zero-mean circularly symmetric complex Gaussian (CSCG) random variable. Furthermore, spatially correlated shadow fading is taken into account [13]. Specifically, we assume that $h_i$ is the shadow fading of the channel between SU$_i$ and the PU in dB and it follows a normal distribution with mean zero and variance $\sigma$ [13]. Then, a normal random variable with zero mean and unit variance, which we denote as $k_i$ (normalized shadow fading), can be obtained as $\frac{h_i}{\sigma}$.

Let SU$_A$ and SU$_B$ be separated by distance $d_{A-B}$. Then, the correlation of the normalized shadow fading between these SUs, i.e., $k_A$ and $k_B$, which we denote as $\rho_{cor}(d_{A-B})$, is modeled as [13]

$$\rho_{cor}(d_{A-B}) = \mathbb{E}[k_A k_B] = e^{-\frac{d_{A-B}}{d_{ref}}},$$

where $d_{ref}$ denotes a reference distance whose value depends on the environment, e.g., rural or urban. Accordingly, the shadow fading of two SUs which are nearby will have a high correlation, i.e., the SUs experience similar shadow fading. The received signal strength (RSS) of the PU at two SUs which are close to each other will be similar due to both path-loss and shadow fading, which makes similar sensing outcomes for these SUs likely.

We assume that the SUs perform individual spectrum sensing in all $N_B$ bands for a time period of $\Delta t$. Therefore, $N_B \times N_{SU}$ individual sensing outcomes can be collected by the CRN
in every time period of $\Delta t$. Each SU can report either the measured RSS for each band (SD) or binary sensing results for each band using a sensing threshold $\gamma$ (HD). In the latter case, the SU reports 1 if the measured RSS exceeds $\gamma$ and 0 otherwise.

The CRN determines the presence or absence of the PU by combining the reported individual sensing results. Although a higher sensing accuracy can be achieved with SD, the signaling overhead will be much larger compared to HD. In this paper, $N_{\text{sample}}$ samples of individual sensing data are used for training and evaluation of our model. This sensing data is generated by computer simulation. Note that each sensing sample corresponds to a different location of the PU and the SUs because they move with velocity $v$, i.e., for every sensing sample, the positions of the PU and the SUs change by $v \Delta t$.

III. DEEP SENSING FOR COOPERATIVE SPECTRUM SENSING

In deep sensing, the CNN model is used to combine individual sensing results to determine the presence of the PU. CNNs have been widely used in deep learning research to classify images by exploiting their spatial characteristics. In CNN, multi-dimensional convolution is applied to an image to extract its spatial features. The recent success of CNNs in visual classification [8], [9] motivates the adoption of CNNs for CSS, because similar to the correlation of nearby pixels, the individual sensing results from nearby SUs and adjacent bands are likely similar due to their spatial and spectral correlation.

In general, a CNN is composed of two parts, which are the convolution part at the front of the network and the fully connected (FC) part at the back of the network, cf. Fig. 2. In the

1In this work, a synthetic data set is used for learning, whereas in a practical system, real data would be used. Nevertheless, the performance evaluation based on the synthetic data can also provide meaningful insights regarding the performance of the proposed scheme [7].
convolution part, the spatial features of the input data are extracted using the convolutional layer and the rectified linear unit (ReLu) layer, which are usually placed in tandem. The convolution layer can be considered as a spatial filter which performs spatial convolution of the input data. For example, let $A$, $B$, and $C$ be the input, weight, and output of the convolution layer. Then, when the size of the convolution is $3 \times 3$, $C[m, n] = \sum_{i=0}^{2} \sum_{j=0}^{2} A[i + m - 1, j + n - 1] B[i, j]$. After the spatial features have been extracted, the ReLu layer introduces non-linearity to the CNN [9]. When the input of the ReLU layer is $x$, the output is given by $\max(x, 0)$, such that negative inputs are blocked. Without the ReLU, the deep sensing scheme would be linear and unable to classify non-linear behavior. The ReLu layers can be followed by the max pooling layer which reduces the size of the data. With the max pooling layer, the computational overhead can be efficiently reduced without significant performance loss as shown in [8].

The FC part performs classification based on the output of the convolution part by gathering the results of feature extraction and making the final decision. To this end, all the data is mixed and fed into the softmax operator in which the final decision on whether the PU is present or not is made using the softmax function. Let index 0 and 1 denote the absence and the presence of the PU, respectively, and $X_S$ be the input of the softmax operation. Then, in our model, the CRN determines that the channel is idle if $\frac{e^{W_0 X_S}}{\sum_{i=0}^{e^{W_1 X_S}}} > \frac{e^{W_1 X_S}}{\sum_{i=0}^{e^{W_1 X_S}}}$, and otherwise it concludes that the channel is occupied by the PU, where $W_i$ is a weight matrix.

The proposed CNN structure for deep sensing is depicted in Fig. 2. We consider a relatively small sized neural network compared to what is normally used for image classification [8], [9] because unlike image classification where typically hundreds of different classes have to be distinguished, in our system model, we only need to classify two classes, namely the presence and absence of the PU. Given that the main aim of our work is to show the usefulness of deep learning for CSS, we do not optimize the set of system parameters, e.g., the size of the convolution layers, and leave this for future work.

To be more specific, in our scheme, individual sensing results from different bands and different SUs constitute the two dimensional input data for the CNN, cf. Fig. 2 which is unlike the DNN based spectrum sensing considered in [7] where the in-phase and quadrature-phase of the temporal signal are used to define a two dimensional matrix. The elements of input matrix can be either binary for HD or continuous for SD. For the convolution part of the network, three $3 \times 3$ convolution layers ($3 \times 3$ conv.) are utilized, because as can be concluded from [9], the $3 \times 3$ convolution is sufficient to extract the spatial features of the input data. Moreover, we have used
a relatively small number[^1] of layers in order to reduce the learning time. Furthermore, the depth of the first, second, and third convolution layers is set to 32, 64, and 128, respectively. The stride, which is the step size used in the convolution filter, is set to 1 and zero padding is used, so that the size of the output remains the same as that of the input. In addition, the cross-entropy loss is considered and the weights of each layer are trained using off-the-shelf stochastic gradient descent algorithms, i.e., Adam (adaptive moment estimation).

IV. PERFORMANCE EVALUATION

In this section, the performance of the proposed deep sensing scheme is examined. Our performance evaluation is implemented using Tensorflow which is an open-source software library for machine intelligence developed by Google. We assume that a single PU and multiple SUs which move at $v = 3 \text{km/h}$ are randomly deployed in a $500 \text{ m x 500 m}$ area. Moreover, the number of bands is set to 16 where $W = 10 \text{ MHz}$ and $N_{\text{SP}}$ is randomly chosen from 1 to 3, i.e., the PU can utilize up to three bands simultaneously. Furthermore, we assume that $\eta = -20 \text{ dB}$, $P = 23 \text{ dBm}$, $\beta = 10^{3.453}$, $\alpha = 3.8$[^14], $\sigma = 7.9 \text{ dB}$, and $d_{\text{ref}} = 50 \text{ m}$[^13].

We assume that each SU performs individual spectrum sensing every 2 seconds, i.e., $\Delta t = 2$ seconds[^3]. Threshold $\gamma$ is set to -107 dBm[^3] such that when HD is used, each SU sends 1 if the RSS exceeds $\gamma$ and sends 0 otherwise. When SD is used, each SU reports the measured RSS. For performance evaluation, 90% of $N_{\text{sample}}$ samples are used to train the model and the remaining samples are used for the evaluation.

For the performance metric, we consider the probability of false alarm ($P_{\text{FA}}$) which is the probability that an SU falsely detects a PU when no PU is present and the probability of miss detection ($P_{\text{MD}}$) which is the probability that an SU fails to detect a PU when one is present[^3]. We examine the sensing accuracy of our proposed scheme for both SD and HD. Moreover, conventional CSS based on the $K$-out-of-$N$ scheme[^2] is considered where the value of $K$ is selected such that $(P_{\text{FA}} + P_{\text{MD}})$ is minimized.

In Fig. 3, we show $P_{\text{MD}}$ and $P_{\text{FA}}$ as functions of $N_{\text{sample}}$ when $N_0 = -174 \text{ dBm/Hz}$ and $N_{\text{SU}} = 32$. We observe that in all cases, the proposed scheme provides a much higher sensing accuracy and the gap between the accuracy of the conventional scheme and that of our proposed scheme.

[^1]: The size of the input data for our application is likely much smaller than image data, because the number of bands and SUs are likely less than 100, respectively. This also justifies the use of a small-size CNN.
increases as the number of samples increases. Even for moderate numbers of sensing samples, e.g., $N_{\text{sample}} = 400$, the performance gap between the proposed scheme and the conventional scheme is considerable, i.e., $P_{\text{MD}}$ and $P_{\text{FA}}$ are reduced by more than 70%. This result suggests that the proposed scheme will outperform the conventional scheme in practical CRNs where the number of training samples will likely be moderate. Furthermore, we can observe that deep sensing with SD achieves a smaller sensing error compared to that with HD because individual sensing results contain more information [4].

In Fig. 4, we show $P_{\text{MD}}$ and $P_{\text{FA}}$ as functions of $N_0$ when $N_{\text{sample}} = 400$ and $N_{SU} = 32$. As expected, the sensing error decreases as the noise density decreases because the individual sensing results become more accurate. Accordingly, when $N_0 = -184$ dBm/Hz, all schemes achieve small sensing errors. However, when the power of the noise increases such that the CRN operates in the low signal-to-noise ratio (SNR) regime where the individual sensing outcomes are inaccurate, the proposed scheme significantly outperforms the conventional scheme. Moreover,
we can observe that even when the power of the noise is large, e.g., $N_0 = -164$ dBm/Hz, deep sensing with SD can achieve a high sensing accuracy, i.e., both $P_{MD}$ and $P_{FA}$ are less than 12 \%, which highlights the benefits of the proposed scheme.

In Fig. 5, we show $P_{MD}$ and $P_{FA}$ as functions of $N_{SU}$ when $N_0 = -174$ dBm/Hz and $N_{sample} = 400$. We observe that CSS becomes more accurate as the number of SUs increases, because the effect of individual sensing errors can be mitigated. In all cases, we find that $P_{MD}$ and $P_{FA}$ are much lower for deep sensing than for the conventional scheme. Especially, when $N_{SU} = 96$, we observe that deep sensing with SD achieves $P_{MD}$ and $P_{FA}$ of almost 1 \%, and deep sensing with HD achieves $P_{MD} = 2.4$ \% and $P_{FA} = 4.3$ \% which are sufficiently low for efficient operation of the CRN.

V. CONCLUSIONS

In this paper, a novel CNN-based CSS scheme for CRN was proposed. In deep sensing, the optimal strategy for combining the binary or real valued individual sensing results of the SUs is learned using a CNN. Through simulations we investigated the performance of the proposed scheme and found that deep sensing significantly outperforms conventional CSS, especially in the low SNR regime. The proposed scheme is the first attempt to use deep learning for CSS in CRN.

REFERENCES

[1] J. Mitola and G. Q. Maguire Jr, “Cognitive radio: Making software radios more personal,” IEEE Personal Communications, vol. 6, no. 4, pp. 13–18, Aug. 1999.
[2] E. Peh, Y.-C. Liang, Y. L. Guan, and Y. Zeng, “Optimization of cooperative sensing in cognitive radio networks: A sensing-throughput tradeoff view,” IEEE Trans. Veh. Technol., vol. 58, no. 9, pp. 5294–5299, Nov. 2009.

[3] W. Lee and D.-H. Cho, “Enhanced spectrum sensing scheme in cognitive radio systems with MIMO antennae,” IEEE Trans. Veh. Technol., vol. 60, no. 3, pp. 1072–1085, Mar. 2011.

[4] J. So and W. Sung, “Group-based multibit cooperative spectrum sensing for cognitive radio networks,” IEEE Trans. Veh. Technol., vol. 65, no. 12, pp. 10 193–10 198, Dec. 2016.

[5] H. Reyes, S. Subramaniam, N. Kaabouch, and W. C. Hu, “A spectrum sensing technique based on autocorrelation and Euclidean distance and its comparison with energy detection for cognitive radio networks,” Computers & Electrical Engineering, vol. 52, pp. 319 – 327, May 2016.

[6] W. Lee and D.-H. Cho, “Improved cooperative spectrum sensing in multiple stages for low-power primary users,” IEEE Wireless Commun. Lett., vol. 2, no. 3, pp. 287–290, Jun. 2013.

[7] T. J. O’Shea, J. Corgan, and T. C. Clancy, “Convolutional radio modulation recognition networks,” in EANN, Aberdeen, U.K., Sep. 2016.

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in NIPS, Stateline, NV, Dec. 2012, pp. 1097–1105.

[9] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” arXiv preprint arXiv:1409.1556, 2014.

[10] T. J. O’Shea and T. C. Clancy, “Deep reinforcement learning radio control and signal detection with KeRLym, a Gym RL Agent,” arXiv preprint arXiv:1605.09221, 2016.

[11] T. J. O’Shea, S. Hitefield, and J. Corgan, “End-to-end radio traffic sequence recognition with deep recurrent neural networks,” arXiv preprint arXiv:1610.00564, 2016.

[12] Y. Cui, X. j. Jing, S. Sun, X. Wang, D. Cheng, and H. Huang, “Deep learning based primary user classification in cognitive radios,” in Proc. of ISCIT, Nara, Japan, Oct. 2015.

[13] A. Algans, K. Pedersen, and P. Mogensen, “Experimental analysis of the joint statistical properties of azimuth spread, delay spread, and shadow fading,” IEEE J. Sel. Areas Commun., vol. 20, no. 3, pp. 523–531, Apr. 2002.

[14] IEEE 802.20, “IEEE 802.20 channel models document,” January 2007.