Wide-band spectrum sensing with convolution neural network using spectral correlation function

Anupama Rajanna1, Srimannarayana Kulkarni2, Sarappadi Narasimha Prasad3
1School of Electronics and Communication Engineering, REVA University, Bengaluru, India
2Department of Electronics and Communication Engineering, BNM Institute of Technology, Bengaluru, India
3Department of Electrical and Electronics Engineering, Manipal Institute of Technology Bengaluru Manipal Academy of Higher Education, Manipal, India

ABSTRACT
Recognition of signals is a spectrum sensing challenge requiring simultaneous detection, temporal and spectral localization, and classification. In this approach, we present the convolution neural network (CNN) architecture, a powerful portrayal of the cyclo-stationarity trademark, for remote range detection and sign acknowledgment. Spectral correlation function is used along with CNN. In two scenarios, method-1 and method-2, the suggested approach is used to categorize wireless signals without any previous knowledge. Signals are detected and classified simultaneously in method-1. In method-2, the sensing and classification procedures take place sequentially. In contrast to conventional spectrum sensing techniques, the proposed CNN technique need not bother with a factual judgment process or past information on the signs’ separating qualities. The method beats both conventional sensing methods and signal-classifying deep learning networks when used to analyze real-world, over-the-air data in cellular bands. Despite the implementation’s emphasis on cellular signals, any signal having cyclo-stationary properties may be detected and classified using the provided approach. The proposed model has achieved more than 90% of testing accuracy at 15 dB.

Keywords: Convolution neural network classification Deep neural network Long-term evolution Wide-band spectrum sensing Wireless communication systems

1. INTRODUCTION
Modern wireless communication systems are able to accommodate a dramatic growth in data traffic. To keep up with the rising demand, wireless communication networks must maximize their use of the available spectrum [1]. A similar objective is sought after by drives like little cell sending, mm-wave band use, powerful range use calculations, gigantic quantities of information, multiple input, multiple output (MIMO) frameworks [2], and mental radio organizations. Accordingly, mental radios attempt to satisfy this reason by progressively doling out range to clients; as a result, signal identification and spectrum sensing have grown in significance as tools for cognitive radio networks. Spectrum efficiency will become even more important for heterogeneous networks as the need for unified communications, sensing, and localization increases with the advent of six generation (6G) and beyond. For instance, dynamic spectrum sharing is a ground-breaking technology that enables 5G and long-term evolution (LTE) to operate together in a corresponding band, introduced in 5G New Radio Release-16. In addition, it is expected that the frequency band used by radar and communications systems would be combined [3], [4].

Journal homepage: http://ijece.iuuescore.com
The amount of second request clamor insights, cyclic frequencies, and certain properties of the pulse shaping filter are a priori data that are required for the energy detection and matched filtering sensing procedures, which are employed in spectrum sensing and signal identification, to work. To finish the sensing process, a statistical decision mechanism should be used following the analysis of the signals that were gathered [5]. Because wireless communications networks need to make quick decisions at all levels of operation, the old sensing paradigm cannot keep up with the demands of this continuously changing operating environment (5G and beyond). When it comes to fixing the problems with parameter adaptation that plague traditional methods, deep learning (DL) has been offered as a viable option. This is because it is well known that DL methods may use the convolutional process to extract the inherent properties of inputs. A statistical decision process is also eliminated when identification is done using a DL-based technique. This is in line with the results of a recent study [6] that shows how DL techniques outperform conventional methods for signal identification in the spectral domain. To meet the requirements for 5G and previous remote organizations, a clever radio plan for range detection and transmission identity is expected. This arrangement might be carried out with the assistance of artificial intelligence/machine learning procedures [7] by using qualities like cyclo-stationarity of signs [8].

2. RELATED WORKS

Convolution neural networks (CNNs) are first prepared for adjustment orders using high-order insights of single transporter signals [9], while the utilization of simulated intelligence procedures for range detection and sign acknowledgment is audited. The “fast Fourier transform (FFT), amplitude-phase representation (AP), and in-stage quadrature (I/Q)” highlights are used to construct a CNN classifier for the purpose of tweak and blockage detection for modern logical clinical groups [10]. Fully connected neural networks were used by a separate group of O’Shea et al. [11] to classify protocols in the industrial, scientific, and medical (ISM) band. The suggested model has good performance, but it uses MATLAB-created synthetic data. The signal characteristics get even more convoluted due to the various phenomena present in real-world channels. Nonetheless, cyclo-stationarity signal investigation has been ready for balance order, boundary assessment, and range detection for over 20 years. In the field of cognitive radio, cyclo-stationary feature detection (CFD) has established itself as a standard method for spectrum sensing and is also used to distinguish between common modulations like M-ary phase shift keying (M-PSK), multiple frequency-shift keying (M-FSK), and M-ary quadrature amplitude modulation (M-QAM) [12], [13].

2.1. Research contribution

The researchers propose the spectral correlation function (SCF) as an element contribution to a CNN for blind remote sign location. To overcome the challenges of range detection and transmission distinguishing proof, a novel CNN model is made and prepared, utilizing the otherworldly SCF of remote signs without bi-recurrence planning. Subsequently, the prescribed technique might be utilized to either check if a sign is in the range or to recognize signs [14], [15]. The presentation of the methodology’s detecting and distinguishing proof is affirmed and approved, utilizing genuine over-the-air signal estimation. The suggested technique uses a two-dimensional procedure to find and pinpoint issues. In method-1, the constructed CNN model is given the SCFs of estimations from global system for mobile communication (GSM), universal mobile telecommunications system (UMTS), and long-term evolution (LTE), as well as the SCF of range, which is entirely made up of clamor. Detection and order are completed at the same time in method-1. In contrast, method-2 has a two-pronged approach: first, spectrum occupancy is measured through a spectrum sensing technique, and then signals are classified.

As a key part of the suggested strategy, it has been seen that using only the relevant half of the input matrices improves classification performance while reducing training time and complexity. The entire dataset, however, is made available to the public in [16]. This dataset was compiled using measurements collected during a large-scale measurement effort that was carried out in a variety of locations and frequency bands. Because of this, the metric-based dataset is made accessible to all scientists so they may use it to improve and validate their own research.

2.2. CNN

CNNs, a type of DNN, are often used for image recognition and classification. Inputs to a CNN are processed in a manner analogous to the human visual system. As opposed to fitting data, it instead extracts features from an input [17]. In this research, we use input matrices that are analogous to images with features placed in certain locations. Nonetheless, it has lately been expanded to a number of new fields of use. Features are extracted and then classified in the first step of a convolutional neural network. After feature extraction
using a convolutional layer, the data from the preceding layers is combined with a pooling layer in the convolutional layer.

2.3. Problem statement

If modern remote organizations want to get the most out of their resources in a fast-changing communication environment, they need to be able to find and separate complex communication mediums quickly, reliably, and in different ways [18]. In this situation, range detection and flag identification become very important tools for the board. We do this by using a two-pronged DL approach to sensing and identifying issues: First, all classes—including GSM, UMTS, LTE, and empty spectrum, commonly known as additive white Gaussian noise (AWGN)—are used to train a new CNN classifier. A CNN classifier that has already been pre-trained on four distinct classes receives the cyclic spectrum after that. Finally, all the sorting has been done. In method-2, a two-stage technique is used, as shown in Figures 1(a) and 1(b). A CNN locator is used to initially decide if a sign happens in the designated band by preparing it with two classes. GSM, UMTS, and LTE signals make up the top notch, though AWGN is the main signal type in the second class. Thus, in the first step, as with traditional spectrum sensing, a determination is made as to whether or not a signal exists in the spectrum. If the second step determines that a signal carrying information is present in the band in question, to identify the kind of signal present in the band, a CNN classifier is activated and fed data from three classes (GSM, UMTS, and LTE).

Figure 1. Two approaches for spectrum sensing, (a) method -1 and (b) method -2

3. METHOD

The proposed strategy depends on an original CNN model. CNN is planned and executed for the characterization of versatile remote correspondence data using the free and open-source artificial intelligence (AI) toolkit Keras [19]. Six layers all out-three convolutional layers and three pooling layers—make up the proposed CNN model. There are 64 filters, 128 filters, and 64 filters in the convolution layers. Two layers, each of which is completely linked, cap off the network. There are 256 neurons in the top-secret layer. For both methods, the quantity of neurons in the second secret layer fluctuates. To extricate highlights that are useful for characterization, every convolution layer utilizes the leaky rectified linear unit (ReLU) initiation capability with an alpha value of 0.1. All things considered, a broken form of ReLU is chosen. Broken ReLU, as opposed to ReLU, maps enormous negative qualities to more modest ones using a line with an unobtrusive incline. Three-by-three channels are used in each convolutional layer. Dimensionality reduction and time spent training are both achieved with the use of 22-max pooling. The 256 neurons and flaws in the ReLU initiation capability give a totally connected layer. The probabilities for each are not entirely settled by the Softmax initiation capability after the completely associated layers. The optimal model parameter values are also determined using the Adam optimizer. Over-fitting is avoided during training by stopping early. For early halting, we set the patience at 10 epochs, and we’ll be keeping an eye on validation loss while we train. If the approval misfortune joins a level and stays at that level for 10 years, preparation is terminated, and the last loads are used for testing. The suggested model begins with the input matrices, which are convolved using filters. Figure 2 shows a block schematic of the suggested CNN paradigm.

When considering the reasoning behind creating this CNN model, it is important to remember that the first convolution layer employs 3×3×64 filters to collect information about shifts in the mapped output’s local areas. A smaller filter is more appropriate for this purpose of recognizing high-frequency characteristics since the SCF causes local and cyclic frequency changes. As a result, the layers account for variations at the local level. The next layer analyses elements like area and size connected to these features after an initial layer selects
the periodic properties of every neighboring word in an interaction. By raising the number of filters to 128, we want to take care of the cyclic properties in more depth. Before transmitting the gathered attributes to the thick layer, the final layer averages them all. As a result, the final layer’s number of filters must be carefully chosen to ensure that sufficient data is gathered while preventing overfitting. For the very last layer, we chose 64 filters because of this. The effectiveness of a classifier model is often evaluated using precision (P), recall (R), and the F1-score. The F1-score gives a general indication of a classifier model’s exhibition since it is the concordant normal of accuracy and review. While recall gives specifics on the percentage of actual positive discoveries, precision assesses the percentage of positive findings.

![Convolution and Pooling Layers](image1)

![Dense Layers](image2)

**Figure 2.** The proposed CNN model with three convolutional layers and two dense layers

### 3.1. Dataset generation and measurement

The data set used to test and evaluate the suggested method was made up of many estimates made in different places and at different rates. Focus is placed on the frequency bands used for cellular communications at 800, 900, 1,800, and 2,100 MHz. A bunch of Yagi-Uda receiving wires and a Rohde Schwarz FSW26 range analyzer are utilized by the collector. The actions have been normalized as follows: There are 16,384 I/Q samples taken for every signal in the spectrum. The tests were conducted using fifteen different signal-to-noise ratio (SNR) settings. There are 4,000 total signals in the game, and each level has its own set of them. To that end, we captured 60,000 signals and added them to the dataset. Figure 3 shows power spectra from Welch’s technique for various signal types. The dataset is divided into test and training information at a proportion of 0.4 to 0.6 while utilizing the proposed system. Due to the varying locations, times, and frequency bands used for the tests, the received power distribution varies greatly.

![Power Spectra](image3)

**Figure 3.** Snapshot of spectrum

Second, Figure 4 shows the phase distribution of four independent recordings of all three signals. The phase of the received signals falls nearly entirely between 0 and 360 degrees. This finding suggests that the amplitude and phase distributions of received signals exhibit Rayleigh-like fading behavior. When considering the measuring area as well as the positions of the transmitters and receivers, this is the predicted outcome. In [20], the data set is made available for download in SCF format. The dataset contains 60,000 SCF matrices, each representing one of the 16,384 I/Q samples received.
4. RESULTS AND DISCUSSION

Using the extensive dataset provided, we test how well the suggested classification model performs. The dataset comprises wideband code division multiple access (WCDMA) for UMTS and LTE signals that were kept in the field from various areas and in more favorable conditions, such as a different total number of channel taps and a different degree of fading. In the training set, there are 9,000 signals per waveform, whereas there are 6,000 signals in the test set. An I/Q signal is 16,384 bytes long. Videos of actual events are used to educate and assess CNN.

We begin with the findings from method-1, which is a four-class classification issue, as was previously indicated. This is because the information was compiled from cellphone signals captured during a thorough measuring effort. We have included the 900 MHz frequency for completeness; however, all cellular bands have been tested. Figure 5 shows that at 11 dB SNR, the model’s test accuracy is over 90%. It has a most extreme exactness of 91% at 15 dB.

The result of the CNN identifier be analyzed using the techniques portrayed at 15 dB and above, assuming the sign is available in the range. In Figure 5, at a 3 dB signal-to-noise ratio, the classification accuracy is found to be more than 85%. The highest efficiency is achieved at 8 dB, and it remains constant up to 15 dB. The disarray networks for the CNN classifier in method-2 demonstrate that the classifier can precisely recognize GSM flags even in low-SNR systems. In any method, the general accuracy of the classifier is low. This truly means that, as opposed to GSM signals, it has trouble perceiving UMTS and LTE signals in the low SNR range. It stands to reason that GSM signals would exhibit characteristics typical of Gaussian distributions in the presence of a high signal-to-noise ratio (SNR), given that Gaussian minimum shift keying (GMSK) is frequently associated with GSM. Due to AWGN, the received signal acquires more Gaussian characteristics as the signal-to-noise ratio (SNR) decreases. When the model is presented with low-SNR UMTS or LTE signals, it is more likely to learn Gaussian features in an effort to reduce its loss function. When put to the test, it is anticipated that the trained model will be able to identify signals with predominant Gaussian properties. Thus, in the lower SNR domain, when UMTS and LTE signals lose their distinctive properties, the model can recognize GSM signals, which intrinsically represent Gaussian qualities.

At this point, we may evaluate method-2 as a whole and see how well it performs. Misdetection during the sensing phase naturally results in a drop in performance. Both the sensing stage’s detection rate and the classification stage’s accuracy are good above 3 dB, so there is no substantial reduction in total performance. Figure 6 demonstrates that, compared to method-1, method-2 has a far better overall performance. It has been shown that the signals left after the CNN detector has first detected and separated noise from the signal set may...
be classified with significantly greater performance, especially at low SNR levels. This allows for improved efficiency in method-2. It is vital to remember that method-2 requires a greater number of assets than case 1 does, both in terms of preparation time and model count.

![Figure 5. Test Accuracy in contrast to signal-to-noise ratio](image)

![Figure 6. Test Accuracy for method-1 and method-2](image)

Most researchers think that traditional ways of sensing the spectrum, like the energy detector and the matching filter, are better at sensing. Note, however, that, unlike some other studies, this one uses actual data from the real world rather than simulated or made-up information. Energy detectors are one example. They can find signals in spectra with high sensitivity, but they can only do this if they know how the noise changes in that spectrum. Even a small mistake in predicting the noise variation has a big effect on how well the sensor works. Spread-spectrum signals, like WCDMA in UMTS, are also close to the noise floor because their power is spread out over a wide range of frequencies. Consider this: Han et al. [21] claim energy detection rate for spread-spectrum signals is inadequate in a fading scenario. However, matching filters are a solution that is particular to a certain waveform and requires an exhaustive understanding of signals.

4.1. Examining the role of variables in the overall effect

This section contrasts the capabilities of SCF with those of the most often used features in sensing applications, such as I/Q, AP, and FFT. Table 1 displays the results of this analysis. Because of the huge blurring impact on the period of the sign, I/Q cannot make a pertinent contribution to the model, as opposed to the balance characterization tests [22], [23]. This is why the results of I/Q tests are so poor.
Table 1. Performance comparison

<table>
<thead>
<tr>
<th>Network</th>
<th>Signal</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convolutional long short-term memory deep neural network (CLDNN) [24]</td>
<td>UMTS</td>
<td>0.337</td>
<td>1.00</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>LTE</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>GSM</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>0.109</td>
<td>0.32</td>
<td>0.20</td>
</tr>
<tr>
<td>Long short-term memory (LSTM) [25]</td>
<td>UMTS</td>
<td>0.337</td>
<td>1.00</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>LTE</td>
<td>0.0000</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>GSM</td>
<td>0.0000</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>0.109</td>
<td>0.32</td>
<td>0.20</td>
</tr>
<tr>
<td>Ours</td>
<td>UMTS</td>
<td>0.78</td>
<td>1.00</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>LTE</td>
<td>0.99</td>
<td>0.81</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>GSM</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>0.94</td>
<td>0.92</td>
<td>0.94</td>
</tr>
</tbody>
</table>

4.2. Analyzing pre-existing deep learning networks

Cellular communication signals are classified using pre-existing DL networks. The classification of modulation begins with these models. The information network and information vector, as proposed in the articles, are used in the examination with little to no changes to the models. Convolutional long short-term memory deep neural network (CLDNN) stores the sufficiency and period of every I/Q test in a 2128-layer lattice. Conversely, the long short-term memory (LSTM) model utilizes a vector-reshaped variant of the CLDNN network. The length of the vector is 256 components. Quadrature components make up the remainder of the vector, while in-phase components make up its first half. Although I/Q vectors and matrices may be used to train LSTM and CLDNN quickly and effectively, doing so results in subpar classification performance.

4.3. Comparison with support vector machine

Previously, we used support vector machines (SVMs) to recognize signals from the actual world. SVM training should be carried out independently for each SNR level, despite the fact that using SCF in SVM improves performance [26]. To put it another way, at the end of training, more models should be made if the dataset contains more SNR values. SVM’s application is restricted and should be worked on because of the necessity of an SNR assessor and the memory stacking of all pre-prepared models, as shown in Figure 7. The CNN-based classifier uses a cheaper feature (no bi-frequency spectrum mapping is required) without sacrificing performance. Since the suggested CNN-based technique uses a training set with equal numbers of signals from each SNR, the resulting model is robust regardless of the SNR. This means that in the testing phase, classification for a wide range of signal-to-noise ratios can be handled by a single model.

Figure 7. Comparison between SVM classifier and CNN classification

5. CONCLUSION

This paper demonstrated a DL-based method for spectrum sensing and signal identification that does not need any prior knowledge and uses SCF as an input to a new CNN model. The first strategy looks at the possibilities of wireless signal detection and classification working together. To continue, we use a sequential method. The findings favor the sequential method over the combined one. In addition, a comparison to other existing features used in DL-based detector models demonstrated SCF’s superiority as a distinguishing feature.

Wide-band spectrum sensing with convolution neural network using spectral ... (Anupama Rajanna)
Also, the findings suggest that the proposed method’s CNN model outperforms other known DL models, SVMs, and conventional CFD in terms of spectral sensing performance under challenging channel circumstances. These findings demonstrate that DL-based approaches are useful in the dynamic communication environment of today’s wireless communication networks. Future research should examine how well the suggested approach works for detecting and identifying different types of cyclic characteristics present in wireless signals or modulation schemes. Further, the suggested method’s performance may be studied considering adversarial assaults, and other methods can be developed to bolster its resistance to incursion. Despite the fact that supervised learning is the primary emphasis of this article, the suggested approach might benefit from the incorporation of unsupervised learning techniques for use in feature expansion, which would further boost its performance.

REFERENCES


BIOGRAPHIES OF AUTHORS

Anupama Rajanna received a B.Eng. degree in electronics and communication engineering and an M.Tech. in signal processing from Visvesvaraya Technological University, Karnataka, India. Currently, she works as a research scholar at REVA University, Bengaluru, India. Her areas of research interest are cognitive radio and compressive sensing. She can be contacted at anupamarajanna@mail.com.

Srimannarayana Kulkarni holds a PhD in VLSI design from I.I.T. Bombay, India. He is currently a professor and additional director at BNMIT, Bangalore, India. His areas of specialization are VLSI, embedded systems, mobile communication satellite communication advanced computer architecture. He can be contacted at sy_kul@yahoo.com.

Sarappadi Narasimha Prasad is a professor in School of ECE at REVA University, India since 2012. His total experience is 22 years, completed graduation from Mangalore University, post-graduation from VTU, and a doctorate from Jain University. He has more than 80 journals/conferences in profile and is presently guiding 8 research scholars. His areas of interest are AI, embedded systems, and signal processing. He can be contacted at sn.prasad@manipal.edu.