

COMBINING TIME-FREQUENCY SIGNAL ANALYSIS AND MACHINE LEARNING WITH AN EXAMPLE IN GRAVITATIONAL-WAVE DETECTION

Jonatan Lerga, Nikola Lopac, David Bačnar, Franko Hržić

Summary

This paper presents a method for classifying noisy, non-stationary signals in the time-frequency domain using artificial intelligence. The pre-processed time-series signals are transformed into time-frequency representations (TFRs) from Cohen's class resulting in the TFR images, which are used as input to the machine learning algorithms. We have used three state-of-the-art deep-learning 2D convolutional neural network (CNN) architectures (ResNet-101, Xception, and EfficientNet). The method was demonstrated on the challenging task of detecting gravitational-wave (GW) signals in intensive real-life, non-stationary, non-Gaussian, and non-white noise. The results show excellent classification performance of the proposed approach in terms of classification accuracy, area under the receiver operating characteristic curve (ROC AUC), recall, precision, F1 score, and area under the precision-recall curve (PR AUC). The novel method outperforms the baseline machine learning model trained on the time-series data in terms of all considered metrics. The study indicates that the proposed technique can also be extended to various other applications dealing with non-stationary data in intensive noise.

Keywords: non-stationary signals; time-frequency representations; artificial intelligence; machine learning; convolutional neural networks; gravitational waves.

1. INTRODUCTION TO TIME-FREQUENCY SIGNAL ANALYSIS

A signal is most commonly presented as a time-dependent function that can be considered a representation characterized by perfect resolution in the time domain with no direct insight into frequency content. The second most common signal representation is the signal's frequency spectrum. It is obtained by calculating a Fourier transform on the time-series signal; however, on the other hand, it provides no information on time instants when frequency components occurred in the original signal (their time localization), thus providing only perfect frequency resolution, losing time resolution [1, 2].

The Fourier transform, and its inverse are defined as follows:

$$X(f) = \mathcal{F}\{x(t)\} = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt, \quad (1a)$$

$$x(t) = \mathcal{F}^{-1}\{X(f)\} = \int_{-\infty}^{\infty} X(f)e^{j2\pi ft} df, \quad (1b)$$

where $x(t)$ is the original signal in the time domain, $X(f)$ is the Fourier transform of the original signal, t is the time variable, f is the frequency variable, \mathcal{F} is the Fourier transform, and \mathcal{F}^{-1} is its inverse.

Time-frequency (TF) signal analysis is a concept that considers both the time and frequency dimensions, allowing for tracking changes in both dimensions at the same time, thus leading to time-frequency representations (TFRs), which display the signal in a 2-dimensional (2D) TF plane [3, 4].

A basic TFR that overcomes some shortcomings of the Fourier transform is the short-time Fourier transform (STFT), obtained by performing the Fourier transform on the analyzed signal inside a sliding time window of a finite duration [5]. For each segment, the signal is assumed to be of a known duration, resulting in the signal spectrum as a function of time [6]. The STFT can thus be defined as follows:

$$STFT_x(t, f) = \int_{-\infty}^{\infty} x(\tau)h(\tau - t)e^{-j2\pi f\tau} d\tau, \quad (2)$$

where τ is the time instant center of the windowing function, and $h(t)$ is the moving windowing function.

The STFT is a linear TFR with well-localized components in time and frequency; however, a known shortcoming of the STFT is the resolution limits in time and fre-

quency domains (due to the fixed-sized window width). It is known that shorter time windows provide better time resolution at the cost of a worse frequency resolution; on the other hand, wider time windows provide better frequency resolution at the cost of a worse time resolution. Thus, it is impossible to adjust the time and frequency resolutions independently when using a single fixed-size windowing function due to Heisenberg's uncertainty principle [7, 8, 9].

However, by using a TFR, we get a better insight into both the time and frequency domains simultaneously when compared to only time-domain or frequency-domain representation. Thus, the TFR decomposes the original time series signal in a way that shows the time instants of changes in frequency content, or rather, it is now possible to monitor and track signals with non-stationary content and to estimate its instantaneous frequency [10, 11, 12, 13].

Several groups or classes of the TFRs exist; the three most important ones are Cohen's, affine, and reassigned classes. Cohen's class is covariant to shifts of the signal in the time and frequency domain [14]. On the other hand, the affine class is covariant to time shifts (or rather translations of the signal) and scaling (or rather dilations of the signal), and are thus known as time-scale representations. In certain cases, a TFR can both belong to Cohen's and affine classes at the same time, depending on its properties. One example of a TFR that belongs to Cohen's class has already been mentioned: the STFT. The most common example of an affine class TFR is the continuous wavelet transform, which linearly expands the signal onto wavelets, which are a set of analysis functions that can be shifted in time and scaled accordingly to fit the original signal [15].

Both Cohen's and affine classes are comprised of mostly bilinear and quadratic TFRs, which due to their non-linear nature, as a consequence, manifest the presence of cross-terms or interference terms in the TF plane [16]. Due to these phenomena, the TFR's understandability and ease of interpretation are hindered because of the overlapping of signal components with the cross-terms in the TF domain (which also can take on negative values due to the bilinear nature of those interactions). However, the quadratic or bilinear TFRs show improved TF resolution and concentration of auto-terms, despite introducing cross-terms [17, 18], and offer a valuable insight into the spectral nature of the signal, the number of frequency components, their time localization, and frequency modulation.

The reassignment class (sometimes referred to as the reassignment method) is a local post-processing method for the TFRs; its goal is to improve the concentration or sharpness of the signal by refocusing the energy distribution, thus achieving better readability of the TFR. The reassignment method is applicable to any bilinear TFR, including Cohen's and affine classes, thus suppressing the interferences and improving the TF localization of the original TFR to which it was applied. The reassignment method

is the second step of a two-step procedure where, first, the interfering components are smoothed out and reduced by using a 2D low-pass filter kernel; however, this also causes the smearing of the rest of the signal in the TF plane. Next, the second step is squeezing, where the signal terms are refocused after the smoothing procedure [19]. This method requires both the magnitude and phase of the observed signal, unlike the usual TFRs that require just the magnitude. The phase is necessary due to the fact that the components representing maximum contribution to the signal have slow phase variation with regard to time (thus, they can be detected by observing the phase's variation rate) [20].

The rest of the paper is structured as follows: Section 2 presents some TFRs from Cohen's class used as input for machine learning based classification. Section 3 describes three state-of-the-art deep learning architectures used for the application of gravitational wave (GW) detection given in this paper. Section 4 provides a short introduction to the gravitational wave detection problem and presents the results of the proposed procedure combining the TFRs and machine learning for their detection in intensive noise. Paper conclusions are found in Section 5.

2. COHEN'S CLASS OF TIME-FREQUENCY REPRESENTATIONS

Besides the already mentioned STFT, being a linear TFR, there are other TFRs in Cohen's class that are bilinear or quadratic. The simplest one is the Spectrogram (SP), which is obtained as the squared modulus of the STFT [5, 21, 22]:

$$SP_x(t, f) = \left| \int_{-\infty}^{\infty} x(\tau)h(\tau - t)e^{-j2\pi f\tau}d\tau \right|^2. \quad (3)$$

The Spectrogram (SP), providing low interference terms in the TFR, is not able to provide good resolution both in time and frequency at the same time. This is due to the fixed-sized window and Heisenberg's uncertainty principle, as mentioned before.

In order to reduce the number of cross-terms between positive and negative frequencies, most TFRs have adopted analytic forms of signals. The analytic signal is a complex signal obtained from the real-valued input signal by removing the negative frequencies from the spectral content (the real part is the original signal, and the imaginary part is the Hilbert transform of the original signal) as follows [4, 23]:

$$y(t) = x(t) + j\mathcal{H}\{x(t)\}, \quad (4)$$

where $y(t)$ is the complex-valued analytic form of the signal, $x(t)$ is the real-valued signal, and \mathcal{H} is the Hilbert transform.

Next, we present a list of other TFRs used in the paper. Let us start with the Margenau-Hill (MH) distribution, which is one of the general Cohen's class quadratic representations, defined as the real part of the product of two STFTs with independently sized time windows, as follows [18, 24]:

$$MH_x(t, f) = \Re\{STFT_h(t, f)STFT_g^*(t, f)\} = \int_{-\infty}^{\infty} \frac{1}{2}(x(t + \tau)x^*(t) + x(t)x^*(t - \tau))e^{-j2\pi f\tau}d\tau, \quad (5)$$

where \Re denotes the real part, h is the first and g is the second windowing function, the $STFT_h(t, f)$ and $STFT_g^*(t, f)$ are the STFTs for the respective windows, and $*$ is the complex conjugate.

The two independent windows approach allows for them to be individually chosen, so that they guarantee adequate time and frequency resolution, thus providing an improvement in resolution with regard to the SP. However, the drawback of this approach is the possibility of generating artifacts when the window lengths are significantly different in size, and when the signal contains components of similar frequencies [16].

The Wigner-Ville (WV) is another one of the fundamental TFRs for Cohen's class. It is defined as the Fourier transform of the signals' autocorrelation function as follows [25, 26]:

$$WV_y(t, f) = \int_{-\infty}^{\infty} y\left(t + \frac{\tau}{2}\right)y^*\left(t - \frac{\tau}{2}\right)e^{-j2\pi f\tau}d\tau, \quad (6)$$

where y is the analytic associate of the original real-valued signal. The WV improves the auto-term resolution and improves the TF resolution; however, it suffers from cross-terms when applied to multi-component signals due to its quadratic nature. The interference terms that affect the WV make a visual interpretation of representation difficult. As a way to reduce the WV interference terms, several modified methods were developed [27, 28], some of which we will showcase next.

The Pseudo Wigner-Ville (PWV) is obtained by applying a time windowing function $h(\tau)$ on the autocorrelation of the WV; it is defined as follows [27]:

$$PWV_y(t, f) = \int_{-\infty}^{\infty} h(\tau)y\left(t + \frac{\tau}{2}\right)y^*\left(t - \frac{\tau}{2}\right)e^{-j2\pi f\tau}d\tau. \quad (7)$$

The time windowing acts as filtering in the frequency smoothing that attenuates the cross-terms oscillating in the frequency direction. However, this decreases the frequency resolution for the auto-terms, while the cross-terms oscillating in the time direction remain unattenuated.

The Smoothed Pseudo Wigner-Ville (SPWV) is obtained by applying an additional smoothing window g on the autocorrelation of the PWV; it is defined as follows [27, 29]:

$$SPWV_y(t, f) = \int_{-\infty}^{\infty} h(\tau) \int_{-\infty}^{\infty} g(u-t)y\left(u + \frac{\tau}{2}\right)y^*\left(u - \frac{\tau}{2}\right) du e^{-j2\pi f\tau} d\tau. \quad (8)$$

The SPWV allows for the independent control of the two windowing function lengths $h(t)$ and $g(t)$ respectively, allowing for the independent control of the smoothing in both the time and frequency domains. However, a tradeoff exists between the TF resolution and the attenuation of interference terms.

Furthermore, suppressing the interference terms of the WV is possible by using kernel functions, which are often designed in the 2D Doppler-lag (ν, τ) domain, also known as the ambiguity domain (where the Doppler variable (ν) represents the frequency shift and is obtained as the Fourier transform of the time (t) variable) [4]. Defining an ambiguity function in the ambiguity domain allows for the applied Doppler-lag kernel functions to be perceived as the WV filtering operations [4]. The TFRs that are obtained through kernel smoothing of the WV, or rather designed to remove interference terms, are referred to as the reduced-interference distribution (RIDs) group of the TFRs; we will next define some of them.

The Born-Jordan distribution (BJ) uses a narrowband kernel in the ambiguity domain that contains the sinc function, which can also be viewed as the Fourier transform of the first B-Spline [30]. The BJ provides good cross-term suppression at the cost of a reduced resolution of the auto-terms [31]; it moreover preserves time and frequency supports [32]. The BJ is defined as follows [29, 33]:

$$BJ_y(t, f) = \int_{-\infty}^{\infty} \frac{1}{|\tau|} \int_{t-\frac{|\tau|}{2}}^{t+\frac{|\tau|}{2}} y\left(u + \frac{\tau}{2}\right)y^*\left(u - \frac{\tau}{2}\right) du e^{-j2\pi f\tau} d\tau. \quad (9)$$

The BJ effectively suppresses the noise and damps the so-called “ghost-frequencies”.

The Choi-Williams distribution (CW) uses an exponential kernel with a width of σ as a way to suppress the cross-terms. The choice of kernel width σ allows for the control of the suppression of interferences and resolution trade-off, with smaller values helping to reduce cross-terms and larger values leading to an increase in auto-term resolution. The CW is defined as follows [22, 34, 35]:

$$CW_y(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\sqrt{\sigma}}{2\sqrt{\pi}|\tau|} e^{-\frac{u^2\sigma}{16\tau^2}} y\left(t + u + \frac{\tau}{2}\right)y^*\left(t + u - \frac{\tau}{2}\right) e^{-j2\pi f\tau} dud\tau. \quad (10)$$

However, the CW does not provide the possibility of independent control over the amount of time and frequency smoothing. The weighing function does not decay along the axis in the ambiguity domain because, otherwise, it would not satisfy the marginal properties. Thus, the CW can only filter the cross-terms that result from interacting components differing both in time and frequency, or rather it shows higher interference levels for signals whose frequency or time supports overlap [36].

The Butterworth distribution (BUD) is a generalized extension of the CW that provides better preservation of auto-term resolution and better suppresses the low-frequency cross-terms [37]. The BUD is defined as follows [38, 39]:

$$BUD_y(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\sqrt{\sigma}}{2|\tau|} e^{-\frac{|u|\sqrt{\sigma}}{|\tau|}} y\left(t + u + \frac{\tau}{2}\right) y^*\left(t + u - \frac{\tau}{2}\right) e^{-j2\pi f\tau} du d\tau. \quad (11)$$

Its kernel acts as a 2D low-pass filter in the ambiguity domain that has variable pass-band and transition region parameters [40, 41].

The Zhao-Atlas-Marks distribution (ZAM) reduces the interference terms while simultaneously providing good time and frequency resolution using a cone-shaped kernel. The ZAM is obtained by frequency smoothing the BJ, as follows [42, 43]:

$$ZAM_y(t, f) = \int_{-\infty}^{\infty} h(\tau) \int_{t-\frac{|\tau|}{2}}^{t+\frac{|\tau|}{2}} y\left(u + \frac{\tau}{2}\right) y^*\left(u - \frac{\tau}{2}\right) du e^{-j2\pi f\tau} d\tau. \quad (12)$$

Finally, a RID with its kernel based on the first kind Bessel function (RIDB) is defined as follows [34, 38, 44]:

$$RIDB_y(t, f) = \int_{-\infty}^{\infty} h(\tau) \int_{t-|\tau|}^{t+|\tau|} \frac{2g(u)}{\pi|\tau|} \sqrt{1 - \left(\frac{u-t}{\tau}\right)^2} y\left(u + \frac{\tau}{2}\right) y^*\left(u - \frac{\tau}{2}\right) du e^{-j2\pi f\tau} d\tau. \quad (13)$$

The RIDB maintains high TF resolution due to preserving auto-terms while suppressing cross-terms; this is due to the low-pass filtering characteristics of the used Bessel kernel.

When applying a kernel based on binomial coefficients to a RID, we obtain a RID-BN, defined as follows [4, 45]:

$$RIDBN_y(t, f) = \sum_{\tau=-\infty}^{\infty} \sum_{u=-|\tau|}^{+|\tau|} \frac{1}{2^{2|\tau|+1}} \binom{2|\tau|+1}{|\tau|+u+1} y[t+u+\tau] y^*[t+u-\tau] e^{-j4\pi f\tau}. \quad (14)$$

If a kernel based on the Hanning window is used in the RID, we obtain a RIDH, defined as follows [33]:

$$RIDH_y(t, f) = \int_{-\infty}^{\infty} h(\tau) \int_{-\frac{|\tau|}{2}}^{+\frac{|\tau|}{2}} \frac{g(u)}{|\tau|} \left(1 + \cos\left(\frac{2\pi u}{\tau}\right)\right) y\left(t + u + \frac{\tau}{2}\right) y^*\left(t + u - \frac{\tau}{2}\right) du e^{-j2\pi f\tau} d\tau. \quad (15)$$

Finally, if a triangular window-based kernel is used in the RID, we obtain the RIDT, defined as follows [33, 34]:

$$RIDT_y(t, f) = \int_{-\infty}^{\infty} h(\tau) \int_{-\frac{|\tau|}{2}}^{+\frac{|\tau|}{2}} \frac{2g(u)}{|\tau|} \left(1 - \frac{2|u|}{|\tau|}\right) y\left(t + u + \frac{\tau}{2}\right) y^*\left(t + u - \frac{\tau}{2}\right) du e^{-j2\pi f\tau} d\tau. \quad (16)$$

Next, we present the deep-learning models used to classify the above described TFR images.

3. DEEP LEARNING METHODS

Machine learning, as a part of artificial intelligence, has experienced rapid and intensive development in recent years [46]. Deep learning, in particular, has spread to various research fields and commercial applications. This subfield of machine learning denotes the algorithms based on computational architectures called deep artificial neural networks (ANNs) and able to extract features from the provided input data automatically. They use them for the learning process, thus eliminating the need for expert-based knowledge and manual feature selection specific for each field [47, 48]. Convolutional neural networks (CNNs) represent a type of the ANNs specially designed to utilize multidimensional input data, and today provide state-of-the-art performances in image classification and recognition tasks [47, 48].

Therefore, due to their above-described mathematical properties beneficial for non-stationary signal analysis, the 2D TFRs have a great potential to be applied in the deep learning-based classification of non-stationary signals by combining them with the 2D CNNs used in pattern recognition. Three state-of-the-art CNN architectures suitable for application with the input data in the form of the TFR images are described next: ResNet, Xception, and EfficientNet.

3.1. ResNet

The ResNet CNN architectures reduce the impact of the vanishing gradient problem on the training process by utilizing the residual learning concept [49]. Thus, very deep

CNN models can be trained, achieving increased classification accuracy. Residual learning is based on the shortcut (residual) connections that transfer the input to the convolutional layers to their output, and then perform the element-wise addition of these feature maps [49]. In this way, non-linear CNN layers do not directly learn the underlying mapping, but rather learn residual functions [49].

The standard residual block consists of two layers with standard convolutions, an intermediate, and an output ReLU activation function. In addition, a bottleneck residual block, consisting of 1×1 , 3×3 , and 1×1 convolutional layers, two intermediate, and an output ReLU activation, is used in ResNet architectures, containing more than 50 layers to reduce the computational cost [49].

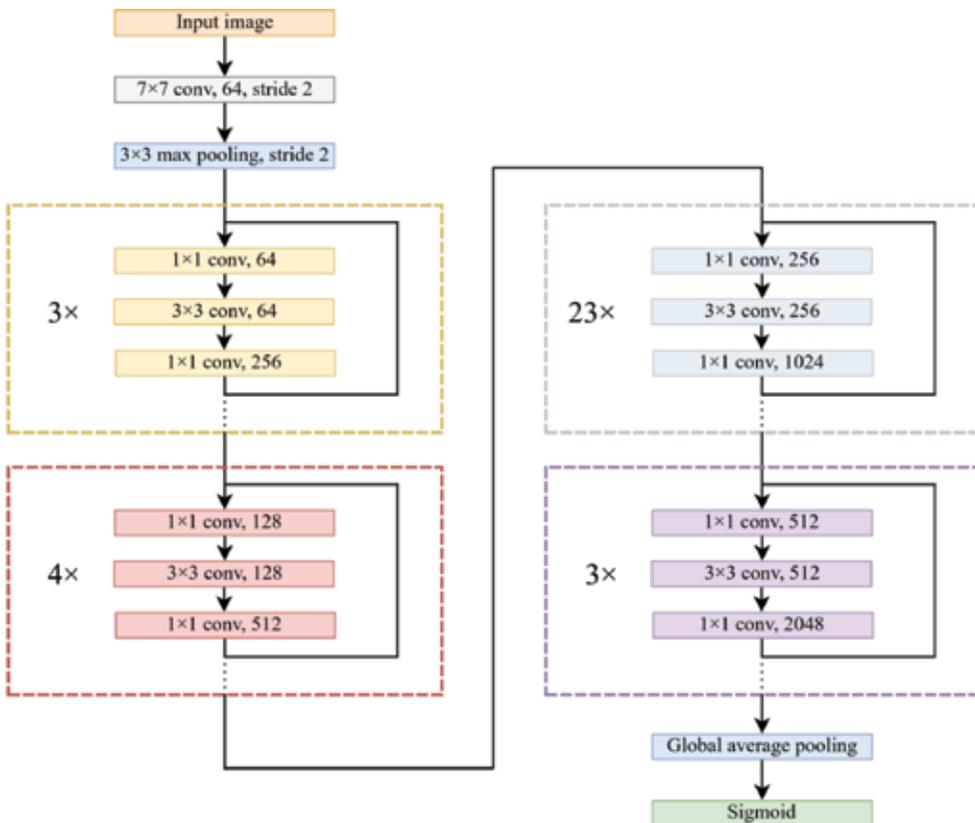


Fig. 1. ResNet-101 architecture.

Sl. 1. ResNet-101 arhitektura.

ResNet-50, ResNet-101, and ResNet-152 are the most often used ResNet architectures, with the number in the name indicating the number of layers. The ResNet-101 architecture is depicted in Fig. 1. The two initial layers of this architecture, including the 7×7 convolutional layer with 64 channels and the 3×3 max-pooling layer, are followed by the four groups of different numbers (3, 4, 23, and 3) of the described bottleneck residual blocks with 256, 512, 1024, and 2048 output channels, respectively. Finally, the global average pooling layer is placed before the fully connected and output layers.

3.2. Xception

The Xception [50] is a CNN architecture utilizing the previously-mentioned shortcut connections, and the depthwise separable convolutions [51] that consist of the point-

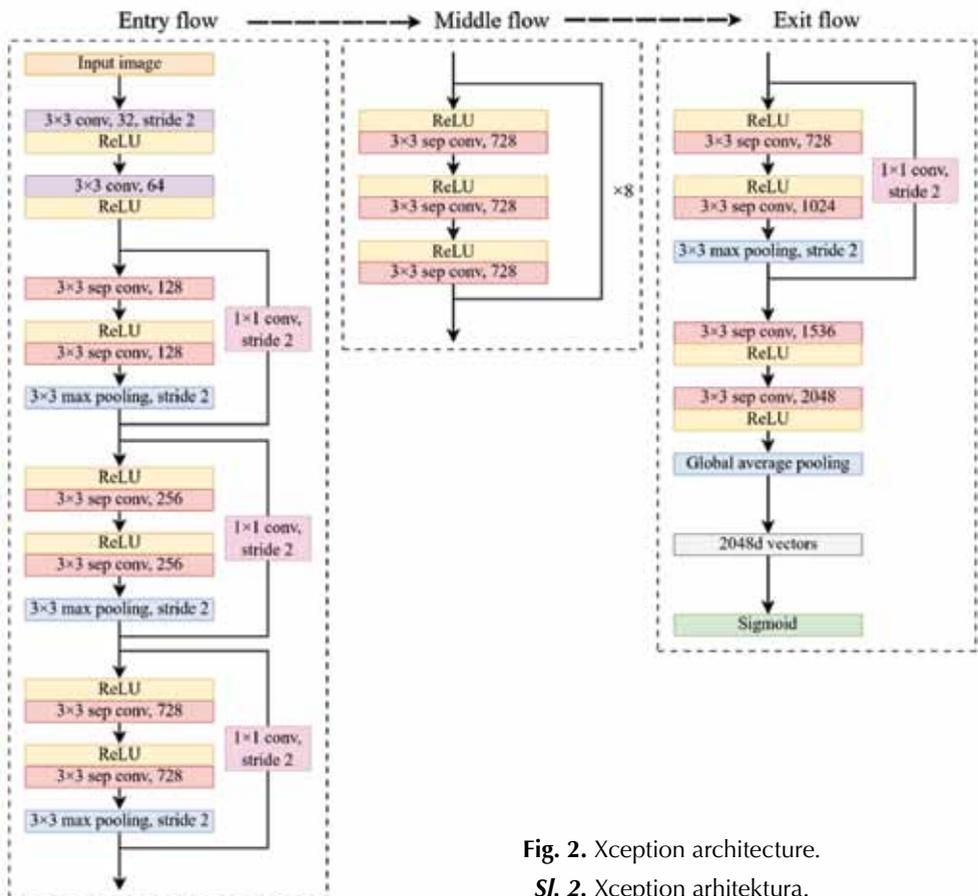


Fig. 2. Xception architecture.

Sl. 2. Xception arhitektura.

wise (1×1) and the depthwise convolution. The pointwise convolution maps the cross-channel correlations by projecting the input channels onto a new channel space, while the depthwise convolution maps the spatial correlations of the channels by performing a spatial convolution on each channel independently [50]. The depthwise separable convolutions significantly reduce computational costs compared to the standard convolutions [52] while simultaneously providing high accuracy [50].

The Xception architecture, consisting of the entry, middle, and exit flows, is illustrated by a schematic overview in Fig. 2. This architecture contains 36 convolutional layers divided into 14 modules, where all modules, except for the first and the last one, have linear residual connections [50]. The first two layers perform the standard, and the others the depthwise separable 3×3 convolutions, followed by the batch normalization [50, 53]. Additionally, 3×3 max-pooling layers with a stride of 2 are utilized too [50]. The output (classification) layer may optionally be preceded by a fully connected layer [50].

3.3. EfficientNet

The EfficientNet architecture utilizes the compound scaling procedure, in which all three CNN dimensions (including network width, depth, and resolution) are scaled simultaneously by fixed values calculated based on the compound coefficient specifying how much extra computational capacity is available [54]. The EfficientNet-B0 represents a novel base CNN architecture obtained by a multi-objective neural architecture search optimizing floating-point operations per second and accuracy [54, 55]. The architecture is characterized by the main building block called mobile inverted bottleneck convolution (MBConv) [52, 55] and squeeze-and-excitation optimization [56].

The MBConv was introduced in the MobileNetV2 CNN architecture developed for mobile platforms to reduce memory requirements [52]. This inverted residual block contains linear bottleneck layers and depthwise separable convolutions, i.e. the pointwise convolution first increases the number of channels of the input data, which is then followed by the depthwise convolution and another pointwise convolution that again reduces the number of channels to match the input [52]. Bottleneck layers are connected by shortcut connections, and do not utilize non-linear activations [52].

The squeeze-and-excitation module performs the feature map recalibration by suppressing those with less information and emphasizing the more informative ones [56]. The squeeze stage of this procedure includes performing global average pooling on each channel [56]. In the excitation stage, the obtained channel descriptors are fed to two fully connected layers containing the ReLU and sigmoid activation functions, respectively

[56]. Finally, the learned weights obtained at the output of these layers are utilized to rescale the input feature maps [56].

The EfficientNet-B0 CNN architecture is shown in Fig. 3. EfficientNet-B0 architecture. The initial standard 3×3 convolutions with 32 output channels are followed by one 3×3 MBConv1 module with 16 output channels, two 3×3 MBConv6 modules with 24 channels, two 5×5 MBConv6 modules with 40 channels, three 3×3 MBConv6 modules with 80 channels, three 5×5 MBConv6 modules with 112 channels, four 5×5 MBConv6 modules with 192 channels, and one 3×3 MBConv6 module with 320 channels [54]. Finally, the output layer is preceded by the standard 1×1 convolution with 1280 channels and the global average pooling layer [54].

EfficientNet-B1 to B7 architectures are obtained by scaling up the base EfficientNet-B0 CNN using different compound coefficient values [54]. These CNN architectures provide state-of-the-art classification performances while, at the same time, being significantly faster and having considerably fewer parameters and operations than the competitive architectures [54].

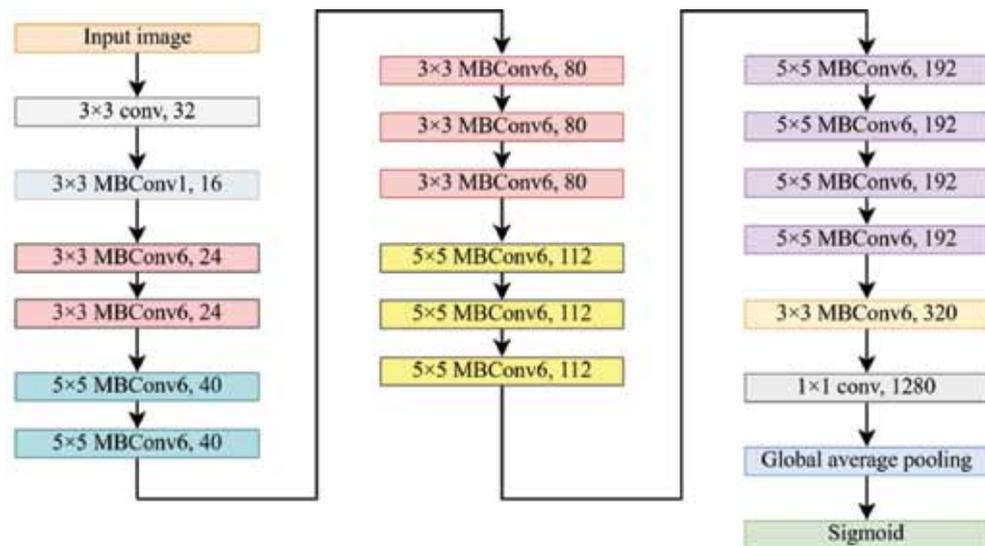


Fig. 3. EfficientNet-B0 architecture.

Sl. 3. EfficientNet-B0 arhitektura.

Next, we present a method combining the TFRs and deep learning models for non-stationary GW signals classification in intensive noise.

4. THE APPLICATION OF THE PROPOSED METHOD COMBINING THE TFRS AND MACHINE LEARNING TO THE GW DETECTION

4.1. An Introduction to the GW Detection

An example of combining the TFRs with the deep learning-based classification to non-stationary signals is gravitational-wave (GW) detection. The GWs are defined as ripples in the space-time curvature, which propagate at the speed of light and are caused by massive accelerating objects, such as coalescing black holes [57]. Although Albert Einstein predicted their existence in 1916 within the general theory of relativity [58], the GWs were not experimentally detected until 2015, when the Advanced Laser Interferometer Gravitational-Wave Observatory (LIGO) detectors [59] made the first direct observation of the compact binary coalescence GW signal originating from the binary black hole (BBH) merger [60]. This detection was immediately awarded the Nobel Prize in Physics in 2017, and it marked the beginning of intensive research in the GW data analysis.

Besides two Advanced LIGO detectors (Livingston, Louisiana, and Hanford, Washington) in the USA, the Advanced Virgo detector near Pisa, Italy, is run by the European Gravitational Observatory (EGO) [61, 62]. Moreover, two new detectors, including KAGRA in Japan and LIGO-India, are expected to join the observation network. The GW detectors are large Michelson interferometers with light-reflecting mirrors located at each end of the two orthogonal arms 4 km long in the Advanced LIGO detectors and 3 km long in the Advanced Virgo detector [59, 61, 63]. The space-time strain caused by the GW coming to the Earth shortens one arm of the detector and simultaneously lengthens the other. Thus produced difference in the arms' lengths is measured as the phase difference of the two reflected laser beams, and the optical signal collected by the output photodetector is proportional to the amplitude of the GW strain [60].

The extremely small magnitudes of the GWs require exceptional measurement sensitivity of the interferometers. However, despite state-of-the-art measurement and noise reduction equipment, various instrumental and environmental noise sources still significantly impact the acquired GW measurements. Some of the most common instrumental noise sources include electrical power system harmonics, noise in electronic components, thermal noise, quantum noise, and optical system noise [64]. On the other hand, some well-known environmental noise sources are seismic noise, vibration noise due to human activity, and electromagnetic noise [65, 66].

Therefore, the collected GW data are characterized by the noise that is non-stationary, non-white, and non-Gaussian [65]. Additionally, the magnitudes of the GWs are

much lower than the background noise, resulting in very low signal-to-noise ratio (SNR) values. Thus, detecting the GW signals is a complex and challenging task, requiring the development of specialized algorithms. There are three main categories of the existing approaches to the GW detection, including matched filtering [67, 68, 69], denoising methods [70, 71, 72], and machine learning-based detection [73].

Matched filtering-based search for a known signal consists of correlating the noisy GW data to a database of simulated waveform templates [74]. Although matched filtering-based methods are used today in LIGO detectors, they are computationally expensive, require large waveform databases, and provide optimal performance only for Gaussian noise [75]. Furthermore, denoising methods are applied to the GW data to reduce the noise and can be beneficial for initial data processing. However, these methods ought to be coupled with other detection algorithms. Finally, machine learning has recently received increased interest in detecting different types of the GWs [73]. Nevertheless, these approaches have been primarily focused on classifying time-series [76, 77, 78] or spectrogram representations, unlike our method, which uses the TFRs from Cohen's class [79, 80, 81].

4.2. The GW Detection from the TFRs Using Machine Learning

A method for detecting the GW signals in intensive noise based on utilizing quadratic TFRs from Cohen's class with 2D deep CNN architectures was proposed and developed by the authors of this paper in [82], and [83]. The flowchart of the proposed method and experimental setup is shown in Fig. 4. The procedure consisted of data preparation, deep learning models, and performance evaluation.

The data generation started with acquiring the real-life noise data from the Advanced LIGO detectors, which are used as background noise in the generated data examples. Using the actual LIGO noise containing detector glitches allowed the generation of the realistic GW dataset. The data were collected from the repositories at the Gravitational Wave Open Science Center (GWOSC) [84] for the O2 observing run [85], checked for the specific quality requirements, and downsampled to 2048 Hz.

The extensive simulations of the GW waveforms originating from the BBH mergers were performed in parallel with LIGO data extraction, utilizing software packages LIGO Algorithm Library (LALSuite) [86], and PyCBC [69, 87]. The simulations were based on the effective-one-body waveform model SEOBNRv4 in the time domain [88], for which a parameter space was defined. These parameters were randomly chosen for each iteration of waveform simulations. The simulated waveforms were then multiplied by the one-sided Tukey window and projected onto the antenna patterns characteristic for the Advanced LIGO detectors, thus obtaining noise-free GW signals.

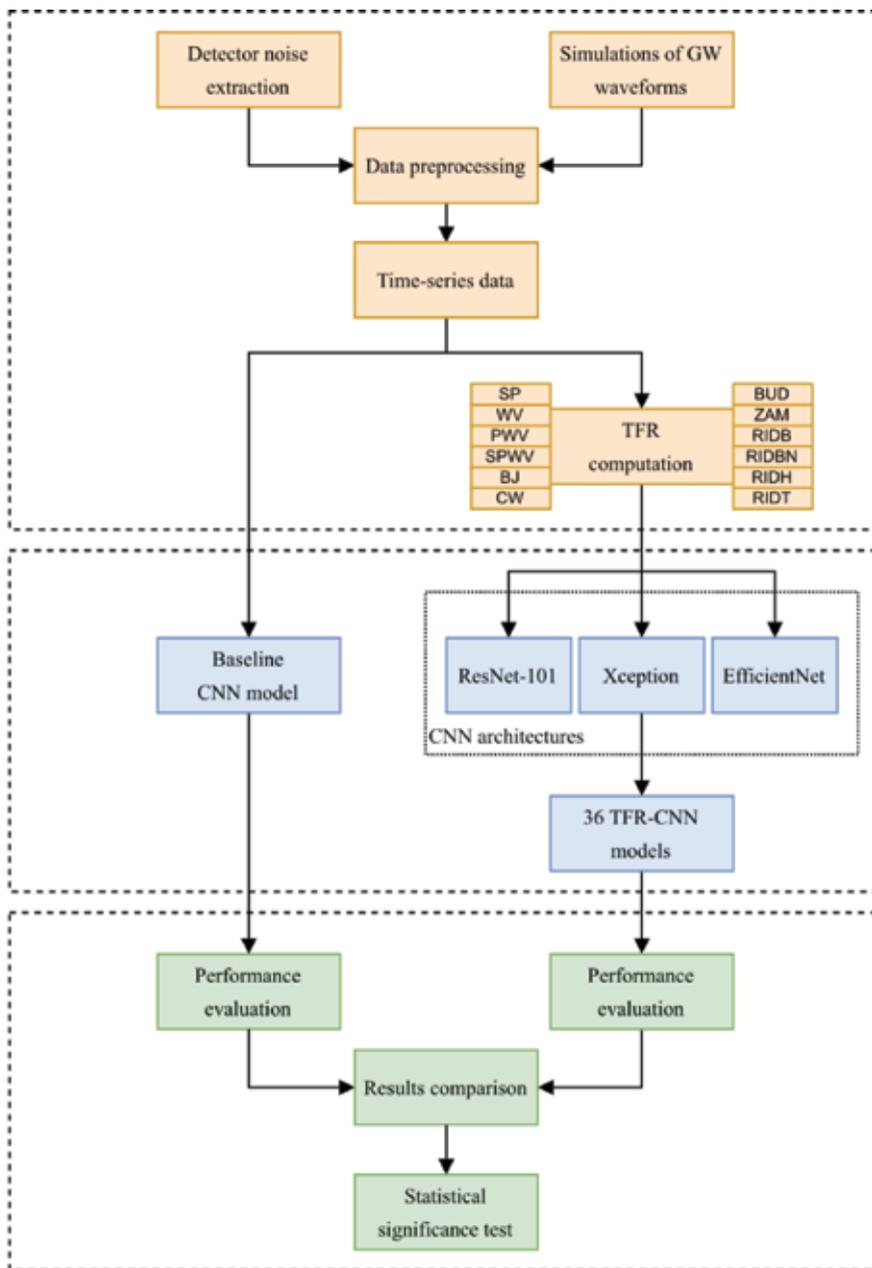


Fig. 4. Experimental setup flowchart.

Sl. 4. Dijagram toka eksperimentalnog postava.

These signals were then injected into the background noise after being scaled to obtain the desired injection SNR, i.e. the desired network optimal matched-filter SNR (NOMF-SNR) [89, 90]. The NOMF-SNR was in the range of $8 - 30$ dB, which corresponded to the values ranging from -123.46 to -2.27 dB according to the standard SNR definition for the raw, noisy time series.

The obtained time-series data examples were then preprocessed by applying the whitening procedure in the frequency domain [91] based on the local estimate of the noise amplitude spectral density (ASD) obtained by the Welch method [92], which was followed by high-pass filtering at 20 Hz. The data examples were then cropped to 0.5 s length, sufficient to observe characteristic BBH chirp waveforms. Finally, $100\,000$ time-series data examples were generated, with one half containing the GW signals in the background noise, and the other half containing only noise.

For illustration purposes, Fig. 5. An example of the time series containing only noise: (a) Raw noise; (b) Preprocessed noise provides a randomly chosen example of the generated time series containing only noise, where the raw data and the data after preprocessing, i.e. whitening and high-pass filtering, are shown. Additionally, Fig. 6 gives an example of the time series containing a GW signal, with the NOMF-SNR of 19 dB, corresponding to the SNR of -77.76 dB, where the raw noise, the noise-free simulated GW signal, and the preprocessed noisy GW signal are shown. As seen in Fig. 6, the amplitudes of the GW signals are much lower than the noise amplitudes, making it very difficult to visually distinguish between the time-series examples with only noise and those containing GW signals.

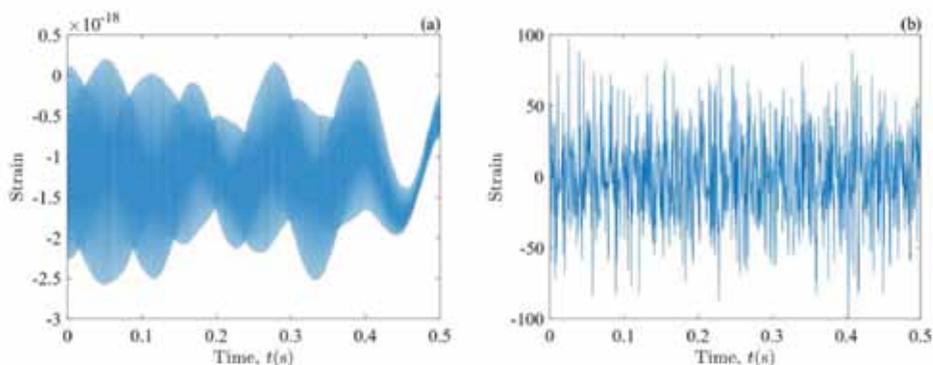


Fig. 5. An example of the time series containing only noise: (a) Raw noise; (b) Preprocessed noise.

Sl. 5. Primjer vremenskog niza koji sadrži samo šum: (a) Neobrađeni šum; (b) Šum nakon predobrade.

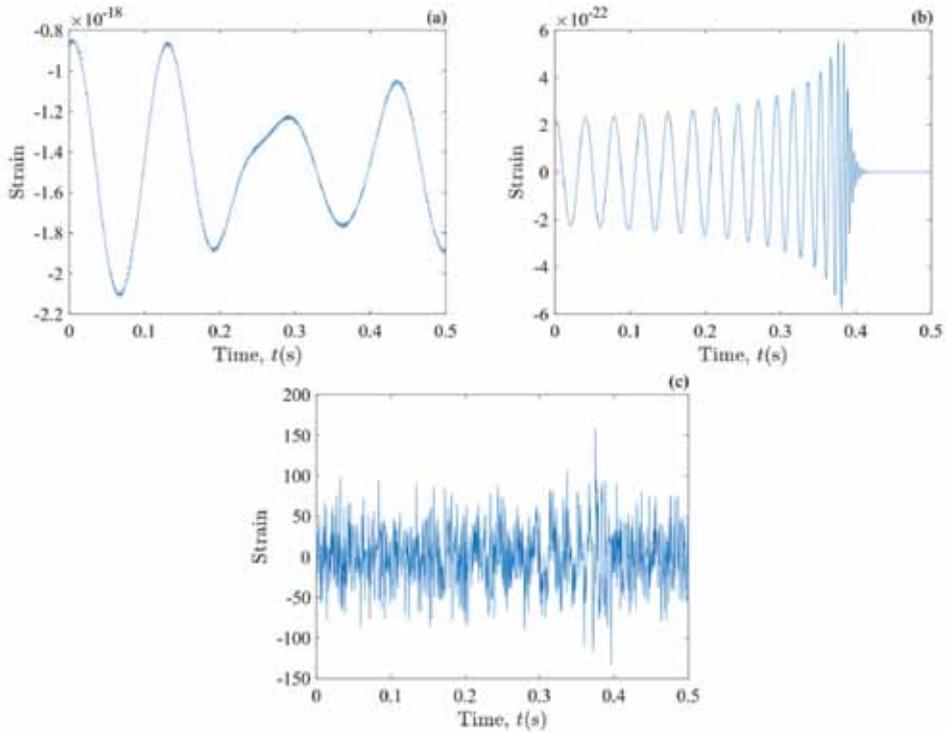


Fig. 6. An example of the time series containing the GW signal in the noise (NOMF-SNR = 19 dB, SNR = -77.76 dB): (a) Raw noise; (b) GW signal without noise; (c) Preprocessed noisy GW signal.

Sl. 6. Primjer vremenskog niza koji sadrži GW signal u šumu (NOMF-SNR = 19 dB, SNR = -77,76 dB): (a) Neobrađeni šum; (b) GW signal bez šuma; (c) GW signal u šumu nakon predobrade.

In the next phase, the generated time series were used to obtain 12 previously described TFRs from Cohen's class, including SP, WV, PWV, SPWV, BJ, CW, BUD, ZAM, RIDB, RIDBN, RIDH, and RIDT. Prior to the TFR computation, analytic signals had been obtained by utilizing the Hilbert transform. In this way, 12 different TFR datasets were obtained, each with 100 000 TFRs, where each TFR contained the 0 – 0.5 s time range and the 0 – 1024 Hz frequency range. Fig. 7 provides the TFRs obtained from the whitened noise-only time series shown in Fig. 5. An example of the time series containing only noise: (a) Raw noise; (b) Preprocessed noise. On the other hand, the computed TFRs of the whitened time series containing the GW signal, whose plots are given in Fig. 6, are shown in Fig. 8. These figures show that the GW-generated chirp pattern can be more easily spotted in the TFRs than in the time-series representations.

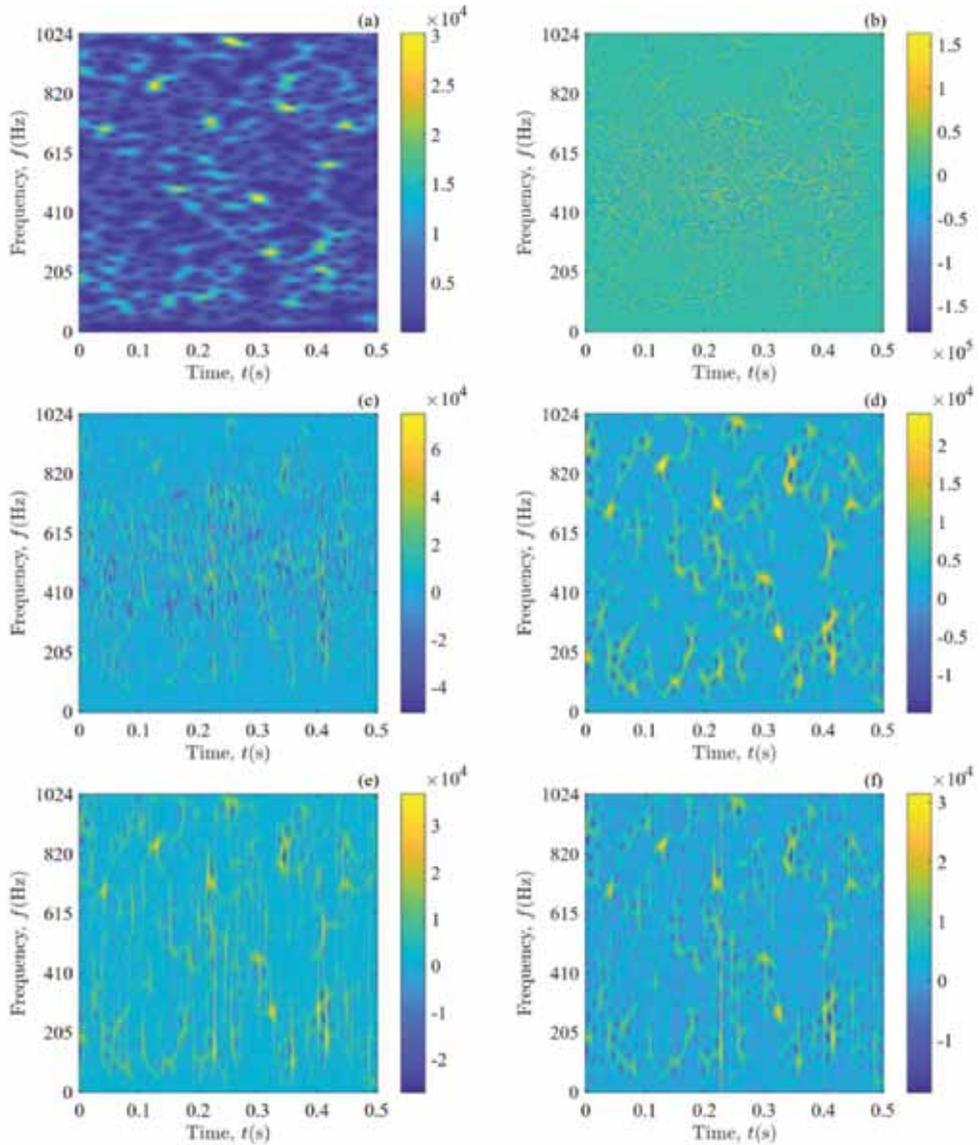


Fig. 7. The TFRs of the time-series example containing only noise: (a) SP; (b) WV; (c) PWV; (d) SPWV; (e) BJ; (f) CW; (g) BUD; (h) ZAM; (i) RIDB; (j) RIDBN; (k) RIDH; (l) RIDT.

Sl. 7. TFR-ovi primjera vremenskog niza koji sadrži samo šum: (a) SP; (b) WV; (c) PWV; (d) SPWV; (e) BJ; (f) CW; (g) BUD; (h) ZAM; (i) RIDB; (j) RIDBN; (k) RIDH; (l) RIDT.

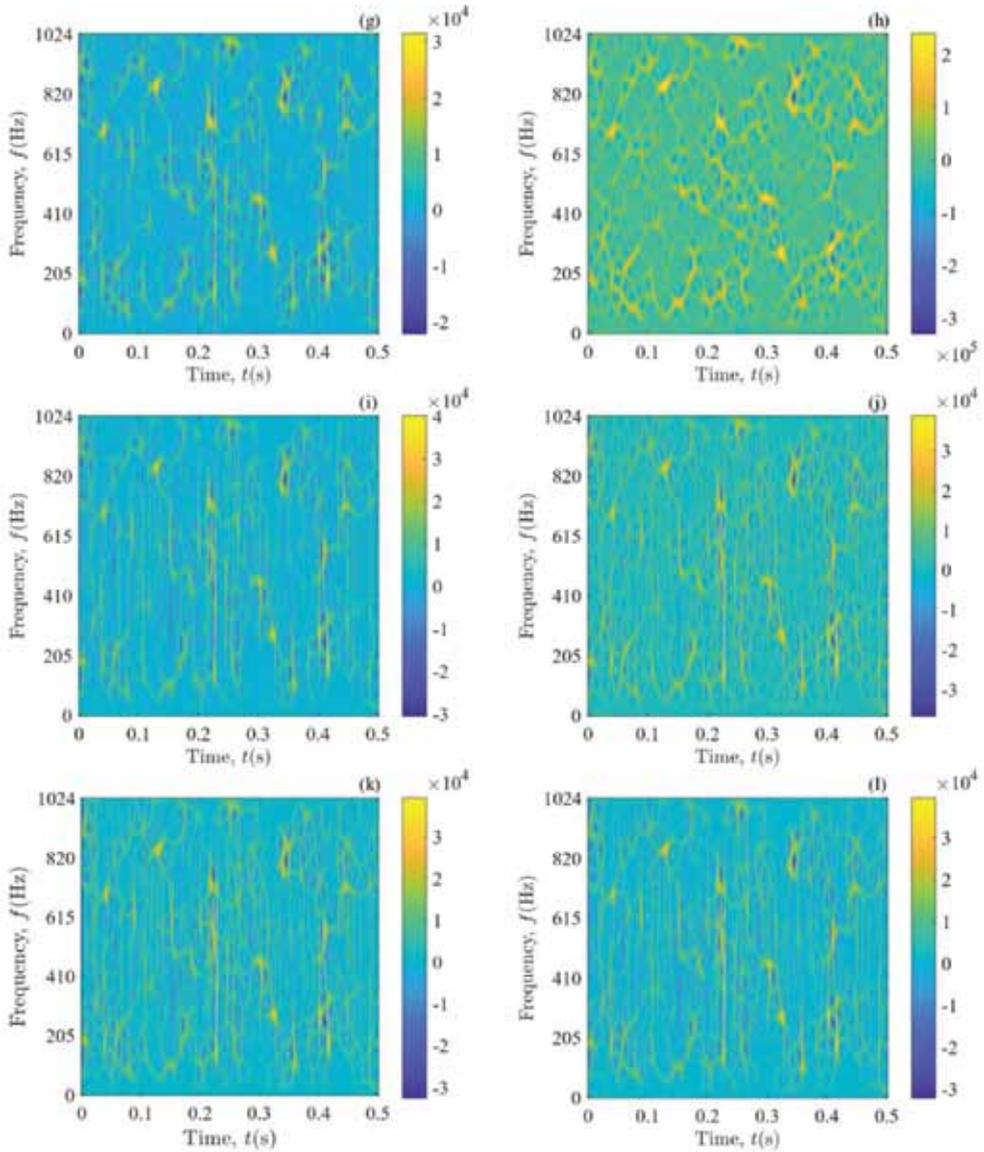


Fig. 7. (cont.) The TFRs of the time-series example containing only noise: (a) SP; (b) WV; (c) PWV; (d) SPWV; (e) BJ; (f) CW; (g) BUD; (h) ZAM; (i) RIDB; (j) RIDBN; (k) RIDH; (l) RIDT.

SI. 7. (nast.) TFR-ovi primjera vremenskog niza koji sadrži samo šum: (a) SP; (b) WV; (c) PWV; (d) SPWV; (e) BJ; (f) CW; (g) BUD; (h) ZAM; (i) RIDB; (j) RIDBN; (k) RIDH; (l) RIDT.

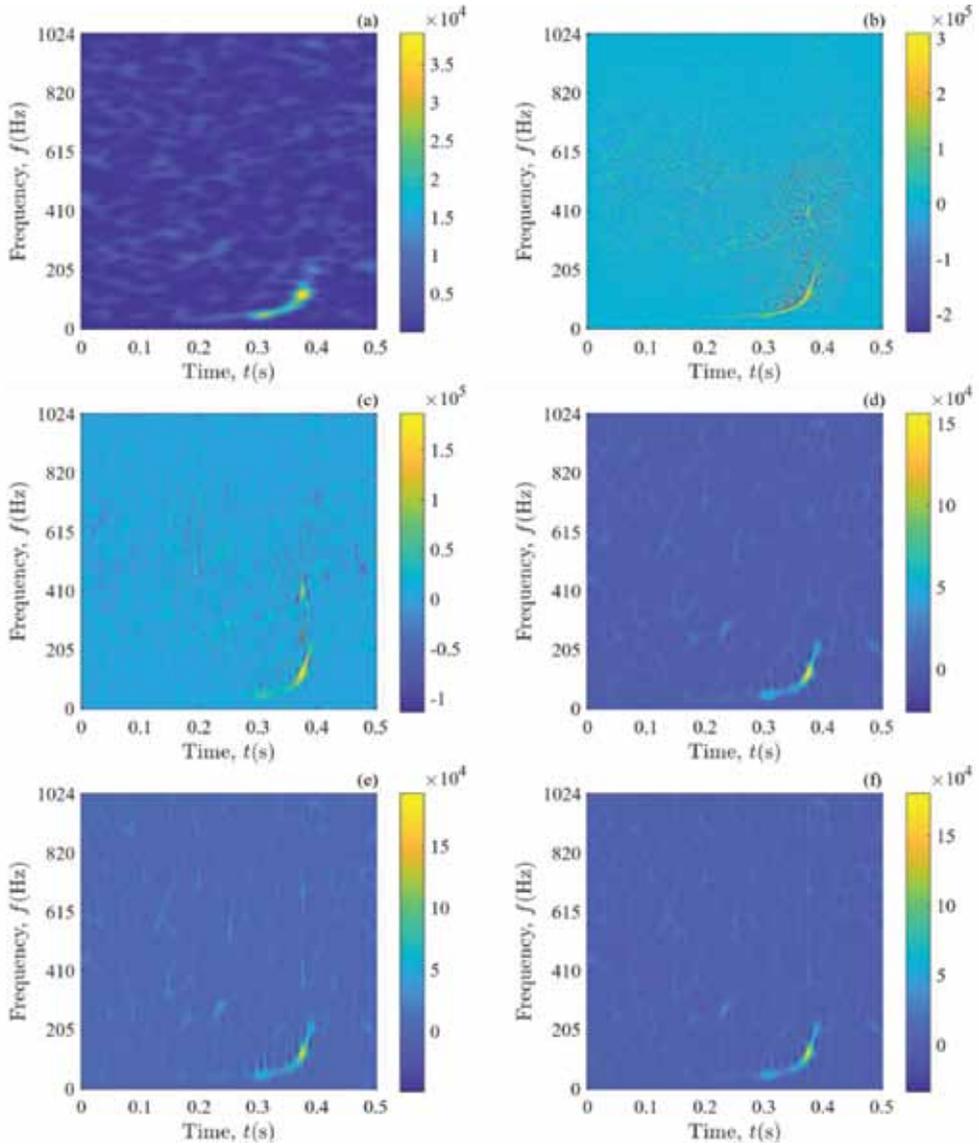


Fig. 8. The TFRs of the time-series example containing the GW signal in the noise (NOMF-SNR = 19 dB, SNR = -77.76 dB): (a) SP; (b) WV; (c) PWV; (d) SPWV; (e) BJ; (f) CW; (g) BUD; (h) ZAM; (i) RIDB; (j) RIDBN; (k) RIDH; (l) RIDT.

Sl. 8. TFR-ovi primjera vremenskog niza koji sadrži GW signal u šumu (NOMF-SNR = 19 dB, SNR = -77,76 dB): (a) SP; (b) WV; (c) PWV; (d) SPWV; (e) BJ; (f) CW; (g) BUD; (h) ZAM; (i) RIDB; (j) RIDBN; (k) RIDH; (l) RIDT.

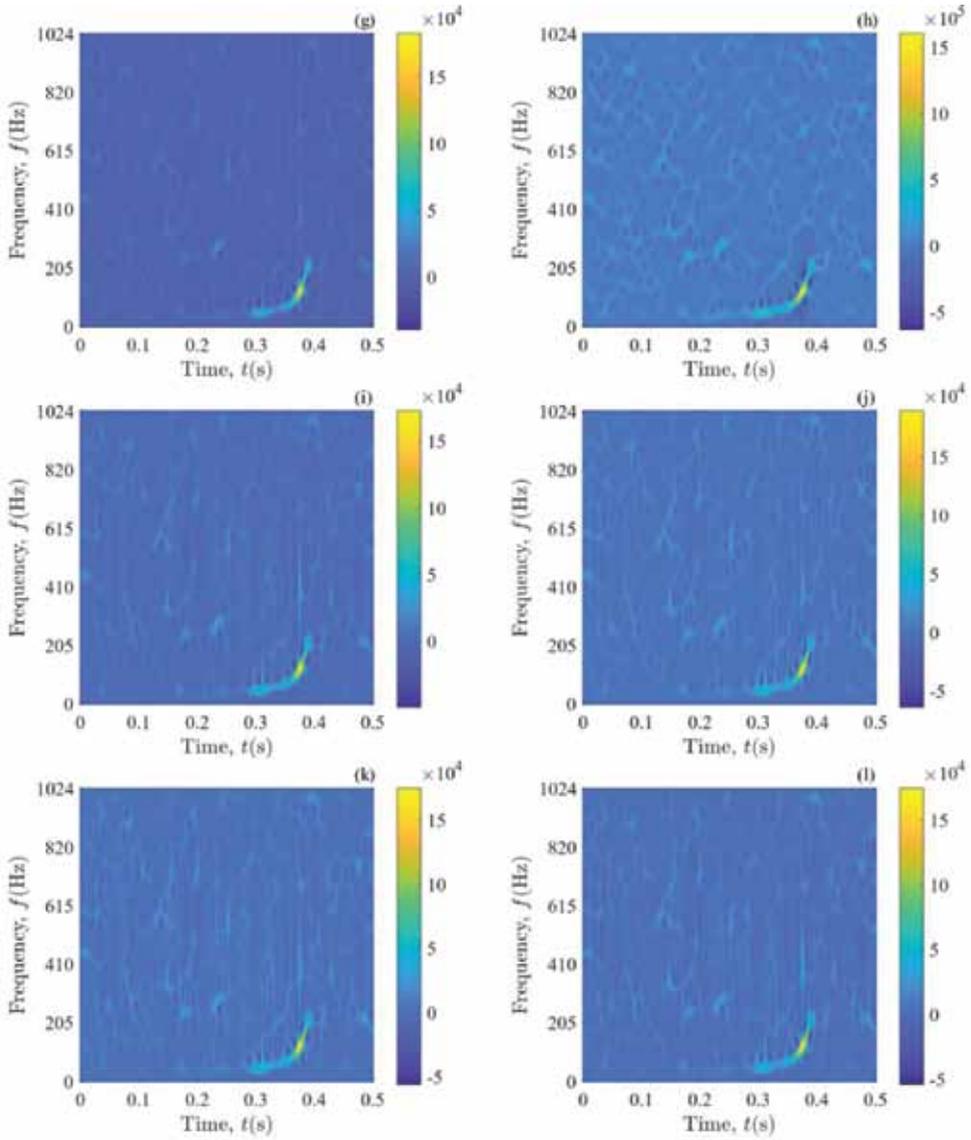


Fig. 8. (cont.) The TFRs of the time-series example containing the GW signal in the noise (NOMF-SNR = 9 dB, SNR = -77.76 dB): (a) SP; (b) WV; (c) PWV; (d) SPWV; (e) BJ; (f) CW; (g) BUD; (h) ZAM; (i) RIDB; (j) RIDBN; (k) RIDH; (l) RIDT.

Sl. 8. (nast.) TFR-ovi primjera vremenskog niza koji sadži GW signal u šumu (NOMF-SNR = 19 dB, SNR = -77,76 dB): (a) SP; (b) WV; (c) PWV; (d) SPWV; (e) BJ; (f) CW; (g) BUD; (h) ZAM; (i) RIDB; (j) RIDBN; (k) RIDH; (l) RIDT.

The obtained TFRs were normalized and stored as 256×256 grayscale PNG images. Each TFR dataset was then divided into the training, validation, and test subset, following the 70/15/15% ratio. Next, the 12 datasets of the TFR images (a total of 1.2 million images) were used as input to the deep learning classification algorithms based on the three 2D CNN architectures described in Section 3 (ResNet-101, Xception, and EfficientNet). Thus, 36 different TFR-CNN models were obtained. The overview of the classification procedure providing the probability that the input TFR contains the GW signal is depicted in Fig. 9.

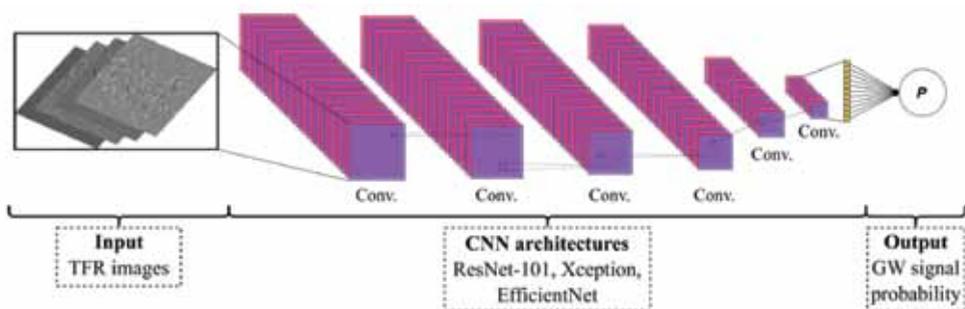


Fig. 9. Deep CNN-based classification of the TFRs of the GW time series.

Sl. 9. Klasifikacija TFR-ova GW vremenskog niza zasnovana na dubokim CNN.

The deep learning models were trained using the binary cross-entropy loss function and choosing optimal training parameters, including learning rate, optimizer algorithm, and batch size. Moreover, all models were trained from scratch, not relying on transfer learning techniques. Finally, the performance of the trained TFR-CNN models was evaluated on the test dataset using several evaluation metrics. Additionally, their performance was also compared to the baseline model trained on the original time series. For this purpose, an adapted 1D deep CNN model presented in [76] was used as a state-of-the-art technique for detecting the GWs in noisy time series.

The study in [82, 83] reported very high classification performances of the proposed method, with the values of the evaluation metrics in the following ranges: 96.540 – 97.100% for the classification accuracy, 0.98505 – 0.98854 for the area under the receiver operating characteristic (ROC) curve (ROC AUC), 97.549 – 99.507% for the precision, 94.147 – 95.867% for the recall, 96.459 – 97.029% for the F1 score, and 0.98989 – 0.99195 for the area under the precision-recall curve (PR AUC). Moreover, the proposed method outper-

formed the baseline model by 3.393 – 3.953%, 1.718 – 2.067%, 0.349 – 2.307%, 5.294 – 7.014%, 3.620 – 4.190%, and 1.269 – 1.475% in terms of classification accuracy, ROC AUC, precision, recall, F1 score, and PR AUC, respectively. Furthermore, the excellent classification performance of the developed approach was also supported by additional, detailed metrics, including confusion matrices, ROC curves, and precision-recall curves. Finally, the statistical significance of the obtained results was confirmed by McNemar’s statistical significance test.

5. CONCLUSION

This paper presents a method for detecting low-intensity signals in intensive noise by combining the deep learning methods and the TFRs from Cohen’s class. The method was demonstrated on a challenging task of the GW detection. The experimental setup consists of data preparation, training and testing of deep learning models, and performance evaluation. Real-life recordings were retrieved from LIGO detectors with and without the GW injections, allowing for testing the method in intensive real-life, non-stationary, non-Gaussian, and non-white noisy scenarios, with the SNR ranging between -123.46 and -2.27 dB. Following the preprocessing procedure, 12 TFRs from Cohen’s class were calculated (SP, WV, PWV, SPWV, BJ, CW, BUD, ZAM, RIDB, RIDBN, RIDH, and RIDT), resulting in 1.2 million TFR images. These were then used as input to the three deep learning algorithms (ResNet-101, Xception, and EfficientNet). Each of the 36 obtained TFR-CNN models was trained and evaluated, delivering an excellent classification performance of the proposed technique, with the classification accuracy up to 97.100%, ROC AUC up to 0.98854, recall up to 95.867%, precision up to 99.507%, F1 score up to 97.029%, and PR AUC up to 0.99195. Furthermore, the proposed technique outperformed the baseline model by up to 3.953%, up to 2.067%, up to 7.014%, up to 2.307%, up to 4.190%, and up to 1.475% in terms of classification accuracy, ROC AUC, recall, precision, F1 score, and PR AUC, respectively. McNemar’s statistical test confirmed the statistical significance of these results. The achieved high performances suggest that the proposed technique is not limited to the GW data only, but that it can easily be extended to other problems dealing with non-stationary signal classification.

REFERENCES

1. P. Gonçalves, P. Flandrin and E. Chassande-Mottin, *Time-Frequency Methods in Time Series Data Analysis*, in *Proceedings of Gravitational Wave Data Analysis Workshop 2*, Paris, 1997.
2. K. Gröchenig, *Foundations of Time-Frequency Analysis*, Birkhauser Verlag, 2013.
3. L. Cohen, *Time-Frequency Analysis: Theory and Applications*, USA: Prentice-Hall, Inc., 1995, p. 320.
4. B. Boashash, *Time frequency signal analysis and processing: a comprehensive reference*, Amsterdam Boston: Academic Press, 2016, p. 1056.
5. M. H. Ackroyd, *Short-Time Spectra and Time-Frequency Energy Distributions*, *The Journal of the Acoustical Society of America*, vol. 50, pp. 1229-1231, November 1971.
6. J. Allen, *Short term spectral analysis, synthesis, and modification by discrete Fourier transform*, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 25, pp. 235-238, June 1977.
7. H. G. Feichtinger and T. Strohmer, Eds., *Gabor Analysis and Algorithms Theory and Applications*, Birkhäuser Boston, 1998.
8. E. Berge, S. M. Berge and F. Luef, *The Affine Wigner Distribution*, 2019.
9. D. Mika and J. Józwiak, *Advanced Time-Frequency Representation in Voice Signal Analysis*, *Advances in Science and Technology Research Journal*, vol. 12, pp. 251–259, March 2018.
10. J. Lerga and V. Sucic, *An Instantaneous Frequency Estimation Method Based on the Improved Sliding Pair-Wise ICI Rule*, in *10th International Conference on Information Science, Signal Processing and their Applications (ISSPA 2010)*, 2010.
11. V. Sucic, J. Lerga and B. Boashash, *Multicomponent Noisy Signal Adaptive Instantaneous Frequency Estimation Using Components Time Support Information*, *IET Signal Processing*, vol. 8, pp. 277-284, 2014.
12. J. Lerga, V. Sucic and B. Boashash, *An Efficient Algorithm for Instantaneous Frequency Estimation of Nonstationary Multicomponent Signals in Low SNR*, *EURASIP Journal on Advances in Signal Processing*, vol. 2011, 2011.
13. J. Lerga and V. Sucic, *Nonlinear IF Estimation Based on the Pseudo WVD Adapted Using the Improved Sliding Pairwise ICI Rule*, *IEEE Signal Processing Letters*, vol. 16, pp. 953-956, 2009.
14. M. Orini, P. Laguna, L. T. Mainardi and R. Bailón, *Time-Frequency Analysis of Cardiovascular Signals and Their Dynamic Interactions*, in *Complexity and Nonlinearity in Cardiovascular Signals*, 2017.

15. Y. Zeevi and R. Coifmana, *Wavelet Analysis and its Applications*, in *Signal and Image Representation in Combined Spaces*, vol. 7, Y. Zeevi and R. Coifmana, Eds., Academic Press, 1998, p. 585.
16. A. Volpato and E. Collini, *Time-frequency methods for coherent spectroscopy*, *Optics Express*, vol. 23, pp. 20040-20050, July 2015.
17. C. Hurley and J. Mclean, *Wavelet: Analysis and Methods*, ETP, 2018.
18. A. Volpato and E. Collini, *Optimization and selection of time-frequency transforms for wave-packet analysis in ultrafast spectroscopy*, *Optics Express*, vol. 27, pp. 2975-2987, February 2019.
19. P. Flandrin, F. Auger and E. Chassande-Mottin, *Time-Frequency reassignment: from principles to algorithms*, in *Applications in Time-Frequency Signal Processing*, vol. 10, A. Papandreou-Suppappola, Ed., CRC Press, 2002, pp. 179-204.
20. I. Shafi, A. Jamil, S. Shah and F. Kashif, *Techniques to Obtain Good Resolution and Concentrated Time-Frequency Distributions: A Review*, *EURASIP Journal on Advances in Signal Processing*, vol. 2009, p. 43, June 2009.
21. P. Flandrin, *Time-Frequency/Time-Scale Analysis*, Volume 10, 1st ed., USA: Academic Press, Inc., 1998, p. 386.
22. F. Hlawatsch and F. Auger, *Time-frequency analysis*, John Wiley & Sons, 2013, p. 472.
23. F. Auger and E. Chassande-Mottin, *Quadratic Time-Frequency Analysis I: Cohen's Class*, ISTE, 2008, pp. 131-163.
24. H. Margenau and R. N. Hill, *Correlation between Measurements in Quantum Theory*, *Progress of Theoretical Physics*, vol. 26, pp. 722-738, November 1961.
25. B. Boashash, *Note on the use of the Wigner distribution for time-frequency signal analysis*, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, pp. 1518-1521, September 1988.
26. J. Ville, *Theorie et applications de la notion de signal analytique*, *Câbles et Transmissions*, 2A, 1948.
27. T. A. C. M. Claasen and W. Mecklenbräuker, *The Wigner distribution—A tool for time-frequency signal analysis*, *Philips JI Research*, vol. 35, pp. 276-300, January 1980.
28. P. Flandrin and B. Escudié, *An interpretation of the Pseudo-Wigner-Ville distribution*, *Signal Processing*, vol. 6, pp. 27-36, January 1984.
29. P. Flandrin, *Some features of time-frequency representations of multicomponent signals*, in *ICASSP '84. IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1984.

30. E. Cordero, M. de Gosson, M. Dörfler and F. Nicola, *Generalized Born-Jordan Distributions and Applications*, *Adv. Comput. Math.*, vol. 46, June 2020.
31. E. Cordero, M. de Gosson and F. Nicola, *On the reduction of the interferences in the Born–Jordan distribution*, *Applied and Computational Harmonic Analysis*, vol. 44, pp. 230-245, March 2018.
32. L. Cohen, *Generalized Phase-Space Distribution Functions*, *Journal of Mathematical Physics*, vol. 7, pp. 781-786, May 1966.
33. J. Jeong and W. J. Williams, *Kernel design for reduced interference distributions*, *IEEE Transactions on Signal Processing*, vol. 40, pp. 402-412, February 1992.
34. F. Auger, P. Flandrin, P. Gonçalves and O. Lemoine, *Time-frequency toolbox tutorial*, *CNRS (France), Rice U.(USA)*, January 2005.
35. H.-I. Choi and W. J. Williams, *Improved time-frequency representation of multicomponent signals using exponential kernels*, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, pp. 862-871, June 1989.
36. F. Hlawatsch, T. G. Manickam, R. L. Urbanke and W. Jones, *Smoothed pseudo-Wigner distribution, Choi-Williams distribution, and cone-kernel representation: Ambiguity-domain analysis and experimental comparison*, *Signal Processing*, vol. 43, pp. 149-168, May 1995.
37. A. Papandreou and G. F. Boudreaux-Bartels, *Distributions for time-frequency analysis: a generalization of Choi-Williams and the Butterworth distribution*, in [*Proceedings*] *ICASSP-92: 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1992.
38. E. Sejdić, I. Djurović and J. Jiang, *Time–frequency feature representation using energy concentration: An overview of recent advances*, *Digital Signal Processing*, vol. 19, pp. 153-183, January 2009.
39. R. V. Mendes, *Non-commutative tomography and signal processing*, *Physica Scripta*, vol. 90, p. 074022, June 2015.
40. A. Papandreou and G. F. Boudreaux-Bertels, *Generalization of the Choi-Williams Distribution and the Butterworth Distribution for Time-Frequency Analysis*, *IEEE Transactions on Signal Processing*, vol. 41, pp. 463-472, January 1993.
41. D. Wu and J. M. Morris, *Time-frequency representations using a radial Butterworth kernel*, in *Proceedings of IEEE-SP International Symposium on Time- Frequency and Time-Scale Analysis*, 1994.
42. Y. Zhao, L. E. Atlas and R. J. Marks, *The use of cone-shaped kernels for generalized time-frequency representations of nonstationary signals*, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, pp. 1084-1091, July 1990.

43. L. Stankovic, M. Dakovic and T. Thayaparan, *Time-frequency signal analysis with applications*, Artech House, 2013.
44. Z. Guo, L.-G. Durand and H. C. Lee, *The time-frequency distributions of nonstationary signals based on a Bessel kernel*, *IEEE Transactions on Signal Processing*, vol. 42, pp. 1700-1707, July 1994.
45. W. J. Williams and J. Jeong, *Reduced interference time-frequency distributions*, *Time-Frequency Signal Analysis: Methods and Applications*, pp. 74-98, 1992.
46. M. I. Jordan and T. M. Mitchell, *Machine learning: Trends, perspectives, and prospects*, *Science*, vol. 349, pp. 255-260, July 2015.
47. I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*, Cambridge, MA: MIT Press, 2016.
48. Y. LeCun, Y. Bengio and G. Hinton, *Deep learning*, *Nature*, vol. 521, pp. 436-444, May 2015.
49. K. He, X. Zhang, S. Ren and J. Sun, *Deep Residual Learning for Image Recognition*, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016)*, Las Vegas, NV, USA, 2016.
50. F. Chollet, *Xception: Deep Learning with Depthwise Separable Convolutions*, in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017)*, Honolulu, 2017.
51. L. Sifre, *Rigid-Motion Scattering For Image Classification*, Paris, 2014.
52. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L.-C. Chen, *MobileNetV2: Inverted Residuals and Linear Bottlenecks*, in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2018)*, Salt Lake City, UT, USA, 2018.
53. S. Ioffe and C. Szegedy, *Batch normalization: Accelerating deep network training by reducing internal covariate shift*, in *32nd International Conference on Machine Learning (ICML 2015)*, Lille, 2015.
54. M. Tan and Q. Le, *EfficientNet: Rethinking model scaling for convolutional neural networks*, in *36th International Conference on Machine Learning (ICML 2019)*, Long Beach, CA, USA, 2019.
55. M. Tan and others, *MnasNet: Platform-Aware Neural Architecture Search for Mobile*, in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2019)*, Long Beach, CA, USA, 2019.
56. J. Hu, L. Shen and G. Sun, *Squeeze-and-Excitation Networks*, in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2018)*, Salt Lake City, UT, USA, 2018.
57. M. Maggiore, *Gravitational Waves: Volume 2: Astrophysics and Cosmology*, 1 ed., Oxford, UK: Oxford University Press, 2018.

58. A. Einstein, *Näherungsweise Integration der Feldgleichungen der Gravitation*, *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften zu Berlin*, pp. 99-108, 1916.
59. J. Aasi and others, *Advanced LIGO*, *Classical and Quantum Gravity*, vol. 32, March 2015.
60. B. P. Abbott and others, *Observation of gravitational waves from a binary black hole merger*, *Physical Review Letters*, vol. 116, February 2016.
61. F. Acernese and others, *Advanced Virgo: a second-generation interferometric gravitational wave detector*, *Classical and Quantum Gravity*, vol. 32, December 2014.
62. F. Acernese and others, *Advanced Virgo Status*, in *Journal of Physics: Conference Series*, 2020.
63. LIGO Caltech, *LIGO's Interferometer*.
64. D. V. Martynov and others, *Sensitivity of the Advanced LIGO detectors at the beginning of gravitational wave astronomy*, *Physical Review D*, vol. 93, no. 11, June 2016.
65. B. P. Abbott and others, *Characterization of transient noise in Advanced LIGO relevant to gravitational wave signal GW150914*, *Classical and Quantum Gravity*, vol. 33, June 2016.
66. A. Effler and others, *Environmental influences on the LIGO gravitational wave detectors during the 6th science run*, *Classical and Quantum Gravity*, vol. 32, January 2015.
67. K. Cannon and others, *GstLAL: A software framework for gravitational wave discovery*, *SoftwareX*, vol. 14, June 2021.
68. A. H. Nitz, T. Dent, T. D. Canton, S. Fairhurst and D. A. Brown, *Detecting Binary Compact-object Mergers with Gravitational Waves: Understanding and Improving the Sensitivity of the PyCBC Search*, *The Astrophysical Journal*, vol. 849, November 2017.
69. S. A. Usman and others, *The PyCBC search for gravitational waves from compact binary coalescence*, *Classical and Quantum Gravity*, vol. 33, October 2016.
70. A. Torres-Forné, E. Cuoco, A. Marquina, J. A. Font and J. M. Ibáñez, *Total-variation methods for gravitational-wave denoising: Performance tests on Advanced LIGO data*, *Physical Review D*, vol. 98, October 2018.
71. N. Lopac, J. Lerga and E. Cuoco, *Gravitational-Wave Burst Signals Denoising Based on the Adaptive Modification of the Intersection of Confidence Intervals Rule*, *Sensors*, vol. 20, December 2020.
72. N. Lopac, J. Lerga, N. Saulig, L. Stanković and M. Daković, *On Optimal Parameters for ICI-Based Adaptive Filtering Applied to the GWs in High Noise*, in *2021 6th*

- International Conference on Smart and Sustainable Technologies (SpliTech 2021)*, Bol and Split, Croatia, 2021.
73. E. Cuoco and others, *Enhancing gravitational-wave science with machine learning, Machine Learning: Science and Technology*, vol. 2, December 2020.
 74. B. S. Sathyaprakash and B. F. Schutz, *Physics, astrophysics and cosmology with gravitational waves, Living Reviews in Relativity*, vol. 12, December 2009.
 75. A. H. Nitz, T. Dal Canton, D. Davis and S. Reyes, *Rapid detection of gravitational waves from compact binary mergers with PyCBC Live, Physical Review D*, vol. 98, no. 2, July 2018.
 76. D. George and E. A. Huerta, *Deep learning for real-time gravitational wave detection and parameter estimation: Results with Advanced LIGO data, Physics Letters B*, vol. 778, pp. 64-70, March 2018.
 77. H. Gabbard, M. Williams, F. Hayes and C. Messenger, *Matching Matched Filtering with Deep Networks for Gravitational-Wave Astronomy, Physical Review Letters*, vol. 120, April 2018.
 78. T. D. Gebhard, N. Kilbertus, I. Harry and B. Schölkopf, *Convolutional neural networks: A magic bullet for gravitational-wave detection?, Physical Review D*, vol. 100, September 2019.
 79. A. Menéndez-Vázquez, M. Kolstein, M. Martínez and L.-M. Mir, *Searches for compact binary coalescence events using neural networks in the LIGO/Virgo second observation period, Physical Review D*, vol. 103, March 2021.
 80. J. D. Álvares and others, *Exploring gravitational-wave detection and parameter inference using deep learning methods, Classical and Quantum Gravity*, vol. 38, July 2021.
 81. A. Iess, E. Cuoco, F. Morawski and J. Powell, *Core-Collapse supernova gravitational-wave search and deep learning classification, Machine Learning: Science and Technology*, vol. 1, May 2020.
 82. N. Lopac, F. Hržić, I. Petrijevčanin Vuksanović and J. Lerga, *Detection of Non-Stationary GW Signals in High Noise From Cohen's Class of Time-Frequency Representations Using Deep Learning, IEEE Access*, vol. 10, pp. 2408-2428, January 2022.
 83. N. Lopac, *Detection of Gravitational-Wave Signals From Time-Frequency Distributions Using Deep Learning*, 2022.
 84. Rich Abbott et al. (LIGO Scientific Collaboration and Virgo Collaboration), *Open data from the first and second observing runs of Advanced LIGO and Advanced Virgo, SoftwareX*, vol. 13, January 2021.

85. B. P. Abbott and others, *GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs*, *Physical Review X*, vol. 9, no. 3, September 2019.
86. LIGO Scientific Collaboration, *LIGO Algorithm Library – LALSuite*, 2018.
87. A. H. Nitz and others, *gwastro/pycbc: PyCBC Release v1.13.6*, Zenodo, 2019.
88. A. Bohé and others, *Improved effective-one-body model of spinning, nonprecessing binary black holes for the era of gravitational-wave astrophysics with advanced detectors*, *Physical Review D*, vol. 95, no. 4, February 2017.
89. B. Allen, W. G. Anderson, P. R. Brady, D. A. Brown and J. D. E. Creighton, *FIND-CHIRP: An algorithm for detection of gravitational waves from inspiraling compact binaries*, *Physical Review D*, vol. 85, June 2012.
90. C. Cutler and É. E. Flanagan, *Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral waveform?*, *Physical Review D*, vol. 49, pp. 2658-2697, March 1994.
91. B. P. Abbott and others, *A guide to LIGO–Virgo detector noise and extraction of transient gravitational-wave signals*, *Classical and Quantum Gravity*, vol. 37, February 2020.
92. P. Welch, *The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms*, *IEEE Transactions on Audio and Electroacoustics*, vol. 15, pp. 70-73, June 1967.

KOMBINACIJA VREMENSKO-FREKVENCIJSKE ANALIZE SIGNALA I STROJNOGA UČENJA UZ PRIMJER U DETEKCIJI GRAVITACIJSKIH VALOVA

Sažetak

Ovaj rad predstavlja metodu klasifikacije šumom narušenih nestacionarnih signala u vremensko-frekvencijskoj domeni korištenjem umjetne inteligencije. Naime, signali u obliku vremenskih nizova transformirani su nakon predobrade u vremensko-frekvencijske prikaze (TFR) iz Cohenove klase, rezultirajući TFR slikama korištenim kao ulaz u algoritme strojnoga učenja. Korištene su tri suvremene metode dubokoga učenja u obliku 2D arhitektura konvolucijskih neuronskih mreža (CNN) (ResNet-101, Xception i EfficientNet). Metoda je demonstrirana na zahtjevnom problemu detekcije signala gravitacijskih valova (GW) u intenzivnom stvarnom i nestacionarnom šumu koji nema karakteristike ni Gaussovog ni bijelog šuma. Rezultati pokazuju izvrsne performanse klasifikacije predloženoga pristupa s obzirom na točnost klasifikacije, površinu ispod krivulje značajke djelovanja prijarnika (ROC AUC), odziv, preciznost, F1-mjeru i površinu ispod krivulje preciznost-odziv (PR AUC). Nova metoda nadmašuje osnovni model strojnoga učenja treniran na podacima u obliku vremenskih nizova s obzirom na razmatrane metrike. Istraživanje pokazuje da se predložena tehnika može proširiti i na različite druge primjene koje uključuju nestacionarne podatke u intenzivnom šumu.

Ključne riječi: nestacionarni signali; vremensko-frekvencijski prikazi; umjetna inteligencija; strojno učenje; konvolucijske neuronske mreže; gravitacijski valovi.

Jonatan Lerga

University of Rijeka
Faculty of Engineering
Department of Computer Engineering
Vukovarska 58, HR 51000 Rijeka, Croatia
Center for Artificial Intelligence and Cybersecurity
Radmile Matejčić 2, HR 51000 Rijeka, Croatia
E-mail: jlerga@riteh.hr

Nikola Lopac

University of Rijeka
Faculty of Maritime Studies
Department of Electrical Engineering, Automation
and Computing
Studentska 2, HR 51000 Rijeka, Croatia
Center for Artificial Intelligence and Cybersecurity
Radmile Matejčić 2, HR 51000 Rijeka, Croatia
E-mail: nikola.lopac@pfri.uniri.hr

David Bačnar

University of Rijeka

Faculty of Engineering

Department of Computer Engineering

Vukovarska 58, HR 51000 Rijeka, Croatia

Center for Artificial Intelligence and Cybersecurity

Radmile Matejčić 2, HR 51000 Rijeka, Croatia

E-mail: dbacnar@riteh.hr

Franko Hržić

University of Rijeka

Faculty of Engineering

Department of Computer Engineering

Vukovarska 58, HR 51000 Rijeka, Croatia

Center for Artificial Intelligence and Cybersecurity

Radmile Matejčić 2, HR 51000 Rijeka, Croatia

E-mail: fhrzic@riteh.hr