A Supervised Approach for the Detection of AM-FM Signals' Cross-Terms in Spectrogram Images

Vittoria Bruni[a*], Domenico Vitulano[a], Silvia Marconi[b]

[a]*Department of Basic and Applied Sciences for Engineering, Sapienza University of Rome, via Antonio Scarpa 16, Rome, 00161, Italy*
[b]*Department of Methods and Models for Economics, Territory and Finance, Sapienza University of Rome, Via del Castro Laurenziano 9, Rome, 00161, Italy*

Abstract

Ridge curves retrieval in time-frequency (TF) domains is fundamental in many application fields as they convey most of information concerning the instantaneous frequencies of non-stationary signals. However, it represents a very hard task in the case of multicomponent signals having non-separable modes because of the presence of cross-terms that generate interference in TF domains. A preliminary detection of these interference regions may be then useful for the definition of effective strategies for ridge curve recovery. This paper introduces a novel approach based on machine learning for the automatic detection of interference regions in spectrogram images. Each spectrogram sample is suitably classified as interference, single mode or background by accounting for its relative information. Some critical problems, such as the training set size and the type of examples to use for populating the training set, are dealt with. Experimental results show that a local linear model for spectrogram image and a small training set can guarantee good classification rates for a wide class of non-stationary signals, even in the presence of moderate noise.

Keywords: multicomponent signals, instantaneous frequency estimation, interfering AM-FM signals, cross-terms identification, machine learning, spectrogram classification, convolutional neural networks, supervised learning

*Corresponding author

1. Introduction

In the last years, there has been an increasing number of approaches dealing with problems involving MultiComponent Signals (MCSs). The reason of this interest stems from the fact that many practical applications are based on MCSs. The latter are usually characterized by a superposition of specific waveforms with given time-dependent frequency content along with an amplitude modulation [1]. That is why they are also referred to as amplitude and frequency modulated signals (AM-FM signals) — or, more simply, chirps. Both frequency and amplitude are fundamental and their instantaneous value i.e., respectively, Instantaneous Frequency (IF) and Instantaneous Amplitude (IA), must be accurately estimated for each component under study. For instance, IF plays a fundamental role in biology and medicine [2, 3, 4], audio processing [5, 6], civil and military air traffic control and security [7, 8, 9], seismology [10, 11], physics [12, 13], etc..

Unfortunately, observed MCSs are a complicated mixture of single components where a correct estimate of each IF and IA passes through modes separation. Even in this case, there is a plethora of approaches dealing with this problem in different domains (as shortly presented in the next section), with different pros and cons. An example is shown in Fig. 1: it refers to a signal representation in the time frequency domain, that is one of the most popular domains for the analysis of non-stationary signals [14, 15]. Specifically, it depicts the spectrogram (squared modulus of the also popular Short Time Fourier Transform (STFT) [16]) of a two component signal. As it can be observed, just two components generate an interference region in the time-frequency plane where their separation can represent a hard task as well as the estimation of the corresponding IF and IA. The task becomes more challenging whenever amplitude modulation concentrates energy signal mainly in the interference region.

Such a difficulty has been mathematically proven in [17, 18, 19], where the spectrogram of both a monocomponent signal and a multicomponent one have been described by a partial differential equation — specifically, an advection equation. In the case of two or more components, such an equation shows some interfering terms that are complicated to manage from both a practical and theoretical point of view. Apart from this theoretical result, the difficulty in managing interference is proven by the plethora of approaches proposed for this specific task that mainly embed ad hoc strategies in more general models. For instance, some approaches treat signal time-frequency

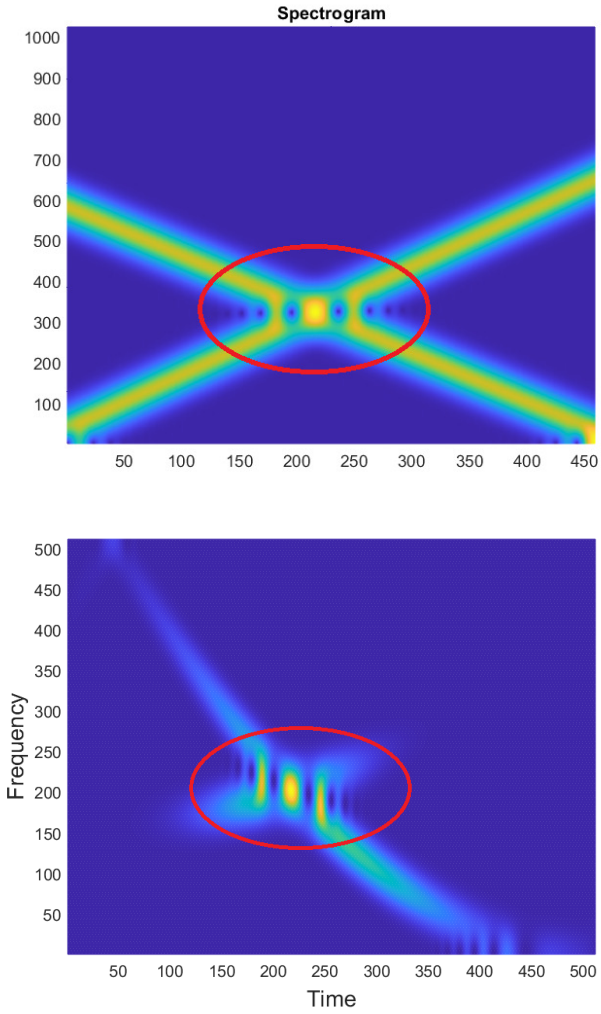

Figure 1: Spectrogram of a two-component frequency modulated signal with: **Top)** constant amplitudes; **Bottom)** time-dependent amplitudes. Red ellipses show the interference region.

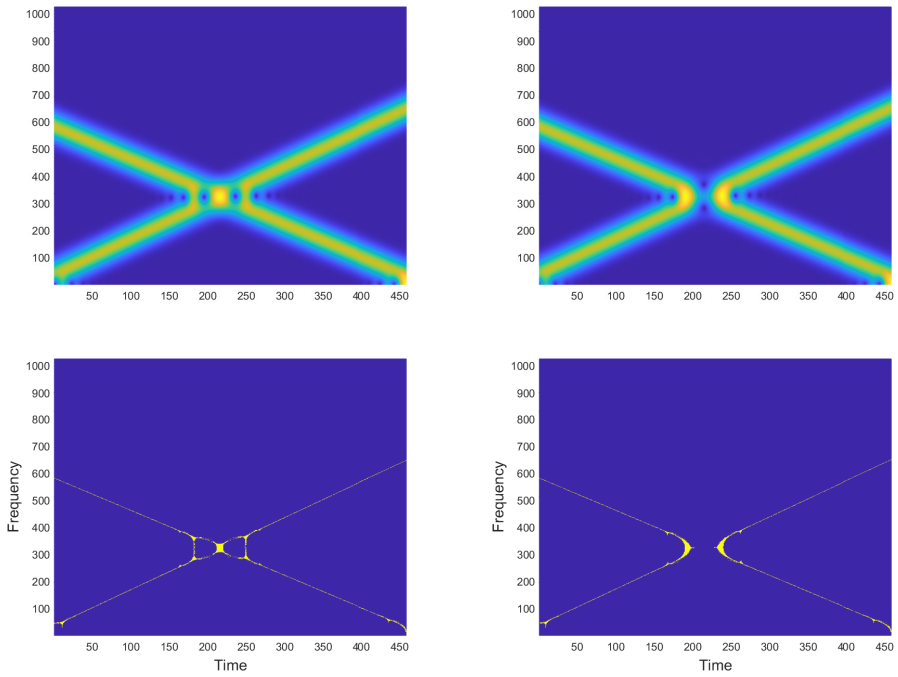

Figure 2: Examples of two-component FM signals. **Top)** Spectrograms; **Bottom)** IF reconstruction via reassignment technique [1]. IF estimation fails because of interference effects: additive *(left)* and subtractive *(right)*.

representation as an image and missing points in the interference region are achieved by interpolating detected IF points [20, 21, 22, 23]. The drawback of this kind of approaches stems from a high computational cost as well as a high sensitivity to noise. Moreover, they often are subjected to the well known switch problem that consists of misalignment of detected ridge points. Other attempts try to use parametric strategies to mitigate missing or wrong IF points [24, 25, 26, 27, 28, 29], including those oriented to define adaptive or optimal support for the analysis windows [30]. More recently, deep learning is used for better reassigning time-frequency distributions over the ridge curves; although promising, several data have to be used for training in order to deal with the high variability of signals [31]. It turns out that an effective preprocessing of the spectrogram image that is able to detect interference regions before spectrogram analysis may be of great interest for the research community working on this topic.

It is also true that interference often is easily detectable by an expert, as

it is evident from Fig. 2 that contains different spectrograms with different MCSs combinations. Though different, interference regions show some non trivial peculiarities that allow to characterize them. A natural question then arises: can a suitably trained machine learning tool recognize interference regions? Machine learning (ML), and in general artificial intelligence (AI), had a huge impact in different fields in the past few years [32]. In particular, Convolutional Neural Networks (CNNs) have been widely used for classification, object detection and recognition, showing very high performance in different contexts of pattern recognition [33, 34, 35]. In addition, recently AI has also been adopted for the solution of classical partial differential equations (PDEs), fractional equations, integral-differential equations as well as stochastic PDEs [36]. If on the one hand, AI has often been criticized "as a very poor caricature of the workings of the human brain" [32], on the other hand, it uses abductive reasoning i.e., it is able to make a probable conclusion from what you know [37]. Taking into account that the mathematical formalism is based on induction and deduction, it turns out that AI can pave the way to different solutions and results. In particular, though very performing in specific case studies, abductive results can be used to produce an alternative and mathematical formalism that can lead to more performing analytical and numerical solutions.

This paper just focuses on the aforementioned observation and then proposes a novel approach oriented to detect interference regions in the spectrogram image by turning it into a problem of image regions classification, that can be addressed using CNN architectures [32]. Specifically, potential and limits of such an approach are studied. Apart from the classical questions tied to a machine learning approach, such as how to build a suitable training set (the minimum number of interfering chirps), how accurate the model is, what is its computational time etc., this approach aims at better understanding the intrinsic difficulty of interference detection that may be useful to find out a different mathematical formalism able to overcome current numerical problems.

The paper is organized as follows. Section 2 offers a short state of the art concerning IF estimation approaches that are closer to the proposed strategy. Section 3 introduces the motivations of the proposed approach along with a technical background useful to understand the rest of the paper. Section 4 presents the proposed model, while Section 5 contains some representative simulations relative to different cases of interest. Finally, Section 6 draws the conclusions.

2. A short state of the art

IF estimation is a very challenging task as proven by the plethora of approaches proposed in the literature. The following state of the art is just oriented to show the major research lines without any claim of being exhaustive. Particular attention will be devoted to those approaches that are closer to the proposed one. A first and broad classification of existing approaches can take into account the working domain. As a result, two classes can be identified: the first class contains strategies working in the time domain, while the second one includes methods defined in the time-frequency plane. A representative example for the first class is the Empirical Mode Decomposition [38, 39, 40], while examples in the more populated second class are those that use spectrogram, syncrosqueezing, Smoothed Wigner Ville Distribution or Adaptive Directional Time Frequency Distribution for estimating IF [29, 41, 42, 43, 44, 45, 46, 47, 48, 49], just to mention some of them.

Approaches belonging to both classes have pros and cons. Time-domain based approaches can directly process the signal in the time domain avoiding problems arising from kernel convolutions and adopted transform; on the other hand they obviously show a greater sensitivity to noise that usually characterizes real world chirps — very often with a non negligible computational effort. Transform-based approaches are highly sensitive to modes separability and robustness to noise represents one of the main challenge [50] — as a result, they often are designed for specific families of signals.

With regard to the first class of methods, an interesting IF estimator designed for wideband non linearly FM signals has been proposed in [51]. The latter is based on removal of signal non-stationarity, by minimizing the bandwidth of the demodulated signal. It is also quite effective for crossing modes (with some limits in the case of subtractive interference) but it is limited to a specific signal family, as it also happens for many approaches belonging to the second class of methods [52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63]. As far as it concerns the second class, it is also worth mentioning those approaches that take advantage of the combination of more than one transform (for instance spectrogram with Hough, Radon and Inverse Radon transform) that have shown a good performance for different kinds of chirp — linear [64, 65, 66], polynomial [67], sinusoidal [53] and non linear ones [68]. There have been some approaches exploiting imaging for IF estimation such as [20, 21], and Viterbi algorithm-based methods [22, 23]. The simplicity of these approaches is often accompanied by a not negligible computational effort, sensitivity to

noise and not correct ridge assignment in the case of interference or close chirps (*switch problem*). Some approaches oriented to minimize this aspect are in [24, 26, 69], that mainly aim at estimating IF direction, and [25], that assumes local monotonicity. Alternative to Viterbi method but similar in spirit approaches are the Ridge Path Regrouping Method (RPRM) [27, 28] (showing a certain sensitivity to noise) and optimization techniques [27, 29], properly designed to mitigate interference effects. As interference among modes is a critical problem also for classical approaches like reassignment method [70], synchrosqueezing [71, 72] and syncroextracting transform [73, 74], some refinement methods have been defined to mitigate this problem, as in [17, 18, 75, 76], as well as alternative approaches like the variational method proposed in [51] and the de-chirping technique in [52, 77, 78]. More recently ML-aided methods have been developed for modes separation other than signal classification. A very recent example is the Time Frequancy Analysis-NET [31] that has been properly built for reassignment purposes. Although promising, it requires populated and representative training set in order to be sufficiently general and robust to noise. To weaken the dependence on modes separability, modes counting methods can be employed to predict eventual nonseparability regions [19, 79, 80]. They are mainly based on a specific measure of signal complexity, as for example, Renyi entropy [79], multiscale entropy [80] or run-length encoding [19]. Unfortunately, independently of the specific constraints and working scenarios, counting methods limit themselves to output the time-dependent number of components; as a result, a successive processing is required for determining the presence of interference or the incoming of a new and time-limited mode.

To overcome some of these problems, the main focus of the proposed method is to define a simple but enough general and accurate model for a specific task (time-frequency plane segmentation and labeling) that makes use of small amount of training data. It is a transform-based method that exploits the spectrogram as an image by modeling interference detection problem as a classification one, as explained in details in the next sections.

3. Motivation of the work

AM-FM signals can be mathematically written as follows [1]:

$$f(t) = \sum_{k=1}^{K} f_k(t) = \sum_{k=1}^{K} a_k(t) e^{i\phi_k(t)}\,, \tag{1}$$

where K represents the number of components while $f_k \in L^2(\mathbb{R})$ is the $k-$th mode with time-dependent amplitude a_k and phase ϕ_k. The time derivative of ϕ_k is the instantaneous frequency $\phi'_k(t)$ of the $k-$th mode.

The Short-Time Fourier Transform of f is:

$$S_f^g(u,\xi) = \int_{-\infty}^{+\infty} f(t)g(t-u)e^{-i\xi t}\,dt, \qquad \forall\,(u,\xi) \in \mathbb{R}\times\mathbb{R}^+, \tag{2}$$

where $g \in L^2(\mathbb{R})$ is a real and symmetric analysis window, while u and ξ represent time and frequency variables, respectively.

The spectrogram P of f is defined as the squared modulus of $S_f^g(u,\xi)$: $P(u,\xi) = |S_f^g(u,\xi)|^2$.

By using the dilated window $g_s(t) = \frac{1}{\sqrt{s}} g\left(\frac{t}{s}\right)$, with $s > 0$, the STFT of a monocomponent signal $f(t) = a(t)\cos\phi(t)$ can be written as [16]:

$$S_f^g(u,\xi) = \frac{\sqrt{s}}{2} a(u) e^{i(\phi(u)-\xi\cdot u)} \left[\hat{g}\left(s(\xi-\phi'(u)) + \epsilon(u,\xi)\right)\right], \tag{3}$$

where $\hat{*}$ denotes the Fourier Transform of $*$, g is a normalized window having support $\left[-\frac{1}{2},\frac{1}{2}\right]$ and frequency bandwidth $\Delta\omega$ [16], while $\epsilon(u,\xi)$ is a corrective term that can be considered negligible for $a(t)$ and $\phi'(t)$ slightly varying in the support of the analysis window. Under these constraints, the spectrogram analytical form becomes:

$$P(u,\xi) = \frac{s}{4}\, a^2(u)\hat{g}^2\left(s(\xi-\phi'(u))\right). \tag{4}$$

Accordingly, the spectrogram of a two-component signal $f(t) = a_1(t)f_1(t) + a_2(t)f_2(t)$, like the one in Fig. 1, can be written as:

$$P(u,\xi) = P_1(u,\xi) + P_2(u,\xi) + 2\sqrt{P_1P_2}\cos(\phi_2(u) - \phi_1(u)), \tag{5}$$

where $P_k = P_k(u,\xi) = |S_{f_k}^g(u,\xi)|^2 = \frac{s\,a_k^2(u)}{4}\hat{g}^2(s(\xi-\phi'_k(u)))$, $k = 1, 2$.

Eq. (5) results less simple than eq. (4) because of the presence of the cross-term, that is responsible for the difficulty in IF estimation in the time-frequency regions where the separability condition

$$|\phi'_1(u) - \phi'_2(u)| \geq \Delta\omega \tag{6}$$

does not hold true. In fact, as Fig. 2 shows, the two modes overlap both in time and frequency. In particular, modes interference is additive whenever

the spectrogram energy increases (i.e., it is greater than the sum of the two isolated chirps energies), and subtractive in the opposite case. In both cases modes overlapping causes wrong ridge points detection (spectrogram maxima), and then incorrect IF estimation. To address this issue, in [17, 54] it has been proven that alternative time-frequency curves (spectrogram isolevel curves for constant amplitude FM signals) contain the same information of ridges but are less sensitive to cross-terms [17, 54]. These curves are characteristic curves $\mathcal{C}_{c,\phi}$ of the following advection equation

$$\frac{\partial P(u,\xi)}{\partial u} \quad +\phi''(u)\frac{\partial P(u,\xi)}{\partial \xi} - \frac{2a'(u)}{a(u)}P(u,\xi) = 0, \quad \forall\, u \in supp\{f\}, \tag{7}$$

that is satisfied by the spectrogram $P(u,\xi)$ of a monocomponent signal. Those characteristic curves $\mathcal{C}_{c,\phi}$ are

$$\xi(u) = \phi'(u) + c, \tag{8}$$

with $c = \xi_0 - \phi'(u_0)$ and (u_0, ξ_0) is a point in the time-frequency (TF) plane [18, 19]. Hence, they are nothing else than shifted copies of the ridge curve. As a result, characteristic curves sufficiently far from ridges correspond to spectrogram points that are less influenced by cross-terms; then, they allow to formally estimate IF with more precision in the case of multicomponent signals — see [18, 19] for a more rigorous explanation.

However, the selection of those good points is not always feasible, especially for fast varying amplitude modulation functions or interference kind/level. In fact, though cross-terms perfectly describe interference effects, they do not help to completely solve IF estimation problem, as it is evident from the evolution law of the spectrogram $P(u,\xi)$ of a two-component signal, i.e.

$$\begin{aligned} &\frac{\partial P(u,\xi)}{\partial u} + \phi_1''\frac{\partial P(u,\xi)}{\partial \xi} + \\ &-\frac{s}{2}\Big[a_1 a_1' \hat{g}_1^2 + a_2 a_2' \hat{g}_2^2 + \hat{g}_1\hat{g}_2(a_1' a_2 + a_2' a_1)\cos\Delta\phi\Big] + \\ &+\frac{s}{2}a_1 a_2 \hat{g}_1 \hat{g}_2 \Delta\phi' \sin\Delta\phi + \frac{s^2}{2}\Delta\phi''\hat{g}_2'\Big[a_2\hat{g}_2^2 - a_1 a_2 \hat{g}_1 \cos\Delta\phi\Big] = 0, \end{aligned} \tag{9}$$

where $\hat{g}_k = \hat{g}(s(\xi - \phi_k'(u)))$, $a_k = a_k(u)$, $\phi_k = \phi_k(u)$, $k = 1, 2$, and $\Delta\phi = \phi_1 - \phi_2$.

That is why refined or alternative approaches are required.

However, independently of the adopted refinement, it would be useful to have prior information concerning TF interference regions. The next section addresses this issue.

4. The proposed model

With reference to eqs. (3) and (5), the time-frequency plane Ω_P can be partitioned into three main regions i.e.,

$$\Omega_P = \Omega_o \cup \Omega_m \cup \Omega_b \tag{10}$$

with

$$\begin{cases} \Omega_o = \{(u,\xi) \in \Omega_P : \exists!\ k \in \{1,2\} : (u,\xi) \in \Omega_{P_k}\} \\ \Omega_m = \{(u,\xi) \in \Omega_P : \exists\ k,j \in \{1,2\}, \quad k \neq j : (u,\xi) \in \Omega_{P_k} \cap \Omega_{P_j}\} \\ \Omega_b = \{(u,\xi) \in \Omega_P : \nexists\ k \in \{1,2\} : (u,\xi) \in \Omega_{P_k}\}, \end{cases},$$

where Ω_* denotes the support of $*$.

Ω_o is the one-mode support and it collects the TF regions where only a single mode contributes to the spectrogram P; Ω_m is the multi-mode support, i.e., the TF interference region among two (or more) chirps as formally written in eq. (5); finally, Ω_b includes the remaining TF points and represents the background i.e., the points where there is no signal.

The main idea is then to define a method that is able to recognize interference regions (Ω_m), i.e., regions that do not satisfy eq. (6). In agreement with eq. (10), the problem then turns into an image classification/segmentation one, where the image is the spectrogram. The image classification problem can be addresed using a CNN with a supervised learning approach: the network learns from a set of input/output examples during the training phase and then makes its own prediction on unseen data in the testing phase. In the present classification problem, square patches of the spectrogram can be considered as input of the CNN and the central point label of each patch is considered as output. In particular, a custom CNN properly designed for hyperspectral data classification purposes [34, 35] has been used and it has been modified to be applied to 2D rather than 3D data in order to classify each point of the spectrogram on the basis of its neighborhood. The adopted architecture conjugates moderate depth and high accuracy rate; similar structures have demonstrated to be able to extract distinctive features for classification requiring a moderate learning time — details concerning the adopted CNN are contained in the next section.

Apart from the selected architecture, the main issue is to properly define the training set to use during the learning process. This is a delicate point especially if one considers that the proposed approach aims at being non

parametric. Independence of transform parameters as well as of chirp family usually is critical and can limit the actual applicability of many approaches. With reference to the purpose of this paper, independence of chirp family is then an issue to address till the definition of the training set. In particular, we would like to define as less examples as possible while mantaining high classification rates in very general cases.

By taking advantage of modeling the problem as a pointwise classification, local approximations of the spectrogram can be considered for restricting the class of IF functions to use in the training phase. In particular, we are interested in approximating a generic IF function $\phi'(u)$ with a piecewise linear IF $\varphi'(u)$ that is defined over a uniform distribution of n points $\{u_i\}_{i=0,\ldots,n-1}$ in the domain $\Omega_{\bar{u}}$ centered at time location $\bar{u}$. More precisely,

$$\phi'(u) = \varphi'(u) + \tau(u), \quad \forall\, u \in \Omega_{\bar{u}} \tag{11}$$

where

$$\varphi'(u) = \sum_{i=0}^{n-2} p_i(u)\chi_{D_i}(u) \tag{12}$$

is the approximating piecewise linear IF with

$$p_i(u) = \phi'(u_i) + \frac{\phi'(u_{i+1}) - \phi'(u_i)}{u_{i+1} - u_i}(u - u_i),$$

$D_i = [u_i, u_{i+1}]$ is the domain of the $i-th$ linear piece, and $\chi_{D_i}(u)$ is the indicator function, while $\tau(u)$ is the pointwise approximation error.

It is worth observing that:

1. independently of the approximation kind, the better the approximation the more $\phi'(u)$ and $\varphi'(u)$ are far from satisfying the separability condition, i.e.

$$|\tau(u)| = |\phi'(u) - \varphi'(u)| << \Delta\omega; \tag{13}$$

2. according to eq. (8), the error for $\phi'(u)$ is the same for the associated characteristic curves;
3. $\tau(u)$ bounds the error for the spectrogram. In particular, let P be the spectrogram of a monocomponent signal having $\phi'(u)$ as IF, and let Q be the spectrogram of a monocomponent signal having $\varphi'(u)$ as

IF (both signals are supposed to have the same amplitude modulation function), then

$$P(u,\xi) = Q(u,\xi) + R(u,\xi), \qquad \forall\,(u,\xi) \in D_{\bar{u},\bar{\xi}}$$

where R denotes the approximation error and $D_{\bar{u},\bar{\xi}} = \Omega_{\bar{u}} \times \Omega_{\bar{\xi}}$ a rectangular domain centered at the TF point to be classified $(\bar{u},\bar{\xi})$.
Eq. (4) and the Lagrange theorem for continuous functions provide

$$|R(u,\xi)| = |Q(u,\xi-\tau(u)) - Q(u,\xi)| = |Q_\xi(u,\eta)|\,|\tau(u)|, \qquad (14)$$

where Q_ξ denotes the derivative of $Q(u,\xi)$ with respect to ξ, and $\eta \in [\xi-\tau(u),\xi]$. Hence, $\forall\,(u,\xi) \in D_{\hat{u},\hat{\xi}}$ it holds

$$|R(u,\xi)| \leq max_{\xi\in\Omega_{\bar{\xi}}}|Q_\xi|(u,\xi)\,|\tau(u)|.$$

For piecewise linear approximations of $\phi'(u)$, the error $\tau(u)$ can be explicitly written as

$$\tau(u) = \frac{1}{2}\sum_{i=0}^{n-2} \phi'''(\bar{u}_i)(u-u_i)(u-u_{i+1})\chi_{D_i}(u), \qquad \bar{u}_i \in D_i \qquad (15)$$

while the global error over the whole time interval $\Omega_{\bar{u}}$ can be defined as

$$\tau = \int_{\Omega_{\bar{u}}} |\tau(u)|du.$$

In particular, using a simple algebra, it can be bounded as follows

$$\tau = \int_{\Omega_{\bar{u}}} |\tau(u)|du \leq M\frac{|\Omega_{\bar{u}}|^3}{12(n-1)^2}, \qquad (16)$$

where $M = max_{u\in\Omega_{\bar{u}}}|\phi'''(u)|$ and $|\Omega_{\bar{u}}|$ is the length of $|\Omega_{\bar{u}}|$.

As it can be observed, the approximation error depends on the second derivative of $\phi'(u)$, that mainly measures the variability of the curvature of the IF. In addition, the larger $|\phi'(u)|$, the larger the number $n-1$ of linear pieces to use in the approximation in order to guarantee a fixed accuracy.

Using eq. (13) for bounding τ in eq. (16), we can conclude that a proper local linear approximation of ϕ' such that $|\phi'(u)-\varphi'(u)| \leq \Delta\omega$ can guarantee negliglible error for classification purposes. In particular, more points are needed for IFs having rapidly varying curvatures (chirp rates).

It is also worth observing that the considerations above have been made in a subset of the TF plane centered at the point to be classified; as a result, the choice of the rectangular window $D_{\bar{u},\bar{\xi}}$ plays a fundamental role: it has to be sufficiently small to guarantee high accuracy of the linear approximation; on the other hand, it has to be enough large to allow to establish the class of the interested point on the basis of its neighborhood. That is why it is expected that it is comparable with the domain of the analysis window used for spectrogram computation.

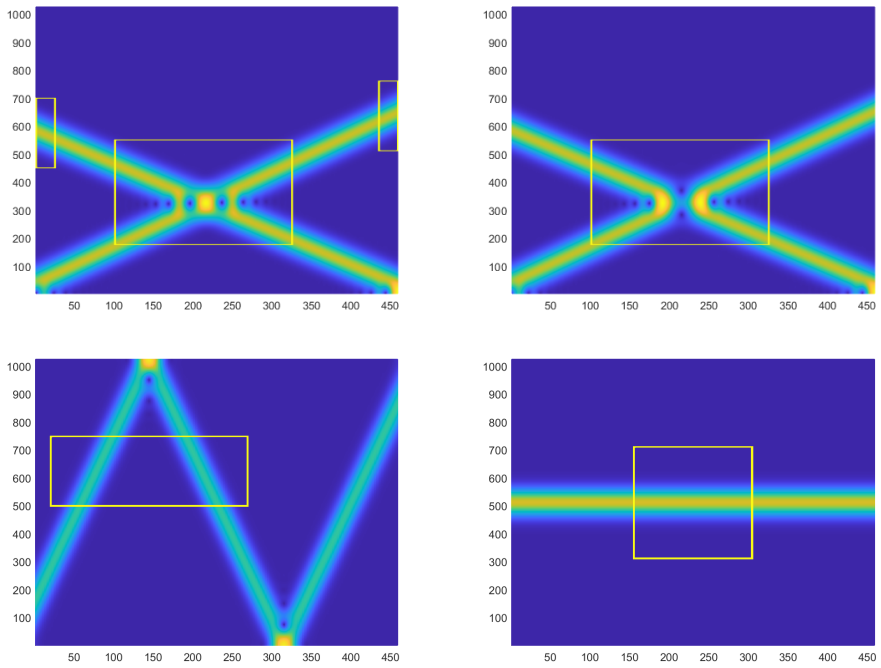

Figure 3: Spectrogram of chirps in eq. (17). **Top**) Additive interference *(left)*, subtractive interference *(right)*. **Bottom**) High slope single chirp *(left)* and constant chirp *(right)*. Rectangles catch peculiar regions of each case.

These observations are fundamental for the definition of the training set that has to contain piecewise linear functions. In particular, two questions arise: how many points should be considered for guaranteeing acceptable accuracies, and what slope the linear chirp should have. Furthermore, as we have to give examples of interference, the additional question is if we have to use all the employed linear monocomponent signals to simulate interference. With regard to the last question, a broad classification of interference

regions may be done in terms of an energy conservation criterion. In fact, as previously discussed, there are usually effects of energy increase (additive inteference) or energy cancellation (subtractive interference) — as depicted in Fig. 3 and in eq. (5). As a result, at least one example for each interference type has to be included in the training set.

Regarding the slopes, they have to be representative of constant, moderate varying and highly varying IFs. Finally, as far as it concerns the number of linear pieces, it is reasonable to use mainly just two-point approximations, while leaving three-point approximations only for IFs having highly varying chirp rate.

Additional details are given in the next section where the performance of the proposed approach under different design choices will be evaluated.

5. Experimental Results

This section provides implementation details related to the proposed approach as well as the evaluation of the results achieved on some selected synthetic AM-FM signals. In all tests AM-FM signals with $N = 512$ samples have been considered. A Gaussian window has been used as STFT analysis window with $M = 54$ samples, the overlapping range in STFT computation has been set equal to $L = M - 1$ while the number of voices equal to $nfft = 2048$. Finally, all tests have been performed in Matlab R2021a on Intel(R) Xenon(R) Gold 6238R CPU 2.20GHz Processor.

Before showing spectrograms classifications, some points concerning the adopted CNN deserve attention and they are listed below:

- definition of the training set so that the selected CNN is able to classify as many (interfering) chirp kinds (in terms of functional classes) as possible by optimizing, at the same time, accuracy and computational time;
- labels definition from spectrograms of single mode chirps;
- selection of the patch size to use as input of the selected CNN;
- definition of the optimal design parameters for the selected CNN;
- noise effects: how noise can influence the final accuracy and how to make CNN robust to it.

These points will be addressed separately in the following subsections.

5.1. Training set definition

Based on the considerations made at the end of the previous section, just linear IFs with different slope (in the time-frequency plane) have been selected, each representative of slow, medium and high IF variation. In addition, additive and subtractive interference between two linear chirps having moderate slopes have been considered.

Specifically, the following signals have been selected:

$$\begin{aligned}
&\textbf{Additive interference}: \ \mathrm{sig}_A(t) = \mathrm{sig}_{A1}(t) + \mathrm{sig}_{A2}(t) \\
&\mathrm{sig}_{A1}(t) = \cos\left(\frac{\pi}{3}Nt^2 + 0.2\pi t\right), \qquad \mathrm{sig}_{A2}(t) = \sin\left(0.3\pi N(1-t)^2\right) \\
&\textbf{Subtractive interference}: \ \mathrm{sig}_S(t) = \mathrm{sig}_{S1}(t) + \mathrm{sig}_{S2}(t) \\
&\mathrm{sig}_{S1}(t) = \cos\left(\frac{\pi}{3}Nt^2 + 1.5\pi t\right), \qquad \mathrm{sig}_{S2}(t) = \sin\left(0.3\pi N(1-t)^2\right) \\
&\textbf{High Slope}: \ \mathrm{sig}_{HS}(t) = \cos\left(\frac{3}{2}\pi Nt^2\right) \\
&\textbf{Constant}: \ \mathrm{sig}_C(t) = \cos\left(\frac{\pi}{2}Nt\right)
\end{aligned} \tag{17}$$

The training set is then composed of their spectrograms. As a matter of facts, to reduce training time only representative spectrogram points belonging to specific TF regions have been selected, as depicted in Fig. 3.

5.2. Labels definition

In order to define the labels, the spectrogram of each single chirp has been binarized. Binarization is performed by means of a threshold equal to 5% of the maximum value of the spectrogram under study — regions close to spectrogram border are not considered. It turns out that background points (Ω_b in eq. (10)) will have a label value equal to '0', while chirp points (Ω_o in eq. (10)) will be set equal to '1'. The map of interfering chirps is achieved as intersection of the maps of the isolated chirps; hence, interference regions (Ω_m in eq. (10)) will be characterized by the label '2', as depicted in Figs. 4 and 5. The labeled training set is shown in Fig. 6.

5.3. Patch size setting

With regard to patch size, as discussed in Section 4, it has been set in agreement with the support of the analysis window adopted for STFT

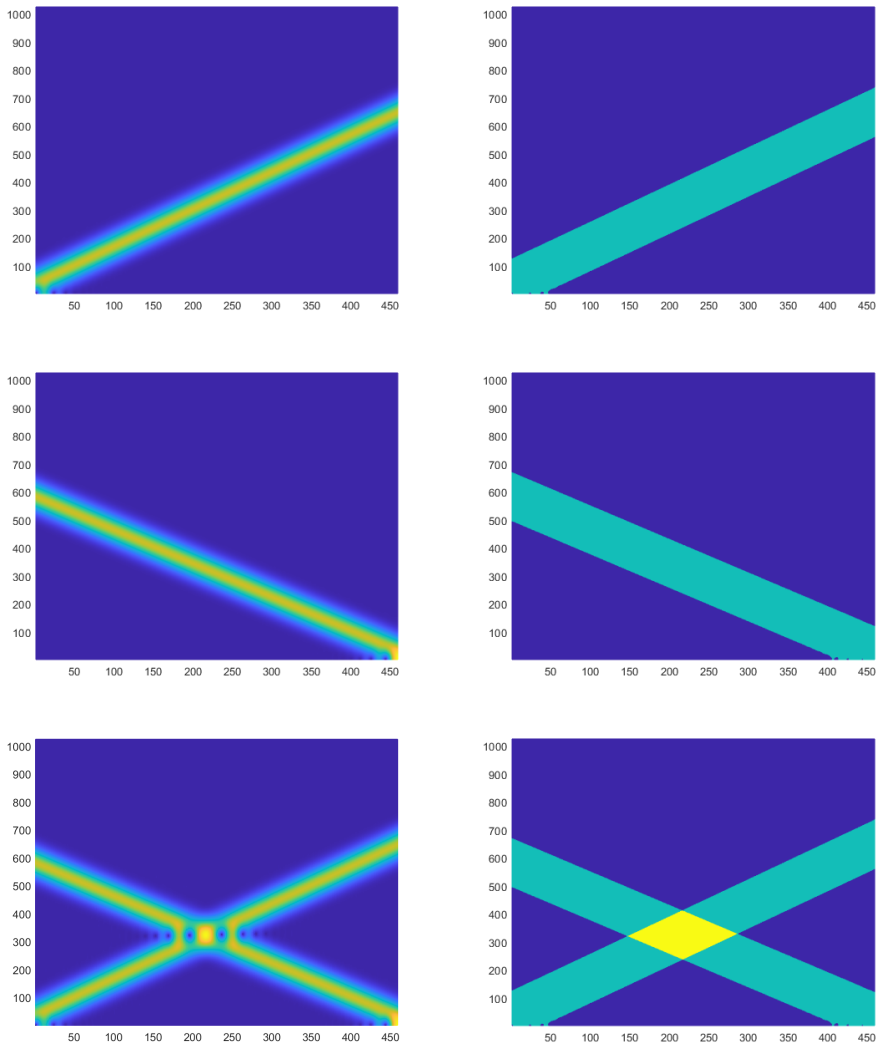

Figure 4: Labels definition in additive interference case. **Top**) Spectrogram of signal sig_{A1} in eq. (17) *(left)* and its domain after binarization *(right)*. **Middle**) Spectrogram of sig_{A2} in eq. (17) *(left)* and its domain after binarization *(right)*. **Bottom**) Spectrogram of sig_{A} in eq. (17) *(left)* and corresponding labels colors: '0' = blue is Ω_b, '1' = cyan is Ω_o, '2'= yellow is Ω_m *(right)*.

computation — $M = 53$ in the test presented in this section. This size allows to minimize misclassification as it gets the best tradeoff between patch size representativeness and localization. This statement has been also confirmed

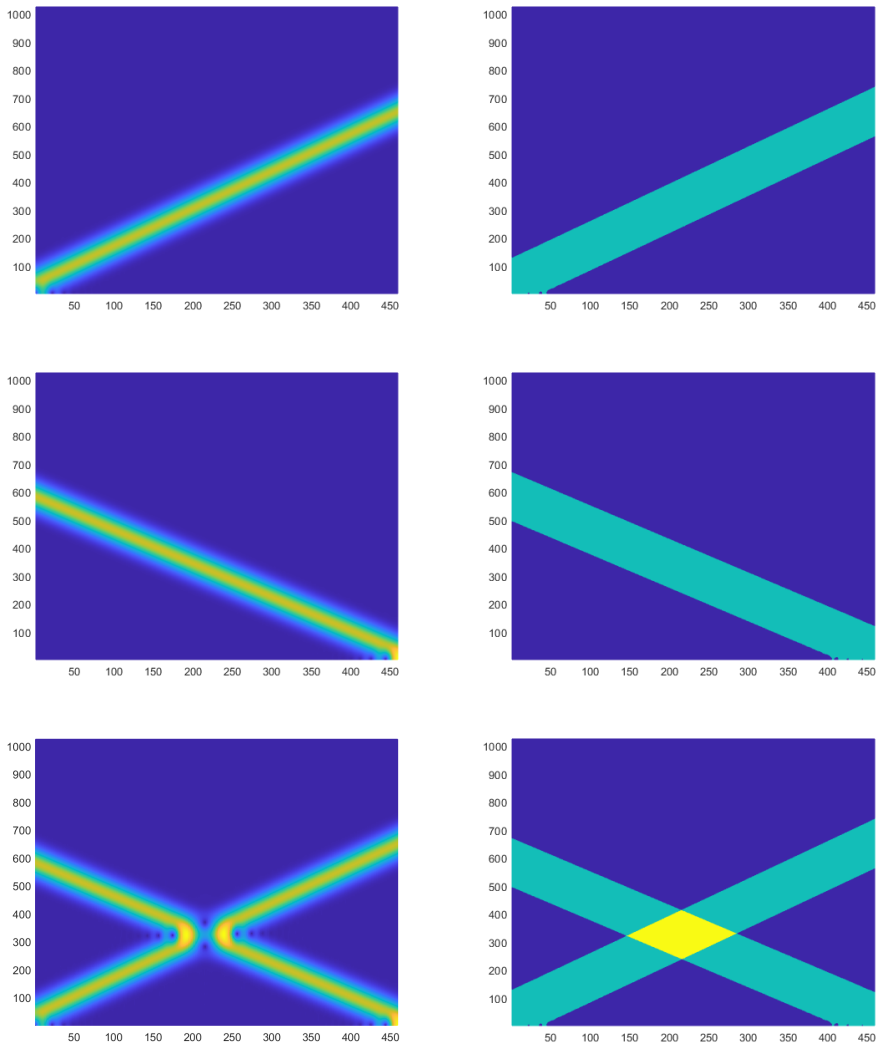

Figure 5: Labels definition in subtractive interference case. **Top**) Spectrogram of signal sig_{S1} in eq. (17) *(left)* and its domain after binarization *(right)*. **Middle**) Spectrogram of sig_{S2} in eq. (17) *(left)* and its domain after binarization *(right)*. **Bottom**) Spectrogram of sig_S in eq. (17) *(left)* and corresponding labels colors, '0' = blue is Ω_b, '1' = cyan is Ω_o, '2'= yellow is Ω_m*(right)*.

by intensive tests using different patch sizes.

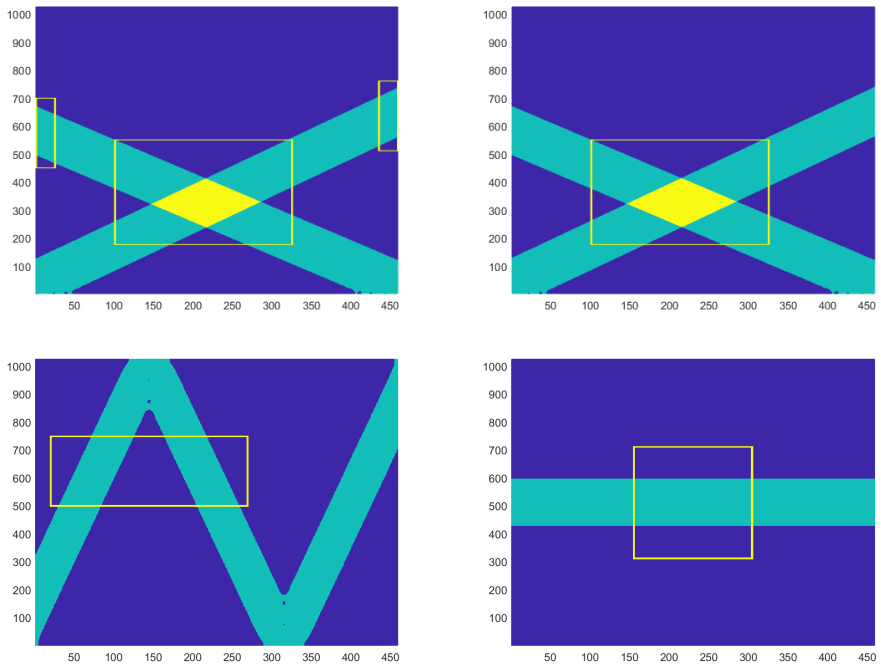

Figure 6: Labels for the selected training set.

5.4. Adopted CNN

The block scheme of the adopted CNN has been shown in Fig. 7. It is composed of four *convolutional layers* interlayed by a *ReLu layer* for feature extraction purposes, three *fully connected layers*, whose output size is 256, 128 and number of expected classes respectively, a *softmax layer* and a *classification layer*. The input layer accepts patches with size $M \times M$ while each convolutional layer employs 8, 16, 32 and 8 fixed size filters (3×3 in this paper), respectively.

All results are the output after 40 epochs using Adam optimizer, 256 as minibatch size, 0.001 as initial learning rate, 0.1 as learning rate factor. Shuffle is performed every epoch.

5.5. Classification Results

To detect interference regions, the proposed network has been trained on overlapping patches falling in the rectangles (ROIs) in Fig. 3 — 60% of the whole points have been used for training and 20% for validation.

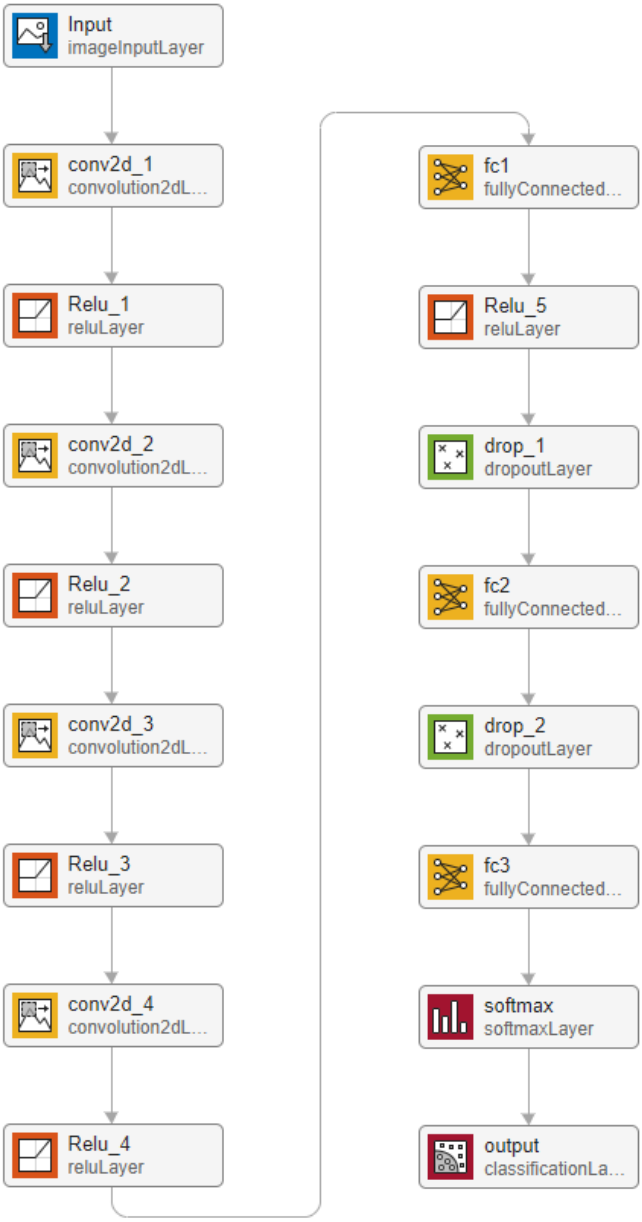

Figure 7: Block scheme of the adopted Convolutional Neural Network.

Validation accuracy [%]	Test accuracy [%]
99.64	99.61

Table 1: Classification results for spectrograms in Fig. 3.

The first classification test has been performed on the remaining 20% of the samples of the same ROIs. The achieved results have been measured in terms of both validation and testing accuracy, and they are presented in Table 1. As it can be observed, classes have been correctly recognized, as expected, since testing points belong to spectrogram referring to linear IFs.

The trained CNN has been then tested on spectrograms containing different kinds of chirp — labels for groundtruth are defined with the same criterion described in Section 5.2. Classification refers to the whole spectrogram except for the external border whose width has been set equal to half of the width of patch dimension. This allows to discard annoying border effects.

The following signals have been used as testing set:

- **Group 1: Constant and linear chirps**

$$\begin{aligned}
&\mathrm{sig}_1(t) = \cos\left(\frac{\pi}{4}Nt\right) && \text{(Constant)}\\
&\mathrm{sig}_2(t) = \cos\left(\frac{\pi}{4}Nt^2\right) && \text{(Low Slope)}\\
&\mathrm{sig}_3(t) = \cos\left(\frac{\pi}{2}Nt^2\right) && \text{(Diagonal)}\\
&\mathrm{sig}_4(t) = \cos(\pi Nt^2) && \text{(Moderate Slope)}\\
&\mathrm{sig}_5(t) = \cos(2\pi Nt^2) && \text{(Very High Slope)}
\end{aligned}$$

- **Group 2: Non linear chirps**

$$\mathrm{sig}_6(t) = e^{i(20\pi\cos\left(\frac{\pi}{128}Nt\right)-1500t)} + e^{i(0.95\pi Nt^2+170t)}$$
(Linear + Sinusolidal)

$$\mathrm{sig}_7(t) = \sqrt{t}\cos(0.5Nt^4+500t) + e^{-2(t-0.4)^2}\cos\left[N\left(\frac{N+1}{N}-t\right)^2+120t\right]$$
(Modulated Cubic + Linear)

$$\mathrm{sig}_8(t) = e^{-2(t-0.4)^2}\cos(0.5Nt^3+500t) + e^{-9(t-0.5)^2}\cos\left[1.245N\left(\frac{N+1}{N}-t\right)^3\right]$$
(Modulated Quadratic Chirps)

$$\mathrm{sig}_9(t) = \cos(40\pi\ln(0.05t+0.002))$$
(Hyperbolic)

$$\mathrm{sig}_{10}(t) = \sin(0.3\pi N(1-t)^2) + \cos(0.2\pi Nt^2+0.2\pi t) + \cos(0.5\pi Nt^2+0.5\pi t)$$
(Three Linear Chirps)

Specifically, Group 1 is composed of a constant IF signal and four chirps having different slopes. The classification accuracy is shown in Table 2, while the corresponding classification maps are shown in Fig. 8. As it can be observed, the proposed method is able to correctly classify spectrogram points related to linear slopes that are different from the one used for training. Misclassifications are observed in the radar-like IF functions, such as sig_4 and sig_5, in correspondence to regions where the sign of the slope changes as they are confused with interference region.

Group 2 contains chirp signals having non linear IFs and it has been defined to check whether the proposed CNN-based method is able to detect interference regions even in the case of chirps belonging to classes different from the ones used for training — various combinations of linear chirps with hyperbolic, amplitude modulated and a sinusoidal one have been included. Group 2 also contains a signal composed of three non-separable linear modes. Classification accuracies are in Table 2 while classification maps are provided in Fig. 9. Even in this case, the method is able to correctly classify most of TF points reaching accuracies greater than 97% for signals sig_7, sig_8, sig_9 and sig_{10}. Some misclassifications occur for the signal sig_6 in correspondence to points where IF curvature rapidly changes.

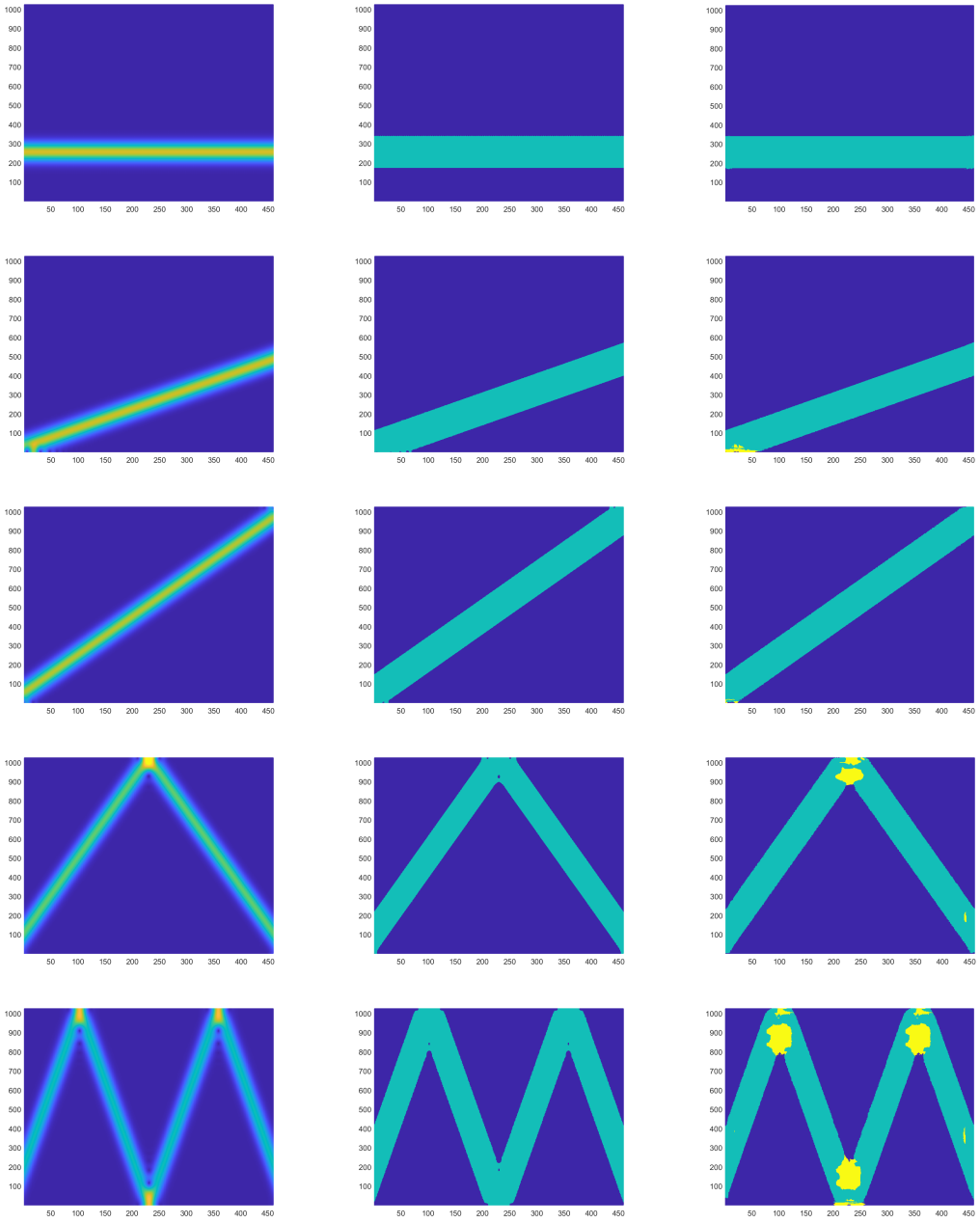

Figure 8: **From top to bottom**) Classification results for the spectrograms of signals in Group 1. Spectrogram *(left)*, spectrogram labels *(middle)*, classification map provided by the proposed method *(right)*.

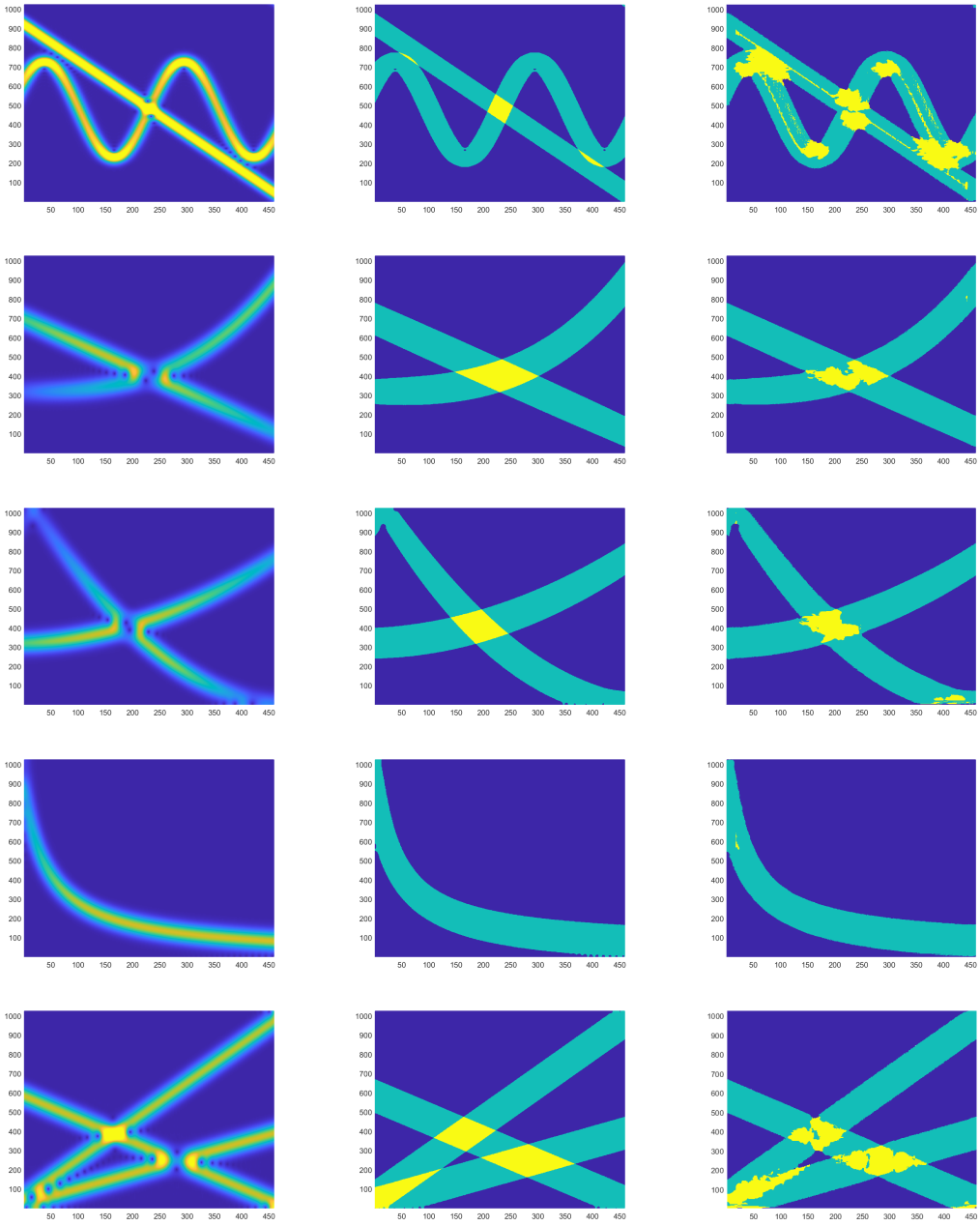

Figure 9: **From top to bottom**) Classification results for the spectrograms of signals in Group 2. Spectrogram *(left)*, spectrogram labels *(middle)*, classification map provided by the proposed method *(right)*.

	Test signal	Accuracy [%]	Refined Accuracy [%]
GROUP 1	sig_1	99.90	99.80
	sig_2	99.93	99.97
	sig_3	99.77	99.78
	sig_4	96.41	96.40
	sig_5	94.85	94.86
GROUP 2	sig_6	88.20	89.71
	sig_7	98.55	98.48
	sig_8	98.63	98.70
	sig_9	99.66	99.64
	sig_{10}	97.43	97.86

Table 2: Classification accuracy, measured in terms of percentage of correct assignments, achieved by the proposed method for spectrograms referring to test signals in Group 1 and Group 2. The third column refers to postprocessed classification maps using median filtering.

5.6. Post processing phase

In order to improve classification results, a median filter has been applied to the classification map in order to correct the classification for sparse and isolated points. In fact, pointwise classification can be misleading in some cases and the local median filter aims at making the classification map spatially homogeneous. The window used for the median filter has size $m \times m$, and $m = 18$ has been used in all tests — it has been fixed by using the 3σ rule to the gaussian analysis window. The improvement in terms of classification rate can be evaluated by looking at the third column of Table 2, while Fig. 10 allows to compare the refined classification maps for signals belonging to Group 2.

5.6.1. Classification under noisy conditions

To evaluate the robustness to noise of the proposed interference detection method, noise having normal distribution and different standard deviations has been added to signals under study. Two kinds of test have been performed:

1. CNN has been trained using clean patches. In this case, we are interested in evaluating to what extent noisy spectrogram can confuse the proposed classification method;

2. CNN has been trained using noisy patches. In this case, the aim is to evaluate if CNN trained on noisy patches increases the classification performance and, if so, the level of such an improvement.

In the first case, chirps in eq. (17) are used for training. In the latter, the same chirps have been corrupted with zero-mean Gaussian noise with predefined Signal to Noise Ratio (SNR), i.e. SNR = 20 db and SNR = 10 db— see Figs. 11 and 12. These SNR values have been selected as they are representative of moderate-high noise.

Fig. 13 shows the results achieved using non noisy and noisy training signals, while some classification maps are reported in Figs. 14 and 15 — post processing has been performed in both cases.

As it can be observed, for test signals having low level of noise (SNR $\geq$ 30 db) recognition rates are almost independent of the presence of noise in the training set: a slight decrease is observed for the noisy training case because of some bad classifications in correspondence to the interference regions. As the noise level increases (SNR < 25 db), classification performed using the noisy-trained CNN gives better results. However, when using the 20db noisy training set, classification rate decreases as SNR decreases, making the method quite robust till SNR values equal to 20 db — for smaller SNR values, classification rates rapidly degrade. On the contrary, classification rates remain high when 10db noisy training set is used, making the method robust to higher noise levels.

6. Conclusions

This paper has presented a novel approach oriented to detect interference regions in the spectrogram image of multicomponent AM-FM signals. To this aim, a machine learning-based strategy has been employed to classify spectrogram samples in three classes: interference, single chirp and background. The proposed method makes use of a small set of spectrogram regions to train the selected CNN architecture. Specifically, the training set is composed of portions of the spectrogram image of a few linear chirps. The rationale under this choice relies on the fact that usually interference regions are very local in the spectrogram. Hence, a local linear model is adequate for the spectrogram of a large class of chirps. Extensive experimental results have shown that the proposed method is able to correctly detect interference regions that involve different kinds of signals, showing a certain robustness to moderate

noise. Future research will be devoted to method improvements and refinements that include optimization of the training set, greater robustness to significant noise, and more accurate detection of very critical interference regions.

7. Acknowledgments

This work has been partially supported by PNRR-CN1_SPOKE_6 Spoke 6 - Multiscale modeling engineering applications B83C22002940006, and by PE0000013-FAIR under the MUR National Recovery and Resilience Plan funded by the European Union - NextGenerationEU. This research has been accomplished within RITA (Research ITalian network on Approximation) and the Italian national research group GNCS (INdAM).

References

[1] Boashash, B. *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*; Academic Press: Cambridge, MA, USA, 2015.

[2] Wang, G.; Teng, C.; Li, K.; Zhang, Z.; Yan, X. The removal of EOG artifacts from EEG signals using independent component analysis and multivariate empirical mode decomposition. *IEEE J. Biomed. Health Inform.* **2015**, *20*, 1301–1308.

[3] Ioana, C.; Gervaise, C.; Stéphan, Y.; Mars, J.I. Analysis of underwater mammal vocalisations using time-frequency-phase tracker. *Appl. Acoust.* **2010**, *71*, 1070-1080.

[4] Chen, H.; Sun, H.; Junejo, N.; Qi, J. Whale Vocalization Classification Using Feature Extraction With Resonance Sparse Signal Decomposition and Ridge Extraction. *IEEE Access* **2019**, *7*.

[5] Guillemain, P.; Kronland-Martinet, R. Characterization of acoustic signals through continuous linear time-frequency representations. *Proc. IEEE* **1996**, *84*, 561–585.

[6] Zeng, F.G.; Nie, K.; Stickney, G.S.; Kong, Y.Y.; Vongphoe, M.; Bhargave, A.; Wei, C.; Cao, K. Speech recognition with amplitude and frequency modulations. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 2293–2298.

[7] Chen, V.C.; Li, F.; Ho, S.S.; Wechsler, H. Micro-Doppler effect in radar: Phenomenon, model, and simulation study. *IEEE Trans. Aerosp. Electron. Syst.* **2006**, *42*, 2–21.

[8] Lyonnet, B.; Ioana, C.; Amin, M.G. Human gait classification using microdoppler time-frequency signal representations. In Proceedings of the 2010 IEEE Radar Conference, Washington, DC, USA, 10–14 May 2010; pp. 915–919.

[9] Zhang, Q.; Yeo, T.S.; Tan, H.S.; Luo, Y. Imaging of a Moving Target With Rotating Parts Based on the Hough Transform. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 291–299.

[10] Shi, Y.; Zhang, D.; Ji, H.; Dai, R. Application of Synchrosqueezed Wavelet Transform in Microseismic Monitoring of Mines. In Proceedings of the 2019 IOP Conference Series: Earth and Environmental Science, Ho Chi Minh City, Vietnam, 25–28 February 2019; IOP Publishing: Bristol, UK, 2019; vol. 384.

[11] Njirjak M, Otovic E, Jozinovic D, Lerga J, Mausa G, Michelini A, Stajduhar I. The Choice of Time–Frequency Representations of Non-Stationary Signals Affects Machine Learning Model Accuracy: A Case Study on Earthquake Detection from LEN-DB Data. *Mathematics* **2022**.

[12] Candes, E.J.; Charlton, P.R.; Helgason, H. Detecting highly oscillatory signals by chirplet path pursuit. *Appl. Comput. Harmon. Anal.* **2008**, *24*, 14–40.

[13] Pham, D.H.; Meignen, S. High-order synchrosqueezing transform for multicomponent signals analysis with an application to gravitational-wave signal. *IEEE Trans. Signal Process.* **2017**, *65*, 3168–3178.

[14] Dragomiretskiy, K.; Zosso, D. Variational mode decomposition. *IEEE Trans. Signal Process.* **2014**, *62*, 531–544, doi:10.1109/TSP.2013.2288675.

[15] Upadhyay, A.; Sharma, M.; Pachori, R.B.; Sharma, R. A Nonparametric Approach for Multicomponent AM–FM Signal Analysis. *Circuits Syst. Signal Process.* **2020**, *39*, 6316–6357.

[16] Mallat, S. *A Wavelet Tour of Signal Processing*; Elsevier: Amsterdam, The Netherlands, 1999.

[17] Bruni, V.; Tartaglione, M.; Vitulano, D. A fast and robust spectrogram reassignment method. *Mathematics* **2019**, *7*, 358.

[18] Bruni, V.; Tartaglione, M.; Vitulano, D. An iterative approach for spectrogram reassignment of frequency modulated multicomponent signals. *Math. Comput. Simul.* **2020**, *176*, 96–119.

[19] Bruni, V.; Tartaglione, M.; Vitulano, D. A Signal Complexity-Based Approach for AM–FM Signal Modes Counting. *Mathematics* **2020**, *8*, 2170.

[20] Rankine, L.; Mesbah, M.; Boashash, B. IF estimation for multicomponent signals using image processing techniques in the time–frequency domain. *Signal Process.* **2007**, *87*, 1234–1250.

[21] Zhang, H.; Bi, G.; Yang, W.; Razul, S.G.; See, C.M.S. IF estimation of FM signals based on time-frequency image. *IEEE Trans. Aerosp. Electron. Syst.* **2015**, *51*, 326–343.

[22] Stankovic, L.; Djurovic, I.; Ohsumi, A.; Ijima, H. Instantaneous frequency estimation by using Wigner distribution and Viterbi algorithm. In Proceedings of the 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, Hong Kong, China, 6–10 April 2003; Volume 6, pp. 6–121.

[23] Djurović, I.; Stanković, L. An algorithm for the Wigner distribution based instantaneous frequency estimation in a high noise environment. *Signal Process.* **2004**, *84*, 631–643.

[24] Li, P.; Zhang, Q.H. IF estimation of overlapped multicomponent signals based on Viterbi algorithm. *Circuits Syst. Signal Process.* **2020**, *39*, 3105–3124.

[25] Li, P.; Zhang, Q.H. An improved Viterbi algorithm for IF extraction of multicomponent signals. *Signal Image Video Process.* **2018**, *12*, 171–179.

[26] Khan, N.A.; Mohammadi, M.; Ali, S. Instantaneous frequency estimation of intersecting and close multi-component signals with varying amplitudes. *Signal Image Video Process.* **2019**, *13*, 517–524.

[27] Chen, S.; Dong, X.; Xing, G.; Peng, Z.; Zhang, W.; Meng, G. Separation of overlapped non-stationary signals by ridge path regrouping and intrinsic chirp component decomposition. *IEEE Sens. J.* **2017**, 5994–6005.

[28] Dong, X.; Chen, S.; Xing, G.; Peng, Z.; Zhang, W.; Meng, G. Doppler Frequency Estimation by Parameterized Time-Frequency Transform and Phase Compensation Technique. *IEEE Sens. J.* **2018**, *18*, 3734–3744.

[29] Zhu, X.; Zhang, Z.; Gao, J.; Li, W. Two robust approaches to multi-component signal reconstruction from STFT ridges. *Mech. Syst. Signal Process.* **2019**, *115*, 720–735.

[30] Stankovic, L.; Katkovnik, V. The Wigner distribution of noisy signals with adaptive time-frequency varying window. *IEEE Trans. Signal Process.* **1999**, *47*, 1099–1108.

[31] P. Pan, Y. Zhang, Z. Deng, S. Fan and X. Huang, TFA-Net: A Deep Learning-Based Time-Frequency Analysis Tool. *IEEE Transactions on Neural Networks and Learning Systems.* **2022**.

[32] Aggarwal Charu C., Neural Networks and Deep Learning: A Textbook, Springer 2018.

[33] Schmidhuber, J. Deep learning in neural networks: An overview. Neural Netw. 2015, 61, 85–117.

[34] H. Firat, M. E. Asker, D. Hanbay, Classification of hyperspectral remote sensing images using different dimension reduction methods with 3D/2D CNN, Remote Sensing Applications: Society and Environment, vol. 25, 2022.

[35] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition. Proc. of ICLR 2015

[36] Cuomo, S., Di Cola, V.S., Giampaolo, F. et al. Scientific Machine Learning Through Physics–Informed Neural Networks: Where we are and What's Next. J Sci Comput 92, 88 (2022)

[37] Josephson, John R.; Josephson, Susan G., eds. (1994). Abductive Inference: Computation, Philosophy, Technology. Cambridge, UK; New York: Cambridge University Press.

[38] Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.C.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. Ser. Math. Phys. Eng. Sci.* **1998**, *454*, 903–995.

[39] Wu, Z.; Huang, N.E. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Adv. Adapt. Data Anal.* **2009**, *1*, 1–41.

[40] Cicone, A. Nonstationary signal decomposition for dummies. In *Advances in Mathematical Methods and High Performance Computing*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 69–82.

[41] P. Flandrin, "Time frequency and chirps," in Wavelet Applications VIII, H. H. Szu, D. L. Donoho, A. W. Lohmann, W. J. Campbell, and J. R. Buss, Eds., vol. 4391, International Society for Optics and Photonics. SPIE, 2001, pp. 161 – 175.

[42] Khan, N.A.; Boashash, B. Multi-component instantaneous frequency estimation using locally adaptive directional time frequency distributions. *Int. J. Adapt. Control Signal Process.* **2016**, *30*, 429–442.

[43] Mohammadi, M.; Pouyan, A.A.; Khan, N.A.; Abolghasemi, V. Locally optimized adaptive directional time–frequency distributions. *Circuits Syst. Signal Process.* **2018**, *37*, 3154–3174.

[44] Carmona, R.A.; Hwang, W.L.; Torresani, B. Characterization of signals by the ridges of their wavelet transform. *IEEE Trans. Signal Process.* **1997**, *45*, 2586–2590.

[45] Carmona, R.; Hwang, W.; Torresani, B. Multiridge detection and time-frequency reconstruction. *IEEE Trans. Signal Process.* **1999**, *47*, 480–492.

[46] Bruni, V., Marconi, S.; Piccoli, B.; Vitulano, D. Instantaneous frequency detection via ridge neighbor tracking. In Proceedings of the 2010 2nd International Workshop on Cognitive Information Processing, Elba, Italy, 14–16 June 2010.

[47] Bruni, V., Tartaglione, M., Vitulano, D., On the time-frequency reassignment of interfering modes in multicomponent fm signals, Proceedings of the 2018 26th European Signal Processing Conference (EUSIPCO), Rome, Italy, 3–7 September 2018. IEEE, pp. 722–726.

[48] S. Meignen and N. Singh, Analysis of reassignment operators used in synchrosqueezing transforms: With an application to instantaneous frequency estimation, *IEEE Transactions on Signal Processing*, 70, 216–227, 2021.

[49] N. A. Khan and I. Djurovic, ADTFD-Ransac for multi-component if estimation, Signal Processing, 2022.

[50] Stankovic, L. Analysis of noise in time-frequency distributions.*IEEE Signal Process. Lett.* **2002**, *9*, 286–289,.

[51] Chen, S.; Dong, X.; Peng, Z.; Zhang, W.; Meng, G. Nonlinear chirp mode decomposition: A variational method. *IEEE Trans. Signal Process.* **2017**, *65*, 6024–6037.

[52] Yang, Y.; Dong, X.; Peng, Z.; Zhang, W.; Meng, G. Component extraction for non-stationary multi-component signal using parameterized de-chirping and band-pass filter. *IEEE Signal Process. Lett.* **2015**, *22*, 1373–1377.

[53] Stankovic, L.; Dakovic, M.; Thayaparan, T.; Popovic-Bugarin, V. Inverse radon transform–based micro-Doppler analysis from a reduced set of observations. *IEEE Trans. Aerosp. Electron. Syst.* **2015**, *51*, 1155–1169.

[54] Bruni, V.; Marconi, S.; Piccoli, B.; Vitulano, D. Instantaneous frequency estimation of interfering FM signals through time-scale isolevel curves. *Signal Process.* **2013**, *93*, 882–896.

[55] Stanković, L.; Mandić, D.; Daković, M.; Brajović, M. Time-frequency decomposition of multivariate multicomponent signals. *Signal Process.* **2018**, *142*, 468–479.

[56] Ding, Y.; Tang, J. Micro-Doppler trajectory estimation of pedestrians using a continuous-wave radar. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 5807–5819.

[57] Chen, S.; Yang, Y.; Wei, K.; Dong, X.; Peng, Z.; Zhang, W. Time-varying frequency-modulated component extraction based on parameterized demodulation and singular value decomposition. *IEEE Trans. Instrum. Meas.* **2015**, *65*, 276–285.

[58] Yang, Y.; Peng, Z.; Dong, X.; Zhang, W.; Meng, G. Application of parameterized time-frequency analysis on multicomponent frequency modulated signals. *IEEE Trans. Instrum. Meas.* **2014**, *63*, 3169–3180.

[59] Ioana, C.; Jarrot, A.; Gervaise, C.; Stéphan, Y.; Quinquis, A. Localization in underwater dispersive channels using the time-frequency-phase continuity of signals. *IEEE Trans. Signal Process.* **2010**, *58*, 4093–4107.

[60] Chen, S.; Dong, X.; Yang, Y.; Zhang, W.; Peng, Z.; Meng, G. Chirplet path fusion for the analysis of time-varying frequency-modulated signals. *IEEE Trans. Ind. Electron.* **2016**, *64*, 1370–1380.

[61] Swärd, J.; Brynolfsson, J.; Jakobsson, A.; Hansson-Sandsten, M. Sparse semi-parametric estimation of harmonic chirp signals. *IEEE Trans. Signal Process.* **2015**, *64*, 1798–1807.

[62] Djurović, I.; Simeunović, M.; Wang, P. Cubic phase function: A simple solution to polynomial phase signal analysis. *Signal Process.* **2017**, *135*, 48–66.

[63] Zhu, X.; Zhang, Z.; Zhang, H.; Gao, J.; Li, B. Generalized Ridge Reconstruction Approaches Toward more Accurate Signal Estimate. *Circuits Syst. Signal Process.* **2019**, 1–26.

[64] Bouchikhi, A.; Boudraa, A.O.; Cexus, J.C.; Chonavel, T. Analysis of multicomponent LFM signals by Teager Huang-Hough transform. *IEEE Trans. Aerosp. Electron. Syst.* **2014**, *50*, 1222–1233.

[65] Barbarossa, S. Analysis of multicomponent LFM signals by a combined Wigner-Hough transform. *IEEE Trans. Signal Process.* **1995**, *43*, 1511–1515.

[66] Wood, J.C.; Barry, D.T. Radon transformation of time-frequency distributions for analysis of multicomponent signals. *IEEE Trans. Signal Process.* **1994**, *42*, 3166–3177.

[67] Alieva, T.; Bastiaans, M.J.; Stankovic, L. Signal reconstruction from two close fractional Fourier power spectra. *IEEE Trans. Signal Process.* **2003**, *51*, 112–123.

[68] Bruni, V.; Tartaglione, M.; Vitulano, D. Radon spectrogram-based approach for automatic IFs separation. *EURASIP J. Adv. Signal Process.* **2020**, *13*, 1–21.

[69] Khan, N.A.; Mohammadi, M.; Djurović, I. A Modified Viterbi Algorithm-Based IF Estimation Algorithm for Adaptive Directional Time–Frequency Distributions. *Circuits Syst. Signal Process.* **2019**, *38*, 2227–2244.

[70] Auger, F.; Flandrin, P.; Lin, Y.-T.; McLaughlin, S.; Meignen, S.; Oberlin, T.; Wu, H.-T. Time-Frequency reassignment and synchrosqueezing: An overview. *IEEE Signal Process. Mag.* **2013**, *30*, 32–41.

[71] Daubechies, I.; Lu, J.; Wu, H.T. Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool. *Appl. Comput. Harmon. Anal.* **2011**, *30*, 243–261.

[72] Daubechies, I.; Wang, Y.; Wu, H. ConceFT: Concentration of frequency and time via a multitapered synchrosqueezed transform. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2015**, *374*.

[73] Yu, G.; Yu, M.; Xu, C. Synchroextracting transform. *IEEE Trans. Aerosp. Electron. Syst.* **2017**, *64*, 8042–8054.

[74] Zhu, X.; Yang, H.; Zhang, Z.; Gao, J.; Liu, N. Frequency-chirprate reassignment. *Digit. Signal Process.* **2020**.

[75] Bruni, V.; Tartaglione, M.; Vitulano, D., A pde-based analysis of the spectrogram image for instantaneous frequency estimation, Mathematics, vol. 9, no. 3, 2021

[76] Bruni, V.; Tartaglione, M.; Vitulano, D., Skeleton-based reassignment of nonstationary signals spectrogram, Digital Signal Processing, vol. 128, 2022

[77] Doweck, Y.; Amar, A.; Cohen, I. Joint model order selection and parameter estimation of chirps with harmonic components. *IEEE Trans. Signal Process.* **2015**, *63*, 1765–1778.

[78] Feng, Z.; Chu, F.; Zuo, M.J. Time–frequency analysis of time-varying modulated signals based on improved energy separation by iterative generalized demodulation. *J. Sound Vib.* **2011**, *330*, 1225–1243.

[79] V. Sucic, N. Saulig, and B. Boashash, Estimating the number of components of a multicomponent nonstationary signal using the short-term time-frequency Renyi entropy, *EURASIP J. Adv. Signal Proc.*, no. 1, 2011.

[80] Bruni, V.; Della Cioppa, L.; Vitulano, D. A Multiscale Energy-Based Time-Domain Approach for Interference Detection in Non-stationary Signals. In *International Conference on Image Analysis and Recognition*; Springer: Berlin/Heidelberg Germany, 2020; pp. 36–47.

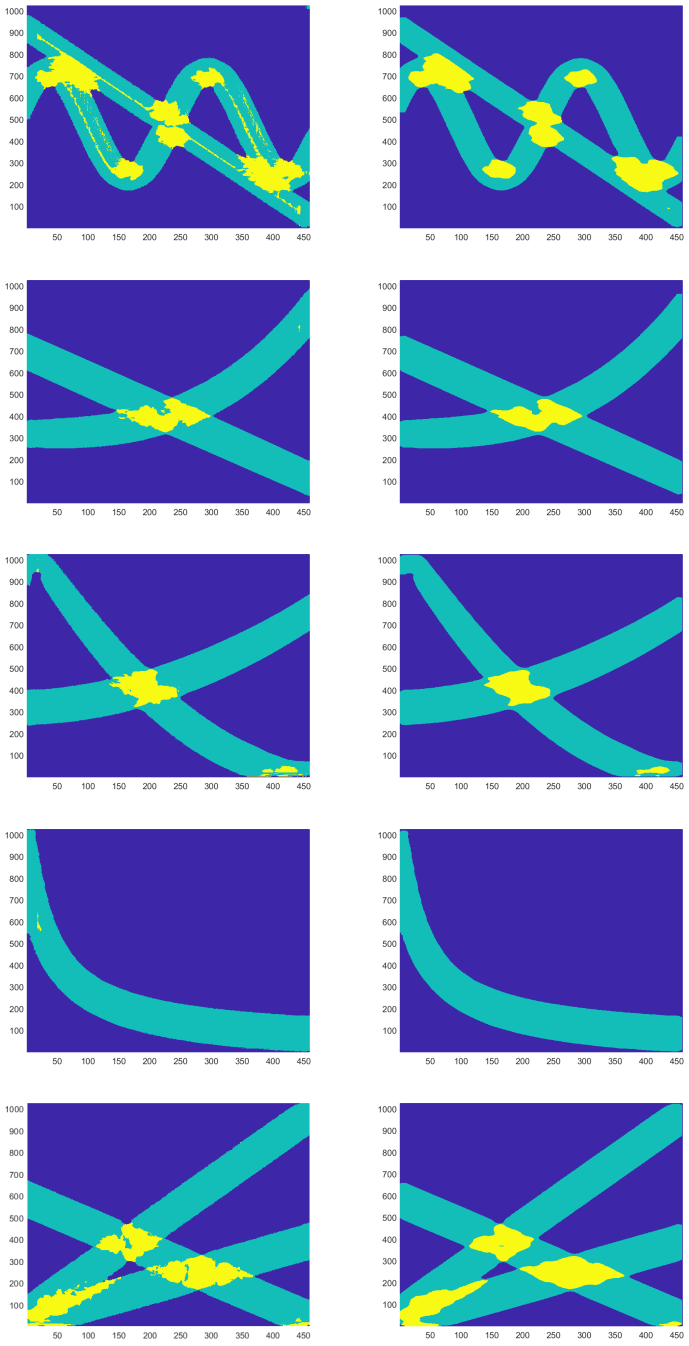

Figure 10: Classification map for the chirp signals in Group 2: without post processing *(left)*, with post processing *(right)* – see text for details.

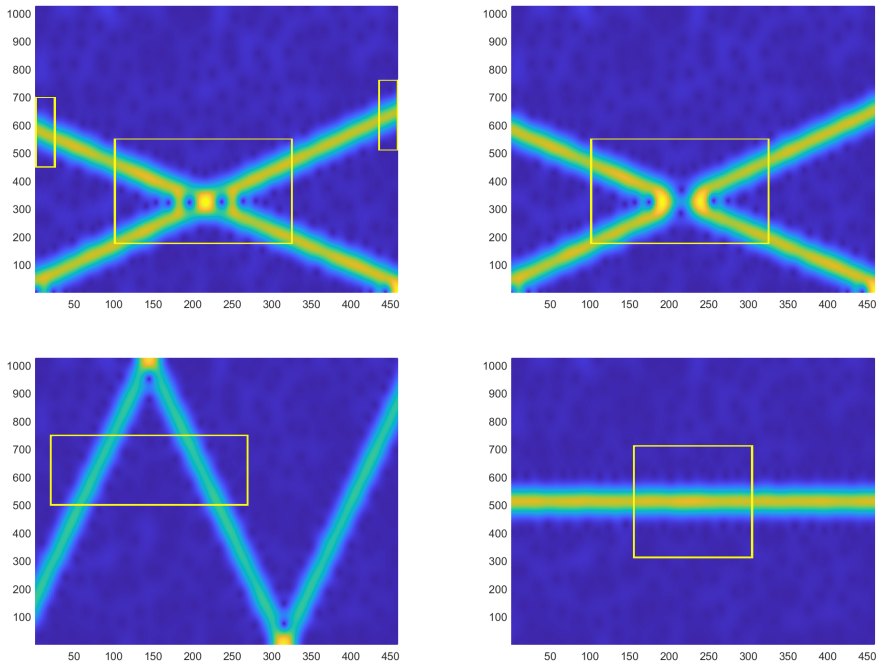

Figure 11: Selected regions in the spectrogram of corrupted training signals defined in eq. (17) — `SNR = 20`db.

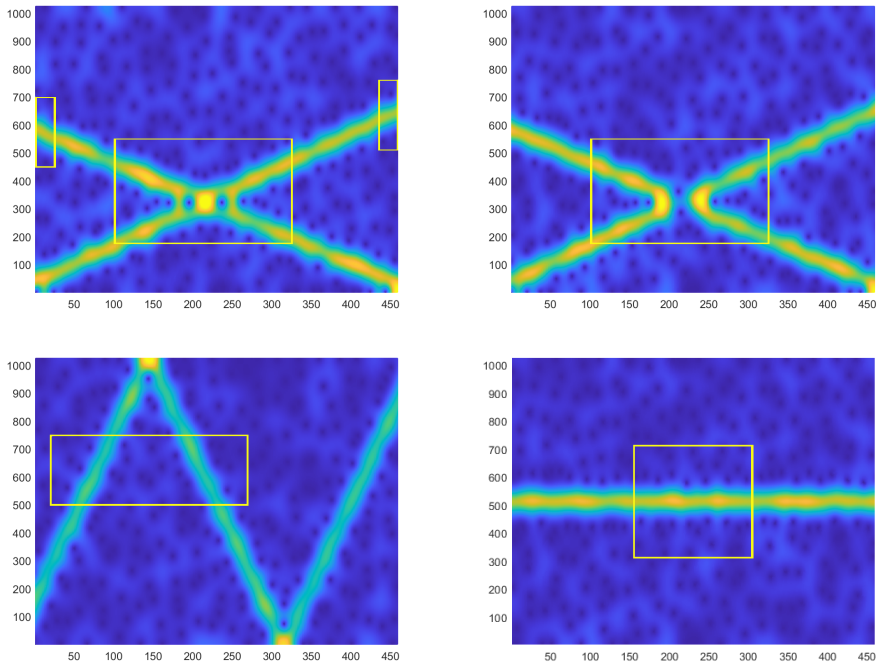

Figure 12: Selected regions in the spectrogram of corrupted training signals defined in eq. (17) – SNR = 10 db.

	Non noisy training	Noisy training (SNR = 20 db)	Noisy training (SNR = 10 db)
SNR	Accuracy [%]	Accuracy [%]	Accuracy [%]
//	97.86	97.44	95.67
30	97.71	97.35	96.45
25	96.99	97.23	96.64
22	93.32	97.12	96.93
20	87.49	96.91	97.24
18	77.56	96.16	97.45
15	57.86	85.60	97.54
10	29.48	44.21	97.14

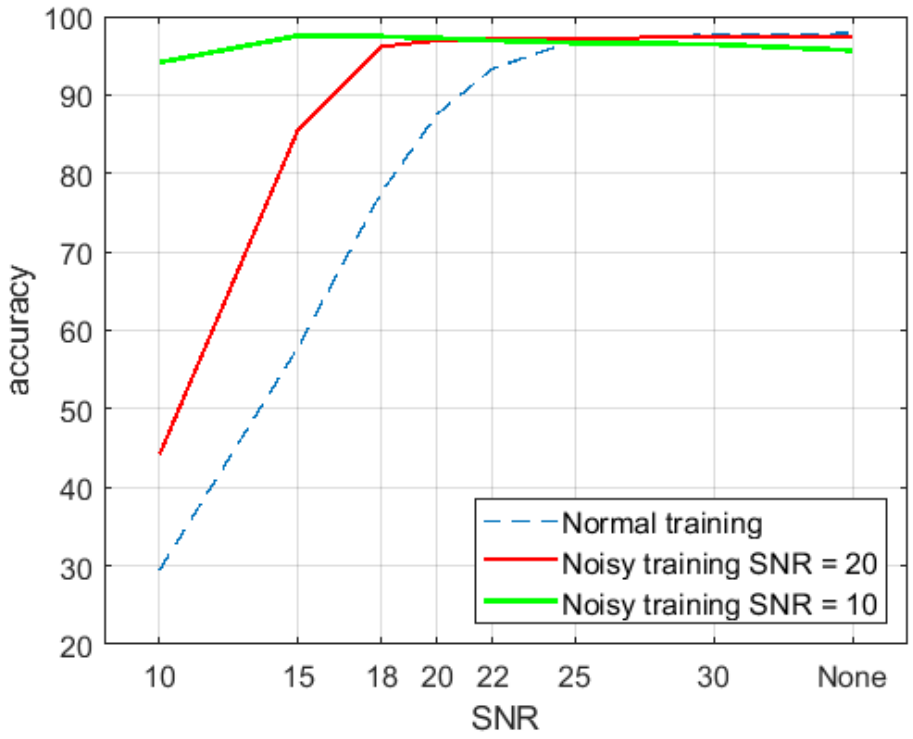

Figure 13: Three linear modes test signal (sig_{10} in Group 2) classification accuracies, measured in terms of percentage of correct assignments, provided by the proposed method using non noisy and noisy training signals with SNR = 20 db and SNR = 10 db. sig_{10} has been corrupted with Gaussian noise with different SNRs.

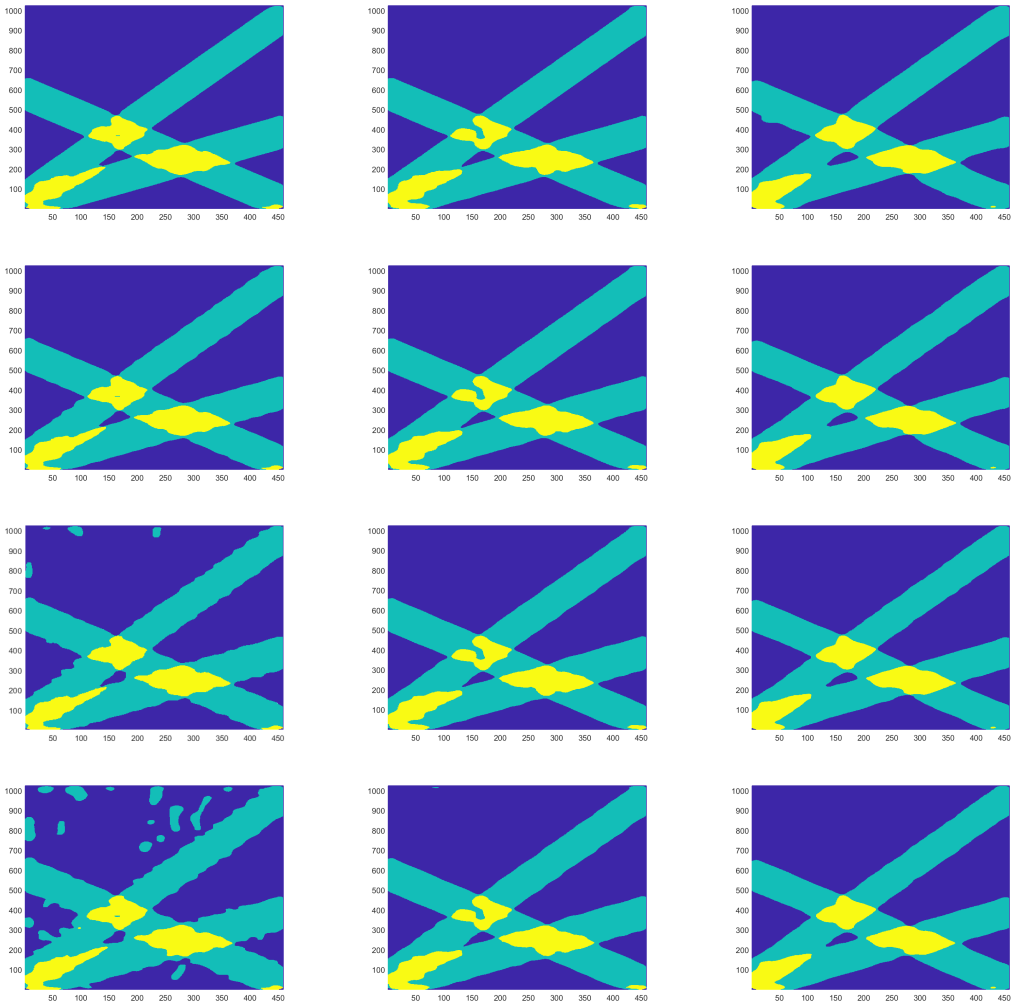

Figure 14: to be continued

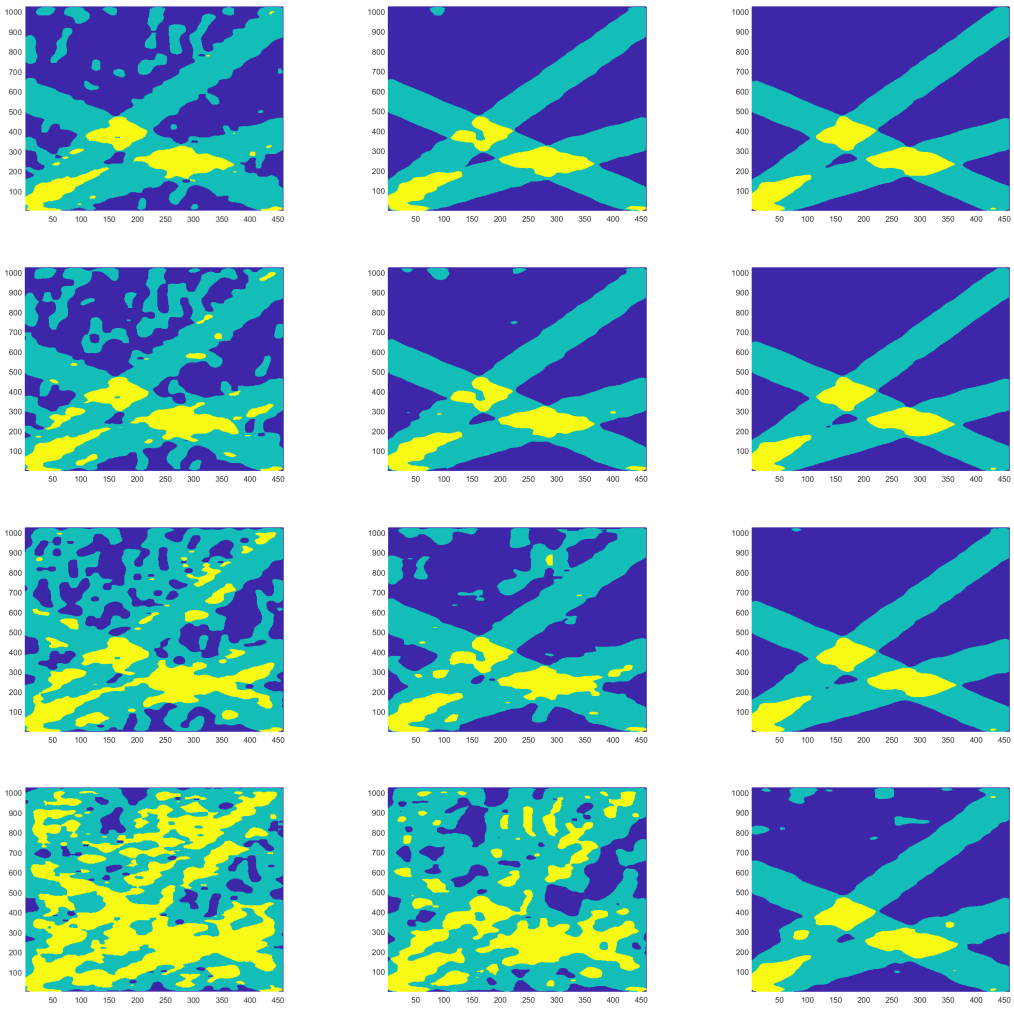

Figure 15: **From top to bottom)** Classification maps for the spectrogram of the three mode interfering signal (sig_{10} in Group 2) and the same signal corrupted with noise with SNR = 30, 25, 22, 20, 18, 15, 10 db. Training set without *(left)* and with noise with SNR = 20db *(middle)*, with noise with SNR = 10db *(right)*.