

Automatic identification of hallucinogenic amphetamines based on their ATR-FTIR spectra processed with Convolutional Neural Networks

Catalin Negoita¹, Mirela Praisler^{1*}, and Iulia-Florentina Darie¹

¹ "Dunărea de Jos" University of Galati, Faculty of Science and Environment, Domnească Street, 47, RO-800008, Galati, Romania

Abstract. New psychoactive drugs that are leading to severe intoxications are constantly seized on the European black market. Recent studies indicate that most of these new substances are synthetic cannabinoids and hallucinogenic amphetamines. In this study, we are presenting the results obtained with an expert system that was built to identify automatically the class identity of these types of drugs of abuse, based on their Attenuated Total Reflection-Fourier Transform Infrared (ATR-FTIR) spectra processed with Convolutional Neural Networks (CNNs). CNNs have been applied with great success in recent years in various computer applications, such as image classification, but little work has been done in using this kind of deep learning models for spectral data classification. The aim of this study was to improve the detection accuracy (classification performance) that we have already obtained with other statistical mathematics and artificial intelligence techniques. The performances of the CNN system are discussed in comparison with those of the later models.

1 Introduction

New powerful detection models are highly needed due to the pressure on European Drug Investigation Departments related to the increasing emergence of new hallucinogenic substances on the European black market [1]. In this paper, we are presenting such a model, which was built in order to improve the detection efficiency in the case of new synthetic 2C-x and DOx amphetamines. The detection is based on the Attenuated Total Reflection-Fourier Transform Infrared (ATR-FTIR) spectra of the targeted drugs of abuse and their automatic analysis by machine learning models that are using representations of the underlying explanatory factors hidden in the spectral data, i.e. One-dimensional Convolutional Neural Networks (CNNs). These models are designed to consider spatial information of the input data and exploit spatially-local correlation by enforcing a local connectivity pattern between neurons of adjacent layers.

* Corresponding author: Mirela Praisler, Mirela.Praisler@ugal.ro

The aim of this study was to increase the accuracy of formerly developed models by choosing an appropriate pre-processing method and the most robust multivariate model. Hence, the efficiency of the CNN model has been assessed in comparison with other artificial intelligence applications screening for drugs of abuse that we have formerly developed, e.g. Partial Least Squares Regression (PLSR) [2], Partial Least Squares Discriminant Analysis coupled with Genetic Algorithms (PLS-DA) [3, 12], and class modelling solutions, such as Soft Class Analog Independent Modelling (SIMCA) [4-6], Principal Component Analysis (PCA) [7, 8] and Artificial Neural Networks (ANNs) [9-11].

2 (1D) One-dimensional Convolutional Neural Networks

CNNs mimic the functioning of the visual cortex in the animal's brain, where it performs hierarchies of characteristics from the input data, such as a signal or an image with broad range of patterns. Mathematically, the convolution is based on filtering the input data and identifying similar places between the chosen filter and the input data in the model. Similar to the brain structure, CNNs are ANNs with different layers of convolutional and alternating samples.

In the case of a spectrum, a one-dimensional convolution practically consists in applying and sliding the windows of a filter over the analyzed spectrum. In a CNN model, these filter values are substituted by learnable parameters (also known as network weights) called “kernels”. The kernel values are determined during the network training and depend on the features of the input data. CNNs can be used for any dimension and size, but the usual dimensions are up to three. Such a model architecture has been successfully used in the case of 1D CNNs built for signals and time series [13], 2D CNNs built for processing images (pixel-level) and matrices [14], as well as 3D CNNs assessing medical imaging databases [15]. A densely connected layer, also termed fully connected layer (FC layer), maps the input vector into another one. Neurons between two adjacent layers are pairwise connected, as shown in Figure 1 [16].

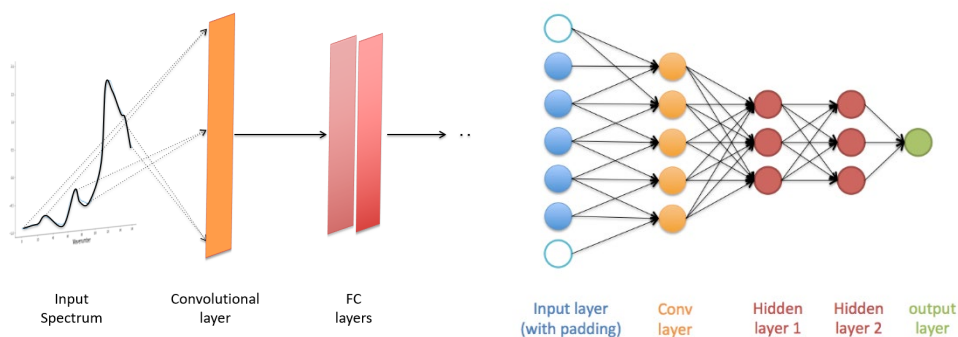


Fig. 1. Architecture of a convolutional neural network [16].

A typical CNN model can extract features from the input data with various levels of abstraction. As a deep neural network, the main components of a hierarchically layered CNN are convolutional blocks, which can extract different types of patterns present in the input by using various filters. Then, pooling layers extract the essential features from the previous layers [17]. This way of extracting features by different abstractions and sharing information from layer to layer creates a receptive field that helps the CNN model to be almost invariant to spatial transformations with a cheaper computational cost.

For classification tasks, this network arrangement is called encoder block. The output of encoders consists of the features extracted from the input data, such as features of the input ATR-FTIR spectra. For this reason, in order to convert encoders into a classifier, it is common to append fully connected layers to the last layer of the network, followed by the use of nonlinear functions such as *SoftMax* or *Sigmoid*. To avoid overfitting, batch normalization and dropout layers are added between layers.

During the training stage of a CNN model, the kernels of the network are first initialized randomly or by using other initialization techniques. This process of initialization can also be accomplished by using weights from a pre-trained model, which is previously trained on a different dataset called transfer learning. Initialized kernels are then optimized by using optimization algorithms (such as *Adam* or *SGD*) and the backpropagation technique.

In conclusion, 1D CNN is a promising method for data mining of 1D signals where only a limited number of data is available and the whole signal information should be considered as the input data. 1D CNNs have recently been considered in successful applications such as personalized medical data classification and fault detection, as well as the identification in power electronics [18]. The main advantage of this technique over the classical ANNs is that 1D CNN extracts the features of a signal by considering local information instead of the whole signal in each network layer. This approach results in a smaller number of trainable parameters and hence a faster training of the network, which requires less computational costs [17].

3 Experimental part

The dataset consisted of the normalized ATR-FTIR spectra of 60 compounds of forensic interest. The spectra have been recorded in the 4000 - 400 cm^{-1} spectral window, 5 cm^{-1} apart, by calculating the average of 1868 scans (see Figure 2). These spectra are made available by the Drug Enforcement Administration (DEA) USA. The samples belong to three sets of drugs: 28 compounds representing 2C-x and DOx amphetamines (Class I), 19 cannabinoids (Class II), and 13 other randomly selected substances (Class III), as presented in Table 1.

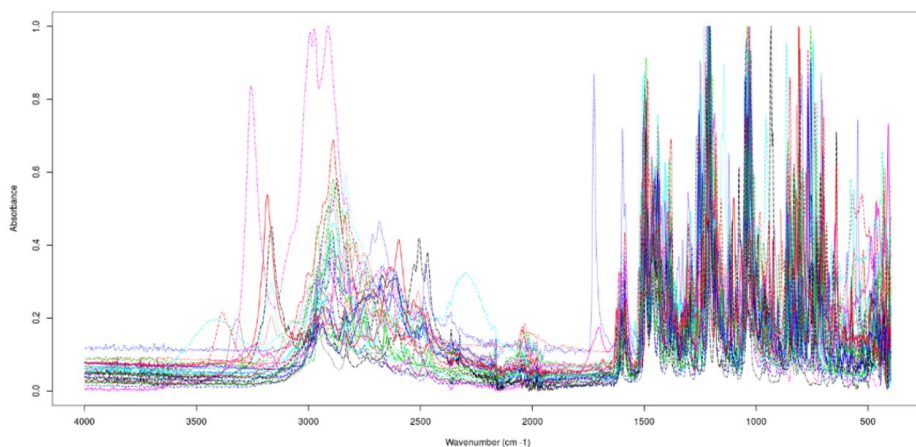


Fig. 2. ATR-FTIR spectra of the analysed compounds.

The main characteristic of the 2C-x amphetamines is that they have methoxy groups on the 2 and 5 positions of the benzene ring. In the case of the DOX amphetamines, this ring is also substituted, but by an alkyl group or a halogen at the 4- position. The dataset was randomly split into a training and a testing subset.

Table 1. Subsets of substances

No	Compound	Compound code		Class
		Training set	Testing set	
1	25B-NBOMe HCl	1		I
2	25C-NB3OMe HCl	2		I
3	25C-NB4OMe HCl	3		I
4	25C-NBOMe HCl		1	I
5	2,5-Dimethoxy-4-Chloro-amphetamine HCl	4		I
6	2,5-Dimethoxy-4-ethylamphetamine HCl	5		I
7	2,5-Dimethoxy-4-methylamphetamine HCl	6		I
8	2,5-Dimethoxyamphetamine HCl	7		I
9	2,5-Dimethoxyphenethylamine HCl	8		I
10	25D-NBOMe HCl	9		I
11	25E-NBOMe HCl	10		I
12	25H-NBOMe HCl	11		I
13	25I-NB3OMe HCl		2	I
14	25I-NB4OMe HCl		3	I
15	25I-NBOMe Base	12		I
16	25I-NBOMe HCl	13		I
17	2C-B BZP diHCl	14		I
18	2C-B HCl		4	I
19	2C-E HCl	15		I
20	2C-I HCl		5	I
21	2C-T-2 HCl		6	I
22	2C-T-7 HCl	16		I
23	3,4-Dimethoxyamphetamine HCl	17		I
24	5-Methoxy-alpha-methyltryptamine HCl		7	I
25	5-Methoxy-Diallyltryptamine HCl	18		I
26	5-Methoxy-N,N-diethyltryptamine HCl	19		I
27	4-Bromo-2,5-dimethoxyamphetamine HCl)		8	I
28	Methoxetamine HCl	20		I
29	AM1220		9	II
30	AM1248		10	II
31	AM2233		11	II
32	JWH-018 adamantyl-carboxamide	21		II
33	JWH-018 Benzimidazole	22		II
34	JWH-018	23		II
35	JWH-018 N-(5-chloropentyl) analog	24		II
36	JWH-019	25		II
37	JWH-022	26		II
38	JWH-073	27		II
39	JWH-081	28		II
40	JWH-122	29		II
41	JWH-200		12	II
42	JWH-203	30		II
43	JWH-210		13	II
44	JWH-250		14	II
45	JWH-307	31		II
46	RCS-4	32		II

47	RCS-8	33		II
48	bk-MDDMA HCl	34		III
49	Bufotenine Oxalate Hydrate	35		III
50	Buphedrone HCl	36		III
51	Butylone HCl		15	III
52	BZP diHCl		16	III
53	Cathine Base	37		III
54	Cathine HCl	38		III
55	Cathinone HCl		17	III
56	CB-13	39		III
57	CP 47,497 C8 homologue	40		III
58	CP 47,497	41		III
59	dihydro-PPP HCl		18	III
60	Diisopropyltryptamine HCl	42		III

4 Results and discussion

The CNN model was built in the *R* environment, by using the *tensorflow* and *keras* libraries. A sequential 1D convolution layer is activated via a *softmax* function and multinomial logistic regression that are used to draw a decision boundary necessary to determine which class is likely to be present. The model was compiled by using accuracy metrics and a loss set as categorical cross-entropy.

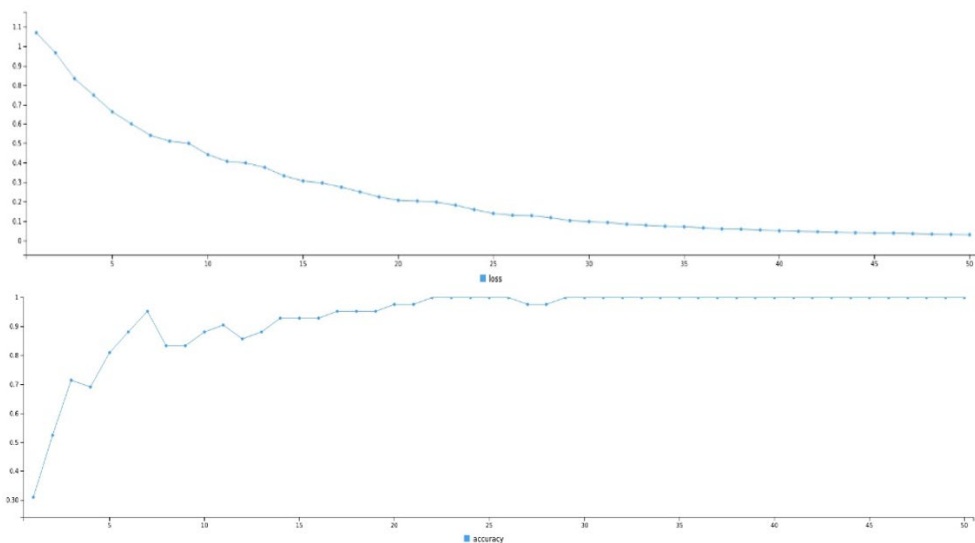


Fig. 3. Compilation and training of the 1D CNN model

After the model compilation, the best accuracy was obtained after approximately 30 epochs (see Figure 3). The overall accuracy is 0.8333. As the Confusion matrix of the 1D CNN model (presented in Table 2) indicates, the most notable performance is that the model correctly classifies all the Class I substances, i.e. all the tested 2C-x and DOx amphetamines. All the negatives (Class III compounds) are also correctly classified. On the other hand, 3 substances belonging to Class II (cannabinoids) are incorrectly classified as members of Class I. In other words, the model is remarkably sensitive, but less selective from the point

of view of discriminating between Class I substances (2C-x and DOx amphetamines) and Class II substances (cannabinoids).

However, this shortcoming is not essential, as both amphetamines and cannabinoids are controlled substances. The characteristic of essence of any forensic tool is its sensitivity, especially in cases of models screening for drugs of abuse such as the one described above. Once the model classifies a substance in one of the classes of positives (Class I or Class II), that compound will be subjected to further laboratory investigation for individual identification.

Table 2. 1D CNN Confusion Matrix

Prediction	Reference		
	Class I	Class II	Class III
Class I	8	0	0
Class II	3	3	0
Class III	0	0	4

The predictions results obtained for each compound included in the testing subset are presented in Table 3. The incorrectly classified cannabinoids are the compounds with the test set code 9, 10 and 11, i.e. AM1220 ((R)-1-((1-Methylpiperidin-2-yl)methyl)-1H-indol-3-yl)(naphthalen-1-yl)methanone), AM1248 (1-[(N-methylpiperidin-2-yl)methyl]-3-(adamant-1-oyl)indole) and AM2233 (1-[(N-methylpiperidin-2-yl)methyl]-3-(2-iodobenzoyl)indole) (see Table 2). We should notice that no AM cannabinoid has been included in the training set, due to the small number of available spectra of these highly potent full agonist for the cannabinoid receptors. It is fair to expect that, once more AM spectra will become available to be included in the training set, a significant improvement of the CNN model selectivity will be obtained.

Table 3. 1D CNN Prediction probability

TESTING SET Compound code	Prediction probability for Class I (%)	Prediction probability for Class II (%)	Prediction probability for Class III (%)	Reference class
1	99.98	0.02	0.01	I
2	74.47	14.69	10.84	I
3	98.59	1.10	0.31	I
4	94.02	5.87	0.11	I
5	89.50	9.94	0.56	I
6	95.66	4.32	0.02	I
7	91.45	8.53	0.02	I
8	98.54	1.43	0.04	I
9	99.88	0.02	0.10	II
10	98.66	1.28	0.06	II
11	90.78	8.71	0.52	II
12	0.23	97.26	2.52	II
13	0.03	99.31	0.66	II
14	25.00	42.21	32.79	II
15	0.03	1.33	98.64	III
16	0.90	24.06	75.04	III
17	1.18	0.90	97.92	III
18	1.75	2.16	96.09	III

5 Conclusions

The availability of rapid and reliable analytical tools screening for controlled substances such as psychedelic amphetamines, cannabinoids or their precursors is a vital necessity for law enforcement teams. ATR-FTIR spectrometers are very adequate for a fast *in situ* screening for drugs of abuse, as these are portable instruments that record the spectra of the analysed compounds in a very short period of time and do not need any sample preparation for the analysis. However, unlike the case of laboratory spectrometers, only few artificial intelligence applications automatizing the detection of amphetamines and cannabinoids with portable spectrometers have been developed. The 1D CNN model presented in this paper can operate handheld ATR-FTIR spectrometers. The results indicate that it may significantly enhance their analytical capacity, by automatizing the spectra processing and improve the screening efficiency in a user-friendly manner.

In comparison with formerly developed artificial intelligence systems screening for similar drugs of abuse, the CNN system described in this paper has a very good efficiency for several reasons. First, as opposed to the systems built with GC-FTIR spectra [5-8, 10, 11], the CNN system had to overcome the disadvantage of the ATR-FTIR spectra in terms of signal to noise ratio. The ATR-FTIR spectra may be easier and faster to record than the GC-FTIR spectra, but they have a significantly weaker intensity than the later. Secondly, as opposed to other systems built with ATR-FTIR spectra [12], the CNN system operates in a simpler (and hence faster) way, as no variable selection is needed before the classification itself. Last but not the least, the CNN system was built with a complex architecture, aiming to test simultaneously the class identity of an unknown against two classes of positives, i.e. Class I (2C-x and DOx amphetamines) and Class II substances (cannabinoids). Consecutive binary testing might further improve the performances of the system. We intend to build and test such alternative systems and the results will be soon reported

References

1. European Monitoring Centre for Drugs and Drug Addiction, *European Drug Report 2020 - Trends and Developments* (Publications Office of the European Union, Luxembourg, 2020)
2. V. Esposito Vinzi, W.W. Chin, J. Henseler, H. Wang, (Eds.), *Handbook of Partial Least Squares - Concepts, Methods and Applications* (Springer, Berlin, 2010)
3. A. L. Pomerantsev, O. Ye. Rodionova, *J Chemom* **32** (8), e3030 (2018)
4. Y. Lee, S.-H. Han, S.-H. Nam, *Appl Spectrosc* **71**, 3702817697337 (2017)
5. M. Praisler, J. Van Bocxlaer, A. De Leenheer, D.L. Massart, *Turk J Chem* **26**, 45-58 (2002)
6. M. Praisler, I. Dirinck, J. Van Bocxlaer, A. De Leenheer, D.L. Massart, *Talanta* **53**, 155-170 (2000)
7. S. Gosav, M. Praisler, D.O. Dorohoi, G. Popa, *J Mol Struct* **744-747**, 821-825 (2005)
8. M. Praisler, S.Ciochina, *J Anal Methods Chem* **2014**, 342497 (2014)
9. M. Hassoun, *Fundamentals of Artificial Neural Networks* (MIT Press, Cambridge USA, 2003)
10. S. Gosav, M. Praisler, D.O. Dorohoi, *J Mol Struct* **834-836**, 188-194 (2007)
11. S. Gosav, M. Praisler, D.O. Dorohoi, G. Popa, *Talanta* **70**, 922-928 (2006)
12. C. Negoita, M. Praisler, A. Ion, *AIP Conference Proceedings* **2075**, 170026 (2019)
13. H. I. Fawaz et al., *Data Min Knowl Disc* **33**, 917-963 (2019)
14. A. Krizhevsky, I. Sutskever, G. E. Hinton, *Communications of the ACM* **60**, 84-90 (2017)
15. S. P. Singh et al., *Sensors* **20** (18), 5097 (2020)
16. C. Chenhao, T. Fearn, *Chemometr Intell Lab Syst* **182**, 9-20 (2018)
17. M. H. Mozaffari, L.-L. Tay, arXiv:2006.10575 (2020)
18. S. Kiranyaz et al., *Mech Syst Signal Process* **151** 107398 (2021)