analytical chemistry

pubs.acs.org/ac Article

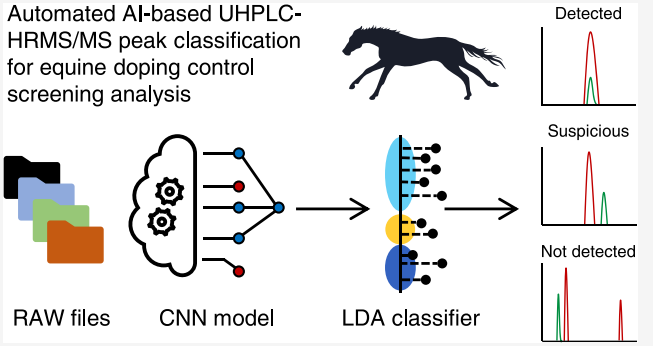
Convolutional Neural Networks Assisted Peak Classification in Targeted LC-HRMS/MS for Equine Doping Control Screening Analyses

Agnès Barnabé,[†] Vivian Delcourt,[†] Benoit Loup, William Montanuy, Stéphane Trévisiol, Marie-Agnès Popot, Patrice Garcia,* and Ludovic Bailly-Chouriberry

Cite This: *Anal. Chem.* 2025, 97, 3236−3241 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Doping control screening analyses usually involve visual inspection of extracted ion chromatograms (EIC) by a trained analytical chemist, followed by further investigations if needed. This task is both highly repetitive and time-consuming, given the hundreds of compounds and metabolites to be screened in tens of thousands of samples per year. With the recent widespread adoption of machine learning in analytical chemistry and the training of high-performance convolutional neural networks (CNN), these operations can be automated with high accuracy and throughput. Applying this technology to doping control is challenging as the false negative rate (FNR) shall be equal to zero. In this study, we demonstrated that implementing a deep learning strategy for chromatogram classification in equine doping control can be feasible and accurate. We illustrated our findings with a CNN scoring model combined with a linear discriminant analysis (LDA) classifier trained on chromatogram images from our ultra-high-pressure liquid chromatography coupled to high-resolution tandem mass spectrometry (UHPLC-HRMS/MS)-based biotherapeutics screening method. We expect that artificial intelligence (AI) will be a valuable tool for doping control laboratories in the near future.



INTRODUCTION

In horseracing and equestrian competitions, the administration of substances to horses is highly controlled, and misuse during an event can lead to severe sanctions. Screening analyses in doping control laboratories search for putative drug's presence in biological samples (plasma, urine, hair), leading them to confirmatory analyses in case of detection following association of racing chemists (AORC) criteria. Modern direct or indirect doping control analyses mostly employ gaseous or liquid chromatography hyphenated to mass spectrometry (MS) to produce multidimensional data. Each target substance's specific signals, characterized by retention times and m/z ratios, are searched in acquired data. This approach requires a trained analytical chemist to visually inspect every single extracted ion chromatogram (EIC) of targeted quasi-molecular ions in the case of MS analysis or fragment ions in the case of MS/MS experiments. However, this task is highly repetitive and time-consuming, given the critical number of samples and the screening of thousands of target substances.

Deep learning approaches have recently been largely adopted in the most challenging chemistry fields, including drug discovery,[1] molecular docking,[2] protein structure prediction,[3] and have been successfully applied to LC-MS/MS analysis for small molecules or peptides.[4−6] More precisely, convolutional neural networks (CNN hereafter) are specifically designed for image classification and have been adapted to chemistry applications such as chemical properties' prediction from Simplified Molecular-Input Line-Entry System (SMILES)[7−9] or chromatography peak detection.[10−12] This breakthrough technology enables us to distinguish accurate signals from noise by analyzing, for instance, the chromatogram's shape, the centering of a peak in an extracted retention time window, or the signal-to-noise ratio. The parameters of the CNN model are highly dependent on the training data provided. Peak classification can depend on the window's width, the correction of the retention time deviation, the selection of minimum intensity, the data quality, or the LC and MS resolutions. Each data set leads to a different trained CNN model and, therefore, to varying predictions for the same batch of EICs.

Received: July 12, 2024
Revised: January 6, 2025
Accepted: January 25, 2025
Published: February 4, 2025

https://doi.org/10.1021/acs.analchem.4c03608
Anal. Chem. 2025, 97, 3236−3241

Applying an AI-based tool to classify EIC in equine doping control requires specific attention and training to avoid false negative predictions. To successfully classify LC-MS/MS signals while complying with doping control high standards, the algorithm presented here relies on a CNN model to evaluate peak shape and analytical parameters such as the number of points in the peak of interest (if there is one), its area, and the deviation of the retention time. The CNN model was designed to analyze two inputs representing EICs of two transitions per targeted substance, just as a trained analytical chemist would during the screening stage. Linear discriminant analysis (LDA) was employed to combine the CNN score and contextual analytical parameters. The combination of CNN scoring and LDA classification could provide a fast and accurate complementary approach to visual inspection by a trained chemist while maintaining a false negative rate (FNR) of 0%. This study demonstrates that an artificial intelligence-based framework could allow the trained chemist to obtain high analysis throughput with the highest quality standard for equine doping control analysis.

MATERIAL AND METHODS

Training and Test Data Set. The data set used to train and test the classifier came from the validation study of a previously published method described by Pinetre et al.[13] to screen monoclonal antibodies (mAbs) and Fc-fusion proteins. Twenty-four plasma samples were collected from nine females, nine geldings, and six males (3−10 years old). Samples were spiked with 8 different concentrations of human mAb "SILuLite", a human IgG reference standard (0, 5, 10, 50, 100, 500, 1000, and 5000 ng/mL), and 1000 ng/mL of "SILuMAB", a stable isotope-labeled universal mAb human standard, used as an internal standard (ISTD). Both "SILuLite" and "SILuMAB" were synthesized by Sigma-Aldrich (St. Louis, MO).

The plasma samples were prepared and then analyzed according to the method described by Pinetre et al. The data was acquired using ultra-high-pressure liquid chromatography coupled to high-resolution tandem mass spectrometry (UHPLC-HRMS/MS) with a hybrid quadrupole-Orbitrap mass spectrometer (Q-Exactive HF, Thermo Fisher). Twelve peptides were targeted, based on two MS/MS transitions, comprising nine targets from "SILuLite" and three from ISTD "SILuMAB". Input of our deep learning model corresponds to EICs of two target transitions of a specific peptide provided as JPEG images.

Considering the 24 samples, the 8 concentrations, and the 12 target peptides, the initial data set size was 2304 inputs. The data curation to keep high-quality chromatograms and the removal of selected concentrations to maximize positive data leads to the suppression of 322 inputs, resulting in a data set of 1982 inputs. In order to artificially extend the size of the data set, data augmentation was performed by flipping chromatograms vertically. The final curated data set consisted of 3964 inputs.

To train and test the classifier, the 1982 inputs were manually annotated by an analytical chemist. Each targeted peptide with a spiked concentration greater than its LOD100% was labeled as "1" (i.e., detected). Blank samples and supplemented samples with concentration below the LOD100% were individually inspected by an expert chemist and annotated as "0" (i.e., not detected), "0.5" (i.e., suspicious), or "1", respectively. The data set was intentionally imbalanced with more positive inputs to promote positive classification and improve the model's performance toward a 0% FNR. Almost 70% of the inputs are labeled as detected (N = 2606) or suspicious (N = 126).

Tocilizumab Administration. The monoclonal antibody administration protocol was established according to the European animal welfare rules with ethics approval number APAFIS 12869. It was performed at the experimental facilities of the "Fédération Nationale des Courses Hippiques" (FNCH) at Goustranville (Normandie, France). The protocol consisted of a subcutaneous administration of 80 mg of Tocilizumab (RoActemra, Roche, Basel, Switzerland) to a 9 year-old French trotter gelding weighing 552 kg. The product was provided as an infusion solution (4 mL) containing 20 mg/mL of Tocilizumab. The product was diluted in a saline solution, and the final 8 mL was administered at two different spots to the horse (4 mL each). Blood was collected in heparin tubes with one preadministration sample (T0), then 30, 60, 90 min, 3h, 6, 10, 24, 34, 48 h, 3, 5, 7d, and finally daily sampling until the substance was not detected anymore.

The samples were prepared and analyzed using the validated screening method described by Pinetre et al.[13] and chromatograms were extracted using an RT window of 0.45 min. The inputs obtained were evaluated by an expert chemist and the proposed CNN-LDA classifier. The proteotypic peptide of the human IGG 1−3−4, "VVSVLTVLHQDWLNGK", was targeted since it is exogenous to the horse. A calibration curve ranging from 0 to 5000 ng/mL (0, 10, 20, 50, 100, 200, 500, 1000, 2000, and 5000) was built by linear regression to determine the concentration in each sample.

Screening Analyses Trial. Four batch analyses obtained from the GIE LCH screening analyses, corresponding to 91 equine urine samples and 4 equine urine samples supplemented with 71 different substances, were prepared, extracted, and analyzed using a screening analytical method described by Trevisiol et al.[14] LC-HRMS/MS analyses were acquired on a liquid chromatography Vanquish Flex system coupled with a quadrupole-Orbitrap mass spectrometer Exploris 120 or Q-Exactive instrument (Thermo Fisher Scientific, Bremen, Germany).

The acquired chromatograms were manually annotated by a trained analytical chemist and classified by the proposed CNN-LDA evaluation pipeline for 71 substances of interest comprising four ISTD. The final data set comprises 6745 inputs (transition 1 + transition 2 chromatograms).

Server Configuration. The artificial intelligence strategy presented in this study was trained and evaluated on an Ubuntu-based (jammy/22.04) HPE ProLiant DL380 Gen10 (Hewlett-Packard Enterprise) server. This server is equipped with two 12-core Intel Xeon-Silver 4213 processors (2.2 GHz), 192 GB of memory, and two NVIDIA Quadro RTX 4000 (8 GB of memory each) for deep learning development.

Input Extraction. The chromatogram's extraction was performed in R (v.4.3.0) based on a CSV file providing the precursor and m/z fragment ions, the polarity, and the theoretical retention time (RT) for two specific transitions per target: transition 1 (most abundant) and transition 2 (second most abundant). For each targeted substance, chromatograms were extracted in MS or MS2 from the raw file using the `Rawrr` package (v.1.8.0)[15] with an RT window of ±0.5 min and a 10 ppm mass tolerance window around the target mass and then saved as JPEG images (200 × 200) using the `ggplot2` package[16] (Figure 1). If the RT deviation exceeds

https://doi.org/10.1021/acs.analchem.4c03608
Anal. Chem. 2025, 97, 3236−3241

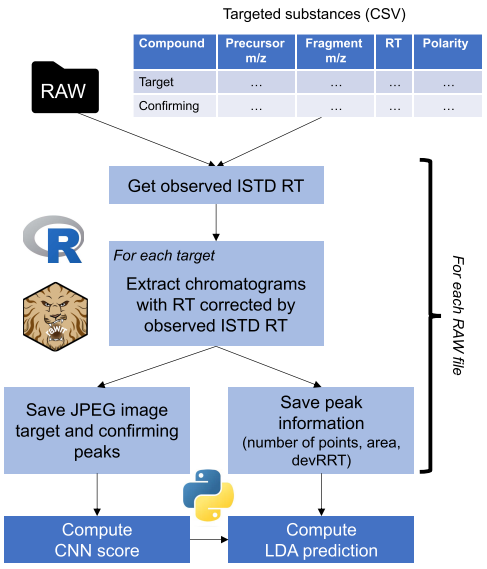


Figure 1. Workflow from the extraction of UHPLC-HRMS/MS chromatograms from raw files to classification using CNN scoring and LDA.

0.2 min, the theoretical RT is adjusted based on the sample's relevant ISTD. The peak of interest was selected as the closest to the corrected RT for each chromatogram. Then, the peak area, the number of data points per peak, and the relative RT deviation (devRRT) were saved in a separate table.

Deep Learning Model. The deep learning model was deployed using `Tensorflow`[17] (v.2.10.0) via `Keras`[18] in Python (v.3.9.16) through a dedicated Conda environment. Input corresponds to the combination of transition 1 and 2 chromatograms, as JPEG images, for a target in a sample. The neural network consisted of two CNN models with almost the same structure: two convolutional layers associated with a two-dimensional (2D) max pooling layer, followed by a flattened layer and three dense layers, all employing a rectified linear activation (ReLU) (Figure 2,A). Two dropout layers at a rate of 0.2 were used between dense layers in the CNN structure reading the transition 2. The concatenated outputs from the CNNs are sent into a multilayer perceptron (MLP) containing six dense layers, each employing a linear activation, to finally return a single value (called CNN score hereafter).

The training parameters, such as the batch size, activation, number of epochs, optimizer, metrics, and loss, were optimized using the kerashypetune function `KerasGridSearchCV` (grid search cross-validation). The model was trained for ten epochs with a batch size of 64 using the Adam optimizer (learning rate of 0.001) on 70% of the curated data set (N = 2774 inputs). Model overfitting was controlled with a validation data set composed of 20% of the training data. The remaining 30% of the curated data set (N = 1190 inputs) was used in test mode for the model's performance evaluation.

The model performance was evaluated by the accuracy metric (1) and the mean squared error (MSE) (2) as losses during the training phase. The accuracy describes the model's efficiency in correctly classifying data as positive or negative.

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \quad (2)$$

Linear Discriminant Analysis. A linear discriminant analysis (LDA) was trained using the `sklearn` Python library[19] to predict the class of the chromatograms (detected, suspicious, or not detected). LDA inputs consisted of the CNN score and three numerical parameters: the peak area, the number of points in the peak of interest, and the devRRT. The test data set from the CNN model evaluation was used to fit and evaluate the LDA, divided as 80% (N = 952 inputs) for the training data set and 20% (N = 238 inputs) for the test data set. A K-Fold cross-validation was performed to evaluate the LDA data set, using the `RepeatedStratifiedKFold` function from the `sklearn` Python library with 5 folds and 10 repeats.

Model Evaluation. Comparing the loss curve of the training data set against the validation data set provides control on the model over and under-fitting. The loss should decrease to the point of stability without increasing afterward, and the gap between the training and validation curves should remain as small as possible. The CNN model's performance was assessed using the `evaluate` method from `Tensorflow`, which returns the loss and accuracy based on the test data set.

LDA performance was evaluated using recall (also known as sensitivity) and specificity measures, which assess the model's ability to classify the data correctly. Recall controls the positive class and specificity the negative class. These metrics and a confusion matrix were computed using the R package `caret`.

RESULTS AND DISCUSSION

This study aims to develop a CNN-LDA classifier for inspecting extracted ion chromatograms mined from LC-HRMS/MS raw data at the screening stage. Since this task is time-consuming for a human reader, the challenge was to provide a high-throughput tool with high accuracy to avoid false negatives. We illustrate here the application of deep learning to equine doping control as a feasible and accurate tool for screening analysis.

Processing Time. The custom R script using `Rawrr` enables direct chromatogram extraction from raw files, thus avoiding time-consuming conversion from a specific vendor file format to a universal open file format (i.e., mzML). This extraction method is a cross-platform, providing that an open-source package is available to read the raw data, which is the case for most suppliers. The durations to process and classify the four batch analyses extracted from the GIE LCH screening method were measured to compare with the actual screening process. The estimated extraction speed is almost 8 inputs per second, which is, to the best of our knowledge, competitive with vendor-targeted data processing software solutions used for sample analyses. The CNN-LDA pipeline classified almost 150 inputs per second. This throughput is compatible with the processing of hundreds of samples per day, each sample being searched for hundreds of targeted compounds.

Model Performance. The deep learning model was designed with convolutional layers followed by an MLP and took as input target-specific transition 1 and 2 chromatograms provided as JPEG images (Figure 2 A). The training phase was performed on 2774 annotated chromatograms (70% of the curated data set) extracted from the mAb screening method validation. The CNN model learned to score peaks centered in

https://doi.org/10.1021/acs.analchem.4c03608
Anal. Chem. 2025, 97, 3236−3241

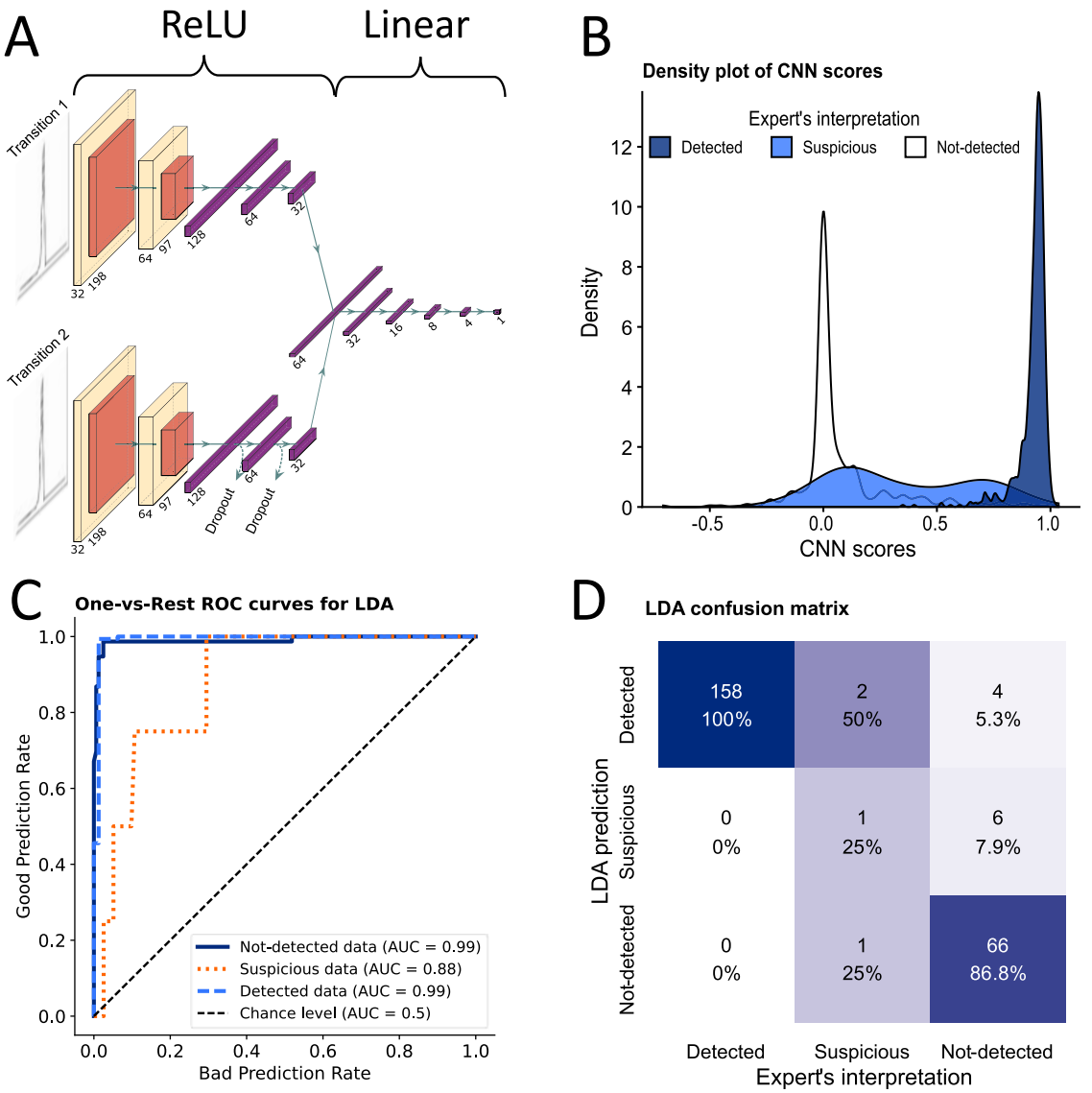


Figure 2. Performances of both the CNN model and LDA training as (A) the CNN model architecture with transition 1 and 2 as inputs computed with PlotNeuralNet,[20] (B) the density plot of the CNN scores depending on the test data set true labels, and the LDA evaluation with (C) the multiclass ROC curves and (D) the confusion matrix. ReLU = rectified linear activation.

the RT window with comparable detection capabilities compared to a trained chemist for both high-concentration samples and suspicious low-concentration samples. The evaluation of the model was based on the accuracy and the MSE computation using the `evaluate` method from `Tensorflow`. The aim is to maximize the accuracy value while minimizing the MSE. In test mode, the trained model provides an accuracy of 0.95 with a loss of 0.025 (Figure 2 B).

Then, the LDA was fitted using 80% of the CNN model's test data set (952 CNN scores and associated peak information). K-fold cross-validation (5 folds and 10 repeats) was performed to evaluate the data set, resulting in a mean accuracy of 0.942 with a standard error of 0.011. The recall and specificity computation with the associated confusion matrix assessed the LDA performance in the test mode. Those metrics were calculated respectively as 100 and 87%. Thus, the model has predicted 0% of false negatives in the test mode (Figure 2 C,D), which is, in doping control screening analyses, a mandatory requirement.

Application to Equine Doping Control. Once the CNN model and the LDA were trained, leading to accurate performances on the test data from the mAb method validation plasma samples, the classifier's efficiency was evaluated in real conditions. First, chromatograms from a Tocilizumab administration study were classified by using the CNN-LDA model. Then, the tool performance was assessed on four batch analyses of urine samples from the GIE LCH screening method described by Trevisiol et al.

Administration Study. Plasma samples were collected at various times after Tocilizumab administration and then prepared and analyzed at GIE LCH using the validated method.[13] Only the "VVSVLTVLHQDWLNGK" peptide and the ISTD were extracted from the raw files and processed by using our deep learning model. The concentration in each sample was determined by using the calibration curve.

The chromatograms were processed in our deep learning classifier, and the predictions are displayed in Figure 3. The CNN model tends to predict high scores, as expected with the imbalanced training data set. The CNN-LDA model correctly

https://doi.org/10.1021/acs.analchem.4c03608
Anal. Chem. 2025, 97, 3236−3241

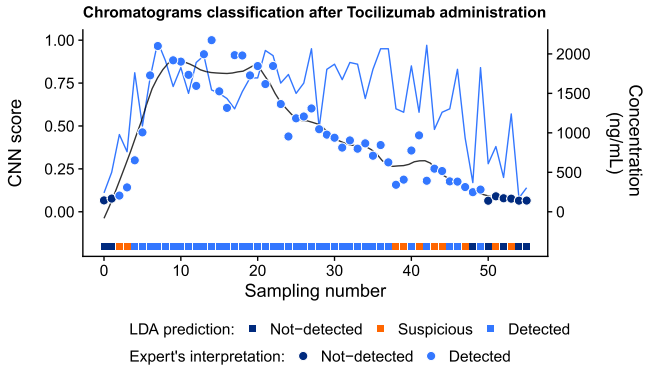


Figure 3. CNN scores (blue line) and LDA classification (squares) of "VVSVLTVLHQDWLNGK" chromatograms extracted from the samples following Tocilizumab administration compared to the concentration of the product (circles and black line).

classified 98% of the chromatograms labeled "detected" and 75% of the chromatograms labeled "not detected".

The discrepancy in "detected" chromatograms corresponds to the nondetection of sample T48. It can be explained by a low signal-to-noise ratio for the T48 target peak combined with an absence of signal for transition 2 (Figure S1). It is noteworthy that T48 refers to almost 11 weeks after administration and corresponds to a low concentration close to the method LOD. Such a result is, therefore, to be expected since the detection of a molecule in those conditions is often challenging for an experienced chemist and, by extension, for our CNN-LDA classifier. Since the classification of the experimental data set was satisfying, we decided to test our model on four batch analyses of urine samples from the LCH screening to challenge the tool with a more complex biological matrix.

Screening Analyses Trial. Four batch analyses, corresponding to 91 racing urine samples and 4 supplemented urine samples, were collected from the LCH screening method described by Trevisiol et al. to test the CNN-LDA classifier on 71 target molecules (N = 6745 inputs). This equals roughly the number of batches manually analyzed daily by an expert analyst at LCH. The four ISTD were detected by both the expert and our classifier in all samples (i.e., 380 inputs out of the 717 "detected"). The other chromatograms labeled as "detected" or "suspicious" correspond to supplemented urines, interferences known by the expert, or trace detection of targeted substances.

The proposed CNN-LDA classifier provides high confidence regarding the positive data prediction with a 0% FNR (Figure 4). It correctly classified more than 78% of negative data, which can significantly reduce visual inspection for the trained analytical chemist. The application of our classifier on screening data illustrates that deep learning strategies can be incorporated into doping control laboratories with high confidence (0% FNR) and rapid classification capabilities. Using urine samples also proves the adaptability of the current classifier to read more complex matrices without requiring transfer learning.

CONCLUSIONS

This study illustrates the potential application of AI-based tools to interpret equine doping control analyses at the screening stage and significantly reduce the time-consuming task of manual inspection performed by trained analytical chemists (67% of reduction in the screening example presented). We designed and trained a CNN model combined with an LDA classifier to screen targeted substances based on the EICs of two specific transitions and analytical parameters directly extracted from raw files. This deep learning strategy proved its accuracy when applied to the Tocilizumab administration study and was evaluated in the context of screening analyses. Even though the training was performed on plasma samples, the classifier proved its adaptability and robustness when applied to 95 routine equine urine samples despite the potential highest matrix complexity. It succeeded, without requiring transfer learning, in reaching a 0% FNR on those urine samples, which is considered as a gold standard in equine doping control. With this proof of concept, we demonstrated that implementing an AI-based strategy is fit for purpose and enables high-throughput screening analysis, as required for doping control. Based on this study, an open-

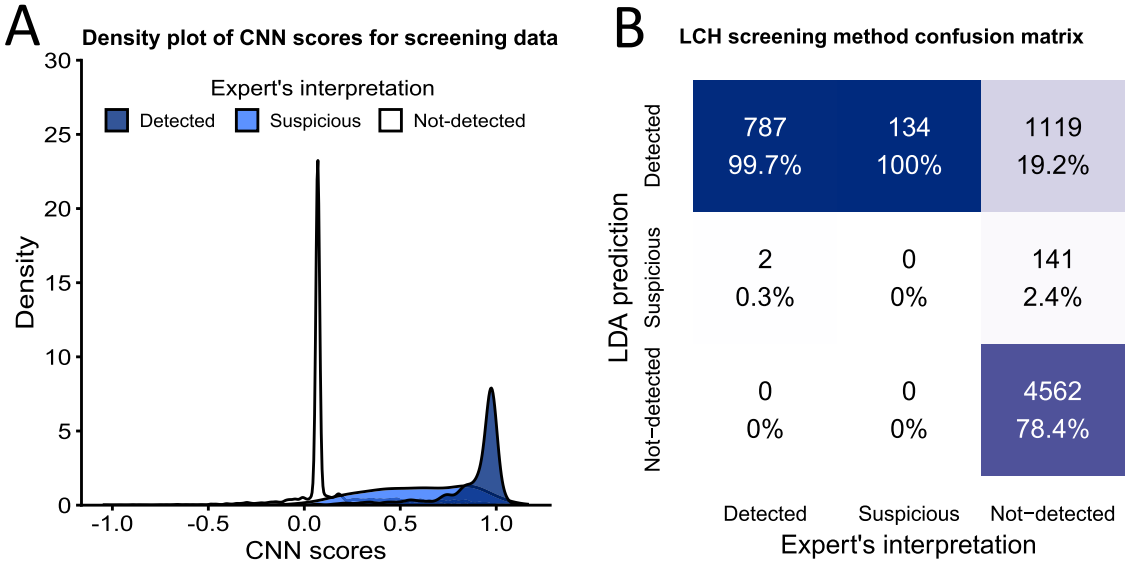


Figure 4. CNN scores (A) and confusion matrix of LDA classification (B) against expert chemist's interpretation for racing samples analyzed by the LCH screening method.

https://doi.org/10.1021/acs.analchem.4c03608
Anal. Chem. 2025, 97, 3236−3241

source and cross-platform tool could be developed that can compete with vendor software. The R script and Jupyter Notebook used to process the raw data, train CNN and LDA models, and predict the classification of EICs are available on GitHub (https://github.com/GIELCH/DeepRead).

For further development, we aim to improve our classifier with chromatograms obtained from other LCH screening methods and technologies using transfer learning. The aim is to provide trained chemist analysts with a tool incorporated into a web interface, which could load a batch of raw data, reprocess it based on a specified method, and classify extracted chromatograms. After being tested in real-life routine conditions, this tool could potentially be validated in accordance with the ISO-17025 standard to perform a preclassification of any LC-HRMS data applied in food safety or drug testing screening methods.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.4c03608.

Chromatograms from the Tocilizumab administration study to illustrate the three categories of model prediction (detected/suspicious/not detected) and the only classification where the analytical chemist and the classifier disagree (T48) (Figure S1) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Patrice Garcia − *GIE LCH, Laboratoire des Courses Hippiques, 91370 Verrières-le-Buisson, France*; Email: p.garcia@lchfrance.fr

Authors

Agnès Barnabé − *GIE LCH, Laboratoire des Courses Hippiques, 91370 Verrières-le-Buisson, France*; orcid.org/0000-0002-8420-7556

Vivian Delcourt − *GIE LCH, Laboratoire des Courses Hippiques, 91370 Verrières-le-Buisson, France*

Benoit Loup − *GIE LCH, Laboratoire des Courses Hippiques, 91370 Verrières-le-Buisson, France*

William Montanuy − *GIE LCH, Laboratoire des Courses Hippiques, 91370 Verrières-le-Buisson, France*

Stéphane Trévisiol − *GIE LCH, Laboratoire des Courses Hippiques, 91370 Verrières-le-Buisson, France*

Marie-Agnès Popot − *GIE LCH, Laboratoire des Courses Hippiques, 91370 Verrières-le-Buisson, France*

Ludovic Bailly-Chouriberry − *GIE LCH, Laboratoire des Courses Hippiques, 91370 Verrières-le-Buisson, France*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.4c03608

Author Contributions

†A.B. and V.D. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank the "Institut Français du Cheval et de l'Equitation" (IFCE) for funding the project "IGGDOP" and are grateful to "Laboratoire Roche" for the generous donation of RoActemra for Tocilizumab administration. The calculation server was partially funded by IFCE thanks to the "SuiviApp" funding. The authors also thank the staff of Goustranville for the Tocilizumab administration protocol.

■ REFERENCES

(1) Vamathevan, J.; Clark, D.; Czodrowski, P.; Dunham, I.; Ferran, E.; Lee, G.; Li, B.; Madabhushi, A.; Shah, P.; Spitzer, M.; Zhao, S. *Nat. Rev. Drug Discovery* **2019**, *18*, 463−477.

(2) Gentile, F.; Agrawal, V.; Hsing, M.; Ton, A.-T.; Ban, F.; Norinder, U.; Gleave, M. E.; Cherkasov, A. *ACS Cent. Sci.* **2020**, *6*, 939−949.

(3) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. *Nature* **2021**, *596*, 583−589.

(4) Domingo-Almenara, X.; Guijas, C.; Billings, E.; Montenegro-Burke, J. R.; Uritboonthai, W.; Aisporna, A. E.; Chen, E.; Benton, H. P.; Siuzdak, G. *Nat. Commun.* **2019**, *10*, No. 5811.

(5) Liu, Y.; De Vijlder, T.; Bittremieux, W.; Laukens, K.; Heyndrickx, W. *Rapid Commun. Mass Spectrom.* **2021**, e9120.

(6) Zeng, W.-F.; Zhou, X.-X.; Willems, S.; Ammar, C.; Wahle, M.; Bludau, I.; Voytik, E.; Strauss, M. T.; Mann, M. *Nat. Commun.* **2022**, *13*, No. 7238.

(7) Weininger, D. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31−36.

(8) Duvenaud, D. K.; Maclaurin, D.; Iparraguirre, J.; Bombarell, R.; Hirzel, T.; Aspuru-Guzik, A.; Adams, R. P. In *Convolutional Networks on Graphs for Learning Molecular Fingerprints*, Advances in Neural Information Processing Systems 28 (NIPS 2015); NeurIPS, 2015.

(9) Goh, G. B.; Hodas, N. O.; Siegel, C.; Vishnu, A. SMILES2Vec: An Interpretable General-Purpose Deep Neural Network for Predicting Chemical Properties, 2017. arXiv:1712.02034. arXiv.org e-Print archive. https://arxiv.org/abs/1712.02034.

(10) Melnikov, A. D.; Tsentalovich, Y. P.; Yanshole, V. V. *Anal. Chem.* **2020**, *92*, 588−592.

(11) Kantz, E. D.; Tiwari, S.; Watrous, J. D.; Cheng, S.; Jain, M. *Anal. Chem.* **2019**, *91*, 12407−12413.

(12) Zeng, J.; Wu, H.; He, M. *Anal. Chim. Acta* **2023**, *1238*, No. 340189.

(13) Pinetre, J.; Delcourt, V.; Becher, F.; Garcia, P.; Barnabé, A.; Loup, B.; Popot, M.-A.; Fenaille, F.; Bailly-Chouriberry, L. *Drug Test. Anal.* **2024**, *16*, 199−209.

(14) Trevisiol, S.; Popot, M.-A.; Garcia, P.; Boyer, S.; Caroff, M.; Drif, L.; Taleb, W.; Tendon, S.; Moulard, Y.; Bailly-Chouriberry, L. In vivo comparative study of hemp straw exposure and cannabidiol oil administration in horse urine *Drug Test. Anal* 2024 DOI: 10.1002/dta.3783.

(15) Kockmann, T.; Panse, C. *J. Proteome Res.* **2021**, *20*, 2028−2034.

(16) Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*; Springer-Verlag: New York, 2016.

(17) Abadi, M.et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems 2015 https://www.tensorflow.org/,software available from tensorflow.org.

(18) Chollet, F. Keras, 2015https://github.com/fchollet/keras.

(19) Pedregosa, F.; Varoquaux, G.; Michel, V.; et al. *J. Mach. Learn. Res.* **2011**, *12*, 2825−2830.

(20) Iqbal, H. lot NeuralNet v1.0.0 P, 2018, https://github.com/HarisIqbal88/PlotNeuralNet.