

Article

Retention Time Prediction with Message-Passing Neural Networks

Sergey Osipenko , Eugene Nikolaev and Yury Kostyukevich *

Center for Computational and Data-Intensive Science and Engineering, Skolkovo Institute of Science and Technology, Nobel Str., 3, 121205 Moscow, Russia

* Correspondence: y.kostyukevich@skoltech.ru

Abstract: Retention time prediction, facilitated by advances in machine learning, has become a useful tool in untargeted LC-MS applications. State-of-the-art approaches include graph neural networks and 1D-convolutional neural networks that are trained on the METLIN small molecule retention time dataset (SMRT). These approaches demonstrate accurate predictions comparable with the experimental error for the training set. The weak point of retention time prediction approaches is the transfer of predictions to various systems. The accuracy of this step depends both on the method of mapping and on the accuracy of the general model trained on SMRT. Therefore, improvements to both parts of prediction workflows may lead to improved compound annotations. Here, we evaluate capabilities of message-passing neural networks (MPNN) that have demonstrated outstanding performance on many chemical tasks to accurately predict retention times. The model was initially trained on SMRT, providing mean and median absolute cross-validation errors of 32 and 16 s, respectively. The pretrained MPNN was further fine-tuned on five publicly available small reversed-phase retention sets in a transfer learning mode and demonstrated up to 30% improvement of prediction accuracy for these sets compared with the state-of-the-art methods. We demonstrated that filtering isomeric candidates by predicted retention with the thresholds obtained from ROC curves eliminates up to 50% of false identities.



Citation: Osipenko, S.; Nikolaev, E.; Kostyukevich, Y. Retention Time Prediction with Message-Passing Neural Networks. *Separations* **2022**, *9*, 291. <https://doi.org/10.3390/separations9100291>

Academic Editor: Rosa María Alonso

Received: 31 August 2022

Accepted: 29 September 2022

Published: 6 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: retention time prediction; untargeted metabolomics; small molecules; message-passing neural networks

1. Introduction

Liquid chromatography–mass spectrometry is the primary technique for untargeted small molecule analysis. High-resolution tandem instruments annotate detected compounds in untargeted metabolomics experiments that rely on experimental or in silico fragmentation data. Mass spectral databases cover less than 1% of known organic compounds [1,2]. In silico fragmentation predictors [3–5] may have a restricted applicability domain due to being trained on small mass spectral datasets. Complementary retention information may refine compound annotation, reduce the search space and eliminate false identities.

In targeted assays, retention time of a detected peak is compared with the value obtained for reference material in the same separation conditions. Wide-scope untargeted experiments, such as forensic or environmental analysis, aim to cover as diverse a range of molecular sets as possible, making such comparisons ineffective because of the lack of experimental reference data. Retention time modeling techniques [6,7] provide an opportunity for high-throughput retention prediction for thousands of candidates retrieved by accurate mass search. These predicted values may be used as reference values to filter false identities.

There have been numerous attempts to estimate retention times of small molecules based on machine learning or non-learning approaches [8–25]. However, such models

lack accuracy because of limited training data and the complicated nature of retention in HPLC. The breakthrough in retention time prediction became possible due to significant progress in developing deep learning techniques. The introduction of the METLIN Small Molecule Retention Time (SMRT) dataset [26] made it possible to train deep learning models to achieve unprecedented performance. Gradient boosting [27], deep neural networks with fully connected layers [26], graph neural networks [28–30] and convolutional neural networks have been deployed with the SMRT data. Recent works have indicated accuracy comparable with the experimental error of retention time measurements [31].

The variety of experimental setups is the major problem associated with retention prediction. Even though a model can provide very accurate predictions for SMRT, these predictions should somehow be transferred to the separation conditions used in a particular experiment. The correlation of retention on various systems is usually unclear. Changes in column dimensions, stationary phase or eluent composition may result in shifted times or even retention order. Complicated regression models may be used to recalculate predictions between two systems, but this requires shared molecules between two datasets [26,27,32–34]. A more intelligent way is to apply the two-step transfer learning approach [35]. This includes model pretraining on a large dataset, and fine-tuning the model for the task of interest on a smaller collection using weights obtained during the first step. Recent work [36] demonstrated that a retention dataset of several hundreds of compounds, not necessarily from SMRT, is needed for transfer learning. A deep learning model is first pretrained on SMRT or even a non-labeled dataset [29,37]. At this step, the model learns low-level molecular features common for different molecular tasks. Then, the model is initialized for target training with pretrained first layers that correspond to such features. Since these first layers are often frozen during fine-tuning, the generalization ability of a model significantly influences the overall performance. Thus, overfitted models may demonstrate low prediction error on the pretraining set but poor performance after transfer learning. Indeed, the prediction error normalized to the chromatographic runtime for small datasets is much worse than that obtained for the SMRT conditions.

Therefore, the evaluation of new architectures may contribute to accurate predictions on small datasets. This work aims to evaluate the message-passing neural network (MPNN), a kind of graph neural network, on retention data. MPNNs were initially designed for handling molecular data [38] and demonstrated outstanding performance on various chemical tasks [39–42]. To our knowledge, MPNN has only been used to model retention time in liquid chromatography on a tiny dataset [43]. We implement a transfer learning pipeline with pretraining MPNN on the SMRT dataset and fine-tuning on several available and in-house retention collections. We also propose a workflow to eliminate false identities based on predicted retention values.

2. Materials and Methods

Datasets. The METLIN SMRT dataset of 80,038 molecules was used for pretraining and was processed with the rdkit library. Retention times were collected in reversed-phase separation conditions with mean and median retention time variability of 36 and 18 s, respectively [26]. SMRT has a bimodal retention distribution (Figure S1) due to non-retained molecules with less than 2 min retention time. Following previous work [30,31], we excluded these molecules from the training set. The final set contained 77,977 compounds.

The model was fine-tuned and evaluated on small reversed-phase datasets from the PredRet database [33] (FEM_long, LIFE_new, LIFE_old, Eawag_XBridgeC18) and the RIKEN Retip dataset [12]. The retention distributions of these sets can be found in Figure S1.

Message-passing neural network. Following the pipeline for constructing the message-passing neural network from the original paper on MPNNs [38], our model included a featurizing step, message-passing, readout and a set of fully-connected layers. We took the implementation from the Keras tutorial on MPNNs with several changes of molecular features and hyperparameters. First, molecules were converted into undirected

graphs, with atoms as nodes and bonds as edges. Graph nodes and edges were encoded with atom and bond features (Table S1) to generate input for the model. The message-passing step updated the node state considering its neighborhood and edge features. Previous states were handled with a gated recurrent unit (GRU) [44]. At the readout phase, the highly-dimensional updated node states were reduced to embedded vectors with a combination of a transformer and average pooling layers. These embeddings were then passed through 3 dense layers with ReLU activation and a linear layer to get predictions. The overview of the workflow is presented in Figure 1. The detailed model architecture with all input/output layer dimensions can be found in Figure S2.

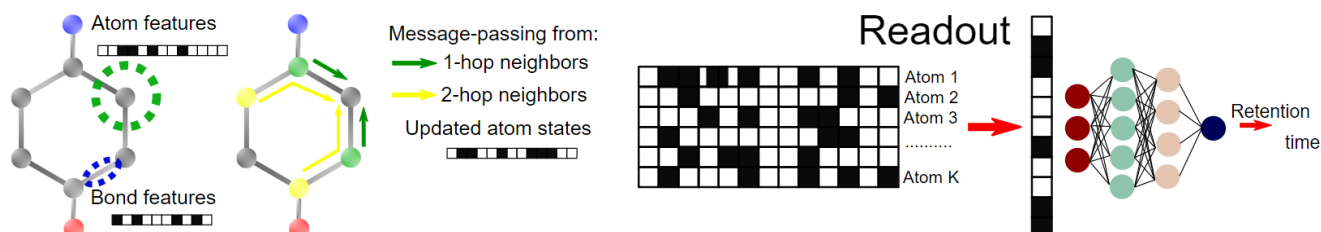


Figure 1. A general scheme of a message-passing neural network.

Model training and evaluation. The model was trained with Adam optimizer, mean absolute error as a loss function, batch size of 64 and initial learning rate 0.0002. The learning rate was decreased by a factor of 0.25 if validation loss did not improve for ten epochs. Overfitting was controlled via an early stopping mechanism. Fitting was stopped if the validation loss did not improve for twenty epochs.

Model evaluation was performed with a 5-fold cross-validation protocol. Additionally, a separate 20% test set was extracted randomly and was common for all cross-validation iterations. Mean (MAE), median absolute error (MedAE), root mean squared error (RMSE), mean absolute percentage errors (MAPE) and R^2 score were used as metrics for evaluation. Datasets were randomly split into training and validation sets at an 80:20 ratio to train the final models. In the transfer learning mode, the weights of all the layers except the dense and the linear were frozen and loaded from a pretrained model before training. The batch size was reduced to 8. All other parameters were kept the same.

Elimination of false identities. To estimate the proposed approach's practical utility, we applied it to distinguish between the isomers of compounds in the evaluation sets by predicted RT values. To choose the threshold to filter compounds, we plotted the receiver–operating characteristic (ROC) curve. For that, for all the molecules in the test sets that were unseen by the model, we downloaded all the isomers from PubChem using PUG requests [45]. For each isomer, RT was predicted by the model and compared with the experimental RT for the initial molecule from the test set to consider its isomer as a true-negative or false-positive for a certain threshold.

For each threshold value taken from the range 0–200% with step 2.5%, we calculated the number of true-positives (TP) (the correct identity's predicted RT was within this threshold from experimentally measured value), false-positives (FP) (the false identity's predicted RT was within this threshold from measured value), the total number of positive cases (P) (equal to the total number of compounds in the evaluation set) and negative cases (N) (total number of isomers of all compounds from test sets without positive cases).

For example, if the relative difference between the experimental and predicted values for a molecule in the test set was 5%, it was considered TP for all thresholds above 5% and FN for all bellow. Its isomers (that are presumed false identities) were considered FP and TN, respectively, for the same thresholds.

To obtain ROC, we plotted TP/P vs. FP/N at different thresholds. The best threshold corresponds to the highest value of TPR–FPR to compromise between the numbers of FN and FP.

All the code was written in Python and is available with the pretrained models via the following link <https://github.com/osv91/MPNN-RT> (Accessed on 3 October 2022).

3. Results and Discussion

Message-passing neural network. Since their introduction in 2017, message-passing neural networks have outperformed other deep neural networks and traditional machine learning methods in many molecular regression and classification tasks [39]. MPNNs are implemented in DeepChem, a popular chemistry-oriented deep learning framework [46]. They are available for general deep learning frameworks in official or non-official tutorials or add-ins. We have chosen an MPNN implementation from Keras [47] code examples and adjusted the model architecture, increasing the number of message-passing steps and adding dense layers to the decoder part of the model.

The results of the 5-fold cross-validation (Figure 2) of the MPNN are listed in Table 1. The MPNN outperforms the recently proposed state-of-the-art models based on 1D-CNN [31] and GNN [35] architectures. More importantly, it provides mean and median absolute errors less than the experimental values for the SMRT dataset, which were reported to be 36 and 18 s, respectively. The difference between the test and the validation metrics is minor, indicating moderate model overfitting.

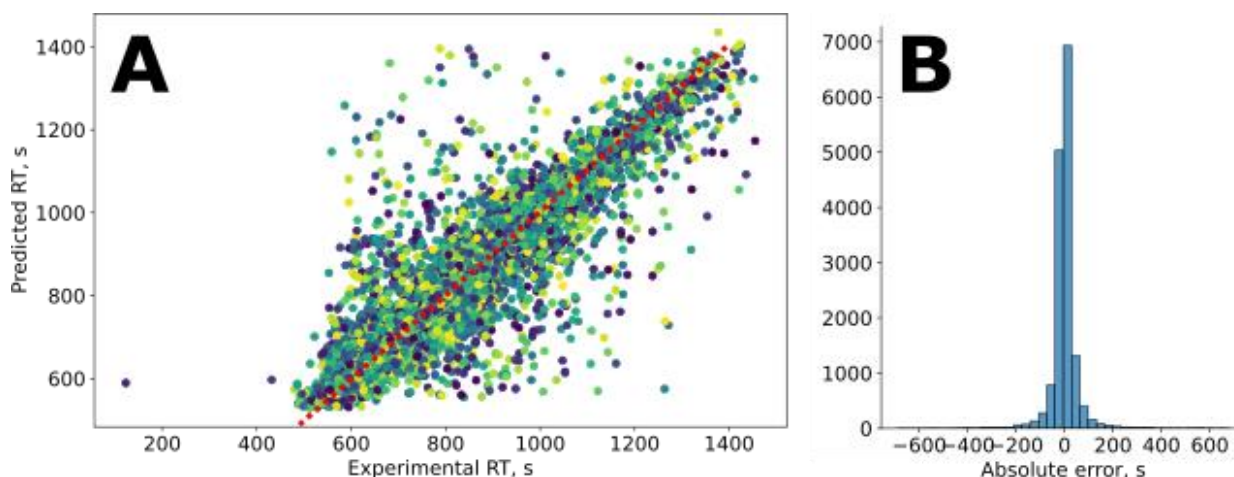


Figure 2. The results of retention predictions on the SMRT dataset: (A) predicted vs. experimental values for the test set; (B) error distribution plot.

Table 1. Mean metrics for 5-fold cross-validation of MPNN on SMRT dataset compared with previously published results.

Model	MAE, s		MedAE, s		MAPE, %		RMSE		R ²	
	Validation	Test	Validation	Test	Validation	Test	Validation	Test	Validation	Test
MPNN	32.1 ± 0.6	31.5 ± 0.1	16.2 ± 0.2	16.0 ± 0.2	4.1 ± 0.1	4.0 ± 0.01	62.8 ± 1.9	60.5 ± 0.3	0.872 ± 0.008	0.879 ± 0.001
1D-CNN [31]	34.7 ± 1.2		18.7 ± 1.3		4.3 ± 0.1		65.5 ± 2.7		No data	
GNN [30]	39.87		25.24		5		No data		No data	

Transferring predictions to small datasets. Predictions made with a model trained on the SMRT dataset cannot be directly applied for small molecule analysis because of the wide variety of chromatographic setups. Changing a column or mobile phase gradient program leads to substantial retention time changes for a given analyte. Therefore, a new collection of retention times for the chosen separation conditions is required to get actual predicted values. These retention times can be obtained with mappings between SMRT and the new dataset, constructing a model from scratch with the new dataset, or using transfer learning. The last approach is the most promising. To reduce overfitting and get more accurate predictions, it may benefit to have a large data collection, such as SMRT or non-labeled public chemistry databases. We estimated the capability of the MPNN pretrained on the SMRT to improve predictions for five publicly available datasets and compared the results

with MPNN trained from scratch and with some previously reported results. We reduced the batch size and froze all the weights except dense and linear layers to account for small dataset sizes. The results are summarized in Tables 2 and 3 and Figure 3. MPNN behaves comparably with or better than 1D-CNN and GNN on most datasets. For example, for the Eawag_XBridgeC18 dataset, the mean and median absolute errors improved by about 30%. In total, MPNN outperforms GNN on 3 of 4 evaluated sets by median absolute error and all datasets by mean absolute error. Only for the RIKEN Retip dataset the mean absolute error of MPNN was higher by more than 2 s than that reported for 1D-CNN.

Table 2. Mean absolute error (in seconds) for 5-fold cross-validation of MPNN compared with previously published results.

Model	RIKEN Retip		FEM_long		Eawag_XBridgeC18		LIFE_new		LIFE_old	
	Validation	Test	Validation	Test	Validation	Test	Validation	Test	Validation	Test
MPNN Transfer Learning	34.7 ± 3.3	38.2 ± 3.2	214.4 ± 25.6	204.6 ± 23.0	79.5 ± 10.3	80.9 ± 6.5	23.3 ± 3.8	22.1 ± 3.3	16.9 ± 2.9	16.9 ± 2.0
MPNN From Scratch	56.2 ± 13.6	57.2 ± 13.2	299.4 ± 51.0	317.7 ± 25.8	137.1 ± 24.7	135.8 ± 14.1	37.5 ± 17.2	39.1 ± 15.2	27.2 ± 8.3	23.2 ± 2.1
1D-CNN Transfer learning [31]	32.4 ± 3.0		No data		No data		23.6 ± 5.1		15.5 ± 2.8	
GNN Transfer learning [30]	No data		235.01		112.78		29.38		17.10	

Table 3. Median absolute error (in seconds) for 5-fold cross-validation of MPNN compared with previously published results.

Model	RIKEN Retip		FEM_long		Eawag_XBridgeC18		LIFE_new		LIFE_old	
	Validation	Test	Validation	Test	Validation	Test	Validation	Test	Validation	Test
MPNN Transfer Learning	22.2 ± 2.4	25.0 ± 3.5	93.5 ± 14.9	125.2 ± 12.0	56.0 ± 4.7	57.4 ± 9.3	12.2 ± 1.1	9.7 ± 3.6	9.9 ± 3.1	9.5 ± 0.7
MPNN From Scratch	38.0 ± 9.5	40.8 ± 12.4	162.4 ± 51.2	193.1 ± 41.3	105.3 ± 20.7	115.1 ± 18.1	20.4 ± 21.5	24.0 ± 22.2	18.11 ± 8.6	19.5 ± 3.6
1D-CNN Transfer learning [31]	22.2 ± 3.6		No data		No data		14.7 ± 4.7		11.8 ± 4.2	
GNN Transfer learning [30]	No data		94.66		83.79		15.16		12.88	

An example of eliminating false identities. The main goal of retention modeling is to provide a filter for non-targeted LC-MS workflows. Compound annotation in untargeted metabolomics is done by accurate mass, providing long isomeric candidates. Their fragmentation spectra may partially distinguish them, but MS/MS data are unavailable for most isomers, and annotation relies on in silico fragmentation. Although these methods are powerful, they require computational resources, and filtering the list of candidates by retention times may reduce the overall time for data processing.

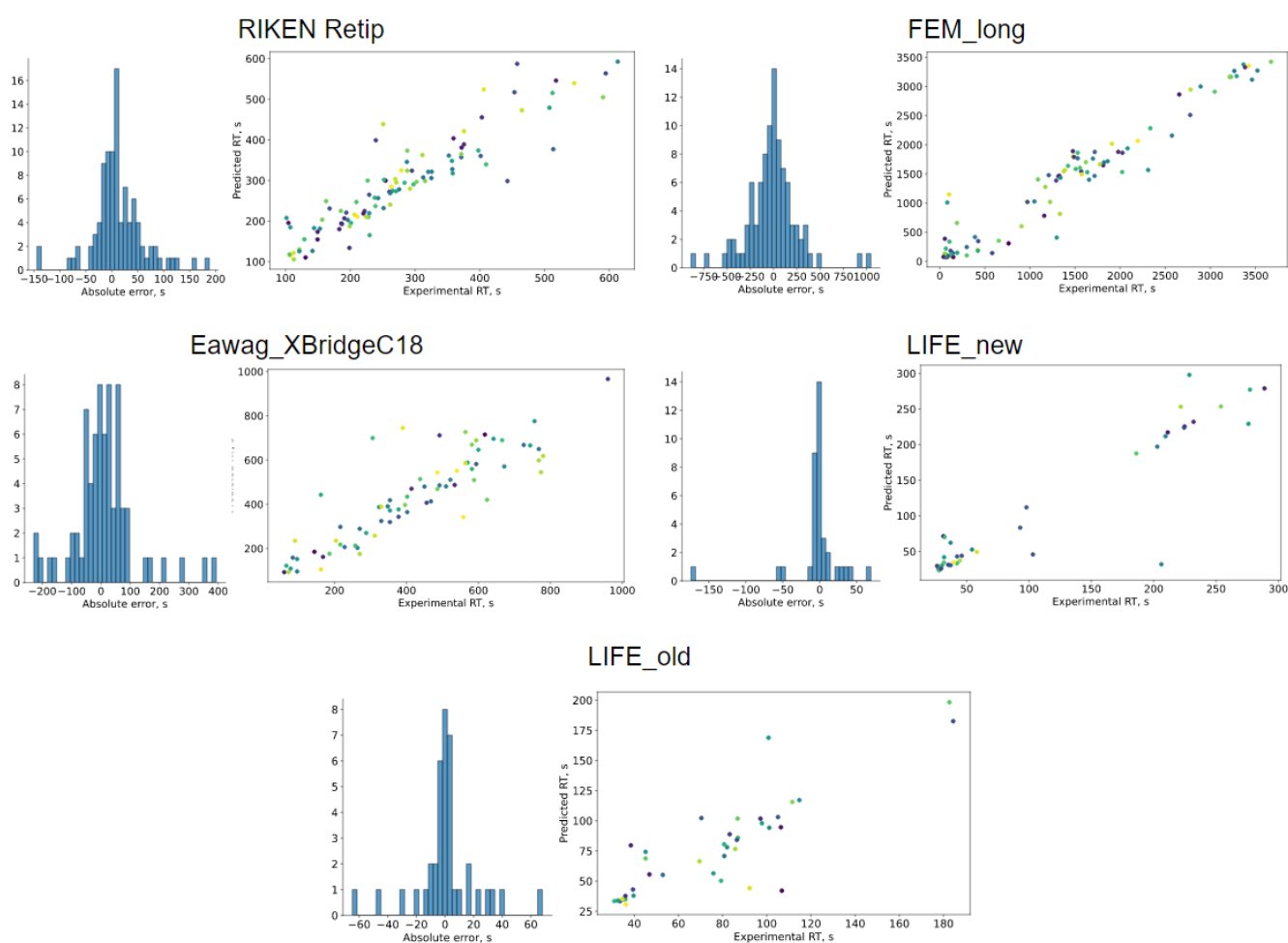


Figure 3. Predicted vs. experimental values and error distribution plots for the evaluated sets.

For example, filtering candidates by retention time may be done using a simple filter based on the prediction standard deviations. However, estimating thresholds from receiver-operation characteristics (ROC) may reduce the number of false results. To test MPNN's ability to aid in compound annotation, we plotted ROC curves with the test splits of small datasets to choose the optimum threshold and filtered out candidates using these values. The results are shown in Figure 4. The best thresholds established for the Retip, FEM_long, Eawag_XbridgeC18, LIFE_new and LIFE_old datasets were 22.5, 10, 20, 17.5 and 20%, respectively, and facilitated eliminating on average 53, 23, 35, 33 and 31% of false identities. Although the average number of eliminated false positives is relatively small, it was obtained for most complete lists of isomeric candidates from a general chemistry database. Targeting the search space to more specific databases, such as HMDB [48], or combining filtration by retention with in silico fragmentation [49] may improve the results.

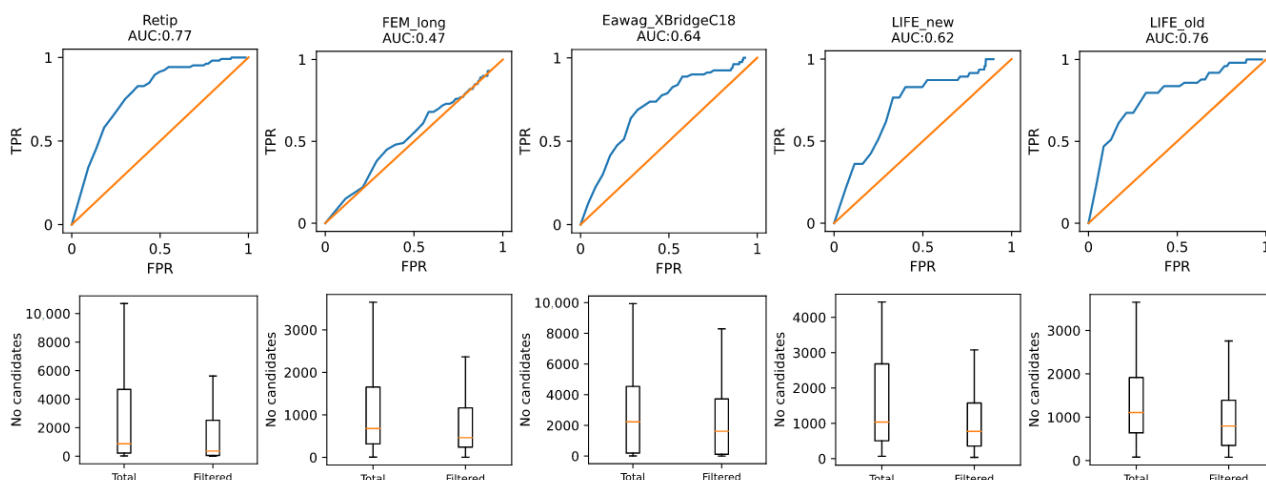


Figure 4. ROC curves and eliminated false identities.

4. Conclusions

Message-passing neural networks applied to the task of predicting retention time in liquid chromatography behave comparably or outperform state-of-the-art graph neural networks and 1D convolutional neural networks when trained on the METLIN SMRT dataset. Transferring the knowledge by two-step training allows accurate predictions on datasets collected from custom HPLC systems benefiting from pretraining on an extensive retention time collection. Investigating ROC curves allows adjusting the thresholds for efficient filtering of isomeric candidates by their predicted retention. Such filtering may facilitate the annotation process in untargeted small molecule applications.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/separations9100291/s1>. Figure S1. Retention time distribution across the investigated datasets. Figure S2. The detailed architecture of the proposed message-passing neural network. Table S1. Atom and bond features.

Author Contributions: Curation, Y.K. and E.N.; writing—original draft preparation, S.O.; writing—review and editing, Y.K. and E.N.; supervision, E.N.; project administration, Y.K.; funding acquisition, Y.K. All authors have read and agreed to the published version of the manuscript.

Funding: The research was supported by the Russian Scientific Foundation grant no. 18-79-10127.

Data Availability Statement: The source code of retention time predictors and the pretrained models are available from GitHub <https://github.com/osv91/MPNN-RT> (Accessed on 3 October 2022).

Conflicts of Interest: The authors declare no conflict of interest.

References

- Xue, J.; Guijas, C.; Benton, H.P.; Warth, B.; Siuzdak, G. METLIN MS2 molecular standards database: A broad chemical and biological resource. *Nat. Methods* **2020**, *17*, 953–954. [[CrossRef](#)] [[PubMed](#)]
- Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B.A.; Thiessen, P.A.; Yu, B.; et al. PubChem 2019 update: Improved access to chemical data. *Nucleic Acids Res.* **2019**, *47*, D1102–D1109. [[CrossRef](#)] [[PubMed](#)]
- Djombou-Feunang, Y.; Pon, A.; Karu, N.; Zheng, J.; Li, C.; Arndt, D.; Gautam, M.; Allen, F.; Wishart, D.S. CFM-ID 3.0: Significantly Improved ESI-MS/MS Prediction and Compound Identification. *Metabolites* **2019**, *9*, 23. [[CrossRef](#)] [[PubMed](#)]
- Dührkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A.A.; Melnik, A.V.; Meusel, M.; Dorrestein, P.C.; Rousu, J.; Böcker, S. SIRIUS 4: A rapid tool for turning tandem mass spectra into metabolite structure information. *Nat. Methods* **2019**, *16*, 299–302. [[CrossRef](#)]
- Ruttkies, C.; Neumann, S.; Posch, S. Improving MetFrag with statistical learning of fragment annotations. *BMC Bioinform.* **2019**, *20*, 14. [[CrossRef](#)]
- Witting, M.; Böcker, S. Current status of retention time prediction in metabolite identification. *J. Sep. Sci.* **2020**, *43*, 1746–1754. [[CrossRef](#)]

7. Haddad, P.R.; Taraji, M.; Szücs, R. Prediction of Analyte Retention Time in Liquid Chromatography. *Anal. Chem.* **2021**, *93*, 228–256. [[CrossRef](#)]
8. Aalizadeh, R.; Nika, M.C.; Thomaidis, N.S. Development and application of retention time prediction models in the suspect and non-target screening of emerging contaminants. *J. Hazard. Mater.* **2019**, *363*, 277–285. [[CrossRef](#)]
9. Aicheler, F.; Li, J.; Hoene, M.; Lehmann, R.; Xu, G.W.; Kohlbacher, O. Retention Time Prediction Improves Identification in Nontargeted Lipidomics Approaches. *Anal. Chem.* **2015**, *87*, 7698–7704. [[CrossRef](#)]
10. Amos, R.I.J.; Haddad, P.R.; Szucs, R.; Dolan, J.W.; Pohl, C.A. Molecular modeling and prediction accuracy in Quantitative Structure-Retention Relationship calculations for chromatography. *TrAC Trends Anal. Chem.* **2018**, *105*, 352–359. [[CrossRef](#)]
11. Bach, E.; Szedmak, S.; Brouard, C.; Böcker, S.; Rousu, J. Liquid-chromatography retention order prediction for metabolite identification. *Bioinformatics* **2018**, *34*, i875–i883. [[CrossRef](#)] [[PubMed](#)]
12. Bonini, P.; Kind, T.; Tsugawa, H.; Barupal, D.K.; Fiehn, O. Retip: Retention Time Prediction for Compound Annotation in Untargeted Metabolomics. *Anal. Chem.* **2020**, *92*, 7515–7522. [[CrossRef](#)]
13. Boswell, P.G.; Schellenberg, J.R.; Carr, P.W.; Cohen, J.D.; Hegeman, A.D. Easy and accurate high-performance liquid chromatography retention prediction with different gradients, flow rates, and instruments by back-calculation of gradient and flow rate profiles. *J. Chromatogr. A* **2011**, *1218*, 6742–6749. [[CrossRef](#)] [[PubMed](#)]
14. Bouwmeester, R.; Martens, L.; Degroeve, S. Comprehensive and Empirical Evaluation of Machine Learning Algorithms for Small Molecule LC Retention Time Prediction. *Anal. Chem.* **2019**, *91*, 3694–3703. [[CrossRef](#)] [[PubMed](#)]
15. Bruderer, T.; Varesio, E.; Hopfgartner, G. The use of LC predicted retention times to extend metabolites identification with SWATH data acquisition. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2017**, *1071*, 3–10. [[CrossRef](#)]
16. Cao, M.S.; Fraser, K.; Huege, J.; Featonby, T.; Rasmussen, S.; Jones, C. Predicting retention time in hydrophilic interaction liquid chromatography mass spectrometry and its use for peak annotation in metabolomics. *Metabolomics* **2015**, *11*, 696–706. [[CrossRef](#)]
17. Codesido, S.; Randazzo, G.M.; Lehmann, F.; González-Ruiz, V.; García, A.; Xenarios, I.; Liechti, R.; Bridge, A.; Boccard, J.; Rudaz, S. DynaStI: A Dynamic Retention Time Database for Steroidomics. *Metabolites* **2019**, *9*, 85. [[CrossRef](#)]
18. Creek, D.J.; Jankevics, A.; Breitling, R.; Watson, D.G.; Barrett, M.P.; Burgess, K.E.V. Toward Global Metabolomics Analysis with Hydrophilic Interaction Liquid Chromatography–Mass Spectrometry: Improved Metabolite Identification by Retention Time Prediction. *Anal. Chem.* **2011**, *83*, 8703–8710. [[CrossRef](#)]
19. Falchi, F.; Bertozzi, S.M.; Ottonello, G.; Ruda, G.F.; Colombano, G.; Fiorelli, C.; Martucci, C.; Bertorelli, R.; Scarpelli, R.; Cavalli, A.; et al. Kernel-Based, Partial Least Squares Quantitative Structure-Retention Relationship Model for UPLC Retention Time Prediction: A Useful Tool for Metabolite Identification. *Anal. Chem.* **2016**, *88*, 9510–9517. [[CrossRef](#)]
20. Feng, C.; Xu, Q.; Qiu, X.; Jin, Y.; Ji, J.; Lin, Y.; Le, S.; She, J.; Lu, D.; Wang, G. Evaluation and application of machine learning-based retention time prediction for suspect screening of pesticides and pesticide transformation products in LC-HRMS. *Chemosphere* **2021**, *271*, 129447. [[CrossRef](#)]
21. Kitamura, R.; Kawabe, T.; Kajiro, T.; Yonemochi, E. The development of retention time prediction model using multilinear gradient profiles of seven pharmaceuticals. *J. Pharm. Biomed. Anal.* **2021**, *198*, 114024. [[CrossRef](#)] [[PubMed](#)]
22. Parinet, J. Predicting reversed-phase liquid chromatographic retention times of pesticides by deep neural networks. *Heliyon* **2021**, *7*, e08563. [[CrossRef](#)] [[PubMed](#)]
23. Pasin, D.; Mollerup, C.B.; Rasmussen, B.S.; Linnet, K.; Dalsgaard, P.W. Development of a single retention time prediction model integrating multiple liquid chromatography systems: Application to new psychoactive substances. *Anal. Chim. Acta* **2021**, *1184*, 339035. [[CrossRef](#)] [[PubMed](#)]
24. Rojas, C.; Aranda, J.F.; Jaramillo, E.P.; Losilla, I.; Tripaldi, P.; Duchowicz, P.R.; Castro, E.A. Foodinformatic prediction of the retention time of pesticide residues detected in fruits and vegetables using UHPLC/ESI Q-Orbitrap. *Food Chemistry*. **2021**, *342*, 128354. [[CrossRef](#)]
25. Liapikos, T.; Zisi, C.; Kodra, D.; Kademoglou, K.; Diamantidou, D.; Begou, O.; Pappa-Louisi, A.; Theodoridis, G. Quantitative Structure Retention Relationship (QSRR) Modelling for Analytes' Retention Prediction in LC-HRMS by Applying Different Machine Learning Algorithms and Evaluating Their Performance. *J. Chromatogr. B* **2022**, *1191*, 123132. [[CrossRef](#)]
26. Domingo-Almenara, X.; Guijas, C.; Billings, E.; Montenegro-Burke, J.R.; Uritboonthai, W.; Aisporna, A.E.; Chen, E.; Benton, H.P.; Siuzdak, G. METLIN small molecule dataset for machine learning-based retention time prediction. *Nat. Commun.* **2019**, *10*, 5811. [[CrossRef](#)]
27. Osipenko, S.; Bashkirova, I.; Sosnin, S.; Kovaleva, O.; Fedorov, M.; Nikolaev, E.; Kostyukevich, Y. Machine learning to predict retention time of small molecules in nano-HPLC. *Anal. Bioanal. Chem.* **2020**, *412*, 7767–7776. [[CrossRef](#)]
28. Kensert, A.; Bouwmeester, R.; Efthymiadis, K.; Van Broeck, P.; Desmet, G.; Cabooter, D. Graph Convolutional Networks for Improved Prediction and Interpretability of Chromatographic Retention Data. *Anal. Chem.* **2021**, *93*, 15633–15641. [[CrossRef](#)]
29. Yang, Q.; Ji, H.; Fan, X.; Zhang, Z.; Lu, H. Retention time prediction in hydrophilic interaction liquid chromatography with graph neural network and transfer learning. *J. Chromatogr. A* **2021**, *1656*, 462536. [[CrossRef](#)]
30. Yang, Q.; Ji, H.; Lu, H.; Zhang, Z. Prediction of Liquid Chromatographic Retention Time with Graph Neural Networks to Assist in Small Molecule Identification. *Anal. Chem.* **2021**, *93*, 2200–2206. [[CrossRef](#)]
31. Fedorova, E.S.; Matyushin, D.D.; Plyushchenko, I.V.; Stavriani, A.N.; Buryak, A.K. Deep learning for retention time prediction in reversed-phase liquid chromatography. *J. Chromatogr. A* **2022**, *1664*, 462792. [[CrossRef](#)] [[PubMed](#)]

32. Bouwmeester, R.; Martens, L.; Degroeve, S. Generalized Calibration Across Liquid Chromatography Setups for Generic Prediction of Small-Molecule Retention Times. *Anal. Chem.* **2020**, *92*, 6571–6578. [[CrossRef](#)] [[PubMed](#)]
33. Stanstrup, J.; Neumann, S.; Vrhovsek, U. PredRet: Prediction of Retention Time by Direct Mapping between Multiple Chromatographic Systems. *Anal. Chem.* **2015**, *87*, 9421–9428. [[CrossRef](#)] [[PubMed](#)]
34. Boswell, P.G.; Schellenberg, J.R.; Carr, P.W.; Cohen, J.D.; Hegeman, A.D. A study on retention “projection” as a supplementary means for compound identification by liquid chromatography–mass spectrometry capable of predicting retention with different gradients, flow rates, and instruments. *J. Chromatogr. A* **2011**, *1218*, 6732–6741. [[CrossRef](#)]
35. Pan, S.J.; Yang, Q.A. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [[CrossRef](#)]
36. Ju, R.; Liu, X.; Zheng, F.; Lu, X.; Xu, G.; Lin, X. Deep Neural Network Pretrained by Weighted Autoencoders and Transfer Learning for Retention Time Prediction of Small Molecules. *Anal. Chem.* **2021**, *93*, 15651–15658. [[CrossRef](#)]
37. Osipenko, S.; Botashev, K.; Nikolaev, E.; Kostyukevich, Y. Transfer learning for small molecule retention predictions. *J. Chromatogr. A* **2021**, *1644*, 462119. [[CrossRef](#)]
38. Gilmer, J.; Schoenholz, S.S.; Riley, P.F.; Vinyals, O.; Dahl, G.E. Neural Message Passing for Quantum Chemistry. *arXiv* **2017**, arXiv:1704.01212.
39. Wu, Z.; Ramsundar, B.; Feinberg, E.N.; Gomes, J.; Geniesse, C.; Pappu, A.S.; Leswing, K.; Pande, V. MoleculeNet: A Benchmark for Molecular Machine Learning. *arXiv* **2017**, arXiv:1703.00564. [[CrossRef](#)]
40. Tang, B.; Kramer, S.T.; Fang, M.; Qiu, Y.; Wu, Z.; Xu, D. A self-attention based message passing neural network for predicting molecular lipophilicity and aqueous solubility. *J. Cheminform.* **2020**, *12*, 15. [[CrossRef](#)]
41. McGill, C.; Forsuelo, M.; Guan, Y.; Green, W.H. Predicting Infrared Spectra with Message Passing Neural Networks. *J. Chem. Inf. Model.* **2021**, *61*, 2594–2609. [[CrossRef](#)] [[PubMed](#)]
42. Withnall, M.; Lindelöf, E.; Engkvist, O.; Chen, H. Building attention and edge message passing neural networks for bioactivity and physical–chemical property prediction. *J. Cheminform.* **2020**, *12*, 1. [[CrossRef](#)] [[PubMed](#)]
43. Xing, G.; Sresht, V.; Sun, Z.; Shi, Y.; Clasquin, M.F. Coupling Mixed Mode Chromatography/ESI Negative MS Detection with Message-Passing Neural Network Modeling for Enhanced Metabolome Coverage and Structural Identification. *Metabolites* **2021**, *11*, 772. [[CrossRef](#)] [[PubMed](#)]
44. Cho, K.; van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. *arXiv* **2014**, arXiv:1406.1078.
45. Kim, S.; Thiessen, P.A.; Cheng, T.; Zhang, J.; Gindulyte, A.; Bolton, E.E. PUG-View: Programmatic access to chemical annotations integrated in PubChem. *J. Cheminform.* **2019**, *11*, 56. [[CrossRef](#)] [[PubMed](#)]
46. Ramsundar, B. Molecular machine learning with DeepChem. *Abstr. Pap. Am. Chem. Soc.* **2018**, *255*, 1.
47. Chollet, F.C. Keras. 2015. Available online: <https://keras.io> (accessed on 30 August 2022).
48. Wishart, D.S.; Feunang, Y.D.; Marcu, A.; Guo, A.C.; Liang, K.; Vázquez-Fresno, R.; Sajed, T.; Johnson, D.; Li, C.; Karu, N.; et al. HMDB 4.0: The human metabolome database for 2018. *Nucleic Acids Res.* **2018**, *46*, D608–D617. [[CrossRef](#)]
49. Hu, M.; Müller, E.; Schymanski, E.L.; Ruttkies, C.; Schulze, T.; Brack, W.; Krauss, M. Performance of combined fragmentation and retention prediction for the identification of organic micropollutants by LC-HRMS. *Anal. Bioanal. Chem.* **2018**, *410*, 1931–1941. [[CrossRef](#)]