TIMS DIA-NN: Deep Learning for DIA Data Analysis

Robin Park[1],Tomasz Siwarga[1], Mateusz Lacki[2], Paolo Ballesteros[1], Vijayaraja Gnanasambandan[1], Jonathan Krieger[1], Tharan Srikumar[1], Dennis Trede[1], Rohan Thakur[1], Stefan Tenzer[2], 1. Bruker Daltonics, 2. University Medical Center of the Johannes-Gutenberg University Mainz

Introduction to TIMS DIA-NN

To enhance the capabilities of the CCS-centric DIA analysis tool, TIMS DIA-NN, we have incorporated deep learning modules. A one feature of this integration is the Collision Cross-Section (CCS) prediction tool, which improves the accuracy of protein identification and quantification. Furthermore, we have embedded a neural network classification mechanism tailored specifically to differentiate between decoy and target precursors. Additionally, we have incorporated models predicting peptide detectability and intensity, enabling more confident library-free data analyses.

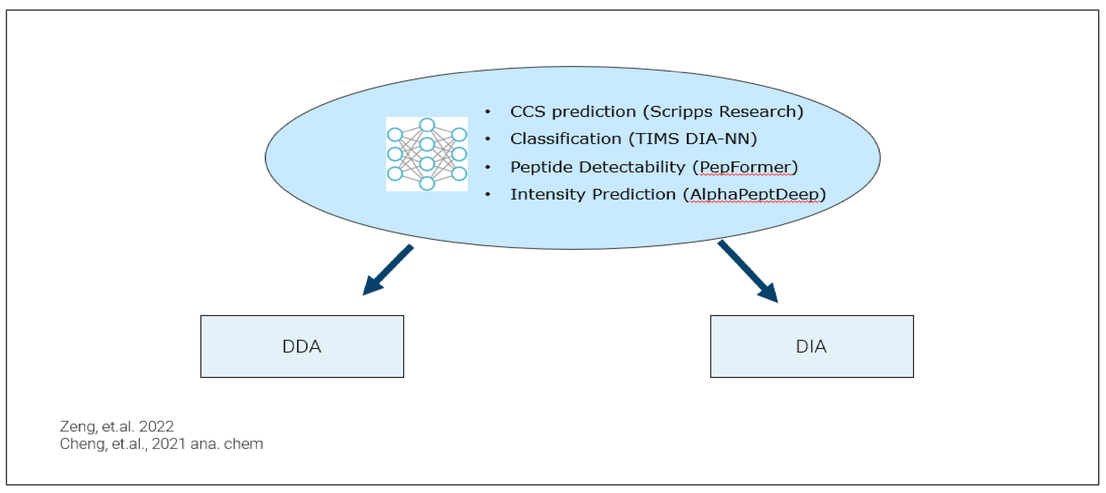

Fig. 1 Deep learning models in ProteoScape can be applicable for both DDA and DIA analysis

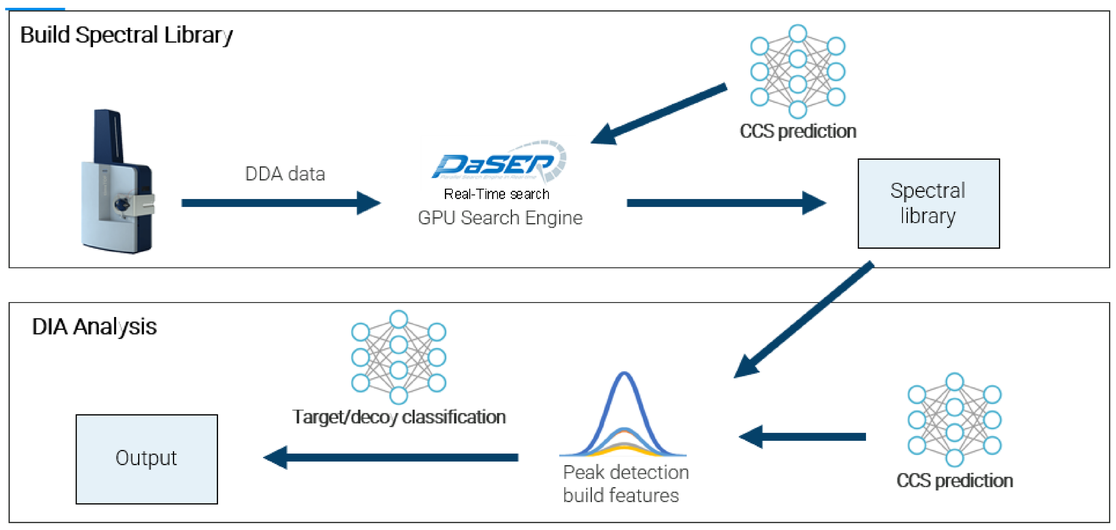

Fig. 2 The Spectral library search in DIA comprises two distinct modules. The first involves constructing a spectral library using the CCS prediction model derived from DDA searches. The second allows TIMS DIA-NN to utilize this spectral library to analyze DIA data via a classification model.

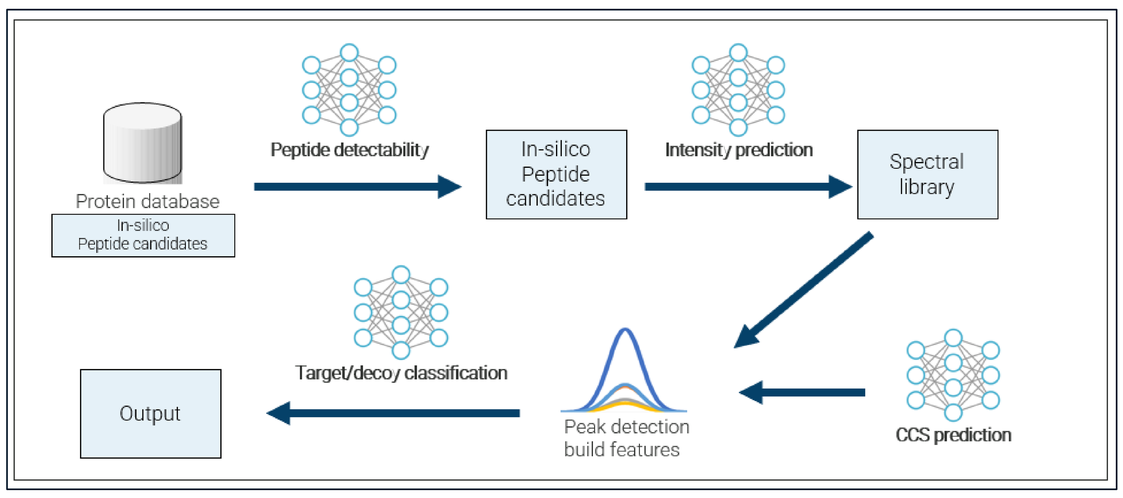

Fig. 3 The spectral library free search includes a peptide detectability module to initially reduce precursor candidates. Next, We apply an intensity prediction model. After generating a library-free spectral library, the subsequent analysis is identical to that of library-based data analysis.

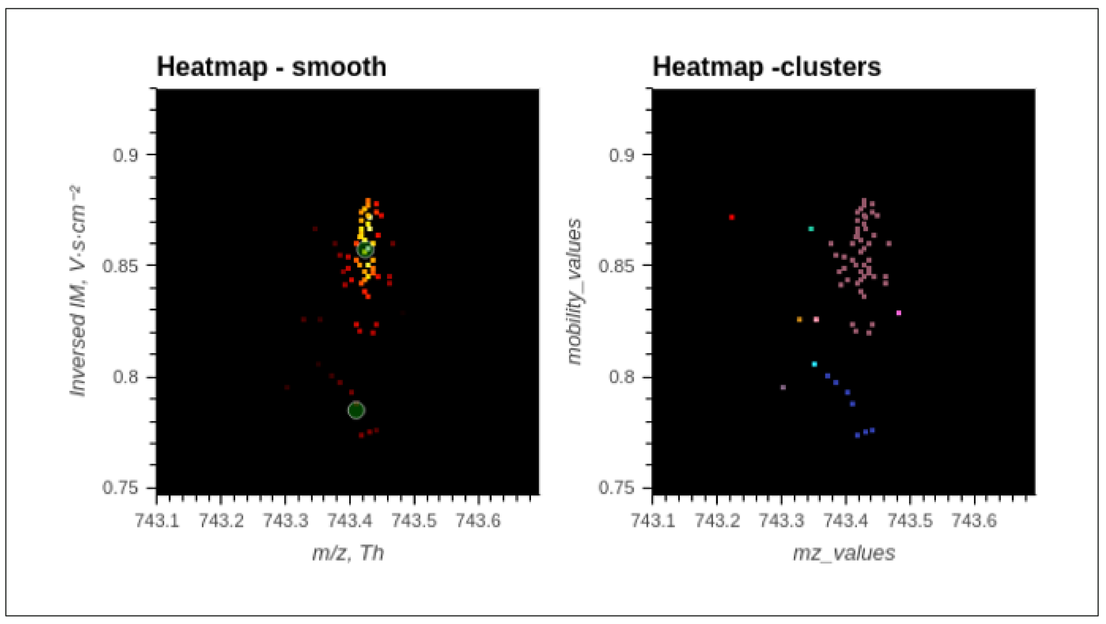

Fig. 4 The 4D peak finding algorithm applies a smoothing algorithm and then utilizes Gaussian fitting. It detects local maxima based on mz and 1/k0 resolutions

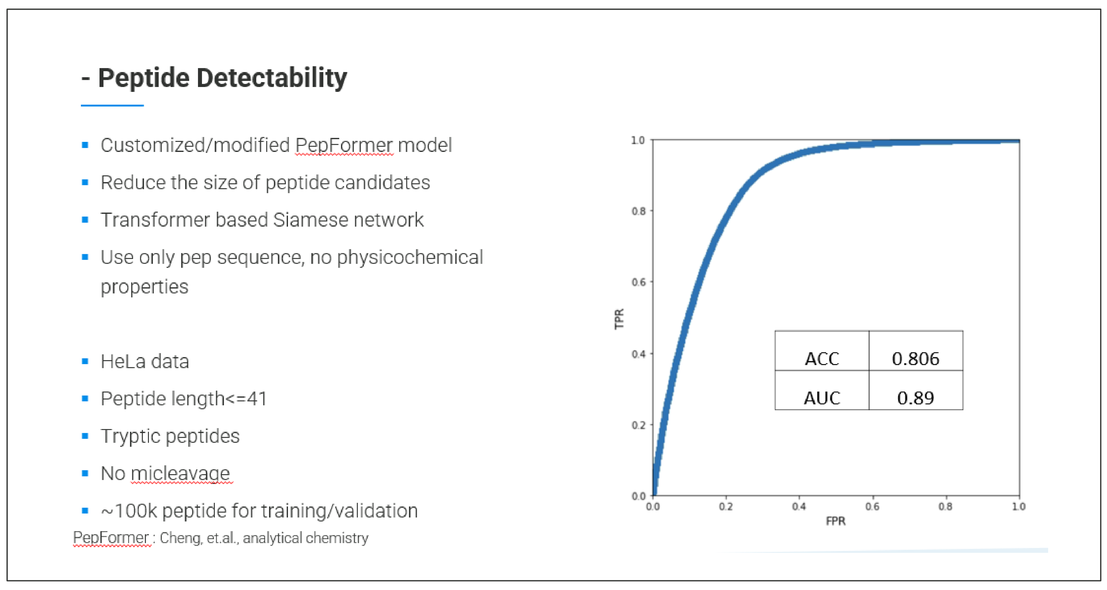

Fig. 5. A customized PepFormer model has been used on approximately 100 non-PTM peptides. The Peptide detectability model reduced the number of peptide candidates from a protein database. HeLa data were used with a maximum peptide length of 41.

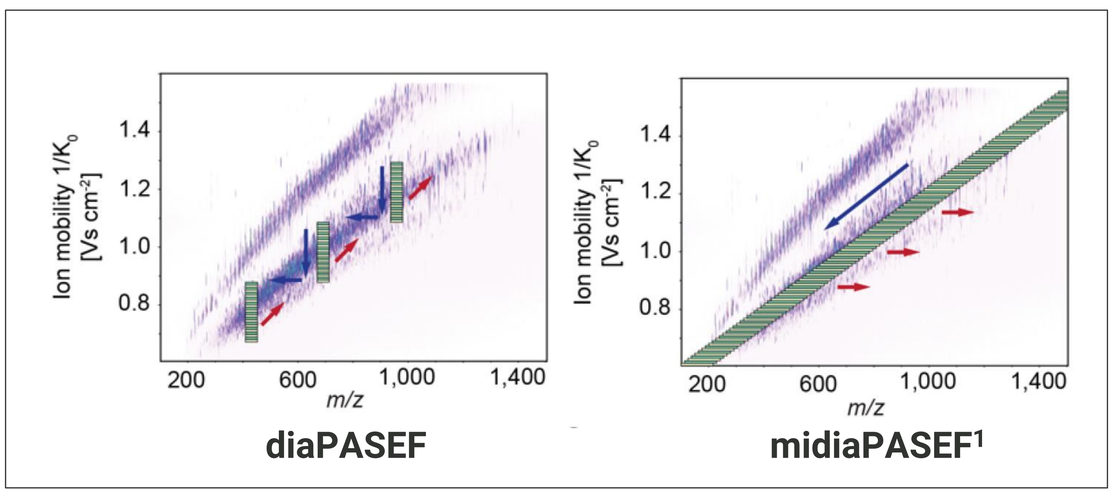

Fig. 6 In a recent publication, Distler et al. presented midiaPASEF, an innovative data-independent acquisition (DIA) scanning mode. This strategy employs diagonal scanning of the quadrupole with overlapping windows, enabling efficient coverage of the precursor ion cloud in both ion mobility and m/z dimensions. Currently, we are developing a new algorithm to analyze mediaPASEF data, aiming to improve both identification and quantification.

Conclusion

- TIMS DIA-NN, CCS-centric DIA analysis software has different types of deep learning modules
- TIMS DIA-NN has been integrated into ProteoScape
- We are working on a new algorithm to analyze midiaASEF data

Technology

References

1. Distler U. et. al,bioRxiv, 2023
2. Cheng, et.al., analytical chemistry, 2021
3. Zeng, et. al, N. comm, 2022