Application of artificial neural network to planar chromatography data

D. Fichou, G. Morlock   Justus Liebig University Giessen

JUSTUS-LIEBIG- UNIVERSITÄT GIESSEN — Food Science

JLU Study life Explore the world

Highlights

- Artificial neural network (ANN) applied for High-Performance Thin-Layer Chromatography (HPTLC), a separation technique in analytical chemistry (Fig. 1).
- Automatization of image evaluation with Restricted Boltzmann Machines (RBM)[1] instead of subjective interpretation of the output pictures.
- First mode of application: **Denoising and feature extraction** improved homogeneity of the background.
- Second mode of application: **Unsupervised sample classification** allowed an automated classification.
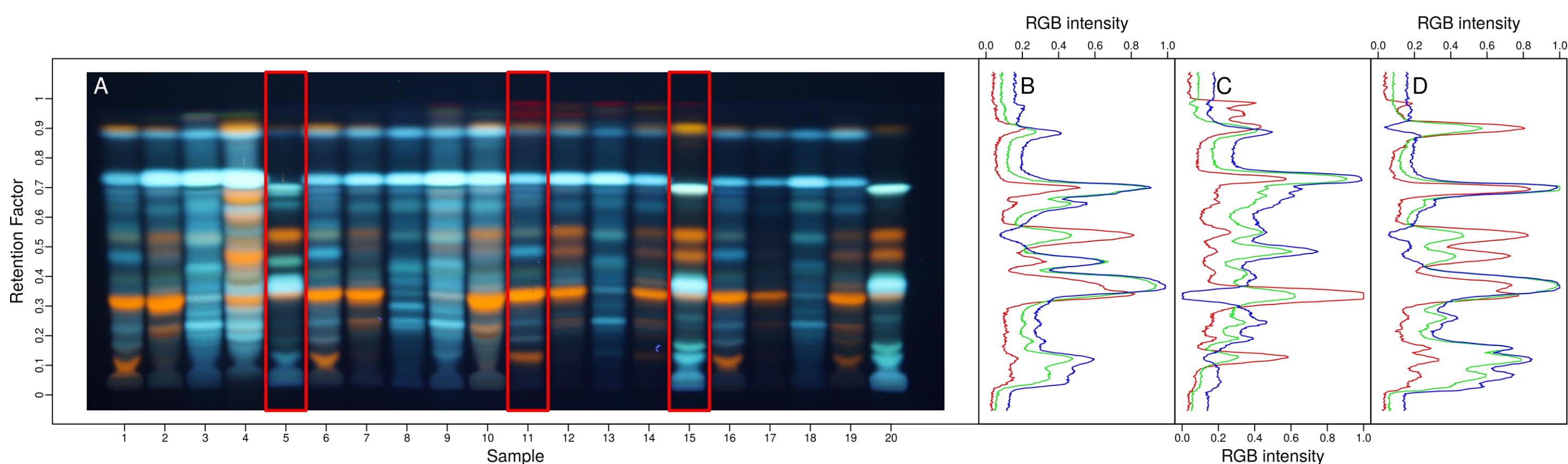

Fig. 1 HPTLC chromatograms with 20 samples of medicinal plant extracts separated in parallel (A); independent variables extracted from Red, Green and Blue channels of tracks 5, 11 and 15 (B to D)

Denoising and feature extraction

Bioautograms contained noise and background irregularities.

- Former approach: Median filter [2]

- New approach: Model trained on pixel patches.
  - → Dataset processed in an alternative way (Figs. 2 and 3).
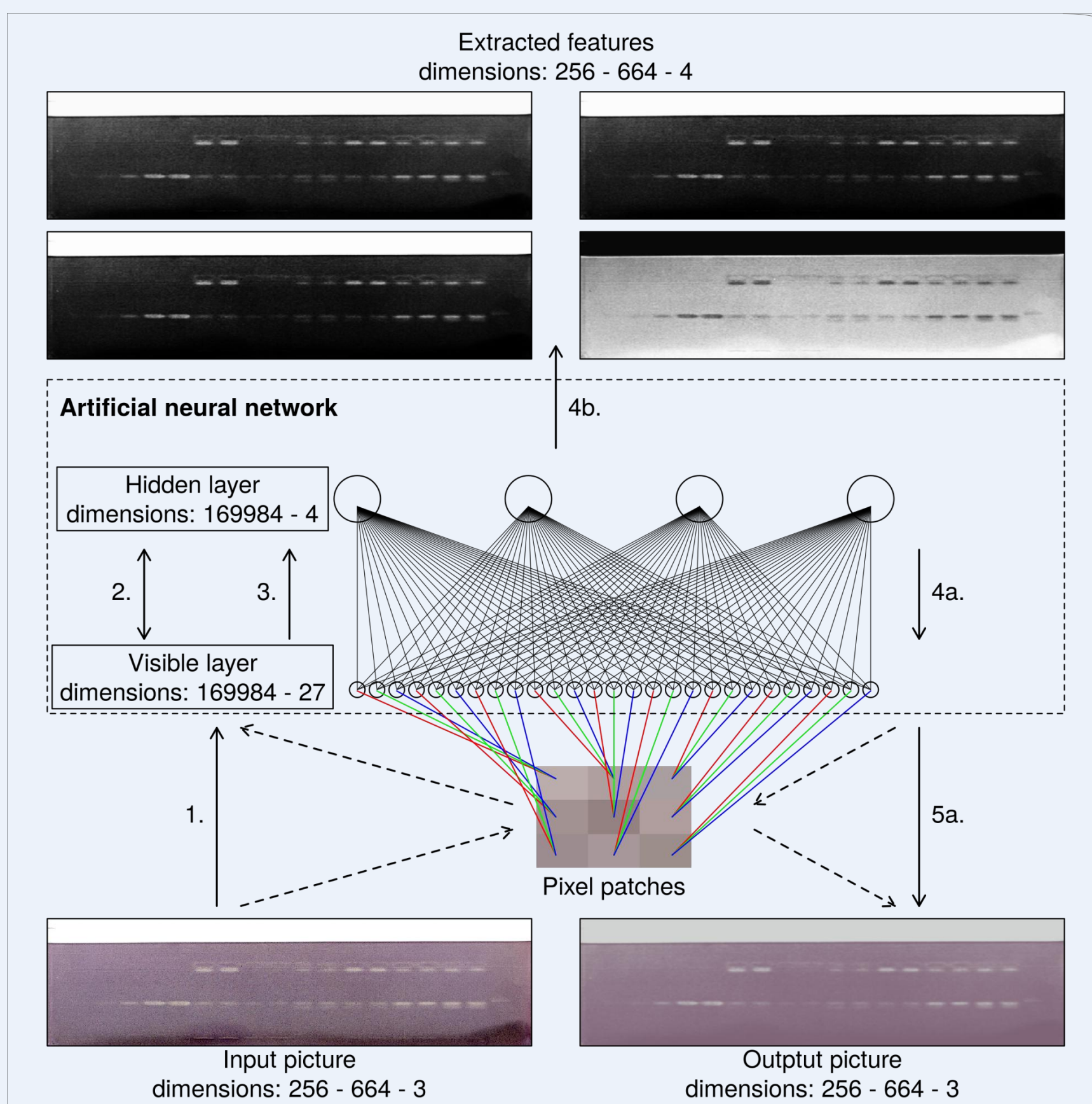 - → Feature extraction gave new point of view on data (Fig. 4).

1. Deconstruction from 3D array to matrix: Each pixel patch becomes one row of the new matrix.
2. Model training: Iteration of the weight matrix to maximize the likelihood.
3. Weight multiplication: Calculation of the hidden unit states.
4a. Weight multiplication: Calculation of the visible unit states.
5a. Reconstruction: Each row of the matrix becomes a pixel patch.
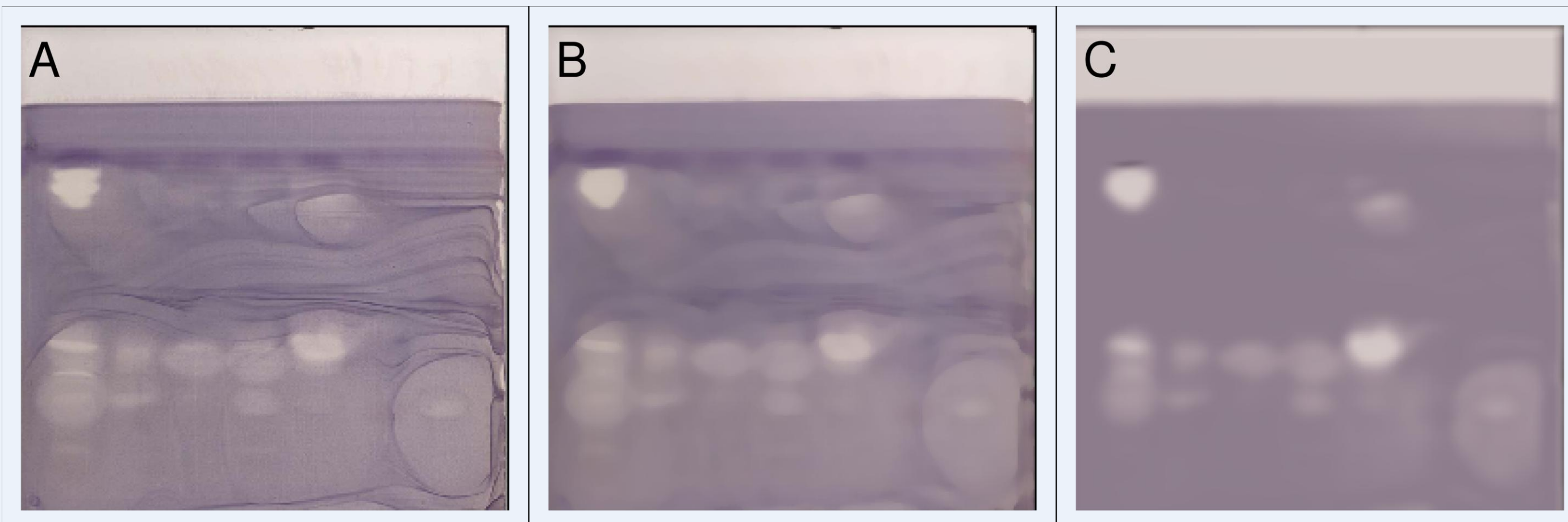4b. Feature extraction: Reconstruction into a new format.

Fig. 2 Denoising pipeline of a bioautographic experiment

Fig. 3 Algorithms' comparison: Original bioautograms (A) processed by median filter (B) and by ANN (C).
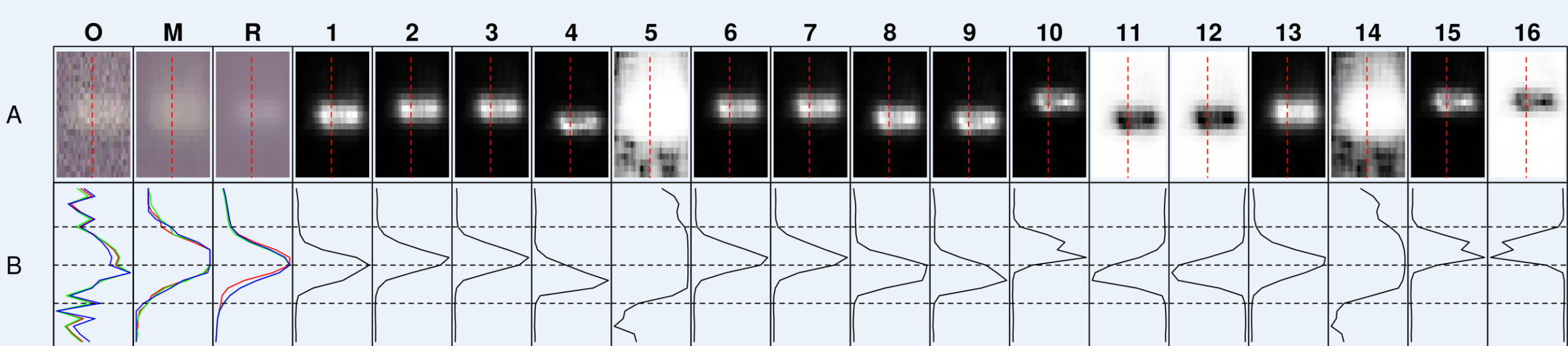

Fig. 4 Zoomed band (A) and normalized video densitograms (B): Original bioautogram (O) processed by median filter (M), by ANN followed by reconstruction (R) and normalized features extracted from each hidden unit states (1 to 16, specialized in discrimination of patterns)

Unsupervised sample classification

Sample classification is a subjective task.

- Former approach: Principal components analysis and hierarchical cluster [3]

- New approach: Model trained on verticale lines.
  - → Online learning, *i. e.* training one batch (chromatograms on one plate) at a time, to learn differencies between classes and to abstract experimental differences
  - → Staked hidden layers learned more abstract features and reproduced human choices (Fig. 5, Table 1).
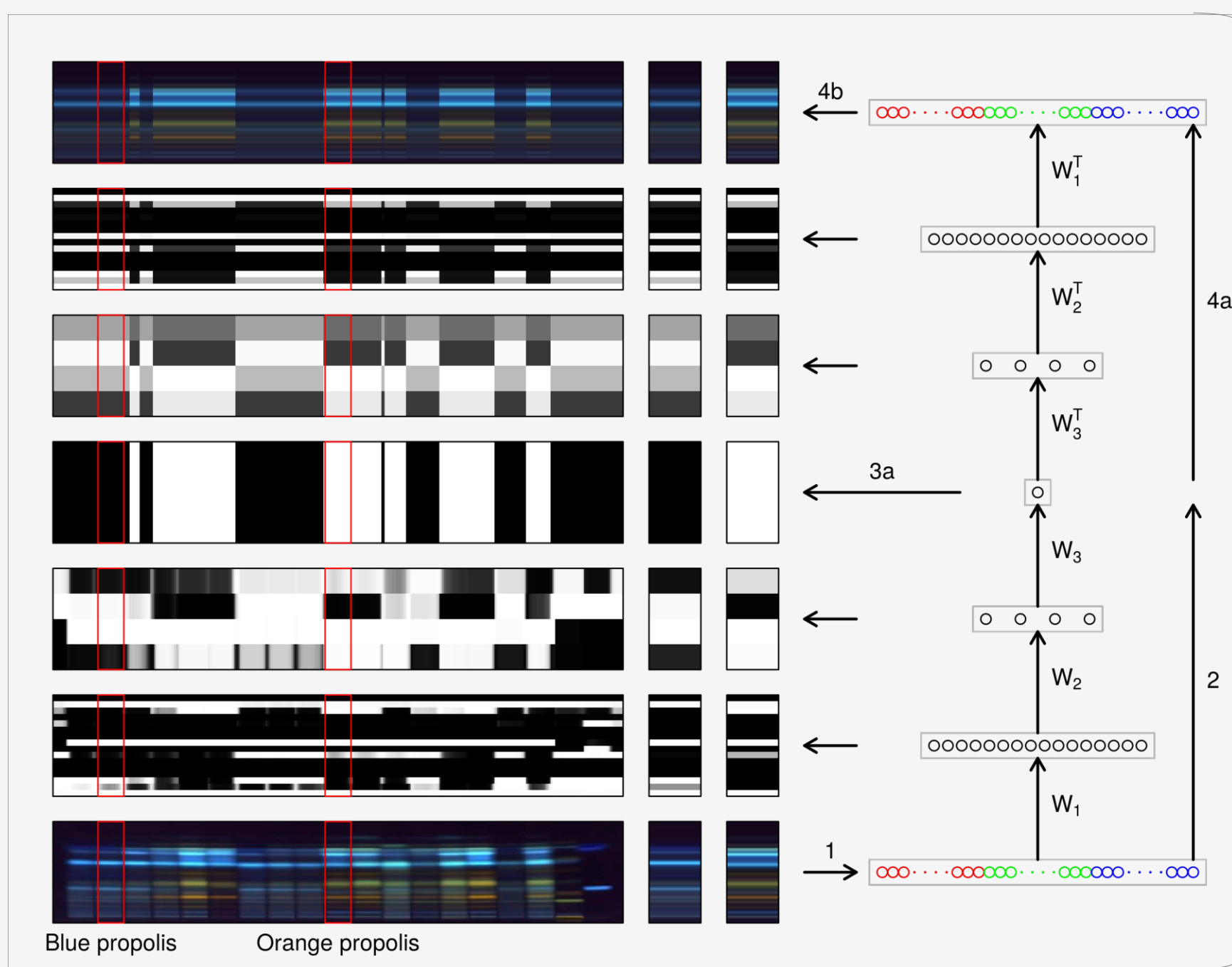 - → ANN generalized between batches and did not need preprocessing (Figs. 6 and 7).

1. Deconstruction: Each vertical line becomes one row of the new matrix.
2. Encoding: Staking of three RBM with decreasing number of units.
3a. Feature extraction: Last layer contained only one unit separating samples into two classes.
3b. Decoding: ANNs are crossed back.
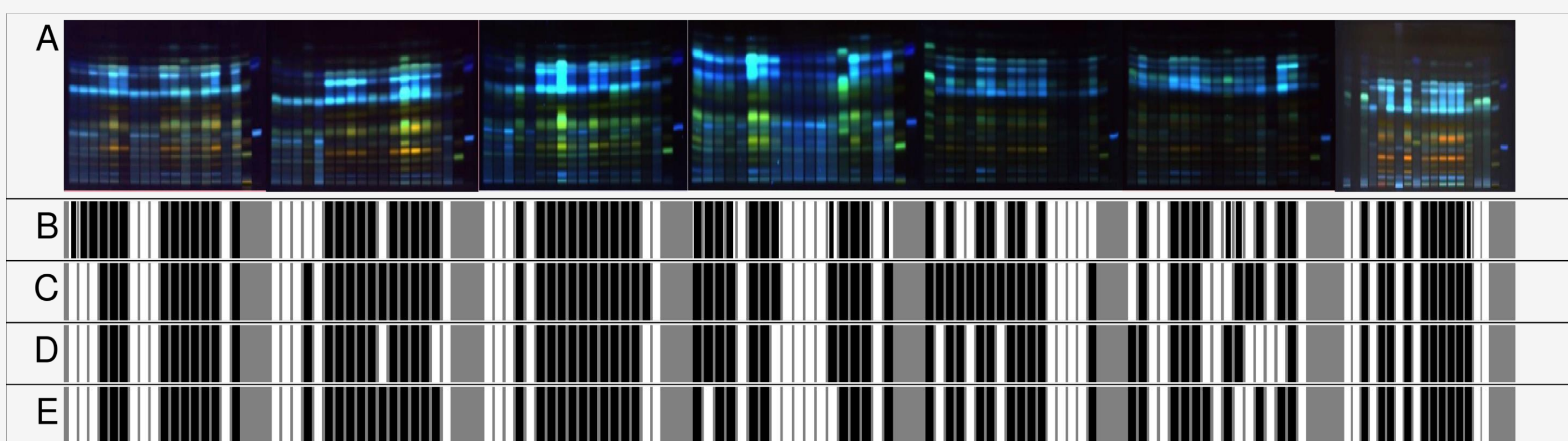4b. Reconstruction into the original format.

Fig. 5 Classification pipeline of one batch of German propolis

Fig. 6 Classification applied to entire dataset: Original batches (A), ANN classification (B), human classification (C to E).
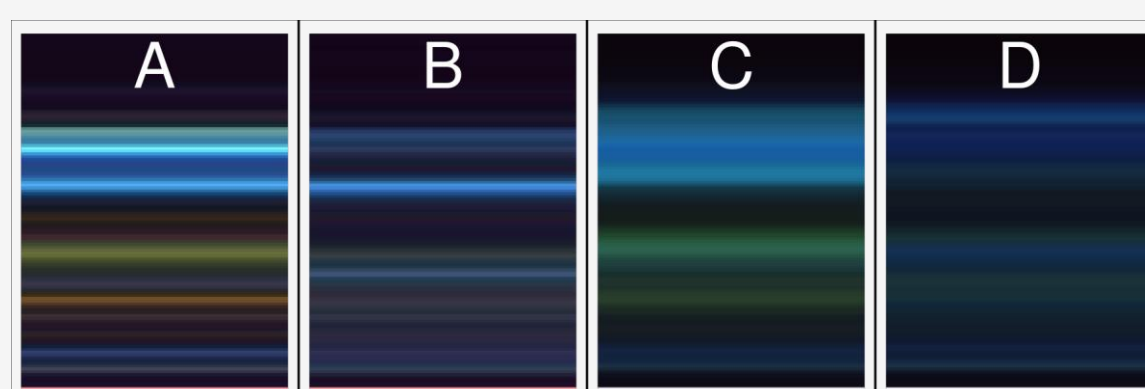

Fig. 7 Real (A, B) and reconstructed (C, D) HPTLC chromatograms of orange (A, C) and blue (B, D) types of propolis

	ANN	Human 1	Human 2
Human 1	86 %		
Human 2	87 %	90 %	
Human 3	88 %	90 %	93 %

Table 1 Accuracy between ANN and human choices

References [1] Hinton, G. E.; Salakhutdinov, R. R. Sciences 2006, 339, 1095–1099. [2] Komsta, Ł. Analytica Chimica Acta 2009, 641, 52–58. [3] Morlock, G. E.; Ristivojevic, P.; Chernetsova, E. S. Journal of Chromatography A 2014, 1328, 104–112.

R packages deepnet ; jpeg ; abind ; EBImage

IFZ Interdisziplinäres Forschungszentrum Giessen Giessen Research Centre for BioSystems, Land Use and Nutrition

Food Science, Justus Liebig University of Giessen, Institute of Nutritional Science, Heinrich-Buff-Ring 26, 35392 Giessen, Germany