Changes of mass spectra patterns on a brain tissue section revealed by deep learning with imaging mass spectrometry data

Changes of mass spectra patterns on a brain tissue section revealed by deep learning with imaging mass spectrometry data

Hidemoto Yamada[1*], Lili Xu[1*], Fumihiro Eto[1], Rei Takeichi[1], Ariful Islam[1], Md. AI Mamun[1], Chi Zhang[1], Ikuko Yao[1,2], Takumi Sakamoto[1,3], Shuhei Aramaki[1,4], Kenji Kikushima[1,3], Tomohito Sato[1,3], Yutaka Takahashi[1,3], Manabu Machida[5,6], Tomoaki Kahyo[1,3 †], Mitsutoshi Setou[1,3,5]

[1] Department of Cellular and Molecular Anatomy, Hamamatsu University School of Medicine, 1-20-1 Handayama, Higashi-ku, Hamamatsu, Shizuoka, 431-3192, Japan.

[2] Department of Biomedical Sciences, School of Biological and Environmental Sciences, Kwansei Gakuin University, 2-1 Gakuen, Sanda, Hyogo, 669-1337, Japan.

[3] International Mass Imaging Center, Hamamatsu University School of Medicine, 1-20-1 Handayama, Higashi-ku, Hamamatsu, Shizuoka, 431-3192, Japan.

[4] Department of radiation oncology, Hamamatsu University School of Medicine, 1-20-1 Handayama, Higashi-ku, Hamamatsu, Shizuoka, 431-3192, Japan.

[5] Department of Systems Molecular Anatomy, Institute for Medical Photonics Research, Preeminent Medical Photonics Education & Research Center, Hamamatsu University School of Medicine, 1-20-1 Handayama, Higashi-ku, Hamamatsu, Shizuoka, 431-3192, Japan.

[6] JST, PRESTO, Kawaguchi, Saitama 332-0012, Japan

* Both authors contributed equally to this work.

† Correspondence: kahyo@hama-med.ac.jp

ABSTRACT: The characteristic patterns of mass spectra in imaging mass spectrometry (IMS) strongly reflect the tissue environment. However, the boundaries formed where different tissue environments collide have not been visually assessed. In this study, IMS and convolutional neural network (CNN), one of the deep learning methods, were applied to the extraction of characteristic mass spectra patterns from training brain regions on rodents' brain sections. CNN produced classification models with high accuracy and low loss rate in any test datasets of mouse coronal sections measured by desorption electrospray ionization (DESI)-IMS, and mouse and rat sagittal sections by matrix-assisted laser desorption (MALDI)-IMS. Based on the extracted mass spectra pattern features, the histologically plausible segmentation and classification score imaging of the brain sections were obtained. The boundary imaging generated from classification scores showed the extreme changes of mass spectra patterns between the tissue environments, with no significant buffer zones for the intermediate state. The CNN-based analysis of IMS data is a useful tool for visually assessing the changes of mass spectra patterns on a tissue section, and it will contribute to a comprehensive view of the tissue environment.

■ INTRODUCTION

Tissues are composed of various molecules, including proteins, nucleic acids, and lipids, and the patterns of these components are a strong reflection of the tissue environment. Attempts have been made to understand the distribution of tissue environments, including comprehensive measurements of biomolecules with a micron-scale spatial resolution.

Imaging mass spectrometry (IMS) is one of the most suitable techniques for extracting mass spectra patterns from tissue. IMS is a two-dimensional mass spectrometry technique that generates information on the intra- and extra-cellular spatial distribution of biomolecules [1]. In the IMS measurements with a micron-scale spatial resolution on a millimeter-scale tissue section, the data with biomolecule mass and spatial distribution information is enormous. Therefore, the IMS researchers have attempted dimension reduction and subsequent advanced algorithms, such as clustering, to analyze the spatial distribution of mass spectra patterns. The autoencoder method enabled to extract the feature of mass spectra patterns and reduced the IMS data into a few components that capture the structural information of the sample [2]. IMS data from whole-body sagittal sections of mice were subject to t-distributed Stochastic Neighbor Embedding (t-SNE) analysis, and whole-body sections were segmented into different organs based on mass spectra patterns [3]. Our research group reported the existence of a cluster segmented into white matter and gray matter in a rat brain [4]. Orthogonal partial least squares discriminant analysis of IMS

data from mouse cerebellum showed that the clusters were segmented into white matter, molecular layer, granular layer, and Purkinje cell layer [5]. While these studies using unsupervised machine learning have indeed been able to classify mass spectra patterns using IMS data with high accuracy, the changes of mass spectra patterns between adjacent regions are not yet well understood.

Supervised machine learning has also been employed to extract the mass spectra patterns from IMS data. Support vector machine (SVM) and its related methodologies were able to segment the anatomical regions on IMS data and to distinguish between different grades of glioma [6, 7]. Moreover, Behrmann's group used a convolutional neural network (CNN) for the identification of lung cancer pathological type [8]. Application of deep neural network-based machine learning to IMS data can distinguish colorectal tumors from normal tissue [9]. Klein's group applied the modified SVM and the CNN to determine histopathological types of epithelial ovarian cancer [10]. While these approaches can be potent applications for diagnosis and clustering of MSI data of normal tissues was also performed [6–10], the boundary imaging between normal tissues has still been a challenging issue since in such regions cellular components are almost similar. To understand how tissue environments change, it is crucial to numerically distinguish regions with similar components.

In this study, CNN-based deep learning was applied to extract characteristic mass spectra patterns from the data of desorption electrospray ionization (DESI)-IMS and matrix-assisted laser desorption ionization (MALDI)-IMS of brain tissue sections. The classification scoring performed by deep learning on IMS data reveals an extreme change of mass spectra patterns.

■ EXPERIMENTAL SECTION

Animals. All experimental procedures were approved by the Ethics Committee of the Hamamatsu University School of Medicine (approval number: #2020027, #2017083) and carried out in accordance with the approved guidelines. C57BL/6 male mice (aged 4 months) and a Wister male rat (aged 8 weeks) were used for the IMS measurement (Table. S1).

Sample preparation. Mouse brains were collected following cervical dislocation and decapitation, quickly frozen in powdered dry ice, and stored at −80 ˚C. A rat brain was sampled under anesthesia using diethyl ether and then quick-frozen with powdered dry ice. The samples were mounted on a sample holder using an optimal cutting temperature compound (Sakura Finetek Japan, Tokyo) and sectioned with a thickness of 10 μm at −20 °C using a Cryostat (CM1950, Leica Microsystems K.K., Tokyo, Japan). For DESI-IMS, the resulting section was mounted onto un-coated glass slides. After cryo-section, the brain sections were kept at room temperature for a few minutes to remove extra water. For MALDI-IMS, the cryosection was mounted onto 100 Ω indium tin oxide (ITO)-coated glass slides (Matsunami Glass Industry Limited, Osaka, Japan), and 40 mg/mL (in 50% methanol) and 610.6 mg of DHB [11] were applied as the matrix onto the mouse and rat tissue sections by spraying using TM-sprayer (HTX Technologies, NC, USA) at 80 ℃, and as a 1.0-μm thick layer by sublimation using iM-Layer (Shimadzu Corporation, Kyoto, Japan) at 180 ℃, respectively.

DESI-IMS analysis. DESI-IMS data were acquired by using Xevo G2-XS quadrupole time-of-flight (Q-TOF) mass spectrometer (Waters, Milford, MA, USA) equipped with a 2D DESI source in negative ion mode following previously described method with slight modification [12]. Prior to measurement, DESI-IMS mass spectra were calibrated using 500 μM sodium formate solution prepared in 90% 2-propanol. As spray solvent, 98% methanol was delivered at a flow rate of 3 μL/min using ACQUITY UPLC binary solvent manager (Waters, Milford, MA, USA). Optimization of DESI source for better ionization and a stable signal from tissue samples was confirmed prior to measurement using the following parameters: capillary voltage, 4.0 kV; cone voltage, 50 eV; source temperature, 120 ℃; nebulizing nitrogen gas pressure, 4.0 bar; spatial resolution, 50 μm; incidence angle of the sprayer, 75 degrees; inlet to sprayer distance, about 7 mm; sample to sprayer distance, about 1.5 mm; scan speed, 200 μm/sec, *m/z* range: 100 –1000 and mass window of 0.02 Da.

MALDI-IMS analysis. Mouse and rat brain sections were subjected to 7 Tesla solarix XR, Fourier-transform ion cyclotron resonance mass spectrometer (FT-ICR-IMS, Bruker Daltonics, Billerica, MA, USA) with *m/z* ranges of 600–1100 and 650–1150, laser shot count at 200 and 500, time of flight to 0.7 ms and 1.0 ms, and laser power of 50% and 75%, respectively. Frequency at 2000 Hz, laser spot raster at 50 μm, the laser diameter value of approximately 25 μm, and the positive mode was used. Sodium trifluoroacetate (TFA, [M + Na]+) was used for the external calibration of FT-ICR-MS [13].

Handling of mass spectrum data. Mass spectrum raw data of DESI-IMS and MALDI-IMS were converted to CSV format, in which intensities from the detected peaks were recorded in each spot (data point), using IMAGE REVEAL (version 1.20.0.10960, Shimadzu, Kyoto, Japan), and SCiLS Lab (version 2015b, SCiLs, Bremen, Germany), respectively. This conversion process took about 1 hour for 18GB of raw IMS data of a rat brain tissue. The total numbers of spots and the detected peaks were listed in Table S1. To reduce the data volume, the top 100 or 300 peaks with the highest intensity in each spot were selected and integrated, which resulted in a net 2,000 to 3,000 peaks, similar to the pixel number of pictures in image recognition with deep learning (Figure. S1). Three or nine brain regions were anatomically defined as labels for supervised learning as follows; basal nucleus, cerebellum, corpus callosum, cortex, hippocampus, hypothalamus, brainstem, midbrain, thalamus (Table. S2). From each of the defined regions, 100 spots were randomly selected as a test data set, and the other spots in the defined region were used as a training data set. The spots in undefined regions were used for the classification by the trained model. These post-conversion processes took 21 hours, and the training and test data sets resulted in 135MB and 6.3MB, respectively, in the case of a rat brain tissue. For dimension reduction analysis using those data sets, principal component analysis (PCA) and t-SNE were conducted by calculating eigenvectors and by using the R package, Rtsne, respectively. Both of them were completed within 30 min.

Deep learning. CNN was applied to the learning of peak patterns. The peak intensities were converted to a 46 x 46 or 54 x 54 matrix in the order of *m/z* increasing. When the number of matrix elements exceeds the number of peaks, lacking elements were filled with zero value to handle mass spectrum data as in

the case of image data. Deep learning was conducted using Chainer frameworks based on Python (Preferred Networks, version 1.19). ResNet with 101 layers was applied to our model [14, 15], and Adam was used as an optimizer [16]. Based on the spots' numbers, the weight decay value and batch size for each sample varied in the range of 0.0001 to 0.02 and 70 to 140, respectively (Table. S1). Finally, CNN returns a classification score for each brain region from that matrix. A classification score for a brain region (i) was defined as a probability score (p_i), which was calculated in the CNN final layer as the following formula:

$$p_i = \exp(y_i) / \sum_{j=i}^{C} \exp(y_j)$$

The feature value (y_i) was from the second layer in CNN from the end. C is the number of brain regions used for the training. In the case of a rat brain tissue, deep learning and imaging took 17 and 9 hours, respectively. In the end, the conversion-to-imaging analysis was completed about 27 hours. The performance of the analysis machine was as follows: GPU, NVIDIA TITAN X 12GB GDDR5X; CPU, Intel Xeon E5-2603v4 1.7GHz with 6 Cores; and Memory, 64GB (16GB DDR4-2400x4). Up to 95% of GPU memory was used during deep learning.

■ RESULTS

Scheme of feature extraction, segmentation, and boundary imaging of mass spectra patterns. The mass spectrum is derived from each point of the IMS data, and consists of m/z peaks and intensity values (Figure. 1). IMS raw data include spot coordinates in addition to mass spectrum data. In our study using rodent brain sections, mass spectrum data were converted to the two-dimensional matrix of intensities in the order of *m/z* increasing in each spot (Figure. S2). Therefore, the converted data can be handled like the ones of greyscale images. CNN, commonly used for image recognition [17], was applied for feature extraction of the converted mass spectrum data. In this study, the feature is synonymous with characteristic mass spectra patterns. The features extracted from training spots were used to classify other tissue regions, resulting in segmentation and score distribution.

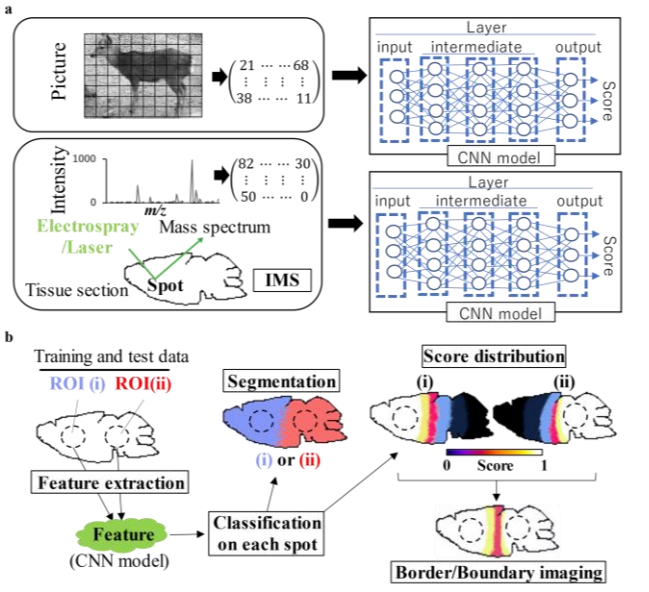

Figure 1. Scheme of feature extraction and classification of a brain section. (a) The mass spectral data acquired from each spot of the tissue section is converted into a matrix (tensor), like image data, and used for training for a convolutional neural network (CNN) model and classification. (b) Features of different regions of interest (ROIs), (i) and (ii), are extracted by a CNN. All spots are classified into a or b, and classification scores are calculated in each spot. Segmentation, the score distribution imaging, and border/boundary imaging are made by converting spot coordination to pixel format.

Classification of DESI-IMS data on a brain section by CNN model.

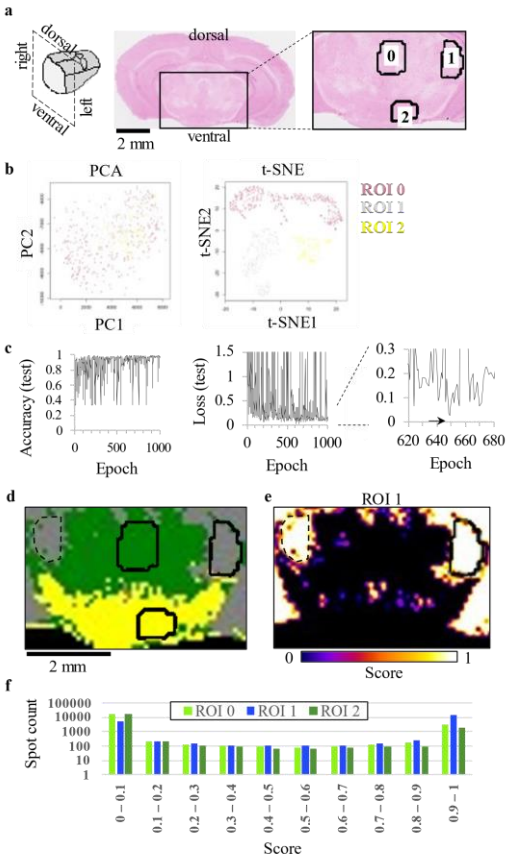

Figure 2. Classification of DESI-IMS data on a mouse brain section by CNN model. (a) A part of a brain region was assigned as ROIs (0–2) as shown in the area surrounded by a square. Spots within an ROI were randomly divided into training and test data sets. One coronal section of mouse brain was stained with Hematoxylin and eosin (H&E) after the measurement of DESI-IMS. (b) PCA and t-SNE were used as dimensionality reduction methods to get an overview of the training data. (c) Accuracy and loss of CNN learning with a test data set. The lowest loss value of the test data set was 0.026653, shown by arrow (epoch = 649). The learned model at that epoch was used for classification. (d) Each spot was classified with the learned CNN model. Areas surrounded by black solid and dotted lines show the selected ROIs for training and test datasets, and the area corresponding to the thalamus contralateral to ROI 1, respectively. (e) Score imaging of ROI 1 was representatively shown. (f) Histogram of score imaging. ROI 0: midbrain, green. ROI 1: thalamus, grey. ROI 2: hypothalamus, yellow.

A mouse brain coronal section was measured by DESI-IMS, which is a soft ionization method by spraying the sample with an electrically charged aqueous mist. Acquisition of training and test datasets is a crucial step of supervised machine learning. Three regions of interest (ROIs) were determined on a brain coronal section as training and tested datasets for CNN (Figure. 2a). The ROIs were selected from the midbrain, thalamus, and hypothalamus referring to a mouse brain atlas [18]. A random selection of one hundred spots from each ROI was used as a test dataset, and the remaining spots within the ROIs were used for training (Table. S2). To know the overview of the dataset, unsupervised machine learning was applied to the training dataset. Dimensional compression by PCA failed to produce the formation of clusters for each ROI, while in t-SNE, clusters formed can be explicitly classified into the three ROIs (Figure. 2b).

To create a versatile model of supervised machine learning using training data with ROIs 0-2, we performed CNN deep learning for 1000 epochs to estimate accuracy and loss. Here, accuracy refers to the ratio of correct values, and loss means a distance between the true values and the values predicted by the model. We generated a model from training data for each epoch (Figure. S3a) and evaluated the model with test data. The model with the lowest loss in the test was selected as the most versatile one here. As a result of learning and testing, it was observed that the convergences to accuracy = 1 with loss < 0.2 (Figure. 2c), and the lowest loss value of the test dataset was 0.026653 at epoch = 649 with an accuracy of 0.990. Therefore, the model of epoch = 649 was a highly versatile model obtained from CNN-based learning in this sample. We applied the CNN-based learned model (epoch = 649) to other areas to investigate whether we could obtain histologically plausible segmentations (Figure. 2d). Most of the spots around each ROI were classified into the same brain region as the ROI. In the coronal section, a region classified to be ROI 1 was identified on the left and right opposite sides of ROI 1, which corresponded to the thalamus. These results indicated that the segmentation of a brain section produced by DESI-IMS data and the CNN-based learned model was histologically plausible.

The classification into three defined brain regions was qualitative, and the certainty of the classification of each spot had not been evaluated. To evaluate the segmentation by the CNN-based learned model numerically, we used the classification scores, which were defined as probability scores of the determined brain regions calculated at the final CNN layer. The classification scores are numerical values from 0 to 1. One value is assigned to one brain region, and the sum of these scores is 1 in each spot. The distribution and histogram of the scores revealed a bimodal distribution with high (0.9 ≤) or almost zero scores for the classified brain regions (Figure. 2e, 2f, and S4a). Similar results were obtained from DESI-IMS measurements on a serial section of the section used in this experiment (Figure. S3b, S4b, S5). These results indicated that the IMS data were classified into brain regions with high certainty in the DESI-IMS analysis by the CNN-learned model.

Classification of MALD-IMS data on a brain section by CNN model. We investigated whether our CNN-based analysis of IMS data was a useful tool to segmentate normal brain regions even for data obtained by another principle for producing ions. A mouse brain sagittal section was measured by MALDI-IMS, which was irradiated with a pulsed laser beam. Three ROIs were determined from basal nuclei, thalamus, and hypothalamus on a brain sagittal section as CNN training and test datasets (Figure. 3a).

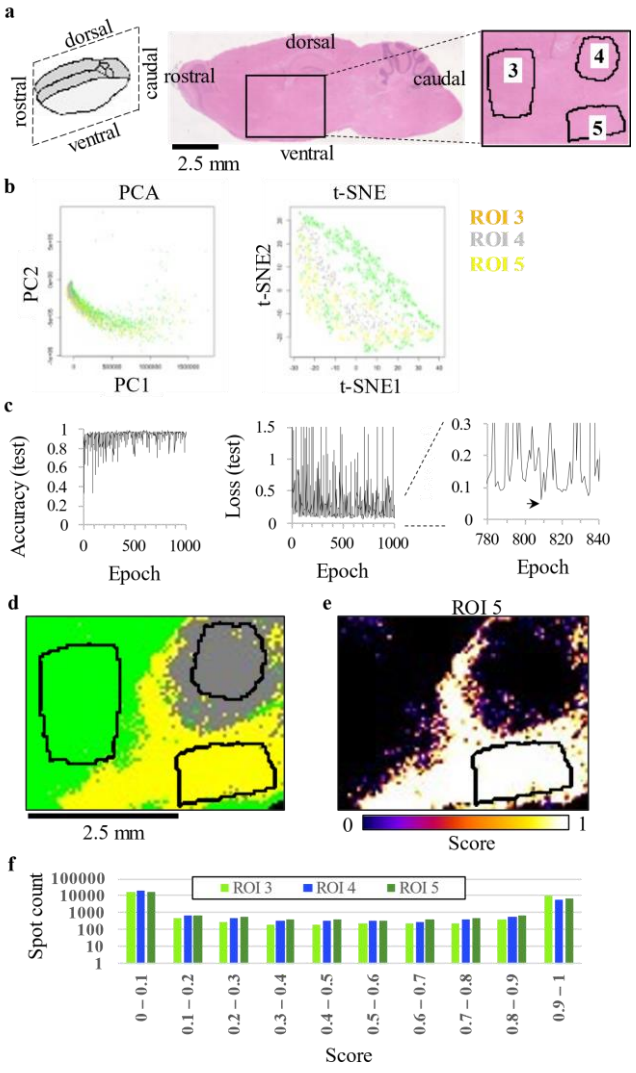

Figure 3. Classification of MALDI-IMS data on a mouse brain section by CNN model. (a) A part of a brain region was assigned as ROIs (3–5) as shown in the area surrounded by a square. Spots within an ROI were randomly divided into training and test data sets. A mouse brain sagittal section was stained with H&E after the measurement of MALDI-IMS. (b) PCA and t-SNE were used as dimensionality reduction methods to get an overview of the training data. (c) Accuracy and loss of CNN learning with a test data set. The lowest loss value of the test data set was 0.060882, shown by arrow (epoch = 809). (d) Each spot was classified with the learned CNN model. Areas surrounded by black solid lines show the selected ROIs for training and test datasets. (e) Score imaging of ROI 5 was representatively shown. (f) Histogram of score imaging. ROI 3: basal nuclei, lime. ROI 4: thalamus, grey. ROI 5: hypothalamus, yellow.

As in the DESI-IMS analysis, training and test data sets were determined from each ROI. Dimensional compression by PCA failed to produce the formation of clusters for each ROI. In the

t-SNE, each ROI roughly formed clusters, but with some ROIs overlapping one another (Figure. 3b).

We conducted CNN deep learning using the training and test datasets. As in the case of DESI-IMS, it was observed that the convergences to accuracy = 1 with loss < 0.2 (Figure. 3c), and the lowest loss value of the test dataset was 0.060882 at epoch = 809, in which accuracy was 0.977. We applied the CNN-based learned model (epoch = 809) to other areas (Figure. 3d). Most of the spots around each ROI were classified into the same brain region as the ROI, and the segmentation of a brain section produced by MALDI-IMS data was histologically plausible.

The distribution and histogram of the scores revealed a bimodal distribution with high (0.9 ≤) or almost zero scores for the classified brain regions (Figure. 3e, 3f, and S4c). As in the DESI-IMS analysis, the IMS data were classified into brain regions with high certainty in the MALDI-IMS analysis by the CNN-learned model.

Classification of MALD-IMS data on a whole-brain section by CNN model. We tried to conduct a CNN-based analysis of IMS data by increasing the target brain regions. A rat whole brain sagittal section was measured by MALDI-IMS. Nine ROIs were determined from basal nuclei, cerebellum, corpus callosum, cerebral cortex, hippocampus, hypothalamus, pons, midbrain, and thalamus as CNN training and test datasets referring to a rat brain atlas (Figure. 4a) [19]. In the PCA method, there were considerable overlaps among the ROIs, while in t-SNE, the number of overlaps between ROIs decreased compared with PCA. In particular, the overlaps of ROI 1 (cerebellum) with other regions, which were noticeable in PCA, were reduced in t-SNE (Figure. 4b). However, the partial overlaps between ROI 5 (hypothalamus) and other ROIs, between ROIs 3 and 4 (cortex and hippocampus), and between ROIs 6 and 7 (brainstem and midbrain) were observed in both PCA and t-SNE.

We conducted CNN deep learning using the training and test datasets from nine ROIs. It was observed that the convergences to accuracy = 1 with loss < 0.001 (Figure. 4c), and the lowest loss value of the test dataset was 0.001052 at epoch = 970, in which accuracy was 1.000. We applied the CNN-based learned model (epoch = 970) to the whole tissue section (Figure. 4d). Most of the spots around each ROI were classified into the same brain region as the ROI, and the segmentation of a whole-brain section was histologically plausible.

The distribution and histogram of the scores revealed a bimodal distribution with high (0.9 ≤) or almost zero scores for the classified brain regions (Figure. 4e, 4f, and S4d). This result showed that the IMS data were classified even on a whole-brain section with high certainty by the CNN-learned model.

From Figure. 4, we can see the regions segmented by the CNN-based learned model had clear boundaries, and to further substantiate this point, we selected the highest scores among the brain regions identified at each spot and integrated the highest score into a single image (Figure. S6). In any of the three measurements of DESI-IMS and MALDI-IMS, there was no buffer zone among brain regions where the highest score gradually decreased with intermediate scores (0.5-0.7) toward adjacent brain regions, and the intermediate scores were mainly observed as thin lines separating the brain regions (Figure. S6b–S6d). The plots of classification scores across the borders showed extreme changes between the ROIs. These results indicated that the CNN-based learned model segmented the brain regions with clear boundaries.

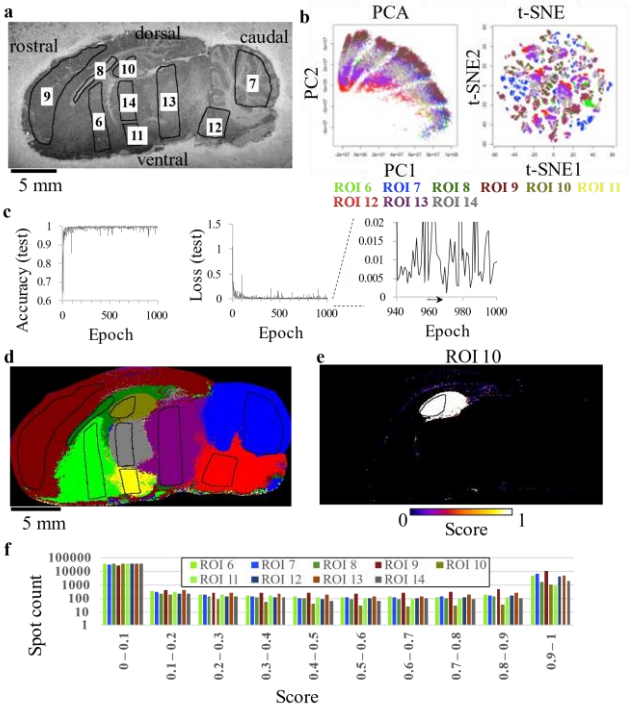

Figure 4. Classification of MALDI-IMS data on a rat brain section by CNN model. (a) A rat brain sagittal section was measured for MALDI-IMS. A part of a brain region was assigned as ROIs (6–14). Spots within an ROI were randomly divided into training and test data sets. An unstained tissue picture was shown. (b) PCA and t-SNE were used as dimensionality reduction methods to get an overview of the training data. (c) Accuracy and loss of CNN learning with a test data set. The lowest loss value of the test data set was 0.001052, shown by arrow (epoch = 970). (d) Each spot was classified with the learned CNN model. Areas surrounded by black solid lines show the selected ROIs for training and test datasets. (e) Score imaging of ROI10 was representatively shown. (f) Histogram of score imaging. ROI 6: basal nuclei, lime. ROI 7: cerebellum, blue. ROI 8: corpus callosum, green. ROI 9: cerebral cortex, maroon. ROI 10: hippocampus, olive. ROI 11: hypothalamus, yellow. ROI 12: pons, red. ROI 13: midbrain, purple. ROI 14: thalamus, grey.

■ DISCUSSION

This research displays a deep neural network-based approach capable of showing extreme changes in mass spectra patterns between the tissue environments. We have measured on various IMS brain datasets and our results fully demonstrate the advantages of employing CNNs.

Higher spatial and mass resolutions are being implemented in IMS analysis, and at the same time, huge amounts of data are being produced. The challenge is how to utilize the ever-increasing amount of data [20]. The mass spectrum acquired for each spot in IMS is a compilation of features [21], but these features do not provide us with direct access to the regional classification information of the sample. It is too difficult to classify

brain regions simply by staring at the mass spectrum of spots in the ROI (Figure. S7). Compared to conventional IMS analysis, in which the data are manually handled, deep learning methods, such as CNN, enable us to extract features for large-scale data analysis. Feature extraction is automatically performed by targeting complex spectral patterns without the need for researchers to worry about selecting biased target ions and without the hassle of selecting from a vast amount of data.

Multivariate analysis, including supervised and unsupervised methods, commits to identifying available complex spectrum patterns. In this study, we used CNN to extract the features of mass spectra patterns for the classification of brain regions and also employed PCA and t-SNE to get an overview of the selected data set with dimensional compression before conducting CNN. t-SNE outperformed PCA with less overlap in the classified areas. t-SNE features a non-linear function and a robust ability to process outliers [22], making it among the best dimension reduction methods and constituting a reason for better performance than PCA analysis. These two algorithms don't extract mass spectra patterns automatically but are easy to introduce compared with the deep learning supervised method because of the relatively low calculation cost. If clusters do not form at all even after dimensional compression, then we have the opportunity to reconsider any issues with the way the data set is taken or measured. It would be better to apply unsupervised machine learning in order to grasp the IMS data.

When comparing our method to other clustering methods that deal with mass spectral patterns, we need to do a trade-off between time and expense cost versus “deep” analysis. IMS data is multi-dimensional, and with the application of IMS techniques to biological materials, the IMS data contains thousands to tens of thousands of peaks in a single irradiation point, including peaks that are difficult to determine whether they are noise or not. When extracting a characteristic mass spectral pattern, it is up to the purpose how "deep" to incorporate these peaks. In the case of comparing tissues with similar cellular components, such as our subject, deep learning would be of high value since it offers a high degree of freedom to design parameters and algorithms and enables more flexible feature extraction, resulting in potentially extracting features "deeper" than conventional clustering methods.

We conducted sample preparations and IMS measurements under the various conditions: coronal and sagittal sections; auto spray and sublimation of matrix; mouse and rat; and DESI and MALDI, and similar results were obtained. Although we will need to test more conditions, our analysis method shows a certain degree of robustness that suggests its universal utility. On the other hand, the *m/z* range and pitch size may still need to be validated. In this experiment, IMS measurements were performed in the *m/z* range suitable for lipid detection, as the brain is rich in lipids. The pitch size was determined by the performance of the IMS device and the physical memory required for deep learning (VRAM 12GB). The *m/z* range and pitch size would need to be adjusted accordingly to the researcher's objectives and the facility's environment.

Although 2000–3000 peaks were selected for the deep learning, the presence of adduct ions and isotopic peaks must be taken into account, so that several hundred molecular species may contribute to the feature extraction. We performed no corrections for the mass shift between spots in this study for feature extraction. One advantage of CNN-based deep learning of IMS data would be that it does not need to consider exact mass accuracy between spots. Image recognition is an area where CNN-based deep learning is in active use [17]. The mass shift between the spots on IMS data corresponds to the spatial shift of pixels on image data. Even if mass shifts are observed between spots, they will be automatically adjusted during the CNN process, in which many parameters are used for weighting and contraction of information, resulting in tolerance to mass shift in the same way as it allows in deep learning for the spatial position and shape of objects on image data. Also, the selection of ROI is determined by the different histological area sizes of the regions of the brain. If we make a similar selection of ROI size for the purpose of balance, we need to scale down data size by the smallest region, ROI 8. Smaller sizes of ROI become lower representatives for the whole region, which may cause difficulty for CNN to learn and extract the feature from large regions, while there may be cases where it is better to balance ROI sizes, especially when all tissue areas are similar in size. A method in which a small number of spots are randomly selected as a data set, even over a large area, could be considered, but scattered spots make it difficult to evaluate them as imaging.

Our method, which allows visual assessment of mass spectra patterns on a tissue section, helps to understand changes in the tissue environment. For example, metaplasia, a transformation of one differentiated cell type to another differentiated cell type, is caused by some sort of abnormal stimulus. Recently, it has been theorized that metaplasia may be caused by the migration of a cell population from a distant tissue region [23]. If an intermediate state region is determined by the mass spectra pattern, it would be caused by a stimulus, not migration. On the other hand, there may be a demand to know the profile of the molecular species that contribute to the extracted features. It is a hot topic in the field of deep learning research to express the extracted feature in a way that researchers can understand [24]. Saliency mapping is one of the methods [25, 26]. In the future, such a method will be a bridge between deep learning and molecular biology.

ASSOCIATED CONTENT

Supporting Information

Supporting figures: Figure S1: Data handling to reduce data volume; Figure S2: Plot for the numbers of selected top *m/z* peaks per spot and total net peaks; Figure S3: Accuracy and loss of CNN learning with a training data set; Figure S4: Score imaging of each ROI provided by the learned CNN model; Figure S5: Classification of DESI-IMS data on a serial brain section by CNN model; Figure S6: Boundary imaging by CNN-based analysis of IMS data; Figure S7: Mass spectra of spots within ROIs in a rat brain tissue section measured by MALDI-IMS.
Supporting tables: Table S1: Sample information and deep learning condition; Table S2: The number of spots used for training and test.

AUTHOR INFORMATION

Corresponding Author

† Correspondence: Tomoaki Kahyo
Email: kahyo@hama-med.ac.jp

Author Contributions

T. K. and M. S. designed the study. H. Y., F. E., A. I., and R. T. got the data and handled imaging mass spectrometry data. T. K. contributed to data conversion and the deep learning. T. K., L. X. and H. Y. prepared the manuscript discussing with A. I., M. A. M., C. Z., Y. I., S. T., K. K., T. S., T. Y., M. M. and M. S.

Notes

The authors declare no conflict of interest.

All experiments in this study were performed in accordance with the guidelines issued by the Institutional Animal Care and Use Committees of Hamamatsu University, School of Medicine, Japan. All experiments were performed with approval from the institutional review board.

ACKNOWLEDGMENT

This work was supported by MEXT Project for promoting public utilization of advanced research infrastructure (Imaging Platform), Grand Number JPMXS0410300222, AMED under Grant Number JP21ak010101179, JSPS KAKENHI Grant Numbers JP18H05268, and Kanehara Ichiro Memorial Medical Foundation, Grant Number 16K147. L. X. owes sincere thanks to the Uehara Memorial Foundation Research Fellowship Program.

REFERENCES

1. McDonnell, L.A., Heeren, R.M.A.: Imaging mass spectrometry. Mass Spectrom Rev. 26, 606–643 (2007). https://doi.org/10.1002/mas.20124
2. Thomas, S.A., Race, A.M., Steven, R.T., Gilmore, I.S., Bunch, J.: Dimensionality reduction of mass spectrometry imaging data using autoencoders. 2016 IEEE Symp. Ser. Comput. Intell. SSCI 2016. (2017). https://doi.org/10.1109/SSCI.2016.7849863
3. Song, W., Wang, L., Liu, P., Choo, K.K.R.: Improved t-SNE based manifold dimensional reduction for remote sensing data processing. Multimed. Tools Appl. 78, 4311–4326 (2019). https://doi.org/10.1007/s11042-018-5715-0
4. Nampei, M., Horikawa, M., Ishizu, K., Yamazaki, F., Yamada, H., Kahyo, T., Setou, M.: Unsupervised machine learning using an imaging mass spectrometry dataset automatically reassembles grey and white matter. Sci. Rep. 9, 1–11 (2019). https://doi.org/10.1038/s41598-019-49819-1
5. Wehrli, P.M., Michno, W., Blennow, K., Zetterberg, H., Hanrieder, J.: Chemometric Strategies for Sensitive Annotation and Validation of Anatomical Regions of Interest in Complex Imaging Mass Spectrometry Data. J. Am. Soc. Mass Spectrom. 30, 2278–2288 (2019). https://doi.org/10.1007/s13361-019-02327-y
6. Eberlin, L.S., Norton, I., Dill, A.L., Golby, A.J., Ligon, K.L., Santagata, S., Graham Cooks, R., Agar, N.Y.R.: Classifying human brain tumors by lipid imaging with mass spectrometry. Cancer Res. 72, 645–654 (2012). https://doi.org/10.1158/0008-5472.CAN-11-2465
7. Luts, J., Ojeda, F., Van de Plas Raf, R., De Moor, B., Van Huffel, S., Suykens, J.A.K.: A tutorial on support vector machine-based methods for classification problems in chemometrics. Anal. Chim. Acta. 665, 129–145 (2010). https://doi.org/10.1016/j.aca.2010.03.030
8. Behrmann, J., Etmann, C., Boskamp, T., Casadonte, R., Kriegsmann, J., Maaß, P.: Deep learning for tumor classification in imaging mass spectrometry. Bioinformatics. 34, 1215–1223 (2018). https://doi.org/10.1093/bioinformatics/btx724
9. Mittal, P., Condina, M.R., Klingler-Hoffmann, M., Kaur, G., Oehler, M.K., Sieber, O.M., Palmieri, M., Kommoss, S., Brucker, S., McDonnell, M.D., Hoffmann, P.: Cancer tissue classification using supervised machine learning applied to maldi mass spectrometry imaging. Cancers (Basel). 13, (2021). https://doi.org/10.3390/cancers13215388
10. Klein, O., Kanter, F., Kulbe, H., Jank, P., Denkert, C., Nebrich, G., Schmitt, W.D., Wu, Z., Kunze, C.A., Sehouli, J., Darb-Esfahani, S., Braicu, I., Lellmann, J., Thiele, H., Taube, E.T.: MALDI-Imaging for classification of epithelial ovarian cancer histotypes from a tissue microarray using machine learning methods. (2018). https://doi.org/10.1002/prca.201700181.
11. Shimizu, Y., Satou, M., Hayashi, K., Nakamura, Y., Fujimaki, M., Horibata, Y., Ando, H., Watanabe, T., Shiobara, T., Chibana, K., Takemasa, A., Sugimoto, H., Anzai, N., Ishii, Y.: Matrix-assisted laser desorption/ionization imaging mass spectrometry reveals changes of phospholipid distribution in induced pluripotent stem cell colony differentiation. Anal. Bioanal. Chem. 409, 1007–1016 (2017). https://doi.org/10.1007/s00216-016-0015-x
12. Takeda, R., Islam, A., Sato, T., Kurita, H., Kahyo, T., Urano, T., Setou, M.: The stability of the metabolic turnover of arachidonic acid in human unruptured intracranial aneurysmal walls is sustained. Clin. Neurol. Neurosurg. 208, 106881 (2021). https://doi.org/10.1016/j.clineuro.2021.106881
13. Moini, M., Jones, B.L., Rogers, R.M., Jiang, L.: Sodium Trifluoroacetate as a Tune / Calibration the Mass Range of 100 – 4000 Da. 0305, (1998)
14. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2016-Decem, 770–778 (2016). https://doi.org/10.1109/CVPR.2016.90
15. Yasunori, K.: Deep Networks with Stochastic Depth implementation by Chainer, https://github.com/yasunorikudo/chainer-ResDrop, (2016)
16. Kingma, D.P., Ba, J.L.: Adam: A method for stochastic optimization. 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc. 1–15 (2015)
17. Traore, B.B., Kamsu-Foguem, B., Tangara, F.: Deep convolution neural network for image recognition. Ecol. Inform. 48, 257–268 (2018). https://doi.org/10.1016/j.ecoinf.2018.10.002
18. Keith, F., George, P.: Paxinos and Franklin's the Mouse Brain in Stereotaxic Coordinates, Compact. Academic Press (2019)
19. George, P., Watson, C.: The Rat Brain in Stereotaxic Coordinates. Academic Press (2013)
20. Race, A.M., Steven, R.T., Palmer, A.D., Styles, I.B., Bunch, J.: Memory efficient principal component analysis for the dimensionality reduction of large mass spectrometry imaging data sets. Anal. Chem. 85, 3071–3078 (2013). https://doi.org/10.1021/ac302528v
21. Guo, D., Föll, M.C., Volkmann, V., Enderle-Ammour, K., Bronsert, P., Schilling, O., Vitek, O.: Deep multiple instance learning classifies subtissue locations in mass spectrometry images from tissue-level annotations. Bioinformatics. 36, I300–I308 (2020). https://doi.org/10.1093/BIOINFORMATICS/BTAA436
22. Li, W., Cerise, J.E., Yang, Y., Han, H.: Application of t-SNE to human genetic data. J. Bioinform. Comput. Biol. 15, 1–14 (2017). https://doi.org/10.1142/S0219720017500172
23. Xian, W., Duleba, M., Zhang, Y., Yamamoto, Y., Ho, K.Y., Crum, C., McKeon, F.: The Cellular Origin of Barrett's Esophagus and Its Stem Cells. Adv. Exp. Med. Biol. 1123, 55–69 (2019). https://doi.org/10.1007/978-3-030-11096-3_5
24. Mahendran, A., Vedaldi, A.: Understanding_Deep_Image_2015_CVPR_paper. Cvpr. (2015)
25. Smilkov, D., Thorat, N., Kim, B., Viégas, F., Wattenberg, M.: SmoothGrad: Removing noise by adding noise. arXiv. (2017)
26. Sundararajan, M., Taly, A., Yan, Q.: Axiomatic attribution for deep networks. arXiv. (2017)

Authors are required to submit a graphic entry for the Table of Contents (TOC) that, in conjunction with the manuscript title, should give the reader a representative idea of one of the following: A key structure, reaction, equation, concept, or theorem, etc., that is discussed in the manuscript. Consult the journal's Instructions for Authors for TOC graphic specifications.

Insert Table of Contents artwork here