

Divide-and-conquer: a flexible deep learning strategy for exploring metabolic heterogeneity from mass spectrometry imaging data

Lei Guo ^{a,#}, Jiyang Dong ^{a,#,*}, Xiangnan Xu ^d, Zhichao Wu ^e, Yinbin Zhang ^f, Yongwei Wang ^g, Pengfei Li ^g, Chao Zhao ^{b,c,*}, Zongwei Cai ^{c,*}

^a National Institute for Data Science in Health and Medicine, Department of Electronic Science, Xiamen University, Xiamen, China

^b Bionic Sensing and Intelligence Center, Institute of Biomedical and Health Engineering, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

^c State Key Laboratory of Environmental and Biological Analysis, Department of Chemistry, Hong Kong Baptist University, Hong Kong SAR, China

^d School of Mathematics and Statistics, The University of Sydney, NSW 2006, Australia.

^e School of Artificial Intelligence, Beijing Normal University, Beijing, China.

^f Department of Oncology, The Second Affiliated Hospital of Medical College, Xi'an Jiaotong University, Shaanxi, China

^g Bruker Scientific Technology Co., Ltd., Beijing, China

[#] Co-first authors

* Corresponding authors,

Jiyang Dong, jydong@xmu.edu.cn

Chao Zhao, chao.zhao@siat.ac.cn

Zongwei Cai, zwcai@hkbu.edu.hk

Abstract

There is growing awareness that metabolic heterogeneity of organism provides vital insight into the disease with molecular mechanism and personalized therapy. The screening of metabolism-related sub-regions that affect disease development is essential for the more focused exploration how disease progress aberrant phenotypes, even carcinogenesis and metastasis. Mass spectrometry imaging (MSI) technique has distinct advantages to reveal the heterogeneity of organism based on the *in situ* molecular profiles. The challenge of heterogeneous analysis has been to perform an objective identification among biological tissues with different characteristics. By introducing the divide-and-conquer strategy to architecture design and application, we establish here a flexible unsupervised deep learning model, called divide-and-conquer (dc)-DeepMSI, for metabolic heterogeneity analysis from MSI data without prior knowledge of histology. dc-DeepMSI can be used to identify either spatially contiguous region-of-interest (ROIs) or spatially sporadic ROIs. We demonstrate that the novel learning strategy successfully obtain sub-regions that are statistically linked to invasion status and molecular phenotypes of breast cancer, as well as organizing principles during developmental phase.

Introduction

Mass spectrometry imaging (MSI) could provide a plethora of metabolic information directly from biological specimens, including spatial distribution, abundance and composition of thousands of biomolecules^{1,2}. Identification from MSI data the region-of-interest (ROIs), which are statistically linked sub-regions or biologically functional regions, is usually used to differentiate cell types from heterogeneous tissue and in turn to contribute to our understanding of the cellular specificity of tissue^{3,4}, and allows better targeting the lesions and distant metastases that are associated with disease diagnosis and prognosis^{5,6}. In particular, ROIs analysis has become a critical foundation, allowing for subsequent detection of known biomarkers and discovery of unknown biomarkers with a major focus in tumor research⁷. Nevertheless, the fundamental question of how to improve accuracy and specificity of ROIs analysis is not crystal clear.

Segmentation is the common method for ROIs analysis in MSI data, which is accomplished by clustering data points (MSI image pixels) with similar characteristics into a cluster (*i.e.*, ROI). An effective segmentation result means that each cluster could link to a sub-region or a molecular phenotype, and the difference between clusters on MSI data can be used to interpret the biological heterogeneity on the tissue⁸.

MSI segmentation by far is a challenging task because of the complexities of MSI data in high dimensionality, low signal-to-noise ratio, and lack of benchmark datasets⁹. The existing methods for MSI segmentation can be roughly divided into supervised and unsupervised depending on whether prior knowledges of ROI label are used. In

69 supervised methods, data from histopathology, pathology or other imaging modals like
70 MRI are often evident to the ROI labels of MSI pixels, then guiding the segmentation
71 of MSI data^{10,11}. However, MSI data is of much more metabolic information which can
72 shape some “hidden” sub-regions that might not be distinguished by histological or
73 other imaging techniques, therefore, segmentation results will be biased if supervised
74 by histological data or other imaging modals. Nevertheless, some MSI studies are lack
75 of prior knowledges of ROI labels because of the extremely precious human tissue
76 specimens, which makes the supervised segmentation unpractical. On the contrary,
77 unsupervised segmentation is exploratory approach in which no prior information is
78 needed for pixels clustering, so the unsupervised segmentation is more practical and
79 gains more extensive attention than supervised one in MSI segmentation.

80 Dozens of unsupervised methods have been proposed for MSI segmentation in the
81 past decades¹²⁻¹⁴. For example, Abdelmola *et al.* use t-distributed stochastic neighbor
82 embedding (*t*-SNE) to reduce the dimensionality of MSI data, then uses *k*-means to
83 segment MSI data into a certain number of clusters that are expected to be in
84 coincidence with the prognostic tumor subpopulations¹²; The widely used vendor
85 software¹³, SCiLS Lab uses *k*-means to conduct MSI segmentation on some selected
86 ions, rather than on extracted features; Cardinal package provides a new unsupervised
87 clustering algorithm, namely spatial shrunken centroids, to produce a smooth MSI
88 segmentation¹⁴; and so on. To our best knowledge, most of the existing unsupervised
89 methods apply statistical model-based clustering algorithms like *k*-means to identify
90 ROIs from MSI data. Since model-based clustering algorithms usually rely on a certain

mathematical hypothesis of ROI¹⁵, for example, *k*-means assume that data points from a same cluster are high-dimensional spherical distribution around the ROI center¹⁶. Model-based clustering algorithm would fail to identify the ROIs that are unsatisfied with its underlying hypothesis. However, as we known, MSI dataset are of highly heterogenous, that is, data points from different sub-regions might distribute inhomogeneously across the MSI dataset. Thus, different sub-regions are of specific discriminate validities under a certain model-based clustering algorithm, making the segmentation results be poor-determined. It is urgent to develop a flexible clustering algorithm which is adaptive to the high heterogeneity of MSI data.

Deep learning is flourishing in recent years and achieved great success in various fields especially for biomedical image analysis. Deep learning features in data-driven strategy and the ability of learning automatically the local structure from the data¹⁷, which allows us to develop a flexible and adaptive clustering algorithm for MSI segmentation. Although deep learning-based methods have been proposed for some contexts of MSI data analysis like classification^{18,19}, the deep learning-based unsupervised segmentation for MSI segmentation is rarely reported because of the high-dimensionality of MSI data and the sensitivity of unsupervised deep learning methods in parameters initialization.

Here, we propose a flexible deep learning-based method called divide-and-conquer (dc)-DeepMSI for segmentation of MSI data by introducing the dc strategy into model designation, training and application. The task of MSI segmentation is divided into two separated sub-tasks, namely dimensionality reduction and feature clustering, then two

independent modules are designed and trained to conquer the two sub-tasks accordingly. In particular, a convolutional neural network (CNN) based deep learning architecture is designed to meet with the high heterogeneity of MSI data, and to achieve a flexible unsupervised MSI segmentation. In addition, to achieve a more accuracy segmentation, dc-DeepMSI provides with two specific modes, namely SPAT-spec and spat-SPEC, for typical ROIs including spatially contiguous ROIs and spatially sporadic ROIs. We illustrate the feasibility of dc-DeepMSI in two typical applications, experimental results show that dc-DeepMSI successfully identify eleven-different organs from a whole-body mouse fetus MSI image, and effectively explore the metabolic heterogeneity from a human breast tumor MSI image. Biomarker screening are performed on the ROIs identified by dc-DeepMSI from the tumor tissue, which further demonstrate that dc-DeepMSI can be used to detect the ROIs connected with clinical diagnosis, and thereby help to illuminate the metabolism associated diseases.

Results

1. dc-DeepMSI: A Divide-and-conquer Strategy Based Model to Segment MSI Data.

By introducing the divide-and-conquer strategy into deep neural network, a deep learning model named dc-DeepMSI is proposed here for unsupervised segmentation of high-dimensional MSI data, in which the task is divided into two independent sub-tasks, dimensionality reduction and feature clustering. Two separate modules are designed and trained accordingly in dc-DeepMSI to meet with the two sub-tasks, as shown in **Fig. 1a**. The dimensionality reduction module (the upper panel of **Fig. 1a**) is

implemented by an autoencoder with the intention of preserving the information and suppressing the noise as much as possible²⁰. The features clustering module (the lower panel of **Fig. 1a**) is designed as two competitive-cooperative CNN and their temporally ensemble copies. Two CNNs are structurally identical with independent parameters initialization, the output of one ensemble CNN feeds into the other CNN network, and vice versa, with intent to reduce the randomness and to achieve stable feature clustering. More architectural details of dc-DeepMSI including loss function, activation function and implementation are presented in the “**Methods**” section.

The CNN networks here play the roles of feature extraction (FE) and argmax classification, where FE block is accomplished by components following by a classifier (**Fig. 1b**). By setting two hyper-parameters, *i.e.*, the convolutional kernel size s and the weight of total variation (TV) loss ω_3 , dc-DeepMSI can switch its working modes between the general mode of SPAT-spec and the specific mode of spat-SPEC to meet with different ROI scenarios in a variety of specimens. Based on the extracted features of hyperspectral data, the general mode SPAT-spec clusters data points (MSI pixels) by both of their spatial closeness and spectral similarity, which is designed for identification of the most common ROIs in which MSI pixels are spatially contiguous across the MSI dataset. The specific mode spat-SPEC, by setting $s = 1$ and $\omega_3 = 0$, clusters MSI pixels by their spectral similarities, which is designed for ROIs identification in which MSI pixels are spatially sporadic across the MSI dataset. Nevertheless, the two modes are not antagonistic. Spatially sporadic ROIs can also be successfully identified by dc-DeepMSI of SPAT-spec mode with a small ω_3 , in which

157 the additional loss item of spatial closeness plays a role of denoising.

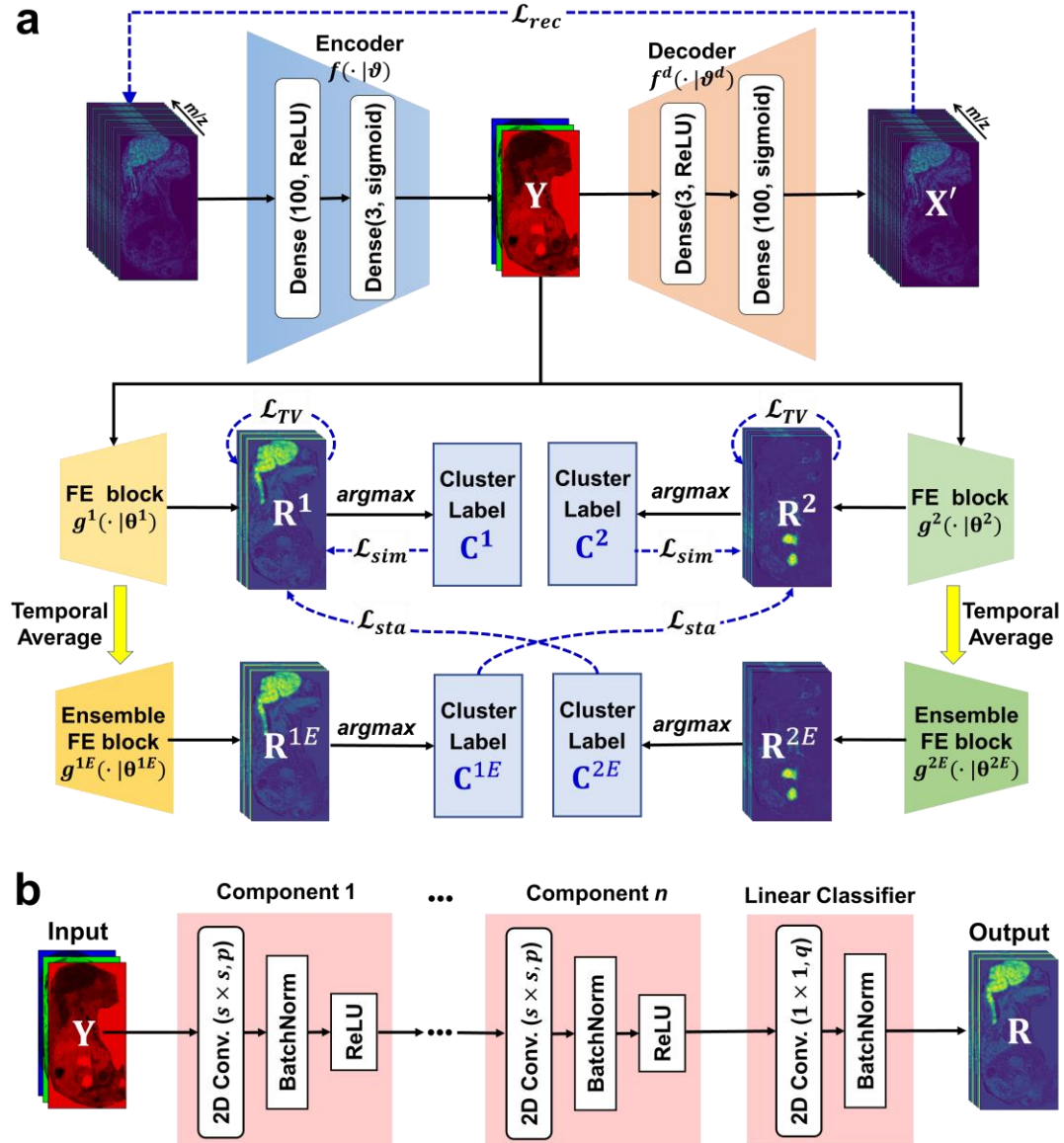


Fig. 1 | Schematic overview of dc-DeepMSI. a, Architecture of dc-DeepMSI. The upper half part is dimensionality reduction module which reduces a high-dimensional MSI data \mathbf{X} to a low-dimensional feature map \mathbf{Y} . The dimension reduction module is implemented by an autoencoder which consists of two fully connection layers in both encoder and decoder blocks. The lower half part is feature clustering module which is consisted of two CNN networks and two ensemble CNN networks. Each CNN network consists of a feature extraction (FE) block and an argmax classification. The cluster label from one ensemble CNN network is feed into its counterpart CNN network by loss function \mathcal{L}_{sta} to stabilize the segmentation result. When dc-DeepMSI reaches

convergence, the four CNN networks will also converge to a similar cluster label. **b**, Architecture of FE block. A FE block consists of n CNN components and a linear classifier, in which a CNN component consists of a 2D convolutional layer with $s \times s$ kernel size and p filters, a batch normalization layer and a ReLU activation function. The n CNN components are used to carry out a deeper feature extraction from the dimension reduced data \mathbf{Y} to an p -dimensional feature map. And the linear classifier, which is consisted of a 2D convolutional layer with 1×1 kernel size and q filters and a batch normalization layer, is used to map and normalize a p -dimensional feature map to a q -dimensional response map \mathbf{R} .

2. dc-DeepMSI Identifies Sub-organs of Mouse Fetus

The ROIs, in which data points contiguously distribute across the dataset, is a common scenario in MSI data of various biological tissues. Among them, MSI images of whole-body mouse fetus are a typical example with such spatially contiguous ROIs. Most noteworthy, molecular features and organs identification of mouse fetus are considered to be the complex and critical preprocessing with applications in areas such as embryological genetics, pathology and pharmacology²¹. Due to limitations in terms of technology, we have not been able to profile the multi-organ structures of mouse fetus from MSI images. To address this issue, we construct a dc-DeepMSI model on the MSI data of fetus mouse (embryonic day18) to identify organs and their sub-organs, in which the general mode SPAT-spec is adopted in view of the spatial continuity of MSI pixels from a same ROI. A total of 11 organs are identified by dc-DeepMSI including brain, orbital cavity, genioglossus muscle, submaxillary gland, sternebra, thymus, deposits of brown fat, heart, adrenal gland, kidney and intestine (**Fig. 2a**). More importantly, functional sub-organ structures are recognized from whole brain organ,

such as, dorsal pallium (isocortex) and hippocampal formation (Hpf) region, midbrain, brainstem and cerebellum (**Fig. 2a**).

To illustrate the performance of dc-DeepMSI on segmentation of MSI data, three commonly used methods are carried out on the MSI dataset of fetus mouse for comparison, including *t*-SNE+*k*-means which is implemented by the Python library Scikit-learn¹², a pipeline provided by commercial SCiLS Lab software²², and a pipeline implemented by Cardinal package²³. Segmentation map of *t*-SNE+*k*-means method shows abundant isolated clusters and obscure boundaries on mouse fetus, especially on fetal brain, which might due to the lack of spatial denoising procedure in the method (**Fig. 2b**). SCiLS Lab software tends to segment some big-size organs, such as brain and thoracic cavity, while fails to identify the sub-organs (**Fig. 2c**). Segmentation result of Cardinal package is a little bit better than that of SCiLS Lab software, but still miss some sub-organs, such as the brain of mouse fetus (**Fig. 2d**). The failure of sub-organs identification might due to the adoption of feature selection instead of feature extraction in dimension reduction procedure in SCiLS Lab and Cardinal package, which may result in severe information loss. These results demonstrate that dc-DeepMSI outperforms the other three methods in more and accuracy organ/sub-organ analysis.

Robustness and stability are two pivotal indicators for deep learning-based method²⁴. Here we illustrate the robustness of dc-DeepMSI on anti-noise in organ segmentation of mouse fetus. Poisson noise is generated and added on the MSI data, then dc-DeepMSI and the other three methods are carried out on the noisy data, respectively. dc-DeepMSI shows that its robustness against the noise by identifying

accurately most of the organs and sub-organs from the noisy fetus data, for example, brain and its sub-organs (**Fig. 2e**). *t*-SNE+*k*-means method whereas delivers too many isolated clusters on the segmentation map (**Fig. 2f**). More experiments on anti-noise ability evaluation by using *k*-means clustering, spectral clustering and Gaussian mixture model (GMM) clustering are detailed on **Supplementary Note 1** and **Supplementary Fig. 1**.

Sensitive to parameters initialization is another nuisance in most of deep learning-based methods²⁵. Here we design a comparative experiment to illustrate the model stability of dc-DeepMSI. Two different deep learning models are constructed. One is an end-to-end architecture model without explicit dimension reduction module (**Supplementary Fig. 2a**). The other is a deep model proposed in Kim's work²⁶, in which dimension reduction module is explicitly designed, but feature clustering module is implemented by a single-CNN structure (**Supplementary Fig. 2b**), which differs from dc-DeepMSI model. Twenty times of training with different parameters initialization are carried out, and the adjusted rand index (ARI) values are calculated (**Fig. 2g**). The end-to-end model has a small ARI mean = 0.68 and a large ARI standard deviation (std = 0.020), which implies the high sensitivity of parameters initialization. Kim's model improves its stability by dimension reduction module architecture (ARI mean = 0.74, std = 0.021). While dc-DeepMSI is of the best model stability (ARI mean = 0.78, std = 0.017) because of the divide-and-conquer strategy and double-CNN structures. More detailed evaluation results can refer to the **Supplementary Note 2** and **Supplementary Table. 1**.

Above all, dc-DeepMSI is more suitable for segmentation on high-heterogeneity and high-dimensional data analysis with outstanding robustness and stability.

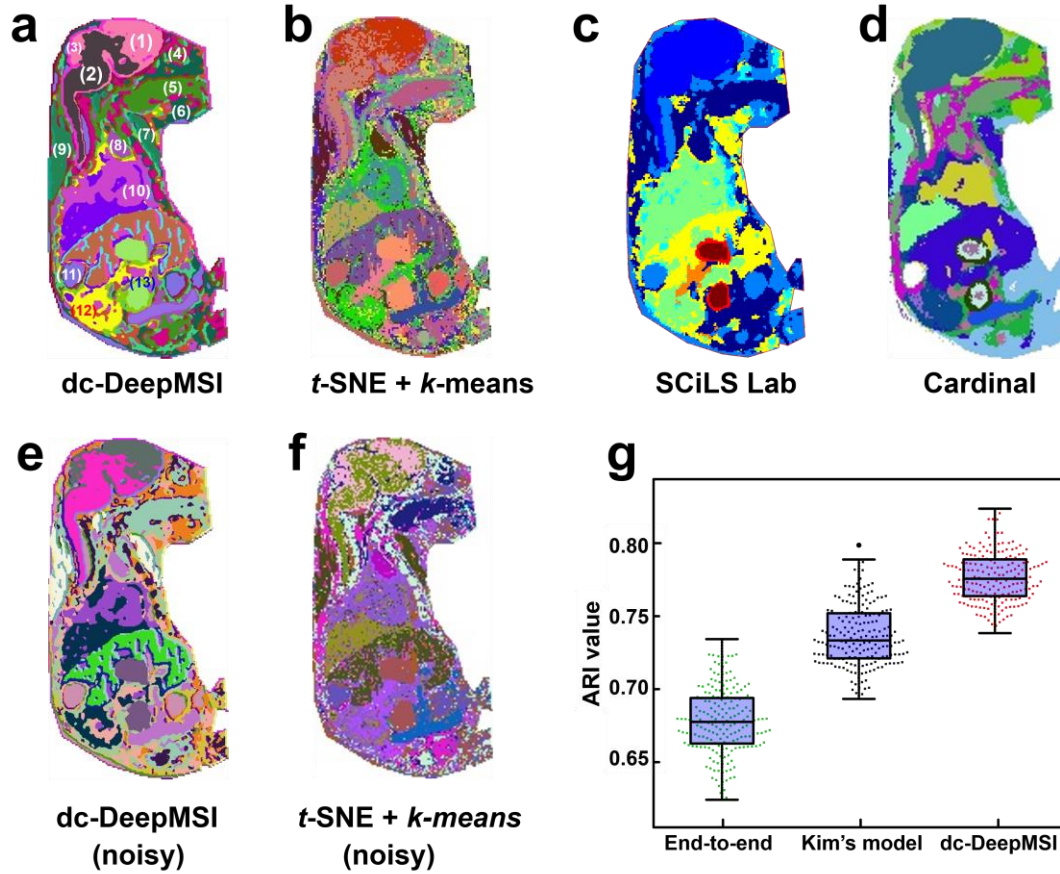


Fig. 2 | Identification of sub-organs of mouse fetus. a-d, Color encoded segmentation maps obtained from dc-DeepMSI, *t*-SNE+*k*-means, SCiLS Lab, Cardinal on original MSI data. Compared with the other three methods, dc-DeepMSI shows a smoothing clustering result as well as a better resolution of sub-organs. e-f, Segmentation maps obtained from dc-DeepMSI and *t*-SNE+*k*-means on a noisy MSI data. g, Comparison ARI values of an end-to-end model, Kim's model and dc-DeepMSI model. The organs and sub-organs are as follows, (1) dorsal pallium (isocortex) and hippocampal formation (Hpf) regions, (2) midbrain and brainstem, (3) cerebellum, (4) orbital cavity, (5) genioglossus muscle, (6) submaxillary gland, (7) sternebra, (8) thymus, (9) deposits of brown fat, (10) heart, (11) adrenal gland, (12) kidney and (13) intestine.

3. dc-DeepMSI Explores Metabolic Heterogeneity of Human Breast Tumor

Being different from the organ identification depending on spatially contiguous ROIs, some MSI datasets suggest that cellular distribution is characterized by the sporadic arrangement, as well as diversity on morphology, such as, human tumors, biofilm and single cell imaging^{6,12,27}. To specify the dc-DeepMSI application on spatially sporadic ROIs detection, taking the human breast sample as an example, the specific mode of dc-DeepMSI is carried out on intratumor regions to explore tumor metabolic heterogeneity. Thus, another divide-and-conquer based strategy is leveraged by dc-DeepMSI on application, in which the MSI dataset of complex tumor sample is divided into cancerous and para-carcinoma regions using the general mode of dc-DeepMSI, then exploring of tumor metabolic heterogeneity is conquered using the specific mode of dc-DeepMSI.

3.1 Cancerous and para-carcinoma discriminating via the general mode

Cancerous cells from solid tumors, *e.g.*, the human breast sample shown in **Supplementary Fig. 3**, possess the pathological characteristics of spatial continuity²⁸. Accordingly, both cancerous ROIs and para-carcinoma ROIs in the MSI data of tumor sample are spatial contiguous. Specially, margins of sub-regions are supposed to be natural edges of sub-populations of tumor samples. In view of this situation, we construct a general mode SPAT-spec of dc-DeepMSI to separate ROIs of carcinoma from para-carcinoma. As expected, the MSI data is successfully segmented into two separate sub-regions with clearly boundary, namely cancerous (blue) and para-carcinoma (light gray) regions (**Fig. 3a**), which shows good consistency with the results

from morphological evaluation (**Supplementary Fig. 3**). Scatter plot shows that data points from cancerous region (colored points) and data points from para-carcinoma region (grey points) can be clearly separated from each other in the cubic embedding space (**Fig. 3m**), or say the feature space of dimension reduced MSI data, which implies that molecular features are significantly different from each other between cancerous and para-carcinoma sub-region, and demonstrates the accuracy and efficiency of dc-DeepMSI in cancerous sub-region detection.

3.2 Tumor intra-heterogeneity exploring via the specific mode

Human tumor has significantly intra-heterogeneity in molecular phenotypes, microenvironment and metabolic regulation²⁹. The morphological analysis of the invasive ductal carcinoma with neuroendocrine differentiation (NED) indicates that breast tumor displays significant intra-tumor heterogeneity that is featured in the sporadic distribution between cancerous regions with different degrees in differentiation and stromal regions. As show in **Supplementary Fig. 3**, at least two typical cancerous regions can be classified by using immunohistochemistry (IHC) analysis according to the chromogranin A expression, including the cancerous region with NED (18%) and cancerous region (15%), as well as respective typical invasive regions.

To explore the molecular phenotypes and microenvironments in tumor sample, we build a model of dc-DeepMSI with the specific mode spat-SPECT on the MSI data of human breast tumor sample. dc-DeepMSI cluster data points in the intact tumor sample into ten-different sub-regions (**Fig. 3b**), in which most of sub-regions are in agreement

with the results of the morphological information. For example, three sub-regions are assigned and associated to two major molecular phenotypes (**Fig. 3c, 3d, 3e**). Additionally, there are one invasive sub-region (**Fig. 3f**) and six stromal sub-regions (**Fig. 3g-3l**). Herein, invasive ductal carcinoma with NED-related segmentation from intact tumor sample is given in **Fig. 3c, 3d**, showing the discrete imaging patterns with obscure boundary between the nests of neoplastic cells, which is basically consistent with morphological results. We also achieve the accurate invasive ductal carcinoma-related segmentation (**Fig. 3e**). The results exhibit the clear boundary between the nests of neoplastic cells according to the morphological results. Continuously, typical invasive region (**Fig. 3f**) and stromal region (**Fig. 3j**) are segmented from the intact sample, demonstrating the sporadic infiltration of neoplastic cells in the fibrous stroma, as well as the randomness of spatial distribution of stromal region, respectively.

Scatter plot shows that data points from a same sub-region are gather together, while data points from different sub-regions are clearly separated from each other in the embedding space, which illustrates the distinct metabolic difference among the ten sub-regions (**Fig. 3m**). Scatter plots of the ten sub-regions are shown in **Supplementary Fig. 4**. Violin plots display the distribution of Euclidean distances between data points of each sub-region and data points of para-carcinoma in the embedding space (**Fig. 3n**). As we can expected, the data points of two molecular phenotypes-related regions are far away from each other in the cubic embedding space, while data points of the stromal sub-regions locate in between the molecular phenotypes and the para-carcinoma, which indicates that lipid profiles of stromal sub-regions are

more similar to para-carcinoma than the invasion and the two major molecular phenotypes-related regions. The results show the ability of dc-DeepMSI in exploring metabolic heterogeneities from MSI data of tumor sample.

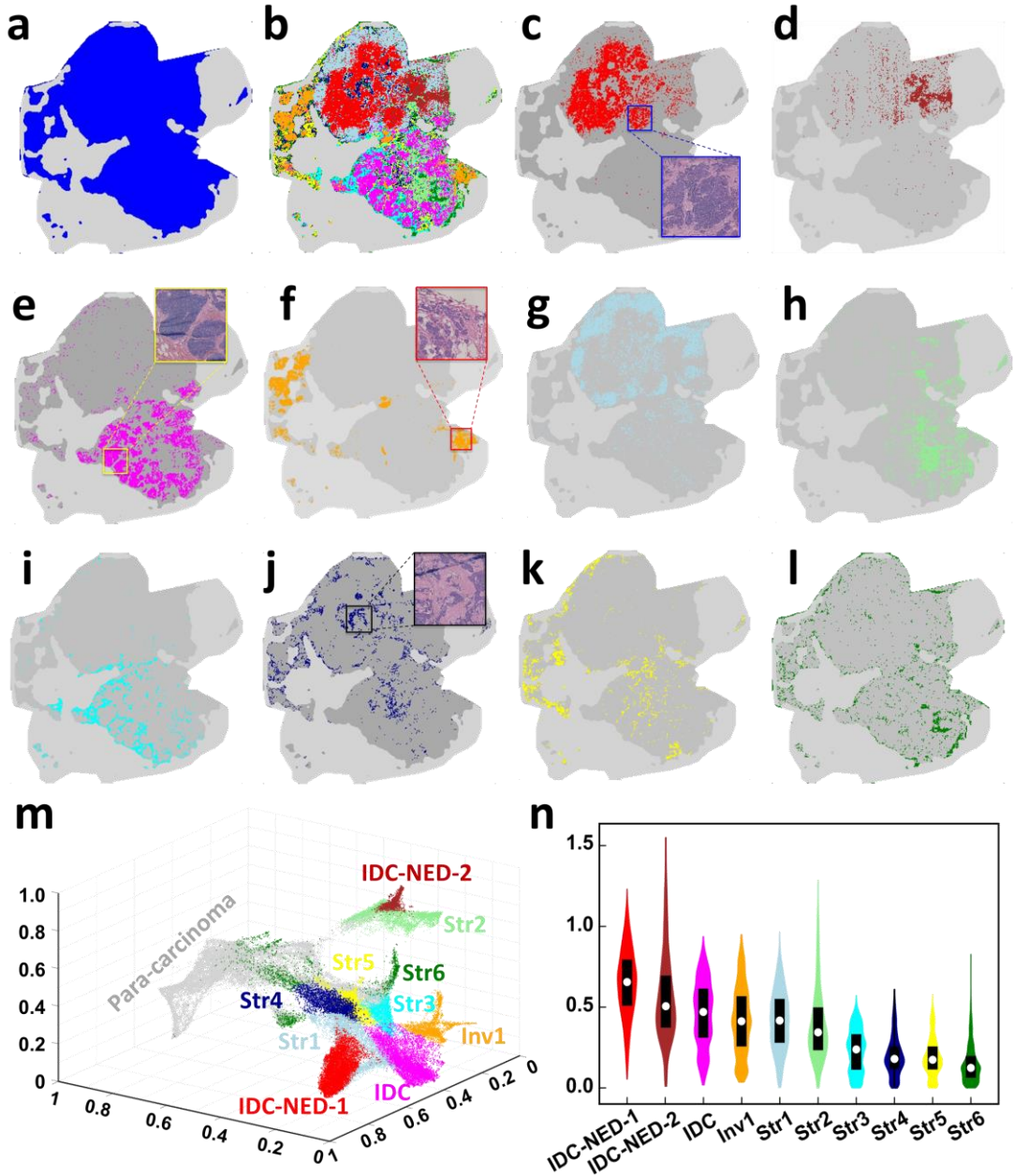


Fig. 3 | Results of dc-DeepMSI on human breast cancer data. **a**, Cancerous and para-carcinoma regions. **b**, Intact tumor sample. **c**, **d**, Invasive ductal carcinoma with NED-related sub-regions, called IDC-NED-1, IDC-NED-2, respectively. **e**, Invasive ductal carcinoma-related sub-regions, called IDC. **f**, Typical invasive region. **g-l**,

Stromal regions 1-6, called str 1-6. **m**, Scatter plot of data points in embedding space corresponding to **b. n**, Violin plot of Euclidean distances between data points of each sub-regions in cancerous and data points of para-carcinoma in the embedding space.

4. Screening of the underlying molecular markers

The underlying molecular markers from ROIs can help us to interpret and validate dc-DeepMSI segmentation results. As a traditional application scenario, MSI is capable of providing the spatial distribution of marker by an expression of single ion. Nevertheless, both multiple molecules and their interaction play an important role in complex biological regulations, making it difficult to use the expression of single ion to elaborate the spatial heterogeneity of bio-samples. To solve this problem, a two-stage screening approach is used here to identify the molecular markers between two given ROIs, namely target ROI and control ROI, as follows:

The first stage uses three univariate statistics to quantify the difference of abundance of a ion between target and control ROIs, that is, fold-change (FC), area under the receiver operating characteristic curve (AUC) and Hedges'g effect size (ES)³⁰. Then ions are defined as markers of the target ROI with respect to control ROI if they satisfy with the criteria as follows,

$$(ES \times AUC) \geq 1.5 \text{ and } |\log_2 FC| \geq 1.$$

If no marker is found in the first stage, we continue the second stage.

The second stage builds a linear regression model on the abundance matrix **X** and the ROIs belonging vector **y** vector³¹,

$$\mathbf{y} = \beta_0 + \beta\mathbf{X} + \varepsilon$$

Where $\beta = (\beta_1, \beta_2, \dots, \beta_i, \dots)$ is the regression coefficients for ions, and \mathcal{E} is the residuals errors. By imposing a least absolute shrinkage and selection operator (LASSO) penalty on the optimization of β , only a few ions are of non-zero coefficients, then these ions are defined as the co-expressive ions of the target ROI, which acts as a marker. More detailed definitions of FC, Hedges' g ES and LASSO optimization can refer to "Methods" section.

The existing evidences have suggested that abundances and spatial distribution of lipids are expressed abnormally in human breast tumor tissues, with a close relationship to aggressiveness and metastatic potentials of tumors. Tumor cells can generate excess lipids to maintain metabolic supplies and support tumor proliferation and invasion^{32,33}. Taking the human breast tumor for instance, the two-stage screening approach is carried out on each sub-region (target ROI) with respect to the other 9 sub-regions (control ROI) to identify the lipid markers or co-expressive lipid ions of the target ROI. Volcano plots and ion images are used to visualize the screening results (**Fig. 4**). And the lists including single- and multi- co-expressive lipid markers from the ROIs of breast tumor sample can refer to the **Supplementary Table 2**.

According to the results of two-stage screening approach, 4 lipid ions are found to be the lipid markers of carcinoma with respect to para-carcinoma regions. For example, the abundant m/z 743.65.73 PE (36:2), which is observed in cancerous regions corresponding on the haematoxylin and eosin (H&E) stain image, is absent in the para-carcinoma regions (**Fig. 4b**). We find that 8 lipids up-regulate in the specific sub-

regions containing invasive ductal carcinoma with NED, such as m/z 839.98 PC (40:3) (**Fig. 4d**) and m/z 795.89 PE (40:4) (**Fig. 4f**).

In the tumor sample at the invasive ductal carcinoma, invasive and stromal sub-regions, the results have shown a series of ions jointly contribute to shape their own molecular profiles. For example, co-expression of 17 ions in invasive ductal carcinoma-associated sub-regions (**Fig. 4h**), which is equivalent to a complex marker with $ES = 1.57$, $AUC = 0.87$ and $\log_2(FC) = -0.15$. Similarly, co-expression of 10 ions is accumulated in invasive sub-regions (**Fig. 4j**), which is equivalent to a complex marker with $ES = 3.49$, $AUC = 0.94$ and $\log_2(FC) = -0.26$. In addition, the one of stromal sub-regions is delineated by the co-expression of 30 ions (**Fig. 4l**), which is equivalent to a complex marker with $ES = 1.66$, $AUC = 0.83$ and $\log_2(FC) = -0.13$. A more detailed result is available in **Supplementary Fig. 5**.

We have demonstrated the significant heterogeneity of spatial distribution of lipid markers in the form of single-ion expression and multi-ions co-expression by the proposed two-stage screening approach, which has an important conductive function to metabolic reprogramming of tumor progression.

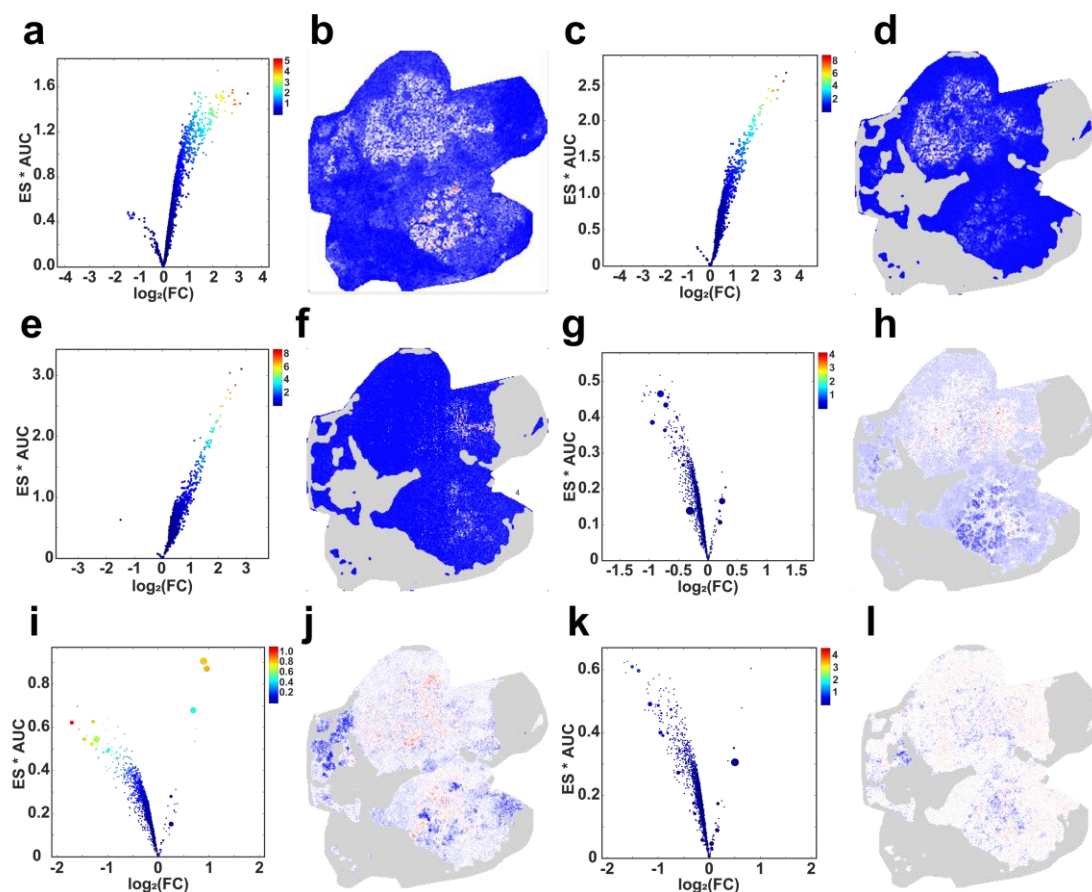


Fig. 4 | Molecular markers among sub-regions identified by dc-DeepMSI. The volcano plots show three measures including ES, AUC and $\log_2(\text{FC})$ between the target sub-region and the control region for all ions. Color encoded ion's images show the normalized abundances of selected markers or co-expressive ions. In volcano plots, the color represents the value of $\text{ES} * \text{AUC} * |\log_2(\text{FC})|$, and the warmer the color, the larger the value. The point size in volcano plots represents the absolute LASSO regression coefficient, the larger the size, the bigger the absolute coefficient. The target and control regions in **a,b** are cancerous and para-carcinoma regions, respectively. The target sub-region is **c, d**, IDC-NED-1. **e, f**, IDC-NED-2. **g, h**, IDC. **i, j**, Invasion. **k, l**, Stromal 5; respectively, and the control region is the remain sub-regions excluded the target sub-region.

Discussion

The screening and identification of metabolism-related sub-regions plays an

important role for better describing the molecular characteristics throughout the biological process and for optimizing the diagnosis and treatment of diseases³⁴. Previous studies reported the potential of MSI for the discovery of metabolic heterogeneity in tumor tissues³⁵. Actually, MSI dataset is appropriate for the metabolic heterogeneous analysis because of MSI provides us with: (1) very rich biological information from molecular level, usually achieve thousands of compounds simultaneously; (2) spatial resolved and (3) (relative) quantitative molecular information for *in situ* analysis of bio-samples. In this paper, for the first time, we introduce a divide-and-conquer strategy into deep neural network, and present a flexible dc-DeepMSI model to screen ROIs of spatially sporadic or spatially contiguous from MSI datasets of complex bio-samples, like human tumor or mouse fetus. dc-DeepMSI provides the possibility to characterize the molecular phenotypes and biomarkers in human tumors, and as well identifies sub-organs in mouse fetus based on spectral similarity and spatial closeness of targeted subpopulations.

The proposed model of dc-DeepMSI outperforms state-of-the-art MSI segmentation methods, which benefits from the following aspects: (1) The adoption of divide-and-conquer strategy greatly reduces complexity of a deep learning model, and as well improve the model stability. (2) The autoencoder based dimensionality reduction leads to a stable and low-dimensional representation of MSI data while minimizing information loss. (3) Feature clustering using two structurally identical but randomly initialized CNNs achieves a robust segmentation, in which the two CNNs work in an adversarial-and-collaborative way. Moreover, two temporally ensemble

CNNs stabilize effectively the segmentation. Several results have proven dc-DeepMSI is a straightforward and more robust approach to identify the presence of sub-regions characterized by similar mass spectrometry profiles, providing results that are not captured by histological technologies.

We provide in this paper a deep learning-based method to identify underlying metabolic heterogeneity from high-dimensional MSI data. Nevertheless, the proposed model is also expected to be broadly applicable in multiple computational tasks with hyperspectral imaging techniques, such as microscopy imaging, remote sensing imaging, and other medical imaging. We believe that our work will facilitate the extensive applications of unsupervised deep learning on high-dimensional data analysis.

Methods

Experimental datasets.

The procedures of animal experiments are approved by the Institutional Animal Care and Use Committee at Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences (Shenzhen, China). All of mice are treated humanely with the consideration of alleviating suffering. Six-week-old C57BL/6 male and female mice are housed under specific pathogen free condition with controlled temperature, humidity and 12 hrs dark: light cycle. One male and two females are bred and observed by a vaginal plug. And then, females are placed in a separate cage after successful mating. We collect the whole-body mouse fetus at embryonic day 18 for MALDI-MSI analysis.

Human tumor samples are collected from patients with breast cancer during the surgical tumor operation at the Second Affiliated Hospital of Medical College, Xi'an Jiaotong University. The patients are recruited with consent in this study and handled in accordance with approved procedures from the Institutional Review Board of the Second Affiliated Hospital of Medical College, Xi'an Jiaotong University and Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences.

The procedures of MALDI-MSI and histological analysis are described in the previous work of Zhao et al ^{22,36}. In short, the mouse fetus and human breast tumor samples are sectioned at a 14 μm -thickness by using CryoStar Nx79 cryostat (Thermo Fisher Scientific, Germany). Then, sections are thaw-mounted onto ITO slides for MALDI-MSI analysis. Subsequently, the serial sections are mounted on 4%

paraformaldehyde coated glass slides then used for H&E staining. MSI datasets are collected by using the RapifleX MALDI Tissue typer (Bruker Daltonics, Germany) with N-(1-Naphthyl)-ethylenediamine dihydrochloride matrix. H&E images are acquired using by Nanozoomer 2.0RS digital pathology scanner (Hamamatsu, Japan) with $0.4 \times$ amplification.

Data Preparation. The raw MSI data is collected using Bruker RapifleX MALDI Tissue typer. SCiLS Lab vendor software is used to read and export MSI data to *.imzML* files. MALDIquant package is then used to carry out data preprocessing including spectral alignment, peak detection, peak binning, *etc*³⁷. Finally, we obtain a data matrix $\mathbf{X}_{M \times N \times H}$, in which M, N are pixel numbers of horizontal and vertical coordinates of MSI image respectively, and H is the hyperspectral dimensionality, or say the ions (m/z) number.

The architecture of dc-DeepMSI. dc-DeepMSI is consisted of two modules, *i.e.*, dimensionality reduction (DR-module) and feature clustering (FC-module), as shown in **Fig. 1a**. DR-module is to learn a nonlinear mapping $f(\cdot | \boldsymbol{\theta})$ to project the high-dimensional data $\mathbf{X}_{M \times N \times H}$ into a low-dimensional data $\mathbf{Y}_{M \times N \times L}$ as follows,

$$\mathbf{Y}_{M \times N \times L} = f(\mathbf{X}_{M \times N \times H} | \boldsymbol{\theta}) \quad (1)$$

Where $\boldsymbol{\theta}$ is the network parameters in DR-module to be trained. FC-module is to learn a nonlinear mapping function $g(\cdot | \boldsymbol{\theta})$ from $\mathbf{Y}_{M \times N \times L}$ to segmentation map/cluster label $\mathbf{C}_{M \times N}$ as follows,

$$\mathbf{C}_{M \times N} = g(\mathbf{Y}_{M \times N \times L} | \boldsymbol{\theta}) \quad (2)$$

where $\boldsymbol{\theta}$ is the network parameters in FC-module to be trained.

To achieve the nonlinear mapping, FC-module is designed with two parallel feature extraction (FE) blocks and two temporally ensemble FE blocks, as shown in **Fig. 1a**.

Firstly, each FE block is implemented by a CNN of n components and a linear classifier (**Fig. 1b**), in which the CNN component is consisted of a 2D convolutional layer of p channels and $s \times s$ kernel size, a batch normalization layer and a ReLU activation function, while the linear classifier is consisted of a 2D convolutional layer of q filters and 1×1 kernel size. The output of FE block is a response map $\mathbf{R}_{M \times N \times q} = (r_{m,n,i})$, on which a segmentation map, or say cluster label $\mathbf{C}_{M \times N} = (C_{m,n})$, will be produced by applying *argmax* classifying,

$$C_{m,n} := \{i \mid r_{m,n,i} \geq r_{m,n,j}, \forall j \neq i \leq q\} \quad (3)$$

Secondly, the temporally ensemble FE block is accomplished by averaging the parameters of its corresponding FE block at each iteration t as follows,

$$\boldsymbol{\theta}^E(t) = \alpha \cdot \boldsymbol{\theta}^E(t-1) + (1 - \alpha) \cdot \boldsymbol{\theta}(t) \quad (4)$$

where $\boldsymbol{\theta}(t)$ and $\boldsymbol{\theta}^E(t)$ are the parameters of FE block and its corresponding temporally ensemble FE block at time t , and $0 \leq \alpha < 1$ is the ensemble momentum.

Specifically, the two FE blocks and two ensemble FE blocks map the input $\mathbf{Y}_{M \times N \times L}$ to 4-different segmentation maps as,

$$\begin{cases} \mathbf{C}_{M \times N}^1 = g^1(\mathbf{Y}_{M \times N \times L} | \boldsymbol{\theta}^1) \\ \mathbf{C}_{M \times N}^2 = g^2(\mathbf{Y}_{M \times N \times L} | \boldsymbol{\theta}^2) \\ \mathbf{C}_{M \times N}^{1E} = g^{1E}(\mathbf{Y}_{M \times N \times L} | \boldsymbol{\theta}^{1E}) \\ \mathbf{C}_{M \times N}^{2E} = g^{2E}(\mathbf{Y}_{M \times N \times L} | \boldsymbol{\theta}^{2E}) \end{cases} \quad (5)$$

The four FE blocks work adversarially and collaboratively to achieve a final segmentation map $\mathbf{C}_{M \times N}$.

Training strategy and implementation. Divide-and-conquer strategy is designed to

train the DR module and FC module, respectively. DR-module is implemented by an autoencoder framework²⁰, which is consisted of two blocks, i.e., the encoder block and the decoder block as follows,

$$\mathbf{Y}_{M \times N \times L} = f(\mathbf{X}_{M \times N \times H} | \boldsymbol{\vartheta})$$

$$\mathbf{X}'_{M \times N \times H} = f^d(\mathbf{Y}_{M \times N \times L} | \boldsymbol{\vartheta}^d)$$

where f and f^d are the mapping functions of encoder and decoder, $\boldsymbol{\vartheta}$ and $\boldsymbol{\vartheta}^d$ are the parameter of encoder and decoder blocks respectively. $\mathbf{Y}_{M \times N \times L}$ is the reduced data.

We use a loss function \mathcal{L}_{rec} to train the autoencoder module as follows,

$$\mathcal{L}_{rec} = \frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N 1 - \left(\frac{\mathbf{X}_{m,n} \cdot \mathbf{X}'_{m,n}}{\|\mathbf{X}_{M \times N}\|_2 \cdot \|\mathbf{X}'_{M \times N}\|_2} \right) \quad (6)$$

where $\|\cdot\|_2$ is l_2 -norm.

The loss function \mathcal{L} in FC-module is a weighted combination of three parts as,

$$\begin{aligned} \mathcal{L} = & \omega_1 \cdot (\mathcal{L}_{sim}(\mathbf{R}^1, \mathbf{C}^1) + \mathcal{L}_{sim}(\mathbf{R}^2, \mathbf{C}^2)) \\ & + \omega_2 \cdot (\mathcal{L}_{sta}(\mathbf{R}^1, \mathbf{C}^{2E}) + \mathcal{L}_{sta}(\mathbf{R}^2, \mathbf{C}^{1E})) \\ & + \omega_3 \cdot (\mathcal{L}_{TV}(\mathbf{R}^1) + \mathcal{L}_{TV}(\mathbf{R}^2)) \end{aligned} \quad (7)$$

where $\omega_1, \omega_2, \omega_3$ are combinational weights, \mathcal{L}_{sim} , \mathcal{L}_{sta} , and \mathcal{L}_{TV} are three loss functions to optimize the network parameters.

Firstly, the similarity loss of \mathcal{L}_{sim} is to make pixels with similar features be assigned to same cluster, which is designed based on cross entropy between the response map \mathbf{R} and segmentation map \mathbf{C} as follows,

$$\mathcal{L}_{sim}(\mathbf{R}, \mathbf{C}) = \frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N \sum_{i=1}^q -\delta(i - C_{m,n}) \cdot \ln r_{m,n,i} \quad (8)$$

where

$$\delta(t) = \begin{cases} 1, & \text{if } t = 0 \\ 0, & \text{Otherwise} \end{cases}$$

Secondly, the stability loss of \mathcal{L}_{sta} is to stabilize the segmentation result, which is calculated using to the response map of one FE model (\mathbf{R}) and the segmentation map of the temporally average of another FE block ($\mathbf{C}^{\sim E}$) as follows:

$$\mathcal{L}_{sta}(\mathbf{R}, \mathbf{C}^{\sim E}) = \frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N \max(0, r_{m,n}^{\text{neg}} - r_{m,n}^{\text{pos}} + \alpha) \quad (9)$$

where α is a margin parameter, and

$$r_{m,n}^{\text{neg}} := \left\{ r_{i,j} \mid \min \left(\|r_{m,n} - r_{i,j}\|_2^2 \right), \forall C_{i,j}^{\sim E} \neq C_{m,n}^{\sim E} \ \& \ m, n \neq i, j \right\}$$

$$r_{m,n}^{\text{pos}} := \left\{ r_{i,j} \mid \max \left(\|r_{m,n} - r_{i,j}\|_2^2 \right), \forall C_{i,j}^{\sim E} = C_{m,n}^{\sim E} \ \& \ m, n \neq i, j \right\}$$

Thirdly, the total variation (TV) loss of \mathcal{L}_{TV} is to make pixels of spatially close be in a same cluster, which is used to decrease the differences between neighboring pixels,

$$\mathcal{L}_{TV}(\mathbf{R}) = \frac{1}{M \times N} \sum_{m=1}^{M-1} \sum_{n=1}^{N-1} \|r_{m+1,n} - r_{m,n}\|_1 + \|r_{m,n+1} - r_{m,n}\|_1 \quad (10)$$

where $\|\cdot\|_1$ is l_1 -norm.

Stochastic gradient descent optimizer is adopted to train both DR-module and FC-module, where the learning rate and the momentum are set to be 0.01 and 0.9 respectively. Network parameters are initialized to be normal distribution $N(0,0.02)$. The proposed model is implemented in Python with PyTorch library and trained the models on a workstation equipped with a GPU Nvidia GTX 2080Ti graphics card.

Lipid ions screening method. Three commonly used metrics including Hedges' g effect size (ES), Fold-change (FC), area under the curve (AUC) are used to screen lipid markers for each sub-region in breast tumor sample,

Hedges' g is a measure of effect size (ES) that tells us how much one ROI differs from another, which can be calculated as,

$$\text{Hedges' } g = \frac{|\mu_1 - \mu_2|}{\sigma_{pooled}^*} \quad (11)$$

where μ_1 , μ_2 are the mean abundances of the target ROI and the control ROI respectively, and σ_{pooled}^* is the pooled and weighted standard deviation,

$$\sigma_{pooled}^* = \sqrt{\frac{(n_1 - 1)\sigma_1^2 + (n_2 - 1)\sigma_2^2}{n_1 + n_2 - 1}} \quad (12)$$

where σ_1 and σ_2 are the standard deviations of the target ROI and the control ROI respectively. The larger the effect size, the greater the difference between two ROIs.

Fold-change (FC) is used to evaluate the abundance difference between two given ROIs, which is calculated as follows:

$$\text{FC} = \frac{\mu_1}{\mu_2} \quad (13)$$

where μ_1 , μ_2 are the mean abundances of the target ROI and the control ROI respectively.

The area under the curve (**AUC**) is a measure of the ability of a classifier to distinguish between classes and is used as a summary of the receiver operating characteristic (ROC) curve. The higher the AUC, the better the performance of the model at distinguishing between the positive and negative classes. Here the positive and negative classes are the target and control ROIs respectively, and logistic regression is adopted to be the classifier model.

Furthermore, least absolute shrinkage and selection operator (LASSO) regression is used to identify co-expressive lipid ions for the target ROI with respect to the control

ROI³¹.

Let $\mathbf{X}_{N \times P}$ be the data matrix of two given ROIs with N data points (pixels) and P lipid ions in each pixel, \mathbf{y} be the ROI belonging vector of the N pixels. We can build a linear regression model on (\mathbf{X}, \mathbf{y}) as follows,

$$\mathbf{y} = \beta_0 + \beta \mathbf{X} + \boldsymbol{\varepsilon} \quad (14)$$

where $\beta = (\beta_1, \beta_2, \dots, \beta_P)$ is the regression coefficients, and $\boldsymbol{\varepsilon}$ is the residuals errors.

Impose LASSO penalty on the optimization of β , we have³¹

$$\hat{\beta}^{lasso} = \underset{\beta}{\operatorname{argmin}} \left\{ \frac{1}{2} \sum_{i=1}^N (y_i - \beta_0 - \sum_{j=1}^P x_{ij} \beta_j)^2 + \lambda \sum_{j=1}^P |\beta_j| \right\} \quad (15)$$

Then most of the regression coefficients will be zero.

The lipid ions of non-zero regression coefficients are defined as the co-expressive lipid ions of the target ROI, which acts as a lipid marker.

Data availability.

All of the datasets analyzed in this paper are public and can be referenced at <https://github.com/gankLei-X/dc-DeepMSI>.

Code availability

Source code is available at <https://github.com/gankLei-X/dc-DeepMSI>.

References

1. Kompauer, M., Heiles, S. & Spengler, B. Autofocusing MALDI mass spectrometry imaging of tissue sections and 3D chemical topography of nonflat surfaces. *Nat. Methods*. **14**, 1156-1158 (2017).
2. Doerr, A. Mass spectrometry imaging takes off. *Nat. Methods*. **15**, 32-32 (2018).
3. Rappez, L. *et al.* SpaceM reveals metabolic states of single cells. *Nat. Methods* **18**, 799-805 (2021).
4. Sun, C. L. *et al.* Spatially resolved metabolomics to discover tumor-associated metabolic alterations. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 52-57 (2019).
5. Aichler, M. & Walch, A. MALDI Imaging mass spectrometry: current frontiers and perspectives in pathology research and practice. *Lab. Invest.* **95**, 422-431 (2015).
6. Yuan, Z. *et al.* SEAM is a spatial single nuclear metabolomics method for dissecting tissue microenvironment. *Nat. Methods* **18**, 1223-1232 (2021).
7. Schwamborn, K. & Caprioli, R. M. Molecular imaging by mass spectrometry - looking beyond classical histology. *Nat. Rev. Cancer* **10**, 639-646 (2010).
8. Verbeeck, N., Caprioli, R. M. & van de Plas, R. Unsupervised machine learning for exploratory data analysis in imaging mass spectrometry. *Mass Spectrom. Rev.* **39**, 245-291 (2020).
9. Rafols, P. *et al.* Signal preprocessing, multivariate analysis and software tools for MA(LDI)-TOF mass spectrometry imaging for biological applications. *Mass Spectrom. Rev.* **37**, 281-306 (2018).

10. Hanselmann, M. *et al.* Active learning for convenient annotation and classification of secondary ion mass spectrometry images. *Anal. Chem.* **85**, 147-155 (2013).
11. Guo, D. *et al.* Deep multiple instance learning classifies subtissue locations in mass spectrometry images from tissue-level annotations. *Bioinformatics* **36**, 300-308 (2020).
12. Abdelmoula, W. M. *et al.* Data-driven identification of prognostic tumor subpopulations using spatially mapped t-SNE of mass spectrometry imaging data. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 12244-12249 (2016).
13. Mallah, K. *et al.* Lipid changes associated with traumatic brain injury revealed by 3D MALDI-MSI. *Anal. Chem.* **90**, 10568-10576 (2018).
14. Jones, M. A. *et al.* Discovering new lipidomic features using cell type specific fluorophore expression to provide spatial and biological specificity in a multimodal workflow with MALDI imaging mass spectrometry. *Anal. Chem.* **92**, 7079-7086 (2020).
15. Dexter, A., Race, A. M., Styles, I. B. & Bunch, J. Testing for multivariate normality in mass spectrometry imaging data: a robust statistical approach for clustering evaluation and the generation of synthetic mass spectrometry imaging data sets. *Anal. Chem.* **88**, 10893-10899 (2016).
16. Tarpey, T. & Flury, B. Self-consistency: A fundamental concept in statistics. *Statistical Science* **11**, 229-243 (1996).
17. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436-444 (2015).

18. Behrmann, J. *et al.* Deep learning for tumor classification in imaging mass spectrometry. *Bioinformatics* **34**, 1215-1223 (2018).
19. Manifold, B., Men, S. Q., Hu, R. Q. & Fu, D. A versatile deep learning architecture for classification and label-free prediction of hyperspectral images. *Nat. Mach. Intell.* **3**, 306-315 (2021).
20. Hinton, G. E. & Salakhutdinov, R. R. Reducing the dimensionality of data with neural networks. *Science* **313**, 504-507 (2006).
21. Liu, Y. *et al.* High-spatial-resolution multi-omics sequencing via deterministic barcoding in tissue. *Cell* **183**, 1665-1681 (2020).
22. Zhao, C. *et al.* MALDI-MS imaging reveals asymmetric spatial distribution of lipid metabolites from bisphenol S-induced nephrotoxicity. *Anal. Chem.* **90**, 3196-3204 (2018).
23. Bemis, K. D. *et al.* Probabilistic segmentation of mass spectrometry (MS) images helps select important ions and characterize confidence in the resulting segments. *Mol. Cell. Proteomics* **15**, 1761-1772 (2016).
24. Bai, X. *et al.* Explainable deep learning for efficient and robust pattern recognition: A survey of recent developments. *Pattern Recogn.* **120**, doi:10.1016/j.patcog.2021.108102 (2021).
25. Erhan, D., *et al.* Why does unsupervised pre-training help deep learning? *J. Mach. Learn. Res.* **11**, 625-660 (2010).
26. Kim, W., Kanezaki, A. & Tanaka, M. Unsupervised learning of image segmentation based on differentiable feature clustering. *IEEE T. Image Process.*

646 **29**, 8055-8068 (2020).

647 27. Spraker, J. E., Luu, G. T. & Sanchez, L. M. Imaging mass spectrometry for
648 natural products discovery: a review of ionization methods. *Nat. Prod. Rep.* **37**,
649 150-162 (2020).

650 28. Foulkes, W. D., Smith, I. E. & Reis, J. S. Triple-negative breast cancer. *N. Engl.*
651 *J. Med.* **363**, 1938-1948 (2010).

652 29. Martinez-Outschoorn, U. E., Peiris-Pages, M., Pestell, R. G., Sotgia, F. &
653 Lisanti, M. P. Cancer metabolism: a therapeutic perspective. *Nat. Rev. Clin.*
654 *Oncol.* **14**, 11-31 (2017).

655 30. Hedges, L. V. Distribution theory for Glass's estimator of effect size and related
656 estimators. *J. Educ. Behav. Stat.* **6**, 107-128 (1981).

657 31. Tibshirani, R. Regression shrinkage and selection via the Lasso. *J. R. Statist.*
658 *Soc. B.* **58**, 267-288 (1996).

659 32. Schulze, A. & Harris, A. L. How cancer metabolism is tuned for proliferation
660 and vulnerable to disruption. *Nature* **491**, 364-373 (2012).

661 33. Morad, S. A. F. & Cabot, M. C. Ceramide-orchestrated signalling in cancer cells.
662 *Nat. Rev. Cancer.* **13**, 51-65 (2013)

663 34. Baker, M. Metabolomics: from small molecules to big ideas. *Nat. Methods* **8**,
664 117-121 (2011).

665 35. Calligaris, D. *et al.* Application of desorption electrospray ionization mass
666 spectrometry imaging in breast cancer margin analysis. *Proc. Natl. Acad. Sci.*
667 *U. S. A.* **111**, 15184-15189 (2014).

36. Zhao, C. *et al.* Airborne fine particulate matter induces cognitive and emotional disorders in offspring mice exposed during pregnancy. *Sci. Bull.* **66**, 578-591 (2021).
37. Gibb, S. & Strimmer, K. MALDIquant: a versatile R package for the analysis of mass spectrometry data. *Bioinformatics* **28**, 2270-2271 (2012).

Acknowledgments

The work was supported by the National Natural Science Foundation of China (91843301, 81871445 and 22176195), the National Key Research Program of China (2017YFC1600505 and 2017YFE0191000), the Natural Science Foundation of Guangdong Province, China (2021A1515010171), and Key projects of basic research in Shenzhen (JCYJ20210324115811031).

Author contributions

Jiyang Dong, Lei Guo, Xiangnan Xu and Zhichao Wu performed the experiments of dc-DeepMSI. Yinbin Zhang and Chao Zhao provided the tumor analysis. Chao Zhao and Zongwei Cai provided the MSI data. Yongwei Wang and Pengfei Li provided consultancy and discussion in data analysis. Jiyang Dong and Chao Zhao co-wrote the manuscript and analyzed the data. Jiyang Dong, Chao Zhao and Zongwei Cai supervised and directed the study.

Additional information

Supplementary Information. The online version of this article contains Supplementary Information (Fig. 1-5, Table 1,2, and Note 1,2), which is available to authorized users.

690 **Conflict statement.** The authors declare there is no conflict of interest in this
691 manuscript.