**质谱成像与HE染色成像综合处理软件使用说明书**

1. **软件目的和用途**

HE染色，即苏木精—伊红染色法 ( hematoxylin-eosin staining )，石蜡切片技术里常用的染色法之一。苏木精染液为碱性，主要使细胞核内的染色质与胞质内的核酸着紫蓝色 ；伊红为酸性染料 ，主要使细胞质和细胞外基质中的成分着红色。HE染色法是组织学、胚胎学、病理学教学与科研中最基本、使用最广泛的技术方法。

对组织切片进行HE染色，可以得到该组织高分辨率的病理学信息。然而HE染色成像无法反映组织切片表面特定代谢物的含量分布。而质谱成像，作为一种非靶向的成像技术，可以对组织切片表面各种代谢物的分布进行成像分析。然而，其缺点是分辨率较低，并且成像结果难以直接得到组织的病理学信息。

因此，本软件的目的主要是通过对同一片组织进行HE染色成像和质谱成像得到的图像数据进行综合处理，从而发挥两种成像方法的优势，对于组织有着更好的了解。

具体来讲，该软件主要实现可以实现以下两个功能：

**功能1：** 通过HE染色图像将组织精确分成不同的病理学区域，映射到质谱成像图中，从而得到组织中各代谢物在各不同的病理学区域的含量分布。

**功能2：** 基于神经网络算法将组织切片质谱成像与HE染色成像进行融合，以达到增强组织切片质谱成像分辨率

1. **软件开发思想**

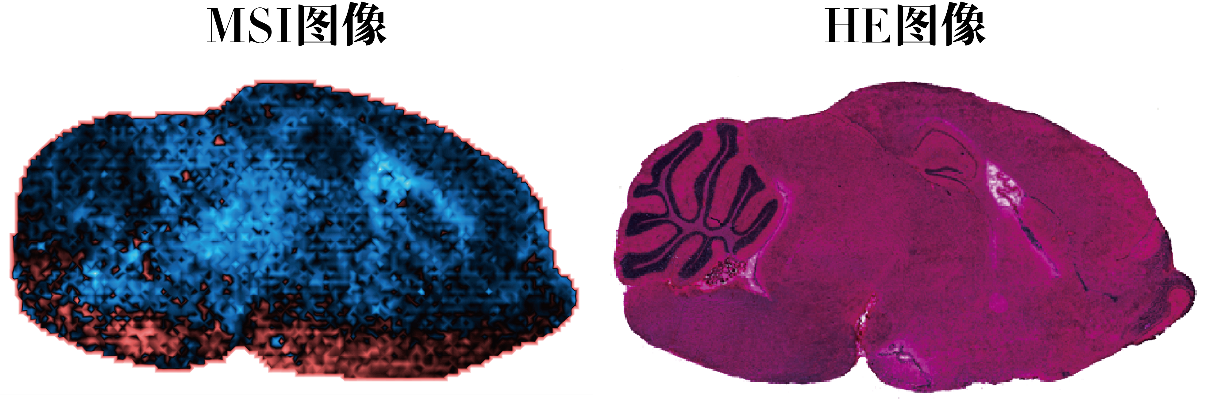


图1

如图1所示，左图是对组织切片进行质谱成像所得到的图像，这个图像可以是质谱成像文件直接导出生成的，也可以是通过各种处理方法生成的如镜像切片差值成像。右图所示为HE染色图像。

* 1. **功能1算法思想**

功能1主要通过HE染色图像将组织精确分成不同的病理学区域，映射到质谱成像图中，从而得到组织中各代谢物在各不同的病理学区域的含量分布。

图2展示了一张小鼠大脑病理学分区的简笔图。如图所示，小鼠大脑在病理学上有非常精细的分区。对每一个分区分别进行研究对于小鼠大脑的病理学研究有着十分重要的意义。

如图1所示，在小鼠大脑切片的HE染色图像中，各病理学分区有着十分明确的区分。相比之下，在质谱成像图中，各病理学分区之间的界限模糊，难以区分。因此，有必要利用HE图像来指导探究各代谢物在各病理学分区中的浓度分布。

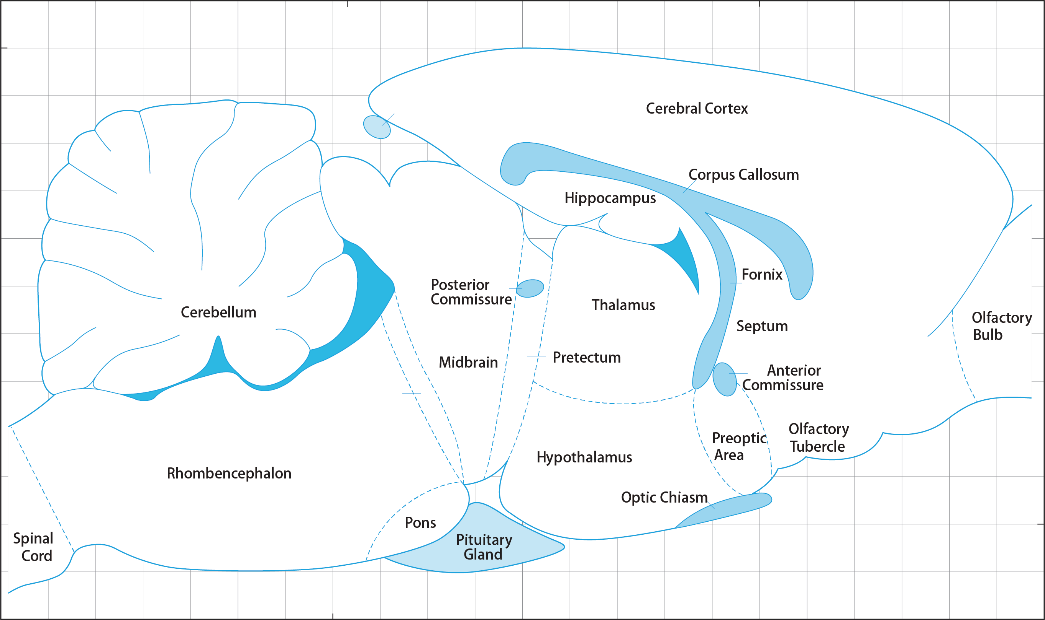


图2

为完成这一工作，首先第一步是将HE图像上的像素点与MSI图像的像素点进行一一对应。

那么如何进行对应？我们的设想是通过在HE图像上进行选区，选区通过一系列的可逆变换与MSI图像上的组织切片选区进行对应。这一变换过程被记录下来，从而达到HE图像上像素点与MSI图像的像素点进行对应的目的。

图3 简要地展示了这一过程。

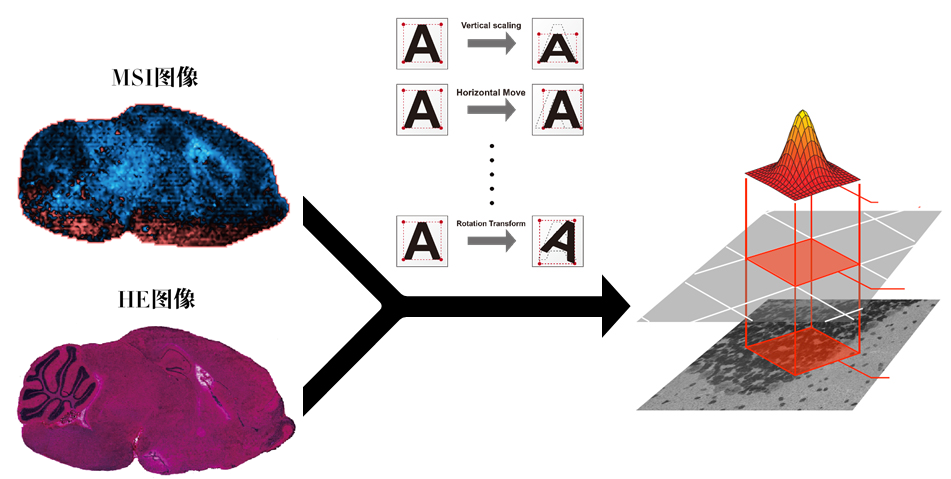


图3

在MSI图像与HE图像进行像素点对应完成之后，我们就可以得到一个HE图像各像素点与MSI图像个像素点对应的列表。

这时候，我们再选择HE图像中的一片病理学区域，就可以精确的映射到质谱成像图中。得到所选定病理学区域所包含的质谱成像像素点。提取这些像素点的质谱数据后，就可以得到各代谢物在该病理学区域的浓度水平数据。

**2.2 功能2算法思想**

质谱成像（MSI）图像和HE图像的获取方式完全不同，其分辨率也相差很远。如果要将MSI图像和HE染色图像相融合，首先第一步是将HE图像上的像素点与MSI图像的像素点进行一一对应。

像素点对应算法思想在功能1中已经介绍，这里不再赘述。

在MSI图像与HE图像进行像素点对应完成之后，我们就可以得到一个HE图像各像素点与MSI图像个像素点对应的列表。

在此，我们设定一些常量

|  |  |
| --- | --- |
| Abbr. | Exp. |
| ***n\_sample*** | The number of pixels extracted from HE image for neural network training |
| ***n\_feature*** | The number of features extracted from HE image |
| ***n\_metabolites*** | The number of metabolites used for training neural network |
| ***Input Value*** | The matrix used as the input value when training the neural network |
| ***Target Value*** | The matrix used as the output value when training the neural network |

我们希望以每个HE像素点作为1个***sample***，每个像素点上R,G,B值作为***feature***；而对应的MSI图像上的像素点所包含的代谢物强度作为作为***Target Value***；来生成神经网络训练数据集。

接下来，我们对HE图像做各种变换，以达到训练数据的多样性。如图4所示。

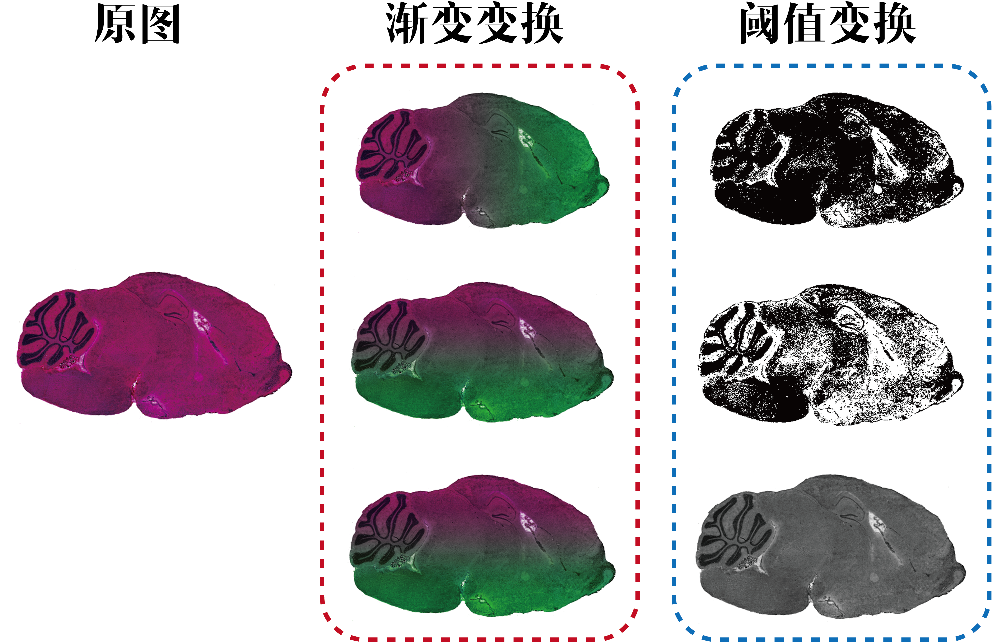


图4

接下来，我们就得到了一个基于HE染色图像的***n\_sample\*n\_feature***的数据矩阵作为***Input Value***； 得到 ***n\_sample\*n\_metabolites***的数据矩阵作为***Output Value***. 我们利用这些数据对于神经网络进行训练。

在对神经网络训练完成后，我们重新选取我们感兴趣的切片区域，同样提取选取区域中的每个像素点组成一个数据集。将该数据集输入到训练好的神经网络中，则可以得到预测出的每个像素点上各个代谢物的强度。

将神经网络预测的代谢物强度按照原来像素点的坐标重建，则可以得到MSI与HE染色图像融合后生成的图像。

1. **软件使用操作指南**

该软件基于Python 3.2编写。软件中所需的Python依赖库包括：PyQt5，Matplotlib，Numpy，Scipy，sklearn，xlrd，xlwt，pyimzml，os，PIL。

该软件需要配合**镜像切片组织质谱成像像素点对应软件**使用。

* 1. **获取质谱成像数据**

对照**镜像切片组织质谱成像像素点对应软件**使用说明书，进行质谱成像选区和对应选区数据导出操作。导出后生成一个Excel文件，如图5所示。

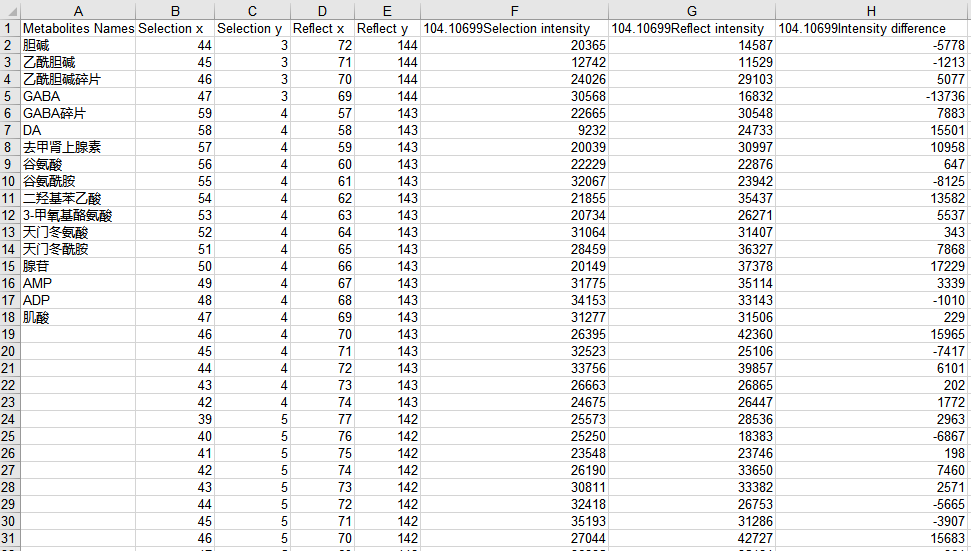
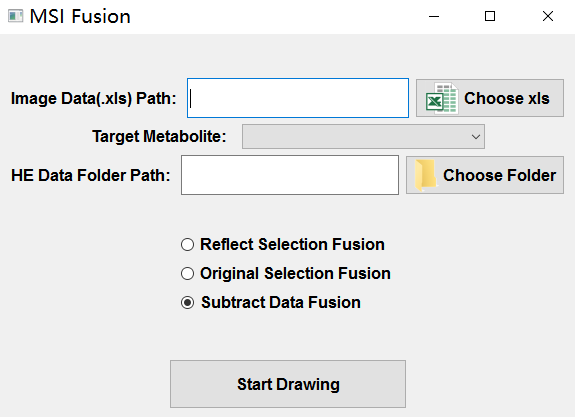


图5

* 1. **选取MSI数据和HE染色图像文件**



框3

框1

框2

图6

如图6所示，框1中选择3.1中生成的Excel文件；选择完后，下拉列表框中会出现目标代谢物名称，选择所需成像的目标代谢物。框2中选择包含HE染色图像的文件夹。那么在该文件夹中，支持多张HE染色图像。包括对于原始的HE染色图像进行各种变换后生成的图像。该文件夹中包含的图像均将作为神经网络训练的原始数据。图像命名格式如图7. （各个图像按照1,2,…,n进行标号，每张图像大小要完全一致，原始HE染色图像必须命名为1.png）

框3中有三个选项。根据需要选择对原始选区进行融合，或是对原始选区镜像选区进行融合，或是对选区差值进行融合。

最后，点击“Start Drawing”按钮开始成像。

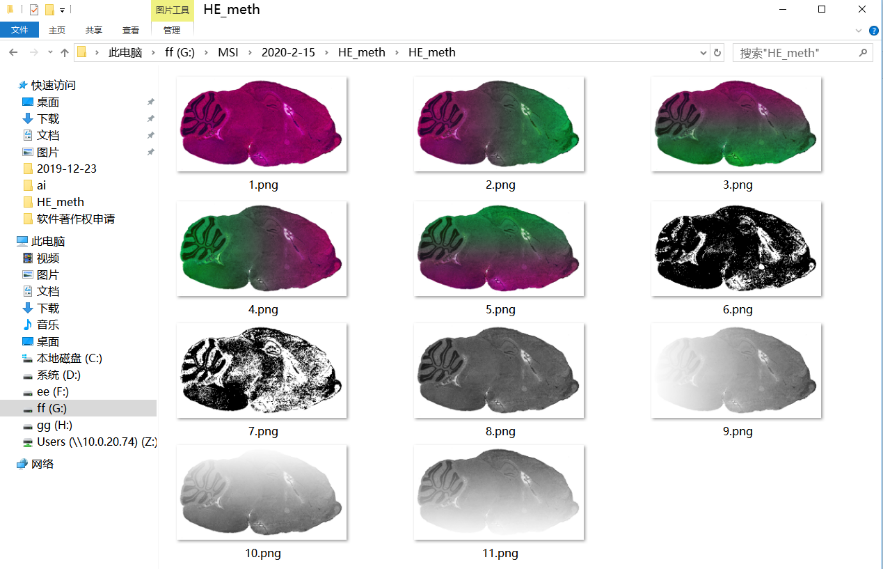
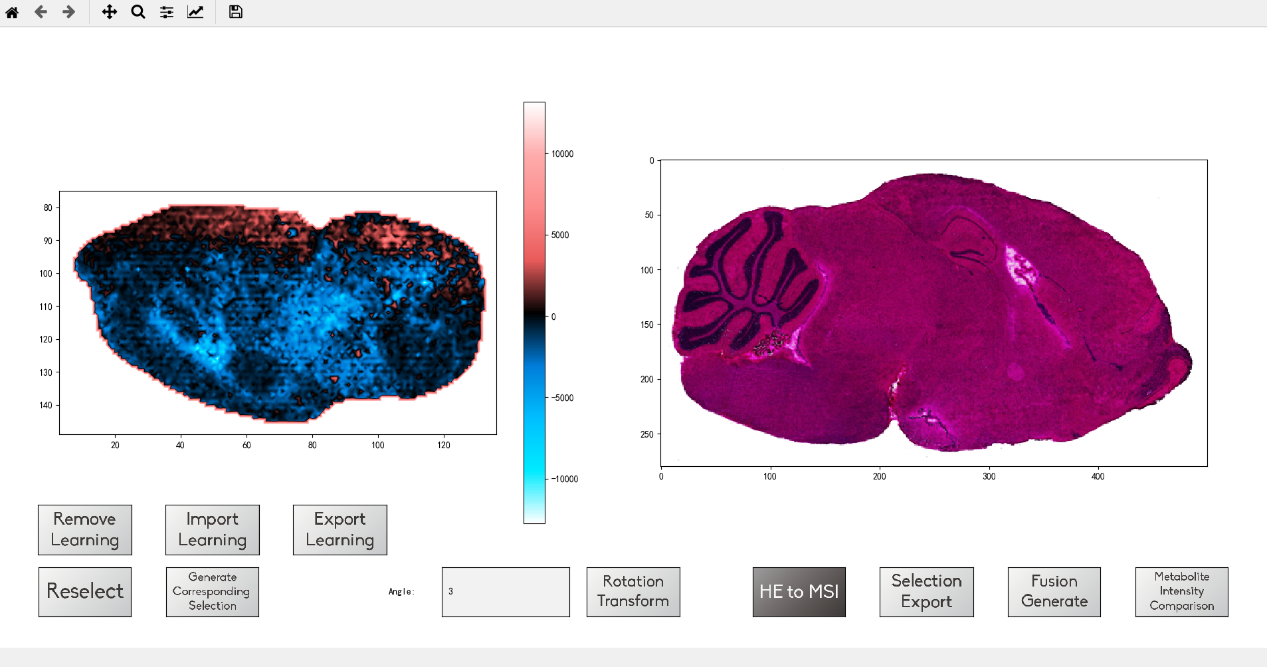


图7

* 1. **MSI与HE染色图像像素点对应界面介绍**

如图8所示为MSI与HE染色图像融合界面。该界面主要分为**MSI图像区,HE图像区和****像素点对应按钮区，功能2按钮区，功能1按钮区**。下面我将对各区的功能以及如何对选区进行基本变换进行介绍。

表一为界面中各个按钮的功能简介。



**MSI图像区**

**HE图像区**

**功能1按钮区**

**功能2按钮区**

**像素点对应按钮区**

图8

|  |  |  |
| --- | --- | --- |
| 名称 | 图样 | 功能 |
| Remove Learning按钮 |  | 移除学习轨迹， 点击后目前状态学习轨迹清除 |
| Angle输入框 |  | 角度输入，输入旋转角度（顺时针为正） |
| Rotation Transform按钮 |  | 旋转变换， 点击后图像区选取框将旋转Angle输入框角度 |
| Import Learning按钮 |  | 导入学习轨迹， 点击将从文件导入学习轨迹至当前状态 |
| Export Learning按钮 |  | 导出学习轨迹， 点击将当前状态学习轨迹导入至文件储存 |
| Generate Corresponding Selection按钮 |  | 生成对应选区，点击将当前图像区选取框按照学习轨迹生成对应选区框 |
| HE to MSI按钮 |  | 将在HE图像区中的轨迹映射到MSI图像区 |
| Selection Export按钮 |  | 将在HE图像区选择的选区内包含的像素点包含的feature以及MSI图像区对应的像素点包含的metabolites信息导出 |
| Fusion Generate按钮 |  | 利用“Selection Export按钮”导出的数据生成融合图像 |
| Metabolite Intensity Comparison按钮 |  | 对于MSI图像区中选区内包含的像素点含有的各代谢物强度进行对比（结合镜像切片组织质谱成像像素点对应软件使用） |

表1

* 1. **软件各功能和使用流程**

该软件工作流程如图9所示。

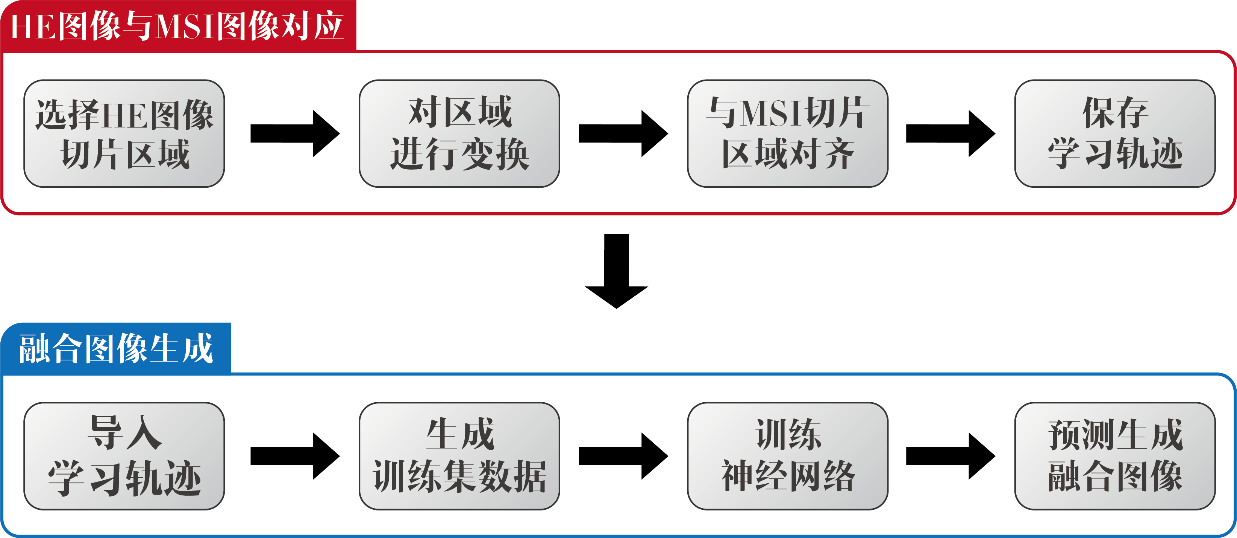


图9

接下来我简要介绍一下软件各功能的使用。

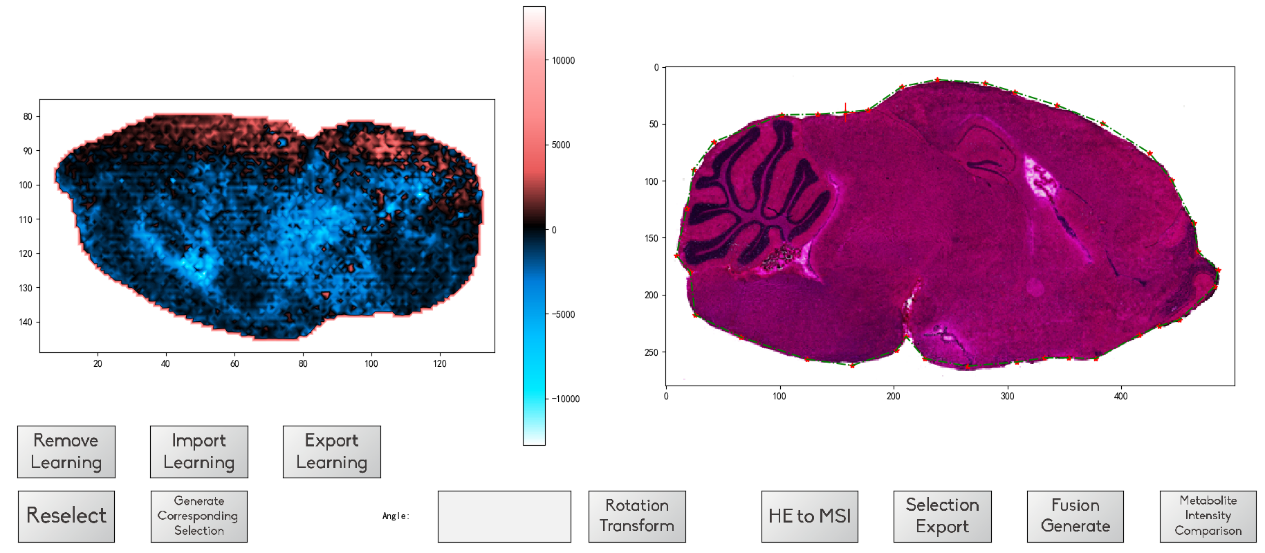
**3.4.1** **HE图像与MSI图像像素点对应**

在实验过程中，由于HE图像和MSI图像是通过两种完全不同的方式获取的，其切片所在位置，分辨率以及像素点个数都存在着巨大的差异。 因此，我们通过选取成像较为清晰的一个代谢物，通过对切片在HE图像和MSI图像中的位置进行人工对齐，并将该学习轨迹记录下来。那么在HE选区中的每个像素点经过该学习轨迹变换都可以映射到MSI图像中切片的对应像素点。对于一次成像，只需要进行一次轨迹学习过程。

那么在轨迹学习过程完毕后，只需要导入学习轨迹，在HE图像中选择一块区域，就可以在MSI图像中生成对应选区。那么HE图像选区中的每一个像素点都可以通过学习轨迹和MSI图像中的像素点进行一一对应。将每个像素点对应的feature导出即可生成训练集数据。

下面详细讲解一下软件的操作。

1. MSI图像区选择选区 在MSI图像区持续点击鼠标左键来创建新的选区控制点。如图10



**持续点击鼠标左键创建新的选区控制点**

**HE图像区**

图10

1. 修改选区范围 将鼠标移动至需要删除的选区控制点附近，点击鼠标右键，即可删除一个选区控制点。
2. 复制选区到MSI图像区

在HE图像区进行选区选取完成后，点击“HE to MSI”按钮就可以将选区复制到MSI图像区。如图11所示

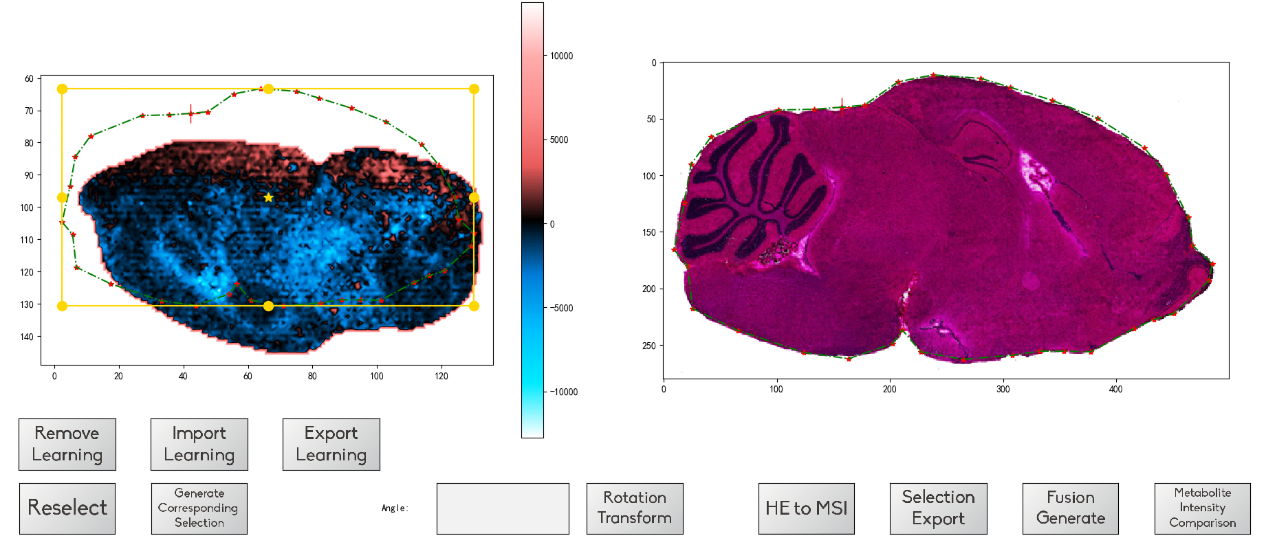


图11

1. MSI图像区选区变换

那么在选区选取完毕后就要对选区进行相应的变换，使得选区与切片对齐；那么这一变换过程软件会自动记录下来。

主要的变换方式包括平移变换，旋转变换和缩放变换。下面分别叙述其在软件中的操作方式。

1o 平移变换

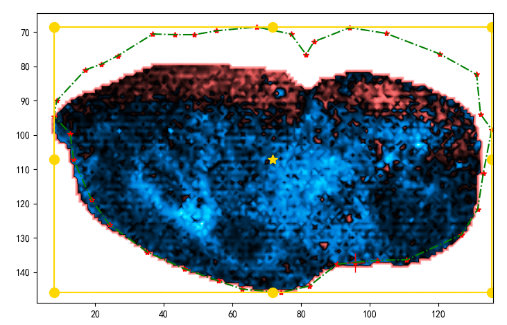
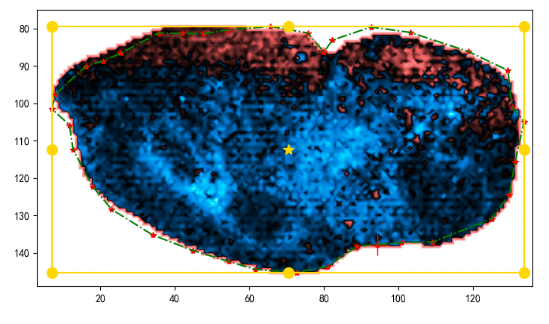
平移变换：将光标移动至MSI图像区（不要按下），按住“↑” “↓” “←” “→”键来进行平移；同时按住“alt”+“↑” “↓” “←” “→”键来进行微调。

 2o 旋转变换

在 中输入需要旋转的角度（顺时针为正），点击“Rotation Transform“按钮，则选区旋转相应的角度。

3o 缩放变换

在MSI图像区按下“Ctrl”+鼠标左键，在选区四周会出现黄色的缩放控制框。按住“Shift”键，将光标移动至选区控制框四个控制轴附近，按住鼠标左键拖动即可完成缩放操作。如图12所示。



按住“shift”健+按住鼠标左键拖动

图12

1. 保存学习轨迹

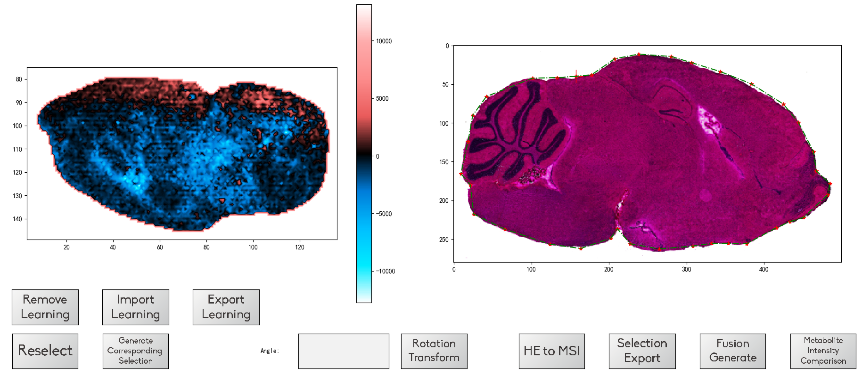
在轨迹学习完成后，点击“Export Learning”按钮，即可将学习的轨迹进行导出，导出的学习轨迹将储存在程序运行文件夹中的Learning Recording Data.xls中。

1. 导入学习轨迹

在学习轨迹保存后，点击“Import Learning”按钮，即可将保存的学习轨迹导入。

1. 生成对应选区

点击“Generate Corresponding Selection”按钮，即可按照学习轨迹，根据HE图像选区,在MSI图像中生成对应的选区。如图13



**点击“Generate Corresponding”按钮**

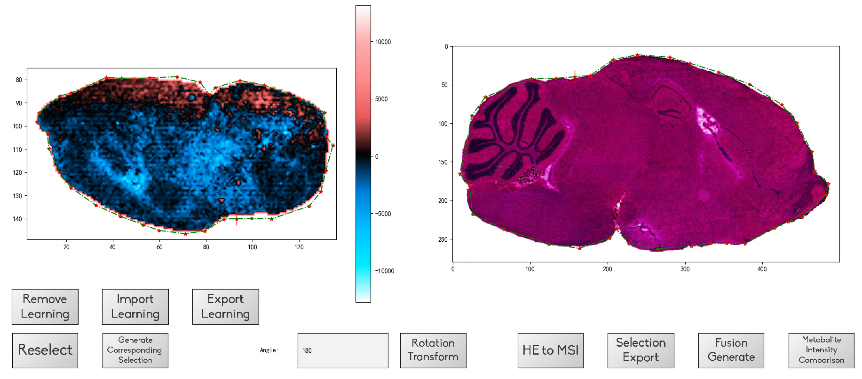


图13

**3.4.2 功能1操作指南**

1) 选择感兴趣的病理学区域

在HE图像与MSI图像像素点对应工作完成后，点击“Reselect”按钮，在HE图像区域重新选择感兴趣的病理学区域，如图14

**选择感兴趣的病理学区域**

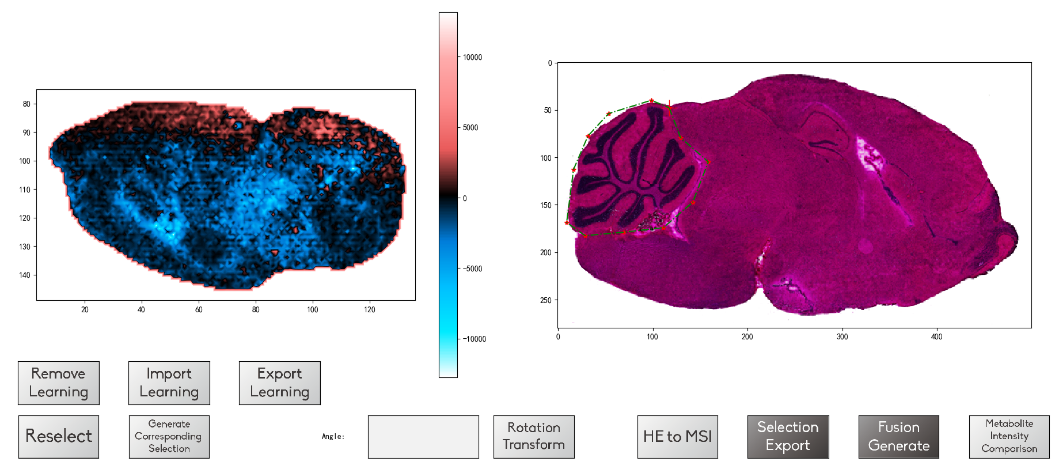


图14

2） 在质谱成像图中映射相应病理学区域

选择完感兴趣的病理学区域后，点击“Generate Corresponding Selection”按钮，在质谱成像图中映射相应病理学区域。如图15.

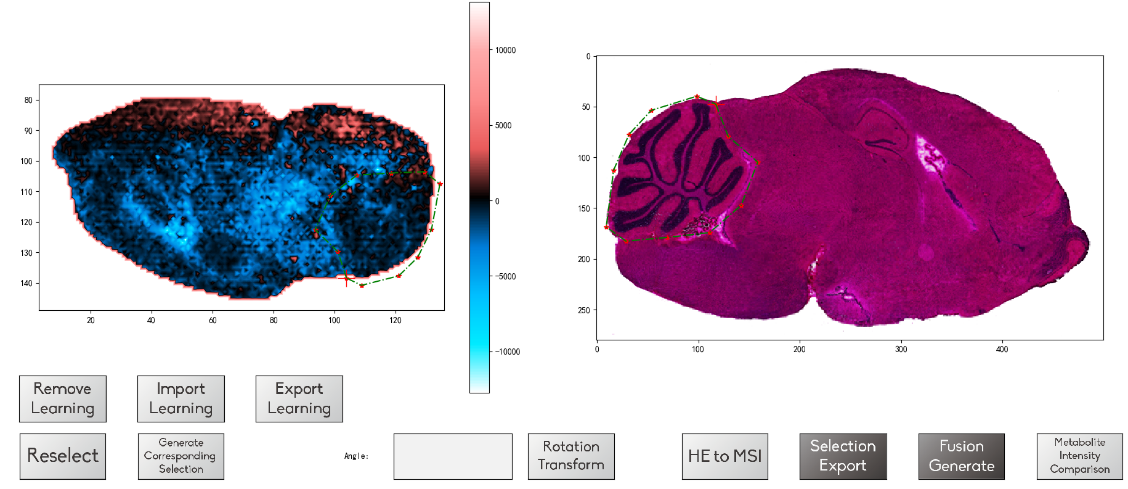


图15

3）生成各代谢物在选定的病理学区域的含量水平

最后点击“Metabolite Intensity Comparison”按钮，生成各代谢物在选定的病理学区域的含量水平。包括一张柱状图，在程序运行目录下会生成包含该区域所有质谱成像像素点数据的表格。如图16.（图16是根据一组镜像切片所生成柱状图。蓝色柱和黄色柱分别代表原选区和镜像选区）

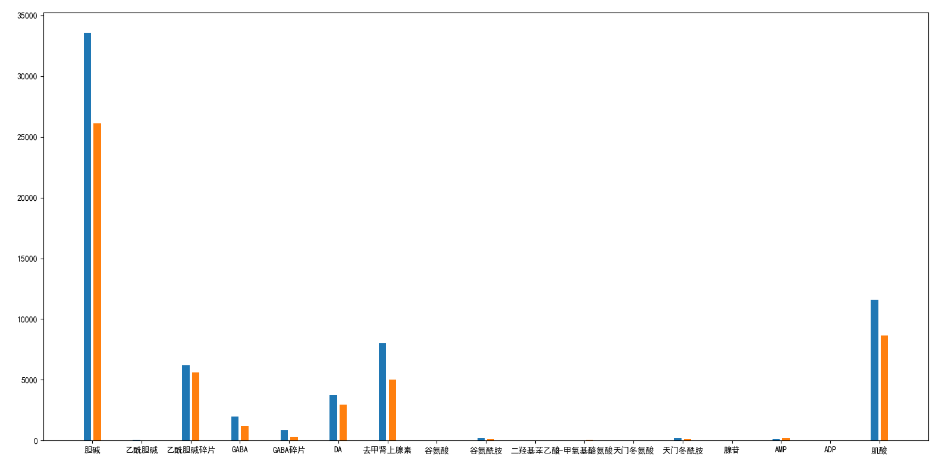


图16

**3.4.3 功能2操作指南**

1） 选区训练集数据导出

在生成对应选区之后，点击“Selection Export”按钮， 会出现“Data Export”对话框。如图17

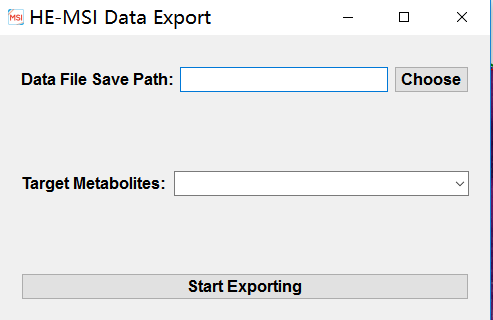
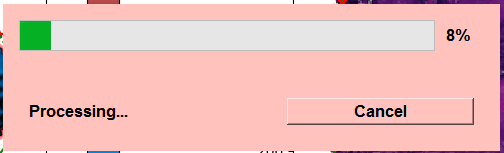


图17

在该对话框中，可以选择导出的代谢物信息文件保存路径，以及需要导出哪些代谢物。点击Start Exporting按钮后开始导出。



导出后的xls文件如图18所示。

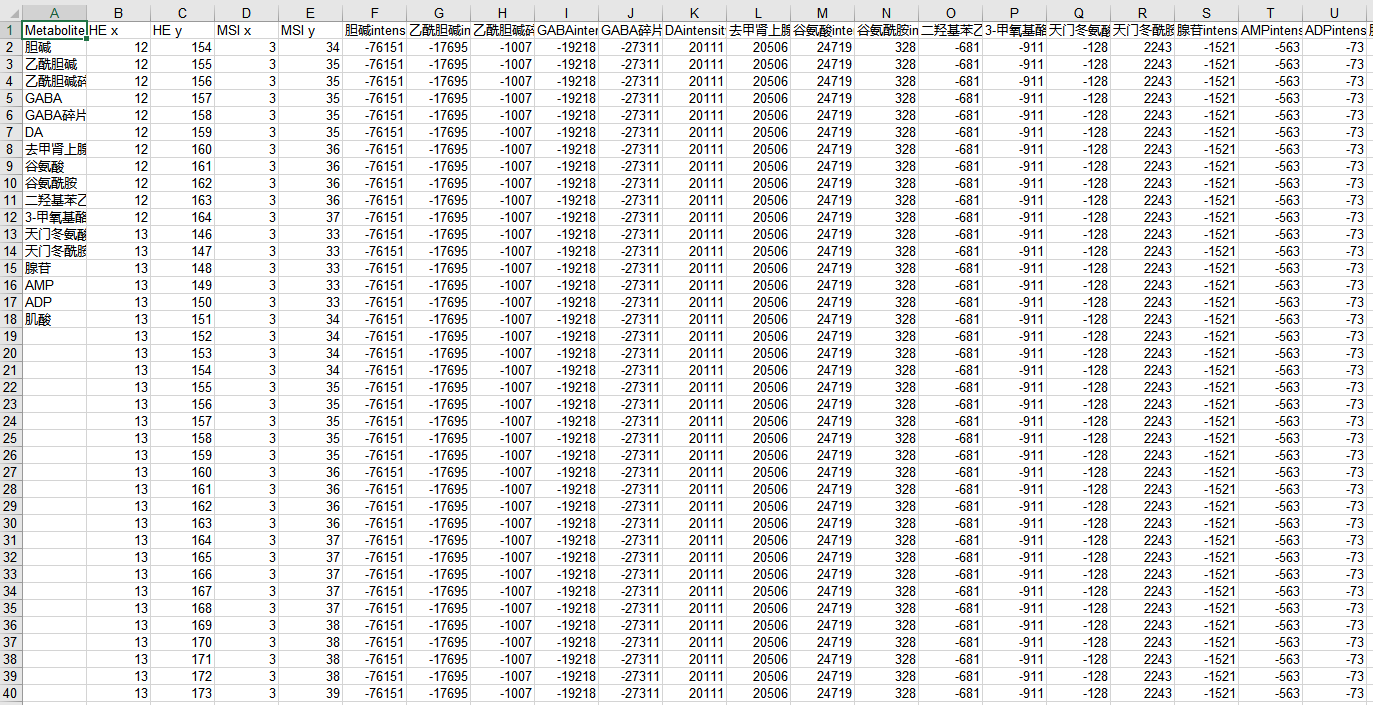


图18

2） 生成融合图像

在数据导出完成后，点击“Fusion Generate”按钮出现对话框.如图19所示.

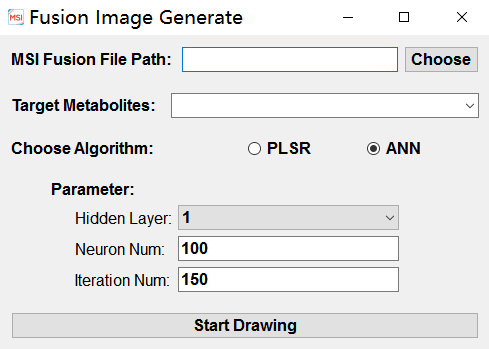


图19

选择之前导出的数据文件，选择需要进行融合成像的代谢物(一次最多选择9个).融合算法提供PLSR(偏最小二乘回归算法)和ANN(人工神经网络算法).在这里我们选择ANN.

在下方ANN算法参数里设置参数.参数范围: Hidden Layer: 1-4; Neuron Num: 100-300; Iteration Num: 100-300(参数设置需要几次摸索以达到最佳效果,设置错误会导致程序崩溃)

点击Start Drawing，经过耐心等待即可对训练集数据进行训练,并自动预测生成目标代谢物的融合图像。如图20。左方按钮用来变化图像着色。下方按钮可拉动来确定着色范围。

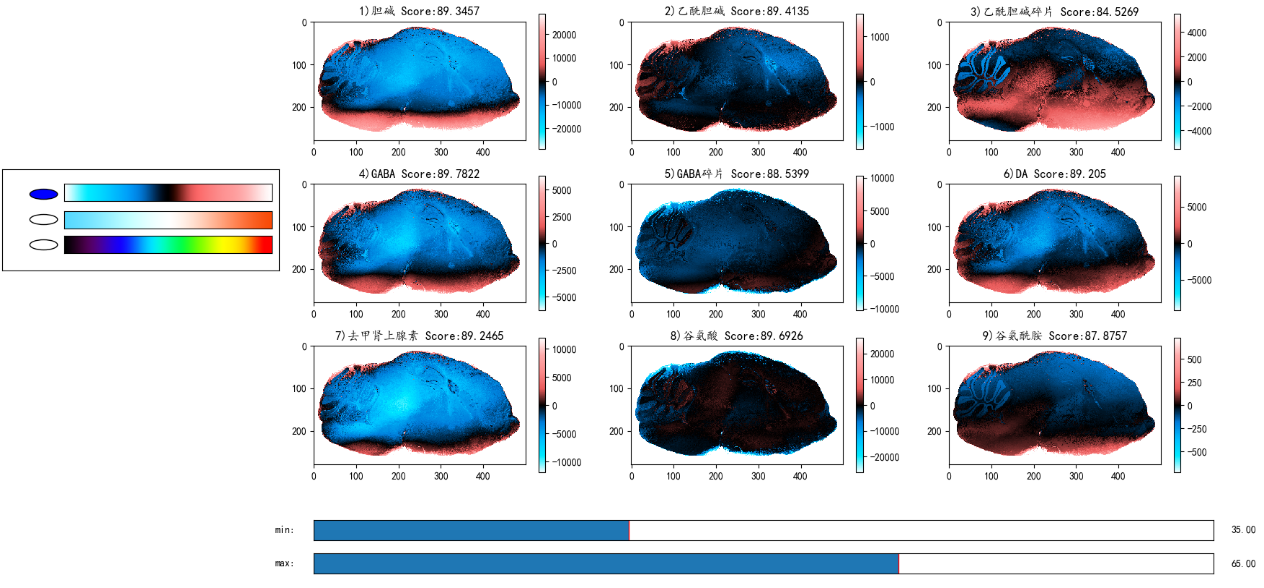


图20

**Tips: 融合图像界面操作指南**

1o 修改着色. 点击左方的颜色条即可变换着色

2o 修改特定代谢物着色强度. 将光标移动至需要进行修改的代谢物图像框中,鼠标左键点击.等到屏幕中出现如下图提示.代表系统已经调整到该代谢物着色强度调整模式.此时拖动屏幕下方的Scollbar,即可对该代谢物进行着色强度调整.

