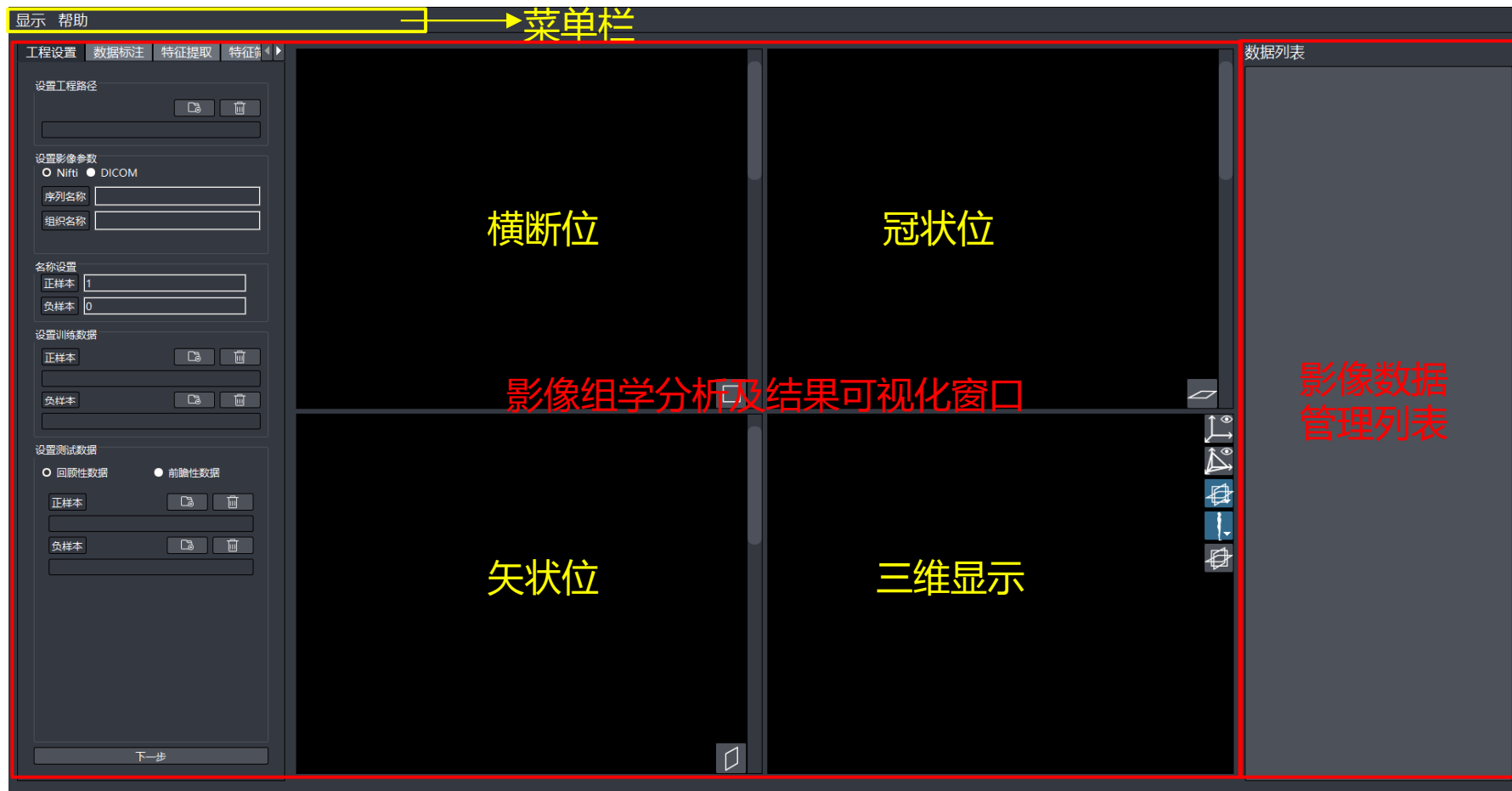


AIMS V1.0 Tutorial

The background of the slide is a dark blue gradient. On the left side, there is a complex network of glowing teal lines and dots, resembling a molecular structure or a data network. On the right side, there is a stylized, glowing teal circuit board or architectural diagram. A white horizontal line with a small white diamond in the center is positioned below the title text.

一. 产品简介

➤ 初始化的软件界面



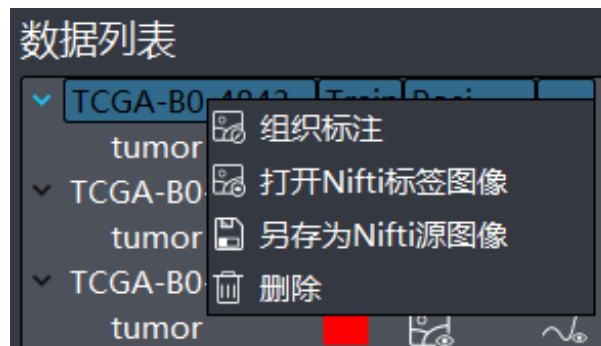
① 菜单栏：包含显示和帮助模块

② 影像组学分析及结果可视化窗口：
执行影像组学分析、切换显示
影像和分析结果

③ 数据管理列表：显示和管理影像
数据及标签数据

二. 图像标注—数据管理

1. 右键单击原始影像名称，弹出原始影像的右键菜单，可新建影像标注以及对影像数据进行管理。



影像右键菜单

2. 右键单击标签图像名称，弹出标签图像的右键菜单，选择相应的工具进行图像标注。

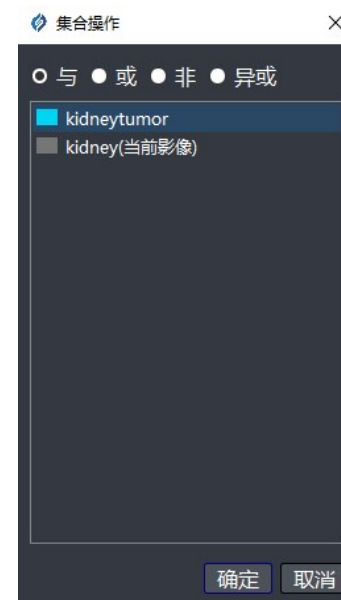
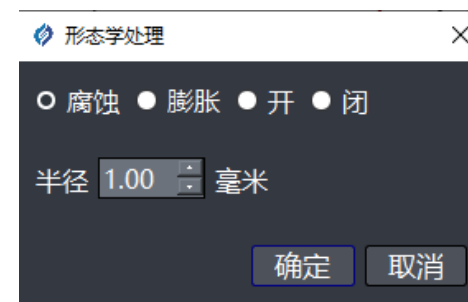


标签图像右键菜单

二. 图像标注--图像标注

3. 标签图像的右键菜单包括:

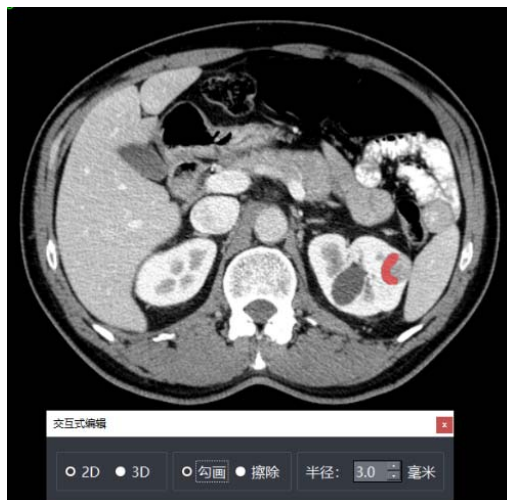
- 交互式编辑: 手动标注图像
- 阈值分割: 通过设置灰度值上下阈值来进行图像分割
- 空洞填充: 对封闭区域内的空洞进行自动填充
- 目标提取: 从不连通的多个区域中提取使用标记点选中的一个或多个区域
- 形态学处理: 对当前标签进行形态学操作, 包括腐蚀、膨胀、开、闭
- 集合操作: 对当前标签与选中标签进行集合操作, 包括与、或、非、异或
- 另存为Nifti标签图像: 将当前的标注结果另存为Nifti格式的标签图像
- 复制标签图像: 将当前标签图像复制一份到对应的数据列表中
- 删除: 删除当前标签图像



标签图像操作

二. 图像标注--标注示例1：2D/3D编辑 + 空洞填充

1. 使用2D/3D编辑，拖动鼠标左键在ROI区域勾画封闭边缘

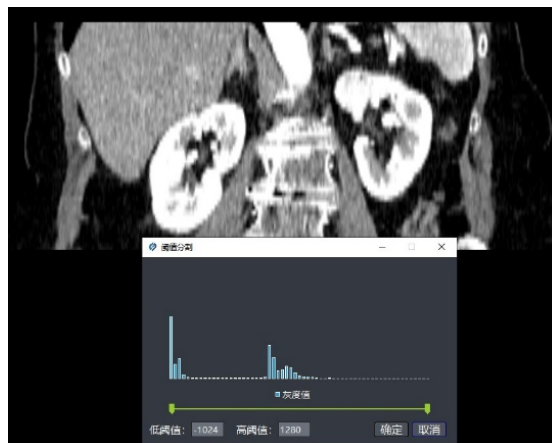


2. 使用空洞填充，对选中视图内封闭边缘内的空洞区域进行自动填充

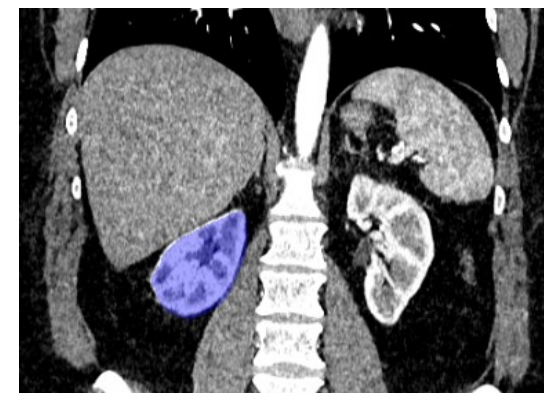


二. 图像标注--标注示例2：阈值分割 + 目标提取

1. 使用阈值分割，拖动滑动条左右两端调节高低阈值，进行阈值分割



2. 使用目标提取，在阈值分割结果的基础上，对ROI区域放置标记点，自动提取标记点所在的联通区域



三. 数据准备

情况1: 原始图像为DICOM格式

仅支持全英文路径 (文件夹和文件名
请勿包含中文字符和空格)

1. 建立工程文件夹, 用于放置所有的训练数据和测试数据
2. 建立正样本训练数据文件夹和负样本训练数据文件夹, 用于训练分类模型
3. 建立正样本验证数据文件夹和负样本验证数据文件夹, 用于评估分类模型
4. 在步骤2和3中的数据文件夹下为每一位患者建立一个相应的患者文件夹, 用于放置该位患者的全部影像数据。
5. 在患者文件夹下建立影像文件夹, 一个影像文件夹下放置一种模态/序列的影像数据。如有多种模态/序列图像, 需建立多个相应的影像文件夹。

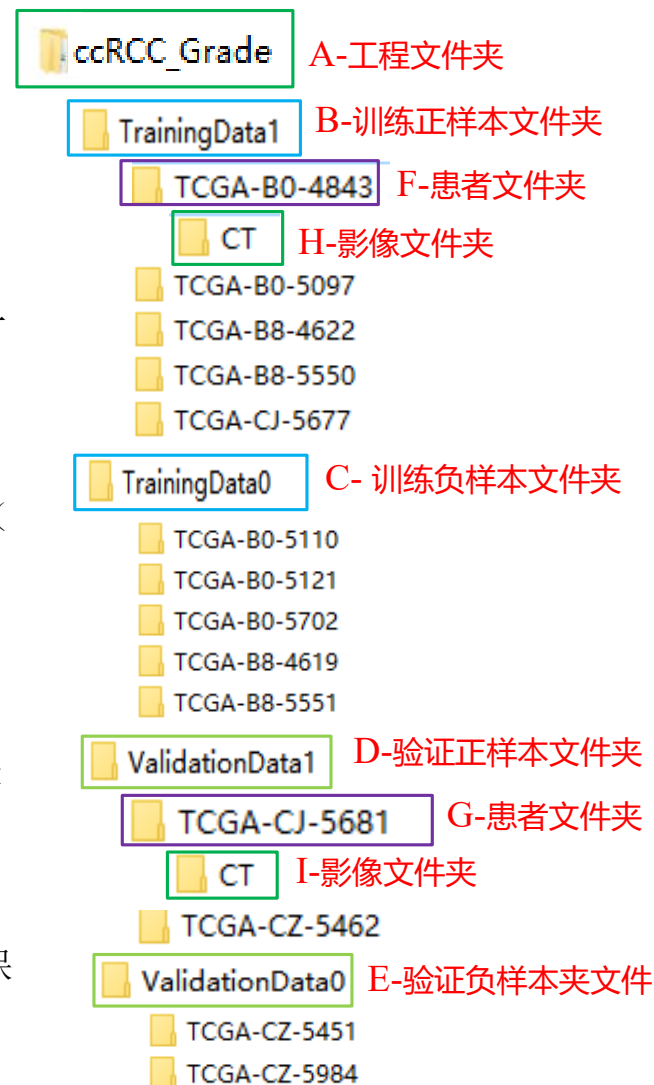
注1: 工程文件夹(A)建议以实际的科研任务命名, 训练正/负样本文件夹(B和C)以**固定的**
TrainingData1和**TrainingData0**命名 (其中1代表正样本, 0代表负样本), 验证/测试数据 (D和E) 文件夹以**固定的****ValidationData1**和**ValidationData0** (**TestData1**和**TestData0**) 命名

注2: 患者文件夹 (F和G) 的名称在满足可追溯原则下可自行命名 (最好与临床资料中的PatientID一致, 方便后续对影像特征与临床特征进行联合分析)

注3: 每位患者的影像数据包含的模态/序列需保持一致。影像文件夹 (H和I) 的名称建议根据实际模态、序列名称命名, 如每位患者仅有单个CT图像序列, 影像文件夹可命名为“CT”。如果有两种不同的序列, 可使用序列名称如“T1”和“T2”命名影像文件夹

注3: DICOM原始图像批量导入后仅用于数据标注 (软件会自动将DICOM原始图像和标签图像保存为Nifti格式), 标注完成后需重新导入Nifti格式的原始图像和标签图像。并且, 所有路径中均不能包含中文字符

DICOM数据准备示例



三. 数据准备

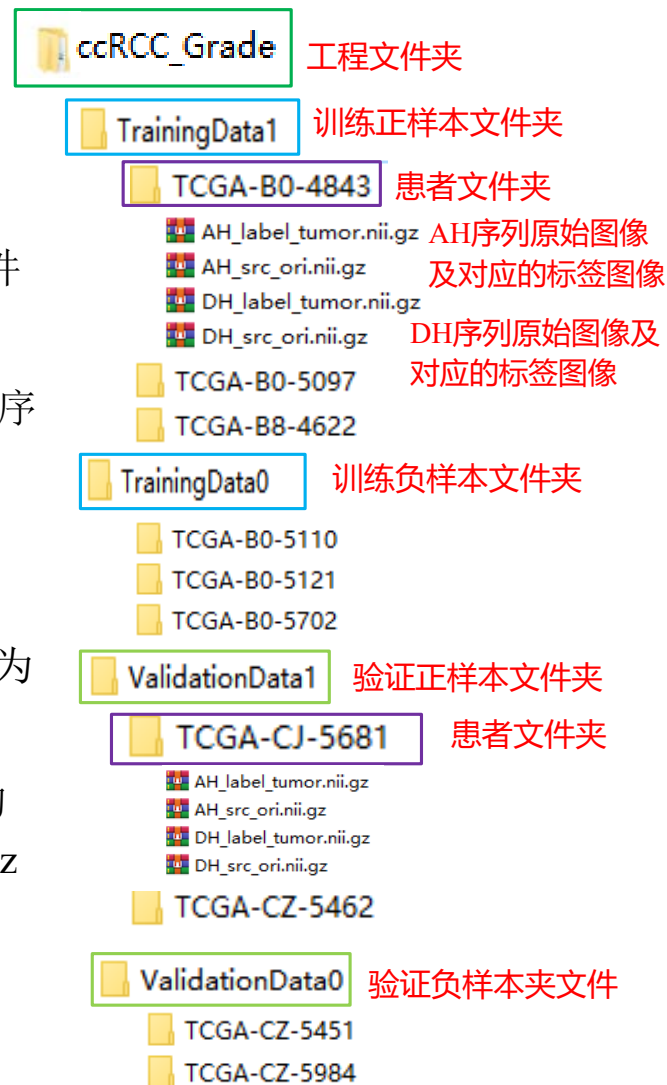
情况2：原始图像和标签图像为Nifiti格式

1. 建立工程文件夹，用于放置所有的训练数据和测试数据
2. 建立正样本训练数据文件夹和负样本训练数据文件夹，用于训练分类模型
3. 建立正样本验证数据文件夹和负样本验证数据文件夹，用于评估分类模型
4. 在步骤2和3中的数据文件夹下为每一位患者建立一个相应的患者文件夹，在该文件夹下放置患者所有的原始图像及其对应的标签图像。
5. 原始图像的命名规则为：序列名称_src_ori.nii.gz, 其对应的标签图像命名规则为：序列名称_label_组织名称.nii.gz。其中，序列名称和组织名称可根据实际情况自行命名，标红字段为固定字段，不可修改。

例1：每位患者仅有单个动脉期CT图像序列，则每位患者的原始图像可统一命名为：**AH**_src_ori.nii.gz，相应的标签图像需统一命名为：**AH_label_tumor**.nii.gz，其中AH为自行设定的序列名称，表示动脉期CT，tumor表示分析的组织为肿瘤区域。

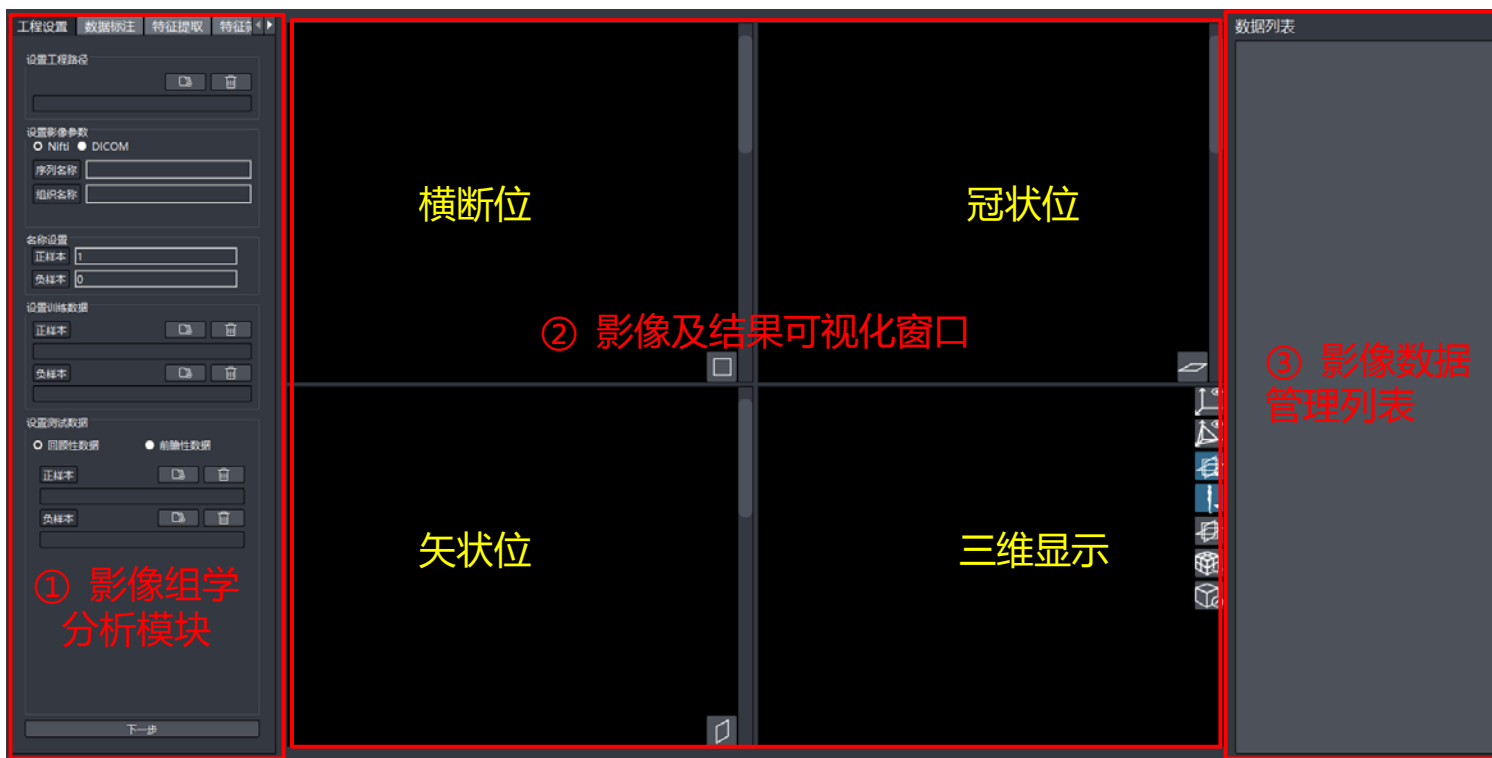
例2：每位患者有动脉期和延迟期（分别用AH和DH表示）的CT序列，且组织名称均为tumor，则两种序列的原始图像可分别命名为：**AH**_src_ori.nii.gz和**AH**_src_ori.nii.gz，相应的标签图像需分别命名为：**DH_label_tumor**.nii.gz和**DH_label_tumor**.nii.gz。

Nifti数据准备示例



四. 影像组学分析

➤ 影像组学模块: 基于影像组学技术进行预测模型构建与应用



① 影像组学分析模块: 基于影像组学技术构建辅助诊断模型

② 影像及结果可视化窗口: 三维影像数据浏览显示、智能分析结果输出及可视化

③ 影像数据管理列表: 对影像数据和标注数据进行管理

四. 影像组学--工程设置

➤ 设置工程路径和实验数据，后续的实验结果将自动保存在工程路径下

- 第1步：点击**A**处，设置工程文件夹（若该路径中已有实验结果，则软件将自动导入）
- 第2步：点击**B**处，选择需导入数据的类型（未进行标注情况下，可选择导入DICOM和Nifti，此时仅需输入序列名称；已标注完成的情况下，必须选择导入Nifti，且需要同时输入序列名称和组织名称，才能进行后续的影像组学分析）
- 第3步：在**B-a**处输入序列名称。输入为DICOM图像时，序列名称即为影像文件夹名称。如存在多个模态/序列图像，名称之间使用英文“,”区分，如ct, mr或者AH,DH。
- 第4步：为进行后续的影像组学分析，需在**B-a**处输入拟分析的感兴趣区域名称（与标注的标签图像同名，如tumor）。输入为DICOM图像时，默认为未经标注的原始图像，不能输入感兴趣区域名称。
- 第5步：为明确正样本和负样本的实际名称，用户可按实际情况在**C**处输入，如正样本处（默认显示1）输入恶性肿瘤，负样本处（默认显示0）输入良性肿瘤
- 第6步：点击**D**导入训练数据集的正样本数据路径（路径选择到训练正样本文件夹）
- 第7步：点击**E**导入训练数据集的负样本数据路径（路径选择到训练负样本文件夹）
- 第8步：在**F**处选择数据类型：回顾性数据（需按照正负样本进行准备）或前瞻性数据（需进行感兴趣区域标注，但无需区分正负样本）
- 第9步：选择回顾性数据，点击**F-a**导入测试正样本（路径选择到测试正样本文件夹），点击**F-b**导入测试负样本（路径选择到测试负样本文件夹）。若选择前瞻性数据，直接导入前瞻性测试数据（路径选择到前瞻性数据文件夹）
- 第10步：点击**G**跳转至“数据标注”页面

注1：第7和8步是非必需步骤，若需要进行回顾性数据测试或前瞻性数据测试，请在此导入相关数据

注2：使用Nifti格式进行数据准备，得到的患者文件夹即可作为单个前瞻性测试数据；如需进行批量前瞻性测试，可以将多个患者文件夹统一放到前瞻性数据文件夹下（文件夹名称可自行设置，建议命名为Prospective data）。

四. 影像组学--工程设置示例

The screenshot shows the '工程设置' (Project Settings) window. The '设置影像参数' (Set Image Parameters) section has the 'DICOM' radio button selected. The '序列名称' (Sequence Name) field contains 'CT'. The '名称设置' (Name Settings) section has '正样本' (Positive Sample) set to '高分级' (High Grade) and '负样本' (Negative Sample) set to '低分级' (Low Grade). The '设置训练数据' (Set Training Data) section has '正样本' (Positive Sample) set to 'D:/ccRCC Grade/TrainingData1' and '负样本' (Negative Sample) set to 'D:/ccRCC Grade/TrainingData0'. The '设置测试数据' (Set Test Data) section has '前瞻性数据' (Prospective Data) selected, with '正样本' (Positive Sample) set to 'D:/ccRCC Grade/ValidationData1' and '负样本' (Negative Sample) set to 'D:/ccRCC Grade/ValidationData0'.

原始图像为DICOM格式，
导入序列名称为CT的图像

The screenshot shows the '工程设置' (Project Settings) window. The '设置影像参数' (Set Image Parameters) section has the 'Nifti' radio button selected. The '序列名称' (Sequence Name) field contains 'CT'. The '组织名称' (Tissue Name) field contains 'tumor'. The '名称设置' (Name Settings) section has '正样本' (Positive Sample) set to '高分级' (High Grade) and '负样本' (Negative Sample) set to '低分级' (Low Grade). The '设置训练数据' (Set Training Data) section has '正样本' (Positive Sample) set to 'D:/ccRCC Grade/TrainingData1' and '负样本' (Negative Sample) set to 'D:/ccRCC Grade/TrainingData0'. The '设置测试数据' (Set Test Data) section has '前瞻性数据' (Prospective Data) selected, with '正样本' (Positive Sample) set to 'D:/ccRCC Grade/ValidationData1' and '负样本' (Negative Sample) set to 'D:/ccRCC Grade/ValidationData0'.

原始图像为Nifti格式，
导入序列名称为CT，标注的
感兴趣区域为tumor的图像

The screenshot shows the '工程设置' (Project Settings) window. The '设置影像参数' (Set Image Parameters) section has the 'Nifti' radio button selected. The '序列名称' (Sequence Name) field contains 'AH,DH'. The '组织名称' (Tissue Name) field contains 'tumor'. The '名称设置' (Name Settings) section has '正样本' (Positive Sample) set to '高分级' (High Grade) and '负样本' (Negative Sample) set to '低分级' (Low Grade). The '设置训练数据' (Set Training Data) section has '正样本' (Positive Sample) set to 'D:/ccRCC Grade/TrainingData1' and '负样本' (Negative Sample) set to 'D:/ccRCC Grade/TrainingData0'. The '设置测试数据' (Set Test Data) section has '前瞻性数据' (Prospective Data) selected, with '正样本' (Positive Sample) set to 'D:/ccRCC Grade/ValidationData1' and '负样本' (Negative Sample) set to 'D:/ccRCC Grade/ValidationData0'.

原始图像为Nifti格式，
导入序列名称为AH和DH，标注的
感兴趣区域为tumor的图像

四. 影像组学--数据标注

➤ 在原始图像上标注待分析的感兴趣区域，后续针对感兴趣区域进行分析建模

第1步：在数据列表中左键单击待标注原始图像的名称，选中后高亮显示(如A所示)

第2步：右键单击图像名称，选择“组织标注”，新建一个默认名称为“tumor”的空白标签图像(如B所示)。

第3步：新建的标签图像默认为选中状态，通过左侧标注工具栏C中的快捷按钮，快速启动“交互式编辑”、“阈值分割”、“目标提取”、“空洞填充”、“形态学处理”和“集合操作”，组合使用不同的标注工具，实现感兴趣区域标注。

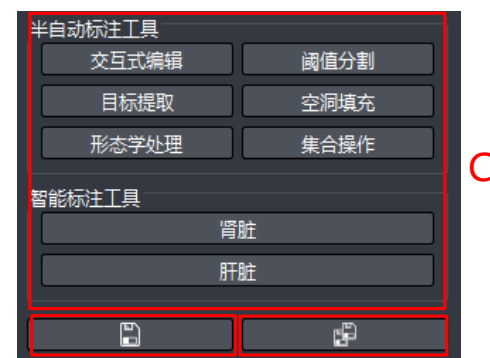
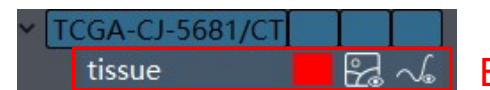
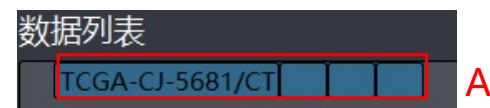
第4步：当数据列表中存在多个图像数据时，标注完一例图像后，建议在数据管理列表直接点击下一个原始图像名称，软件将自动按照命名规则保存当前的原始图像和标签图像，无需用户进行手动保存。

第5步：重复步骤1~4, 待全部数据标注完毕后，点击E处进行保存，完成数据标注。

第6步：如果导入的图像数据是DICOM格式的，完成所有的数据标注后，回到工程设置，设置影像参数为“Nifti”，输入相应的序列名称和组织名称，并重新设置训练数据和测试数据，点击“下一步”，再次点击“下一步”跳转到“特征提取”。

注：点击D处可手动将当前选中的标签图像另存为Nifti文件，用户可设置标签图像文件名。

注：如需批量修改组织标注名称，请参考第四部分“工具”中的“组织重命名”。



D

E

四. 影像组学--特征提取

➤ 针对感兴趣区域提取高通量影像组学特征

第1步：在A处单击选中需要提取特征的数据类型，可根据需要选中训练数据和/或测试数据，以减少不必要的特征提取时间

第2步：根据需要，在B处单击选中相应的图像滤波器，每选中一个图像滤波器，都将对经过滤波变换后的图像提取所有类型的特征

第3步：根据需要，在C处单击选中相应的特征类型。其中，如果感兴趣区域为二维，需勾选“二维形状特征”。D处将根据选择的图像类型和特征类型，自动显示将提取的特征数量。

第4步：点击E，将进行特征提取。特征提取过程中，请勿操作或关闭软件。特征提取完成后，将自动在工程路径下生成“Feature”文件夹，以CSV和mat（Matlab软件）两种格式保存相应的特征文件。并自动跳转至“特征筛选”。

注：特征提取阶段较为耗时，执行过程中用户可最小化弹窗及软件。针对同一批数据以及相同的特征提取设置，特征提取仅需执行一次即可。

图像类型：

Original：原始图像

Log：经过高斯拉普拉斯滤波（边缘增强滤波）的图像

Wavelet：经过小波变换的图像

Square：经过图像强度值平方变换的图像

SquareRoot：经过图像强度值平方根变换的图像

Logarithm：经过图像绝对强度对数变换的图像

Exponential：经过图像强度值指数变换的图像

Gradient：经过图像局部梯度变换的图像

LocalBinaryPattern3D：经过局部二值模式3D变换的图像

特征类型：

firstorder：一阶统计学特征

shape：形状特征，如果输入为二维图像，需勾选“二维形状特征”

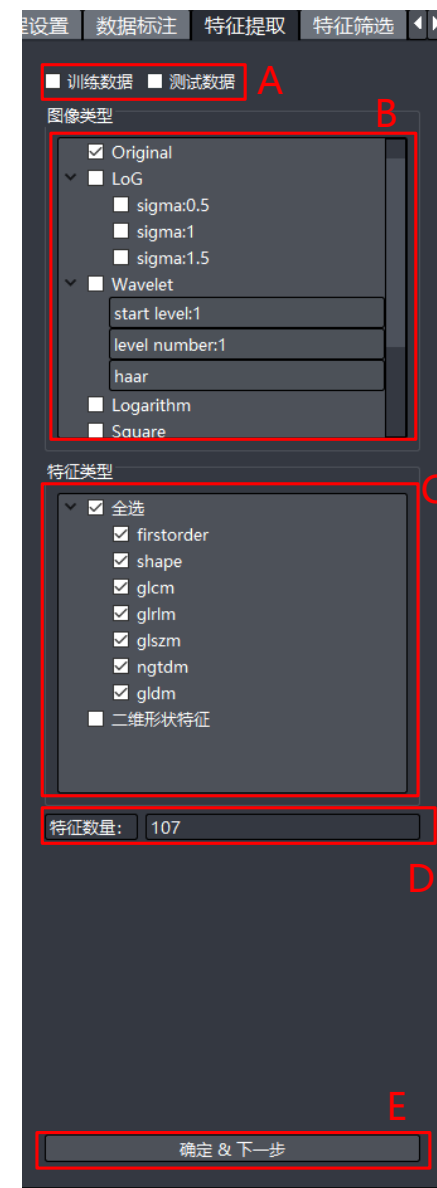
glcm：灰度共生矩阵

glrlm：灰度游程矩阵

glszm：灰度区域大小矩阵

ngtdm：邻域灰度差分矩阵

gldm：灰度相关矩阵



四. 影像组学--特征筛选

➤ 针对具体任务从高通量影像组学特征中筛选出关键特征

第1步：在A处选择特征归一化算法（Standardization、Centering、UnitRange），防止特征之间量级不同导致对分类效果的不利影响

第2步：在B处选择一个特征选择算法（或多个算法的组合），以针对具体任务筛选出具有较高鉴别力的特征子集，利于后续分类模型的精度和效率。当选择多个特征选择算法时，将按照显示的顺序依次执行，下一个特征选择算法将在前面的特征选择算法筛选出的特征中进一步选择特征（None表示空，即没有选择特征选择算法）。

第3步：在C处对每一个当前选中的特征选择进行相应的参数设置。

第4步：点击D，跳转至“训练”。

注：若当前工程已进行了特征提取，在没有修改图像数据和特征提取参数的情况下，可在“工程设置”后直接左键单击E，进行特征筛选。

归一化方法

- Standardization: 标准化（将每个特征变换到均值为0，标准差为1）
- Centering: 中心化（每个特征减去训练样本中该特征的平均值）
- UnitRange: 归一化（将每个特征归一化到[0,1]）

特征选择算法

- VarianceThreshold: 特征方差
- t-test: t检验算法
- Correlation analysis: 相关性分析算法（包括Pearson、Spearman和Kendall）
- mRMR: 最大相关最小冗余算法（包括MIQ和MID）
- Lasso: 最小绝对值收敛和选择算法
- SequentialFeatureSelect: 序列化特征选择算法（SFS:序列前向选择算法、SBS:序列后向选择算法、SBFS:序列后巷浮动选择算法、SFFS:序列前向浮动选择算法）



四. 影像组学--分类模型训练

➤ 基于筛选后的特征子集，选择相应的分类算法，在训练数据上构建分类器，得到针对具体任务的预测模型

第1步：在**A**处左键单击方框按钮，选择相应的分类算法，训练完成后将得到所有选中算法的分类器（默认选中C-SVC分类器）

第2步：在**B**处设置被选中的分类算法的参数，选中Grid Optimization可以针对具体任务优化相关分类算法的超参数，这将会增加训练时间，但通常将获得更优的分类模型。

第3步：在**C**处设置交叉验证或者划分数据集的参数（自动将剩余比例的训练数据作为内部验证数据）以及重复实验的次数，训练完成后将输出在验证集上的结果，用于评估分类器性能

第4步：在**D**处选择交叉验证的准则，选择保存在验证数据上取得最高AUC（ROC曲线下面积）对应的分类器参数，或者保存在验证数据上取得最高ACC（分类准确率）对应的分类器参数，以构建最终的分类器。

第5步：点击**E**，开始进行训练。训练完成后将生成面向该任务筛选出的影像组学特征子集（自动保存在工程路径下的Feature文件夹下）以及预测模型（自动保存在工程路径下的Model文件夹下），用于后续的回顾性测试和前瞻性测试。

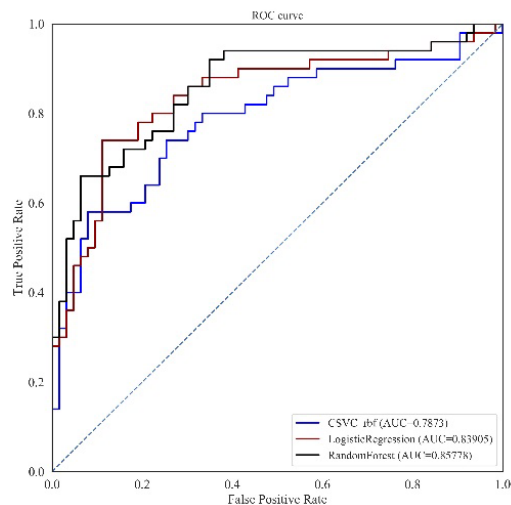
注：训练耗时与训练数据量、分类器个数及参数设置和实验设置相关，通常可以在数十分钟内完成。

The screenshot shows a software interface for training a classification model. It is divided into several sections, each highlighted with a red box and a corresponding letter label (A-E) on the right side:

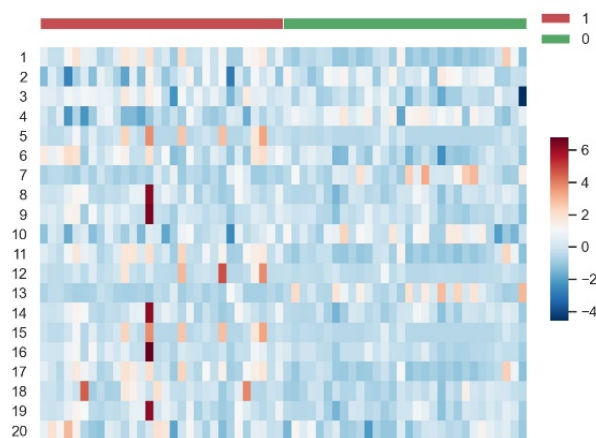
- A (Classifier Selection):** A list of classifiers with checkboxes. ☒ C-SVC, ☐ Nu-SVC, ☐ LogisticRegression, ☐ Random Forest, ☐ AdaBoost, ☐ XGBoost.
- B (C-SVC Parameter Settings):** A section for configuring C-SVC parameters. It includes a slider for 'C' (set to 1.0), a dropdown for '核类型' (Kernel Type) set to 'rbf', a dropdown for '自由度' (Degree of Freedom) set to '3', a checkbox for 'gamma' set to '自动' (Automatic) with a value of '0.00010', and a dropdown for 'coef0' set to '0.00'. The 'Grid Optimization' checkbox is also checked.
- C (Experiment Settings):** A section for configuring experimental parameters. It includes a dropdown for '重复实验次数' (Repeat Experiment Times) set to '5', a radio button for 'K折交叉验证' (K-fold Cross Validation) which is selected, a dropdown for 'k-折交叉验证' (k-fold Cross Validation) set to '5', a radio button for '自动划分数据集' (Automatic Dataset Division) which is selected, and a dropdown for '训练数据比例' (Training Data Ratio) set to '0.80'.
- D (Cross-validation Criterion):** A section for selecting the cross-validation criterion. It has two radio buttons: 'AUC' (selected) and 'ACC'.
- E (Train & Next Step):** A large button at the bottom labeled '训练 & 下一步' (Train & Next Step).

四. 影像组学--训练结果

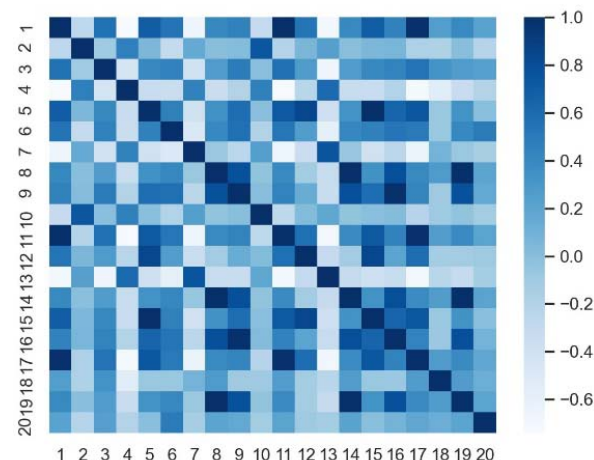
- 训练完成后，将依次输出验证统计结果图（用户手动关闭当前图像后，自动显示下一幅图像，所有图像会自动保存在工程路径下的Evaluation文件夹中）



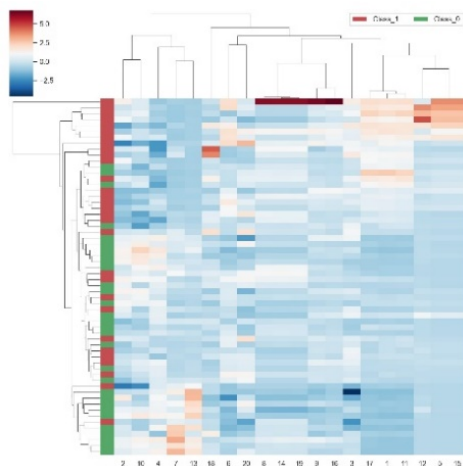
ROC曲线图（交叉验证）



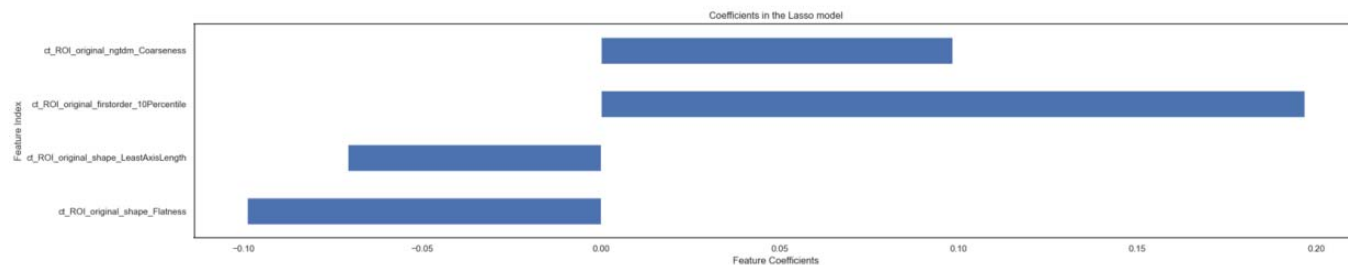
特征数值热图



特征间相关系数热图



特征聚类热图

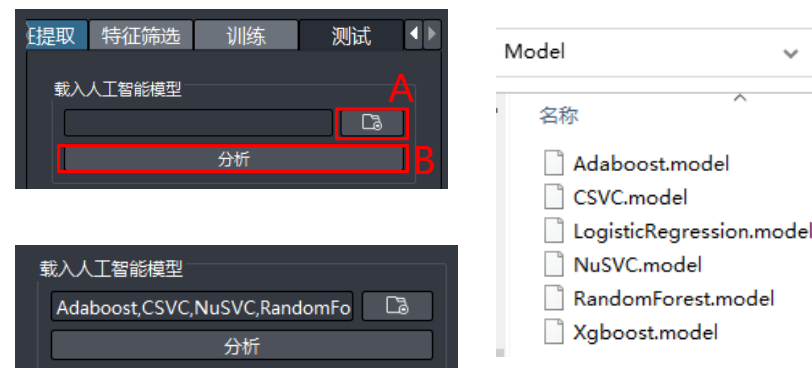


Lasso特征选择结果图

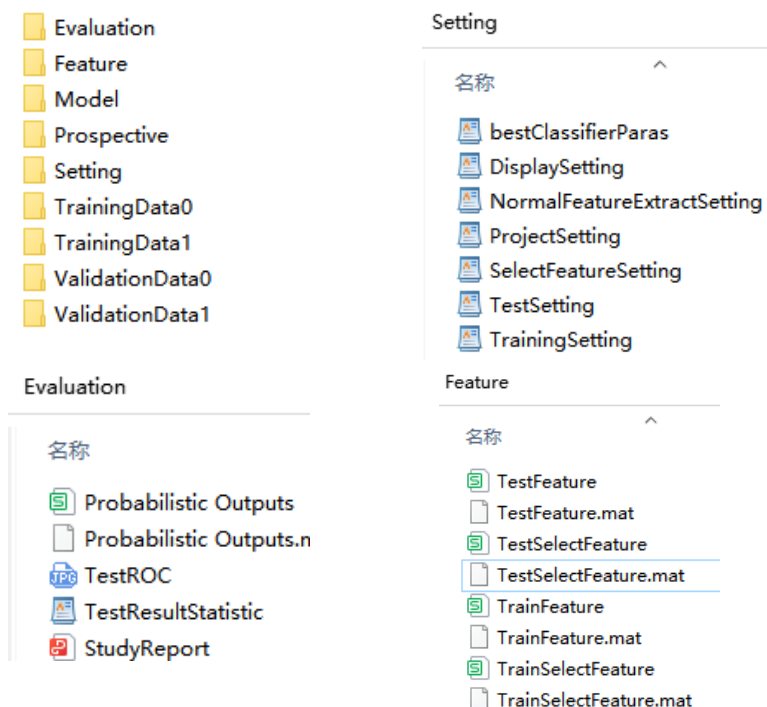
四. 影像组学--分类模型预测

➤ 使用训练得到的预测模型，对回顾性/前瞻性测试数据进行预测

第1步：点击**A**自动打开工程文件夹下的model目录，左键单击选择已训练的模型文件（.model文件，按住ctrl单击模型文件可同时选择多个模型文件）。



第2步：点击**B**进行分析，对设置好的回顾性测试数据或前瞻性测试数据（需经过数据标注和特征提取）进行预测。



工程路径下生成的配置文件和实验结果

载入人工智能模型

Adaboost,CSVC,LogisticRegression,

分析

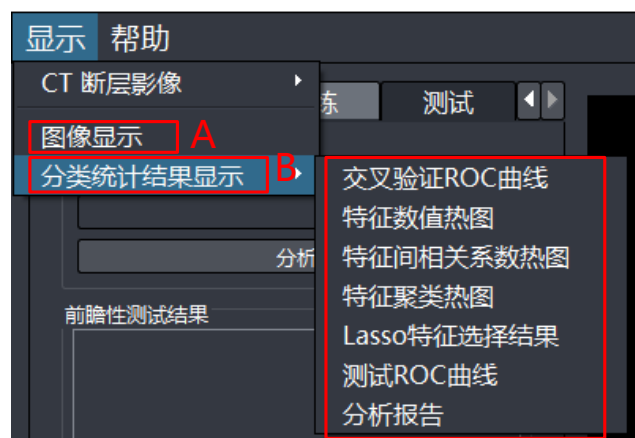
Study ID	Adaboost	CSVC
TCGA-B0-46...	高分级	高分级
TCGA-B0-48...	高分级	低分级
TCGA-B0-48...	低分级	低分级
TCGA-B0-48...	高分级	高分级
TCGA-B0-50...	低分级	低分级
TCGA-B0-51...	低分级	低分级
TCGA-B0-51...	低分级	低分级
TCGA-B0-57...	高分级	高分级
TCGA-B0-57...	低分级	低分级
TCGA-B8-51...	低分级	低分级

前瞻性数据预测示例

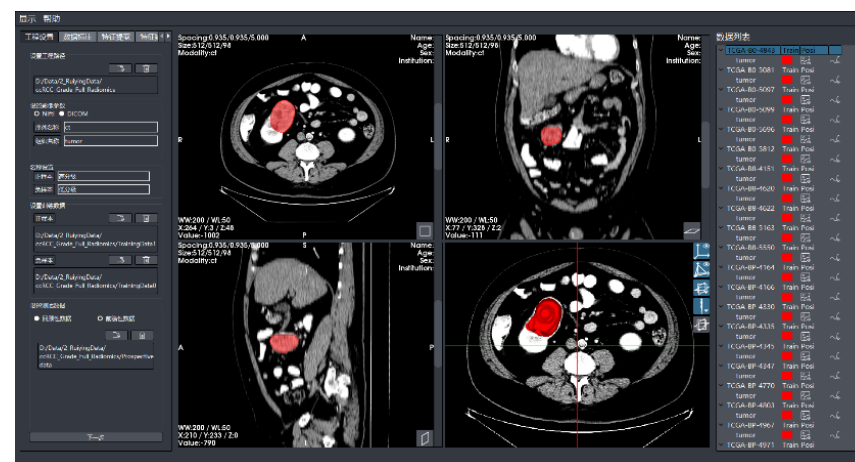
四. 影像组学--分类结果可视化

第1步：点击**A**“图像显示”，可从分类统计结果显示切换为图像显示；

第2步：点击**B**“分类统计结果显示”中的任一项，可从图像显示切换回相应的统计结果显示。
。如果选择“分析报告”，将使用默认的PDF阅读器打开自动生成的实验统计报告



显示工具栏



图像显示

AIMS实验统计报告

生成日期: 2023/11/26

工程路径: D:/ccRCC_Grade

序列名称: ct

组织名称: tumor

训练集中正样本个数: 44

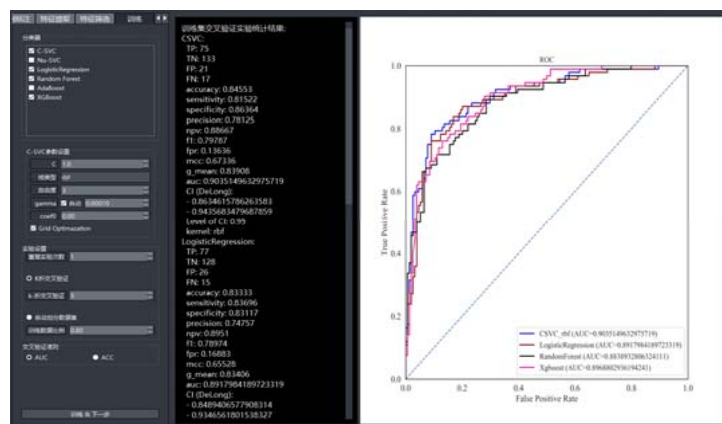
训练集中负样本个数: 67

测试集中正样本个数: 28

测试集中负样本个数: 45

前瞻性样本个数: 10

实验统计报告



分类统计结果显示