

ICC 计算结果报告

2019.9.11

计算数据：

我们对 2d 与 3d 数据的 group1（从 306 例数据中相应抽取出的 25 例）、group2（第二次绘制的 25 例数据）、group3（第二次绘制的 25 例数据）之间相互计算了 ICC 值，分别得到了 ICC_{12} 、 ICC_{13} 、 ICC_{23} 、 ICC_{123} ，并将他们的分布绘制成直方图。

计算结果：

3d 数据：

3d 数据计算得到的各 ICC 直方图如图 1 所示

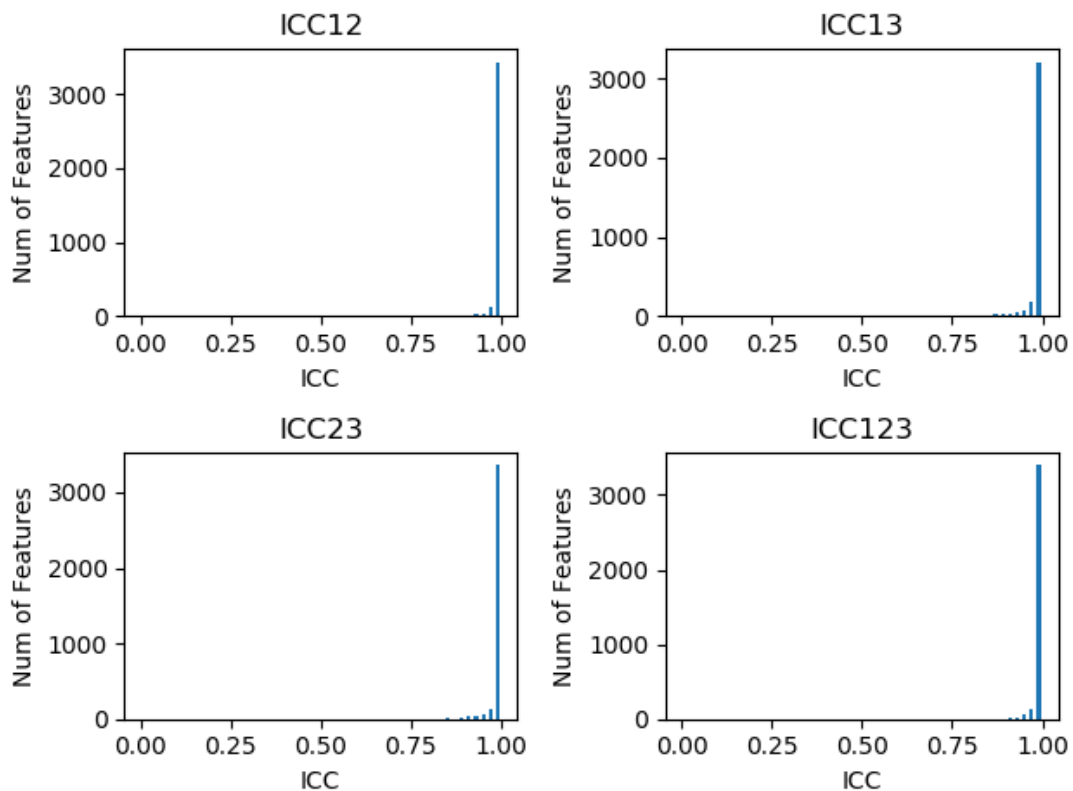


图 1 3d 数据计算所得各 ICC 直方图一览

由于绝大部分的 ICC 值都落在 0.9 到 1.0 之间，所以我们将该区间的直方图单独展示（特征数量超过 500 的部分略去）如图 2。

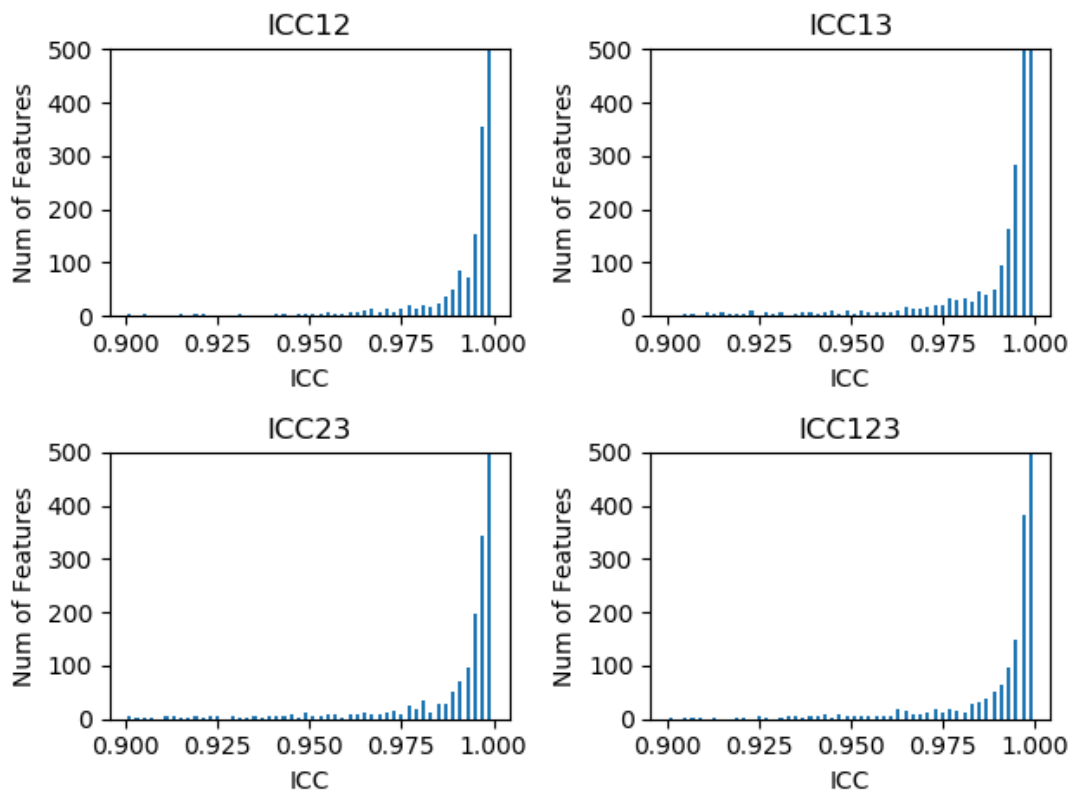


图 2 ICC 值在 (0.900, 1.000) 之间的 3d 数据各 ICC 直方图一览 (略去特征数大于 500 部分)

其中 ICC_{12} 、 ICC_{13} 、 ICC_{23} 、 ICC_{123} 的最小值分别为 0.258、0.389、0.220、0.650，最大值都可以达到 1.000。我们计算了在设定不同数值为阈值时，可以筛去的特征数（即 ICC 值低于阈值的特征的数量），具体可见表 1 所示。

表 1 3 数据 ICC 值的一些统计数据

	ICC_{12}	ICC_{13}	ICC_{23}	ICC_{123}
min	0.258	0.220	0.389	0.650
max	1.000	1.000	1.000	1.000
under 0.90	23	97	35	20
under 0.92	39	133	66	29
under 0.95	74	228	128	87

另外，我们还计算了 ICC_{12} 、 ICC_{13} 、 ICC_{23} 两两之间的皮尔逊相关系数，如表 2 所示。

表 2 2d 数据 ICC_{12} 、 ICC_{13} 、 ICC_{23} 两两之间的皮尔逊相关系数

	ICC_{12}	ICC_{13}	ICC_{23}
ICC_{12}		0.692	0.175
ICC_{13}	0.692		0.782
ICC_{23}	0.175	0.782	

2d:

由于 ROI 区域是二维的单层形状, 所以以下 6 个特征没有提取得到, 它们分别表示平扫期、动脉期、静脉期数据的 ROI 区域的平面度、包含 ROI 的椭球体最小轴长特征。

2d 数据计算得到的各 ICC 直方图如图 3 所示

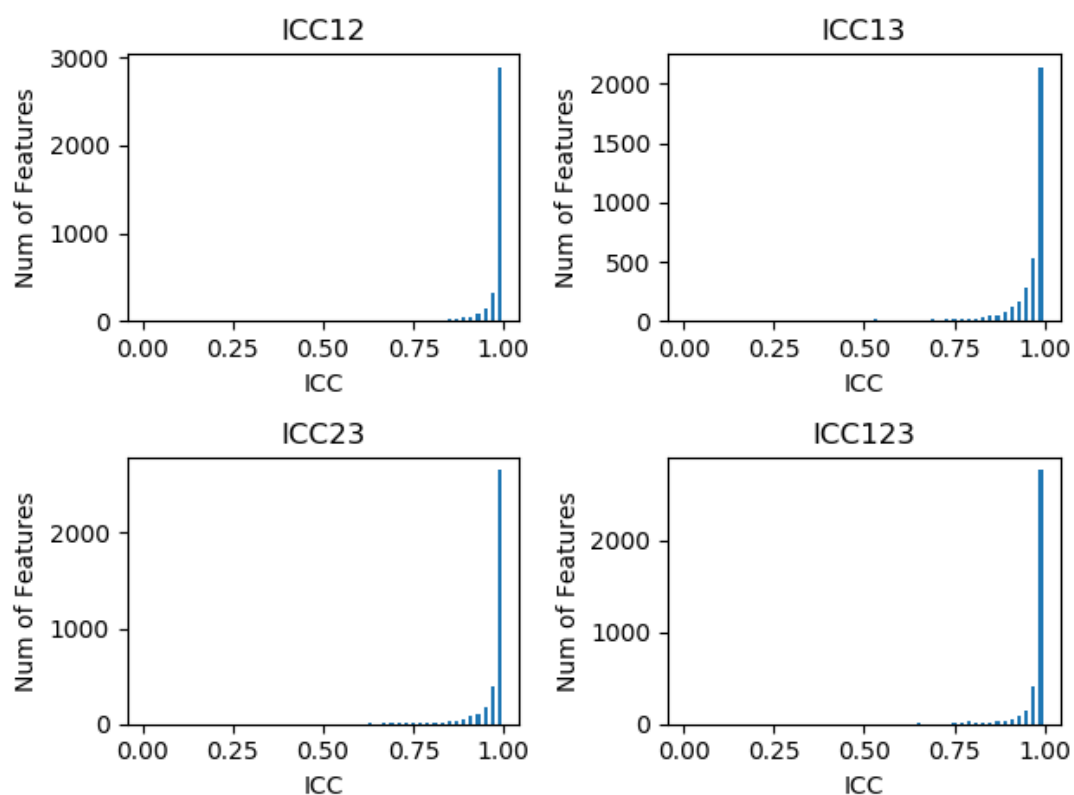


图 3 2d 数据计算所得各 ICC 直方图一览

其中 ICC_{12} 、 ICC_{13} 、 ICC_{23} 、 ICC_{123} 的最小值分别为 0.089、-0.109、0.411、0.291, 最大值都可以达到 1.00, 我们计算了在设定不同数值为阈值时, 可以筛去的特征数 (即 ICC 值低于

阈值的特征的数量), 具体可见表 2 所示。

表 3 2d 数据 ICC 值的一些统计数据

	ICC ₁₂	ICC ₁₃	ICC ₂₃	ICC ₁₂₃
min	0.089	-0.109	0.411	0.291
max	1.000	1.000	1.000	1.000
under 0.90	154	425	245	176
under 0.92	197	538	324	228
Under 0.95	354	807	502	376

由于绝大部分的 ICC 值都落在 0.9 到 1.0 之间, 所以我们将该区间的直方图单独展示 (特征数量超过 800 的部分略去), 如图 4。

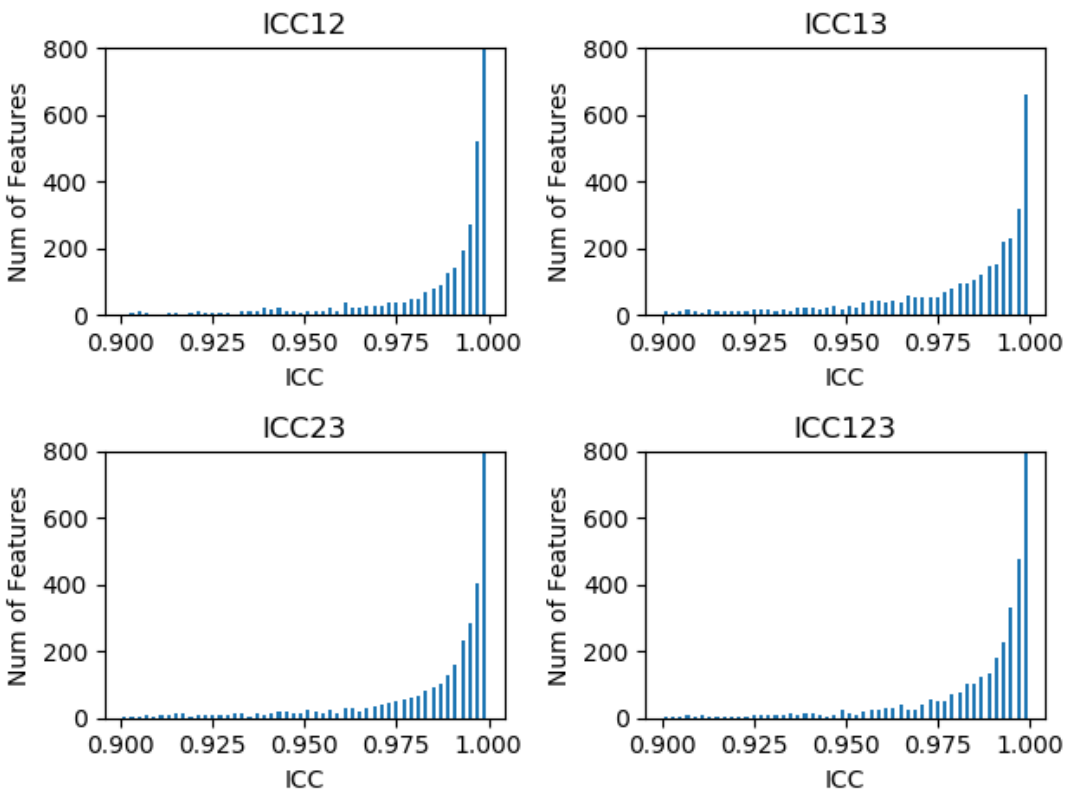


图 4 ICC 值在 (0.900, 1.000) 之间的 2d 数据各 ICC 直方图一览 (略去特征数大于 800 部分)

另外, 我们还计算了 ICC₁₂、ICC₁₃、ICC₂₃ 两两之间的皮尔逊相关系数, 如表 4 所示。

表 4 3d 数据 ICC_{12} 、 ICC_{13} 、 ICC_{23} 两两之间的皮尔逊相关系数

	ICC_{12}	ICC_{13}	ICC_{23}
ICC_{12}		0.820	0.487
ICC_{13}	0.820		0.857
ICC_{23}	0.487	0.857	

结果讨论：

由以上结果可以看出，无论是 2d 数据还是 3d 数据，三组数据所提取到的特征所计算得到的 ICC 都相当高，也就是说特征的可重复度与信度很高。另一方面，也表示了我们无法通过计算 ICC 的方式达到很好的降维目的，因为可以筛去的特征十分有限。

目前所需要解决的问题是：

1. 2d 与 3d 数据的 ICC 值各选取那个数为阈值较好
2. 如果有一些特征的某一 ICC 值高，但是另一 ICC 的值低，我们是否应该删去这个特征。

2d以0.9为阈值，3d以0.95为阈值，其实最好是能把阈值设置成一样的，只是这个结果和我推想的有点不一样。另外，因为三期需独立建立模型，所以要把三期分开

比如 2d 数据中静脉期“小波变换后得到的像素大小平均值”这一特征， ICC_{12} 、 ICC_{13} 、 ICC_{23} 、

ICC_{12} 的数值分别为 0.089、0.952、-0.109、0.569。

这个是需要排除的，我们所要用的特征是 ICC_{12} 和 ICC_{13} 所共有的。但是2D和3D是相互独立的，需分开筛选

3. 如果有一些特征在 3d 数据中 ICC 值较高，但在 2d 数据中较低，我们应该只剔除 3d 数据中的该特征，还是在 3d 和 2d 数据中都删去。

2D数据与3D数据相互独立，互不干涉

比如还是第 2 点中距离的特征，它在 3d 数据中的 ICC 值分别为 0.985、0.924、0.913、0.960，明显高于在 2d 数据的 ICC，应该怎么看待这种特征。

4. 如果要同时考虑两个 ICC 值来进行筛选的话，是否直接根据 ICC_{123} 来筛选更加方便，而且解释性更好？

可以，但是不严谨，还是推荐用两个 ICC 值来筛选

5. 为何 group1 与 group2 的 ICC 值表示两个医生对同一组影像数据绘制 ROI，group1 与 group3 的 ICC 值表示同一个医生间隔 2 个月后重新对同一组数据绘制 ROI。

我个人的理解是因为如同医生上次所说由于绘制 ROI 的人手与时间有限，所以只能将本来想相隔两个月绘制的 ROI 作为不同医生绘制的来使用，而相隔较短时间绘制的 ROI 作为相隔 2 个月重新绘制的来使用。

不知以上理解是否正确。

理解是正确的

6. 有些 ICC 值出现负数，经过查询以后发现它确实可以是负值，但是负数的意义尚没有查询得到一个明确解释。我认为本研究中暂时不需要了解负值的意义，打算将其优先级放

不需要讨论负值的问题

低。