**Open data for global multimodal land use classification - outcome of the 2017 IEEE GRSS data fusion contest**

**内容：**本文总结了2017 Data Fusion Contest（DFC17）上各参与者针对LCZ分类问题的成果。

LCZ通常应用于较粗的空间尺度（分辨率100m、200m），以便能获取城市的空间结构，而这种城市结构很难在像素分辨率很高的情况下获得。

They are generally applied at a coarse spatial scale (typically grids of resolution 100 or 200 m), in order to be able to catch this sense of urban structure that cannot be perceived when working at single-pixel scale at very high resolution.

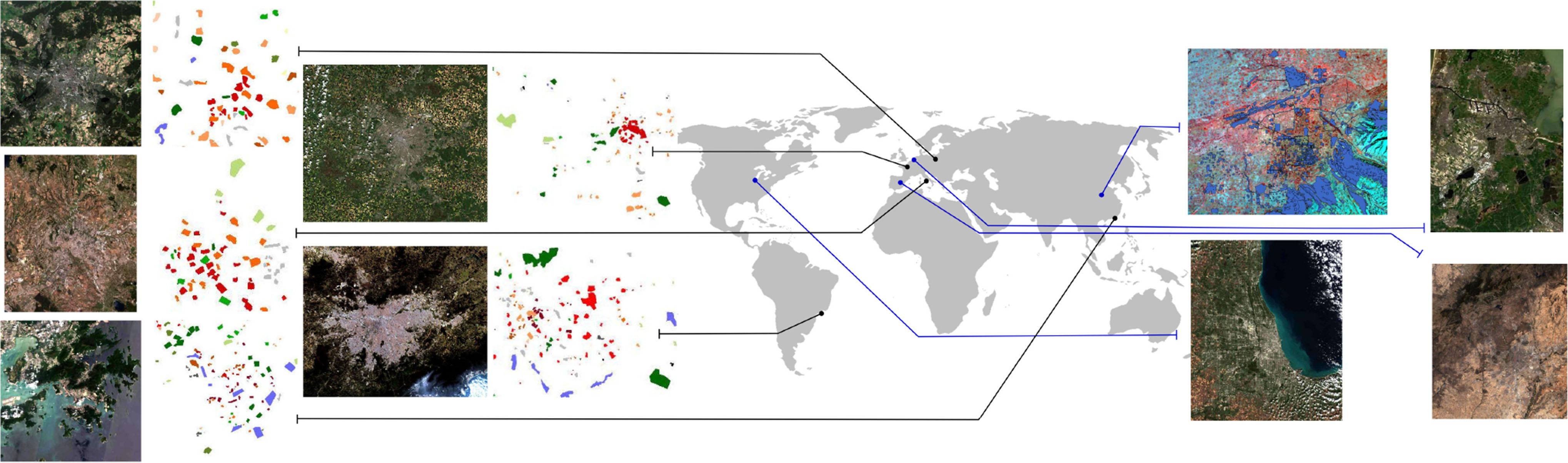
由无任何经验的志愿者提供的LCZ训练样本可能会产生精确度不够高的分类结果。

Moreover, there is evidence that training samples from unexperienced volunteers can result in inaccurate classification results.

**包含数据：**Landsat8多光谱影像、Sentinel-2多光谱影像及OSM数据。

**训练样本城市：**Berlin, Hong Kong, Paris, Rome, and Sao Paulo

**测试样本城市：**Amsterdam, Chicago, Madrid, and Xi’an



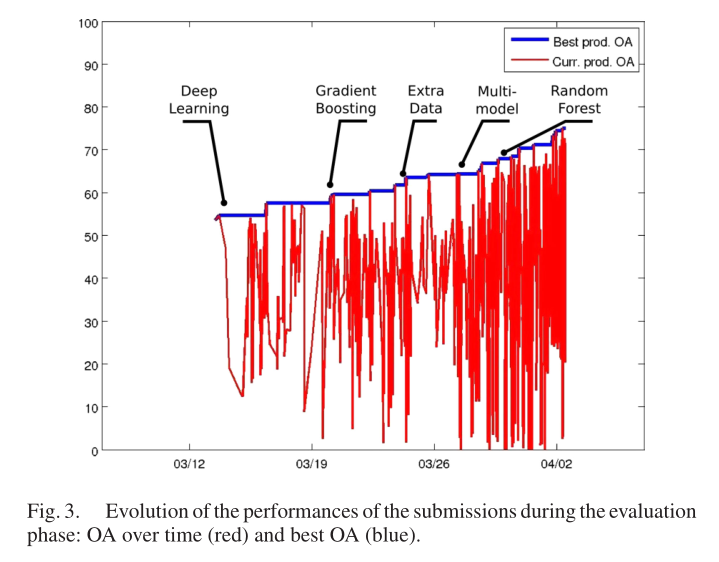
**Problems：**训练数据数量非常不均衡，其中各类数量从323到17716变化不等。

**几种类型的分类方法：**

1. Random Forest type methods
2. Boosting
3. Deep learning
4. Expert handcrafted features

**结果及分析：**





其中采用随机森林方法的几组总体精度较高，而DL方法结果并不理想。

**原因可能是：**

实际上DFC17的数据集中训练样本的数量不少，但也并不十分地多，这可能就是为什么纯粹的DL方法所得结果的排名不高 （DL方法在面对海量训练数据时可以体现出其独特优势）。

Indeed, the number of training samples in the dataset of DFC17 was not small but not extremely large either, which was probably the reason why no pure DL approaches ranked among the winners.

从Fig. 3看出，最开始提交的DL方法在开始时划定了总体精度51.4%的基准线，结果尚可接受；但第一周之后，优势地位被Extreme gradient boost及RFs这些方法占据，结果表明，一旦调整好超参数，这些方法便能够充分利用数据的不均衡性及离散性。

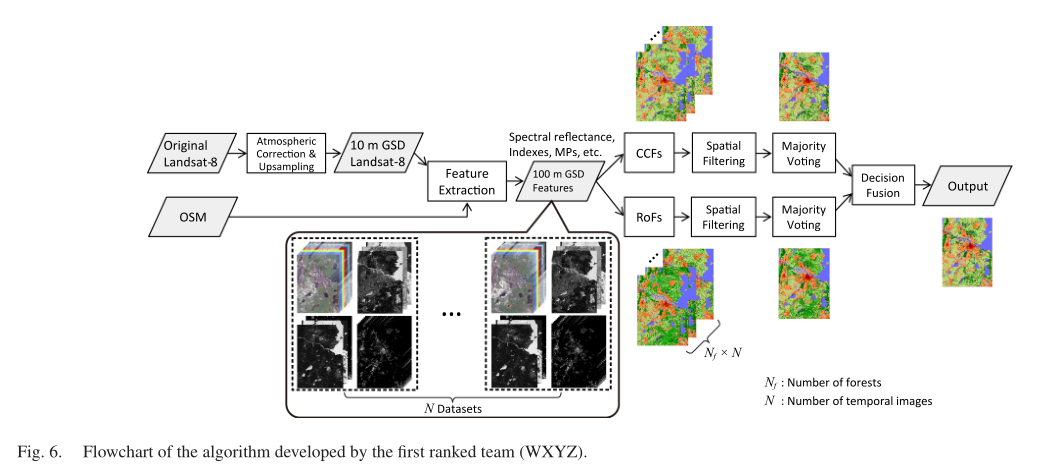
It shows that DL models were ready to use soon after the release of the test data (first submissions were received less than 12h after opening the server) thus establishing an acceptable baseline with 51.4% of OA. After the first week however, the lead was taken on by teams exploiting ensemble methods like extreme gradient boost or RFs, showing that these approaches can take full advantage of imbalanced and sparse data once the right hyperparameters have been found by tuning.

**优胜组方法介绍：**

**第一名组：WXYZ**

**算法：**基于决策树集成分类器(Decision tree ensemble classifier)，即Canonical correlation forests (CCFs)和Rotation forests (RoFs)，利用了从Landsat8和OSM数据中提取的空间与光谱特征。

**框架：**



分为预处理、特征提取、分类和分类后处理几步。

预处理：先对Landsat8影像进行大气校正等处理，并将全色与长波红外波段数据标准化至0-1。然后采用三次插值法对所有波段上采样至10m，以使空间分辨率为100m时特征提取更容易进行。将OSM栅格数据采样至10m。

特征提取：共43个特征类型，包括反射亮度、光学指标、OSM特征和空间特征。具体包括10x10窗口内的平均值与标准差，所有波段共22个特征值；3个光学指标包括NDVI、NDWI、BSI，同样计算三个指标的平均值与标准差，共6个特征值；OSM数据降采样至100m，且分别包含“建筑”、“土地利用”和“水体”3种类型。空间信息则通过从NDVI与OSM的建筑图像中计算形态轮廓 (Morphological profiles, MPs)，结构单元大小分别为3x3、5x5、7x7，共得到12个特征值。

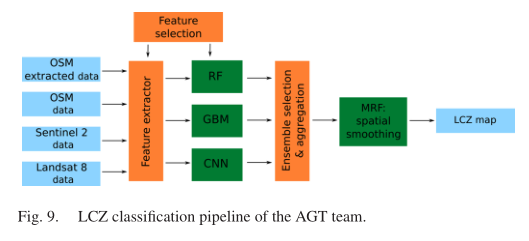
分类：两种方法（RoFs和CCFs）的多数表决结果

分类后处理：对比两种方法对不同地区的分类效果，选择更准确的结果作为最终结果

**第二名组：AGT：**

**算法：**集成了三种分类方法，分别是CNN、RF和Gradient boosting machines (GBM)，其中RF与GBM利用人工采集的特征进行训练，而CNN直接利用原始数据。最后采用马尔可夫随机场模型进行平滑，以增强结果表现。

**框架：**

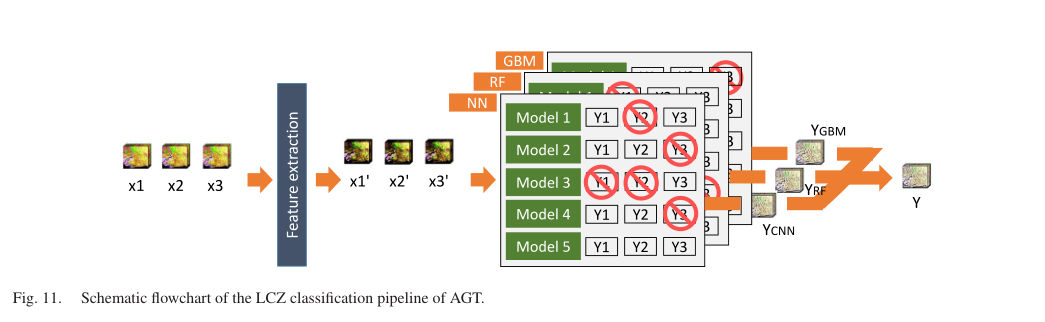


分为数据采集/增强、特征提取与选择、机器学习模型学习与检验、模型继承及分类后处理。

分类时，通过RF模型进行重要特征选择，选择相对重要的特征进行更进一步的模型训练。

为了实现交叉验证，模型将训练数据集分为5个部分，每一部分使用四个城市进行训练，使用第五个城市进行验证。在每次迭代中，每个分类器的模型都通过这两种数据进行训练和验证。

分类器集成分为三部分：the site temporal samples combiner, the first-level model combiner, and the second-level model combiner，如下图，其中x1, x2, x3表示原始多时相多光谱影像，x1’, x2’, x3’表示多时相特征，Y1, Y2, Y3表示单个模型的LCZ预测map（交叉验证时，每个分类器产生了5组模型），Y­CNN, YRF, YXGB表示每个模型产生的LCZ map，Y是最终结果。



三种分类器各有优势，RF is a well-known robust and stable classifier, GBM is the state of the art in many classification tasks based on boosting approaches and is not prone to overfitting, and CNN because they can efficiently capture spatial relationships.

**RF方法改进：**为了改善数据中类别的不均衡性问题，对每个类别添加权重，使数据中占比较低的类权重更高。

**CNN架构：**为防止出现过拟合，只包含三个层。

第一层是一个使用了batch normalization (BN)算法的 1x1卷积过程，选择tanh或ReLU作为激活函数。第二层具有高度参数共享特性，最大限度地降低过拟合现象。第三层是softmax层，其直接连接到卷积层，为所有LCZ分类产生后验概率。

The first layer was a 1 × 1 convolution with batch normalization and a tanh or ReLU nonlinearity. For the second layer, we designed a special type of convolutional kernel with high degree of parameter sharing. The kernel used was symmetric with respect to the origin in order to share more parameters and achieve rotation invariance (as shown in [57]; see Fig. 12). With these kernel topologies, we could achieve the highest performance minimizing the risk of overfitting. The third layer was the softmax layer that was connected directly to the convolutional layer and produced posterior probability estimates for all the LCZ classes.

**马尔可夫随机场平滑处理**

**结论**

在比赛中多组队伍引入了DL方法，但受有限的训练数据及遥感影像的一些独有问题影响。但通过多数据融合的方式，部分数据问题可以解决。