# Planet: Understanding the Amazon from Space, 1st Place Winner’s Interview 星球：从太空了解亚马逊，第一名获奖者访谈

|10.17.2017  
|2017年10月17日

In our recent competition, challenged the Kaggle community to label satellite images from the Amazon basin, in order to better track and understand causes of deforestation.  
在我们最近的竞争中，为了更好地跟踪和了解森林砍伐的原因，Kaggle社区向亚马逊流域的卫星图像贴上标签提出了挑战。

The competition contained over 40,000 training images, each of which could contain multiple labels, generally divided into the following groups:  
比赛共有40000多张训练图片，每张图片都可能有多个标签，通常分为以下几组：

* Atmospheric conditions: clear, partly cloudy, cloudy, and haze  
  大气条件：晴朗，部分多云，多云，霾
* Common land cover and land use types: rainforest, agriculture, rivers, towns/cities, roads, cultivation, and bare ground  
  常见土地覆盖和土地利用类型：雨林、农业、河流、城镇、道路、耕地和裸地
* Rare land cover and land use types: slash and burn, selective logging, blooming, conventional mining, artisanal mining, and blow down  
  稀有土地覆盖和土地利用类型：刀耕火种、选择性伐木、开花、传统采矿、手工采矿和排污



img  
img公司

We recently talked to user , the winner of the competition, to learn how he used an ensemble of 11 finely tuned convolutional nets, models of label correlation structure, and a strong focus on avoiding overfitting, to achieve 1st place.  
我们最近与比赛的获胜者user交谈，了解他如何使用11个精细调谐的卷积网、标签相关结构模型以及避免过度拟合的强大焦点来获得第一名。

img  
img公司

img  
img公司

# Basics 基础

#### What was your background prior to entering this challenge? 在参加这次挑战之前你的背景是什么？

I majored in computer science and have more than 10 years of experience programming in Java and working on large-scale data processing, machine learning, and deep learning.  
我主修计算机科学，有超过10年的Java编程经验，从事大规模数据处理、机器学习和深度学习。

#### Do you have any prior experience or domain knowledge that helped you succeed in this competition? 你是否有任何先前的经验或领域知识帮助你在这场比赛中取得成功？

I entered a few deep learning competitions on Kaggle this year. The experiences and the intuition I gained helped a lot.  
今年我参加了一些关于Kaggle的深度学习比赛。我获得的经验和直觉帮助很大。

#### How did you get started competing on Kaggle? 你是怎么开始和卡格尔比赛的？

I’ve been reading a lot of books and papers about machine learning and deep learning since 2010, but I always found it hard to apply the algorithms I learned on the kinds of small datasets that are usually available. So I found Kaggle a great platform, with all the interesting datasets, kernels, and great discussions. I couldn’t wait to try something, and entered the “Predicting Red Hat Business Value” competition last year.  
自2010年以来，我阅读了很多关于机器学习和深度学习的书籍和论文，但我总是发现很难将我学到的算法应用到通常可用的各种小数据集上。所以我发现Kaggle是一个很棒的平台，有所有有趣的数据集、内核和很棒的讨论。我迫不及待想尝试一下，去年参加了“预测红帽商业价值”比赛。

#### What made you decide to enter this competition? 是什么让你决定参加这次比赛的？

I entered this competition for two reasons.  
我参加这次比赛有两个原因。

First, I’m interested in nature conservation. I think it’s cool to use my skills to make our planet and life better. So I’ve entered all the competitions of this kind that Kaggle has hosted this year. And I’m especially interested in the Amazon rainforest since it appears so often in films and stories.  
首先，我对自然保护感兴趣。我觉得用我的技能让我们的星球和生活变得更好很酷。所以我参加了卡格尔今年举办的所有此类比赛。我对亚马逊雨林特别感兴趣，因为它经常出现在电影和故事中。

Second, I’ve entered all kinds of deep learning competitions on Kaggle using algorithms like segmentation and detection, so I wanted a classification challenge to try something different.  
其次，我参加了各种关于Kaggle的深度学习竞赛，使用了诸如分割和检测之类的算法，所以我想要一个分类挑战来尝试一些不同的东西。

# Let’s Get Technical 让我们从技术上

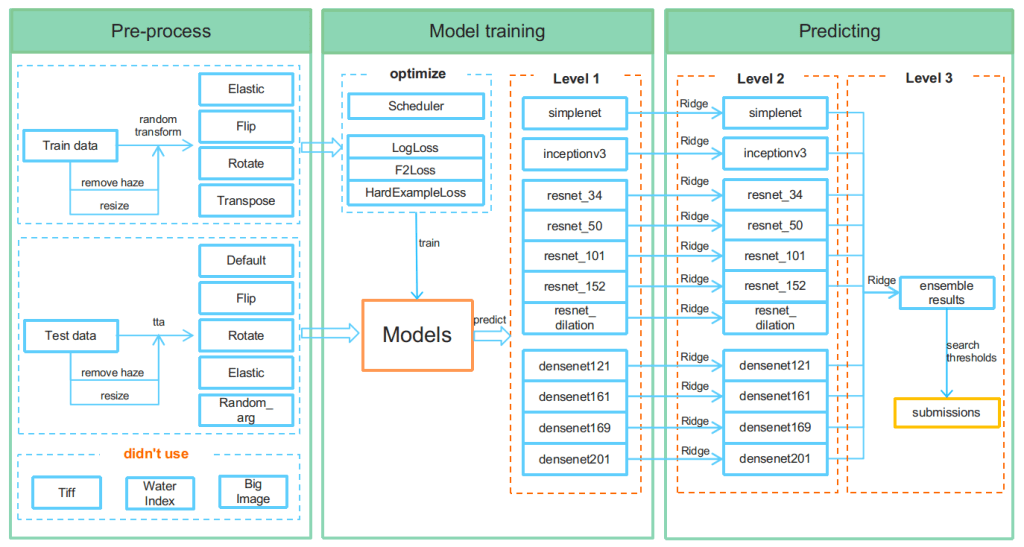
#### Can you introduce your solution briefly first? 你能先简单介绍一下你的解决方案吗？

This is a multi-label classification challenge, and the labels are imbalanced.  
这是一个多标签分类的挑战，标签是不平衡的。

It’s a hard competition, as image classification algorithms have been widely used and built upon in recent years, and there are many experienced computer vision competitors.  
近年来，图像分类算法得到了广泛的应用和发展，有许多经验丰富的计算机视觉竞争对手。

I tried many kinds of popular classification algorithms that I thought might be helpful, and based on careful analysis of label relationships and model capabilities, I was able to build an ensemble method that won 1st place.  
我尝试了许多我认为可能有用的流行分类算法，在仔细分析标签关系和模型功能的基础上，我构建了一个获得第一名的集成方法。

This was my model’s architecture:  
这是我模型的架构：



img  
img公司

In words:  
换句话说：

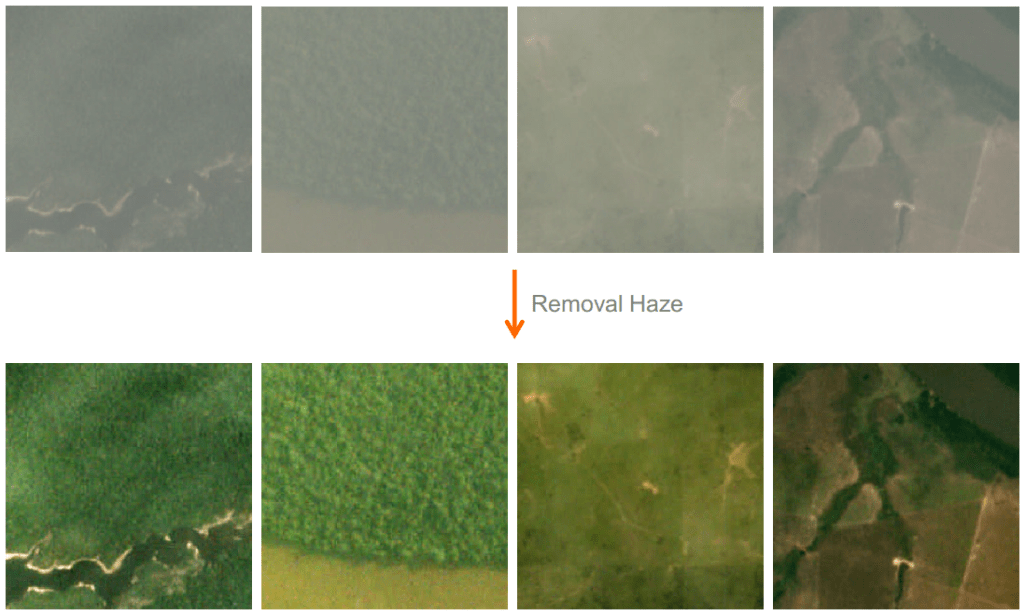
* First, I preprocessed the dataset (by resizing the images and removing haze), and applied several standard data augmentation techniques.  
  首先，我对数据集进行了预处理（通过调整图像大小和消除雾），并应用了几种标准的数据增强技术。
* Next, for my models, I fine-tuned 11 convolutional neural networks (I used a variety of popular, high-performing CNNs like ResNets, DenseNets, Inception, and SimpleNet) to get a set of class label probabilities for each CNN.  
  接下来，对于我的模型，我微调了11个卷积神经网络（我使用了各种流行的、高性能的CNN，如ResNets、DenseNets、Inception和SimpleNet），以获得每个CNN的类标签概率集。
* I then passed each CNN’s class label probabilities through its own ridge regression model, in order to adjust the probabilities to take advantage of label correlations.  
  然后，我通过自己的岭回归模型传递每个CNN的类标签概率，以便调整概率以利用标签相关性。
* Finally, I ensembled all 11 CNNs, by using another ridge regression model.  
  最后，我使用另一个岭回归模型整合了所有11个cnn。
* Also of note is that instead of using a standard log loss as my loss function, I used a special soft F2-loss in order to get a better score on the F2 evaluation metric.  
  另外值得注意的是，我没有使用标准的对数损失作为损失函数，而是使用了一个特殊的软F2损失，以便在F2评估指标上获得更好的分数。

#### What preprocessing and feature engineering did you do? 你做了什么预处理和特征工程？

I used several preprocessing and data augmentation steps.  
我使用了几个预处理和数据扩充步骤。

* First, I resized images.  
  首先，我调整了图片的大小。
* I also added data augmentation by flipping, rotating, transposing, and elastic transforming images in my training and test sets.  
  我还通过翻转、旋转、转置和弹性变换图像来增加数据。
* I also used a haze removal technique, described in this paper, to help my networks “see” the images more clearly.  
  我还使用了本文中描述的烟霾消除技术，帮助我的网络更清晰地“看到”图像。

Here are some examples of haze removal on the dataset:  
以下是数据集上消除雾的一些示例：



img  
img公司

As we can see in the following chart, haze removal improved the F2 score of some labels (e.g., water and bare\_ground), but decreased the F2 score of others (e.g., haze and clear). However, this was fine since ensembling can select the strongest models for each label, and the haze removal trick helped overall.img  
如下图所示，除雾改善了某些标签（如水和裸地）的F2得分，但降低了其他标签（如雾和透明）的F2得分。不过，这很好，因为ensembling可以为每个标签选择最强大的型号，而且消除雾霾的技巧有助于整体效果

#### What supervised learning methods did you use? 你使用了什么有监督的学习方法？

The base of my ensemble consisted of 11 popular convolutional networks: a mixture of ResNets and DenseNets with different numbers of parameters and layers, as well an Inception and SimpleNet model. I fine-tuned all layers of these pre-trained CNNs after replacing the final output layer to meet the competition’s output, and I didn’t freeze any layers.  
我的集成的基础包括11个流行的卷积网络：具有不同数量参数和层的resnet和densenet的混合，以及Inception和SimpleNet模型。在更换了最终输出层后，我对这些经过预训练的cnn的所有层进行了微调，以满足竞争对手的输出，我没有冻结任何层。

The training set consisted of 40,000+ images, so would have been large enough to even train some of these CNN architectures from scratch (e.g., resnet\_34 and resnet\_50), but I found that fine-tuning the weights of the pre-trained network performed a little better.  
训练集由40000多张图片组成，因此足够大，甚至可以从头开始训练一些CNN架构（例如resnet\_34和resnet\_50），但我发现，对预先训练的网络的权重进行微调会更好一些。

#### Did you use any special techniques to model the evaluation metric? 你有没有使用特殊的技术来建立评估指标的模型？

Submissions were evaluated on their F2 score, which is a way of combining precision and recall into a single score – like the F1 score, but with recall weighted higher than precision. Thus, we needed not only to train our models to predict label probabilities, but also had to select optimum thresholds to determine whether or not to select a label given its probability.  
提交的报告是根据他们的F2分数进行评估的，这是一种将精确性和回忆结合成一个单一分数的方法，就像F1分数一样，但回忆权重高于精确性。因此，我们不仅需要训练我们的模型来预测标签概率，而且还必须选择最佳阈值来确定在给定标签概率的情况下是否选择标签。

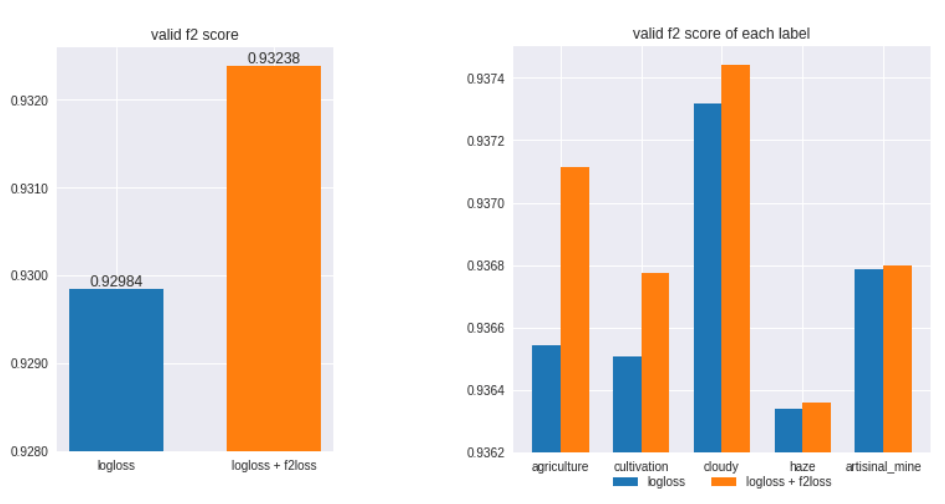
At first, like many other competitors, I used log loss as my loss function. However, as the chart below shows, lower log losses don’t necessarily lead to higher F2 scores.  
一开始，像许多其他竞争对手一样，我使用日志丢失作为我的丢失函数。然而，如下图所示，较低的对数损失不一定会导致较高的F2分数。

img  
img公司

img  
img公司

This means we should find another kind of loss function that allows our models to pay more attention to optimizing each label’s recall. So with the help of code from the forums, I wrote my own Soft F2-Loss function.  
这意味着我们应该找到另一种损失函数，让我们的模型更加注重优化每个标签的召回。所以在论坛代码的帮助下，我编写了自己的软件F2 Loss函数。

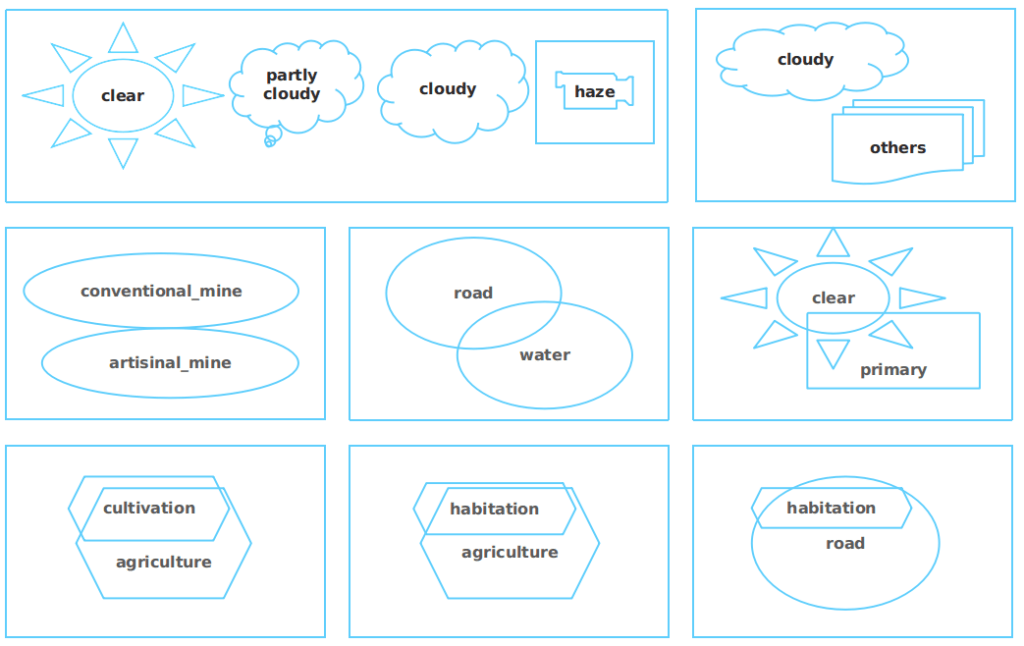
This did indeed improve the overall F2 score, and in particular, the F2 score of labels like agriculture, cloudy, and cultivation.  
这确实提高了F2的总分，特别是像农业、阴天和栽培这样的标签的F2分。



img  
img公司

#### What was your most important insight into the data and models? 你对数据和模型最重要的见解是什么？

I analyzed the correlation between labels, and found that certain labels coexist quite frequently, whereas others do not. For example, the clear, partly cloudy, cloudy, and haze labels are disjoint, but habitation and agriculture labels appear together quite frequently. This meant that making use of this correlation structure would likely improve my model.  
我分析了标签之间的相关性，发现某些标签非常频繁共存，而其他标签则不共存。例如，晴朗、部分多云、多云和霾标签是不相交的，但居住和农业标签一起出现的频率相当高。这意味着利用这种相关性结构可能会改进我的模型。



img  
img公司

For example, let’s take my resnet-101 model. This predicts probabilities for each of the 17 labels. In order to take advantage of label correlations, though, I added another ridge-regularized layer to recalibrate each label probability given all the others.  
例如，让我们以resnet-101模型为例。这可以预测17个标签中每一个的概率。不过，为了利用标签相关性，我添加了另一个脊正则化层，以便在所有其他标签概率的情况下重新校准每个标签概率。

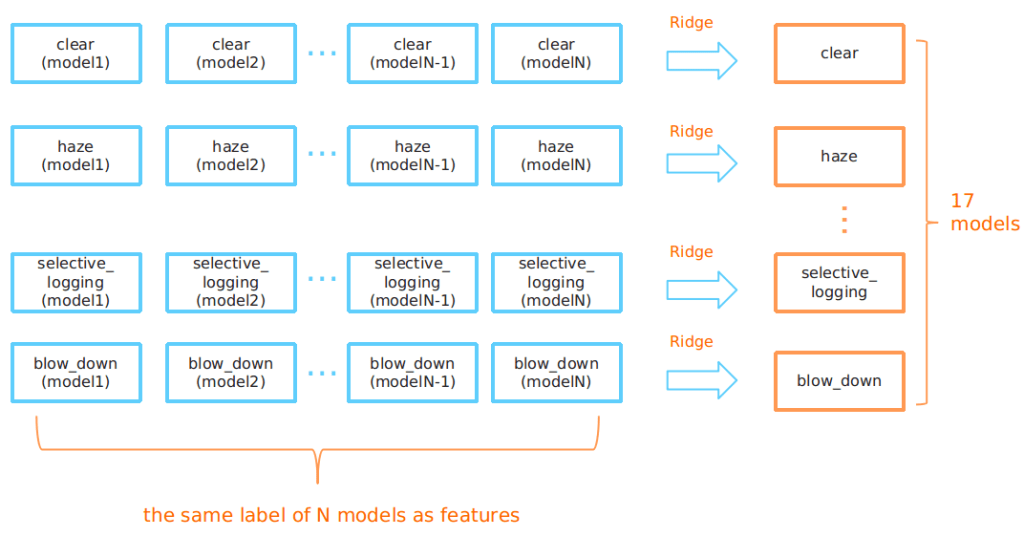
In other words, to predict the final clear probability (from the resnet-101 model alone), I have a specific clear ridge regression model that takes in the resnet-101 model’s predictions of all 17 labels.  
换句话说，为了预测最终的清除概率（仅从resnet-101模型），我有一个特定的清除岭回归模型，它接受resnet-101模型对所有17个标签的预测。

img  
img公司

img  
img公司

#### How did your ensemble your models? 你的模特们是怎么合拍的？

After we get predictions from all N models, we have N probabilities of the clear label from N different models. We can use them to predict the final clear label probability, by using another ridge regression.  
在我们得到所有N个模型的预测后，我们得到N个来自N个不同模型的清晰标记的概率。我们可以使用它们来预测最终的清晰标记概率，通过使用另一个岭回归。



img  
img公司

This kind of two-level ridge regression does two things:  
这种两级岭回归有两个作用：

1. First, it allows us to use the correlation information among the different labels.  
   首先，它允许我们使用不同标签之间的相关信息。
2. It allows us to select the strongest models to predict each label.  
   它允许我们选择最强的模型来预测每个标签。

#### Were you surprised by any of your findings? 你对你的发现感到惊讶吗？

Even though I’d predicted the final shakeup of the leaderboard (where the public and private leaderboard scores differed quite a bit), I was still surprised.  
尽管我已经预测了排行榜的最终调整（公共和私人排行榜的得分相差很大），但我仍然感到惊讶。

Essentially, at the last stage of the competition (10 days before the end), I found that the public scores were very close, and I couldn’t improve my local cross-validation or public scores any more. So I warned myself to be careful to avoid overfitting on what could just be label noise.  
基本上，在比赛的最后阶段（比赛结束前10天），我发现公众分数非常接近，我无法再提高我的本地交叉验证或公众分数。所以我提醒自己要小心，避免过度贴标签噪音。

To understand this pitfall better, I simulated the division into public and private leaderboards by using different random seeds to select half of the training set images as new training sets. I found that as the seed changed, the difference between my simulated public and private scores could grow up to 0.0025. But the gap between the Top 1 and Top 10 entries on the public leaderboard was smaller than this value.  
为了更好地理解这个陷阱，我用不同的随机种子选择一半的训练集图像作为新的训练集，模拟了公共和私人排行榜的划分。我发现，随着种子的变化，我的模拟公共和私人分数之间的差异可能会增长到0.0025。但在公共排行榜前1名和前10名之间的差距小于这个值。

img  
img公司

img  
img公司

This meant that a big shakeup could very likely happen in the real competition as well.  
这意味着，在真正的竞争中，也很可能发生一场大的变革。

After carefully analyzing, I found that this kind of variation arose with difficult images where labels were prone to confusion from humans as well, like whether an image should be labeled haze vs. cloudy, road vs. water, or blooming vs. selective logging.  
经过仔细分析，我发现这种变化是由一些难以理解的图像引起的，这些图像的标签也容易与人类混淆，比如一幅图像是否应该被标记为霾与多云、道路与水，或者开花与选择性采伐。

Because of this, I persuaded myself that the public leaderboard scores weren’t a perfect metric of model capability. This was unexpected: since the public test set contains 40,000+ images, it seems like the leaderboard should be pretty stable.  
正因为如此，我说服自己，公众排行榜的分数并不是衡量模特能力的完美标准。这是出乎意料的：因为公共测试集包含40000多个图像，所以排行榜应该相当稳定。

So I adjusted my goal to simply keep myself in the top 10, and decided not to care about my exact position on the public leaderboard in the last week. Instead, I tried to find the most stable way to ensemble my models, I threw away any models that would likely lead to overfitting, and in the end I used voting and ridge regression.  
所以我调整了自己的目标，仅仅是保持自己在前10名，并决定不关心我在公众排行榜上的确切位置在上周。相反，我试图找到最稳定的方法来整合我的模型，我抛弃了任何可能导致过度拟合的模型，最后我使用了投票和岭回归。

#### Why so many models? 为什么有这么多模特？

The answer is simple: diversity.  
答案很简单：多样性。

I don’t think the number of models is a big problem, for several reasons:  
我不认为模特的数量是个大问题，原因有几个：

1. First, if we want a simple model, we can simply choose 1-2 of them, and it will still get a decent score on both the public and private leaderboards (top 20).  
   首先，如果我们想要一个简单的模型，我们可以简单地选择其中的1-2个，它仍然会在公共和私人排行榜（前20名）上获得不错的分数。
2. Second, we have 17 labels, and different models have different capabilities on each label.  
   其次，我们有17个标签，不同的型号在每个标签上有不同的功能。
3. Third, our solution will be used to replace or simplify the human labeling job. Since computational resources are relatively cheaper than humans, we can predict unlabeled images by using strong models, modify any incorrectly predicted images, and then use the expanded data set to train stronger or simpler models iteratively.  
   第三，我们的解决方案将被用来取代或简化人类的标签工作。由于计算资源比人类相对便宜，我们可以通过使用强模型来预测未标记的图像，修改任何错误预测的图像，然后使用扩展的数据集迭代地训练更强或更简单的模型。

#### What tools did you use? 你用了什么工具？

Python 3.6, PyTorch, PyCharm community version.  
Python 3.6，PyTorch，PyCharm社区版。

#### What does your hardware setup look like? 你的硬件设置是什么样子的？

A server with four NVIDIA GTX TITAN X Maxwell GPUs.  
一个有四个NVIDIA GTX TITAN X Maxwell GPU的服务器。

# Words of wisdom 智慧的言语

#### What have you taken away from this competition? 你从这次比赛中得到了什么？

As we discussed above, I found that using a soft F2-loss function, adding a haze-removal algorithm, and applying two-level ridge regression were important in achieving good scores.  
如上所述，我发现使用软F2损失函数、添加烟霾去除算法和应用两级岭回归对获得好的分数很重要。

Also, due to label noise, we must trust our local cross-validation.  
此外，由于标签噪声，我们必须信任我们的本地交叉验证。

#### Do you have any advice for those just getting started in data science? 你对那些刚开始学数据科学的人有什么建议吗？

1. Learn from good courses like Stanford’s CS229 and CS231n.  
   从斯坦福大学的CS229和CS231n这样的好课程中学习。
2. Learn from Kaggle competitions, kernels, and starter scripts.  
   从Kaggle竞赛、内核和启动脚本中学习。
3. Enter Kaggle competitions and use them to get feedback.  
   网络错误
4. Read papers everyday and implement some of them.  
   网络错误

Planet：从太空中了解亚马逊，获得第一名获奖者Edwin Chen | 10.17.2017在我们最近的行星：从太空竞赛中了解亚马逊，Planet挑战Kaggle社区标记亚马逊流域的卫星图像，以便更好地跟踪和了解砍伐森林的原因。 比赛包含40,000多张训练图像，每张图像可以包含多个标签，一般分为以下几组：大气条件：晴朗，晴天，阴天和阴霾共同的土地覆盖和土地利用类型：雨林，农业，河流，城镇/城市，道路，耕种和裸地稀有土地覆盖和土地使用类型：刀耕火种，选择性采伐，开花，常规采矿，手工采矿和排污我们最近采访了比赛的获胜者，了解他如何使用11个精细调整卷积网的集合，标签相关结构模型，以及强调避免过度拟合，以获得第一名。 基础知识在进入此挑战之前，您的背景是什么？ 我主修计算机科学，拥有超过10年的Java编程经验，从事大规模数据处理，机器学习和深度学习。 您是否有任何先前的经验或领域知识帮助您在本次比赛中取得成功？ 今年我参加了一些关于Kaggle的深度学习比赛。我获得的经验和直觉帮助了很多。 你是如何开始参加Kaggle的比赛的？ 自2010年以来，我一直在阅读大量关于机器学习和深度学习的书籍和论文，但我总是发现很难将我学到的算法应用于通常可用的小型数据集。因此，我发现Kaggle是一个很棒的平台，包含所有有趣的数据集，内核和精彩的讨论。我迫不及待地尝试了一些东西，去年参加了“预测红帽商业价值”竞赛。 是什么让你决定参加这个比赛？ 我参加本次比赛有两个原因。 首先，我对自然保护感兴趣。我认为用我的技能让我们的星球和生活变得更好是很酷的。所以我参加了Kaggle今年举办的所有比赛。我对亚马逊热带雨林特别感兴趣，因为它经常出现在电影和故事中。 其次，我已经使用分段和检测等算法参加了Kaggle的各种深度学习竞赛，所以我想要一个分类挑战来尝试不同的东西。 让我们来技术您可以先简要介绍一下您的解决方案吗？ 这是一个多标签分类挑战，标签是不平衡的。 这是一场艰难的竞争，因为近年来图像分类算法已被广泛使用和建立，并且有许多经验丰富的计算机视觉竞争者。 我尝试了许多我认为可能有用的流行分类算法，并且基于对标签关系和模型功能的仔细分析，我能够构建一个赢得第一名的集合方法。 这是我的模型的架构：在文字中：首先，我预处理数据集（通过调整图像大小和消除雾霾），并应用了几种标准数据增强技术。 接下来，对于我的模型，我微调了11个卷积神经网络（我使用各种流行的，高性能的CNN，如ResNets，DenseNets，Inception和SimpleNet）来获得每个CNN的一组类标签概率。 然后，我通过自己的岭回归模型传递了每个CNN的类标签概率，以便调整概率以利用标签相关性。 最后，通过使用另一个岭回归模型，我合并了所有11个CNN。 另外值得注意的是，我没有使用标准的日志丢失作为我的损失函数，而是使用了特殊的软F2丢失，以便在F2评估指标上获得更好的分数。 您做了哪些预处理和功能工程？ 我使用了几个预处理和数据增强步骤。 首先，我调整了图像大小。 我还通过在训练和测试集中翻转，旋转，移调和弹性转换图像来添加数据。 我还使用了一种雾霾去除技术，在“使用暗通道先前单张图像雾霾去除”中描述，以帮助我的网络更清晰地“看到”图像。 以下是数据集中雾霾去除的一些示例：如下图所示，雾霾去除改善了某些标签（例如水和裸地）的F2分数，但降低了其他标签的F2分数（例如，雾度和清晰度） ）。然而，这很好，因为集合可以为每个标签选择最强的模型，并且雾霾去除技巧有助于整体。 你使用了哪些监督学习方法？ 我的整体基础由11个流行的卷积网络组成：ResNets和DenseNets的混合，具有不同数量的参数和层，以及Inception和SimpleNet模型。在更换最终输出层以满足比赛的输出后，我对这些预训练的CNN的所有层进行了微调，并且我没有冻结任何层。 训练集由40,000多个图像组成，因此可能足够大，甚至可以从头开始训练一些CNN架构（例如，resnet\_34和resnet\_50），但我发现微调预训练网络的权重进行了好一点。 您是否使用任何特殊技术来建模评估指标？ 提交的评分是根据他们的F2评分，这是一种将精确度和召回率组合成单一评分的方法 - 如F1评分，但回忆加权高于精确度。因此，我们不仅需要训练我们的模型来预测标签概率，而且还必须选择最佳阈值来确定是否在给定其概率的情况下选择标签。 起初，和许多其他竞争对手一样，我使用日志丢失作为我的损失函数。但是，如下图所示，较低的对数损失不一定会导致较高的F2分数。 这意味着我们应该找到另一种损失函数，它允许我们的模型更加注重优化每个标签的召回。所以在论坛代码的帮助下，我编写了自己的Soft F2-Loss功能。 这确实改善了整体F2得分，特别是农业，阴天和栽培等标签的F2得分。 您对数据和模型最重要的见解是什么？ 我分析了标签之间的相关性，发现某些标签经常共存，而其他标签则没有。例如，透明的，部分混浊的，混浊的和阴霾的标签是不相交的，但是居住和农业标签经常出现在一起。这意味着利用这种相关结构可能会改善我的模型。 例如，让我们采用我的resnet-101模型。这预测了17个标签中每个标签的概率。然而，为了利用标签相关性，我添加了另一个脊形正则化层来重新校准给定所有其他标签的每个标签概率。 换句话说，为了预测最终的清晰概率（仅来自resnet-101模型），我有一个特定的清晰岭回归模型，它接受了所有17个标签的resnet-101模型的预测。 你的模特你的模特怎么样？ 在我们从所有N个模型得到预测之后，我们有来自N个不同模型的清晰标签的N个概率。我们可以使用它们来预测最终的明确标签概率，使用另一个岭回归。 这种两级岭回归做了两件事：首先，它允许我们使用不同标签之间的相关信息。 它允许我们选择最强的模型来预测每个标签。 你的任何发现都让你感到惊讶吗？ 即使我预测排行榜的最终重组（公共和私人排行榜得分相差很多），我仍然感到惊讶。 基本上，在比赛的最后阶段（结束前10天），我发现公共分数非常接近，我无法再提高我当地的交叉验证或公共分数。所以我警告自己要小心避免过度拟合标签噪音。 为了更好地理解这个陷阱，我通过使用不同的随机种子来选择一半的训练集图像作为新的训练集来模拟公共和私人排行榜的划分。我发现随着种子的变化，我模拟的公共和私人分数之间的差异可能会增长到0.0025。但公共排行榜的前1名和前10名参赛者之间的差距小于这个值。 这意味着在真正的竞争中很可能会发生重大变革。 在仔细分析之后，我发现这种变化出现了困难的图像，其中标签也容易与人类混淆，例如图像是否应该标记为混浊与阴天，道路与水，或开花与选择性伐木。 因此，我说服公众排行榜得分并不是模型能力的完美衡量标准。这是出乎意料的：因为公共测试集包含40,000多个图像，所以排行榜看起来应该非常稳定。 所以我调整了我的目标，只是让自己进入前10名，并决定不关心我在上周的公共排行榜上的确切位置。相反，我试图找到最稳定的方式来合奏我的模型，我扔掉了任何可能导致过度拟合的模型，最后我使用了投票和岭回归。 为什么这么多型号？ 答案很简单：多样性。 我不认为模型的数量是一个大问题，原因如下：首先，如果我们想要一个简单的模型，我们可以简单地选择其中的1-2个，并且它仍然会在公众和私人排行榜（前20名）。 其次，我们有17个标签，不同的模型在每个标签上都有不同的功能。 第三，我们的解决方案将用于替换或简化人类标签工作。由于计算资源比人类便宜，我们可以通过使用强模型来预测未标记的图像，修改任何错误预测的图像，然后使用扩展的数据集来迭代地训练更强或更简单的模型。 你用了什么工具？ Python 3.6，PyTorch，PyCharm社区版。 您的硬件设置是什么样的？ 带有四个NVIDIA GTX TITAN X Maxwell GPU的服务器。 智慧之言你从这次比赛中拿走了什么？ 正如我们上面所讨论的，我发现使用软F2损失函数，添加雾霾去除算法，并应用两级岭回归对于获得良好的分数非常重要。 此外，由于标签噪音，我们必须相信我们的本地交叉验证。 对于那些刚开始使用数据科学的人，您有什么建议吗？ 学习斯坦福大学CS229和CS231n等优秀课程。 学习Kaggle比赛，内核和入门脚本。 进入Kaggle比赛并使用它们获得反馈。 每天阅读论文并实施其中一些。  
网络错误