# 皮革纹理匹配

## 项目需求

皮革厂生产大概2000种纹理的皮革，这些皮革种类都有样品和编号。

客户上门时，会带着他们的所需的皮革样品。这个时候，厂家需要将客户的样品跟库里的样品的纹理进行匹配，找出近似度最高的6种皮革，显示其编号及样品图。目前这个匹配的过程完全是靠人工，工作量很大，并且有些纹理非常接近，找出最近似的6种皮革也很困难。

因此，厂家要求开发一款产品，替代这个人工过程。有时候厂家的业务员会上门到客户那里去，这个时候客户的样品是在客户那里的，业务员通过产品可以拍摄样品图片或录制样品视频来进行匹配。

*（备注：产品仅需匹配皮革的纹理，不需要考虑颜色，要求越方便越准确越好。）*

## 方案设计

通过深度学习构建一个分类器模型，输入待匹配的样品图片或视频，给出置信度最高的前6个种类的编号。

* **数据集**：通过实拍+网络上收集一定种类的皮革纹理图片。由于有的皮革纹理较细，有的纹理较粗，在进行分辨或训练时需要采用不同的比例尺，使得局部纹理与整体的纹理具有相似性。  
  收集到的皮革通常一张就是一个种类，有的可能是多张皮革的纹理相同或相似，仅仅是颜色不同。因此在处理数据集时，先是得到一整张皮革大分辨率的图片，之后根据纹理的粗细程度选择合适的比例尺，截取出若干分辨率较低的局部图片组成最终的数据集。
* **训练模型**：由于需求较简单，训练数据也不需要太复杂，可对ImageNet相关的深度神经网络进行迁移学习来实现主要目标。

**产品初步流程为**：拍摄/接收用户提供的样品图片/视频、通过模型得到置信度最高的6个种类的编号、展示编号及样品图。

**优化**：

1. 考虑到提供给模型的图片可能存在模糊、高曝光、过暗等问题，会影响到最终的结果，因此可对传入的图片进行初步判断。
2. 需要对数据集进行预处理：彩色图转灰度图、统一尺寸。在进行模型推理时，也需要对输入图片进行同样的处理。

## 具体实施方案

### 数据集的收集与预处理

数据集决定了最终模型的上限，需要尽可能清晰、平滑、光照均匀。先收集了40个类的训练数据，按纹理粗细初步分为3个规格：粗纹理、中等纹理、细纹理，在截取局部图片时使用不同的比例尺，即截取图片的尺寸不同，分别为：800\*800、600\*600、300\*300。

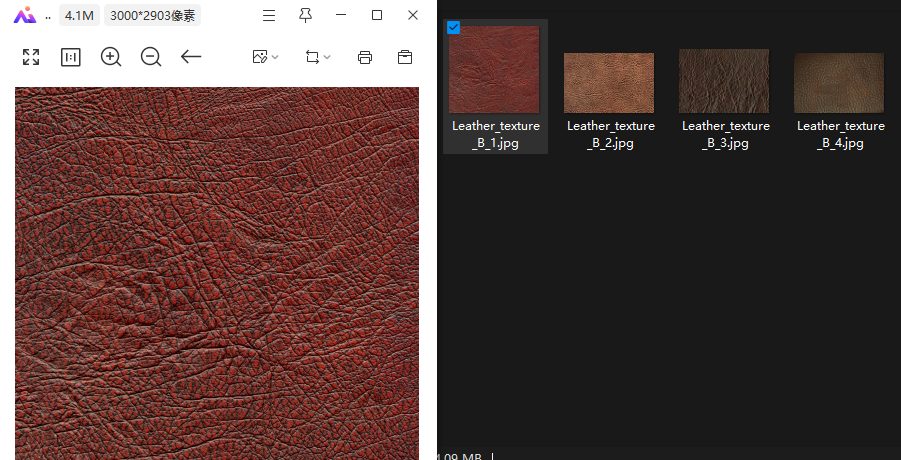


图 1-1 粗纹理的数据集

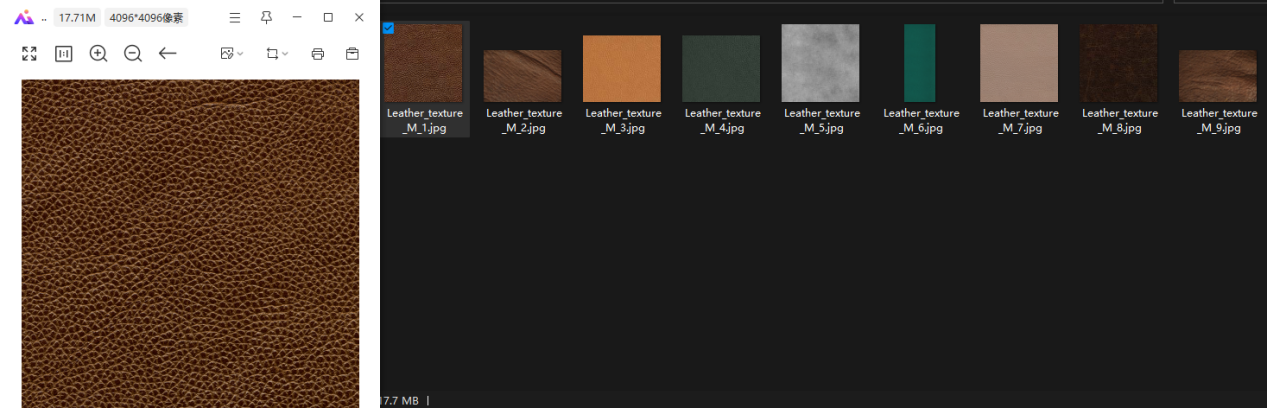


图 1-2 中等纹理的数据集

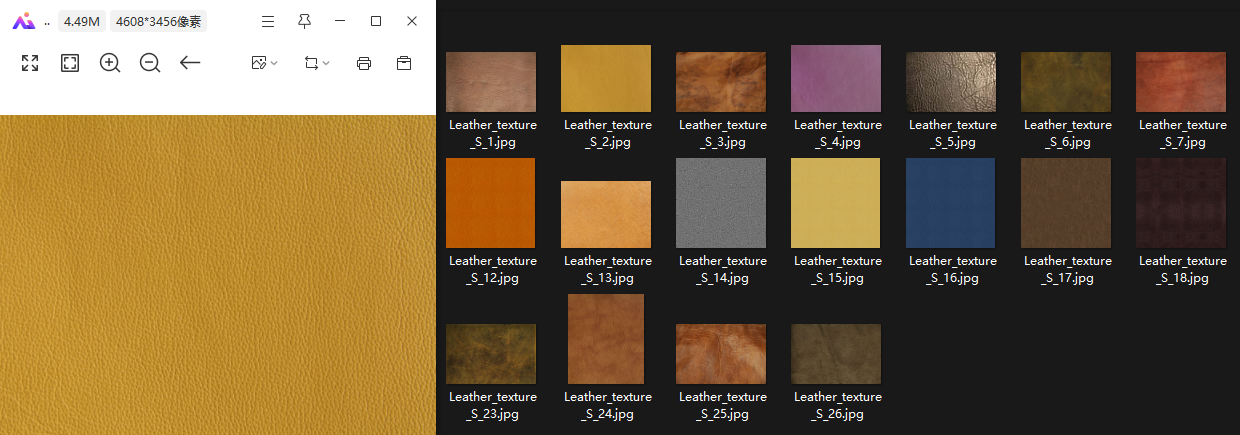


图 1-3 细纹理的数据集

最后统一处理成灰度图，并将尺寸设置为224\*224，与ImageNet数据集的尺寸一样，在提高模型质量上可能存在帮助。

每个类准备30张图片，按训练集与测试集比例为7:3来划分数据集，即每个类训练集大致有19个数据，测试集大致有11个数据。在尝试每个种类不同数量的数据进行训练后，每个类有30张数据是模型质量较好。数据数量过少容易使某个类的特征占比过大，数量过多容易导致过拟合。

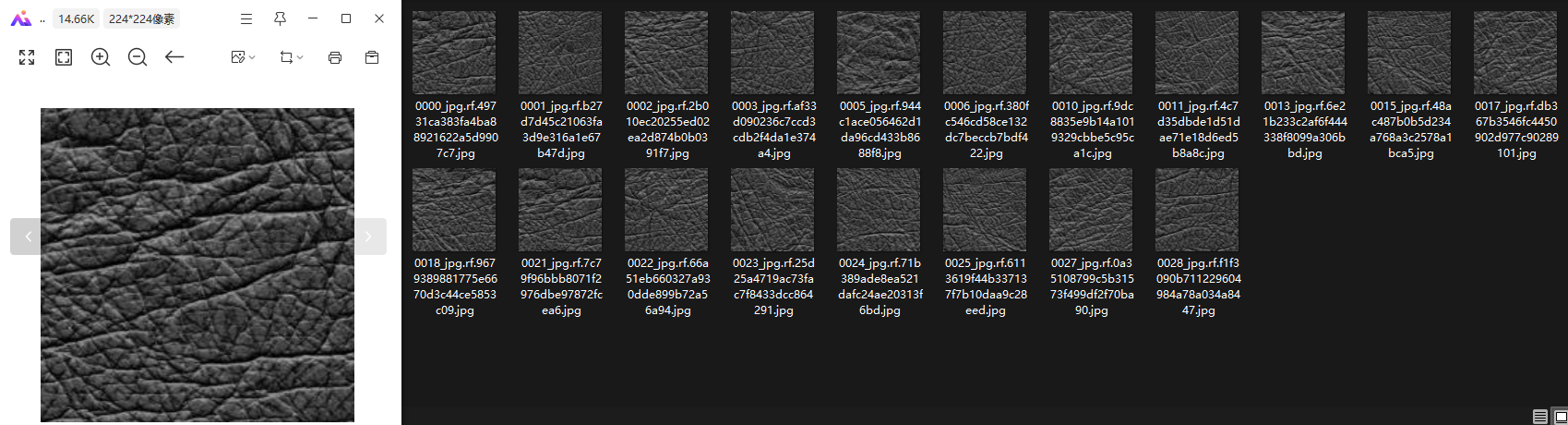


图 1-4 预处理后每个类的训练集数据

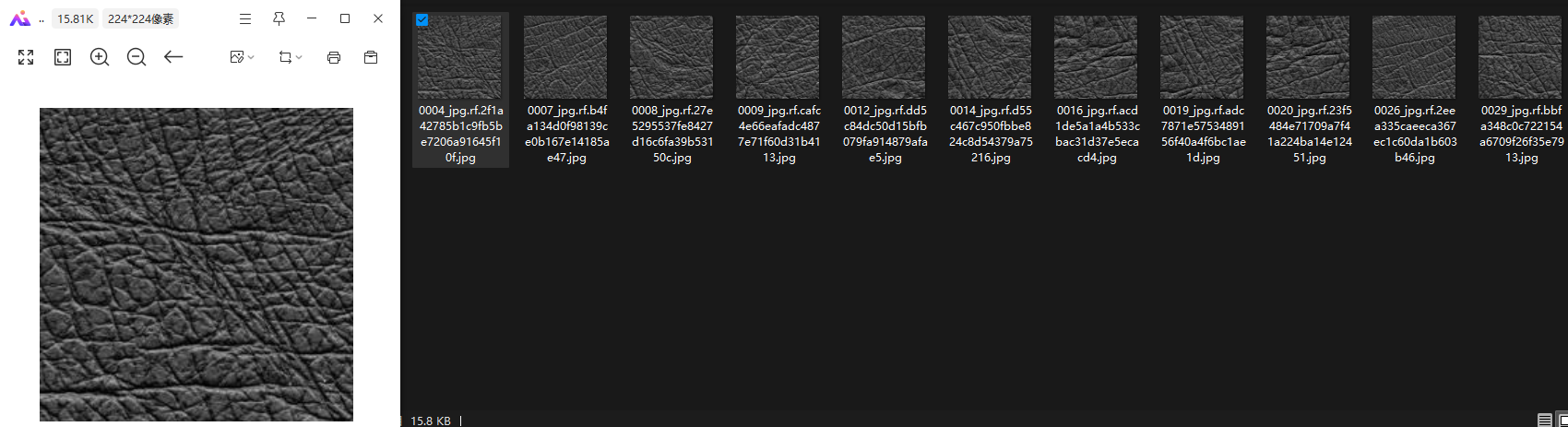


图 1-5 预处理后每个类的测试集数据

后续有了客户数据，可使用该预训练模型再进行迁移学习。或者保存每个类的特征向量，后续添加新的训练数据时，直接通过模型得到新的特征向量，在推理阶段只要推理出输入图片的特征向量，然后找出最相似的6个特征向量即可。

### 训练模型

训练模型采用resnet18，其拥有较高的训练速度、对本项目更加适中的模型复杂度。过于复杂的模型容易导致过拟合，例如resnet50、googlenet等。而vgg16\_bn等模型的训练时间相对较长，最终的模型质量与resnet18的差不多。因此采用resnet18模型来进行迁移学习。

### 优化

为了提高匹配的准确性，对输入图片质量进行评估判断，过滤掉质量较差的图片。

可以采用opencv对图片各种指标进行计算，自定义标准来进行评估，也可以训练一个模型来进行评估。由于各指标的计算及标准不好定，故本项目通过一个模型来进行评估，同样使用resnet18进行迁移学习得到模型。  
理论上可以按清晰而光照均匀的、模糊的、对比度过高的等标准对图片进行分类，组成数据集。但对于实际应用环境来说，过高的标准可能导致输入图片被判定为不合格的概率提高，需要用户重新拍摄输入的次数增加。故此预训练模型仅先针对模糊的图片做限制，因为图片能否正确表达出纹理的重要性是最高的，其他类似光照不均匀、有阴影遮挡等影响图片质量的因素，对于模型推理结果的干扰程度是低于模糊这一因素对推理结果的干扰程度的。该判断模型只是辅助人工进行判断，提高输入图片质量的下限，在进行人工判断与输入时就需尽可能提高输入图片的质量。

对该判断图片质量模型的数据集进行同匹配皮革纹理模型的数据集的预处理：转灰度图、统一尺寸为224\*224。其中从网上收集的图片通过高斯模糊等方法得到对应的模糊图片，每个种类取出少量且数量相同的图片共同组成新的数据集。数据集中目前包含两个类：质量较好的、模糊的，后续可视情况优化升级。

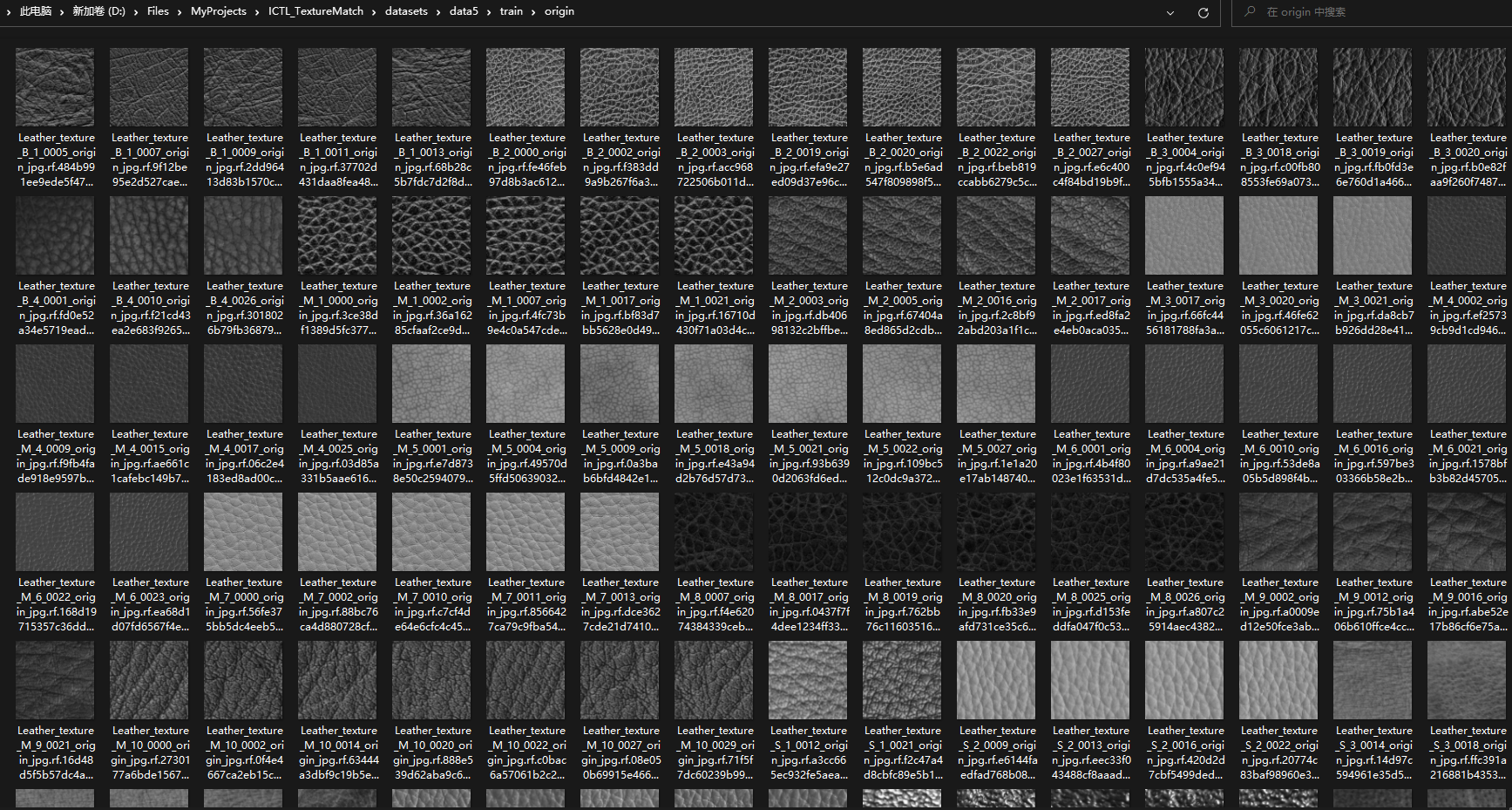


图 6 质量较高的图片数据集

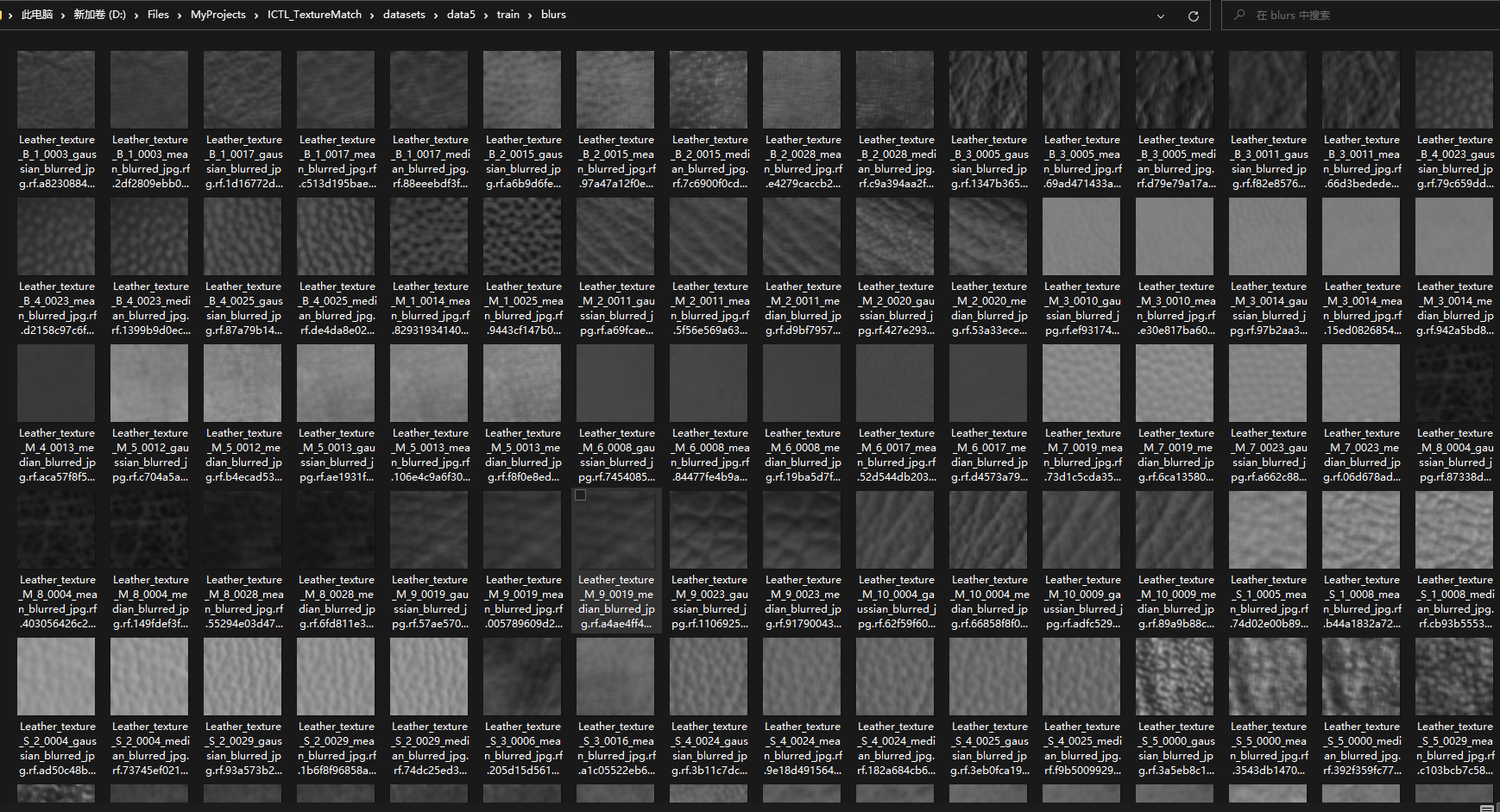


图 7 模糊的图片数据集

### 部署

由于训练阶段使用的是PyTorch，得到的模型文件为pth文件。通过Python转换成ONNX文件，之后编写C++代码在Linux平台生成动态库，通过API接口来接收图片数据并使用ONNX文件进行推理后返回置信度最高的6个类的编号。后续更新升级主要是更新ONNX文件及各种类编号对应的样品图片集。