

**研究生课程考核试卷**

**（适用于课程论文、提交报告）**

**科 目： 人工智能与知识工程 教 师： 叶俊勇**

**姓 名： 王涛 学 号： 20190802025t**

**专 业： 仪器科学与技术 类 别： 学术**

**上课时间： 2020年4月至2020年6月**

**考 生 成 绩：**

|  |  |  |
| --- | --- | --- |
| **卷面成绩** | **平时成绩** | **课程综合成绩** |
|  |  |  |

**阅卷评语：**

**阅卷教师 (签名)**

重庆大学研究生院制

**深度学习在病理图像癌症分类中的应用**

摘要：随着人工智能智能技术的不断发展，其在各个领域中的应用不断拓展加深，深度学习在众多领域越来越展现与发挥出巨大的作用和潜力。其中，将深度学习应用于基于影像组学的癌症分类是智能医疗中的重要领域之一。将深度学习应用于医疗，不仅能极大辅助医生的诊断工作，同时对于提高医疗质量和患者体验有着重要意义。本文介绍了深度学习在基于病理图像的癌症诊断中的重要应用。在一个完整的癌症分类系统中，首先对数据集进行预处理，剔除重复图像，将数据集提取为适宜深度学习模型使用的数据集。然后利用数据集对深度学习模型进行训练。然后利用测试集对模型进行验证和评价，对模型进行修正改进，最后将其应用于病理图像中对癌症进行分类预测。

关键字：病理组学、神经网络、深度学习、乳腺癌

**1.引言**

随着人们生活水平的不断提高，人们对健康的要求和重视也在不断提高。更好的医疗条件对居民而言具有重要意义。但是受限于医疗技术的发展，现有的医疗条件往往不能满足患者的需求。特别是在我国，由于人口基数大以及经济发展阶段限制的原因，导致医生和患者的比例严重失衡。医疗资源的分布也并不平衡，往往集中在大城市以及大医院中。基层一线的医生往往承担着巨大的工作量，就影像科医生而言，不包括其他诊断，但就因为人群中的癌症发病人数一直居高不下，使得影像科的医生一致超负荷运转。因此计算机辅助诊断对于提高医疗质量和减轻医生负担具有重要意义。

**2.利用深度学习进行癌症诊断概述**

相比于人类诊断，机器对于繁重的任务不会表现出疲劳，且能从不断增加的病理数中进行学习，进一步提高诊断准确率。近年来，随着人工智能技术的发展，越来越多的科学家开始研究将人工智能技术，特别是深度学习技术应用于医疗诊断当中。已经有机器模型应用在癌症的分类诊断中，有些AI模型的分类水平甚至超过了职业医师的水平。基于影像组学，利用深度学习进行诊断的任务多种多样，本文主要介绍其在病理组学中进行癌症分类的应用。

在医疗诊断中，病理是诊断、预后和指导治疗的金标准，病理医生也被称为是“医生的医生”。在一个典型的诊断流程中，CT、MRI可以筛查病灶的位置，进行基本定性。如果要确诊，实际上要取出一些组织采样，然后制成切片，通过显微镜下观察组织和细胞结构的变化来对癌症进行定性、分级、分类和分型，以此为后续的手术和化疗、放疗等治疗方案提供指导。因此病理是诊疗中非常重要的环节。现在我国病理医生极其缺乏，仅满足需求的15%。而且，病理诊断实际上是一个非常复杂的过程，主观性很强，一致率比较低。因此利用人工智能技术，可以帮助医生提高诊断的可重复性、准确率、效率，从一定程度上缓解病理医生不足的状况。

近年来，人工智能尤其是深度学习技术，在自然图像的识别和理解上取得了很大成功，但是自然图像和病理图像有很大的区别。自然图像的特点是数据量大、获取容易，手机就可以获取自然图像，普通人可以标注，不需要专业训练，很容易可以获得很多标注数据。自然图像的尺寸一般很小，计算机处理起来比较容易。另外，待识别的物体在图像中的面积比较大。

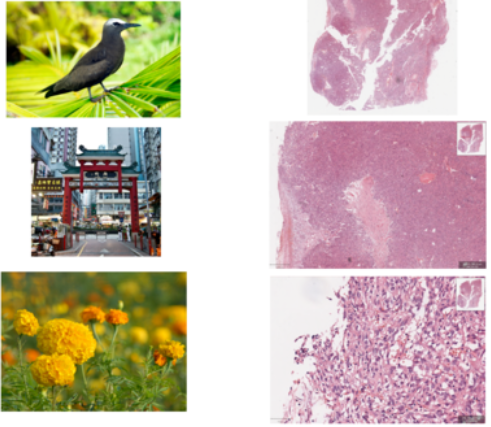


图2.1 自然图像与病理图像

图2.1是一张病理图像在不同放大倍数下的图像，第一张是全局，是一个采样的轮廓。第二张放大4倍以后可以看到一些纹路。第三张放大40倍以后可以看到一些细胞和组织的细微结构。病理图像的特点是数据量比较少，获取成本比较高，需要昂贵的专门的扫描仪把切片扫描成图像，扫描一张高分辨率图像一般需要20-30分钟。病理数据需要专家和医生标注，一张高分辨率的图像有几万乘几万，也就是上亿的像素，因此计算机处理起来有很大挑战。

**2．深度学习算法**

近年来，一系列机器学习的经典算法在AI医疗领域得到应用。但就目前而言，应用较多的方法主要集中在深度学习算法上。虽然已经有很多文献都论证各种方法的优缺点以及提出了相应的解决办法。但是大多都是从理论出发，而且大部分是在对数据分布、分类任务的特征等各种各样的假设前提之下。在现实中，这些假设常常难以满足。因此，对给定的分类任务而言，首先是选择适当的模型，再通过实验比较这些算法。这里我们简单介绍四种广泛使用的深度学习算法：卷积神经网络（CNNs），深度置信网络（RBM），栈式自编码网络和递归神经网络。

2.1 卷积神经网络

卷积神经网络（Convolution NerualNet，CNNs）通过结合局部感知区域、共享权重、空间或时间上的池化降采样三大特点来充分利用数据本身包含的局部性等特征，优化网络结构，并且保证一定程度上的位移的不变性。它的权值共享网结构相似于生物神经网络，降低了网络模型的复杂度，减少了权值的数量。因为这样的结构特点使其尤其适合大图像数据的机器学习，可以使数量庞大的图像识别问题不断降维。CNNs的主要结构包括卷积部分和全连接部分。卷积部分包括卷积层、激活层、下采样层，通过卷积部分的叠加提取特征；全连接部分连接特征提取和输出计算损失和完成识别分类等功能。CNNs是第一个真正成功地采用多层层次结构网络的 具有鲁棒性的深度学习方法。CNNS适应性强，善于挖掘数据局部特征，使得CNNs已经成为众多科学领域的研究热点之一，在模式识别中的各个领域得到应用并取得了很好的结果。随着海量标记数据和GPU并行计算的发展，使得卷积神经网络研究大放异彩。

2.2 深度置信网络

深度置信网络的基本结构单元是受限玻尔兹曼机 （RestrictedBoltzmann Machine,RBM）。受限玻尔玆曼机通过学习数据的概率密度分布提取抽象特征。深度置信网络通过 RBM 的叠加，逐层学习数据 概 率 分 布，即提取多种概率特征。DBN训练过程分为无监督训练阶段和微调阶段。无 监督贪婪逐层训练方法获得各层的条件概率分布，微调阶段包括利用带标签数据和 BP 算法 对 网 络 参数进行调整。DBN 具有高灵活性，容易拓展，不仅在图像处 理、语音识别领域网络预训练中应用广泛，也是构建新型深度学习网络的结构基础。

2.3 栈式自编码网络

栈式自编码网络的基本单元是自编码器（Au-to-encoder,AE）。自编码器包括编码层和解码层，通过编码－解码的方式复现输入信号。堆栈式自编码网络是累加自编码器，通过无监督贪婪逐层训练得到每层自编码器的权重，在每一层得到了关于输入数据的另一种表达形式，这些不同的表示就是特征，在原有的特征基础上加入自动学习到的特征可以提高学习能力，是深度神经网络预训练的重要方法之一。

2.4 递归神经网络

上述几种模型，隐藏层之间彼此独立，逐层的提取特征，但这样的连接方式忽略了数据间的关联性。递归神经网络会对上一层的信息进行记忆，添加到当前层的计算输出中，在结构中体现为隐层间有节点相连。递归神经网络是非常强大的动态系统，通常用于描述动 态 时 间 行 为 序 列，使状态数据在网络中循环传递，它更擅长语音文本信息，一次处理一个输入序列元素每个节点同时包括过去时刻序列元素的历史信息，一般采用时间反传（backpropagationthroughtime,BPTT）训练算法来解决非长时依赖问题。

**3 病理切片乳腺癌分类**

接下来部分主要通过介绍在乳腺癌分类中的典型处理流程来说明利用病理数据进行癌症分类的处理过程。图3所示为病理分类任务的流程图。第一步，根据

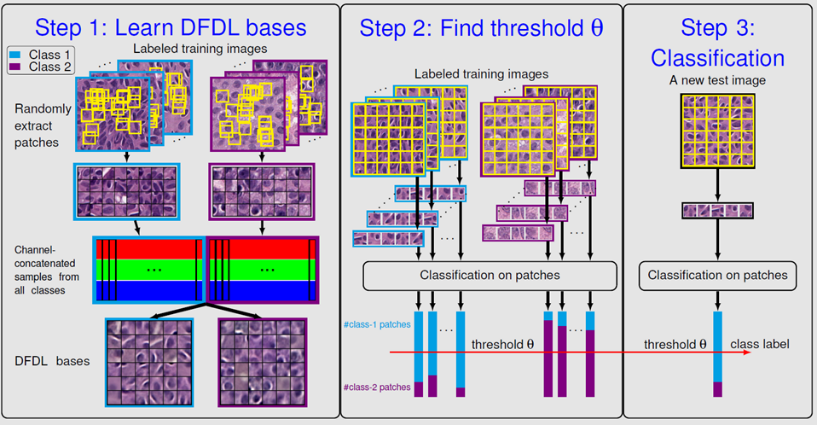


图3 病理图像分类任务流程

**3.1病理切片处理**

如前文中所说病理切片的扫描数据非常大，往往都是上万乘以上万维，即上亿的像素量，很难像一般数据（如自然图像）一样在简单预处理之后直接在神经网络上进行训练。必须要先进行数据预处理。病理数据集的处理主要包含三个步骤：剔除重复数据，根据MASK掩膜提取ROI，病理图像分割。

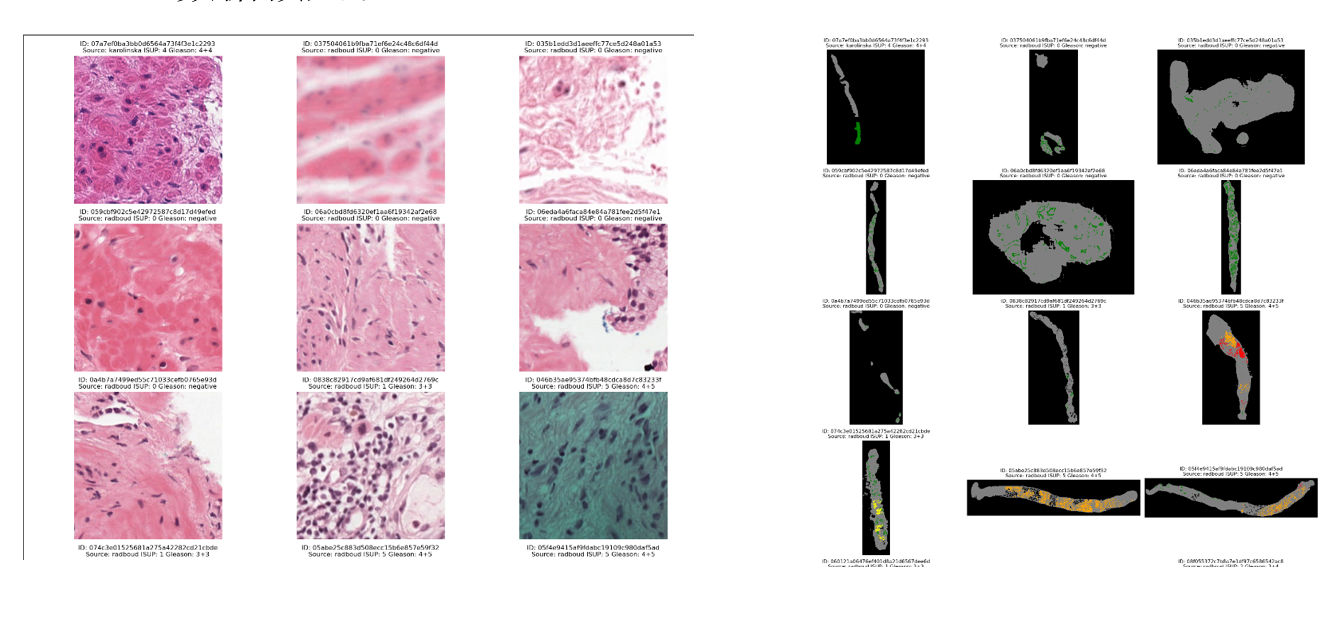


图3.1.1 病理数据集

医学影像数据集和自然图像数据不同的重要一点就是其数据量较少，难以直接获得，且因为隐私等原因，医院之间的病理数据往往并不相通，很难有单一机构独立发布大的、质量高的病理图像数据集。因此现在有的病理数据集并不多，而且大都是由多个机构中心联合发布，组成一个较大的数据集。这样新组合的数据集中很可能就会有重复的病理数据出现。对于要剔除这样的重复数据一般采用hash等算法来检测重复图像进行剔除。但同样因为病理数据大，因此首先又要生成数据集的缩略图，然后利用hash算法进行检测。

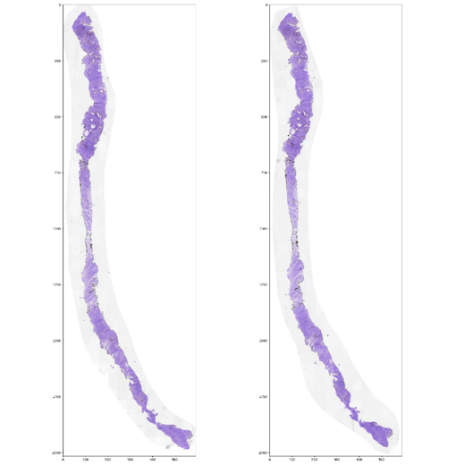
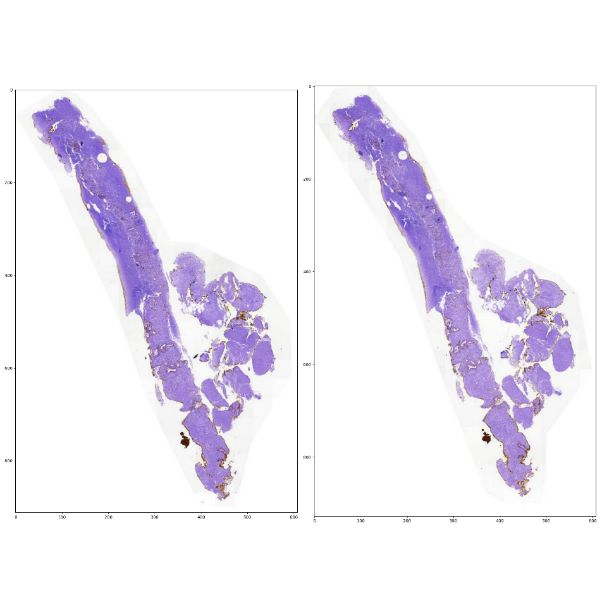
 

图3.1.2 hash算法检测重复病理图像

用以进行分类训练的数据集，一般由专家进行了标定，标明了病灶区（ROI）所在区域。所以第二步要根据ROI的标签值定位病灶区域，将病理图像按照标签值分类进行提取。在这一步主要有随机滑窗提取patch和固定分割提取patch两种，一般采用固定滑窗进行分割，这样可以最多的从病理图像上获取信息。在损失图像数据较少的情况下构建数据集，尽量确保在下一步的数据处理过后保留更多的数据。这一步比较重要，对后面分类效果的影响非常大，因为对于病理切片染色后，病理区域中并非全是病灶，在标记的病灶区域周围就是正常组织。在提取patch的时候，patch当中很可能有病灶组织和正常组织混合，统计两者之间的比例进行分割。对于正常类，因为图像区域非常大，所以在提取的时候可以设定非常严格的要求，保留全为正常区域的patch。在只区分良恶的任务下，保存tumor类的时候，如果要求设定的要求太严格，比如只保留全为tumor的patch，则最后提取出来的patch数量可能就会严重不足。而如果这种比例设的太低，虽然能更多的扩充tumor数据集，但是不能将tumor和normal两类分开。直接影响最后模型的分类精度。特别是在多分类任务的情况下，面临的这个问题就更加严重。选取合适的阈值和分割的patch尺寸大小，对于模型分类效果影响非常的大。

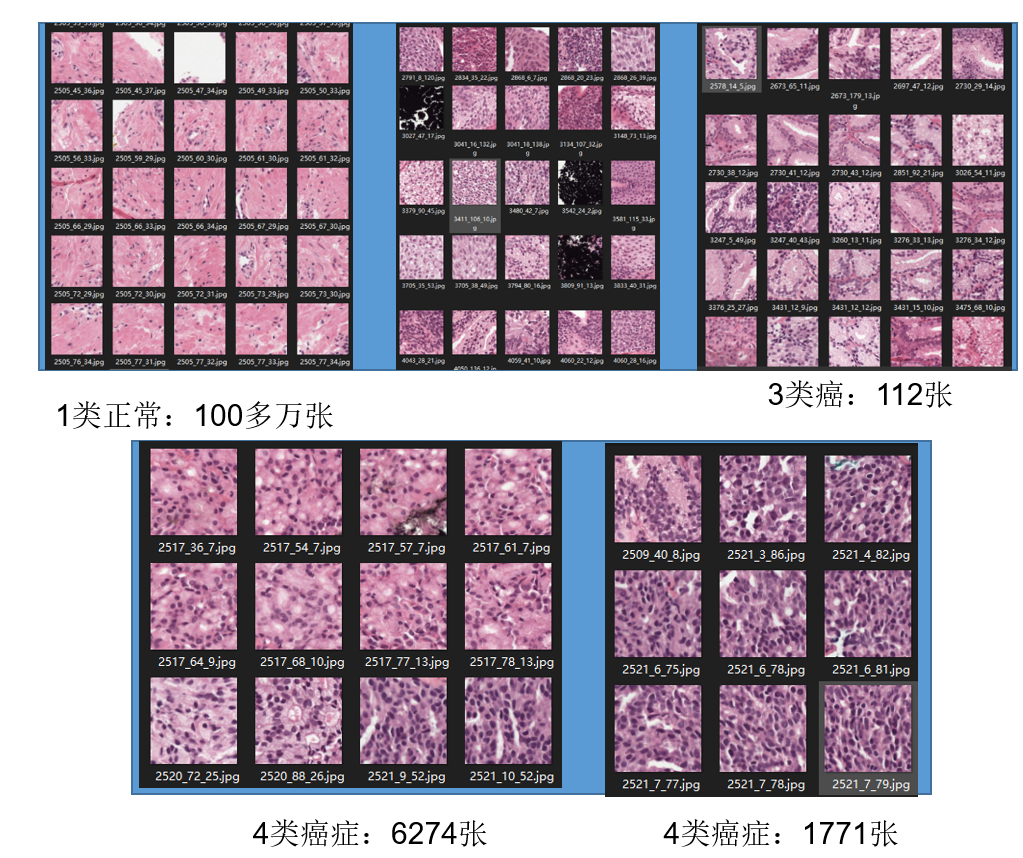


图3.1.3 panda数据集提取patch

图3.1.3所示为在处理panda时，设定比较严格的提取要求（每张patch中最多只有两类，且主要类占比超过0.96）时的处理结果。可以看出3类癌的patch分割数量非常少，数据不平衡。对这种问题一般的处理步骤为，先设定一个比较严格筛选标准，如果得到的数据集的大小不够则考虑先降低patch尺寸的大小。如果得到的数据集大小还是不够，则考虑不断尝试考虑降低筛选标准。实在不行，那就证明该数据集不适宜进行该分类任务的训练，这时因该放弃寻找其他满足要求的数据集。加上数据量非常大，所以在这一步的耗时和处理周期非常长，但是对于后续处理非常重要。

**3.2过滤和处理**

此步骤由过滤和预处理模块（FPM）执行。该模块执行以下转换：

•RGB到灰度转换。为了减少要处理的数据量，将24位的RGB图像转换为8位的灰度图像。

•缩放。脸部图像被缩放到固定数量的行和列。每个人脸的输出分辨率可以由用户根据要求的精度设定。

•直方图均衡化。这种方法增加了全局对比度，因此可以更好地分布在图像直方图上。此过程有助于提高面部识别精度。

当FPM完成后，将执行一个名为表格数据集构建模块（TDBM）的进程。这个模块获取图像像素，并生成表格数据集，其中行是人脸目标的个数，列是图像像素。最后一列表示类属性，即当执行SALM时用户分配的标签。TDBM通过M\*（N + 1）维的矩阵，其中M是数据库上检测到的人脸的总数，N是取决于面部图像分辨率的总像素数（即对于25×25图像，存储图像像素需要625列）。第（N + 1）列用来存储给人脸图像的标签。使用ARFF（属性关系文件格式）格式存储数据。ARFF文件是一个ASCII文本文件，用于描述共享一组属性的实例列表。 ARFF文件由怀卡托大学计算机科学系的机器学习项目组开发，与Weka ML软件一起使用。

**3.3训练和分类**

对于分类算法的训练，需要以下步骤：

•分割数据集。 该操作由随机排列和分割模块（RPSM）执行。基本上，执行表格数据集中包含的样本的随机排列，并且结果数据集分为两个数据集：训练数据集和测试数据集。

•训练分类算法 每个分类算法将由RPSM生成的训练数据集作为输入，并对每个分类器执行模型构建。然后，每个分类器的模型都存储起来，以便稍后在分类步骤中使用。

在分类模块中，借助于先前训练好的分类器，将来自测试集的人脸图像作为输入，应用过滤器和预处理，并对经过训练的分类器生成的每个模型中的人脸图像进行评价。进行比较后，人脸图像被分类，每个分类器生成的都存储在磁盘中。

**4.总结与展望**

本文详细探讨了深度学习在基于视频流的人脸检测与识别中的应用，对计算机视觉和深度学习技术在解决人脸检测和识别问题上的适用性进行了评价，并构建了一个典型的视频流人脸检测和识别的实现框架。作为未来的工作，必须解决两个主要问题：第一步，必须考虑一个更加可靠、鲁棒性更好的人脸检测器，以解决Haar-Like特征人脸检测器的不足；第二，要用更多的机器学习算法来验证哪些方法的分类结果最好；第三，要考虑和解决在复杂背景下如何提高人脸识别精度的机器学习方法。

**5.感谢**

不知不觉中，伴随着期末的到来，人工智能课程要结束了。叶老师上课认真负责，一丝不苟却又能以诙谐的方式调动课堂氛围，寓教于乐，让我们在欢快的气氛中有所收获，感谢老师的辛勤付出！

**参考文献**

1. [E. García Amaro](http://www.baidu.com/link?url=aI30BaOTIxHzqpr_4Q3_AYBuYv5zQ3jRGvZpHH1HLmXL4uBKK1B88DUw_JNS0EXn7jdF4UTAjKMx0ycBM7S_DTUfm59Xr61HTkvjYIQfADZW_TKWd2H0B5oZMERq-yvQQV--jZHyOvccTfGm0FlIC-7Nfw7Rymwrg5OIrY0V9kP9bBo9_7b4rqp1VkE4y1qE), [M.A.Nuño-Maganda](http://www.baidu.com/link?url=apQz2zZY3xnfxtaEFgGl80xyrrZAKkzy20YbGKJ0nvBcuksuXYiNSptS7MMZrApNdxE2feNTc4KKBym_yid1YmlXlYhOob49Y5Ftk_gFfDJnzStJmoEDN-y4OBS-bRdPTmztTmS_0K-yaI8azG6U0W2WAITGgsGIQRVKOpbMXxvOmqvH22oLi35VoOb6Vqnl), *Evaluation of Machine Learning Techniques for Face Detection and Recognition* - [International Conference on Electrical Communications & Computers](https://www.baidu.com/link?url=OGqXe13ZcaD62OSyUaaq4vRwdrXVKjqZ971HPzxGGz7PR297FATw9MNK30Ux45SjxH8ubihu3rOPtwVEudYEga9JKC2nao344jvyGozpGJBz28qpZXQAiNWYGeE7We-tjISNqfg29DnM9CS5MwxT9CHpREgWxEqEGSKk-Ogp2KuOTeu7lptN0xa9ktDlOrCHP76ZEsRfZAIdj6PTyd5aFxgUSW3ng-cW3b6ZJG_JU_j5R5dkkPGRwqTOgNK0kPtwIo2GswwkQgR4marjh6JSGFc4aNNS1u6k1Nf1ydn2f3u&wd=&eqid=ebe96a44000179f8000000065965b910), pp: 213-218, 2014.
2. Patil H, Kothari A, Bhurchandi K, *Expression invariant face recognition using semidecimated DWT, patch-LDSMT, feature and score level fusion*. Appl Intell, vol. 44, no.4, pp: 913-930, 2014
3. Moghaddam B, Jebara T, Pentland A. *Bayesian face recognition*. Pattern Recognition, vol. 20, no. 33. pp: 1171-1182, 2015.
4. Jufu Feng, Xiao Ma, Wenjing Zhuang. *Collaborative representation Bayesian face recognition.* [Science China Information Sciences](http://xueshu.baidu.com/usercenter/data/journal?cmd=jump&wd=journaluri%3A%28247511c3d87ad72b%29%20%E3%80%8AScience%20China%20Information%20Sciences%E3%80%8B&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dpublish&sort=sc_cited),vol. 60, no. 4, pp: 81-101, 2017.
5. D. S. Bolme, J. R. Beveridge, and B. A. Draper, *Facel: Facile face labeling*, Lecture Notes in Computer Science. vol. 5815, pp. 21–32, Springer, 2016.