# 基于特征融合的宫颈癌前病变识别

**摘要**：宫颈癌是威胁全世界女性的健康的重大疾病，而阴道镜筛查是检测宫颈癌的是否发生病变的不可缺少手段。在临床上，人工筛查不可避免的会出现误诊和漏诊的现象。为了辅助医生，提高宫颈癌前病变的识别率，我们提出了一种双线性卷积网络来对阴道镜图像进行识别：1、使用相同的卷积网络来对碘图像和醋酸图像来提取空间特征。2、使用双线性池化将提取到的空间特征进行融合，获得更高层次的语义信息。为了训练和评估我们的模型，我们一共使用了2400张临床数据，并且在测试集上达到了81%的准确度，72.13%的敏感性和89%的特异性，AUC达到了0.86。结果表明，相比于临床医生和仅使用醋酸图像的计算机辅助算法，我们的提出的方法可以更好的对宫颈癌前病变进行识别。

**关键字：双线性池化、宫颈癌、深度学习、特征融合**

**一、Introduction**

宫颈癌是世界上威胁女性生殖系统健康的第二大肿瘤，占世界上所有癌症的12%[1,2]。在中低收入国家，它有80%的致死率，每年都有超过26万女性死于该病[3]。现已证明，宫颈癌的发生是由病毒引起的，而高危人乳头瘤病毒（HR-HPV）感染是引发的必须因素[1]。子宫颈上皮内病变是宫颈表面鳞状上皮的异常增生和癌前病变,需要经过漫长的时间才能进展到子宫颈癌，专家们往往都是根据这一特性，进行癌前病变筛查，阻断宫颈癌的发展。

目前，阴道镜下活检是诊断子宫颈上皮内病变及宫颈癌的金标准[4]。阴道镜可以在放大3-7倍的情况下检查子宫颈，在应用稀释至3%的醋酸溶液以及卢戈碘溶液，使得异常病变组织与正常组织出现明显分化，其典型变化情况如fig.1所示。但是，临床检查的准确性还是很大程度上依赖医生的主观判断，部分有经验的医生在临床上也仅有48%的特异性[5]。因此找到一种科学、准确、快速的阴道镜诊断方法可以缓解临床上需要大量的经验丰富医生的尴尬局面。

近年来，机器学习算法在医学领域已经有了广泛的应用，例如分类[6]，分割[8]，图像恢复[7]等。机器学习算法能够提取到常规算法难以检测的信息，因此可以向临床医学中提供有效的诊断。此外，也已经有部分学者将机器学习算法运用到阴道镜图像分析[6][10][11]。然而，已经存在的这些算法大多仅对醋酸图像进行分析，忽略了临床上，医生是将碘实验图像和醋酸实验图像对比观察的这一先验知识，具有一定的局限性。

不同等级病变的特征差异非常小，只能通过转化区的颜色，纹理等细微差异来区分。双线性池化是一种非常有效的特征融合技术，它已经在计算机视觉领域有了广泛的应用[12][13][14][15]。它能捕获特征之间的关系，将不同特征进行交互，从而生成更具有表达力的全局关系。使用双线性池化技术把醋酸图像和碘图像相关特征融合，可以提高信息的交互，获得更有价值的分类信息。

在这篇论文中，我们介绍了所设计BF-CNN的结构和如何同步处理两张图片。BF-CNN是一种辅助医生进行宫颈癌前病变筛查的计算机辅助算法。它首先使用两个相同结构的CNN去提取两种状态图像的空间特征，为了使特征更具有更多的属性信息，使用不同线性结构将获得特征映射到更高维的空间，然后进行双线性池化获得具有丰富信息的特征描述子，然后对病变的种类进行区分。此外，为了使我们的结果更具有临床意义，我们将病变种类划分为CIN1-和CIN2+。实验结果和其他算法进行比较，并通过准确率，特异性，灵敏性和AUC等指标来评估性能。

我们的主要贡献可以被总结为以下几点：

1. 提出了一个新的自动划分宫颈癌前病变的方法，该方法可以结合两种不同状态图像的空间特征，生成更具有丰富信息的特征描述子。
2. 为了获得更多的属性信息，我们设计两个线性结构，将各自特征映射到更高维度的空间。
3. 我们提出了方法在2400张临床数据上进行训练和评估。我们的结果远超于其他方法，准确率到达81%，特异性89%，灵敏性73%。此外，ROC曲线的面积达到0.86。

**二、related work**

1、计算机辅助算法对阴道镜图像识别

已经有部分学者致力于使用机器学习来对宫颈疾病进行筛查和分析。早期，有研究者使用手动提取特征，分类器训练和验证的方式证明了传统方法的可行性[16][17][18][19]。但是手工提取的特征需要一定的专业知识，而且存在一定的人为误差具有一定的局限性。近几年，相继有学者使用神经网络、支持向量机、KNN、决策树等方式来对宫颈图像自动进行分析[20]。Xu等[21][22]人使用 PLAB,PHOG和PLBP三种互补类型特征来进行训练多种不同类型的分类器。Asiedu[23]将提取到的醋酸图像和碘图像特征以串行或者并行的方式结合，训练支持向量机分类器。该方式虽然结合了两种图像特征，但是他们的分类结果不具有一定的临床意义。Zhang等[24]使用了不同的预训练模型的DenseNet对宫颈图像自动分析，尽管受限于数据量，但是迁移学习还是能得到一个满意的结果。Li等[25]在vgg的基础上设计一个基于注意力机制的特征金字塔卷积网络用来分类阴道镜图像，准确度和灵敏度都略有提高。而我们提出的方法是使用两个不同的卷积网络作为两种图像关键属性的征提取器，再以直观高效的方式对他们进行交互建模。深度学习的相关技术可以极大提高自动分析图像的性能，并且使用深度学习的相关技术来分析处理医学图像已经慢慢成为一种研究趋势[26][27][28][29]。

2、双线性池化技术的应用

双线性池化最早是被提出用来模拟类似于图像的“风格”和“内容”这种双因素变化的模型[30]。在近几年，Lin等[12]人发现双线性池化技术可以很好的将特征进行融合，因此设计了一个双线性卷积网络架构，聚合了两个线性结构中的特征，扩大细粒度类别中的细微特征。这项工作的发表又引起了双线性池化技术的关注。Gao等[13]提出采用 Random Maclaurin (RM)和Tensor Sketch (TS)两种方法对双线性池化获得的特征进行降维。Kong等[33]在假设分类器是支持向量机下，为了进一步的简化计算，对协方差矩阵进行低秩近似。Fukui等[32]在Gao等[13]基础上提出了Multimodal compact bilinear pooling，将其应用到视觉问答领域，融合不同模态的特征。Chao等[15]为了增强层间特征的交互，减少卷积过程中的信息丢失，提出了Hierarchical Bilinear Pooling，用于细粒度图像分类。Suh等[34]将bilinear pooling用于ReID领域，将行人的局部特征和全局特征进行融合，再进行检索。Hu等[35]使用同源双线性池化技术将提取到的多视角特征融合，用于动作识别领域。

三、Methods and Materials

1 data and preprocessing

我们的数据集一共包含1200例病例，每个病例中包含一张醋酸图像、一张碘图像和一张病理报告单。这些数据全部来在浙江邵逸夫医院2017年1月到2018年3月的临床实践数据。这些病例没有受限制条件的筛选，完全是随机选区，一共有2400张临床数据被选择使用。在临床上，宫颈图像可分为四类：正常，CIN1,CIN2,CIN3和Cancer[36]。正常和CIN1可依靠人类的自身免疫而恢复，而CIN2 、CIN3和Cancer需要手术进行切除。为了使我们的结果具有临床意义，我们结合医生的意见将样本划分为两类一类为LSIL(包含正常和CIN1),另一类为HSIL(包含CIN2、CIN3和Cancer)。训练集一共使用1000例病患，其中LSIL有503例，HSIL为497例。测试集LSIL和HSIL各包含100例。我们在训练集上进行5折交叉验证，在测试集上来检测模型的性能。数据预处理也是实验的一个必不可少的一个环节。原数据的尺寸大小为640\*480, 为了有利于训练，我们统一把图片resize到448\*448。数据增强是深度学习常用的一个技巧之一，可以提高模型的泛化能力和鲁棒性。在训练过程中，选择了三种数据增强方式：随机裁剪、水平翻转、颜色抖动。其各个参数可以如表1所示。

2、Method overview

我们的实验目标是通过我们设计的卷积网络来识别是否发生高等级的病变，整体的网络架构图如图1所示。两个特征提取网络为resnet[39]，两个结构是相互独立的，其结构示意图如图2所示 。而双线性池化是为了融合两个卷积网络提取到的特征，其内容在3.4节详细介绍。

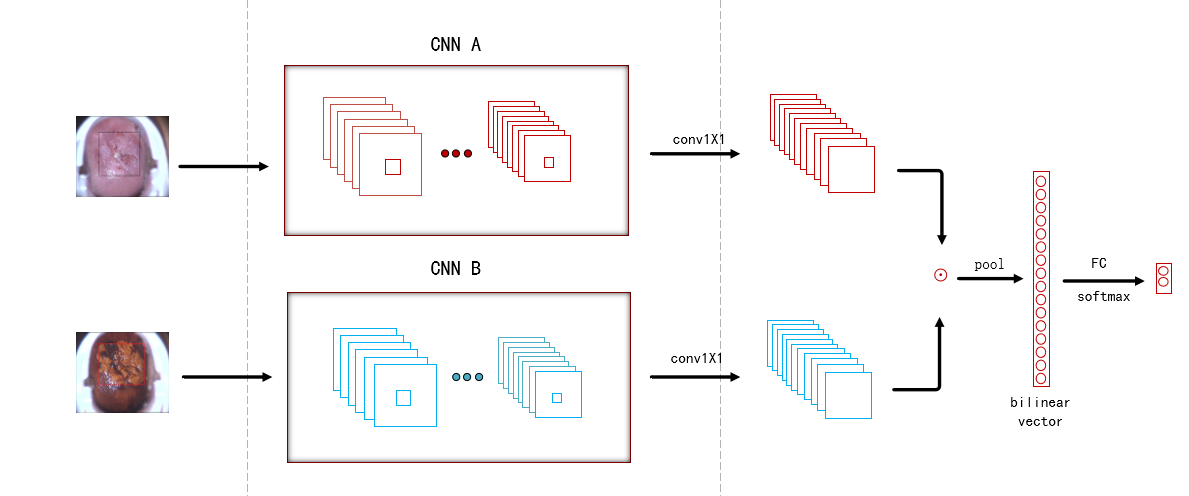


图 1 BF-CNN 结构图

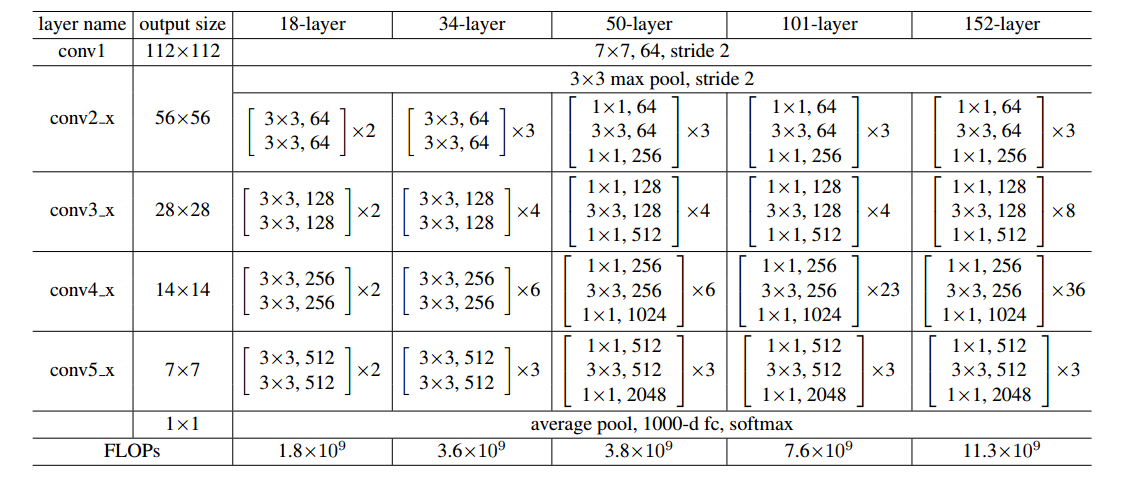


图 2 resnet 网络结构

3、fine-tuning pretrained networks

在我们的研究中，我们使用resnet作为我们的主干网络结构，该网络曾在2015年Imgnet分类比赛中取得冠军。每个resnet由16个残差单元组成，残差块的组成由图3所示，其残差计算公式如eq1所示。Resnet在每一层都引入了一个跳跃连接，这可以使得上一个残差块的信息没有阻碍的流入到下一个残差块，提高了信息的流动能力，并且也因此避免了由于网络过深而引起的梯度消失的状况。Resnet实际上可以视为多个浅层网络的融合，它相比于VGG等卷积网络的可以更快的收敛。由于我们的数据量是有限的，深层的resnet并没有比浅层的resnet有明显的性能提升，所以我们采用resnet34作为我们的特征提取网络。

(1)

其中，代表要学习的残差映射。和为输出和和输入的卷积层处理过的特征图。

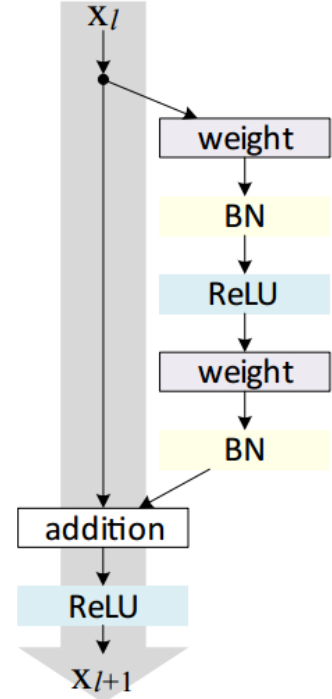


图 3 残差单元结构图

Fine-tuning pretrained networks（迁移学习）是在深度学习中经常被使用的一种技术，尤其是在计算机视觉和自然语言处理中。众所周知，训练一个神经网络需要大规模的数据、算力和时间。在数据量有限的情况下，随机初始神经网络的参数进行训练，很容易得到局部最优解，不能得到一个满意的结果。Fine-tuning把一个训练好的模型迁移到一个全新的任务上，并且以这个模型作为新的起点对参数进行细微的调整，即便当新任务和原始任务的相似性不是很高时，也是可以取得一个相对较好的结果，提高模型的泛化性能[37]。所以在我们的数据量有限的情况下，使用迁移学习可以更快的收敛，同时并获得一个较好的结果。现阶段的迁移学习大多都是从ImageNet或其他大型数据集学习到参数运用到其他基于图片特征的任务上。参考迁移学习在相似任务上的表现[24]，使用ImageNet的预训练模型要比kaggle的预训练模型的性能稍好一些，所以我们采用resnet34在ImageNet中的预训练模型来初始化我们特征提取网络的参数。

4、bilinear pooling for prediction

宫颈各种病变种类通常具有相似的外观，分类主要是通过医学试剂处理后感兴趣区域（在临床上称为转化区）的细微特征的差别来区分病，例如颜色，纹理，血管等。感兴趣区域在经过醋酸试剂和碘试剂理过后，图像的特征分化往往是存在一定的关联的。在临床上，专家也是通过对比两种图像的感兴趣区域的特征，来判断该病例是否发生病变的。而双线性池化是一种特征融合的技巧，可以成对捕获特征间的联系。该方法是模拟人类的视觉处理过程的两个pathway，一个进行物体识别，另一个发现物体的位置，基于这样的思想，将两个不同的特征共同发挥作用，提高细微特征的差异[12]。在这个思想的基础上，我们使用两个卷积网络分别提取两种图像的最具有区分度的特征，使用双性化池化的将两种特征融合，来增强对细微特征的感知。值得注意的是，这刚好和人类专家也是对两种图像的关键特征的结合是相似的想法。

分解双线性池化最早被应用于视觉问答领域，而Kim此基础[38]上提出了基于Hadamard product的分解双线性池化用于多模式学习注意力机制。我们结合该方法在细粒度图像识别中的应用[15]，将其设计到我们的架构中。这里，将介绍该模块在我们架构中的实现的基本公式。假设有来自同一个病例的醋酸图像A和碘图像B, 设A被卷积网络处理后，输出的特征图为,B被卷积网络处理过后，输出的特征图为,h特征图的高度，w为特征图的宽度，c为特征图的通道数，我们把c定义空间特征的描述维度,可把和写为 和 。所以双线性模型就可以有如下定义：

(2)

其中是一个可被学习的投影矩阵，为双线性模型的输出。然后把投影矩阵分解成两个一阶向量，可被写作如下：

(3)

其中和。我们需要维可学习矩阵去获得输出向量。因此，输出向量可被写作如下：

(4)

其中, 是投影矩阵， 是分类矩阵，d是一个控制输出向量维度的超参数，○是Hadamard product。需要说明的是，为了使feature map包含更多的属性信息，我们和Chao[15]一样,对卷积层提取的特征使用独立的线性结构将其映射到更高维度的空间。

最终对癌前病变分类的预测分数是把输出向量放入Softmax层来获得的。Softmax函数在是分类任务中常用的输出层，其计算公式如下：

(5)

其中代表属于每一类的概率，和为权重矩阵和偏移矩阵，为双线性向量。

相对应的损失函数选择交叉熵函数，该损失函数也是分类任务中常用的，其计算方式如下：

(6)

其中n代表病例的数量，代表病例的真实标注，代表我们设计模型所预测出的标签。

6 implementations

为了提高我们模型的泛化能力和鲁棒性，我们采用5折交叉验证法来训练模型。训练集被划分为5份，其中四份用来训练，剩余的部分作为验证集，来计算模型的精度。也就是说模型需要被重复训练5次。最终取每个模型在测试集上的预测结果的平均值来评估模型的最终性能。

在训练过程中，迭代次数的值设置为100，学习率设为0.003，batch\_size设为16。所有的实验都是在windows10平台上，AMD 3600 3.7HZ CPU，NVIDIA RTX 2060s 8G GPU 和 16G RAM，使用pytorch框架编码实现。

**4、result and discusion**

1、模型的评估标准

对模型的训练，实际是为了解决判别阴道镜图像是否含有高等级病变的二分类的问题。图像的标签为0和1。0代表的是阴性样本，1代表的是阳性样本。为了分析我们所提出模型的性能，我们采用4项在常用的指标，如敏感性，特异性，准确度和AUC。前三项指标的计算公式如下：

(7)

(8)

(9)

特异性和灵敏性医学领域常用的评价指标。其中TP,TN,FP,FN在测试集中真阳性，真阴性，假阳，假阴性的数量。真阳性是指属于阳性类别并被正确分类的那一组患者，假阴性是指属于阳性类别但被错误分类为阴性的那一组患者，真阴性和假阳性的定义是类似的。ROC（receiver operating characteristic）和ROC曲线下的面积（AUC）是分类问题中常用来评价分类器性能的指标，它的x轴为假阳性，y轴为真阳性率。

2、结果分析

模型在训练的时候，学习率设置的相对较小，若学习率过大，则会陷入过拟合，造成一些性能损失。图4和图5显示了我们的模型的在训练过程训练集和验证集的损失和准确度变化。模型基本在训练60个epoch以后，验证集的损失就基本趋于稳定，达到0.42，准确率达到0.82。为了评估我们模型的性能，我们使用200例病人作为测试集。图6显示了不同的训练模型在测试集上预测结果的平均值。对醋酸图像使用resnet34进行测试，其分类准确度75.50%，灵敏度为71.16%，特异性79.61%。对碘图像使用resnet34进行测试，其分类准确度达到了76.5%，灵敏度为72.13%，特异性为80.58%。我们在此基础上，增加双线性模块融合两个状态图像的相关联信息，最终分类准确度达到81%，灵敏度为72.13%，特异性89.32%,灵敏度相比单一结构的网络略有提升，特异性有了较大的提高。我们拿resnet34对醋酸图像的roc曲线做对比，如图7所示,可以观察到我们提出的方法的ROC曲线更靠近左上角，AUC的值为0.86，证明了我们的模型相比于仅使用单一图像的网络结构是有效的。

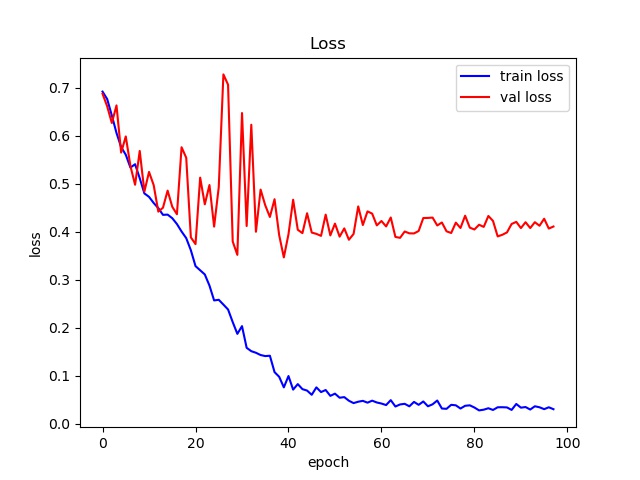
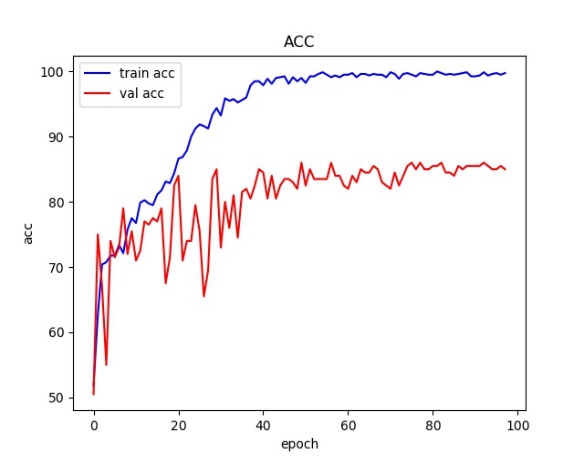


图 4 损失曲线 图 5 准确率曲线

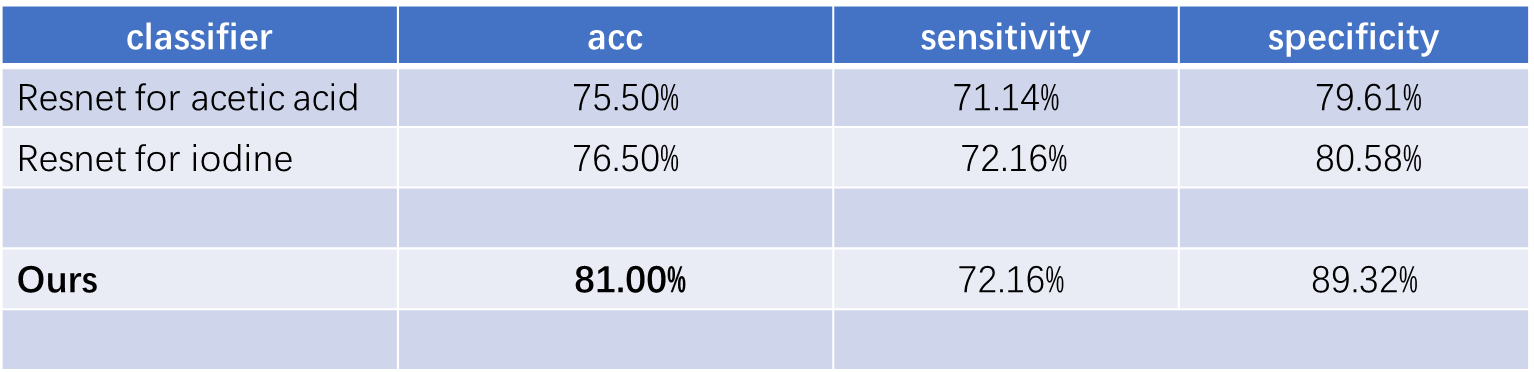


图 6 实验结果对比

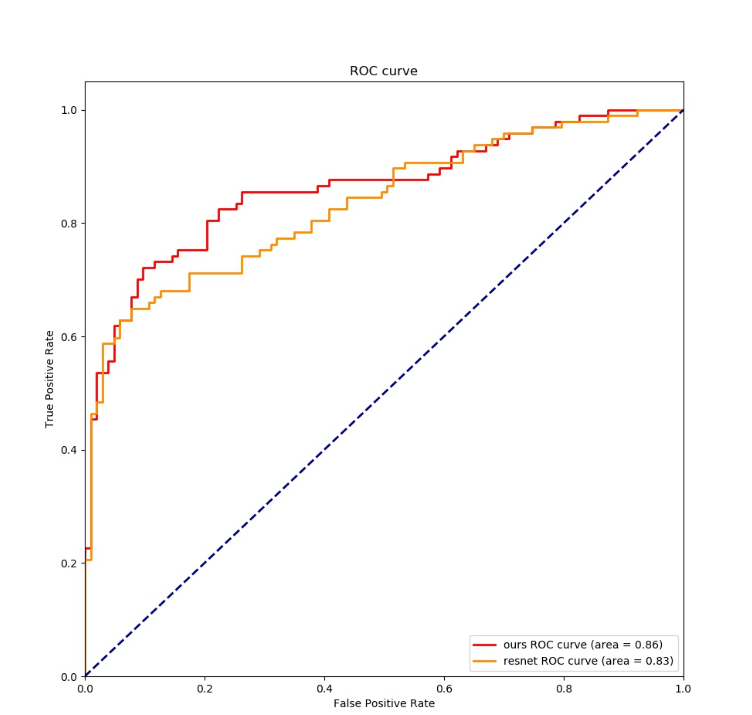


图 7 ROC曲线

3 类激活映射

为了验证模型是否和我们设想的一样能够提取到两类图像相关联的信息，我们可视化了两个独立卷积网络在我们数据集的提取到的特征图，同时邀请了专家对疑似病变区域进行了标注。我们使用grad-cam[40]算法获得类激活映射图。在图8中，我们展示了一些病例的两类图像以及它们相应的可视化结果。这些病例原始图像的黑色框，是我们邀请专家对病变位置的标注。图像高亮区域是把该病例判别为阳性密切相关的部分。从可视化的结果来看，两个卷积网络在提取的特征在宫颈上响应的位置基本是一致的，与专家标注的结果相比，是存在一定比例重合的。可视化的结果表明了我们提出的模型是有能力从两类图像中提取到相关联的信息的。

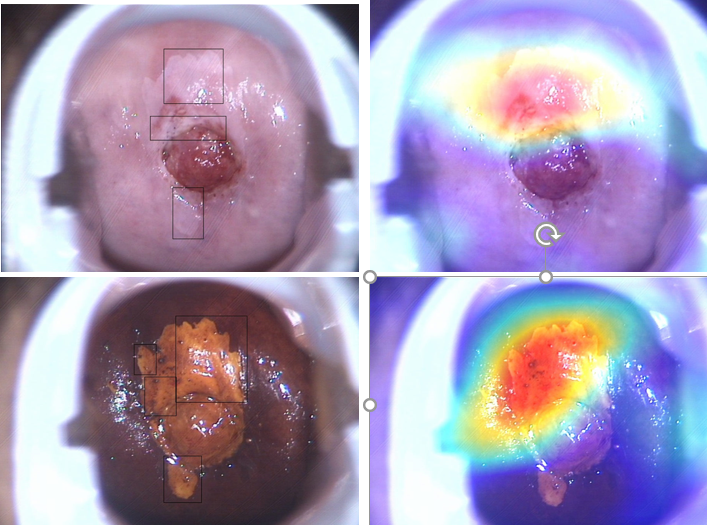


图 8 可视化结果对比图

五、Conclusion

我们提出了计算机辅助算法-BF-CNN，该算法融合结合医生的经验融合两类图像的特征用于宫颈图像的癌前病变的识别。

**参考文献**

1. Siegel, R.L., K.D. Miller and A. Jemal, Cancer Statistics, 2017.

CA Cancer J Clin, 2017. 67(1): p. 7-30

1. J. Ferlay, I. Soerjomataram, M. Ervik, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D. M. Parkin, D. Forman and F. Bray 2013 GLOBOCAN 2012 v1.0, Cancer incidence and mortality worldwide: IARC CancerBase No. 11, Lyon, France: IARC [internet]. Available from:http://globocan.iarc.fr, . Last accessed on 28-07- 2018.
2. Tsu, V. and J. Jeronimo, Saving the World's Women from Cervical

Cancer. N Engl J Med, 2016. 374(26): p. 2509-11.

1. Azvolinsky, A., Screening Guideline for Cervical Cancer Recommends Against Human Papillomavirus-Pap Cotesting. J Natl Cancer Inst, 2017. 109(12).
2. M. F. Mitchell, D. Schottenfeld, G. Tortolero-Luna, S. B. Cantor,and R. Richards-Kortum, “Coloposcopy for the diagnosis of squamous intraepithelial lesions-A meta-analysis,” Obstet. Gynecol., vol. 91, no. 4, pp. 626–631, 1998
3. T. Xu et al., “Multi-feature based benchmark for cervical dysplasia

classification evaluation,” Pattern Recognit., vol. 63, no. January

2016, pp. 468–475, 2017.

1. Anavi, Y., Kogan, I., Gelbart, E., Geva, O., Greenspan, H.:Visualizing and enhancing a deep learning framework using patients age and gender for chest x-ray image retrieval. In: Medical Imaging 2016: Computer-Aided Diagnosis, vol. 9785,p. 978510. International Society for Optics and Photonics (2016)
2. Kamnitsas, K., et al.: Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Med. Image Anal. 36, 61–78 (2017)
3. Kim, E., Huang, X.: A data driven approach to cervigram image analysis and classification. In: Celebi, M., Schaefer, G. (eds.) Color Medical Image analysis, pp. 1–13. Springer, Heidelberg (2013). [https://doi.org/10.1007/978-94-007-5389-1 1](https://doi.org/10.1007/978-94-007-5389-1%201)
4. Kudva V., Prasad K., Guruvare S. (2020) Transfer Learning for Classification of Uterine Cervix Images for Cervical Cancer Screening. In: Kalya S., Kulkarni M., Shivaprakasha K. (eds) Advances in Communication, Signal Processing, VLSI, and Embedded Systems. Lecture Notes in Electrical Engineering, vol 614. Springer, Singapore
5. Kudva, V., Prasad, K. & Guruvare, S. Hybrid Transfer Learning for Classification of Uterine Cervix Images for Cervical Cancer Screening. *J Digit Imaging* (2019).
6. Lin,T.-Y.; RoyChowdhury, A.; and Maji, S. Bilinear cnn models for fine-grained visual recognition. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), 1449– 1457, 2015.
7. Gao, Y.;Beijbom, O.; Zhang, N.; and Darrell, T. Compact bilinear pooling. InProceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2016, 317–326.
8. Li, Y.;Wang, N.; Liu, J.; and Hou, X. Factorized bilinear models for image recognition. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017, 2079–2087.
9. Chao jian Yu, Xinyi Zhao, Qi Zheng, Peng Zhang, and Xinge You. Hierarchical bilinear pooling for fine-grained visual recognition. In The European Conference on Computer Vision (ECCV), 2018.
10. V. Pallavi and K. Payal, “Automated analysis of cervix images to grade the severity of cancer,” in Proc. Annu. Int. Conf. IEEE Eng. Med. Biol.Soc., 2011, pp. 3439–3442.
11. S. Y. Park et al., “Automated image analysis of digital colposcopy for the detection of cervical neoplasia,” J. Biomed. Opt., vol. 13, no. 1.Jan./Feb. 2008, Art. no. 014029.
12. H. Greenspan et al., “Automatic detection of anatomical landmarks in uterine cervix images,” IEEE Trans. Med. Imag., vol.28, no.3, pp.454–468, Mar.2009.
13. W. Li et al., “Using acetowhite opacity index for detecting cervical in-traepithelial neoplasia,” J. Biomed. Opt., vol. 14, no. 1, Jan./Feb. 2009, Art. no. 014020.
14. Y. Jusman, S. C. Ng, and N. A. Abu Osman, “Intelligent screening

systems for cervical cancer,” Sci. World J., vol. 2014, 2014.

1. T. Xu, C. Xin, L.R. Long, S. Antani, Z. Xue, E. Kim, X. Huang, A new image data set and benchmark for cervical dysplasia classification evaluation, in: International Workshop on Machine Learning in Medical Imaging, Springer, Cham, 2015, pp. 26–35,

<http://dx.doi.org/10.1007/978-3-319-24888-2_4>

1. T. Xu, E. Kim, X. Huang, Adjustable adaboost classifier and pyramid features for image-based cervical cancer diagnosis, in: 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI), IEEE, 2015, pp. 281–285, <http://dx.doi.org/10.1109/isbi.2015.7163868>.
2. M. N. Asiedu *et al*., "Development of Algorithms for Automated Detection of Cervical Pre-Cancers With a Low-Cost, Point-of-Care, Pocket Colposcope," in *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 8, pp. 2306-2318, Aug. 2019.
3. T. Zhang, Y.-m. Luo, P. Li, P.-z. Liu, Y.-z. Du, P. Sun, B. Dong, H. Xue, "Cervical precancerous lesions classification using pre-trained densely connected convolutional networks with colposcopy images", *Biomedical Signal Processing and Control*, vol. 55, pp. 101566, 2020.
4. Li H., Zhao J., Zhang L., Zhao J., Yang L., Li Q. (2020) Feature Pyramid Based Attention for Cervical Image Classification. In: Li Q., Leahy R., Dong B., Li X. (eds) Multiscale Multimodal Medical Imaging. MMMI 2019. Lecture Notes in Computer Science, vol 11977. Springer, Cham
5. G. Litjens et al., “A Survey on Deep Learning in Medical Image

Analysis,” Med. Image Anal., vol. 42, pp. 60–88, 2017.

1. R. Zhang et al., “Automatic Detection and Classification of

Colorectal Polyps by Transferring Low-Level CNN Features from

Nonmedical Domain,” IEEE J. Biomed. Heal. Informatics, vol. 21,

no. 1, pp. 41–47, 2017.

1. J. X. Qiu, H. J. Yoon, P. A. Fearn, and G. D. Tourassi, “Deep

Learning for Automated Extraction of Primary Sites from Cancer Pathology Reports,” IEEE J. Biomed. Heal. Informatics, vol. 22, no.1, pp. 244–251, 2018.

1. A. Esteva et al., “Dermatologist-level classification of skin cancer with deep neural networks,” Nature, vol. 542, no. 7639, pp. 115–118, 2017.
2. J. B. Tenenbaum and W. T. Freeman. Separating style and content with bilinear models. Neural computation, 12(6):1247–1283, 2000. 2
3. Gao, Y., Beijbom, O., Zhang, N., Darrell, T.: Compact bilinear pooling. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 317–326 (2016)
4. Fukui, Akira et al. “Multimodal Compact Bilinear Pooling for Visual Question Answering and Visual Grounding.” *EMNLP* (2016).
5. Kong, S., Fowlkes, C.: Low-rank bilinear pooling for fine-grained classification. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 7025–7034. IEEE (2017)
6. Suh Y, Wang J, Tang S, et al. Part-aligned bilinear representations for person re-identification[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 402-419.
7. Hu J F, Zheng W S, Pan J, et al. Deep bilinear learning for rgb-d action recognition[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 335-351.
8. WHO: Human Papillomavirus and Related Cancers in the World, Summary Report, ICO Information Centre on HPV and Cancer, 2014.
9. Yosinski J, Clune J, Bengio Y, et al. How transferable are features in deep neural networks?[C]//Advances in neural information processing systems. 2014: 3320-3328.
10. Kim, J.H., On, K.W., Lim, W., Kim, J., Ha, J.W., Zhang, B.T.: Hadamard product for low-rank bilinear pooling. arXiv preprint arXiv:1610.04325 (2016)
11. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
12. Selvaraju R R, Cogswell M, Das A, et al. Grad-cam: Visual explanations from deep networks via gradient-based localization[C]//Proceedings of the IEEE international conference on computer vision. 2017: 618-626.