**Expression recognition based on convolutional neural network, capsule network model and ResNet network**

Hualing Sha Yuhan Li Jingxue Chen Zhan Chen

Wuhan Textile University

**Abstract**

*面部表情作为表达内心情绪的一种直观方式，在人类的日常活动和沟通中扮演着举足轻重的角色。随着计算机视觉与图像模式识别技术的迅速发展，人脸的脸部特征信的辨识日益引起人关注。表情识别系统分别用三种方法实现: Vgg，ResNet和胶囊网络模型，再分别对fer2013与ck+数据集进行对比实现。*

# **Introduction**

现如今人类用于表达自身情感的方式多种多样，但是最 为直接的还是通过人脸的面部表情去判断。即使是相隔两地，人们也可以通过视频观察到他人的表情，从而判断出对方的心理状态,加深对彼此的理解,所以人脸表情的识别一直都是计算机视觉方向的一大热点。人脸表情识别的关键在于对人脸不同表情特征的提取.在 早期人们一般使用的都是基于传统人工设计的表情提取方法,但是这些方法一般设计都比较困难,并且最后得到的识别准确率并不高。为了解决传统方法的不足，另一种基于深度学习的方法随之诞生了出来。目前深度学习的应用场景十分广泛，包括有图像、语音、自然语言处理等领域,为了适应这些不同的场景,越来越多的深度学习模型被提出。 例如：Vgg，ResNet和胶囊网络模型。

# **Related Technologies**

## **Convolutional neural network**

卷积神经网路与一般的卷积网路不同之处是网路。在卷积神经网路的回旋层面，神经元仅连接于邻近神经元。CNN的旋流往往包括一些关键因素。各特征面包含若干长方形的神经元。在相同的属性平面上，每个个体都具有相同的权值，而所分享的权值为卷积核心。一般采用一个随机十进制的方法对卷积核心进行初始化。在神经网络的训练中，通过对卷积核的分析，得到一个适当的权重。分享加权（卷积核心）的最大优点是减小了网络层间的联结，并减小了过多的拟合。二次取样一般分为二次平均取样和最大二次取样。子抽样是一种特定的卷积法。卷积法和卷积法极大地降低了建模的计算量，降低了建模的参量。

在一般的卷积神经网络中，在每一次卷积操作中，都会反复地考虑采用局部均值或极大特征池层次的构造。采用多特征性提取网络，能显著地减少多特征图像的分辨能力。将完整的连接层的输出结果输入到softmax层，可以实现基于节点的多重特性抽取和特征图的对比。

## **ResNet network**

ResNet是一种残差网络，见图2.1。咱们可以把它理解为一个子网络，这个子网络经过堆叠可以构成一个很深的网络。

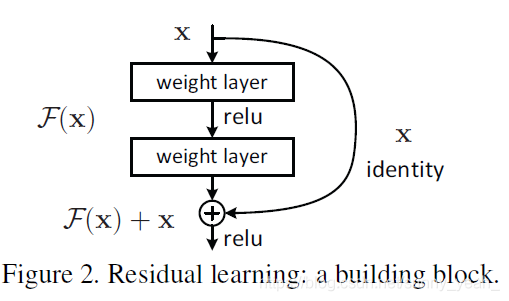


图2.1 残差学习构建块图

Resnet（残差网络）的特点是容易优化，并且能够通过增加相当的深度来提高准确率。其内部的残差块使用了跳跃连接，缓解了在深度神经网络中增加深度带来的梯度消失问题。

## **Capsule network model**

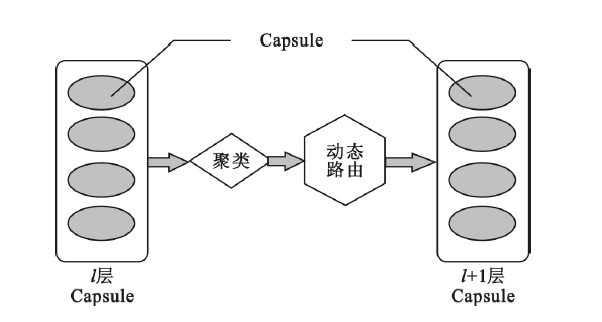
胶囊网络是一种用于深度学习的高性能网络架构，见图2.2。在该构造中，每一特性都采用了一个矢量（也就是一个胶囊）。通过对这种结构进行训练，能够有效地进行图像的特征抽取。每个空间的特性都是一个矢量，并用聚类法将这些特性逐渐结合起来。简而言之，胶囊网络不是神经元而是由胶囊组成的。胶囊就是一组能够在图片的特殊部位，比如长方形的物体。其结果为矢量（比如8维矢量）。每一矢量的长短代表了一个对象的存在的估算可能性，该可能性被用来记录对象的位置参数（诸如精确位置、旋转等）。当对象稍微改变（例如移动，旋转，大小等）时，该容器将会产生一个具有同样的长度但稍微改变的向量值。因此，胶囊是等变的。

图2.2 胶囊神经网络结构图

CapsNet和常规的神经网络一样，也是一个多层次的结构。下面的胶囊叫做矢量包。他们各自利用图片中很少的一部分来作为一个被称作“知觉区域”的输入，并试图探测出某一特殊模式（比如长方形）的出现及其构造。在更高的层次上，被称作“路由胶囊”的空间可以被用来探测更大更复杂的目标，例如船舶。

## **Technology differences**

卷积网络Vgg由卷积层、池化层、激活层、全连接层、输出层等构成，与多种不同的基础神经网络类似，其基础构造比较单一。

随着网络层数的加深，网络的表达能力会更强，这是因为卷积核的作用是提取图像的特征，然而一个卷积核是不够的，一个卷积核只能反应图像的某一个特征，所以我们需要多个卷积核，这些不同的卷积核可以提取到图像不同的特征，从而让我们的模型学习图像特征的能力更强。因此有足够的卷积核和足够的参数才可以更好表述原始图像的特征。Resnet可以很好的解决梯度消失或者爆炸、退化问题。

# **Emotion Recognition Dataset**

系统在Fer2013数据集"对人脸表情进行7分类实验,且在CK+数据集上也进行了7分类的实验。

Fer2013数据集总共有35886张人脸表情图片.其中,训练集28708张,验证集和测试集都是3589张。每张图片都是48\*48的灰度图，总共包含有7类表情，分别为愤怒.厌恶.恐惧、快乐、悲伤、惊讶、平静。见表3-1。

CK+数据集是123名测试者录制的593个表情序列，带有标签的序列只有327个,且每一个表情都只对应一个标签,总共包含7类表情，分别为愤怒。蔑视.快乐，悲伤.恐惧.厌恶、惊讶。见表3-2。

表3-1 FER2013原始数据集数据配置

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 生气 | 厌恶 | 恐惧 | 开心 | 伤心 | 惊讶 | 中性 | 总数 |
| 标签 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | - |
| 训练集 | 705 | 717 | 472 | 281 | 182 | 1290 | 254 | 3589 |
| 测试集 | 162 | 160 | 1185 | 74 | 478 | 329 | 680 | 28708 |
| 总数 | 867 | 877 | 5957 | 355 | 2460 | 1619 | 3204 | 35886 |

表3-2 CK+原始数据集数据配置

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 愤怒 | 厌恶 | 高兴 | 恐惧 | 悲伤 | 惊讶 | 轻蔑 | 总数 |
| 标签 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | - |
| 训练集 | 1230 | 1590 | 1860 | 660 | 750 | 880 | 480 | 8820 |
| 测试集 | 120 | 180 | 210 | 90 | 90 | 60 | 60 | 990 |
| 总数 | 1350 | 1770 | 2070 | 750 | 840 | 540 | 540 | 9810 |

# **Experimental Analysis**

本文将预处理后的fer2013数据集与CK+数据集分别放入VGG、ResNet、Capsule模型中进行训练和测试，可以从下表中看出，ResNet 在Fer2013上准确率比其他两种高，Capsule在CK+上准确率比其他两种高。以此数据可以说明在面对庞大数据量时ResNet模型训练后对测试集的准确率比较高，在对一定量数据时Capsule模型训练后对测试集的准确率最高，VGG模型的训练效果都不太理想。说明了ResNet模型和Capsule模型的训练收敛更快，模型速度也更快，模型正确率更加稳定，一定程度上提升了识别的准确率。

|  |  |  |
| --- | --- | --- |
| 数据集 | 模型 | 准确率% |
| Fer2013 | VGG | 62.39% |
| Fer2013 | ResNet | 89.01% |
| Fer2013 | Capsule | 60.41% |
| CK+ | VGG | 58.88% |
| CK+ | ResNet | 77.16% |
| CK+ | Capsule | 96.45% |

# **Conclusion**

本文用于表情识别实验，通过不同的三种模型对两个数据集进行测试，在进行训练的过程中，虽然各个参数的调整对最后的损失和准确率会有影响，但在训练时使用哪种神经网络的模型才是决定性因素，即使训练后的准确率不太理想，但还是可以发现ResNet和Capsule模型可以更加稳定的进行识别。在训练期间，可以迅速地进行训练，并且变得更为平稳。

# **References**

1. U. Panniello and M. Gorgoglione, “A contextual modeling approach to context-aware recommender systems,” in Proc. the 3rd Workshop on Context-Aware Recommender Systems, October 2013.
2. A. Rajaraman, J. Leskovec, and J. D. Ullman, “Recommendation systems,” in Mining of Massive Datasets, Cambridge University Press, 2012, pp. 303-323.
3. G. Badaro, H. Hajj, W. El-Hajj, and L. Nachman, “A hybrid approach with collaborative filtering for recommender systems,” in Proc. 9th International Wireless Communications and Mobile Computing Conference, 2013, pp. 349-354.
4. R. Burke, “Hybrid recommender systems: Survey and experiments,” User Modeling and User-Adapted Interaction, vol. 12, no. 4, pp. 331-370, November 2002.
5. J. Fan, W. Pan, and L. Jiang, “An improved collaborative filtering algorithm combining content-based algorithm and user activity,” in Proc. International Conference on Big Data and Smart Computing, 2014, pp. 88-91.
6. F. Fouss and M. Saerens, “Evaluating performance of recommender systems: An experimental comparison,” presented at the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, 2008.
7. H. Ji, J. Li, C. Ren, and M. He, “Hybrid collaborative filtering model for improved recommendation,” in Proc. IEEE International Conference on Service Operations and Logistics, and Informatics, 2013, pp. 142-145.
8. Roy Deepjyoti,Dutta Mala. An Improved Cat Swarm Search-Based Deep Ensemble Learning Model for Group Recommender Systems[J]. Journal of Information & Knowledge Management,2022,21(03).
9. Chen Yida,Qiu Xiaoyu,Ma Chuanjiang,Xu Yunfeng,Sun Yang. A recommender system fused with implicit social information through network representation learning[J]. Computers and Electrical Engineering,2022,100.
10. Fernández-García Antonio Jesús,Rodriguez-Echeverria Roberto,Preciado Juan Carlos,Perianez Jorge,Gutiérrez Juan D.. A hybrid multidimensional Recommender System for radio programs[J]. Expert Systems With Applications,2022,198.
11. Khoali Mohamed,Laaziz Yassin,Tali Abdelhak,Salaudeen Habeeb. A Survey of One Class E-Commerce Recommendation System Techniques[J]. Electronics,2022,11(6).
12. Horasan, Fahrettin. Latent Semantic Indexing-Based Hybrid Collaborative Filtering for Recommender Systems[J]. Arabian Journal for Science and Engineering,2022(prepublish).
13. Tavakoli Mohammadreza,Faraji Abdolali,Vrolijk Jarno,Molavi Mohammadreza,Mol Stefan T.,Kismihók Gábor. An AI-based open recommender system for personalized labor market driven education[J]. Advanced Engineering Informatics,2022,52.
14. Sumathy B.,Kumar Anand,Sungeetha D.,Hashmi Arshad,Saxena Ankur,Kumar Shukla Piyush,Nuagah Stephen Jeswinde. Machine Learning Technique to Detect and Classify Mental Illness on Social Media Using Lexicon-Based Recommender System[J]. Computational Intelligence and Neuroscience,2022,2022.