# 第二章 GIF相关特性及现有的情感分析技术

GIF 图像由一系列的视频帧组成，在社交网络中的 GIF图像有以下特点：

1. 语义信息覆盖范围广， GIF 图像之所以含有情感倾向是因为 GIF 的内容是多样的。例如我们看到美丽的夕阳会产生积极的情感，看见可爱的婴儿会产生积极的情感，看到可口的食物会产生积极的情感。产生情感的因素很多，我们需要一种方法将他们识别出来才能完成最终的情感判断。如下图是我们在在线 GIF 图像网站 Giphy 中爬取到的标注动词数量分布，可见GIF 的语义信息分布十分广泛。

图2.1 GIPHY 数据集动词频率表（部分） 按照出现次数降序排列

1. GIF 是由一串时间上连续的视频帧组成的，情感的检测不是对孤立的视频帧进行的，仅仅考虑视频帧中存在哪些语义信息是不够的，还需要考虑视频帧之间的时序关系。

静态图像情感分析技术

目前现有的图像情感分析技术主要有三种，第一种是使用复杂的模型（通常是深度神经网络）加上大量训练数据，期待机器学习到情感分析问题所需要的全部中层表示，第二种是使用简单的模型加上少量的训练数据，通过人工定义的中层表示来实现情感分类。第三种是直接使用底层特征，套上传统的分类器直接得到分类结果，因为情感分析问题十分复杂，直接使用这种方法效果不好。而第一第二种思路各有优缺点，第一种的优点是中层特征的抽取过程自动进行，无需人工干预，缺点在于情感分析是一个系统工程，需要大量的要想得到好的效果，需要大量的带标注训练数据和复杂的模型[32]。第二种思路是使用人工设计的中层特征，先在待分类的图像上检测出中层特征的存在，再根据中层特征得到最后的情感分类。这种思路的优势是模型简单，只需要单独训练每一个中层特征的检测器。训练过程简单并且可以并行化。这种思路的劣势在于需要设计一个覆盖面足够广并且有效的中层特征体系需要一些领域知识。

自动学习中层特征体系的图像情感分析技术

[32]文给出了一种自动学习图像情感的框架。 如下图所示，他们设计了一种渐进式优化的图像情感分析框架。

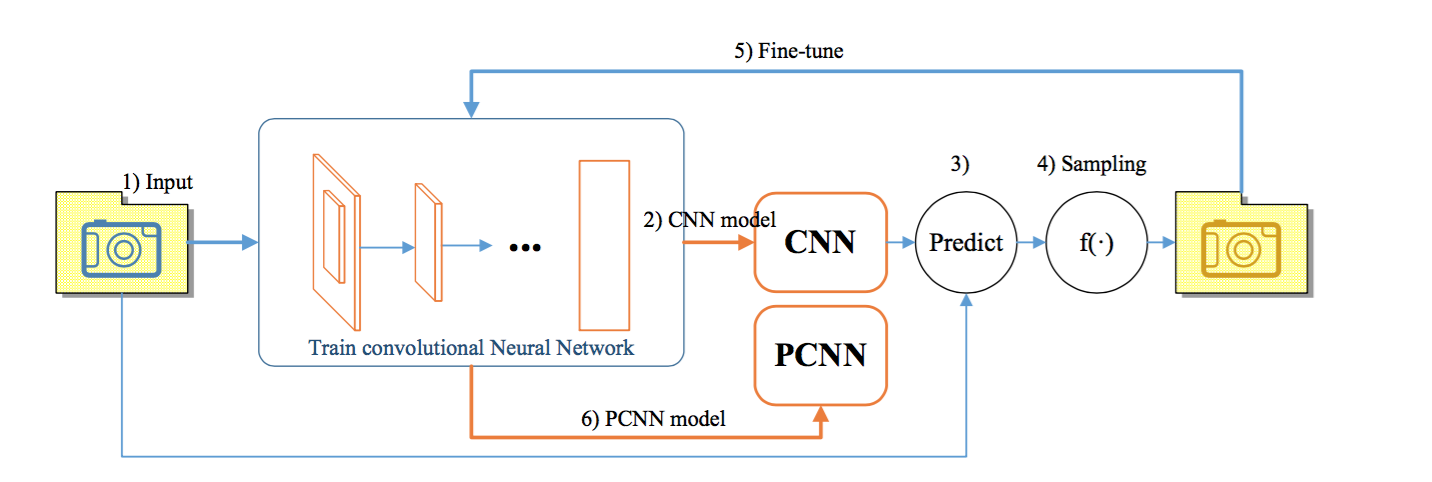


图2.2 [32]文设计的渐进式优化情感分析结构

这种框架在五十万张粗标注的 Flickr数据集上进行了训练，之后为了训练中层特征，使用了少量人工标注的 Twitter 数据。

如下图所示， 传统的用于图像情感分析的卷积神经网络结构

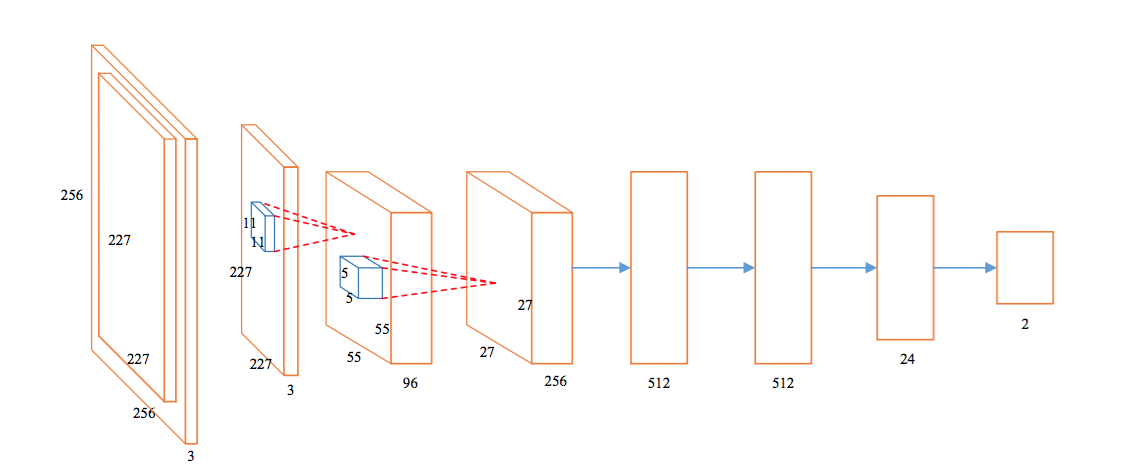


图2.3 常见的采用5个卷积层的卷积神经网络结构

在该结构中，每张待分类的图片被缩放为256\*256大小，然后中间227\*227的区域被裁剪出来输入到网络中。首先是两个卷积层，每个卷积层后都有一个 max-pooling 层。第一个卷积层有96个卷积核，每个卷积核的尺寸是11\*11\*3.卷积层进行卷积的步长是4.第二个卷积层有256个卷积核组成，每个卷积核的尺寸是5\*5，步长为2.

根据他们的结果

使用底层特征直接学习的情感分析技术

好的特征应具有不变性（旋转，尺寸，光照），和分类精度。在[27]文中，作者使用四种图像特征表示，它们分别是：

**颜色直方图：** 作者在 HSV 颜色空间内对每帧的颜色直方图进行计算。采用颜色直方图的原因是在以往的很多静态图片情感分析工作（例如[1][28]等等）中，颜色直方图都有应用。从宏观上来看一帧的颜色分布和情感倾向存在关联性。

**面部表情：** [1]文的作者认为，面部表情在 GIF 图片的情感分析中存在显著影响。作者使用的面部表情检测方法如下：构建一个输入层和隐含层都如图2.3所示的卷积神经网络，与图2.3不同之处在于，[1]文作者构建的卷积神经网络的输出层是一个具有平方合页损失函数的支持向量机模型。这个网络在28，708张48\*48的人脸图片训练集中进行训练，这些训练集分为七类，分别是：愤怒，恶心，恐惧，开心，沮丧，惊讶和中性。[1]文的作者使用 OpenCV提供的 基于Haar-like特征和AdaBoost方法的人脸检测工具进行人脸检测。用上文的方法进行面部表情特征的抽取结果是一个6维的向量。

**图片美学特征：** 前人的工作证明视觉材料的情感与图片美学特征之间存在着内在的联系[35]. [1] 文的作者提取了在[35]文中描述的图片美学特征的一个子集。每一个 GIF 帧被切分成3\*3的小单元。在每个小单元中提取暗通道，亮度，锐度，对称性，白平衡，色彩丰富度，色彩和谐度以及目光敏感度等特征。正规化后的帧中占主导地位的物体里图片中心的距离也被作为图片美学特征的一个方面提取.以上所有这些特征构成了一个149维的向量。GIF 图片的每帧需要提取一次。

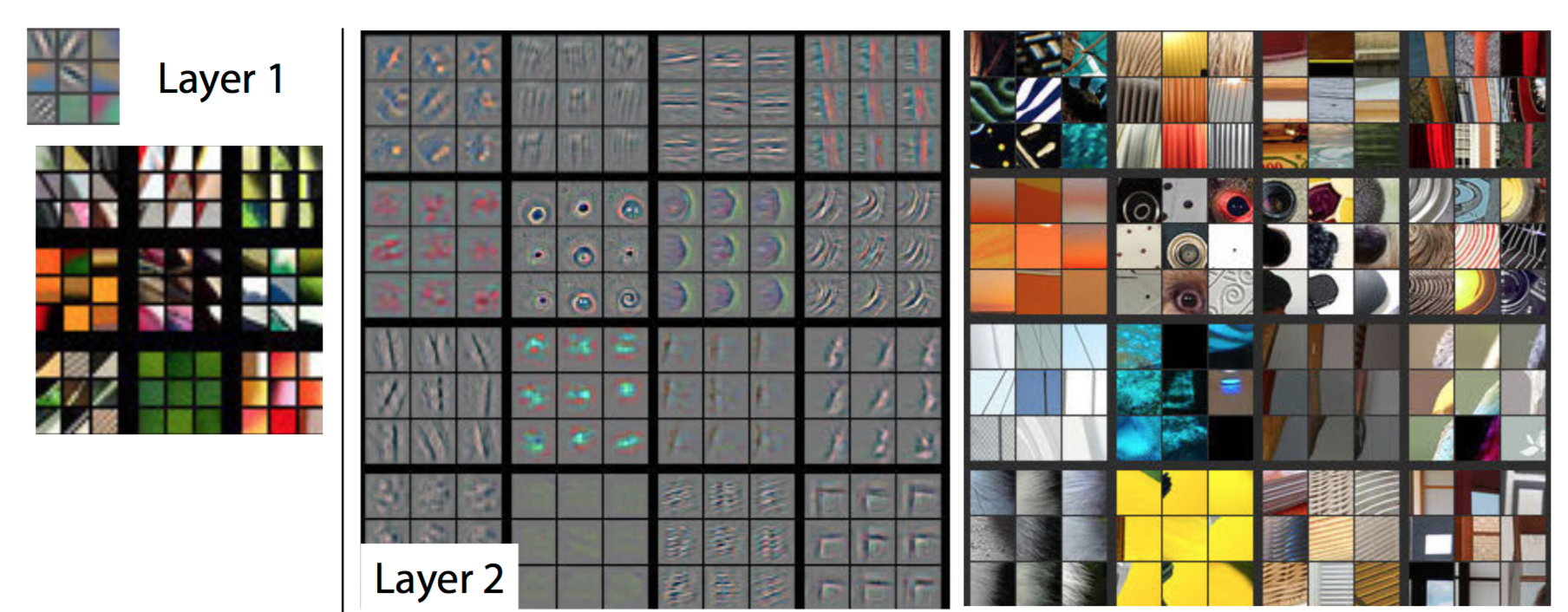
**SentiBank中层特征：** SentiBank中层是由 Borth[28]等人提出的一个针对静态图像情感分析的中层特征，[1] 文的作者直接将静态图像情感分析的特征用于 GIF 情感分析。SentiBank 中层特征由1200维向量构成，该向量的每一个维度都是一个线性支持向量机的分类结果，表示的是该图像包含某一个“形容词名词对”概念的可能性。SentiBank 特征在[36][37]等情感分析任务中都表现的很好。根据[36]文，SentiBank 特征在受控测试集上识别形容词名词对的 F 值已经超过了0.6。

现有方法的主要问题

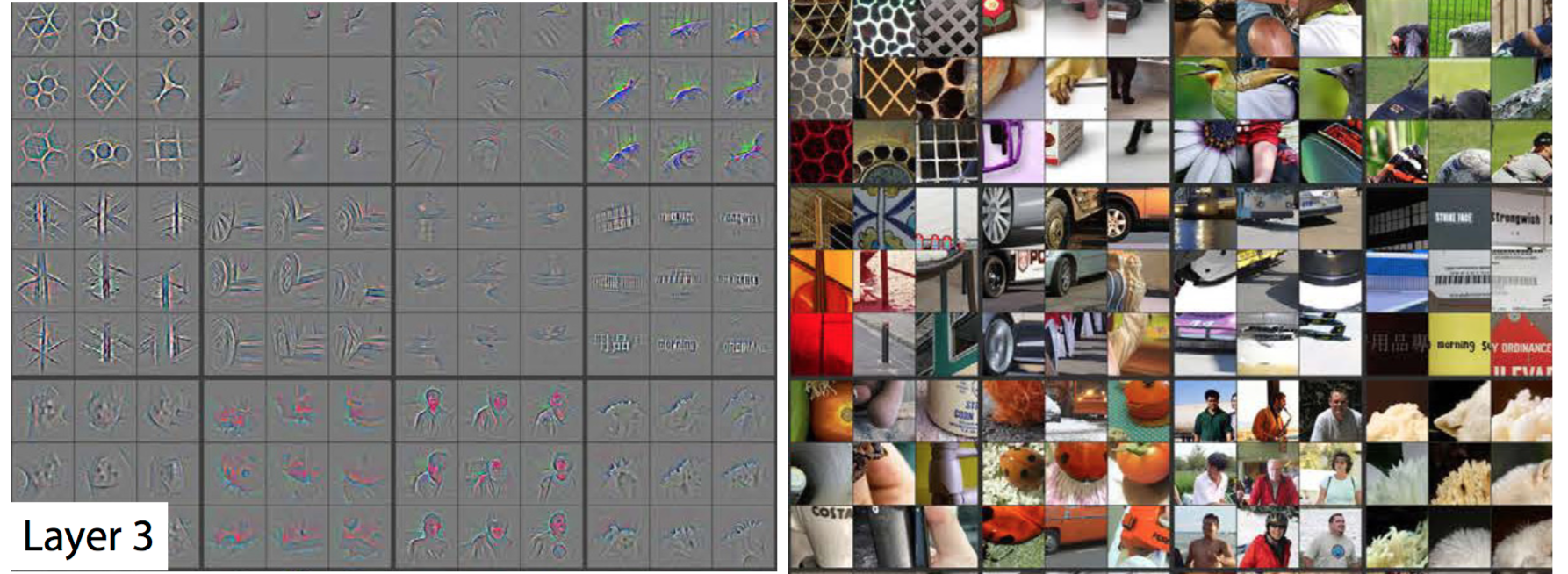
以上我们介绍了两种现有的视觉情感分析技术，

第一种技术基于深度神经网络（卷积神经网络）。希望能够通过深度神经网络自动学习特征的特点自动学习情感分析需要的中层特征。例如 GIF 视频中包含小孩摔倒，暴风雨天气等概念。该方法希望深度神经网络能通过训练数据学习到这些概念。然而，根据文献[38]如图2.4，图2.5，图2.6所示，目前五层卷积层能学到的还只是图片的局部，并不能学到类似于情感倾向等复杂的语义概念。从想要学习到情感倾向性特征需要更加复杂的模型和更多的训练数据。

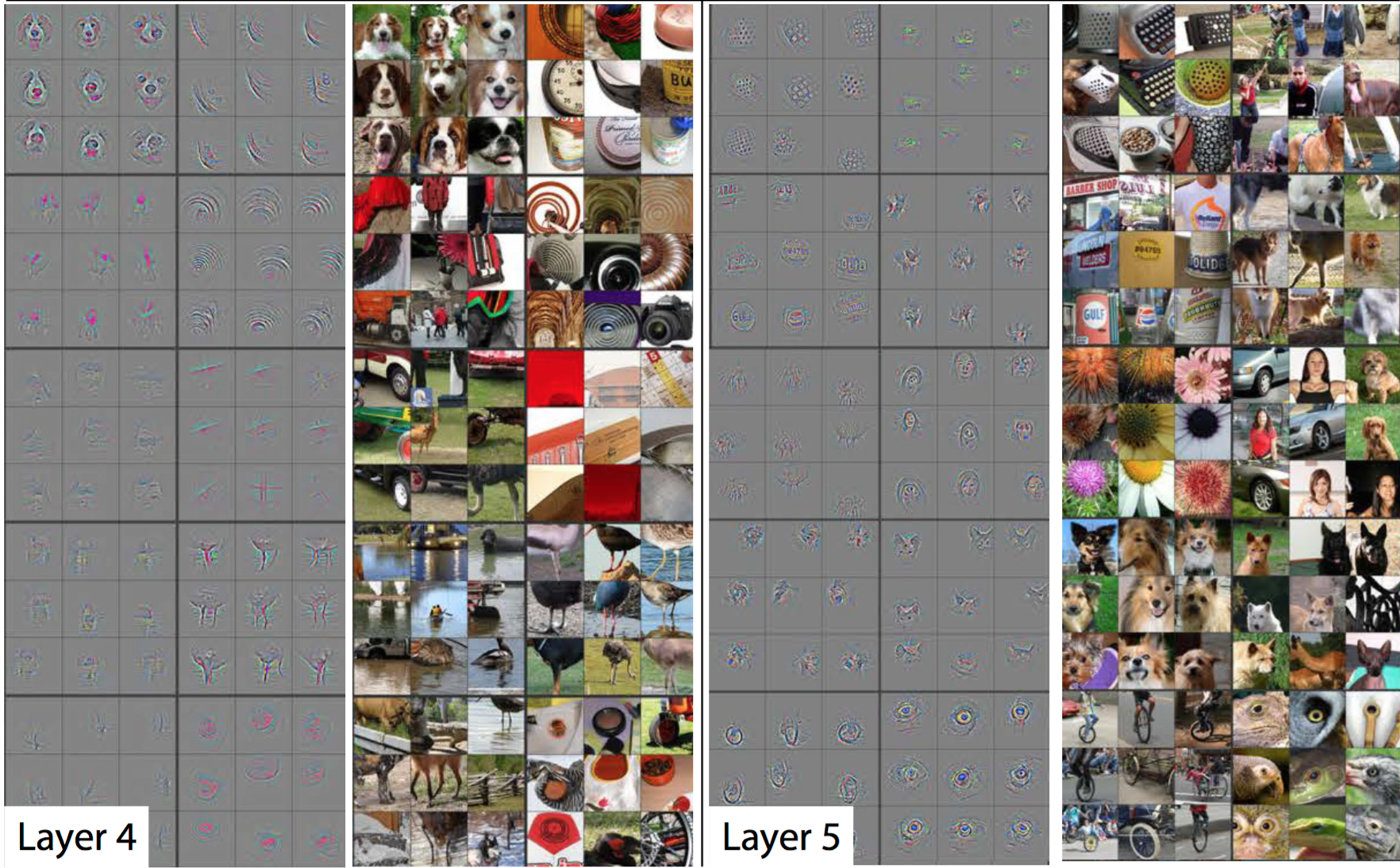
然而，即使是如图2.3所示的常见的采用五个卷积层的卷积神经网络在训练时也会遇到标注训练数据不足的问题。如果应用到 GIF 情感分析问题上，增加网络中隐含结点的数量势必需要更多的训练数据。更复杂的模型训练过程也会更加容易出现问题。但是，虽然现有的卷积神经网络无法可靠学习到 GIF 视频的所有中层特征以及中层特征到情感倾向的关系。如果设计好了可靠的中层特征，我们依然可以使用卷积神经网络来从 GIF 视频中检测它们。具体细节我们将在第三章中探讨。此外，这种方法并没有考虑帧和帧之间的时序关系。



**图2.4 第一第二卷积层学习到的特征表示示例**

****

**图2.5 第三卷积层学习到的特征表示示例**

****

**图2.4 第四第五卷积层学习到的特征表示示例** (以上三张图片来源：Matthew D. Zeiler and Rob Fergus, Visualizing and Understanding Convolutional Networks,ECCV 2014, Part I, LNCS 8689, pp. 818–833, 2014)

第二种技术基于底层特征（颜色直方图，面部表情，美学特征）或静态图像的中层特征（SentiBank）直接套用常见分类器进行情感预测。对于底层特征，这些人工选择的底层特征虽然与情感倾向存在一定关联。但是对于内容包罗万象的 GIF 视频来说，单纯依靠底层特征肯定是不够的。对于SentiBank中层特征，虽然在静态图像的情感分析中表现不错，但是GIF 动画是视频帧的时间序列。具有诸如摔倒，跳跃，舞蹈等动作信息，仅仅依靠形容词名词对不能很好的进行描述。同时使用底层特征和中层特征组合进行描述会由于不同特征的维度相差太大导致分类器对某些特征“忽略”的问题。

由于形容词和名词对不能很好的描述 GIF 视频中出现的中层概念，我们提出了动词名次对的表示方法。我们将形容词名词对和动词名次对合称为情感对。使用情感对来描述 GIF 视频中出现的中层概念具有描述准确，可检测性好等特点。具体内容将在第三章中进行阐述。和第一种方法一样，这种方法也没有考虑 GIF 动画前后帧之间的时序关系。

本章小结

在本章中，我们首先分析了 GIF 情感分析任务的两个基本特点。即 GIF 动画中语义信息和概念的分布范围很广，同时GIF动画是一个视频帧的时间序列。然后介绍了两种目前常见的视觉情感分析方法，第一种方法使用深度学习直接学习情感分类，然而，由于情感分析问题比较复杂，使用7层网络的效果并不能直接学到。同时卷积神经网络也不能处理时序信息。第二种方法采用颜色直方图，面部表情，美学特征等底层特征联合图像情感分析领域的“形容词名词对”中层特征。然而，这些特征并不足以描述 GIF 视频中分布广泛的语义和概念。不满足第一点要求，该方法也没有考虑时序。在这个基础上，我们将在第三章介绍我们的方法，通过构建 GIF 情感语义体系结构来解决第一个问题，通过构建情感对序列来解决第二个问题。

参考文献

1. Brendan J, Subhabrata B, Shih-Fu C. Predicting Viewer Perceived Emotions in Animated GIFs, ACM MM’14, November 3-7, 2014, Orlando, Florida, USA
2. Jianbo Yuan, Quanzeng You, Sean McDonough, and Jiebo Luo, Sentribute: Image Sentiment Analysis from a Mid-level Perspective, ACM SIGKDD, Workshop on Issues of Sentiment Discovery and Opinion Mining (WISDOM), Pages: 10:1--10:8, 2013
3. 赵妍妍, 秦兵, 刘挺. 文本情感分析[J]. 软件学报, 2010, 21(8): 1834-1848.
4. Feldman R. Techniques and applications for sentiment analysis[J]. Communications of the ACM, 2013, 56(4): 82-89.
5. Riloff E, Shepherd J. A corpus-based approach for building semantic lexicons[J]. arXiv preprint cmp-lg/9706013, 1997.
6. Hatzivassiloglou V, McKeown K R. Predicting the semantic orientation of adjectives[C]. Proceedings of the 35th annual meeting of the association for computational linguistics and eighth conference of the european chapter of the association for computational linguistics. Association for Computational Linguistics, 1997: 174-181.
7. Esuli A, Sebastiani F. Sentiwordnet: A publicly available lexical resource for opinion mining[C]. Proceedings of LREC. 2006, 6: 417-422.
8. 朱嫣岚, 闵锦, 周雅倩, 等. 基于 HowNet 的词汇语义倾向计算 [J]. 中文信息学报, 2006, 20(1): 14-20.
9. Turney P D. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews[C]. Proceedings of the 40th annual meeting on association for computational linguistics. Association for Computational Linguistics, 2002: 417-424.
10. Yu E, Kim Y, Kim N, et al. Predicting the Direction of the Stock Index by Using a Domain-Specific Sentiment Dictionary[J]. Journal of Intelligence and Information Systems, 2013, 19(1): 95-110.
11. Pang B, Lee L, Vaithyanathan S. Thumbs up?: sentiment classification using machine learning techniques[C]. Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10. Association for Computational Linguistics, 2002: 79-86.
12. Pang B, Lee L, Vaithyanathan S. Thumbs up?: sentiment classification using machine learning techniques[C]. Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10. Association for Computational Linguistics, 2002: 79-86.
13. Riloff E, Patwardhan S, Wiebe J. Feature subsumption for opinion analysis[C]. Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 2006: 440-448.
14. McDonald R, Hannan K, Neylon T, et al. Structured models for fine-to-coarse sentiment analysis[C]. Annual Meeting-Association For Computational Linguistics. 2007, 45(1): 432.
15. Nakagawa T, Inui K, Kurohashi S. Dependency tree-based sentiment classification using CRFs with hidden variables[C]. Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics. Association for Computational Linguistics, 2010: 786-794.
16. Deng Z H, Luo K H, Yu H L. A study of supervised term weighting scheme for sentiment analysis[J]. Expert Systems with Applications, 2014, 41(7): 3506-3513.
17. 唐慧丰, 谭松波, 程学旗. 基于监督学习的中文情感分类技术比较研究[J]. 中文信息学报, 2007, 21(6): 88-94.
18. Zheng L, Wang H, Gao S. Sentimental feature selection for sentiment analysis of Chinese online reviews[J]. International Journal of Machine Learning and Cybernetics, 2015: 1-10.
19. 陈俊杰, 张大炜, 李海芳. 融入模糊理论的 SVM 在图像情感识别中的应用研究[J]. 计算机科学, 2009, 36(8): 288-290.
20. Davidov D, Tsur O, Rappoport A. Enhanced sentiment learning using twitter hashtags and smileys[C]. Proceedings of the 23rd International Conference on Computational Linguistics: Posters. Association for Computational Linguistics, 2010: 241-249.
21. Barbosa L, Feng J. Robust sentiment detection on twitter from biased and noisy data[C]. Proceedings of the 23rd International Conference on Computational Linguistics: Posters. Association for Computational Linguistics, 2010: 36-44.
22. Go A, Bhayani R, Huang L. Twitter sentiment classification using distant supervision[J]. CS224N Project Report, Stanford, 2009: 1-12.
23. Bravo-Marquez F, Mendoza M, Poblete B. Combining strengths, emotions and polarities for boosting twitter sentiment analysis[C]. Proceedings of the Second International Workshop on Issues of Sentiment Discovery and Opinion Mining. ACM, 2013: 2.
24. 谢丽星, 周明, 孙茂松. 基于层次结构的多策略中文微博情感分析和特征抽取[J]. 中文信息学报, 2012, 26(1): 73-83.
25. Yan B, Zhang B, Su H Y, et al. Chinese microblog sentiment classification considering users’ reviews based on Naive Bayesian algorithm[J]. Future Communication, Information and Computer Science, 2015: 115.
26. Lu T J. Semi-supervised microblog sentiment analysis using social relation and text similarity[C]. Big Data and Smart Computing (BigComp), 2015 International Conference on. IEEE, 2015: 194-201.
27. Y.-Y.Chen,T.Chen,W.H.Hsu,H.-Y.M.Liao,andS.-F.Chang.Predictingviewer affective comments based on image content in social media. In ACM ICMR, 2014.
28. Borth, D.; Ji, R.; Chen, T.; Breuel, T.; and Chang, S.-F. 2013b. Large-scale visual sentiment ontology and detectors using adjective noun pairs. In ACM MM, 223– 232. ACM.
29. Cees G.M. Snoek , etc. Early versus Late Fusion in Semantic Video Analysis
30. J.MachajdikandA.Hanbury.Affectiveimageclassificationusingfeaturesinspired by psychology and art theory. In ACM Multimedia, 2010.
31. X.Wang, J.Jia, P.Hu, S.Wu, J.Tang, and L.Cai. Understanding the emotional impact of images. In ACM Multimedia, 2012.
32. Quanzeng You, Jiebo Luo, Hailin Jin, and Jianchao Yang, Robust Image Sentiment Analysis using Progressively Trained and Domain Transferred Deep Networks, the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI), Austin, TX, January 25-30, 2015.
33. George A. Miller (1995). WordNet: A Lexical Database for English. Communications of the ACM Vol. 38, No. 11: 39-41.
34. Christiane Fellbaum (1998, ed.) WordNet: An Electronic Lexical Database. Cambridge, MA: MIT
35. Bollen, J.; Mao, H.; and Pepe, A. 2011. Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In ICWSM.
36. Bollen, J.; Mao, H.; and Zeng, X. 2011. Twitter mood pre- dicts the stock market. Journal of Computational Science 2(1):1–8.
37. Borth, D.; Ji, R.; Chen, T.; Breuel, T.; and Chang, S.-F. 2013b. Large-scale visual sentiment ontology and detectors using adjective noun pairs. In ACM MM, 223–232. ACM.
38. Matthew D. Zeiler and Rob Fergus, Visualizing and Understanding Convolutional Networks, ECCV 2014, Part I, LNCS 8689, pp. 818–833, 2014.