第三章 GIF 情感分析概念语义体系构建

前文讲到，GIF 视频有两个特点， 第一，视频内容包罗万象。没有弄懂这些内容的语义之前无法分析 GIF 情感。第二，视频中各个概念的语义信息存在时序关系，概念的顺序会影响最终的情感判断。而从第二章的内容我们知道现有的 GIF 情感分析方法没有处理好这两个特点，导致分类效果不佳。因此，一个好的GIF 情感分析方法应该既能够考虑GIF 视频中包罗万象的概念语义，例如美丽的夕阳，跌倒的儿童，又能对概念之间的时序关系进行建模，例如哭泣的儿童在前，欢笑的儿童在后。由此，我们引出了本文的主要工作:

1. GIF 情感分析概念语义体系的构建
2. GIF 情感分析时序模型的设计

【此处应该有一张表现 GIF 视频中广泛分布的概念语义的图片】

**3.1 GIF情感分析概念语义体系的构建**

在这部分工作中，我们的目标是尽可能全面的概括在 GIF 视频中可能出现的导致情感倾向变化的概念语义。为了达到这个目标，我们首先要定义概念语义体系中概念的标准。

**3.1.1 概念语义体系的评价标准**

**情感相关性：**概念语义体系中的概念将被用于 GIF视频情感分析，因此只有和情感相关的概念才应该被纳入概念体系。

**广泛性：**前文提到，GIF 视频中的概念是包罗万象的，因此概念语义体系中的概念来源也应该来自于生活的方方面面。这样才能保证对任意的含有情感倾向的 GIF 视频，都能找到合适的概念来描述它。

**可检测性：**我们构建概念语义体系的目的是划定 GIF 情感识别中所需要的概念框架。因此概念语义体系中概念应该是具体的可检测的**。**

**3.1.2 概念语义体系的构建**

在定义了上面三个概念标准之后，从前人的工作[37]中我们得知，在静态图像情感分析领域，使用“形容词名词对”可以同时符合以上三个标准。但是，在动态视频情感分析领域，“形容词名词对”在广泛性上有所欠缺，这主要体现在 GIF视频中有很多动作信息，使用形容词名词对不能很好描述。



图3.2 ANP与 SentiPair 的比较

3.1.3 形容词名词对与动词名词对

为此，我们提出了动词名词对的概念。与[37]文中的形容词名词对类似，动词名词对包含一个动词和一个名词，它描述的是一个具体的动作和动作的对象。为了表述方便我们将形容词名词对和动词名次对合称为“情感对”



图3.3 情感对的构成

至此，我们有了比较好的表示 GIF 情感分析中概念语义的方法。但是还缺少层次体系结构。在自然语言处理领域，我们参考了[39][40]的工作.我们引入了WordNet作为我们概念语义体系结构的基础。WordNet是由普林斯顿大学提出的一个词典项目。与常见的词典项目不同，WordNet不仅把单词以字母顺序排列，而且按照单词的意义组成一个“单词的网络”。普林斯顿大学的心理学家，语言学家和计算机工程师们将 WordNet设计成一个覆盖范围宽广的英语词汇语义网。在这个语义网中包含了大量的概念语义。具有作为 GIF 情感分析概念语义体系的先决条件。

但是，仅仅使用 WordNet 的语义网络不能满足我们的要求。主要原因是在 GIF 情感分析的过程中，WordNet 中的很多词表示的概念是很少在 GIF 动画中出现的。因此，我们需要从 WordNet 的网络中筛选出在 GIF 动画中经常出现的概念。同时，构造“形容词名词对”，“动词名词对”的过程中，我们只需要 WordNet 中的形容词，动词和名词三种类型的词。其他的例如副词，介词，助词均需要删除。

3.1.4 无关概念语义的筛选

如上文所述，我们选择了 WordNet 作为概念语义体系的基础。在 WordNet 的基础之上，我们删除了数词，介词，副词等和“情感对”无关的概念语义词。同时，我们还需要从 WordNet 中筛选在 GIF 动画中经常出现的概念语义项。为了达到这个目标，我们爬取了 GIF 收藏网站 Giphy.com 中的 GIF 标注（tag）.如下图所示：

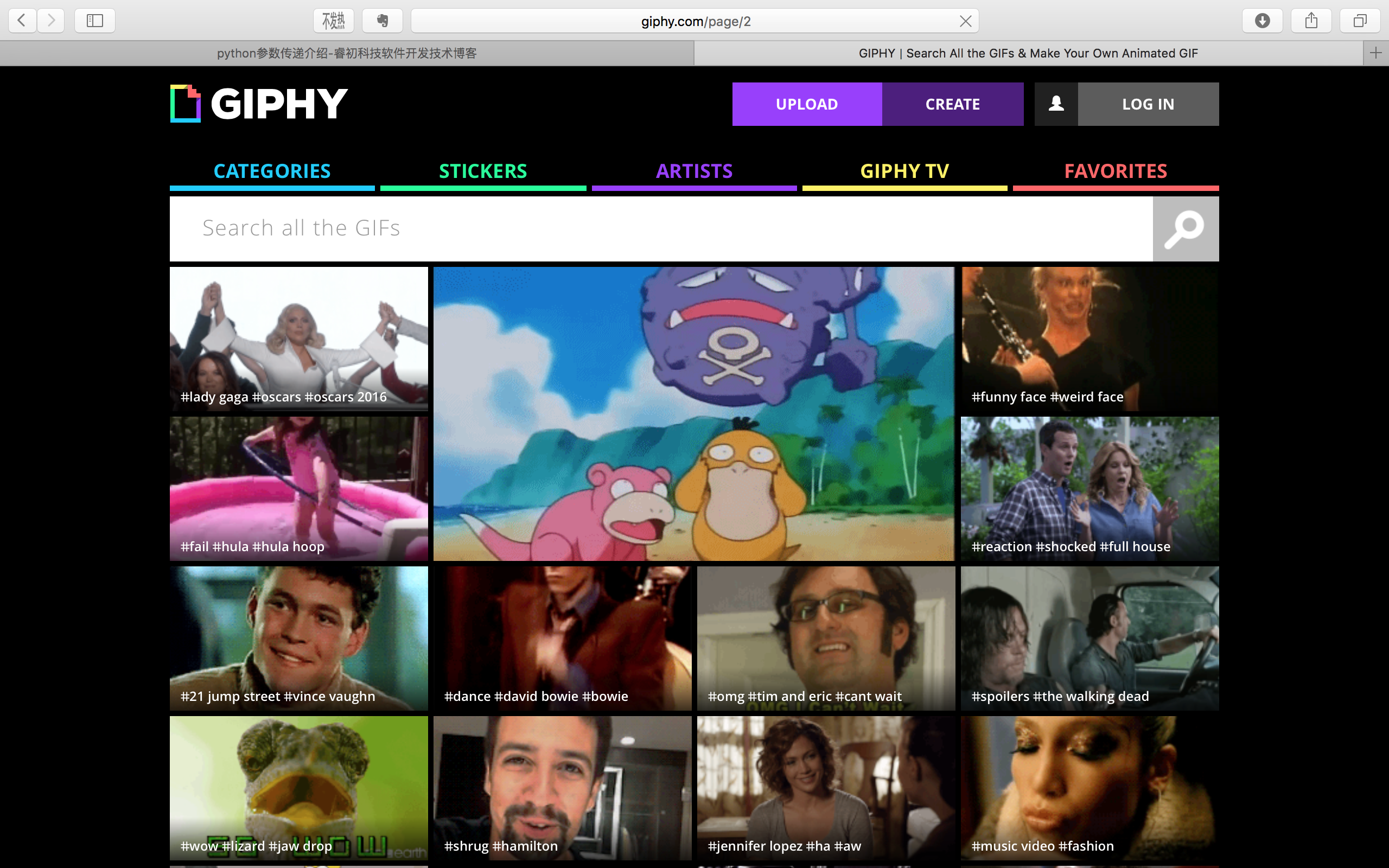


图3.4 Giphy.com 示例

如上图所示， Giphy.com 是一个 GIF 动画的收藏网站。根据网站介绍， 它包含超过十万张 GIF 动画，同时，对每张 GIF 动画，会有一个标签描述 GIF 动画的内容。例如，第一列第二张 GIF 动画上一个小女生在玩呼啦圈，在转过几圈之后呼啦圈掉在了地上，而这张 GIF 动画对应的标签有三个，第一是失败，第二和第三分别是呼啦圈的缩写和呼啦圈。我们使用爬虫技术采集了 Giphy.com 中的英文标签。因为这些标签都是描述 GIF 内容的，我们使用这个集合对 WordNet 进行筛选，保留在 Giphy.com 的标签中出现数量多的词语。

图3.5 Giphy.com 标签中概念语义词分布（筛选出10000次以上的词）

由于搜索技术的限制，我们删除了其中词语长度小于三的词，因为这些词会匹配到其他无关的语义项。至此，我们可以构建出 GIF 动画的概念语义体系。该体系的概括性由 WordNet 保证。 同时通过筛选步骤，该语义体系中都是和 GIF 动画相关的概念。

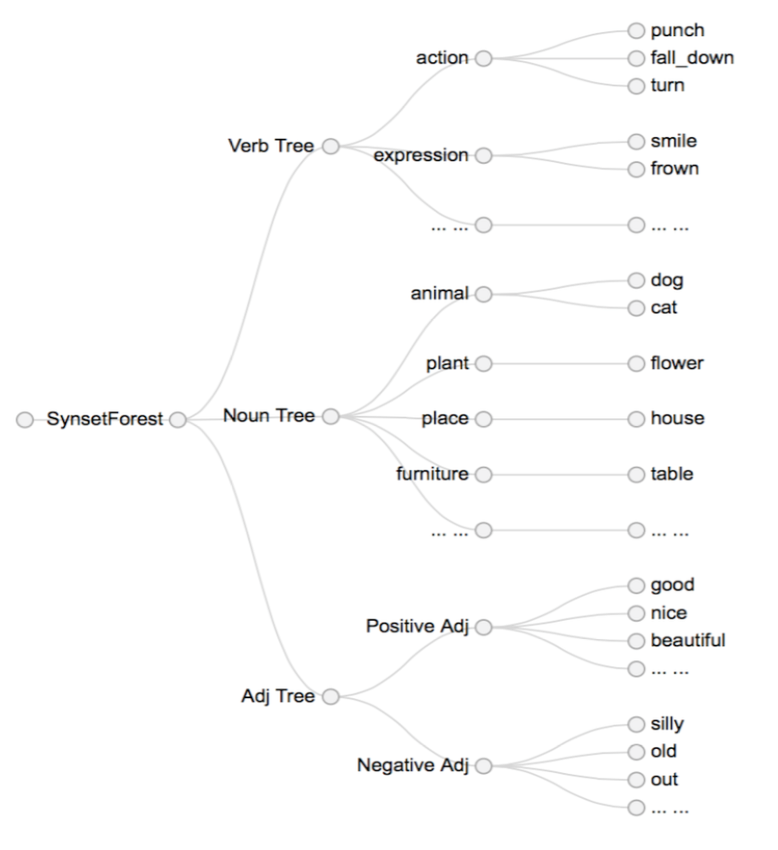


图3.6 GIF 情感分析概念语义体系示例

我们提出的GIF 情感分析感念语义体系是一个词的森林，该森林包括形容词树，名词树，动词树三棵树。在这三棵树的基础上，我们可以构建出形容词名词对和动词名词对的集合。

**3.2 情感对检测器的训练**

上文提到，GIF 情感分析的两个主要工作是：

1. 构建 GIF 情感分析概念语义体系。在该语义体系中规定 GIF 情感分析需要检测的概念。
2. 对 GIF 动画中包含的概念语义进行时序处理。

在构建好GIF 情感分析概念语义体系后， 我们要开始训练情感对的检测器。这些检测器的输入是 GIF 动画，输出是情感对的概率值. 在这些概率值将作为中层特征的特征值。

**3.2.1 多任务情感回归**

类似于逻辑回归和支持向量机的传统独立回归算法忽略了每个分类项目之间的关联性。例如，“惊讶的男孩”实际上也可以是“惊讶的人”，“晴朗的天空”也可以是“美丽的景色”。分类结果并不应该是一个一对多（one V.S. all）的结果. 因此，我们的情感对检测器应该是一个“多任务”回归问题。

**3.2.2 模型选择**

为了选择一个合适的模型进行情感对的检测，我们考察了多种方法。一种途径是类似于Borth [37]等人的方案，对每一个中层概念语义，也就是“情感对”训练一个支持向量机分类器。这种方法的优点是检测算法的鲁棒性取决于多个独立的检测器，某一个检测器性能不佳并不会造成全局的问题。但是缺点也很明显，由于我们的中层特征维数很高，需要训练大量的中层检测器，运行效率也会受影响。

受到深度学习尤其是卷积神经网络在视觉识别方面的巨大成功的启发[41]，我们选择了卷积神经网络作为我们情感对检测器的检测模型。正如前文所述，深度学习不适合用于直接学习 GIF 视频的情感倾向，因为他不能学习到情感分析所需要的全部中层特征，但是七层卷积网络学习视觉底层特征的能力已经被大量研究证实[41][42][43]。因此我们使用卷积神经网络来进行中层特征分类器的训练。所用网络结构如下图所示：

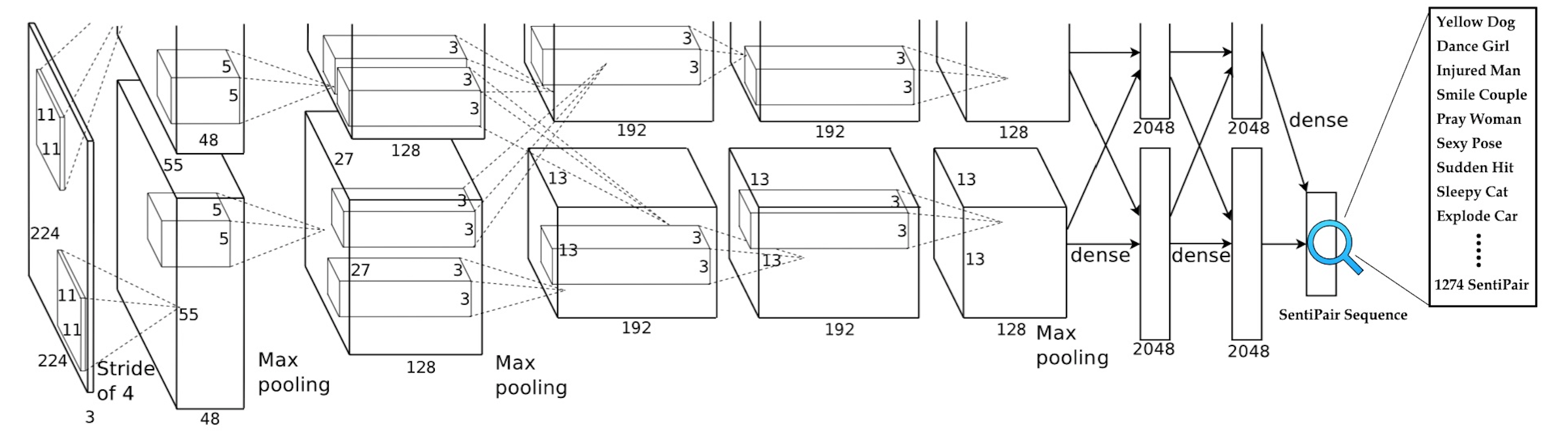


图3.7 我们所用的网络结构

第一个卷积层含有96个卷积核，每个卷积核的尺寸为11\*11\*3，扫描的步长为4像素。第二个卷积层的输入是第一层的输出，拥有256个5\*5\*48的卷积核，第三第四第五个卷积层都是互相直接连接，之间没有加入 Max Pooling 或正则化。 第三个卷积层有384个卷积核，每个卷积核的尺寸为3\*3\*256.第四个卷积层有384个卷积核，每个尺寸为3\*3\*192，第五个卷积层有256个卷积核，每个尺寸为3\*3\*192，两个全连接层拥有4096个神经元。

如上图所示，我们主要对传统卷积神经网络的最后一层做了修改。将原先的Softmax 层改为了 Multi-Task 约束。网络的输入时150，528维，每一层的神经元个数为：253，440-186，624-64，896-64，896-43，264-4096-4096-1274。 最后一层的个数为1274是因为我们在检测器的训练过程中选择了1274个权值最高的情感对。

【此处应该加上一张神经元训练过程中 Cost下降速率和准确率提升速率曲线】

情感对检测器输出结果：

【此处配合一张情感对检测器检测出的图片结果】

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.
39. George A. Miller (1995). WordNet: A Lexical Database for English. Communications of the ACM Vol. 38, No. 11: 39-41.
40. Christiane Fellbaum (1998, ed.) WordNet: An Electronic Lexical Database. Cambridge, MA: MIT Press.
41. Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks
42. Jeff Donahue, Yangqing Jia, Oriol Vinyals, Eric Tzeng. DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition
43. Andrej Karpathy, George Toderici, Sanketh Shetty. Large-scale Video Classification with Convolutional Neural Networks