第三章 GIF 情感分析概念语义体系

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. 构建了GIF 情感分析概念语义体系
2. 设计了GIF 情感分析时序模型

**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，受到 Borth.等人工作[37]的启发，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。如下图所示，图的左边部分是我们提出的情感对的检测结果，右侧是使用现有的形容词名词对的检测结果。绿色方框中的图片用情感对可以准确描述，却不能用形容词名词对准确描述。例如第二排第三张图片中“狂奔的蜗牛”。需要动作信息才能表现视频的内容。红色圆圈中的图片为我们分类器错误分类的样本，第三行第一列那个样本对应的“情感对”为“许多人”，“行走的人”。第三行第五列中只用“形容词名词对”描述时，只有“许多人”的信息。因为信息的缺失导致分类出错。



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。

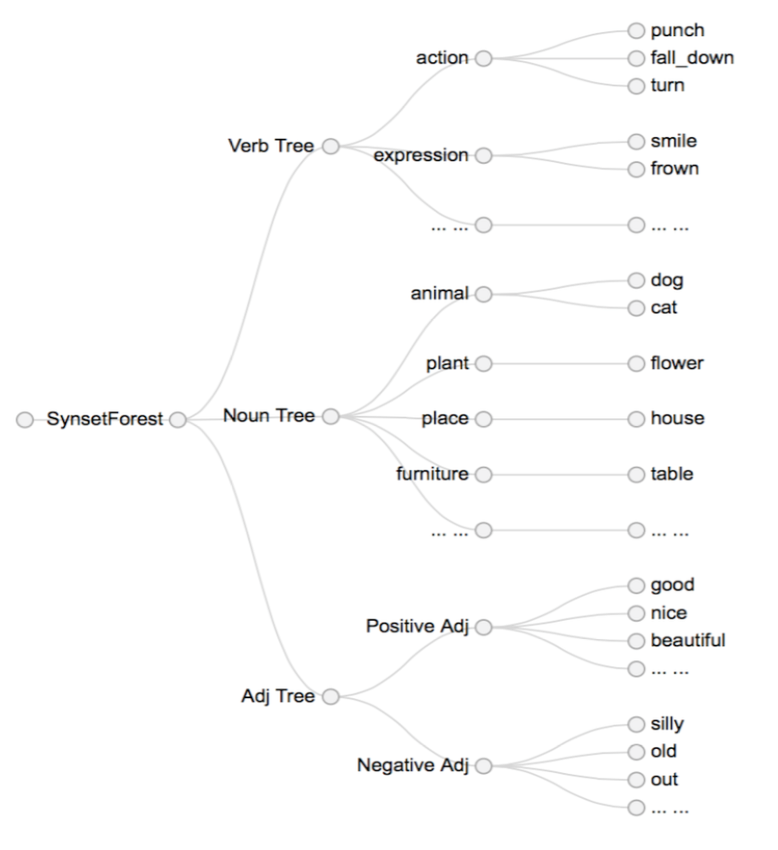


图 构建出的概念语义体系示意图

**3.2 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。

筛选的目标有两个：

一、找出情感倾向性强的概念语义。在概念语义体系中不同的概念语义所含有的情感倾向是不同的，因此，我们需要筛选出那些情感倾向性强的概念语义。

二、找出 GIF 视频中经常出现的概念语义。因为那些很少出现在 GIF 视频中的概念语义项目会增加概念语义的检测难度，并且这些项目对 GIF 情感分析没有什么贡献。

目前条件下，我们无法对概念语义体系中所有的概念语义进行检测，我们的实验也证明，不同筛选方案会对情感对检测的准确率产生很大影响。

**SentiWeight情感权值**

为满足第一个筛选目标，我们设计了 SentiWeight情感权值。构建过程如下：在 SentiWordNet中，每个词的情感倾向性(SentiScore)分为若干个等级，如下图所示。正数表示积极情感倾向，负数表示消极倾向。SentiWordNet的情感权值绝对值越大表示情感倾向越强烈。

SentiScore 就是 SentiWordNet 中该概念语义的情感得分，显然，SentiWeight 的取值范围是[0,1]。

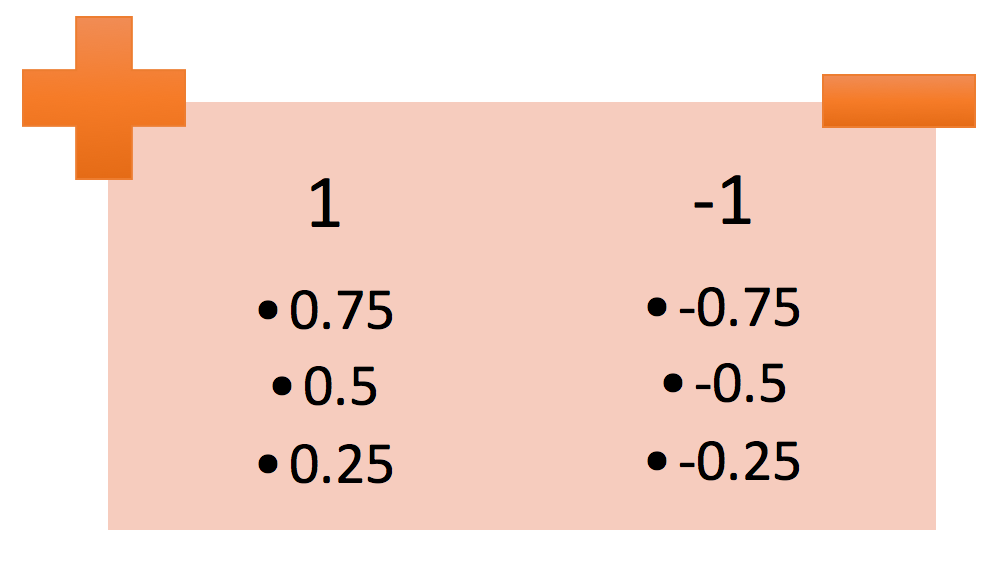


图 SentiWordNet 中的情感权值

**GiphyWeight 情感权值**

为满足第二个筛选目标，我们设计了 GiphyWeight 情感权值。构建过程如下：我们在GIF 视频网站Giphy.com 中搜索情感词，统计 Giphy.com 搜索结果中 GIF 图像的数量Count, 每一个概念语义的GiphyWeight 权值依据下式得出：

在上式中，Counti 是第 i 个概念语义在 Giphy.com 中对应的 GIF 动画个数。分母则是全部概念语义对应的 GIF 动画个数的最大值。因此，GiphyWeight 情感权值的取值范围是[0,1]。为了得到 Counti 的数据，我们爬取了 GIF 收藏网站 Giphy.com 中的 GIF 标注（tag）.如下图所示：

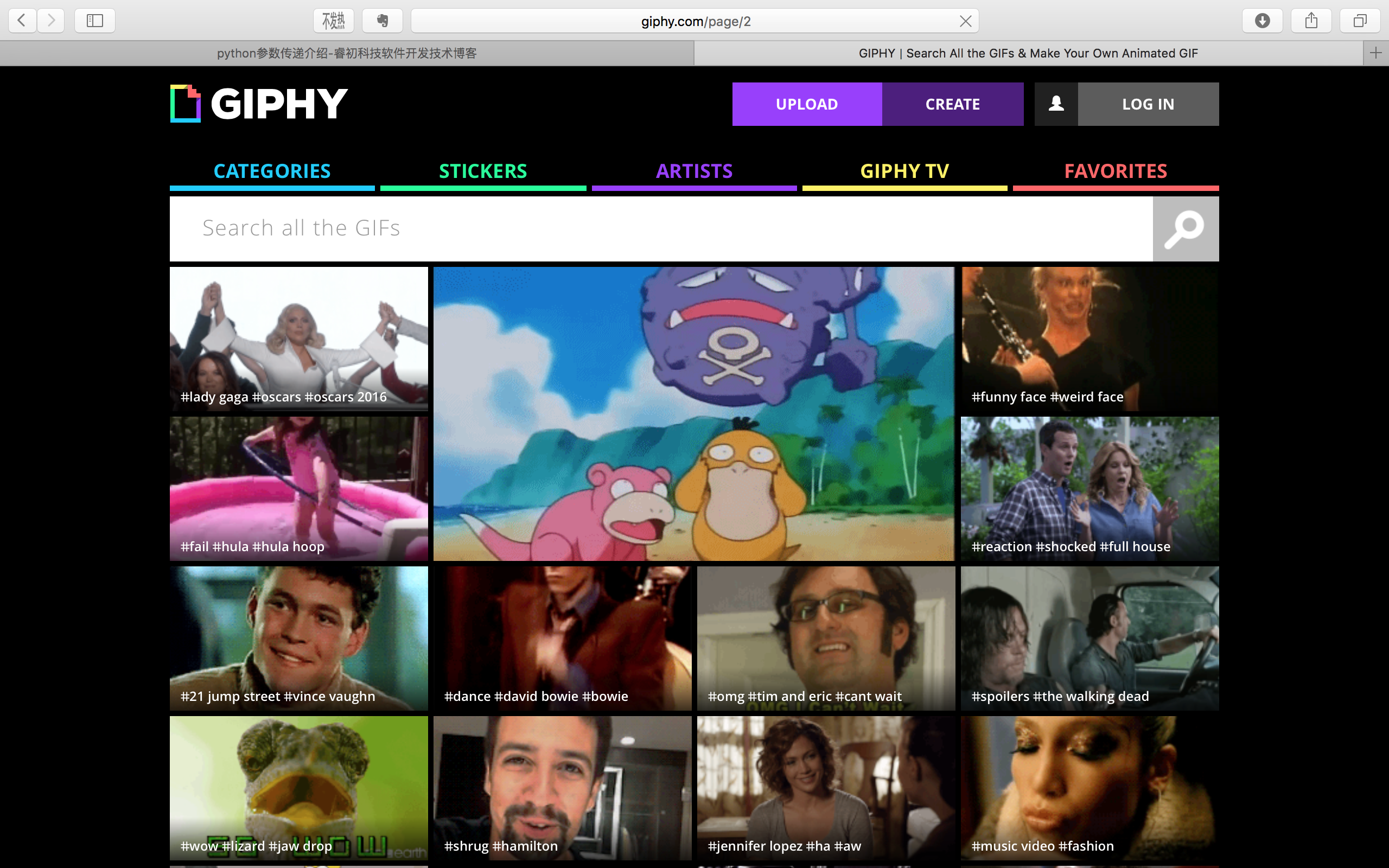


图3.4 Giphy.com 示例

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3的词，因为这些词会匹配到其他无关的语义项。

**筛选权值 FilterWeight**

在得到 SentiWeight 情感权值和 GiphyWeight 情感权值后，我们根据下式计算出一个筛选权值 FilterWeight:

显然，由于 SentiWeight情感权值和 GiphyWeight 情感权值的取值范围均为[0,1], 筛选权值 FilterWeight 的取值范围为[0,1]. 由于 SentiWeight 情感权值描述的是概念语义的情感丰富程度，GiphyWeight 情感权值描述的是概念语义在 GIF 视频中分布的广泛程度，按照上式得出的筛选权值既能描述概念语义的丰富程度，又能描述概念语义在GIF 视频中出现的概率。我们对筛选权值定义了一组阈值，下表展示的是不同阈值下筛选出的 SentiPair 数。

表 不同阈值下筛选出的“情感对”数

|  |  |
| --- | --- |
| 阈值 | 筛选出的“情感对”数 |
| > 0.7 | 31463 |
| > 0.8 | 5111 |
| > 0.9 | 1274 |

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。

**3.3 概念语义体系用于挖掘情感相关性**

上文提到，我们提出的概念语义体系为 GIF 视频中出现的概念语义划定了范围。实际上，作为语义体系，概念语义体系的另一个功能是可以用于挖掘情感相关性。如下图所示。 GIF 视频内容是三只可爱的金毛犬，但是金毛犬作为犬类的一种，我们可以将“可爱的金毛犬”这个情感对扩展成概括性更好的“可爱的狗”和“可爱的动物”。这种高层抽象的能力是传统基于底层特征的情感分析方法所不具备的。我们在训练情感对检测器的时候使用了这种相关性。如果情感对检测器检测到了某一个概念语义，例如“可爱的金毛犬”，我们会对检测结果进行扩展，将“可爱的金毛犬”扩展成“可爱的动物”。



图 概念语义体系用于挖掘情感相关性的示例

**3.4 情感对检测器的训练**

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. 而这些概率值将作为中层特征的特征值。

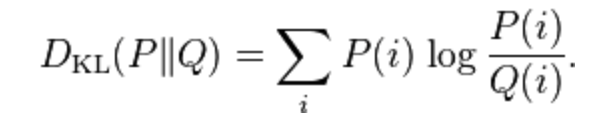
**3.4.1 多任务情感回归**

多任务学习又被称为多示例学习或多标签学习。传统的单标签分类学习是从一个只属于一个标签l的样本集合中学习，其中每一个标签属于一个互斥的标签集合L 。

在多标签分类中，每个样本属于一个L样本集合的一个子集。在过去，多标签分类由文本分类和医学分析而产生和推动的。现在，我们发现现代的许多应用对多标签分类方法需求持续增长，比如蛋白质分类，音乐归类，和语义场景分类。

传统单任务学习方法（SVM,LR 等等）忽略了每个分类项目之间的关联性。在 GIF 视频情感分析中, 检测到的概念语义存在关联性。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”。因此，我们的情感对检测器应该是一个“多任务”回归问题。

我们在多任务情感分析中使用的损失函数是交叉熵损失函数。使用 KL 距离来计算标签与分类结果之间的相似度。



对于两个离散型分布 P,Q，KL 距离可以用上式进行计算。

**3.4.2 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：

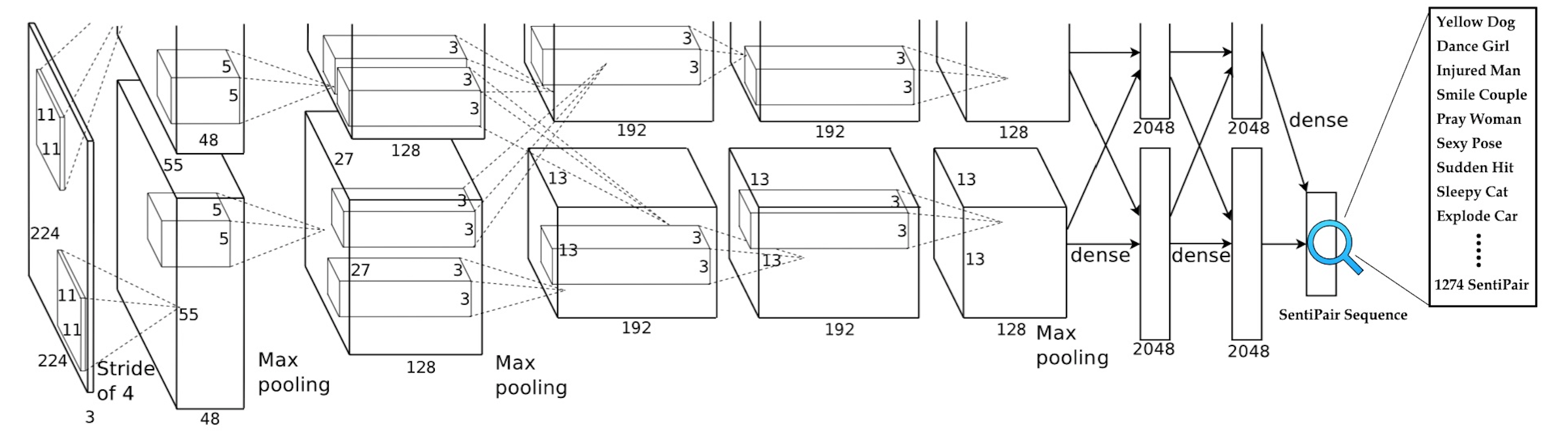


图3.7 我们所用的网络结构

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分别为：253，440-186，624-64，896-64，896-43，264-4096-4096-1274。

我们训练的情感对检测器的性能如下图所示，横坐标是不同的实验配置，纵坐标是情感对检测器的准确率。横坐标中 TopK 指的是在分类的结果中，分类置信度最高的前 K 个结果用于计算准确率。横坐标中 Relative 指的是使用我们的概念语义体系进行关联性挖掘。横坐标中 MTL 指的是使用多任务损失函数。在图中的三种颜色代表三种不同筛选结果。从左至右分别使用0.7，0.8和0.9作为筛选阈值。从实验结果中我们发现，大体上筛选阈值越高，分类精度越高，当然，筛选过多会造成中层特征的表达能力下降。此外，TopK 中，K 值越大，准确率越高。值得注意的是，对于低阈值的筛选结果，加入相关性挖掘后，检测子的准确率反而下降了，这可能是由于低阈值的类别数过多，分类准确率太低，相关性挖掘后扩展的词往往和真实的“情感对”相差甚远。

也就是说，加入 Multi-Task 后，对所有的分类类别数，情感对检测器的检测效果均有提高。其中在1274个情感对的情况下，提高的更明显（3%）。加入情感相关性之后，大部分分类类别数都可以提升效果，但是在分类类别数较高的情况下分类准确率反而下降。这是因为加入关联性后，扩展的概念语义是和分类结果直接相关的概念语义，如果本身的分类准确率不高，加入相关性约束后会适得其反。此外，不同筛选阈值的筛选结果对分类结果的影响也是显著的。

图 SentiPair 数目对检测器准确率的影响

**3.5 本章小结**

在本章中，我们致力于解决GIF视频情感分析中概念语义分布的广泛性问题。通过提出 GIF 情感分析概念语义体系，我们为GIF 情感分析中的概念语义划定了范围。该体系的完整度由已经被广泛验证过的 WordNet 保证。同时，由于 GIF 视频内容的特色，直接照搬 WordNet 体系结构显然是不合适的。需要对 WordNet 进行筛选。为此我们提出了两条筛选原则，一是概念语义要有足够的情感丰富性，二是概念语义在 GIF 视频中出现的频率要足够高。为了满足这两条原则，我们提出了 SentiWeight 情感权值和 GiphyWeight 情感权值。在概念语义体系的基础之上，我们使用深度神经网络训练了“情感对”检测器，实验表明，使用多任务学习并利用概念语义体系结构挖掘情感相关性后，检测性的准确率得到了大幅度提升。

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