

报告正文

参照以下提纲撰写，要求内容翔实、清晰，层次分明，标题突出。
请勿删除或改动下述提纲标题及括号中的文字。

(一) 立项依据与研究内容 (4000-8000 字):

1. 项目的立项依据 (研究意义、国内外研究现状及发展动态分析, 需结合科学研究发展趋势来论述科学意义; 或结合国民经济和社会发展中迫切需要解决的关键科技问题来论述其应用前景。附主要参考文献目录);

1.1 项目研究的意义

无人驾驶汽车、智能机器人等诸多人工智能前沿科技产品的迅速发展使海量、多变的图像和视频数据对目标分类与识别技术提出了更高的性能要求。形状是目标的基本特征。人类可以轻易根据形状来分类和识别目标, 然而这样的“简单”任务对基于形状目标分类与识别技术的产品来说仍然存在巨大的困难。根本原因在于“人”在对目标的形状识别时具有语义学习核判别能力。因此基于形状语义的目标分类与识别技术急需突破。

形状匹配和分类是基于形状的目标分类与识别的基本科学问题。申请人作为主研人员在执行国家自然科学基金项目“形状的语义结构表示及其分类学习研究(61173121)”过程中对形状几何变换带来的几何差异、目标的自遮挡与互遮挡、非刚性物体的运动、形状的类内变化、轮廓错误分割和噪声的干扰[1]等进行了探讨, 这些问题仍然是目前形状匹配和分类面临的重大技术挑战, 这些技术难点限制了基于形状的目标分类与识别的应用。上述问题并非局限于形状分析领域, 诸多计算机视觉任务如目标检测、运动跟踪、行人再识别等也面临类似甚至相同的问题。对于目标几何差异、遮挡、非刚性运动等同类形状几何变换多样性问题, 部分相关研究工作通过刻画目标的非刚性结构来避免这些不利因素的影响[2][3]。针对形状的类内变化, 比较流行的解决方案是利用深度神经网络学习高维非线性样本空间以线性区分各类别样本。此外, 深度神经网络对噪声和遮挡等不利因素也具有一定适应性。针对形状分析问题的特点, 本项目拟采用深度学习技术模拟人脑形状分析的认知机理, 探索具有非刚性结构保留特性的深度形状语义网络多层表示及其判别性等相关科学问题。研究结果一

方面可丰富和发展现有形状匹配及分类的理论，推动基于形状的目标分类与识别技术的进步，另一方面也为智能产品研发提供先进的技术支撑。

1.2 国内外研究现状分析

近年来，新的形状匹配及分类方法不断出现，并发表在计算机视觉领域的顶级期刊上，如 PAMI, TIP 等[4][5][6][7]。如图 1 所示，这些研究主要围绕形状表示、形状匹配、形状分类、度量学习等科学问题而展开。

1.2.1 形状表示

形状表示是提取具有区分能力形状特征，其区分能力强弱将直接影响后续应用的性能。现有的形状表示方法主要通过形状基本特性来人工设计形状描述器，如轮廓结构特征、轮廓的空间位置关系、轮廓的尺度空间特征等。根据基本特性不同形状表示可以分为三类。第一类是结构轮廓形状表示方法。这类方法将形状的轮廓看作若干基元的组合，使用链码[8][9][10]、边缘逼近、边缘分解[11][12][13][14]、句法分析等方法设计形状表示方法。由于特征设计的考虑了轮廓的局部性，使得这类方法可有效地处理目标的遮挡问题。第二类方法是利用空间位置关系来表示形状。这类方法通过统计轮廓序列上的点的空间位置分布关系来设计形状描述子，有效地整合了形状轮廓的全局信息和局部信息，具有良好的几何不变性和较强的形状描述能力。许多方法是从 Belongie[15]提出的形状上下文直接或间接发展而来，包括内距离形状上下文[16][17]，距离集[18]，弦上下文[19]，高函数[20]，轮廓点分布直方图[21]，信息特征[22][23]，距离内比[24]等方法。第三类是多尺度形状描述方法[25]。通过充分利用形状在不同尺度下的特性来刻画形状的凹凸特征，如曲率尺度空间[26]，多尺度凹凸表示[27]，三角形区域描述子[28][29]，视觉部分[30][31]，形状树[32]，多分辨率多边形[33]，多尺度距离矩阵[34]等等，它们很好地解决了轮廓噪声的影响。

上述形状表示方法具有诸如算法简单、计算复杂度低、应用灵活、针对性强等优点。如图 1 所示，它们存在两个问题。首先，这些方法提取的形状特征都是单层的，而单层特征只能保留形状的部分信息，与人脑对形状的多层认知机理有较大差别，这不利于充分使用形状携带的信息。其次，它们需要人工参与设计，涉及较多的领域知识(Domain knowledge)。当面对海量、多变的数据时，这在一定程度上限制了它们在实际项目中的应用。

1.2.2 形状匹配

形状匹配以形状表示作为输入，通过设计算法寻找形状之间的整体及局部的对应关系[1]，并达到度量形状相似度的目的。形状匹配的准确性决定了对形状间相似度的辨别能力。基于轮廓序列点集顺序关系的形状匹配算法是现在最常见的方法。按照是否知道轮廓序列点集顺序关系，这些形状匹配方法可以分为两类[1]。对于轮廓序列点集顺序关系已知的形状表示，通常采用动态规划(Dynamic programming)算法对形状距离求取全局最优解。而对于轮廓序列点集顺序关系未知的形状表示，通常使用优化方法，如 TPS-RPM，局部邻居结构，双向图匹配，或者概率方法，如产生式模型，层次 Procrustes 分析等[1]。值得注意的是，如图 1 所示，对于现有的形状匹配方法，形状匹配过程仅在已经得到人工设计的形状表示后进行。形状表示限制了形状匹配方法的效果，而形状匹配结果却不能反过来影响形状表示。

1.2.3 形状分类

形状分类也以形状表示作为输入，通过设计分类器对目标进行分类。现有的形状分类方法主要基于形状匹配。这类方法将形状间的匹配分数作为形状间的相似度，由此设计分类器对形状进行分类。除基于形状匹配的形状分类方法之外，还存在许多各具特色的方法。它们不再需要通过形状匹配方法取得形状间的对应关系。例如，利用隐马尔科夫模型对形状类进行建模[35]，将形状轮廓分解成相互允许重叠的轮廓片段并利用贝叶斯模型进行建模[36]，将轮廓片段和骨架路径转化为向量特征的形式，用高斯混合模型来建模[37]，使用空间金字塔匹配(Spatial Pyramid Matching, SPM)和最大池化(Max Pooling)方法构造形状中层表示[14]，构造二层的骨架联合树模型[38]等。

如图 1 所示，形状分类方法的分类器设计主要以浅层机器学习方法为主，如隐马尔科夫模型[35]，最近邻模型[39]，贝叶斯模型[36]，高斯混合模型[37]，支持向量机[14] [38]等。而浅层机器学习方法对具有复杂流形结构的高维特征空间的分类能力是有限的。此外，与形状匹配方法相同，现有形状分类方法使用的浅层机器学习方法独立于形状表示。形状表示限制了机器学习方法的应用，而机器学习方法并不能反过来帮助改善形状表示。

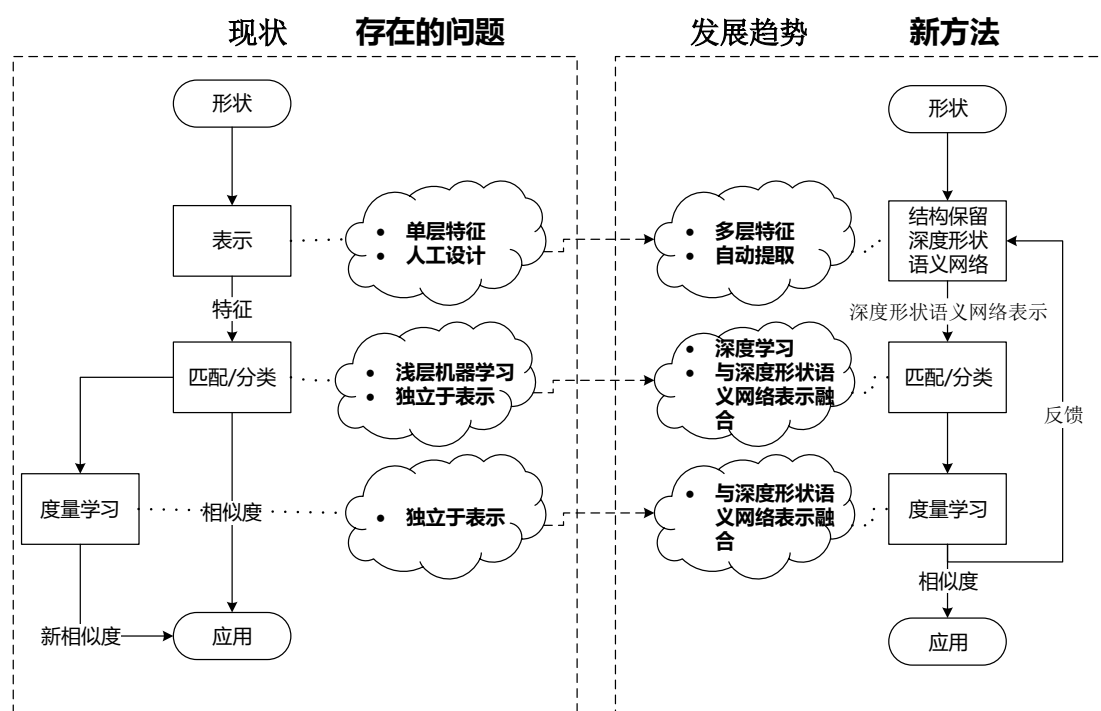


图 1 国内外研究现状分析及项目研究思路

1.2.4 度量学习

形状相似性度量形成的形状距离矩阵中存在着不满足三角形不等式的情况[1]。这是影响检索和分类精度的一个重要因素。现有形状检索和分类方法尝试挖掘形状相似性度量空间流形结构中潜在的信息，为每个形状找到最优检索路径。主要研究成果包括 Yang 提出的图转导算法(Graph transduction)[40]，局部约束传播过程(Locally constrained diffusion process, LCDP)[41]，Kontschieder 提出的相互 k-最近邻图 (Modied mutual k-NN graph)方法[42]，白翔提出的协同转导 (Co-transduction)方法[43]等。这些方法已经大幅提升形状检索的性能。然而，如图 1 所示，现有的度量学习方法独立于之前的形状表示、形状匹配、形状分类等步骤，局限于度量空间中挖掘潜在的流形结构，并未涉足匹配空间中的结构挖掘[1]。

1.3 存在的问题及发展动态分析

综上所述，通过对国内外研究现状的分析，可以看出形状表示、匹配及分类方法存在如下问题：在形状表示方面，以人工设计为主，提取的特征为单层特征；在形状匹配和分类方面，以浅层机器学习方法为主，学习过程独立于形状表示；在度量学习方面，度量学习独立于之前的形状表示及形状匹配和分类。

近年来，深度学习技术在各前沿领域得到迅猛发展。由于深度学习技术具

所示,从深度形状语义网络表示机理入手,利用深度学习技术结构灵活性的特点,模拟人类大脑形状认知过程(如图 2b),搭建具有结构保留能力的深度形状语义学习网络,统一形状的特征表示、分类及匹配各个步骤,建立形状表示、匹配及分类各步骤间的反馈机制,从而达到在全局上学习最优形状语义表示的目的。

形状分析中几何变换、目标的自遮挡与互遮挡、非刚性物体的运动、形状轮廓分割错误与噪声等常见问题可以归结为形状的非刚性变换问题。为此,本项目将形状分析问题转化为形状语言处理问题,探索深度形状语义网络的结构学习层的搭建与设置;通过学习视觉形状单词的非刚性结构关系,保留形状固有的结构特征,从而利用形状结构特征提升深度形状语义网络表示的鲁棒性,使深度形状语义网络表示能够解决形状匹配及分类的若干研究难点。针对形状类内变化问题,在结构学习层的基础上,构建并设置卷积神经网络层,对形状的结构信息进一步抽象,学习形状结构中的显著特征。同时在深度网络的决策层引入度量学习,融合形状的可利用信息,最大限度地探索深度形状语义网络表示的判别性。本项目的研究将开拓形状分析的新思路,为处理海量、多变的图像和视频数据提供技术支撑,提升以目标分类、识别为核心技术智能产品的形状感知能力。

主要参考文献

- [1] 周瑜,刘俊涛, and 白翔,“形状匹配方法研究与展望,”自动化学报, vol. 38, no. 6, pp. 889–910, 2012.
- [2] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, “Object Detection with Discriminatively Trained Part-Based Models,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1627–1645, Sep. 2010.
- [3] A. J. Rodríguez-Sánchez, G. L. Dudek, and J. K. Tsotsos, “Detecting, Representing and Attending to Visual Shape,” in *Shape Perception in Human and Computer Vision*, London: Springer London, 2013, pp. 429–442.
- [4] B.-W. Hong and S. Soatto, “Shape Matching Using Multiscale Integral Invariants,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 1, pp. 151–160, Jan. 2015.
- [5] X. Bai, C. Rao, and X. Wang, “Shape vocabulary: a robust and efficient shape

- representation for shape matching,” *IEEE Trans. Image Process.*, vol. 23, no. 9, pp. 3935–3949, Sep. 2014.
- [6] L. M. Lui, W. Zeng, S. Yau, and X. Gu, “Shape Analysis of Planar Multiply-Connected Objects Using Conformal Welding,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 7, pp. 1384–1401, Jul. 2014.
 - [7] H. Van Nguyen, F. Porikli, H. Van Nguyen, and S. Member, “Support Vector Shape: A Classifier-Based Shape Representation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 4, pp. 970–982, Apr. 2013.
 - [8] H. Freeman, “On the Encoding of Arbitrary Geometric Configurations,” *IEEE Trans. Electron. Comput.*, vol. EC-10, no. 2, pp. 260–268, Jun. 1961.
 - [9] H. Freeman and A. Saghri, “Generalized chain codes for planar curves,” in *in: Proceedings of the Fourth International Joint Conference on Pattern Recognition, Kyoto, Japan, November 7–10, 1978*, pp. 701–703.
 - [10] J. Iivarinen and A. Visa, “Shape Recognition of Irregular Objects,” in *Intelligent Robots and Computer Vision XV: Algorithms, Techniques, Active Vision, and Materials Handling, Proc. SPIE 2904*, 1996, pp. 25–32.
 - [11] W. I. Grosky and R. Mehrotra, “Index-based object recognition in pictorial data management,” *Comput. Vision, Graph. Image Process.*, vol. 52, no. 3, pp. 416–436, Dec. 1990.
 - [12] W. I. Grosky, P. Neo, and R. Mehrotra, “A pictorial index mechanism for model-based matching,” *Data Knowl. Eng.*, vol. 8, no. 4, pp. 309–327, Sep. 1992.
 - [13] R. Mehrotra and J. E. Gary, “Similar-shape retrieval in shape data management,” *Computer (Long. Beach. Calif.)*, vol. 28, no. 9, pp. 57–62, 1995.
 - [14] X. Wang, B. Feng, X. Bai, W. Liu, and L. Jan Latecki, “Bag of contour fragments for robust shape classification,” *Pattern Recognit.*, vol. 47, no. 6, pp. 2116–2125, Jun. 2014.
 - [15] S. Belongie, J. Malik, and J. Puzicha, “Shape matching and object recognition using shape contexts,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 4, pp. 509–522, Apr. 2002.

- [16] H. Ling and D. W. Jacobs, "Shape classification using the inner-distance.," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 2, pp. 286–299, Feb. 2007.
- [17] D. W. W. Jacobs, "Using the Inner-Distance for Classification of Articulated Shapes," in *2005 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2005, vol. 2, pp. 719–726.
- [18] C. Grigorescu, S. Member, and N. Petkov, "Distance sets for shape filters and shape recognition.," *IEEE Trans. Image Process.*, vol. 12, no. 10, pp. 1274–1286, Jan. 2003.
- [19] Y. Mingqiang, K. Kidiyo, and R. Joseph, "Shape Matching and Object Recognition Using Chord Contexts," in *2008 Int. Conf. Vis.*, 2008, no. 1, pp. 63–69.
- [20] J. Wang, X. Bai, X. You, W. Liu, and L. Jan, "Shape matching and classification using height functions," *PATTERN Recognit. Lett.*, vol. 33, no. 2, pp. 134–143, 2012.
- [21] X. Shu and X.-J. Wu, "A novel contour descriptor for 2D shape matching and its application to image retrieval," *Image Vis. Comput.*, vol. 29, no. 4, pp. 286–294, Mar. 2011.
- [22] Z. Tu and A. L. Yuille, "Shape Matching and Recognition - Using Generative Models and Informative Features," in *Eur. Conf. Comput. Vis.*, 2004, pp. 195–209.
- [23] Z. Tu, S. Zheng, and A. Yuille, "Shape matching and registration by data-driven EM," *Comput. Vis. Image Underst.*, vol. 109, no. 3, pp. 290–304, Mar. 2008.
- [24] N. Kaothanthong, J. Chun, and T. Tokuyama, "Distance Interior Ratio: a new shape signature for 2D shape retrieval," *Pattern Recognit. Lett.*, vol. 78, pp. 14–21, 2016.
- [25] D. Zhang and G. Lu, "Review of shape representation and description techniques," *Pattern Recognit.*, vol. 37, no. 1, pp. 1–19, Jan. 2004.
- [26] F. Mokhtarian, S. Abbasi, and J. Kittler, "Efficient and Robust Retrieval by

- Shape Content through Curvature Scale Space,” in *Efficient robust Retr. by shape content through curvature scale space.*, 1996, pp. 35–42.
- [27] T. Adamek and N. E. E. O’Connor, “A Multiscale Representation Method for Nonrigid Shapes With a Single Closed Contour,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 5, pp. 742–753, May 2004.
 - [28] N. Alajlan, I. El Rube, M. S. Kamel, and G. Freeman, “Shape retrieval using triangle-area representation and dynamic space warping,” *Pattern Recognit.*, vol. 40, no. 7, pp. 1911–1920, Jul. 2007.
 - [29] N. Alajlan, M. S. Kamel, and G. H. Freeman, “Geometry-based image retrieval in binary image databases,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 6, pp. 1003–1013, Jun. 2008.
 - [30] L. J. J. Latecki and R. Lakamper, “Shape similarity measure based on correspondence of visual parts,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 10, pp. 1185–1190, 2000.
 - [31] L. J. Latecki and R. Lakämper, “Convexity Rule for Shape Decomposition Based on Discrete Contour Evolution,” *Comput. Vis. Image Underst.*, vol. 73, no. 3, pp. 441–454, Mar. 1999.
 - [32] P. F. Felzenszwalb and J. D. Schwartz, “Hierarchical Matching of Deformable Shapes,” in *2007 IEEE Conference on Computer Vision and Pattern Recognition*, 2007, pp. 1–8.
 - [33] E. Attalla and P. Siy, “Robust shape similarity retrieval based on contour segmentation polygonal multiresolution and elastic matching,” *Pattern Recognit.*, vol. 38, no. 12, pp. 2229–2241, Dec. 2005.
 - [34] R. Hu, W. Jia, H. Ling, and D. Huang, “Multiscale distance matrix for fast plant leaf recognition,” *IEEE Trans. Image Process.*, vol. 21, no. 11, pp. 4667–72, Nov. 2012.
 - [35] M. Bicego and V. Murino, “Investigating Hidden Markov Models’ capabilities in 2D shape classification,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 2, pp. 281–286, Feb. 2004.
 - [36] K. B. Sun and B. J. Super, “Classification of Contour Shapes Using Class

- Segment Sets,” in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*, 2005, vol. 2, pp. 727–733.
- [37] X. Bai, X. Yang, D. Yu, and L. J. Latecki, “Skeleton-based shape classification using path similarity,” *Int. J. Pattern Recognit. Artif. Intell.*, vol. 22, no. 4, pp. 733–746, Jun. 2008.
 - [38] A. Erdem and S. Tari, “A similarity-based approach for shape classification using Asian skeletons,” *PATTERN Recognit. Lett.*, vol. 31, no. 13, SI, pp. 2024–2032, Oct. 2010.
 - [39] L. Chen, J. J. J. McAuley, R. S. S. Feris, T. S. S. Caetano, and M. Turk, “Shape classification through structured learning of matching measures,” in *2009 IEEE Conf. Comput. Vis. Pattern Recognit.*, 2009, pp. 365–372.
 - [40] X. Yang, X. Bai, L. J. Latecki, and Z. Tu, “Improving Shape Retrieval by Learning Graph Transduction,” in *Comput. Vis. – ECCV 2008*, 2008, pp. 788–801.
 - [41] X. Yang, S. Koknar-tezel, and L. J. Latecki, “Locally constrained diffusion process on locally densified distance spaces with applications to shape retrieval,” in *Computer Vision and Pattern Recognition*, 2009, pp. 357–364.
 - [42] P. Kotschieder, M. Donoser, and H. Bischof, “Beyond Pairwise Shape Similarity Analysis,” in *Asian Conference on Computer Vision*, 2009, vol. 5996, pp. 655–666.
 - [43] X. Bai, B. Wang, C. Yao, W. Liu, and Z. Tu, “Co-transduction for shape retrieval,” *IEEE Trans. Image Process.*, vol. 21, no. 5, pp. 2747–2757, May 2012.
 - [44] S. M. A. Eslami, N. Heess, C. K. I. Williams, and J. Winn, “The shape boltzmann machine: A strong model of object shape,” *Int. J. Comput. Vis.*, vol. 107, no. 2, pp. 155–176, 2014.
 - [45] Zhirong Wu, S. Song, A. Khosla, Fisher Yu, Linguang Zhang, Xiaoou Tang, and J. Xiao, “3D ShapeNets: A deep representation for volumetric shapes,” in *2015 IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 1912–1920.
 - [46] J. Yang, Q. Zhang, K. Xian, Y. Xiao, and Z. Cao, “Rotational contour

- signatures for both real-valued and binary feature representations of 3D local shape,” *Comput. Vis. Image Underst.*, 2017.
- [47] Jin Xie, Yi Fang, Fan Zhu, and Edward Wong, “Deepshape: Deep learned shape descriptor for 3D shape matching and retrieval,” in *2015 IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 1275–1283.
- [48] X. Bai, S. Bai, Z. Zhu, and L. J. Latecki, “3D Shape Matching via Two Layer Coding,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 12, pp. 2361–2373, Dec. 2015.
- [49] Z. Zhu, X. Wang, S. Bai, C. Yao, and X. Bai, “Deep Learning Representation using Autoencoder for 3D Shape Retrieval,” *Secur. Pattern Anal. Cybern. (SPAC), 2014 Int. Conf.*, vol. 204, p. 6, Sep. 2014.
- [50] M. Masoumi and A. Ben Hamza, “Spectral shape classification: A deep learning approach,” *J. Vis. Commun. Image Represent.*, vol. 43, pp. 198–211, Feb. 2017.
- [51] Z. Zhu, X. Wang, S. Bai, C. Yao, and X. Bai, “Deep learning representation using autoencoder for 3D shape retrieval,” in *Proceedings 2014 IEEE International Conference on Security, Pattern Analysis, and Cybernetics (SPAC)*, 2014, vol. 204, pp. 279–284.
- [52] S. J. Thorpe, “NEUROSCIENCE: Seeking Categories in the Brain,” *Science* (80-.), vol. 291, no. 5502, pp. 260–263, Jan. 2001.

2. 项目的研究内容、研究目标，以及拟解决的关键科学问题（此部分为重点阐述内容）；

2.1 项目的研究目标

针对当前形状表示需要根据特定任务进行人工设计的局限性以及匹配分类过程独立于表示无法进行统一优化等问题，本项目以形状分析已有研究成果为基础，借助循环神经网络（Recurrent Neural Network, RNN）、卷积神经网络（Convolutional Neural Network, CNN）等有效深度学习技术手段，系统研究深度形状语义网络表示机理及其判别性，从形状语义解析角度模拟人类大脑形状认知机能，针对形状分析存在的几何变换、目标的自遮挡与互遮挡、非刚性物体的

运动、形状轮廓分割错误与噪声等固有问题，搭建具有形状非刚性结构保留机制的深度形状语义网络，形成以深度学习为基础的形状语义分析理论体系，形成基于形状的目标分类与识别的新成果。研究内容与研究子问题关系如图 3 所示。

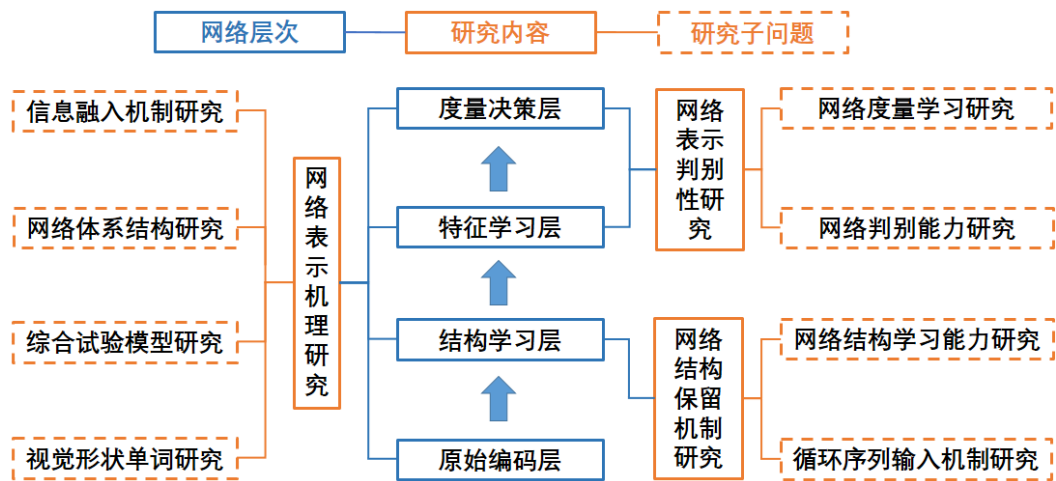


图 3 研究内容组织结构图。围绕深度形状语义网络四个层次的搭建与设置展开，归纳为三方面研究内容八个子问题，研究内容涉及形状几何变换、形状表示判别性等形状分析难点及核心问题。

2.2 研究内容

研究内容一：深度形状语义网络表示机理研究

模拟人类大脑视觉认知原理，从图像语义角度，结合领域知识，研究形状轮廓分解以及视觉形状单词(Visual Shape Words)的定义与描述，在此基础上构建具有形状非刚性结构保留功能的深度形状语义网络，研究深度形状语义特征形成与表示机理。主要研究内容如下：

1) 视觉形状单词研究：探索影响形状轮廓提取的关键影响因素，从视觉语义角度研究形状轮廓最优分解方案，分析不同轮廓片段的个性与共性，结合形状领域知识，设计易于提取且对于局部几何特征（如曲率，梯度，凹凸性等）具有优秀刻画能力的视觉形状单词，定义鲁棒的视觉形状单词字典，形成视觉形状语义表示体系，针对深度网络结构特征，从视觉形状单词在深度形状语义网络中的适用性以及它们之间融合方式的角度，研究关于视觉形状单词的最优形状轮廓编码方式。

2) 深度形状语义网络体系结构研究：分析不同深度学习网络结构的特点，

针对形状分析所面临的不同问题，研究以**原始编码层、结构学习层、特征学习层、度量决策层**为基本功能单元的**深度形状语义网络**结构的搭建，从模拟人类大脑认知形状的角度，以视觉形状语义解析的形式，将形状语义网络中低层神经元到高层神经元传递过程模拟为形状单词到形状短语再到形状语句的形状语言学习过程。

3) 深度形状语义网络影响因素综合试验模型研究：在确定深度形状语义网络框架的基础上，分析影响深度学习性能的关键因素，如神经元之间的链接方式、神经元的激励函数、正则化因子、递归机制等。对影响因素之间的相互耦合作用进行深入研究，从语义特征学习质量，泛化能力和训练速度方面等入手，分析研究引起深度网络学习性能波动的主要因素并对其进行建模分析，根据分析结果对深度形状语义网络在网络结构、神经元设置与训练模式等方面进行进一步优化。

4) 深度形状语义网络的额外信息融入机制研究：形状图像训练样本除了具有类别标签信息外，往往具有一些额外信息（Extra or side information）可以加以利用。本项目从度量决策层入手，围绕目标优化以及深度度量学习，研究深度形状语义网络的额外信息融入机制，包括有监督情况如形状子类别信息（如计算机视觉领域流行的细粒度分类情况， Fine-grained categorization）、形状图片标题信息、形状属性信息等，无监督情况包括样本之间流形结构信息、相似度信息、形状角点信息等。以进一步提升深度形状语义网络的判别能力与泛化能力。

研究内容二：深度形状语义网络结构保留机制研究

由于目标的自遮挡与互遮挡、非刚性物体的运动、拍摄角度等原因，同类形状样本在几何上变换多端，这种几何变换多样性严重影响形状匹配与分类精度。形状几何变换是形状分析的难点问题之一，解决该问题关键是形状非刚性结构的提取，因为形状非刚性结构对几何变换具有良好的鲁棒性。本项目拟研究深度形状语义网络结构学习能力以及形状单词循环序列输入机制，使得深度形状语义网络能够在学习过程中保留形状非刚性结构，从而提升深度形状语义网络对几何变换学习的鲁棒性。主要研究内容如下：

1) 深度形状语义网络结构学习研究：目标的非刚性运动、目标拍摄的角度、

遮挡等常见因素都很容易导致形状几何变换的多样性，从而增加同类样本的差异以致影响形状分析的性能。本项目拟利用循环神经网络，在深度形状语义网络中设计多层记忆神经单元，构建**结构学习层**，把形状分析问题转化为形状语义分析问题，利用循环神经网络在序列学习上的优势，学习视觉形状单词之间相关性从而保留形状碎片之间非刚性结构关系以避免形状几何变换带来的负面影响。

2) 视觉形状单词循环序列输入机制研究：有别于自然语言单词序列或其他视频序列，形状单词序列是一个首尾相连的循环序列。针对上述特征，本项目将从深度形状轮廓编码形式、循环池化、循环切片、循环堆叠、循环滚动以及数据扩张等多个方面探索视觉形状单词循环序列输入机制。

研究内容三：深度形状语义网络表示判别性研究

形状分类、形状检索等诸多形状分析任务都属于有监督学习或半监督学习问题，而这类问题的解决关键在于所学习的形状特征表示对不同形状的优秀判别能力。显然，深度形状语义网络表示的判别性研究是深度形状语义网络学习性能提升的核心问题。主要研究内容如下：

1) 深度形状语义网络判别能力研究：拟采用卷积神经网络架构在深度形状语义网络高层框架中构建**特征学习层**，利用卷积神经网络在显著性特性学习上优势，突出形状非刚性结构的显著特征部分，从而提升整个形状语义网络的判别能力。此外，本项目还将分析训练模式、神经元设置、激活函数、卷积层设置等因素之间耦合关系，从网络架构与设置层面进一步优化深度形状语义网络的判别能力。

2) 深度形状语义网络度量学习研究：在优化深度形状语义网络结构的同时，本项目拟在深度形状语义网络中融入度量学习理论，从而在深度形状语义网络顶层构建**度量决策层**，充分利用形状额外信息与先验知识，设计最优形状表示度量目标，通过神经网络反馈调节机制，保证形状轮廓样本在高层形状语义度量空间得到最佳分类。

2.3 拟解决的关键科学问题

科学问题 1：研究形状的多层语义表示机理

当前形状分析采用的形状表示均以手动设计特征为主，这些特征大多数是

根据特定任务而特别设计的，多数表现为单层低语义特征，并没有表达出各类形状内在的高层语义信息，受旋转、尺度、形变、视角等因素的影响较大，因此需要模拟大脑对形状图像的认知过程，构建从形状单词到形状短语再到形状语句多层形状语义体系，研究适合形状分析的深度学习网络架构，对深度形状语义网络的低语义特征到高语义特征的表示机理进行探索并建模分析。

科学问题 2：研究形状非刚性结构保留机制

作为形状分析难点问题，形状的几何变换是造成类内样本之间的差异性的主要原因，从而直接影响形状分析的精确度。形状非刚性结构对形状的几何变换有天然的鲁棒性。因此，可以构建视觉形状语义表示体系，从形状单词序列学习角度重新审视形状分析的问题，探索深度形状语义网络结构保留机制，从而提取形状非刚性结构提升网络对几何变换的鲁棒性。

科学问题 3：研究基于形状语义判别的形状分类与匹配

诸多形状分析任务实质上是一个形状分类或形状匹配问题，而决定这些任务的分类精度或匹配性能的关键在于形状表示的判别性能。在传统形状分析工作中，形状匹配分类过程通常独立于形状的表达，并不能根据形状匹配分类的结果对形状的表达进行反馈与自动调节，导致整个形状分析的学习框架的判别能力不能完全地被激发，因此需要构建集表示、匹配分类以及度量学习于一体的深度学习网络框架，使得形状表示、匹配分类相互反馈与自动调节，从总体上优化深度形状语义网络的判别能力，提升形状语义表示的判别性能，从而保证最终学习到的形状特征不仅具有高层形状语义特征，且在度量空间下具有丰富的判别信息。

3. 拟采取的研究方案及可行性分析（包括研究方法、技术路线、实验手段、关键技术等说明）；

3.1 研究方法和技术路线

本项目以深度形状语义网络多层表示机理为研究核心，围绕形状分析中几何变换与同类样本多样性这两类难点问题，重点研究深度形状语义网络的非刚性结构保留机制以及深度形状语义网络表示的判别性。本项目主要研究思路：通过形状轮廓分解与形状字典学习技术，构建原始编码层，奠定形状语义表示基础；通过循环神经网络技术，构建结构学习层，保留形状非刚性结构，提升形状语义表示的几何变换鲁棒性；通过卷积神经网络技术，构建特征学习层，突

出形状语义表示的显著特征，提升深度形状语义网络的判别能力；利用深度度量学习技术，构建度量决策层，融合额外信息与先验知识，完成整个网络优化反馈，提升度量空间的判别性，促进形状的分类与匹配。最终，在深度形状语义网络中，形状表示的学习过程被转化为从形状单词到形状短语再到形状语句的形状语义归纳过程，从而达到模拟人类大脑对形状认知的目的。本项目的技术路线以及所涵盖的研究方法如图 4 所示。

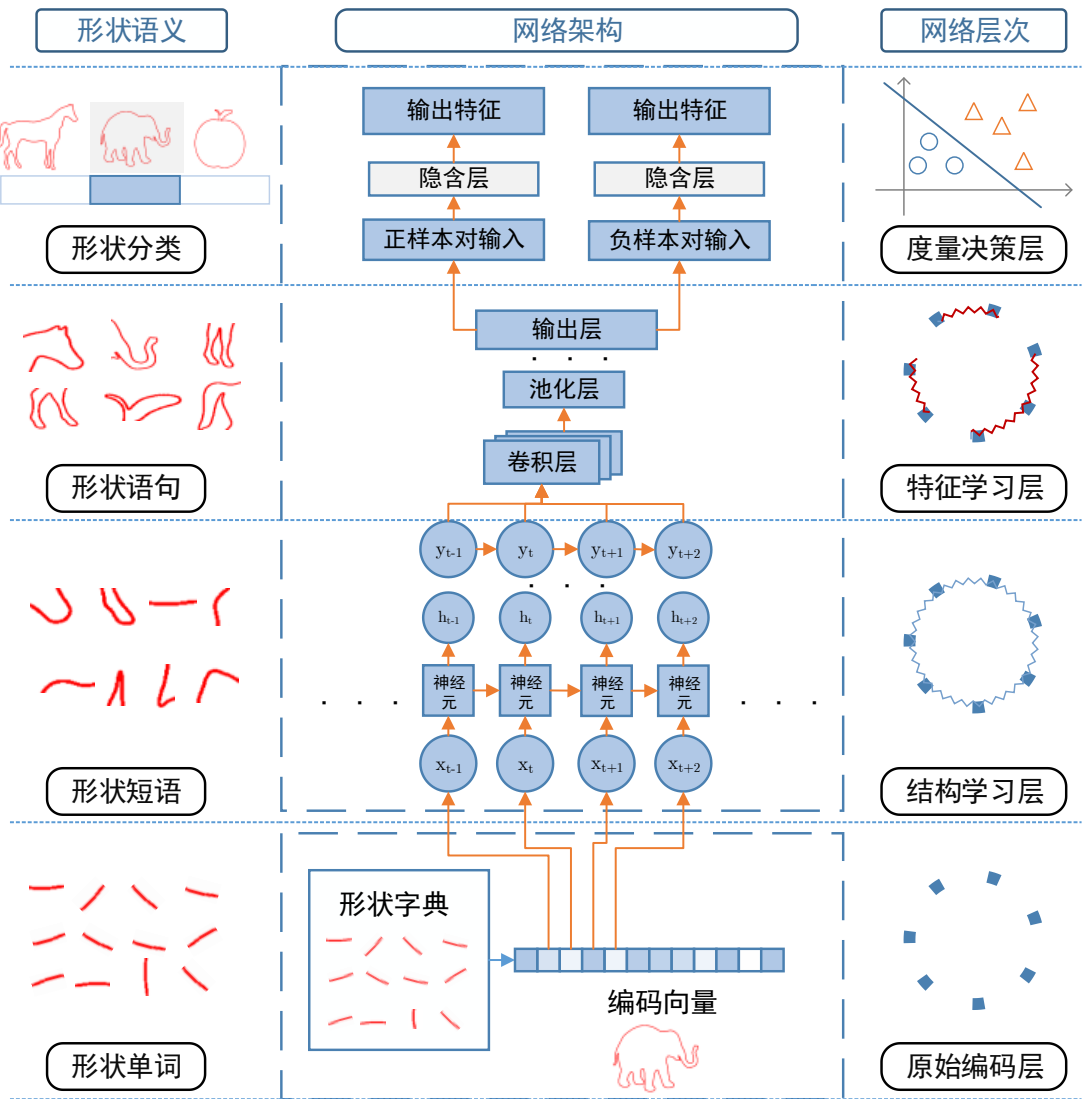


图 4 总体研究方案与技术路线图。由底向上分别从原始编码、结构学习、特征学习与度量决策角度对形状单词、形状短语、形状语句、形状分类进行形状语义建模。

3.2 研究方案

根据本项目的研究目标，结合研究内容，提出如下研究方案：

1) 深度形状语义网络表示机理研究

现有的形状表示方法大多依赖人工设计，主要表现为单层特征，不足以描述形状的语义信息，限制了形状分类与匹配的性能。为了有效获取形状的语义信息，本项目拟研究形状语义的多层表示机理，探索由形状到语义的映射模式，学习如下形状多层语义特征学习目标：

$$O_1 = f_1(I, \Theta_1),$$

$$O_i = f_i(O_{i-1}, \Theta_i), i = 2, 3, \dots, n,$$

其中， $f_i, i = 2, 3, \dots, n$ 为待学习的形状语义 n 级映射函数； $\Theta_i, i = 1, 2, \dots, n$ 为待学习的函数参数； I 表示输入形状的最低层特征； $O_i, i = 1, 2, \dots, n$ 表示对输入形状由低级到高级的 n 层语义表示。为了使 f 具备提取由低级到高级的多层形状语义特征的能力，研究适合形状语义分析的深度学习层次结构，搭建深度形状语义网络，探寻形状的多层语义表示机理。如图 4 所示，项目将由底向上分别从原始编码、结构学习、特征学习与度量决策角度研究形状单词、形状短语、形状语句、形状分类等形状语义的建模方式与表示机理。

首先，通过检测角点位置信息对形状轮廓进行分解，研究轮廓碎片的提取方法，并探索各类形状轮廓片段的关键特性和规律。分析形状轮廓的不同底层特征之间的共性和个性，如梯度、曲率、坐标、区域等。深入分析计算机视觉领域常用的融合技术，将现有特征表示框架和局部形状描述方法，如词袋模型、空间金字塔匹配、局部约束线性编码等，改进为适用于深度形状语义网络的底层特征，构建基于轮廓片段的视觉形状编码方法，形成深度形状语义网络的底层视觉单词特征。

其次，深入研究循环神经网络的建模过程，设计适用于形状结构保留的循环网络搭建方式，通过学习形状单词之间相关关系，提取可以保留形状非刚性结构的形状短语特征以提升形状几何变换的鲁棒性。同时，深入研究卷积神经网络对显著特征的学习过程，搭建适用于形状高层语义认知的卷积网络结构，提取形状非刚性结构的显著判别信息，获得描述高层语义信息的形状语句特征，为形状度量学习与判别奠定基础。

2) 深度形状语义网络结构保留机制研究

形状匹配与分类的精度受目标非刚性运动、遮挡、拍摄角度等几何变换因素严重制约。由于形状非刚性结构对形状几何变换的天然鲁棒性，本项目拟从网络结构学习与形状单词循环序列输入机制角度，着重探索深度形状语义网络的形状非刚性结构学习与记忆功能。主要研究思路如下：

形状语义网络结构学习：轮廓经过分解后，产生的形状片段具有潜在的顺序发生关系。本项目拟对形状片段之间这种关联性进行建模，从而刻画形状轮廓的非刚性结构，提升深度形状语义网络对几何变换的鲁棒性。为此，基于形状片段获取底层视觉单词后，进一步搭建结构学习层。假设形状单词构成集合 X ，结构学习层将学习由形状单词集 X 到形状短语集 P 的映射函数

$$P = f_1(X, \xi),$$

其中， ξ 是结构学习函数 f_1 的参数。为了学习 f_1 及其参数 ξ ，本项目拟借鉴传统的循环神经网络结构，探索极大化保留轮廓非刚性结构的网络连接模式以及动态门控结构学习机制，设计适应于以形状序列的单词表示为输入的记忆神经网络结构。

视觉形状单词循环序列输入机制：尽管传统的循环神经网络对自然语言单词这种首尾明确的序列数据具有优异的建模能力，却缺乏对形状单词序列这种首尾相连循环性序列的建模与处理相应技术，从而限制整个网络对形状语义的建模能力以及非刚性结构的捕获能力。针对上述问题，本项目拟从循环序列角度探究深度形状轮廓编码的形式，并着重分析形状视觉单词的循环池化、循环切片、循环堆叠、循环滚动等机制，从而设计适用于形状单词循环序列输入的记忆神经网络结构。

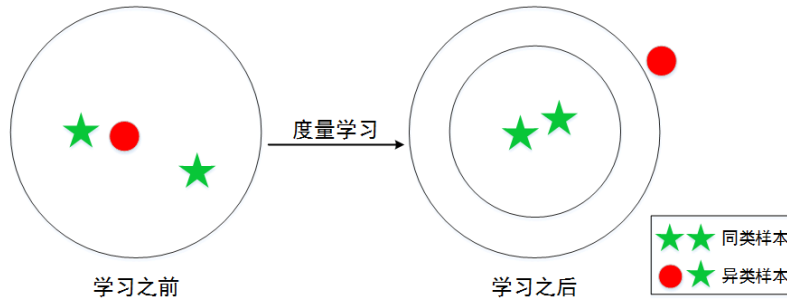
3) 深度形状语义网络表示的判别性研究

本项目拟从深度网络结构的设置以及深度度量学习两个方面入手，研究深度形状语义网络表示的判别性：

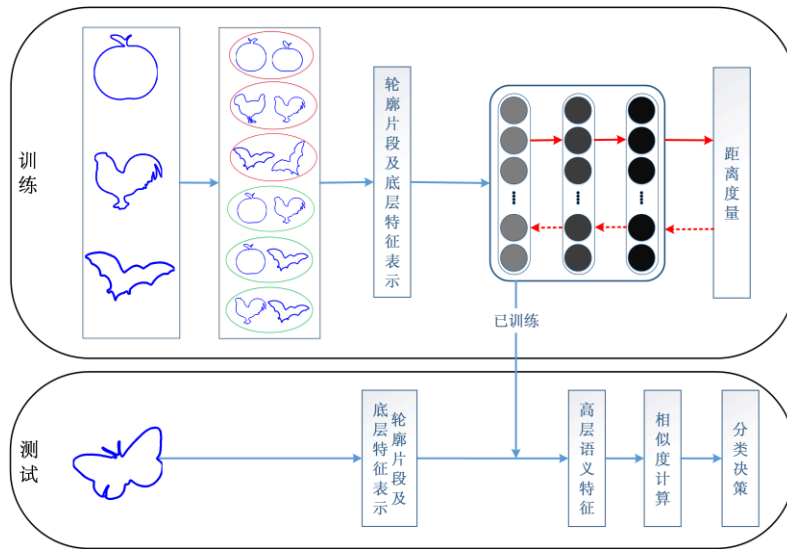
网络判别能力：卷积神经网络是目前计算机视觉领域主流深度学习网络架构，在诸多分类问题上表现强劲。本项目拟借鉴卷积神经网络结构，在结构学习层输出的基础上，搭建特征学习层。具体地，特征学习层将学习由形状短语集 P 到形状语句集 S 的映射函数

$$S = f_2(P, \eta),$$

其中， η 是特征学习函数 f_2 的参数。为了学习 f_2 及其参数 η ，搭建适用于形状高级语义特征学习的卷积神经网络，同时探索特征学习层面中激活函数、层数设置、反馈方式等对形状轮廓分类的核心因素，从而学习形状结构的高级语义表示，即形状语句，保证后续的形状度量决策。



(a) 度量学习前后样本距离对比示意图



(b) 形状多层深度学习网络的表示及分类

图 5 形状多层深度网络与度量学习框架

深度度量学习：项目拟融合度量学习思想于深度学习网络框架中，构建度量决策层，赋予网络多层特征更多的判别信息，使学习到的同类形状样本特征的距离足够近，异类形状样本特征的距离尽量远，如图 5 (a)。首先构建不同的形状正样本对和负样本对，通过深度形状语义网络逐层学习，在网络的最后层计算距离度量 d ，根据度量的结果，前向反馈调整多层特征表示，如图 5 (b)。假设深度学习网络有 $K+1$ 层，则优化目标函数为：

$$J = \frac{1}{P} \sum_{l_{ij}=1} d(H(x_i), H(x_j)) - \frac{\alpha}{Q} \sum_{l_{ij}=-1} d(H(x_i), H(x_j)) + \beta \sum_{k=1}^K (\|W^{(k)}\|_F^2 + \|b^{(k)}\|_2^2),$$

其中， H 是对整个网络的复合映射函数，涵盖了结构学习函数 $P = f_1(X, \xi)$ 与特征学习函数 $S = f_2(P, \eta)$ 的复合； α 是正负样本重要性平衡参数； β 是正则化参数； P 是正样本对数； Q 是负样本对数；对于 $k \in \{1, 2, \dots, K\}$ ， $W^{(k)}$ 与 $b^{(k)}$ 包含了整个网络的学习参数。

通过把样本语义特征的距离度量嵌入深度网络中进行监督式的学习可以使学习到的特征不但具有丰富的语义信息，而且包含强判别信息，可以提高形状匹配及分类的精度。此外，本项目拟在度量决策层中，从度量目标优化及约束角度，以属性相关性约束、样本邻接矩阵等形式，融入形状其他可利用信息，进一步提升深度形状语义网络判别能力。

3.3 可行性分析

3.3.1 理论与技术基础

深度学习技术近年来在计算机视觉研究中得到广泛应用并取得大量优秀研究成果，而形状匹配与分类也一直是计算机视觉的研究热点，每年都有新的研究成果发表在计算机视觉的顶级期刊或会议上，如 PAMI、TIP、CVPR、ICCV、ECCV 等。这两方面的研究成果为本项目探索基于深度学习的形状匹配与分类方法奠定了坚实的理论和技術基础。

3.3.2 研究基础与研究经验

项目人员配备合理，具有扎实的相关工作基础。项目主要成员主研了国家自然科学基金项目两项：“主动视觉注意的语义认知计算模型研究（60975015）”和“形状的语义结构表示及其分类学习研究”（61173131）。承担了重庆市攻关项目“全景图像制作与基于嵌入式的漫游关键技术（2009AC2057）”，“嵌入式 WEB 监控视频摘要生成及其快速浏览关键技术研究 and 开发（CSTC2009AB2230）”，重庆市重点项目“图像处理技术及其应用基础研究（CSTC2005BA2002）”。承担了深圳华尊科技公司的企业委托项目“视频大数据行人检测和搜索”。在视觉特征提取与描述、曲率尺度空间技术、轮廓特征提取技术、描述器的降维、图像语义建模等技术领域积累了扎实的研究基础，部分结果已发表于国际和国内知名

刊物。同时，采用深度学习技术解决大数据条件下的海量视频行人检测与搜索问题取得了较好的实践经验。因此，本项目的研究是前期研究项目的深入和扩展。项目主要成员具有应用数学或软件工程背景，项目团队具有很强的理论、算法研究能力和很强的软件开发能力，这为项目的理论推导、算法设计和软件开发提供了强有力的保障。

3.3.3 实验条件储备

项目组部分成员具有扎实的编程基础和丰富的编程经验（C/C++，Matlab，Python），能够独立实现了大量相关算法。在形状分析方面，项目组积累了大量的形状分析的主流数据库、实验数据和相关理论的实现经验。在深度学习方面，项目组搭建了形状深度学习实验平台，能够熟练使用常见深度学习框架以及在图形工作站上进行 GPU 并行编程。这些都为本项目的顺利实施提供了支撑和保障。

4. 本项目的特色与创新之处；

1) 项目特色

本项目采用深度学习理论框架研究形状的多层语义表示，在此基础上深入研究形状非刚性结构保留机制与网络表示的判别性问题，并把度量学习融入到深度学习网络之中，构建了集形状表示、形状分类、形状匹配与度量学习为一体的深度形状语义网络。其特色在于：形状表示从人工设计转变为自主学习和提取，从原始编码单层特征转变为多层语义特征；深度形状语义网络不仅能够保留并利用形状非刚性结构提升形状表示对几何变换的鲁棒性，还能利用卷积神经网络与度量学习技术在特征学习上的优势提升形状表示的判别性。

2) 项目创新点

① 本项目模拟人类大脑的视觉认知机理，探明形状的多层语义表示原理。低层表示了各式各样的视觉形状单词（形状碎片），高层逐步表示形状单词组成的视觉形状短语（形状轮廓片段）直至表示视觉形状语句（形状轮廓）。在整个学习表示过程中，形状轮廓整体结构以及具有明确类别信息的显著轮廓片段被有效地提取。

② 本项目模拟自然语言处理模式求解形状分析问题，设计适用于处理形状循环序列输入的循环深度网络架构，捕获视觉形状单词之间循环顺序发生关系，从而刻画形状碎片对于形状轮廓的非刚性结构，从而最大限度地降低几何变换

对形状表示带来的负面影响。

③ 本项目在高层网络结构中采用卷积神经网络框架，降低网络训练参数，学习形状显著特征，同时在网络优化目标中引入度量学习，融入额外信息，从而从多方面提升深度形状语义网络的判别能力。

5. 年度研究计划及预期研究结果（包括拟组织的重要学术交流活动、国际合作与交流计划等）。

5.1 年度研究计划

时间阶段	研究内容
2018.01~2019.08	① 收集整理相关数据库，搭建课题研究实验平台 ② 研究视觉形状单词提取与描述 ③ 研究深度形状语义网络体系结构 ④ 研究深度形状语义网络影响因素综合试验模型 ⑤ 研究深度形状语义网络额外信息融入机制
2019.09~2020.02	① 研究深度形状语义网络的循环序列输入机制 ② 研究深度形状语义网络结构学习能力
2020.03~2020.08	① 研究深度形状语义网络判别能力 ② 研究深度形状语义网络度量学习
2021.09~2021.12	① 邀请国内外软件质量研究专家，进行短期学术交流 ② 总结研究成果，整理材料并结题

5.2 预期研究成果

- 1) 提出一套结构保留深度形状语义网络多层表示学习理论；
- 2) 实现建立配套的形状匹配和分类软件模块，为后续研究提供实验基础；
- 3) 参加 3 次国际或国内学术会议；
- 4) 通过本项目实施，培养计算机视觉研究领域青年教师 2 名，博士生 2 名，3 名以上的硕士生；
- 5) 拟在国内外重要学术期刊和会议上发表与本项目相关的学术论文 12 篇，其中至少 9 篇被 SCI 检索。

(二) 研究基础与工作条件

1. 研究基础（与本项目相关的研究工作积累和已取得的研究工作成绩）；

项目组主要成员长期从事计算机视觉与模式识别领域的研究工作，已完成众多与本项目紧密相关的科研项目，如国家自然科学基金项目“形状的语义结构

表示及其分类学习研究(61173121)”,国家自然科学基金项目“主动视觉注意的语义认知计算模型研究(60975015)”,“重庆市自然科学基金“形状特征行为分析和特征描述研究”(CSTS2010BB2061)”,重庆市攻关项目“全景图像制作与基于嵌入式的漫游关键技术(2009AC2057)”,“嵌入式 WEB 监控视频摘要生成及其快速浏览关键技术研究 and 开发(CSTC2009AB2230)”,重庆市自然科学基金重点项目“图像处理技术及其应用基础研究(CSTC2005BA2002)”和“工业 CT 图像重建与处理系统”。主要相关工作体现在如下三个方面:

(1) 形状语义表示及形状分类。国家自然科学基金项目“形状的语义结构表示及其分类学习研究(61173121)”研究形状局部底层特征的表示和描述和形状语义表示,建立形状语义映射学习框架。研究取得的创新性成果已发表在《IEEE Transaction on Pattern Analysis and Machine Intelligence》等国内外著名期刊上,同时积累了许多常用典型的形状分析数据库,如 MPEG-7, ETH80, Kimia99, Kimia216, Swedish Leaf, Tari 等。

(2) 图像特征检测、描述与匹配。重庆市自然科学基金“形状特征行为分析和特征描述研究(CSTS2010BB2061)”,重庆市攻关项目“全景图像制作与基于嵌入式的漫游关键技术(2009AC2057)”,“嵌入式 WEB 监控视频摘要生成及其快速浏览关键技术研究 and 开发(CSTC2009AB2230)”,重庆市自然科学基金重点项目“图像处理技术及其应用基础研究(CSTC2005BA2002)”和“工业 CT 图像重建与处理系统”,着重开展了图像特征的检测与分析研究,包括图像边缘检测、角点检测、人脸识别、图像配准、图像分割、图像重建、尺度空间技术、多尺度曲率乘积技术、基于流形学习的降维技术、多尺度协方差矩阵、基于小波分析的特征检测与描述技术、基于 SIFT 算法的图像描述、图像特征点匹配、图像拼接技术、基于条件数的特征点稳定性分析与检测技术等。在这些研究中取得了一些创新性成果,其中部分研究成果在国内外著名期刊上发表论文,《Pattern Recognition》、《Pattern Recognition Letters》、《IEEE Transaction on Broadcasting》、《Journal of Electronic Imaging》、《Neural Computing& Applications》、《自动化学报》、《电子与信息学报》、《中国图像图形学报》、《光电工程》等。

(3) 基于统计学习的图像语义相关工作基础积累。项目组通过国家自然科

学基金项目“主动视觉注意的语义认知计算模型研究(60975015)”，本项目提出的图像语义认知计算，涉及了大量的统计学习和数据挖掘技术。项目组在前期图像的语义研究中，阅读了大量的国内外的研究文献，积累关于图像语义的基础理论和方法，熟悉了图像语义分析的开源 Matlab 软件包，并掌握和重写了该软件包所包含的 EM、PLSA、LDA、K-Means 等大量的统计学习技术。

(4) 深度学习相关工作基础积累。项目组主要成员与专业从事视频大数据技术研究的高新技术企业合作建立了良好的合作关系。承担了深圳华尊科技公司委托开发项目“视频大数据行人检测和搜索”中，项目目标是在无约束的视频检测和搜索感兴趣的行人。项目组主要成员将深度学习技术，如 CNN、SRBM 等，应用在大数据平台下针对海量视频进行行人检测与搜索。摄像头视角的无约束以及行人的非刚性形变导致深度学习技术在训练行人分类器时需要消耗大量计算机资源。针对这一问题，项目组主要成员利用 MapReduce 技术实现训练任务的分布式计算。这为项目的实施提供了充分的实践参考。

总之，上述与本项目紧密相关的工作积累和在深度学习方面丰富的实践经验为项目的成功实施提供强有力的保障。

项目组近年的代表性论文如下：

- [1] X Zhang, Y Qu, D Yang, H Wang, J D Kymer, “Laplacian scale space behavior of planar curve corners”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.37, no.11, pp.2207-2217 (2015) (SCI)
- [2] M Yan, Y Fu, X Zhang, D Yang, L Xu, JD Kymer, “Automatically Classifying Software Changes via Discriminative Topic Model: Supporting Multi-category and Cross-project”, Journal of Systems and Software, vol.113, no.3, pp.296-308 (2016) (SCI)
- [3] M Yan, X Zhang, D Yang, L Xu, JD Kymer, “A component recommender for bug reports using Discriminative Probability Latent Semantic Analysis”, Information and Software Technology, vol.73, pp.37-51 (2016) (SCI)
- [4] Y Fu, M Yan, X Zhang, D Yang, Xu L, JD Kymer, “Automated Classification of Software Change Messages by Semi-supervised Latent Dirichlet Allocation”, Information and Software Technology, vol.57, pp.369-377 (2015) (SCI)
- [5] X Zhang, H Wang, A Smith, L Xu, B Lovell, D Yang, “Image Corner Detection Algorithm Based on Gradient Correlation Matrix of Planar Curves”, Pattern Recognition, vol.43,

pp.1207-1223 (2010) (SCI)

- [6] X Zhang, H Wang, M Hong, L Xu, D Yang, L Brain, “Robust Image Corner Detection Based on Scale Evolution Difference of Planar Curves”, Pattern Recognition Letters. Vol.30, pp.449-455. (2009) (SCI)
- [7] X Zhang, M Lei, D Yang, “Multi-scale Curvature Product for Robust image Corner Detection in Curvature Scale Space”, Pattern Recognition Letters, vol.28, pp.545-554 (2007) (SCI)
- [8] Y Ge, D Yang, X Zhang, J Lu, “Image registration based on Subpixel Localization and Cauchy-Schwarz Divergence” Journal of Electronic Imaging, vol.19, no.3, pp.033001 (2010) (SCI)
- [9] Q J Chen, Y C Zhong, X Zhang, “A pseudo genetic algorithm”, Neural Computing & Application, vol.19, pp. 77-83 (2010) (SCI)
- [10] Y Qu, X Zhang, D Yang, “2D Shape Classification via Sparse Semantic Representation”, IEEE Transactions on Pattern Analysis and Machine Intelligence, (Under review, SCI)
- [11] S. Huang, A. Elgammal, J. Lu, D. Yang: “Cross-Speed Gait Recognition Using Speed-Invariant Gait Templates and Globality-Locality Preserving Projections”, IEEE Transactions on Information Forensics and Security, vol.10, no.10, pp. 2071-2083, 2015 (SCI)
- [12] S. Huang, D. Yang, J. Zhou, X. Zhang, “Graph regularized linear discriminant analysis and its generalization”, Pattern Analysis and Application, vol.18, no.3, pp.639-650, 2015 (SCI)
- [13] S. Huang, D. Yang, H. Zhang, L. Huangfu, X. Zhang, “Shape primitive histogram: low-level face representation for face recognition”, IET Biometrics, vol.3, no.4, pp.325-334, 2014 (SCI)
- [14] S. Huang, A. Elgammal, “On the effect of hyperedge weights on hypergraph learning”, Image and Vision Computing, vol.57, pp.89-101, 2017 (SCI)
- [15] Q Huang, X Zhang, S Huang, D. Yang. “Sparse graph-based inductive learning with its application to image classification”, Journal of Electronic Imaging, vol.25, no.5, pp. 050502-050502, 2016 (SCI)
- [16] S. Huang, D. Yang, Y. Ge, X. Zhang, “Discriminant Hyper-Laplacian Projections and its scalable extension for dimensionality reduction”, Neurocomputing, vol.173, pp.145-153,

2016 (SCI)

- [17] S Huang, D Yang, B Liu, X Zhang, “Regression-based Hypergraph Learning for Image Clustering and Classification”, arXiv preprint arXiv:1603.04150 (2016)
- [18] H. Wang, Y. Kawahara, C. Weng, J Yuan, “Representative Selection with Structured Sparsity”, Pattern Recognition, vol.63, pp.268-278, 2017 (SCI)
- [19] J. Yang, H. Wang, J. Yuan, Y. Li, J. Liu, “Invariant multi-scale descriptor for shape representation, matching and retrieval”, Computer Vision and Image Understanding, vol.145, pp.43-58, 2016 (SCI)
- [20] H. Wang, J. Fan, Y. Li, “Research of shoeprint image matching based on SIFT algorithm”, Journal of Computational Methods in Sciences and Engineering, vol.16, no.2, pp.349-359, 2016 (EI)
- [21] B. Li, H. Wang, M. Liu, H. Hu, Z. Mao, “Applicability of non-line-of-sight ultraviolet single-scatter approximation model” Photonic Network Communications, vol.31, no.1, pp.147-154, 2016 (SCI)
- [22] Z. Chen, H. Wang, Z. Zhao, X. Liu, “Novel PSWF-Based Multidimensional Constellation Modulation for Broadband Satellite Communication”, Wireless Personal Communications, vol.88, no.3, pp.493-524, (2016) (SCI)
- [23] J. Meng, H. Wang, J. Yuan, Y.P. Tan, “From Keyframes to Key Objects: Video Summarization by Representative Object Proposal Selection”, in Computer Vision and Pattern Recognition (CVPR), pp.1039-1048, (2016) (SCI)
- [24] Y. Ge, C. Peng, M. Hong, S. Huang, D. Yang, “Joint Local Regressors Learning for Face Alignment”, Neurocomputing, vol.208, pp.262-268, (2016) (SCI)
- [25] Y. Ge, X. Ren, C. Peng, X. Wang, “Extended Robust Cascaded Pose Regression for Face Alignment”, in Chinese Conference on Biometric Recognition (CCBR), pp.50-58, (2016) (EI)
- [26] X. Gu, Y. Ge, “Weighted Local Metric Learning for Person Re-identification”, in Chinese Conference on Biometric Recognition (CCBR), pp.686-694, (2016) (EI)
- [27] Z. Feng, F. Liu, H. Guo, Z. Chen, M. Jiang, M. Hong, Q. Jia, “Distributed Compressed Sensing MRI Using Volume Array Coil”, International Journal of Distributed Sensor Networks, vol.9, no.9, pp.989678, (2013) (SCI)
- [28] 张小洪,李博, 杨丹, “一种新的 Harris 多尺度角点检测”, 电子与信息学报, vol.29, no.7,

pp.1735-1738 (2007) (EI)

- [29] 张小洪, 雷明, 杨丹, “基于多尺度曲率乘积的图像角点检测”, 中国图象图形学报, vol.12, no.7, pp.1270-1275 (2007) (EI)
- [30] 张小洪, 杨丹, “基于梯度相关矩阵行列式的角点检测”, 中国图象图形学报, vol.14, no.8, pp.1601-1608 (2009) (EI)
- [31] 杨丹, 游磊, 张小洪, “参数优化模型的正交逼近算法”, 电子与信息学报, vol.31, no.11, pp.2610-2613 (2009) (EI)
- [32] 杨丹, 游磊, 张小洪, 李博, “基于区域生长的鱼眼图像轮廓提取算法”, 计算机工程, vol.36, no.8, pp.217-221 (2008) (CSCD)
- [33] 杨丹, 王洪星, 张小洪, 闫卫杰, “轮廓曲线的 LoG 变换及图像共变区域的检测”, vol.36, no.6, pp.817-822 (2010) (EI)
- [34] 李博, 杨丹, 张小洪, “一种新的基于梯度方向直方图的图像配准方法”, 计算机应用研究, vol.24, no.3, pp.312-314 (2007) (CSCD)
- [35] 李博, 杨丹, 张小洪, “基于 Harris 多尺度角点检测的图像配准新算法”, 计算机工程与应用, vol.42, no.35, pp.37-40 (2006) (CSCD)
- [36] 葛永新, 杨丹, 张小洪, “基于特征点对齐度的图像配准方法”, 电子与信息学报, vol.29, no.2, pp.425-428 (2007) (EI)
- [37] 葛永新, 杨丹, 张小洪, “基于边缘特征点对对齐度的图像配准方法”, 中国图象图形学报, vol.12, no.7, pp.1291-1295 (2007) (EI)
- [38] 马丽涛, 杨丹, 张小洪, 李博, “一种新的基于条件数的图像配准算法”, 中国图象图形学报, vol.13, no.2, pp.277-283 (2008) (EI)
- [39] 王玉珠, 杨丹, 张小洪, “Robust Corner Detection Based on Multi-scale Curvature Product in B-spline Scale Space”, 自动化学报, vol.33 no.4, pp.414-417 (2007) (EI)
- [40] 徐小明, 杨丹, 张小洪, 周小龙, “基于局部不变映射的特征描述器算法”, 自动化学报, vol.34, no.9, pp.1174-1177 (2008) (EI)
- [41] 徐小明, 杨丹, 张小洪, 周小龙, “极坐标下基于差分统计的描述器设计”, 中国图形图像学报, vol.14, no.5, pp.961-966 (2009) (EI)
- [42] 雷明, 杨丹, 张小洪, “基于协方差矩阵的 B-样条多尺度表示的角点检测”, 光电工程, vol.35, no.2, pp.45-50 (2008) (CSCD)
- [43] 雷明, 杨丹, 张小洪, “条件理论控制下良态特征的匹配算法”, 光电工程, vol.35, no.5,

pp.124-128. (2008) (CSCD)

2. 工作条件（包括已具备的实验条件，尚缺少的实验条件和拟解决的途径，包括利用国家实验室、国家重点实验室和部门重点实验室等研究基地的计划与落实情况）；

本项目所需研究和实验条件都已经基本具备：项目组所在单位重庆大学软件工程学院拥有软件工程实验室、重庆市信息安全中心以及重庆大学软件中心等实验基地。另外，重庆大学数字图书馆拥有较齐全的国内外期刊电子杂志。研究过程中所需计算机、服务器和云平台等硬件条件都已具备。项目组的成员不仅有专门从事算法和理论研究的人员，同时项目成员中还拥有丰富经验的软件工程开发人员。且目前项目组成员已经在国内和国际相关领域期刊及会议上发表了大量论文，完全具备完成项目的能力。

3. 正在承担的与本项目相关的科研项目情况（申请人和项目组主要参与者正在承担的与本项目相关的科研项目情况，包括国家自然科学基金的项目和国家其他科技计划项目，要注明项目的名称和编号、经费来源、起止年月、与本项目的关系及负责的内容等）；

4. 完成国家自然科学基金项目情况（对申请人负责的前一个已结题科学基金项目（项目名称及批准号）完成情况、后续研究进展及与本申请项目的关系加以详细说明。另附该已结题项目研究工作总结摘要（限 500 字）和相关成果的详细目录）。

项目名称：形状的语义结构表示及其分类学习研究（编号：61173121），起止年月：2012.1-2015.12。

项目完成情况：该项目按照预订计划完成了研究计划，得到了若干有特色的研究成果。项目取得的成果已经总结并发表为 15 篇学术论文，其中 SCI 检索 12 篇，EI 检索 14 篇。项目在执行期间共培养研究生 10 名，其中毕业博士研究生 2 人，硕士研究生 4 人。因此，该项目发表论文数目及人才培养情况均达到研究计划的要求。

与本项目关系：本项目研究内容是在其成果的基础上的进一步研究，该项目为本项目提供了重要的理论支撑、宝贵的项目实施经验和合理的研究规划。

工作总结摘要：形状描述、推理和理解仍然是计算机视觉与模式识别领域

一个极具挑战性的热点问题。为跨越形状认知的"语义鸿沟",本项目研究了形状特征的尺度空间行为和稳定性,形状局部底层特征的表示和描述,高维局部特征描述空间的各种特性及其低维子空间的构造方法,形状单词在形状语义空间中的分布特征和语义映射学习框架等内容,得到了角点的轮廓拉普拉斯尺度空间行为理论、基于 Chebyshev 多项式拟合的连续曲率估计方法、基于弧长的角度估计方法、基于二维非负偏最小二乘方法的图像特征提取方法、基于非负矩阵分解的图像特征提取算法、形变物体的相似性度量图嵌入匹配方法、基于稀疏图的直推算法、基于形状信息的特征提取方法、基于稀疏表示的图像特征提取方法、判别超拉普拉斯投影算法、基于局部保留映射的子空间学习方法、类特定稀疏表示方法、基于形状语义稀疏表示的形状分类框架、骨架模型的自动化构建及薄板样条拟合方法等研究成果。本项目提出的形状分类框架有效地建立了形状底层特征与语义间的映射关系。本项目取得的成果已经总结并发表在 TPAMI、IEEE Signal Processing Letters、Neurocomputing、CVPRW 2014、ICMEW 2014、自动化学报等重要期刊或会议上,共计论文 15 篇,其中 SCI 检索 12 篇, EI 检索 14 篇。综上所述,本项目的研究不仅具有理论意义,而且具有广阔应用前景。

成果目录:

- [1] *Zhang, Xiaohong、Wang, Hongxing、Smith, Andrew W. B.、Ling, Xu、Lovell, Brian C.、Yang, Dan, Corner detection based on gradient correlation matrices of planar curves , Pattern Recognition, 43(4), pp 1207-1223, 2010/4. SCI
- [2] *Feng, Xin、Liu, Tao、Yang, Dan、Wang, Yao, Saliency Inspired Full-Reference Quality Metrics for Packet-Loss-Impaired Video, IEEE Transactions on Broadcasting, 57(1), pp 81-88, 2011/3. SCI
- [3] Ma, Peng、Yang, Dan、Ge, Yongxin、Zhang, Xiaohong、Qu, Ying, Face recognition using two-dimensional nonnegative principal component analysis, Journal of Electronic Imaging, 21(3), 2012/9JUL-SEP. SCI
- [4] *Peng Ma、Dan Yang、Yongxin Ge、Xiaohong Zhang、Ying Qu、Sheng Huang、Jiwen Lu, Robust Face Recognition via Gradient based Sparse Representation, Journal of Electronic Imaging, 2013, 22(1):3018. SCI
- [5] *Ge, Yongxin、Yang, Dan、Zhang, Xiaohong、Lu, Jiwen , Improved similarity measure-based graph embedding for face recognition , Journal of Electronic Imaging, 21(1), 2012/3JAN-MAR. SCI

- [6] 杨丹、王洪星、张小洪、闫卫杰, 轮廓曲线的 LoG 变换及图像共变区域的检测, 自动化学报, 06 期, pp 817-822, 2010/6/15. EI
- [7] 方蔚涛、马鹏、成正斌、杨丹、张小洪, 二维投影非负矩阵分解算法及其在人脸识别中的应用, 自动化学报, 09 期, pp 1503-1512, 2012/9/15. EI
- [8] 李博、杨丹、雷明、葛永新, 基于近邻消息传递的自适应局部线性嵌入, 光电子.激光, 05 期, pp 772-778, 2010/5/15. EI
- [9] 李博、杨丹、邓林, 移动机器人闭环检测的视觉字典树金字塔 TF-IDF 得分匹配方法, 自动化学报, 37(06), pp 665-673, 2011. EI
- [10] 杨丹、李博、赵红, 鲁棒视觉词汇本的自适应构造与自然场景分类应用, 电子与信息学报, 09 期, pp 2139-2144, 2010/9/15. EI
- [11] 冯欣、杨丹、张凌, 基于视觉注意力变化的网络丢包视频质量评估, 自动化学报, 11 期, pp 1322-1331, 2011. EI
- [12] 霍东海、杨丹、张小洪、洪明坚, 一种基于主成分分析的 Codebook 背景建模算法, 自动化学报, 04 期, pp 591-600, 2012. EI
- [13] 马鹏、杨丹、方蔚涛、葛永新、张小洪, 基于光滑性和主成分的非负矩阵分解算法, 计算机应用, 05 期, pp 1362-1365, 2012/5/1. ISTP
- [14] 林晓泽、周絮语、李相军, 基于轮廓的旋转和尺度不变区域的检测, 计算机应用研究, 05 期, pp 1951-1953+1958, 2012/5/15. ISTP
- [15] 张小洪、李相军、林晓泽, 基于知识的部分匹配方法研究, 计算机应用研究, 06 期, pp 2397-2400, 2012/6/15. ISTP

(三) 其他需要说明的问题

1. 申请人同年申请不同类型的国家自然科学基金项目情况(列明同年申请的其他项目的项目类型、项目名称信息, 并说明与本项目之间的区别与联系)。

2. 具有高级专业技术职务(职称)的申请人或者主要参与者是否存在同年申请或者参与申请国家自然科学基金项目的单位不一致的情况; 如存在上述情况, 列明所涉及人员的姓名, 申请或参与申请的其他项目的项目类型、项目名称、单位名称、上述人员在该项目中是申请人还是参与者, 并说明单位不一致原因。

3. 具有高级专业技术职务（职称）的申请人或者主要参与者是否存在与正在承担的国家自然科学基金项目的单位不一致的情况；如存在上述情况，列明所涉及人员的姓名，正在承担项目的批准号、项目类型、项目名称、单位名称、起止年月，并说明单位不一致原因。

4. 其他。