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（建议 8000 字以下）：

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 研究意义

在海量数据的支撑下，机器学习尤其是深度学习算法在诸如计算机视觉、自然语言处理等领域取得了巨大的成功。机器学习的理想场景是有大量带标记的训练实例，并且训练数据与测试数据具有相同的分布。然而在许多现实情况下，收集足够的带标记的训练数据通常是耗时、代价昂贵甚至无法实现的。并且在机器学习被使用的诸多领域中，训练集与测试集数据独立同分布的假设往往并不能够成立。面对数据分布的偏移，传统的机器学习算法训练得到的模型不能在相似的新领域中取得预期的结果，这限制了机器学习模型的泛化能力和知识复用能力。

从另一个角度来看，目前机器学习算法取得成功多是在封闭和静态的环境中，即假设训练数据和测试数据来自相同的语义和特征空间。实际上在很多现实应用中，我们所面临的是动态和开放的环境，比如可能的语义信息随着时间会发生变化，不同来源的数据具有不同的特征表达等等。在这样的环境下，目前已有的机器学习算法也会失效。

本项目旨在通过一个统一研究的目标和技术框架，提出新颖的机器学习算法来应对上述挑战。一方面，既然高质量的有标记数据是代价高昂和不易获取的，一个自然的想法是我们是否可以将其他任务中已有的标记数据重复利用，把其中学到的知识迁移到目标任务中？另一方面，既然动态环境和开放环境可能带来语义空间和特征空间的分歧，我们是否可以提出新的方法去应对从未在训练集中见过的场景？在本项目中，我们以领域自适应和零样本学习为着手点，开展上述两方面的研究，借助跨领域跨模态的监督信息，进行自适应迁移学习，应对标记短缺，动态和开放环境下机器学习中的诸多挑战。

领域自适应（Domain Adaptation, DA）技术可以被用来改善机器学习模型在跨领域间的性能。当目标领域中无法获得大量带标签数据用于训练具有良好性能的机器学习模型时，可以考虑在不同但相关的有大量带标签数据可以被获取的辅助领域进行模型的预训练，然后将训练好的模型进行调整后应用于目标领域，这克服了实际应用中目标域带标签数据稀缺、难以获取的困境。然而，跨域的数据分布差异成为了模型迁移的障碍。领域自适应旨在学习一个模型使得在辅助领域获取的知识能够在目标领域得到很好地泛化，引入领域自适应技术可以减小辅助领域与目标域的数据分布差异，从而实现领域不变知识的跨域迁移和复用。领域自适应是机器学习与计算机视觉范畴内前沿的研究方向之一，并在计算机视觉、生物信息学等方面有极大的应用前景。迁移学习和领域自适应技术有望处理目标领域标注数据稀缺的问题，避免从头进行模型训练的高额成本，提高机器学习模型的普适性和知识迁移复用的能力，因而具备较大的理论研究价值和宽广的应用前景。

更进一步，假设我们无法找到一个合适的源域进行学习，以及我们可能面对的是开放的场景，即测试集中出现的类别从未在训练集中出现过，我们又该如何应对？传统的机器学习模型都假设训练集与测试集中可能的类别信息是一致的。但是对于复杂的现实应用场景，该种假设并不能保证。比如，在生物信息领域，信息安全领域和个性化推荐领域等等，算法需要对未知类别的处理足够鲁棒。甚至在一些类似冷启动推荐的场景中，要专门研究未知类别。这就导致传统机器学习的假设不能成立，从而使得已有的机器学习算法无法在这些场景使用。基于此，本项目拟进一步对零样本学习（Zero-shot Learning, ZSL）展开研究，去解决训练集与测试集不匹配的开放化机器学习场景。

本项目的研究紧密围绕国务院《新一代人工智能发展规划》的总体路线，聚焦人工智能重大科学前沿问题，兼顾当前需求与长远发展，以突破人工智能应用基础理论瓶颈为重点，促进学科交叉融合，为人工智能持续发展与深度应用提供强大科学储备。重点突破自适应学习、零样本学习等理论方法，实现具备高可解释性、强泛化能力的人工智能。一直以来，通用人工智能都被业界视为镶在 AI 皇冠上的一颗明珠，其虽是未来 AI 的必然趋势，但由于技术的欠缺导致这一图景久久不能实现。在走向通用人工智能的路上，模型的迁移能力，泛化能力和小样本学习能力都是必须要面对的挑战，因为不具备这些能力的模型和算法不可能通用。因此，本项目的研究将对摘取 AI 皇冠上的明珠提供一些思路和尝试。

1.2 国内外研究现状和发展动态

1.2.1 领域自适应（Domain Adaptation）

迁移学习（Transfer Learning）试图让机器学习人类类比学习、举一反三的能力，迁移学习的研究受到了人类可以智能地应用在之前学习到的知识来更快更好地解决新的问题这一事实的启发。最早在 NIPS 1995 关于“学会学习”的研讨会上讨论了机器学习领域的迁移学习的基本动机，自 1995 年之后，迁移学习以“学会学习”、“知识迁移”、“终生学习”、“多任务学习”、“归纳迁移”、“增量学习”等不同的名称出现，逐渐引起了关注。2005 年，美国国防部高级研究计划局(DARPA)的信息处理技术办公室(IPTO)对迁移学习给出了一个新的定义：一个系统将先前学习到的知识或技能进行识别并将其在新的任务中加以运用的能力。此后，2010 年发表于 TKDE (IEEE Transactions on Knowledge and Data Engineering) 的综述论文《A Survey on Transfer Learning》系统阐述了迁移学习的研究历程，提出了迁移学习的形式化定义及分类，并将领域自适应划分为迁移学习的子领域之一。

目前常见的领域自适应算法主要分为传统方法（浅层学习算法）和深度学习算法。浅层域适应的常用算法主要分为基于实例的 DA 和基于特征的 DA[1]。在深度学习与 DA 相结合的研究中，文献[1]将深度 DA 分为基于差异、基于对抗和基于重建三大类，文献[2]将其分为基于实例、基于映射、基于网络和基于对抗四大类。文献[3][4]从数据和模型的角度对迁移学习和领域自适应的多种代表性方法进行了概述。文献[5]关注于单源无监督的域适应场景，特别是该设定下的深度域适应方法，根据域偏移损失和生成/判别设定的不同，将深度域适应方法归类为基于差异的方法、基于对抗生成的方法、基于对抗判别的方法、基于自监督的方法四类。文献[6,7]侧重于从特征选择、特征空间对齐的角度对域适应算法进行研究，文献[8,9]基于对抗学习的思想进行了算法的拓展和改进。文献[10]结合元学习、对抗学习、正则化的思想，提出了基于元学习的权重时序正则化域对抗网络。文献[11]从领域分布差异、对抗、重构和样本生成四个角度对深度域适应方法进行了综述，并对跨域标签空间不同的复杂场景进行了概述。

在域适应的应用方面，文献[12]总结了域适应在诸如图像分类、目标检测、语义分割、姿态估计、视频动作检测等计算机视觉领域中的应用。文献[3]总结了域适应方法在医学影像与计算机辅助诊断、生物序列分析、交通场景识别、推荐系统等领域的应用。此外域适应在文本分类、情感分析、相关性提取、机器翻

译等自然语言处理领域也得到了广泛的应用[13]。

从是否有监督，参与的领域数量和特征空间异同等角度出发，领域自适应算法可以做如图 1-1 所示的分类：

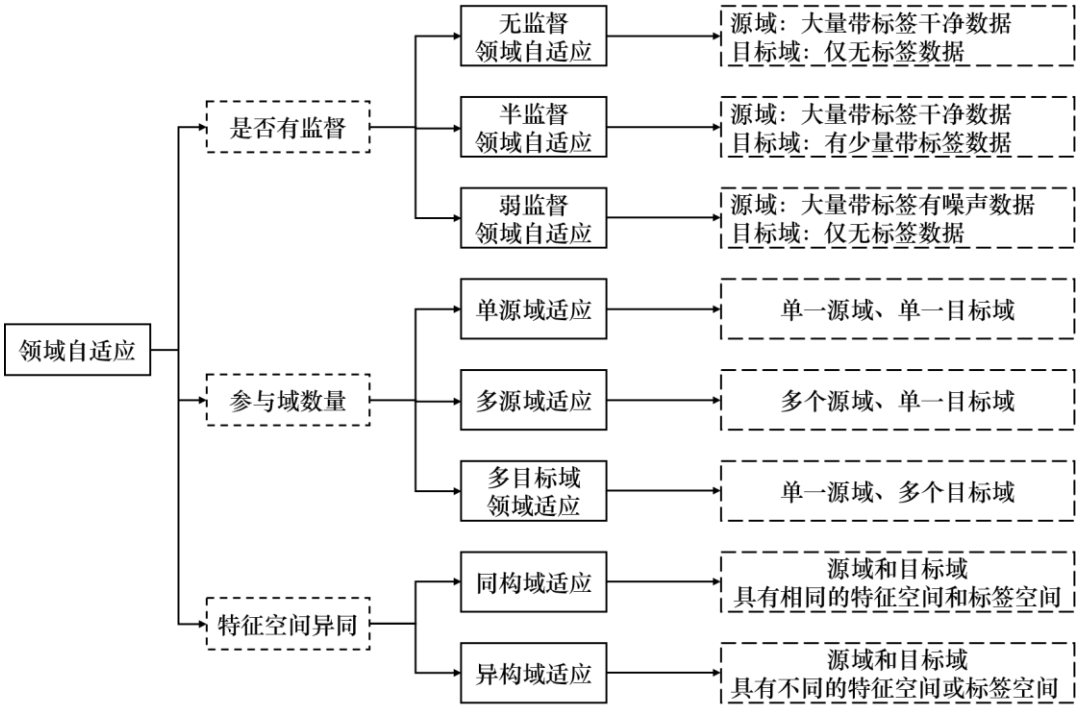

图 1-1 领域自适应算法分类

领域自适应技术是当前人工智能技术迈向通用人工智能的一小步，也是在有限标记数据的情况下最大化数据利用率的行之有效的技术手段。当前主流的领域自适应方法主要有基于度量学习的方法和基于对抗学习的方法。随着领域自适应研究的不断深入，现实鸿沟、语义鸿沟、数据隐私和负迁移将成为今后的研究和实践中需要进一步解决的问题。

1.2.2 零样本学习（Zero-shot Learning）

由于监督学习需要大量的高质量有标记数据，而现实中有的数据标注工作不易进行，有的数据并不存在足够的数据去进行模型训练，为了让模型能对训练时“不可见类”同样有效，零样本学习应运而生。根据文献[15]-[17]，对零样本学习的界定方法有几种，首先根据最终分类时是否加入“可见类”，分为传统 ZSL 和一般化 GZSL（Generalized ZSL）；根据不可见类的实例和语义信息的有无，可以分为 Transductive (G)ZSL, Transductive Semantic (G)ZSL 和 Inductive (G)ZSL；在文献[16]中，将处理图像识别问题的零样本学习称为狭义 ZSL，将应用于更广阔领域的零样本学习定义为广义 ZSL；根据方法的不同可以分为基于分类器与基

于实例的方法，而根据语义空间选择的不同又可以有不同分类。

零样本学习方法的核心思想是利用可见类的视觉信息与语义信息建立语义信息到视觉信息的某种关联，再利用不可见类的语义信息对不可见类进行分类。具体实现方式主要有两种，分别为基于嵌入(Embedding)的方法以及基于生成(Generative)的方法。

基于嵌入的方法是一种比较直观的方法，其通过学习一个投影矩阵或者嵌入函数来关联可见类的视觉特征和其对应的语义向量，之后使用投影方程在嵌入空间中计算不可见类原型表示和预测表示的相似度，由此来识别不可见类。根据嵌入空间的不同，可分为几种方法：语义嵌入方法，包括最早的 DAP 模型（三层 MLP）[18]，以及后来的 ALE（label-embedding，标签映射到相应的属性特征）[19]，SAE（通过恢复到输入层的方法解决域漂移）[20]，以及 SCoRE 中 Deep-RIS（为每个属性建立 CNN，做分类器）和 Deep-RULE（建立语义空间）[21]；还有就是视觉嵌入方法，在[22]中作者认为嵌入到视觉空间要比嵌入到语义空间更好，主要是可以减轻枢纽度问题。同时，嵌入空间也可以是一个潜在空间（Latent Space），这种方法下语义向量和视觉特征共同映射到一个潜在空间当中（或是建立与语义空间平行的潜在空间），目的主要是语义空间和视觉空间可能差距过大，直接映射效果不好，同时通过潜在空间还可以提供一些人为定义的语义信息之外的隐藏语义信息用来支持后续的分类，这类的方法包括结合语义向量与隐向量的 LDF[23],以及 MLSE[24]。

基于生成的方法思路有所不同，它通过生成不可见类的实例（视觉特征）将 ZSL 问题转换成了传统的监督学习问题。这个方法的关键步骤就是生成不可见类样本的实例，主要方法有 GAN，VAE，结合 GAN 和 VAE 的方法（我们的方案主要是基于 GAN，所以此处主要讨论基于 GAN 的方法）。首先是基于 WGAN 的方法，f-CLSWGAN[25]通过 Conditional-WGAN 合成不可见类的视觉特征，在训练时通过分类损失函数来增强合成视觉特征的可分类能力；Cycle-CLSWGAN[26]在[25]基础上，通过添加 Multi-modal Cycle Consistency Loss 来使得合成的视觉特征可以很好的表达其对应的语义信息；Lis-GAN[27]在[25]基础上，通过设置多个真实灵魂样本并使得生成的语义特质接近其中之一，使得生成的样本更加接近真实；AFC-GAN[28]针对 GZSL 中分类时对可见类的偏好问题，提出边界损失函数来减轻该问题。

关于发展趋势，首先，从零样本学习解决的问题角度，我们现在使用的训练

数据通常是与测试实例具有相同特征空间并且具有相同语义类型的训练实例，而现实中我们有时需要在训练数据与实际使用数据的特征空间与语义类型不完全一致的情况下实现零样本学习，处理动态和开放场景是未来的一个发展方向；同时，深度学习中还有许多图像识别之外的任务，在这些任务中使用 ZSL 也能大大提升效率降低成本，在针对不同任务的时候，亦可以注意输入数据的不同，如精细图像处理时，可以尝试特征金字塔融合的方法；在一些任务中，多模态数据的特征提取也是一个不错的选择。

1.3 迫切需要解决的问题及应用前景

通过前期已经开展的研究，我们认为目前在面向跨领域跨模态迁移的自适应机器学习算法的研究中还存在如下一些需要迫切解决的问题：

（1） 动态场景下的现实鸿沟问题

在已有的算法研究中通常认为可在同一时刻获取到大量的源域带标签数据，并且这些数据具有相似的分布。但是在现实应用中，源域的数据往往是以流式形态（Streaming）不断产生的，并不能够在某一时刻获得大量的数据。同时，不同时间产生的源域数据也可能具有不同的数据分布。如何识别源域数据中的概念漂移（Concept Drift），以及如何进行增量域适应和在线域适应(Online Domain Adaptation)是有待迫切研究的问题。此外，当数据不断进化产生时，如何处理不同时期模型的灾难性遗忘（Catastrophic Forgetting）也成为迁移学习中的难点。

（2） 开放场景下的语义鸿沟问题

目前大多领域自适应算法研究封闭集中的跨域知识迁移，即通常假设源域和目标域共享类别标签，但是在现实场景中，源域和目标域具有相同的类别标签空间这一假设往往不能够成立，存在源域数据类别多于目标域、目标域数据类别多于源域等情况。针对前者，已有文献尝试了对源域实例加权来增强跨域共享标签实例的重要性。后者所面临的挑战更加复杂，也更贴近现实问题。在本项目中，我们尝试使用零样本学习和小样本学习等技术学习关于新的类别的知识，并针对开放集的跨域迁移开展进一步研究。

（3） 分布式场景下的数据隐私问题

目前的迁移学习方法大多假设带标签的源域数据可以不受限制地获取，在现实场景中，与目标域相关的源域或辅助领域数据可能来自于另外的个人和机构，可能无法访问数据的全部信息。此类情况下，如何在进行跨域知识迁移的同时保护好数据的隐私是一个重要的问题。开发基于模型参数而不是基于数据特征的域

适应技术以及开发基于加密数据的域适应技术是需要迫切研究的方向。在该领域，联邦学习已经做了一些尝试。本项目中，我们将分布针对源域数据不可访问和目标域数据不可访问的情形进行研究。

在应用前景方面，香港科技大学的杨强教授是迁移学习方面的领军人物，他以“昨天、今天、明天”的比喻来说明深度学习、强化学习与迁移学习的关系。深度学习领域的著名人物，斯坦福大学吴恩达教授在接受采访时也表示迁移学习将是他未来五年的重要研究方向。本项目的研究成果可广泛应用于图像分类、目标检测、语义分割、姿态估计、视频动作检测等计算机视觉和多媒体领域中，以及在医学影像与计算机辅助诊断、生物序列分析、交通场景识别、推荐系统等领域中得到应用。此外，还可应用在文本分类、情感分析、相关性提取、机器翻译等自然语言处理领域。通过提升模型的泛化能力、迁移能力和小样本学习能力，为通用人工智能的发展提供一定的推动，并提升数据的利用率，显著降低人工智能系统的部署成本。

附：主要参考文献

[1] Wang M, Deng W. Deep visual domain adaptation: A survey[J]. Neurocomputing, 2018, 312: 135-153.

[2] Tan C, Sun F, Kong T, et al. A survey on deep transfer learning[C]//International conference on artificial neural net-works. Springer, Cham, 2018: 270-279.

[3] 庄福振, 罗平, 何清 等.迁移学习研究进展[J].软件学报, 2015, 26(1):26-39.

[4] 刘建伟, 孙正康, 罗雄麟. 域自适应学习研究进展[J].自动化学报, 2014, (8):1576-1600.

[5] Zhao S, Yue X, Zhang S, et al. A Review of Single-Source Deep Unsupervised Visual Domain Adaptation[J]. IEEE Trans-actions on Neural Networks and Learning Systems, 2020.

[6] 袁晨晖, 程春玲. 基于 PE 散度实例过滤的深度域适应方法[J].计算机科学, 2020, 47(8): 151-156.

[7] 毕朝阳. 基于特征选择的领域自适应方法研究[D]. 广州：华南理工大学, 2019.

[8] 王格格, 郭涛, 余游 等. 基于生成对抗网络的无监督域适应分类模型[J]. 电子学报, 2019, 48(6): 1190-1197.

[9] 毛潇锋. 基于对抗学习的深度视觉域适应方法研究[D]. 哈尔滨：哈尔滨工程大学, 2019.

[10] 陈迪. 基于对抗和正则化方法的域适应算法研究[D].成都：电子科技大学,2020.

[11] 范苍宁, 刘鹏, 肖婷 等. 深度域适应综述:一般情况与复杂情况[J].自动化学报, 2020, 46: 1-34.

[12] Csurka G. Domain adaptation for visual applications: A comprehensive survey[J]. arXiv preprint arXiv:1702.05374, 2017.

[13] Wilson G, Cook D J. A survey of unsupervised deep do-main adaptation[J]. ACM Transactions on Intelligent Systems and Technology (TIST), 2020, 11(5): 1-46.

[14] Pan S J, Yang Q. A survey on transfer learning[J]. IEEE Transactions on knowledge and data engineering, 2009, 22(10): 1345-1359.

[15] Pourpanah, F., Abdar, M., Luo, Y., Zhou, X., Wang, R., Lim, C., & Wang, X. (2020). A Review of Generalized Zero-Shot Learning Methods. ArXiv, abs/2011.08641.

[16] W. Wang, V. W. Zheng, H. Yu, and C. Miao, "A survey of zero-shot learning: Settings, methods, and applications," ACM Transactions on Intelligent Systems and Technology, vol. 10, no. 2, pp. 1–37, 2019.

[17] Paul, A., Krishnan, N.C., & Munjal, P. (2019). Semantically Aligned Bias Reducing Zero Shot Learning. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 7049-7058.

[18] Lampert, C.H., Nickisch, H., & Harmeling, S. (2009). Learning to detect unseen object classes by between-class attribute transfer. 2009 IEEE Conference on Computer Vision and Pattern Recognition, 951-958.

[19] Akata, Z., Perronnin, F., Harchaoui, Z., & Schmid, C. (2013). Label-Embedding for Attribute-Based Classification. 2013 IEEE Conference on Computer Vision and Pattern Recognition, 819-826.

[20] Kodirov, E., Xiang, T., & Gong, S. (2017). Semantic Autoencoder for Zero-Shot Learning. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 4447-4456.

[21] Morgado, P., & Vasconcelos, N. (2017). Semantically Consistent Regularization for Zero-Shot Recognition. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2037-2046.

[22] Zhang, L., Xiang, T., & Gong, S. (2017). Learning a Deep Embedding Model for Zero-Shot Learning. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 3010-3019.

[23] Li, Y., Zhang, J., Zhang, J., & Huang, K. (2018). Discriminative Learning of Latent Features for Zero-Shot Recognition. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 7463-7471.

[24] Ding, Z., & Liu, H. (2019). Marginalized Latent Semantic Encoder for Zero-Shot Learning. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 6184-6192.

[25] Xian, Y., Lorenz, T., Schiele, B., & Akata, Z. (2018). Feature Generating Networks for Zero-Shot Learning. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 5542-5551.

[26] Felix, R., Kumar, B.V., Reid, I., & Carneiro, G. (2018). Multi-modal Cycle-consistent Generalized Zero-Shot Learning. ECCV.

[27] Li, J., Jing, M., Lu, K., Ding, Z., Zhu, L., & Huang, Z. (2019). Leveraging the Invariant Side of Generative Zero-Shot Learning. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 7394-7403.

[28] Li, J., Jing, M., Lu, K., Zhu, L., Yang, Y., & Huang, Z. (2019). Alleviating Feature Confusion for Generative Zero-shot Learning. Proceedings of the 27th ACM International Conference on Multimedia.

[29] Chen Y, Li W, Sakaridis C, et al. Domain adaptive faster r-cnn for object detection in the wild[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 3339-3348.

[30] Hoffman J, Wang D, Yu F, et al. Fcns in the wild: Pix-el-level adversarial and constraint-based adaptation[J]. arXiv preprint arXiv:1612.02649, 2016.

[31] Inoue N, Furuta R, Yamasaki T, et al. Cross-domain weak-ly-supervised object detection through progressive domain adaptation[C]//Proceedings of the IEEE conference on com-puter vision and pattern recognition. 2018: 5001-5009.

[32] Zhang Y, Barzilay R, Jaakkola T. Aspect-augmented ad-versarial networks for domain adaptation[J]. Transactions of the Association for Computational Linguistics, 2017, 5: 515-528.

[33] Liu P, Qiu X, Huang X. Adversarial multi-task learning for text classification[J]. arXiv preprint arXiv:1704.05742, 2017.

[34] Li, J., Jing, M., Zhu, L., Ding, Z., Lu, K., & Yang, Y. Learning Modality-Invariant Latent Representations for Gener-alized Zero-shot Learning[C]// In Proceedings of the 28th ACM International Conference on Multimedia . 2020:1348-1356.

[35] Fu L, Nguyen T H, Min B, et al. Domain adaptation for relation extraction with domain adversarial neural net-work[C]//Proceedings of the Eighth International Joint Confer-ence on Natural Language Processing (Volume 2: Short Papers). 2017: 425-429.

[36] Yang Z, Hu J, Salakhutdinov R, et al. Semi-supervised qa with generative domain-adaptive nets[J]. arXiv preprint arXiv:1702.02206, 2017.

[37] Britz D, Le Q, Pryzant R. Effective domain mixing for neural machine translation[C]//Proceedings of the Second Conference on Machine Translation. 2017: 118-126.

[38] Li, J., Jing, M., Lu, K., Ding, Z., Zhu, L., & Huang, Z. Leveraging the invariant side of generative zero-shot learn-ing[C]// In Proceedings of the IEEE/CVF Conference on Com-puter Vision and Pattern Recognition. 2019:7402-7411.

[39] Li, J., Jing, M., Lu, K., Zhu, L., Yang, Y., & Huang, Z. From zero-shot learning to cold-start recommendation[C]// In Proceedings of the AAAI Conference on Artificial Intelligence 2019: 4189-4196.

[40] Pan W, Xiang E W, Yang Q. Transfer Learning in Collabo-rative Filtering with Uncertain Ratings[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2012, 12: 662-668.

[41] Pan W, Xiang E W, Liu N N, et al. Transfer learning in collaborative filtering for sparsity reduction[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2010, 10: 230-235.

[42] Li J, Lu K, Huang Z, et al. On both cold-start and long-tail recommendation with social data[J]. IEEE Transactions on Knowledge and Data Engineering, 2019, 33(1): 194-208.

[43] Li, J., Lu, K., Huang, Z., & Shen, H. T. Two birds one stone: on both cold-start and long-tail recommendation[C]// In Proceedings of the 25th ACM international conference on Mul-timedia, 2017:898-906.

[44] Li, J., Jing M., Lu K., Zhu L., Shen, H.T., Investigating the Bilateral Connections in Generative Zero-shot Learning, IEEE Transactions on Cybernetics, 2021.

[45] Li, J., Chen, E., Ding Z., Zhu L., Shen, H.T., Maximum Density Divergence for Domain Adaptation, IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2021.

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 研究内容

本项目的研究动机和出发点主要有两个，第一是尽可能减少深度模型训练过程中对目标任务有标记数据的需求，第二是提升深度模型在开放和动态场景下的自适应能力以及泛化能力。

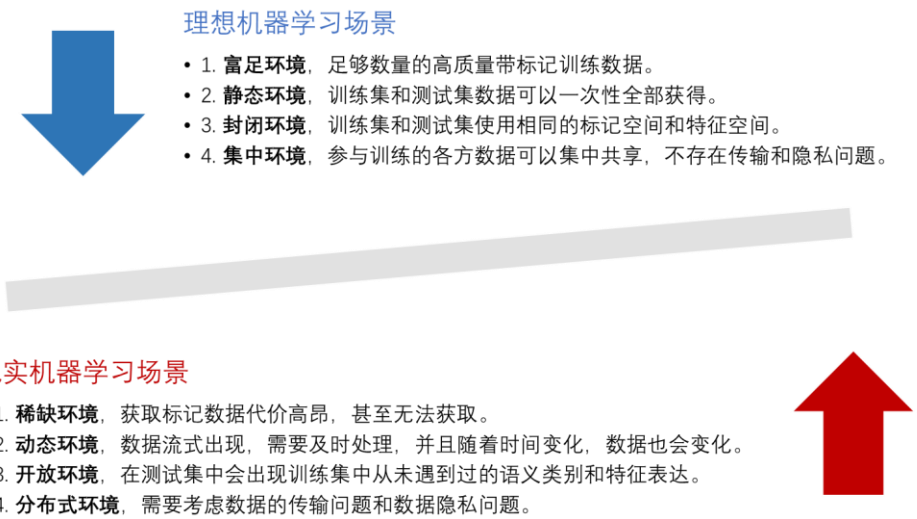

图 2-1 现实机器学习场景中存在的挑战与本项目的研究动机

为了应对这两个问题，我们从领域自适应和零样本学习两个角度出发，进行具体的研究。在我们前期的工作中，我们发现目前存在的主要挑战包括：1）动态场景下的现实鸿沟问题，即在在线（Online）和进化（Evolving）环境下，我们无法一次性获得所有的目标域数据，并且随着时间的推移目标域的数据可能发生显著变化，这与当前已有的工作所假设的静态封闭的环境是完全不一样的，存在很大的挑战。2）开放场景下的语义鸿沟问题，在一个开放的场景中，我们几乎不可能在训练集中枚举到所有可能的语义类别和特征表达，这就要求我们的模型必须要有零样本学习能力。3）分布式场景下的数据隐私问题，在很多现实应用中，往往会涉及到多个参与者，这多个参与者之间协同训练一个共享的模型，但是彼此又不愿意共享其原始训练数据，这对机器学习模型的训练带来了显著挑战。为了更好地说明问题，我们将理想机器学习场景中的一些假设与现实机器学习场景中的实际情况进行对比，结果如图 2-1 所示，这些挑战也对应了本项目的研究动机。为了应对上述三大挑战，我们在本项目中开展面向跨领域跨模态迁移

的自适应机器学习算法研究，按照侧重点的不同，本项目的研究内容主要包括以下部分：

（1） 在线和进化环境下的领域自适应研究

已有的领域自适应算法多通过学习新的特征表达来对齐源域和目标域的数据分布，从而将源域上训练到的模型迁移到目标域上。在这一学习范式中，带标签的源域和不带标签的目标域数据都参与训练，并且整个数据集都是假设可以被完整访问的。但是，对于实际应用系统，比如一个在线场景下的业务，目标域的数据是流式产生的，无法在某一时刻被完整访问。同时，在线业务又要求及时反馈，因此我们也不能等积攒一定量的数据后再做知识迁移。而目前已有的一些算法，比如基于距离度量的方法和基于对抗学习的方法都会受到源域和目标域数据不均衡的影响。在极端情况下，假设目标域只有一个样本（样本逐个到达），那么目前已有的领域自适应算法几乎都将失效。因此，处理在线场景下的领域自适应更符合现实应用场景，且存在非常大的技术挑战。

另一方面，现实应用中的很多场景并非一成不变。现实应用中涉及到的实例往往会随着时间发生进化。比如对于自动驾驶中的视觉处理系统，其需要应对的业务场景会随着时间、季节和气候等因素不断发生变化。与此同时，我们又不能频繁地重新训练和更改模型。因此，我们不仅要求模型能够处理在进化环境下样本数据分布发生的变化，同时还要要求模型不能发生灾难性遗忘，不仅要能处理环境进化后的新业务，还要能应对之前的老业务。因此，要求模型是能持续学习，并且能克服遗忘的。

目前，我们已在传统的领域自适应算法方面取得了一些显著的成果，积累了大量的经验。但是这些已有的算法几乎都关注的是静态环境。为了加快迁移学习在国民经济主战场的落地，研究在线和进化环境下的领域自适应算法迫在眉睫。因此，本项目拟研究的第一个关键内容为在线和进化环境下的领域自适应。

（2） 基于对抗生成模型的零样本学习研究

在第一个主要研究内容中，我们重点关注了现实应用中的动态环境挑战。事实上，在现实应用中还存在另一大问题，即开放环境挑战。在已有的大多数机器学习模型中，我们都假设训练集和测试集采样自相同的语义空间和特征空间，而这一假设在很多现实应用中是难以满足的。比如，对于一个有用户参与的在线App，在上线之前无法完整预测用户会产生哪些内容。因此，一个对开放环境鲁

棒的机器学习算法，必须要有能处理在训练集中没有见过的类别样本的能力。在本项目中，我们拟从零样本学习的角度对此开展研究。

零样本学习主要是为了解决欠标注场景下的机器学习问题。因为缺少训练数据和高质量标签，传统的有监督机器学习算法无法适用。目前，已有的工作多是通过引入一个新的辅助空间，比如文本描述，来帮助目标空间（多为视觉空间）的学习。在本项目中，我们提出基于生成模型直接生成新类别的训练样本，进而使用新生成的样本（Synthesized/Generated Samples）进行有监督学习，并将学习到的模型应用到真实的新类别样本上。特别地，由于生成模型可以使用随机噪声作为输入，因此在理论上可以生成任意数量的辅助样本用于训练。从而将极富挑战的零样本学习转换为一个普通的有监督学习问题。

但是在零样本学习中，由于我们没有不可见样本（Unseen Samples），模型的训练只能依赖可见样本（Seen Sample）。由此引发的一个难以避免的问题是新类别的生成样本会与可见样本比较相似，但是我们的目标又是希望新类别的生成样本更接近于不可见样本，从而导致特征混淆问题。在本项目的研究中，我们将重点研究如何避免特征混淆，并提出可行的解决方案。同时，由于目前流行的生成模型基于生成对抗网络（GAN），生成的样本可能会比较相似，缺乏多样性，即产生模型崩塌（Mode Collapse）问题。本项目也将展开研究尝试解决生成式零样本学习中的模型崩塌问题。

（3） 领域数据受限访问场景下的迁移学习研究

当前已有的迁移学习算法多假设数据是集中管理，可任意访问的。但是对于很多现实应用来讲，无论是源域数据还是目标域数据都可能是分布式存储的，并且在很多场合分布式的参与方并不愿意共享原始训练数据。比如，对于一个在线医学辅助诊断系统，多家医院可共同训练该模型以提高系统的准确率和可靠性，但是参与的多家医院的数据分布式地存储在不同的地方，并且这多家医院并不愿意将自己的原始数据贡献出来。在这样的场景下，我们不仅仅要考虑数据的传输开销问题，还要考虑数据受限访问的问题。这些问题的解决对于迁移学习应用到实际应用具有强烈的现实意义。

在本项目中，我们以领域数据受限访问场景下的迁移学习研究为主题，对上述问题开展研究。在数据受限访问的情况下，我们将重点放在模型上，尝试去最大化模型的泛化能力。特别地，我们考虑两种情况，第一种情况是源域数据受限访问，第二种情况是目标域数据受限访问。在源域数据受限访问的情况下，我们

仅仅提供在源域数据上训练好的有监督模型，此后源域数据不可再访问，我们仅有该训练好的模型和无监督的目标域数据。在此情况下，我们尝试引入自监督学习（Self-supervised Learning）的思想，利用没有标签的目标域数据去更新在源域上训练好的模型，并尝试最大化该模型的泛化能力。在目标域受限访问的情况下，往往需要多个源域共同学习，这一设定通常被称为领域泛化（Domain Generalization）。由于我们没有目标域数据，因此很自然的想法是去学习多个源域中的共享信息，并假设这些共享信息也将被嵌入在目标域中，从而将模型顺利迁移到目标域中。可以看出，在目标域受限访问的情况下，我们的目标依然是最大化模型的泛化能力。

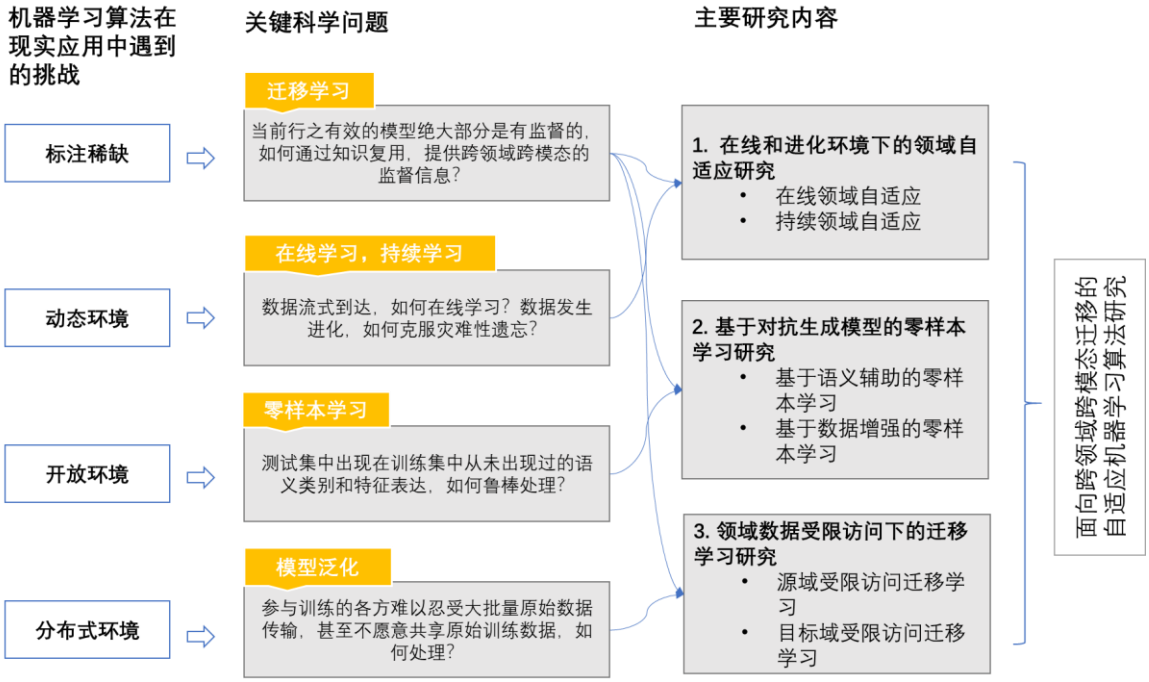

图 2-2 项目的研究动机、研究内容和相互关系

最后是对上述三个研究内容的综合运用。由于目前在工业界和学术界行之有效的机器学习算法几乎都是有监督学习算法。若能提出革命性的无监督学习算法当然是我们所渴望完成的。但是这一目标需要长期的探索和持续的投入。在本项目中我们基于已有的研究基础，尝试通过跨领域和跨模态的迁移学习来应对领域内所存在的一些挑战。在本项目的研究中，我们以领域自适应和零样本学习为着手点，重点关注他们在现实应用中遇到的瓶颈问题。最后，我们借助计算机视觉和多媒体为应用场景，开展对算法有效性的验证。我们的研究思路、动机和对应关系如图 2-2 所示。可以看出，我们的研究内容不是各自隔离的，而是相互关联、彼此支撑、从而共同组成一个有机的整体。我们为了解决标注稀缺问题，开展面

向跨领域和跨模态监督的迁移学习。但是已有的迁移学习算法存在诸多理想化的假设，因此我们针对动态环境，开放环境和分布式环境，分而治之，并最终形成一套整体学习框架，在国际上占领迁移学习领域的研究高地，服务国家重大需求。

2.2 研究目标

当前，尽管我们处于大数据时代，但是代价高昂、不易获取的高质量标记数据依旧是机器学习发展的瓶颈。为了突破这一瓶颈，本项目开展面向跨领域和跨模态监督的迁移学习研究。其目标是通过迁移相关领域中的有标记数据到目标任务中，辅助完成目标任务。一方面可以减少目标任务对有标记数据的依赖，另一方面提高了数据的利用率，从整体上降低了机器学习的代价。在我们已有的研究中，已经对迁移学习做了一些探索，但是这些探索还依赖于一些较为理想化的假设。在本项目的研究中，我们将重点突破这些假设的限制，全面应对迁移学习在现实中的挑战。结合具体研究内容，本项目的研究目标可以细化为如下几个方面：

（1） **提出新颖的领域自适应算法和模型应对动态场景所带来的分布差异和不均衡挑战。**传统的基于度量学习的方法和基于对抗学习的方法均无法直接应用到在线场景。因此，本项目的研究中需提出全新的领域自适应思路来处理源域和目标域之间存在的巨大不平衡性。另一方面，在进化场景中，基于神经网络的模型会产生灾难性遗忘（Catastrophic Forgetting），本项目的研究需要保证既能鲁棒适应进化环境，又能克服遗忘。

（2） **提出新颖的零样本学习算法和模型应对开放场景所带来的语义分歧挑战。**我们一方面基于生成模型，从根源上挑战零样本学习，利用生成模型扩充训练数据。另一方面，解决生成式零样本学习中存在的特征混淆问题和模式坍塌问题，提升广义零样本学习准确率。

（3） **提出新颖的模型泛化策略应对分布式场景带来的数据隐私和数据传输挑战。**以往的迁移学习算法多关注数据分布的对齐，但是在数据访问受限的情况下，该思路无法直接应用。在本项目中，我们将研究的焦点从数据转移到模型，尝试从自监督学习和对抗攻击等角度出发，最大化模型的泛化能力。

最终，本项目综合运用上述研究成果，形成一套整体学习框架，为解决标注稀缺问题提供行之有效的方案。同时，在该过程中发表一系列高水平研究成果，争取在该领域国际学术界的领先优势，并培养一批研究生和本科生。

2.3 拟解决的关键科学问题

本项目所涉及的领域内现有的算法多在一个理想化的假设框架下开展，本项目尝试突破这些理想化的假设，通过跨领域和跨模态的监督信息，使迁移学习可以自然应对现实应用中的问题。因此，本项目的研究具有较高的难度，虽然我们有一些前期研究基础，但是为了实现本项目的研究目标，我们认为尚有如下一些关键问题有待应对：

（1） 数据流式到达的情况下，源域和目标域之间数据的数量分布会非常不均衡，应该如何应对？（**在线领域自适应**）

（2） 为了应对不断进化的环境，我们的目标模型需要不断自我更新，在更新的过程中，如何保证不会产生灾难性遗忘？即学习了新的知识后，忘记了之前学习的知识怎么办？（**自适应持续学习**）

（3） 生成式零样本学习方法虽然可以生成任意数量的样本，但是生成的过程是以可见类作为参照的，这就导致生成的不可见类别极易与可见类混淆，如何克服这种混淆，并保证生成样本的多样性？（**生成式零样本学习**）

（4） 在数据访问受限制的情况下，我们将面临“巧妇难为无米之炊”的困境，如何从模型设计和训练出手，使得模型的泛化能力最大化？（**无源领域自适应与领域泛化**）

这些关键科学问题均对应于本项目的具体研究内容，对于这些问题的研究方案和可行性分析，我们在下一节详细介绍。

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 研究方案

下面我们就本项目要解决的四个关键科学问题，分别介绍各自拟采用的研究方案。

（1） 在线领域自适应（Online Domain Adaptation）研究方案

现有的领域自适应方法大多假设目标数据是提前收集好的（静态性假设），通过减少源域和目标域之间的分布差异，将知识从有标记的源域转移到无标记的目标域，实现跨域目标识别。但在实际应用中，理想的情况下也仅有少量的目标数据是能提前收集得到，大量目标数据往往是后续在线到达的（动态性现状），

如图 3-1 所示。在这样的情况下，不仅源域与目标域数据之间存在差异，随着时间和周围环境的变化，后续在线到达的数据往往还会与提前收集到的数据之间出现差异。由于少量的目标数据难以反映目标域的整体分布，现有的域适应方法训练的模型仅能学习到收集的部分目标域数据的分布，面对在线到达的数据难免出现性能下降的问题。

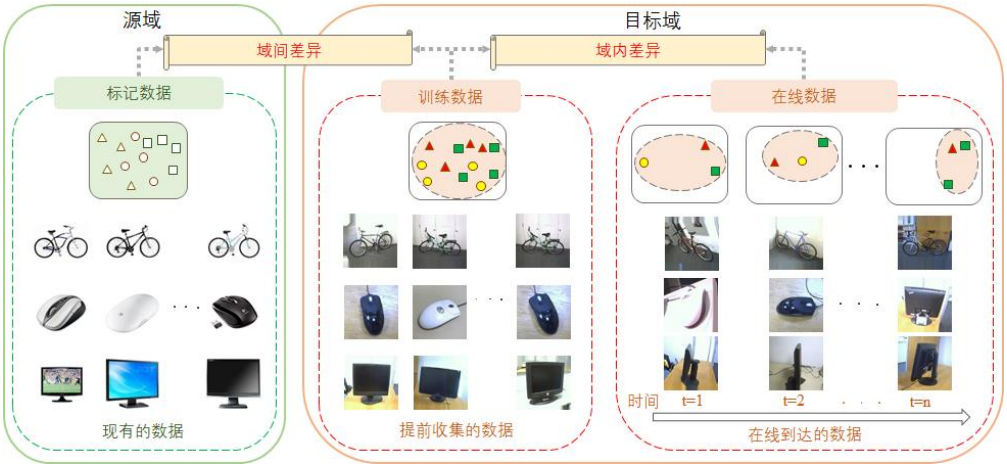

图 3-1 在线领域自适应问题示意图

针对在线领域自适应所面临的挑战，现有的解决思路主要有：1）调整位于分类边界的样本；2）利用主动学习进行模型更新；3）每隔一定时间重新训练模型。但是，第一种方案面对目标数据变化较大的场景效果不佳，而且调整时需计算新样本与现有样本的相似度，随着时间的推移，计算量会不断加大，难以满足实际需要。第二种方案在实际应用中难以实时的获得新来样本的标签以满足更新的需要，同时随着时间的变化，可能会出现重复的数据，而变化后的模型可能需要重新适应重复数据来满足实际需求，这会加大实际应用的难度。第三种方案每隔一段时间进行一次域适应和模型训练，需要消耗大量的计算资源，很难适用于许多实际应用，而且实际应用会面临一些其它问题：如这段时间到达的数据过少，到达的数据类别不均衡或出现类别缺失等，这些问题使得方法难以有效的进行域适应，无法满足实际应用的需要。

针对在线域适应的挑战和现有研究思路的不足，我们提出了两种面向不同在线场景下的域适应方案：1）面向样本批量到达的在线无监督域适应方案；2）面向样本逐个到达的在线无监督域适应方案。现分别说明如下：

针对样本批量到达的情况，我们提出联合减少源域数据与目标域数据之间的域间差异和目标域中训练数据与在线数据之间的域内差异来实现在线域适应。该

方案包括分类器训练阶段和在线识别阶段。在分类器训练阶段，和现有的无监督域适应方法类似，为了减少训练分类器的域间差异，我们提出将提前收集到的源域数据和目标域数据投影到一个共享特征空间上，使得源域数据和目标域数据在这个共享特征空间上分布差异最小。然后在该特征空间上基于带标记的源域数据训练一个适应于目标域数据的分类器，这个分类器能很好的对目标域数据进行分类。在在线识别阶段，我们提出通过减少域内差异来对齐新到达的数据和训练数据，使得新来的数据尽可能地服从已学习到的共享特征空间的分布，这样基于共享特征空间训练的分类器就能很好地识别在线数据。基于这两个阶段，我们提出的方案能够很好地处理无监督环境下的在线跨域识别任务。综上所述，我们的目标函数可以表达为：

$$\min D_1(P_s^T X_s, P_t^T X_t) + D_2(P_t^T X_t, P_i^T X_{t_i})$$

其中，D_1表示域间差异，D_2表示域内差异。在分类器训练阶段，我们提出利用联合策略来减少源域和目标域之间的域间差异，首先通过最小化边缘概率分布和条件概率分布来减少域偏移并以此为损失函数学习共享的特征空间。然后，我们提出在共享特征空间中训练概率分类器。根据标记概率，基于信息熵对样本进行权重选择，并根据局部重要性保留几何结构。最后，重复优化直到模型收敛。

在在线识别阶段，当目标数据以流的方式到达时，我们首先通过分类器训练阶段学习到的特征表达函数P_t将新到达的数据映射到共享空间中。根据输入的数据，从训练的共享特征空间中提取潜在子空间，使得潜在的子空间与新到达的数据类别几乎一致。接下来，通过最小化新到达数据的子空间和潜在子空间之间的差异，将新到达的数据对齐到潜在子空间。在线域对齐的目标函数为：

$$\min\left|P_i^T X_{t_i} - Z\right|_F$$

其中Z表示利用潜在语义分析从共享子空间中提取到的潜在子空间。最后，利用分类器训练阶段训练好的分类器就可以得到在线识别的结果。

更进一步，我们考虑更富挑战性的场景，即目标域样本逐个在线到达的情形。面对数据一个一个到达的情况，我们无法获得目标域数据的分布信息，需要根据在线到达的数据对模型进行在线更新以适应在线数据。如图 3-2 所示，我们提出在训练过程中，利用超图嵌入和分布匹配的联合优化来学习潜在分布空间，然后在该空间上基于聚类策略训练基本分类器。在在线识别过程中，我们使用学习到

的分类器对在线数据进行分类。然后通过连接新到达的样本来更新超图和聚类中心。

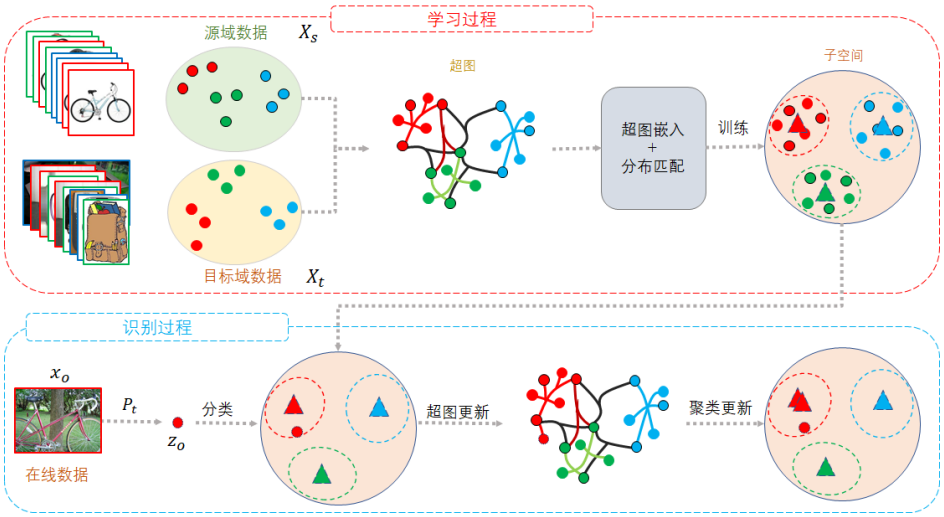

图 3-2 利用超图嵌入和分布匹配的联合优化方案示意图

在训练过程中，利用超图模型能反映数据高阶相关信息的特点，根据源域和目标域的数据结构、源域的标签信息和目标域的聚类信息来构建超图。通过最小化源域和目标域之间的分布差异和超图结构损失来实现域适应。

（2） 自适应持续学习（Adaptative Continual Learning）研究方案

在动态环境中，随着时间的推移，目标模型所要处理的主要任务会发生一定的变化。比如对于一个面向自动驾驶的机器学习模型，最常见的训练数据是正常天气下采集的样本，遇到极端天气后，我们希望模型能鲁棒应对，这就需要对模型参数做一定的自适应更新。但是在参数更新的过程中，我们希望模型还是能处理之前已经熟悉的任务，即依然能处理正常天气环境下的事务。近年来，深度神经网络(DNNs)在计算机视觉领域取得了巨大的成功。然而，最近的研究表明，深度神经网络会产生灾难性遗忘（Catastrophic Forgetting），即深度神经网络在当前任务训练时，会忘记从前一个任务中学到的知识。为了解决这个问题，已有一些工作提出了通过持续学习来增强深度神经网络的长期记忆。但是由于持续学习非常具有挑战性，现有的工作极大地简化了设置以模拟连续在线多任务学习这一设定。具体来说，现有的工作通常将一个数据集分成多个不相交的子集，从而得到多个任务，这些任务遵循相同的边缘分布。我们认为这种设置过于简单，无法近似真实的应用程序。在真实世界的场景中，顺序到达的任务的数据分布会不时发生显著的变化，图 3-3 给出了一个示例，其中包含三个任务。在已有的方法中，

如左图所示，不同的任务从相同的数据集中分离出来，忽略了不同任务之间的分布差距。我们的方案，如右图所示，更接近实际情况，考虑到不同任务中不同的样本分布可能会发生显著变化。本项目开展自适应在线持续学习(Adaptative Online Continual Learning，AOCL)的研究。具体来说，我们的方案拟采用对抗性学习的思想，通过一个回放缓冲区来对齐不同任务的分布并记忆之前的任务。

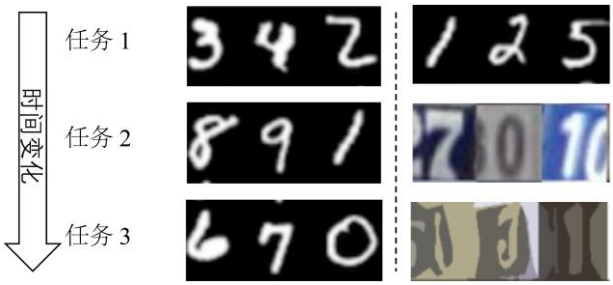

图 3-3 简化的持续学习设定和实际应用中的持续学习设定对比

我们的方案的目标不仅要克服灾难性遗忘，同时要对齐不同任务的特征表示。具体来说，我们利用一个非常小的内存缓冲区（Memory Buffer）来存储以前任务的一些样本，因为基于回放的方法已经被证明对灾难性的遗忘问题是有效的。与此同时，我们训练一个对抗性的网络来减少分配差距。对抗性网络由一个特征生成器和一个鉴别器组成。生成器学习不同任务样本的特征表示，同时，该鉴别器可以区分样本是来自回放缓冲区还是来自当前任务。通过优化二者之间的极大极小博弈，对抗性地训练生成器和判别器。值得注意的是，我们方案中的对抗性网络有两个优点。一方面，由于回放缓冲区中的样本可能与当前任务中的样本有不同的数据分布，混淆鉴别器可以对分布进行对齐。另一方面，由于回放的样本来自旧的任务，混淆鉴别器会促使生成器学习新任务和旧任务之间无法区分的特征，从而记忆旧任务，即克服遗忘。为了便于理解，我们在图 3-4 中给出我们的方案的一个想法示意。为了方便问题描述，我们假设一共有三个任务，在实际应用中可能会存在更多数量的任务需要处理。

在本项目拟采用的方案中，我们将当前任务流中的样本称为目标域样本，而在回放缓冲区中的样本称为源域样本。则我们的解决方案可以看作是持续学习和领域自适应的融合。我们的模型有如下特点：1）使用有监督学习损失来处理有标签的样本；2）使用软标签来更新回放缓冲区中的目标域样本；3）使用对抗学习思想来对齐源域和目标域之间的分布差异。

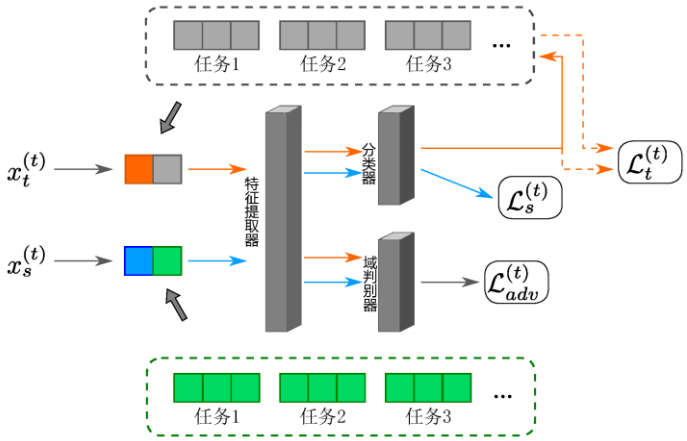

图 3-4 自适应在线持续学习示意图

（3） 生成式零样本学习（Generative Zero-shot Learning）研究方案

零样本学习当前的主要方案有嵌入式（Embedding）模型和生成式（Generative）模型。其中，嵌入式模型一般包括特征-属性嵌入、属性-特征嵌入、中间潜在层嵌入三类。其主要是通过将可见类和不可见类的样本/特征与语义属性都映射到同一嵌入空间中，以完成样本的识别。而生成式模型一般可以分为基于 GAN（生成对抗网络）的零样本学习模型和基于 VAE（变分自编码器）的零样本学习模型。生成式模型通过直接从语义属性中生成不可见类的样本，不断完善合成的样本的真实性，并利用这些样本直接训练有监督分类器，从而完成零样本学习分类。其中，基于 GAN 的零样本学习模型，近年来由于 WGAN 的发展，逐渐成为了生成式模型中的主流。由于生成式模型可以将零样本学习转换为一个有监督的学习问题，并且扩充不可见样本的数量，因此其具有很大的发展潜力。在本项目的研究方案中，我们将主要以生成式零样本学习为基础模型。

在生成式零样本学习模型中存在两个很大的挑战，一是特征混淆，二是模式坍塌。其中特征混淆主要出现在广义零样本学习中，在该设定下，大量的可见类样本中夹杂少量不可见类样本。需要模型同时识别出可见类和不可见类的样本。然而，在模型的训练过程中，更多的是采用可见类样本和语义属性来训练，因此，模型将会更加倾向于将不可见类样本也标注为可见类样本，从而造成识别率大幅下跌。为了减轻广义零样本学习中不可见类容易偏向于可见类样本的问题，以往的工作如 f-CLSWGAN 提出采用 WGAN 方法直接利用语义属性生成以假乱真的不可见类特征。然而，由于 WGAN 模型是在可见类特征和语义属性上训练的，基于 WGAN 生成的不可见类特征依然会与可见类特征有较高的相似度。我们所

提出方案的主要思路有：1）既然存在特征混淆问题，是否可以提出一种新的度量标准以量化特征混淆？一旦特征混淆问题可以被量化，我们就可以直接通过最小化该度量来减轻特征混淆；2）特征混淆是指可见类和不见类内不容易被区分，即他们之间的决策边界比较模糊。因此，我们尝试提出一种边界损失函数，该损失函数可以最大程度地明确可见类和不可见类的决策边界；3）受 CycleGAN 的启发，我们尝试引入一种多模态的循环一致损失，以减轻 GAN 中的模式崩溃并保持合成特征的语义一致性。与现有的基于 GAN 的方法相比，鼓励通过深度非线性映射将合成特征转换回原始语义嵌入信息，从而进一步减轻模式崩溃问题和特征混淆问题。

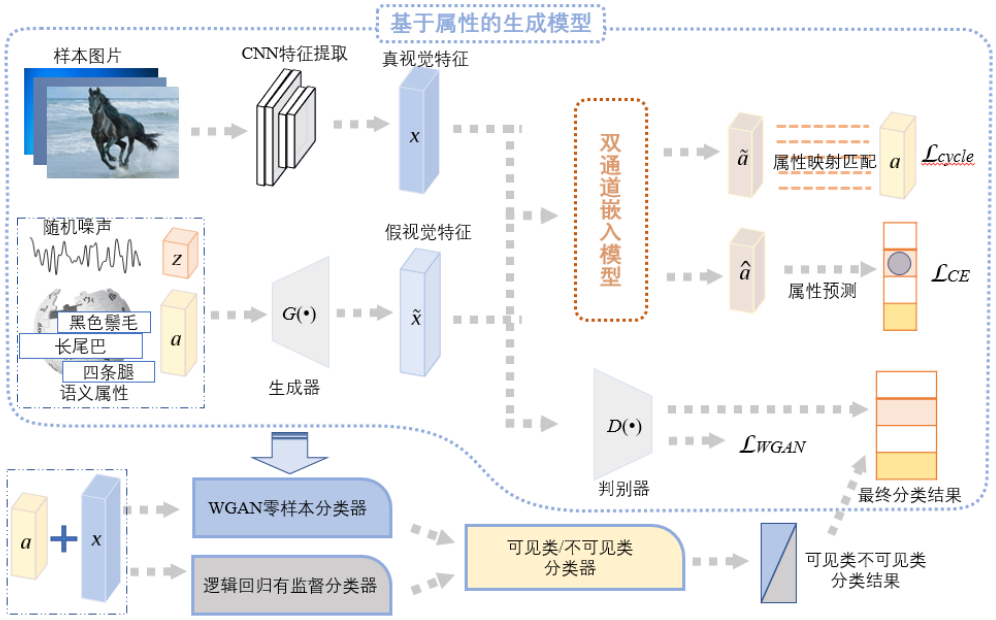

图 3-5 生成式零样本学习方案示意

图 3-5 给出了本方案的一个示意。为了解决零样本学习中的域偏移问题和广义零样本学习识别中不可见类容易偏向于可见类样本的问题，我们认为仅靠原始语义属性不足以判别视觉特征，因此直接从视觉特征中提取潜在语义属性作为原始语义属性的补充，以增强 GAN 模型生成样本的判别能力。为了保持语义一致性，我们的方案中借鉴了 CycleGAN 中多模态的循环一致损失，使得从原始语义属性中生成的视觉特征能够映射回原始语义属性，一方面保证生成样本的可靠性，另一方面防止模式坍塌；此外，我们选择一定的验证集，即取出训练集中的部分样本作为验证集样本，并保证其中含有不可见类样本，基于验证集样本训练出一个可见类/不可见类二元分类器，在广义零样本识别时作为预分类器事先对测试样本二元分类。

（4） 无源领域自适应与领域泛化（Domain Generalization）

传统的领域自适应通常基于度量学习或者对抗学习两种范式来缩小源域和目标域之间的数据分布差异，这些方法都假设在领域适应学习阶段，带标签的源域数据和无标签的目标域数据都是完全可访问的。但是在一些实际场景下，我们无法或很难同时获得源域和目标域的数据。第一，由于训练数据大规模、分布式的特点，预训练模型的使用已经越来越成为大众的形势。第二，出于数据隐私性和安全性考虑，欧盟和一些政府已经相继出台了一系列数据保护政策，因此在很多领域自适应场景中，源域数据会因为上述原因而访问受限，相对地，我们只能使用在源域上预训练的模型。对于目标域而言，在模型实际部署运用时，任何环境、任务、图像方面的细微不同都会导致不同的目标域分布，我们希望模型在任何场景下都表现良好，而这些场景又无法在训练时预先确定，因此目标域样本的收集变得十分困难。传统的领域自适应方法无法解决这些单个领域的数据不可访问的场景，因此一些新的方法需要被提出用于无源或无目标域的领域自适应场景。

由于在源域数据不可见或目标域数据不可见的场景下，领域之间的数据分布差异变得无法或者很难估算，这直接导致了传统的基于数据分布对齐的方法无法用于该类场景。因此在单个领域数据无法访问的场景下，领域自适应面临着无法直接优化的问题。

针对目前无源领域自适应和域泛化方法存在的不足，我们通过深入分析引起领域偏差的根本原因，提出基于对抗攻击的方法来解决这类单个领域数据无法访问的域适应问题。对抗样本是指通过向原始样本添加一个微小扰动，使得在某种度量下（例如人眼），扰动后的样本和原始差距很小，同时模型对于对抗样本会造成显著的误判。这实际上定义了一个绝对正确的第三方（在视觉任务中一般为人眼），并认为在偏差小于一定程度时，样本不应该显著改变模型的输出结果。而在真实世界中，并没有严格的扰动范围限制，当我们训练模型时，我们真正需要的是模型的泛化能力，即模型能够在一定程度上应对没有见过的新样本。从这个角度来看，所有测试集里被分错的样本都是没有范围限制的、天然存在的对抗样本，这也是领域偏移的本质——它们都是由于训练集和测试集分布不同造成的。我们认为，模型的鲁棒性（即模型正确应对人为构造的对抗样本的能力）与泛化性（即模型正确应对自然生成的样本的能力）之间是存在紧密联系的，由此

我们提出在领域自适应中利用对抗攻击和对抗学习的方法来提升模型的鲁棒性，进而提升模型的泛化性能，使得模型能够迁移到目标域。另外，由于对抗攻击的性质，在单个领域的数据缺失的情况下，也完全可以进行，十分适合用于无源或无目标场景下的领域自适应。我们的方案能够用统一的框架来解决无源或无目标的领域适应。在无源领域自适应设定下，我们的问题是如何利用源域预训练的模型以及无标签的目标域样本。而在领域泛化设定下，我们的目标是如何将利用多个源域来学习能够泛化到不可见的目标域的模型。首先我们从对抗攻击出发，考虑在训练集分布P_0下的最坏情况问题：

$$\min_{\theta} \sup_{P:D(P,P_0)\leq\rho} E_P\left[\mathrm{l}(\theta;(X,Y))\right]$$

其中θ表示训练的模型参数，D为距离度量，l为损失函数。解决上述最坏情况问题能够保证模型在应对与训练数据分布距离小于ρ的所有分布下都能够取得较好的效果，而满足$D(P,P_0)\leq\rho$的分布能够表示真实世界中的保留着原语义的领域偏移。因此，我们希望学到的模型在相同的语义空间中即使存在领域偏移也具有良好的性能。

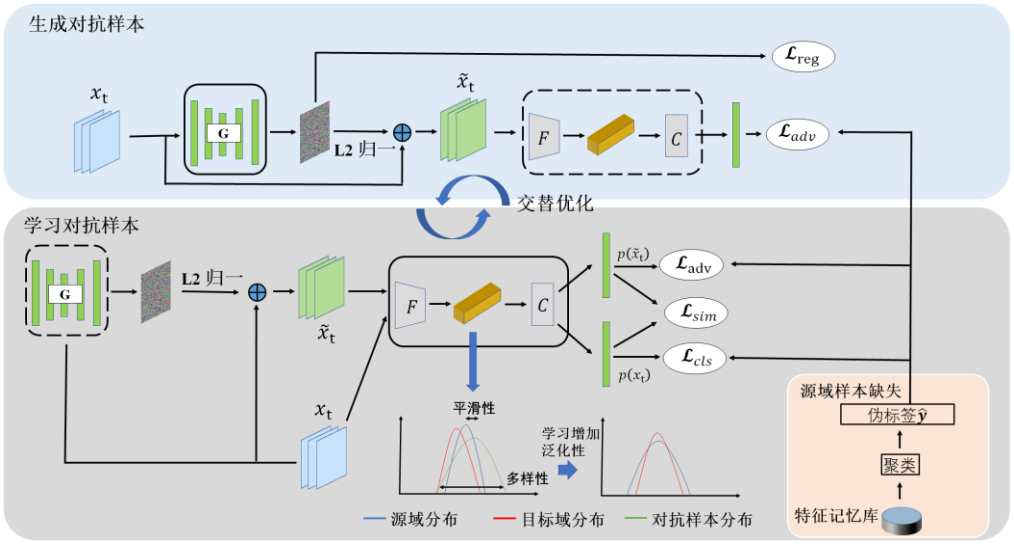

图 3-6 基于对抗攻击的无源无目标领域自适应框架

根据以上观点，我们提出了一个迭代框架来解决上述问题，如图 3-6 所示。我们将模型的迁移分为两个阶段：（1）生成对抗样本以及（2）学习对抗样本。在生成对抗样本阶段，我们通过一个神经网络 G 来产生扰动，使得图片在添加扰动后能够迷惑分类器，对于扰动生成器 G，其主要目标如下：

$$\max_{\theta_G} \mathcal{L}_{adv} = \frac{1}{n_t}\sum_{i=1}^{n_t} l\left(f\left(x_t^i + \epsilon\left(G\left(x_t^i\right)/|G\left(x_t^i\right)|_2\right)\right), y_t^i\right)$$

其中x_t^i为训练样本（在无源领域自适应中，我们使用目标域样本作为训练，反之亦然），n_t为样本数量，在无源领域适应中我们使用在特征空间聚类后的伪标签y_t^i作为攻击的目标类别。通过上述过程，生成器能够在训练集分布附近不断地开拓潜在的数据分布，因此我们的方案将不再受限于可见的训练分布。在学习对抗样本阶段，我们使得模型 F 能够对生成器产生的对抗样本进行抵御，从而增加模型的泛化性。

3.2 可行性分析

针对上一节所给出的研究思路和解决方案，我们已经做了一些初步尝试，开发测试了一些原型算法，以保证所提的方案在技术手段上是合理的，在实验设置上是可行的。详细分析，本项目的可行性主要体现在以下方面：

（1） **项目研究目标明确。**本项目的主要目标是利用跨领域跨模态的监督信息进行迁移学习，来缓解目标域对标注数据的依赖。尽管我们处在大数据时代，但是我们周围的数据几乎都是无标记的数据，而目前行之有效的机器学习算法几乎都是有监督学习算法，这就形成了一个矛盾。迁移学习为解决这一矛盾提供了可行的思路。但是目前已有的迁移学习算法多存在过于理想化的假设，比如静态环境，封闭环境和集中环境。而现实应用中需要面对的是动态环境，开放环境和分布式环境。在这样的环境下，已有的机器学习算法性能将受到严重影响，甚至无法工作。解决理想假设与现实问题之间的冲突为我们提供了一个强有力的出发点。为了应对这些冲突，本项目提出从在线领域自适应，自适应持续学习，生成式零样本学习和领域数据受限访问下的领域自适应四个方面开展研究。每一个研究内容都对应一个明确的问题和目标。同时，各个研究内容形成一个有机的整体，服务于同一个目标。

（2） **项目研究基础扎实**。项目负责人李晶晶现为电子科技大学计算机科学与工程学院研究员（校百人计划），澳大利亚昆士兰大学博士生指导教师。2017 年“博新计划”博士后。博士学位论文获得 2018 年 ACM 成都优秀博士论文奖和 2018 年中国电子学会优秀博士论文奖。入选 2019 年电子科技大学“学术新人奖”，2020 年电子科技大学“人才托举计划”项目，2020 年

电子科技大学“百人计划”。近五年来，项目负责人围绕领域自适应和零样本学习在 TPAMI，TIP，TKDE，TNNLS 和 CVPR，ACM MM 等 JCR 一区期刊及 CCF A 类会议上发表长文四十余篇，获得授权专利六项。担任 CCF 推荐期刊 MTAP 迁移学习主题的客座编委, TPAMI, TIP, TCYB, TNNLS, TKDE, CVPR, ICCV, AAAI, ACM MM 等期刊和会议审稿人。在迁移学习领域取得了一系列较为突出的研究成果，为本项目的顺利开展提供了扎实的研究基础。在此，选择其中比较有代表性的两组成果进行简单说明，以证明在研究基础方面的优势。

主要研究基础一：针对目标域训练数据少、已有有标记数据利用率低、数据分布不一致和多源异构数据难融合的问题，开展了迁移学习研究，提出了通用的迁移学习新方法，贡献了新的分布差异度量标准，阐明了领域不变特征的属性，探索了面向边缘智能和分布式智能的迁移学习。申请人在该方向的研究成果相继发表在 IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), IEEE Transactions on Cybernetics (TCYB)，IEEE Transactions on Image Processing (TIP)， IEEE Transactions on Neural Networks and Learning Systems (TNNLS)，IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)，International Joint Conference on Artificial Intelligence (IJCAI) 和 ACM International Conference on Multimedia (ACM MM) 等本领域内著名期刊和会议上。其中一篇论文还入选了 ESI 热点论文和高被引论文。成果发表后吸引了广泛的关注，多位国际知名学者引用和评价了这些工作。引用者中包括了多国国家科学院院士，IEEE Fellow，知名国际期刊主编和国家杰出青年基金获得者等等。在迁移学习领域，申请人的一个代表性成果是提出了独立共同迁移框架，该框架是一个通用的模型，其综合考虑了源域和目标域的数据对齐问题，以及源域和目标域的结构保留问题。特别地，申请人所提出的独立共同迁移框架突破了已有迁移学习算法要求源域和目标域必须使用相同维度特征表达的限制，具有很强的适用能力。相关研究成果发表在 2019 年出版的 IEEE Transactions on Cybernetics 上，该期刊是中科院 JCR 认定的自动化与控制系统和人工智能领域的一区 TOP 期刊。申请人的研究成果发表后，成功入选 ESI 热点论文榜，并从 2019 年 9 月开始迄今一直在高被引论文榜。申请人在一般领域自适应方面的研究基础为本项目开展在线领域

自适应，自适应持续学习以及无源领域自适应和领域泛化提供了基础。

主要研究基础二：针对数据多模态、数据缺失、样本污染和大规模快速响应等挑战，开展了多视角学习和哈希学习研究，提出了自适应哈希新方法，揭示了高维多模态数据的低秩结构，实现了大规模数据的近实时处理。申请人在该方向的研究成果相继发表在 IEEE Transactions on Knowledge and Data Engineering (TKDE)，IEEE Transactions on Multimedia (TMM)，IEEE Transactions on Image Processing (TIP)，IEEE Transactions on Cybernetics (TCYB)，International Joint Conference on Artificial Intelligence (IJCAI)，AAAI Conference on Artificial Intelligence (AAAI) 和 ACM International Conference on Multimedia (ACM MM) 等本领域内著名期刊和会议上。申请人的研究成果不仅进入 ACM MM 2019 最佳论文候选列表，同时这些成果还得到了许多学界学者的关注和认可。申请人在多模态学习方面的成果为本项目开展跨模态迁移学习，尤其是零样本学习打下了基础。

（3） **项目研究时机成熟。**近年来深度学习的快速发展为很多机器学习方向取得突破提供了强大的动力，为了尽快取得一些基础性研究成果，比如基于距离度量的领域自适应和基于对抗学习的领域自适应，研究者们可以在相对比较理想的假设下去尝试解决一些基础问题。但是随着技术的不断发展，机器学习算法服务于真实应用场景的需求日益凸显，这就催生了本项目的研究动机。当前，国内外的研究者们已经在迁移学习的基础理论上取得了一些突破，结合我们已有的研究基础，开展更加富有挑战且针对实际应用的研究正当时机。

（4） **项目研究保障充分。**首先是项目的人员保障比较充分。本项目的负责人长期从事领域自适应和零样本学习的研究，在此过程中，逐渐培养起了一支专业的、有活力的和有凝聚力的研究队伍。研究队伍中包括了教授，副教授，博士后和研究生。尽管受项目申请名额的限制，其中有的团队成员无法出现在本项目的研究成员列表中，但是在实际的研究过程中，本团队的所有成员将齐心协力全力以赴，共同为本项目的研究出谋划策。此外，申请人在长期的科研过程中，与国内外的相关同行建立了较为紧密的合作关系，这也为项目的开展提供了有益保障。最后，项目的科研条件有保障。项目的依托单位电子科技大学是国家 985、211 和双一流 A 类建设高校，

学校为项目的执行提供支持。申请人所在实验室拥有相对完备的实验硬件条件，为项目顺利开展解决了后顾之忧。

综上所述，本项目目标明确，思路清晰，基础扎实，时机得当，保障充分，完全有能力应对项目拟解决的关键科学问题，有望取得有国际影响力的重大原创性成果。

4. 本项目的特色与创新之处；

4.1 本项目的特色

与已有的研究成果相比，本项目的特色主要体现在如下方面：

（1） 紧密结合实际应用需求，攻克现实挑战，具有很高的应用价值

已有的研究多集中在理想化假设的实验室环境中，即静态环境，封闭环境和集中环境。本项目紧密结合实际应用需求，研究在动态环境，开放环境和分布式环境中的机器学习问题，尝试通过迁移学习缓解目标域标注数据稀缺的问题。本项目所研究的每一个场景都有坚定的现实应用基础，这些问题的突破将使得迁移学习算法可以无缝应用到现实场景中，服务于国民经济建设和国家社会重大需求。

（2） 研究内容层层递进，拟解决的关键科学问题具有较高创新价值

本项目的研究内容主要包含三部分，这三个部分分别对应于现实中的动态环境，开放环境和分布式环境。实际上，这三部分内容在技术手段上是层层递进的。我们以领域自适应为例，目前已有的领域自适应算法研究的是理想假设下的设定，我们的研究内容一探讨了在线场景下目标域样本分批或者逐个达到的情形；研究内容二进一步考虑了目标域中含有源域所没有见过的语义类别；研究内容三更进一步，考虑如果源域数据不可访问以及目标域数据不可访问的情形下又该如何进行领域适应。这些的问题都是迁移学习所面临的棘手问题，解决这些问题具有较高的创新价值。

（3） 积极探索前沿技术，独辟蹊径，将有可能产生引领性学术成果

以自适应持续学习为例，该主题实际上是一个前人未曾涉足的新问题，但是它又是现实中客观存在的问题。由于我们处在一个开放的动态环境，随着时间的发展应用的主要矛盾会产生一定变化，我们要求机器学习模型不仅能自适应地应

对这些变化，同时又不产生灾难性遗忘。对该问题的研究有助于唤起学术界同行的持续关注，产生引领性的学术成果。

4.2 本项目的创新之处

根据本项目的研究内容和技术方案，总结本项目的创新之处主要包括理论创新，技术创新和应用创新。分别说明如下：

（1） 理论创新

以零样本学习为例，目前大多数零样本学习算法都是学习视觉空间与语义空间之间的映射，比如学习视觉空间到语义空间的映射，然后在语义空间中对视觉样本进行分类；以及学习视觉空间与语义空间的共享空间，然后在共享空间中分类。这些模型可以统称为嵌入模型（Embedding Model）。虽然在引入辅助的语义空间后，嵌入模型可以进行零样本学习。但是需要注意的是，嵌入模型是间接的方法。他们并没有从根本上挑战零样本学习场景，因为嵌入模型并不会对零样本学习中的数据稀缺性带来改观。在理论上，本项目拟直接挑战零样本学习的根本性问题——数据稀缺性。通过生成模型直接合成样本，从而将零样本学习转化为一个普通的有监督学习问题。此外，本项目将从理论上讨论生成模型应用到零样本学习中可能会引发的问题，即特征混淆和模型崩塌。我们将对特征混淆进行理论分析，形式化定义，研究量化指标以及优化方案。

以领域自适应为例，通过学习新的特征表达并在新的特征空间最小化源域和目标域之间的分布差异是目前主流的方法。但是在基于距离度量的方案中，研究者们都在尝试从不同的角度和不同的粒度去应用已有的几个距离度量，比如最常用的 MMD，但是在距离度量本身的研究上长期处于停滞状态，没有什么新的进展。需要注意的是，MMD 等度量并不适合于源域和目标域的样本分布非常不均衡的情况，比如目标域样本逐个到达的情况。因此，本项目在不均衡情况下的距离度量理论上也将开展研究，有望取得一定突破。

（2） 应用创新

常见的机器学习方法都是基于有监督模型的。有监督模型必然面临着训练样本稀缺的问题以及应用场景特定的限制。要打破样本稀缺问题和应用场景的限定，就需要研究少样本学习甚至零样本学习。因此，零样本学习自身就是一个创新应用。以本项目的零样本学习应用场景为例，本项目特别关注如下三类创新应

用：1) 零样本跨模态识别与检索。传统的搜索引擎是基于关键字与检索目标的标签相匹配进行检索的，本项目的研究成果可以直接用于跨模态检索，且无需目标样本中有检索关键字对应的标签。2) 零样本动态目标检测。在诸如自动驾驶等场景中，目标检测是一项核心技术。但是已有的目标检测方法只能检测出经过训练的类别，无法检测新类别。本项目的研究成果可以应对这一挑战。3) 零样本医疗辅助诊断。由于医疗图像处理中训练样本的获取和标注代价高昂，已有医疗辅助诊断系统只能检测常见病。本项目所研究的技术可以用于少样本病种的辅助诊断。

（3） 技术创新

首先，自适应持续学习是一个技术上的创新。其次，对生成模型中的特征混淆进行研究，提出量化指标和优化方案是之前尚未开展的研究，有望产生重要的学术成果和广泛的学术影响。第三，对生成对抗网络中的模型崩塌问题进行研究，可以带来全新的应对模型崩塌的技术，进一步提升生成对抗网络的稳定性和可用性。最后，我们尝试使用对抗样本去提升模型的泛化能力，该方案不仅仅是领域数据受限访问下迁移学习的贡献，也是对对抗攻击领域的贡献。

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

本项目预计四年完成，即从 2022 年 1 月至 2025 年 12 月。研究工作大体分为三个阶段：第一阶段为文献调研，第二阶段为理论研究与创新，第三阶段为实验验证与验收。详细安排如下：

2022-01 至 2022-08 ： 文献调研并完成领域自适应和零样本学习理论分析。

2022-09 至 2023-02 ： 积极开展学术交流及合作，提出适合在线领域自适应的新颖算法，并证明其有效性。

2023-03 至 2023-12 ： 完成自适应持续学习模型的设计，并证明其在克服灾难性遗忘上的比较优势。

2024-01 至 2024-07 ： 完成生成式零样本学习模型的设计和优化，完成特征混淆和模式坍塌应对方案的设计，并证明其有效性、灵活性、可扩展性，同时开展模型的扩展应用研究。

2024-08 至 2025-02 ： 完成无源领域自适应算法和领域泛化算法的设计与实现，在理论上论证对抗样本对模型泛化能力的影响。

2025-03 至 2025-09 ： 建立分布式计算平台及应用测试平台，采集并构建应用测试数据库，实现用于演示的原型软件系统，并完成应用测试验证。

2025-10 至 2025-12： 项目验收准备、鉴定及相关成果发布。

5.2 预期研究成果

本项目预期在以下方面产生研究成果：

（1）**学术论文**。本项目作为基础研究项目，学术论文是重要的成果组成部分。我们计划在 JCR 二区及以上期刊及 CCF-A 类会议和期刊上共发表原创性研究成果不少于 10 篇。具体计划是：2022 年发表至少 2 篇中科院 JCR 二区及以上期刊论文；2023 年在高水平国际会议上发表至少 3 篇论文；2024-2025 年在高水平国际会议及期刊上发表至少 5 篇一流论文和不少于 2 篇额外的中文期刊论文。本课题研究成果将在总体上达到国际先进水平，部分研究方向上达到国际领先水平。上面提到的高水平国际会议及期刊主要包括：ACM Conference on Multimedia (MM), AAAI Conference on Artificial Intelligence (AAAI), International Joint Conference on AI (IJCAI), IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), IEEE International Conference on Computer Vision (ICCV), IEEE Transactions on Pattern Analysis & Machine Intelligence (TPAMI), IEEE Transactions on Image Processing (TIP), IEEE Transactions on Multimedia (TMM), IEEE Transactions on Circuits and Systems for Video Technology (TCSVT), IEEE Transaction on Knowledge and Data Engineering (TKDE), IEEE Transactions on Neural Networks and Learning Systems (TNNLS), Pattern Recognition (PR)等。

（2）**应用验证**。本项目的主要研究动机指向了现实应用中的诸多挑战，因此，本项目预期研究成果的另一重要组成部分为应用验证。我们将具有动态性、开放性和小样本性的推荐系统以及医工结合为具体应用，综合应用在线领域适应，持续领域适应和零样本学习方法应对这些问题。同时，并利用该课题的研究成果，申请面向应用的纵向和横向课题，服务国家重大需求。

（3）**人才培养**。在人才培养方面，坚持立德树人，依托本项目预期培养博士研究生 2 名，硕士研究生 6 名。

（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）；

申请人一直致力于欠标注场景的机器学习算法的研究和实践中。从博士研究生阶段开始，就在紧密围绕迁移学习和小样本学习开展原创性研究和创新性应用。迄今共发表 ACM/IEEE 汇刊长文 20 多篇，高水平会议论文（CCF 推荐 A 类会议论文长文）20 多篇。其中，作为第一作者或通讯作者发表高水平论文（CCF 推荐 A 类会议以及 ACM/IEEE 汇刊长文）21 篇。申请人针对欠标注场景的机器学习相关研究取得的成果受到学术界的关注和肯定。比如，申请人作为第一作者发表的题为“独立共同迁移”的论文自 2019 年 6 月正式发表后，从 9 月开始进入 ESI 热点论文榜单，并至今一直在榜 ESI 高被引论文。申请人 2019 年发表于 CCF A 类会议 ACM Multimedia 的论文进入最佳论文候选。此外，申请人的研究成果被多位国际著名学者引用，其中不乏欧洲，加拿大，澳大利亚，芬兰和波兰的国家院士。此外，还有 30 余位 ACM/IEEE/IAPR/IET Fellow，国际期刊主编，以及国家杰出青年基金获得者在其论文中引述了申请人的研究成果。申请人的研究成果在 2018 年获得了国家一级学会中国电子学会的优秀博士论文奖和 ACM 成都地区的优秀博士论文奖。申请人凭借所取得研究成果，在 2017 年成功入选人社部博士后创新人才支持计划。

申请人长期坚持对学术社区的贡献。申请人曾担任 CCF 推荐期刊 Multimedia Tools and Applications 的在迁移学习方向的责任客座编辑。受邀担任 IEEE TPAMI，TIP，TNNLS，TKDE，TMM，TCYB 和 PR 等期刊的评审人，担任 ACM MM，CVPR，ICCV，ICDE 和 AAAI 等会议的程序委员会委员。申请人研究成果绝大部分在发表时公开源代码，以方便社区的研究者。同时，申请人注重与海内外学者的交流，鉴于申请人在欠标注场景取得的研究成果，世界排名前 50 的昆士兰大学邀请申请人作为电子与信息工程系校外博士生指导老师。同时，申请人受邀在 ACM 成都地区年会和兄弟院校开展学术讲座。

以下是申请人在本项目相关的主题发表的一些研究成果：

[1] **Li, Jingjing**, Erpeng Chen, Zhengming Ding, Lei Zhu, Ke Lu, and Heng Tao Shen. "Maximum Density Divergence for Domain Adaptation." *IEEE Transactions on Pattern Analysis and Machine Intelligence (**TPAMI**)* (2020).(中科院 JCR 一区)

[2] **Li, Jingjing**, Ke Lu, Zi Huang, Lei Zhu, and Heng Tao Shen. "Transfer independently together: A generalized framework for domain adaptation." IEEE Transactions on Cybernetics (**TCYB**) 49, no. 6 (2018): 2144-2155. (**ESI 热点，高**

被引，中科院 JCR 一区)

[3] **Li**, **Jingjing**, Mengmeng Jing, Ke Lu, Lei Zhu, Heng Tao Shen, Faster Domain Adaptation Networks, *IEEE Transactions on Knowledge and Data Engineering (**TKDE**)*, 2021 (CCF A)

[4] **Li, Jingjing**, Ke Lu, Zi Huang, Lei Zhu, and Heng Tao Shen. "Heterogeneous domain adaptation through progressive alignment." *IEEE transactions on neural networks and learning systems (**TNNLS**)* 30, no. 5 (2018): 1381-1391. (中科院 JCR 一区)

[5] **Li, Jingjing**, Mengmeng Jing, Ke Lu, Lei Zhu, and Heng Tao Shen. "Locality preserving joint transfer for domain adaptation." *IEEE Transactions on Image Processing (**TIP**)*, 28, no. 12 (2019): 6103-6115. (中科院 JCR 一区)

[6] **Li, Jingjing**, Mengmeng Jing, Ke Lu, Zhengming Ding, Lei Zhu, and Zi Huang. "Leveraging the invariant side of generative zero-shot learning." *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (**CVPR**)*, pp. 7402-7411. 2019. (CCF A)

[7] **Li, Jingjing**, Ke Lu, Zi Huang, and Heng Tao Shen. "On both Cold-Start and Long-Tail Recommendation with Social Data." *IEEE Transactions on Knowledge and Data Engineering (**TKDE**)* (2019). (CCF A)

[8] **Li, Jingjing**, Yue Wu, and Ke Lu. "Structured domain adaptation." *IEEE Transactions on Circuits and Systems for Video Technology (**TCSVT**)* 27, no. 8 (2016): 1700-1713.

[9] **Li, Jingjing**, Jidong Zhao, and Ke Lu. "Joint Feature Selection and Structure Preservation for Domain Adaptation." *In International Joint Conference on Artificial Intelligence (**IJCAI**)*, pp. 1697-1703. 2016. (CCF A)

[10] **Li, Jingjing**, Ke Lu, Zi Huang, and Heng Tao Shen. "Two birds one stone: on both cold-start and long-tail recommendation." *In Proceedings of the 25th ACM international conference on Multimedia (**ACM MM**)*, pp. 898-906. 2017. (CCF A)

[11] **Li, Jingjing,** Mengmeng Jing, Ke Lu, Lei Zhu, Yang Yang, and Zi Huang. "From zero-shot learning to cold-start recommendation." *In Proceedings of the AAAI Conference on Artificial Intelligence (**AAAI**)*, vol. 33, pp. 4189-4196. 2019. (CCF A)

[12] **Li, Jingjing**, Lei Zhu, Zi Huang, Ke Lu, and Jidong Zhao. "I read, I saw, I tell: Texts assisted fine-grained visual classification." *In Proceedings of the 26th ACM international conference on Multimedia (ACM MM)*, pp. 663-671. 2018. (CCF A)

[13] **Li, Jingjing**, Mengmeng Jing, Ke Lu, Lei Zhu, Yang Yang, and Zi Huang. "Alleviating Feature Confusion for Generative Zero-shot Learning." *In Proceedings of the 27th ACM International Conference on Multimedia (**ACM MM**)*, pp. 1587-1595. 2019. (CCF A)

[14] **Li, Jingjing**, Erpeng Chen, Zhengming Ding, Lei Zhu, Ke Lu, and Zi Huang. "Cycle-consistent Conditional Adversarial Transfer Networks." *In Proceedings of the 27th ACM International Conference on Multimedia (**ACM MM**)*, pp. 747-755. 2019. (CCF A)

[15] **Li, Jingjing**, Mengmeng Jing, Lei Zhu, Zhengming Ding, Ke Lu and Yang Yang, Learning Modality-Invariant Latent Representations for Generalized Zero-shot Learning, *In Proceedings of the 28th ACM International Conference on Multimedia (**ACM MM**)* 2020, CCF A.

[16] **Li, Jingjing**, Mengmeng Jing, Ke Lu, Lei Zhu, Heng Tao Shen, Investigating the Bilateral Connections in Generative Zero-shot Learning, *IEEE Transactions on Cybernetics (**TCYB**)*, 2021. (中科院 JCR 一区)

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）；

申请人所在单位电子科技大学 1960 年被中共中央列为全国重点高等学校，1961 年被中共中央确定为七所国防工业院校之一，1988 年更名为电子科技大学，1997 年被确定为国家首批“211 工程”建设的重点大学，2000 年由原信息产业部主管划转为教育部主管，2001 年进入国家“985 工程”重点建设大学行列，2017 年进入国家建设“世界一流大学”A 类高校行列。2019 年教育部和四川省签约共同推进我校世界一流大学建设。经过 60 余年的建设，学校形成了从本科到硕士研究生、博士研究生等多层次、多类型的人才培养格局，成为一所完整覆盖整个电子信息类学科，以电子信息科学技术为核心，以工为主，理工渗透，理、工、管、文、医协调发展的多科性研究型大学，成长为国内电子信息领域高新技术的源头，创新人才的基地。

申请人所在的计算机科学与工程学院现有计算机科学与技术一级学科和网络空间安全一级学科博士学位授予权，其中计算机科学与技术一级学科具有博士后流动站。计算机科学与技术一级学科在 2017 年全国第四轮学科评估中评为 A；2020 US News 全球计算机学科排名 26 位；计算机学科位于 ESI 前 0.64‰，位列全球第 32 位（至 2020 年 7 月）。现有教职工 200 余人，专任教师 140 余人，拥有一支包括 1 位中科院院士（双聘），1 位国家科技进步一等奖获得者，1 位新世纪百千万人才工程入选者，4 位教育部新世纪人才，2 位四川省教学名师，20 余位国家高层次人才，已形成以中青年学术专家和学术骨干为主，专业结构和年龄结构合理、富有活力的教师队伍。

申请人所在的团队为“四川省计算机网络技术及应用重点实验室”，该实验室始建于 1995 年，是四川省教委重点实验室，也是西南地区最早成立的计算机网络应用技术实验室之一。实验室现有专职教师 18 人，计算机应用技术博士生 20 余人，计算机应用技术硕士研究生 100 余人。所有专职人员均具备硕士及以

上学历，其中教授 3 人、副教授 4 人、高级工程师 1 人、讲师 7 人，助教 3 人。实验室总面积约 300 平方米，主要从事下一代互联网、新型网络信息系统、人工智能等领域的应用基础研究和系统开发工作。同时，实验室还负责电子科技大学计算机应用技术专业本科、硕士和博士的培养及教学工作。

团队目前已经积累了大量的硬件设备也可用于本项目的研究，团队所拥有的设备，包括 GPU 服务器和个人电脑等，总价值近千万元。团队成员具有很强的相关理论积累、承担国家级项目的实践经验与良好的团队合作精神，能够有力保障本项目理论研究的顺利进行。此外，团队还与美国，新加坡和澳大利亚等国家的同行建立了密切的合作关系，为本项目的顺利开展提供了强有力的专家资源。

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）：

申请人目前作为负责人承担一项自然科学基金青年项目，项目批准号：61806039，项目名称：基于异构域适配的欠标注场景迁移学习研究，项目执行年限：2019.01-2021.12，项目直接费用：27 万元。

以上项目的研究内容为本项目提供了基础。在以上项目中，我们基于异构域适配，即基于源域和目标域采样自不同的特征空间或不同的数据模态的域适配问题，研究深度异构迁移学习模型，将已取得的迁移学习成果与深度网络融合。但是这些研究仍然是处于传统的迁移学习研究范畴，存在理想化的假设。在本项目中，我们尝试突破这些假设，研究适用于动态、开放和分布式环境的迁移学习算法，并将这些算法应用到实际问题中。同时，上述项目将于 2021 年 12 月结题，本项目的开始时间为 2022 年 1 月，两个项目资助时间不存在重叠，可以顺利衔接，在此前的基础上开展更深入研究。

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）。

尚无已完成项目。

（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）。

无。

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。

无。

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。

无。

4. 其他。

无。