

基于卷积条件神经过程的地统计插值实验

李锦韬

2201213292 lijintao@stu.pku.edu.cn

1. 选题背景

地统计学 (Geostatistics) 是统计学在地学领域的应用, 主要处理点参考数据 (Point Referenced Data, PRD), 其不但关注插值的结果, 还重视所插入值的不确定性 [1]。传统地统计采用高斯过程回归 (即各类克里金插值方法) [2, 3], 具有独特的预测结果不确定度衡量优势, 但问题在于计算是高度密集的, 每个新插入点的复杂度达到立方级别 [4]。

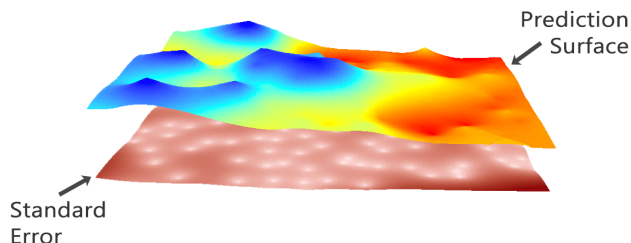


图 1. 地统计: 兼顾插值预测与不确定估计 [5]

近年来, 深度学习快速发展, 神经网络模型展现出强大的函数逼近能力, 能够一定程度上逐步学习数据的分布 [6]。使用神经网络达到高斯过程效果的融合研究逐渐兴起, 本项目因此选择了神经过程 (Neural Process) 作为研究对象。

神经过程基于元学习的思想, 将神经网络应用于对高斯过程的模拟, 能够实现类似的效果, 目前已经成为一个具有诸多成员的模型家族 [7]。本文聚焦于卷积条件神经过程, 对地统计插值进行了实验, 分析其预测效果, 并给出未来的改进方向。

2. 相关研究

高斯过程处理的是函数的分布, 而神经网络模型处理的是一个函数的逼近。为了使神经网络能像高斯

过程那样预测均值和方差, 需要引入元学习 (Meta-Learning) 的概念。神经过程在图2右下象限, 可将其视为使用神经网络来元学习从数据到预测随机过程的映射, 这也同时是神经过程名称的由来。

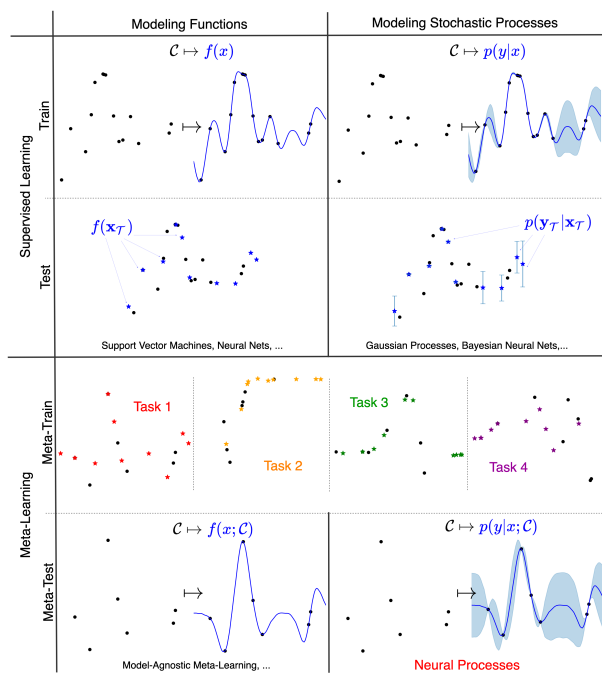


图 2. 概念对比: 元学习与监督学习 (上下两栏); 建模单个函数与建模随机过程 (左右两栏)。图上黑色圆点是训练点, 而星星点是测试点 [8]。

神经过程来源于 DeepMind 团队在 2018 年的开创性研究, 首先提出的是条件神经过程模型 (Conditional Neural Process, CNP), 意为其在给定一系列观察数据时能够定义函数的条件分布。该模型实现了深度学习灵活性和概率模型不确定性的结合, 能够进行端到端 (end-to-end) 的训练 [9]。

在 CNP 模型之后, DeepMind 为了改进 CNP

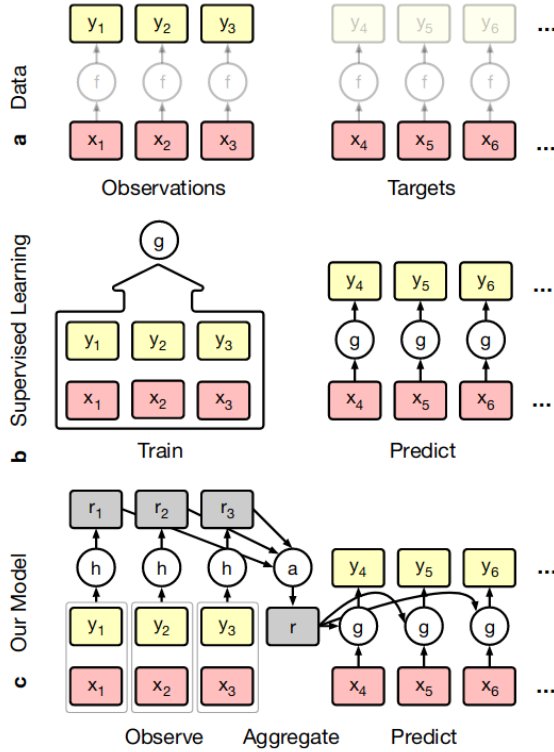


图 3. CNP 条件神经过程 [9]: (a) 子图上 f 代表从输入数据到标签的映射 (可能是固定的、或是某个随机函数的一次实现)。(b) 子图上的 g 代表监督学习的目标近似函数, 其在每个新任务下是重新随机初始化参数化的, 通过最小化近似函数 g 与 f 之间的损失, 能在大量数据支撑下完成 g 对 f 的拟合。(c) 子图上的 h 是多层全连接神经网络 (MLP), 完成对背景点观测编码, 使用对称聚合器 a 聚合成 r , 嵌入作为解码器 g 的条件, 在同时给定目标测试点情况下给出预测值和方差。

的采样生成不连续问题, 在 CNP 基础上增加了一个类似于 VAE 瓶颈的隐变量 z , 其每一个随机样本都对应于随机过程的一个具体实现, 这样就可以通过多个样本在解码器网络中的前向传递, 生成目标处的预测分布。该模型在文献中被命名为神经过程 (Neural Process, NP) [10, 11]。但为了将其与 CNP 更好的区分, 本项目称其为隐变量神经过程 (Latent Neural Process, LNP)。

CNP 和 LNP 模型共同构成了神经过程家族 (Neural Process Family, NPF) 的基类 [8, 7]。后续有许多研究者将神经过程模型与其他神经网络模型结合, 逐渐使成员扩充。其中, Kim 等提出了注意力

神经过程 (Attentive Neural Process, AttnCNP), 其修订了聚合器, 引入基函数增加参数来刻画不同背景点所贡献的权重, 能实现注意力效果 [12]; Gordon 等提出了卷积神经过程, 将编码器部分由 MLP 改为卷积神经网络, 根据卷积核感受野理论其能够增加一定自相关建模能力 [13]。

3. 项目内容

本项目选择使用卷积神经过程模型来解决地统计空间预测问题, 即给定一系列地理要素的观测值, 预测目标地理要素的值和不确定度。由于地学数据来源极其广泛, 数据类型和模态丰富多样, 目前还没有统一的标准数据集, 因此本项目将以 DEM (数字高程模型) 栅格数据和其他一些能够体现神经过程特性的数据集为主要数据。

数据集。 本项目选择下方罗列的 3 个数据集进行实验。数据集构建为标准的元学习数据集, 从图像上的采样点数量分布为 $[0, n_{pixels} \times 0.3]$, 即每次使用不超过每张图像上的 30% 像素点作为背景点训练。数据集的图示情况见图5。

- TfaSR 数据集 [14]: DEM 数字高程数据, 格式为 GeoTIFF, 每个像素上存储地形高程值, 使用时进行归一化处理, 每张图像大小为 $(1 \times 64 \times 64)$, 训练集大小为 25088, 测试集大小为 6272。相比于普通的直接按照经纬度划分的地形数据, 该数据集中的河流和山脉部分占比较大, 相当考验模型对空间自相关性的学习和理解。
- FashionMNIST 数据集 [15]: 与 MNIST 数据集在格式上完全相同, 均为 $(1 \times 32 \times 32)$ 的灰度图像, 但图像内容是时装鞋帽等产品的图像, 比手写数字更加复杂, 更考验模型的泛化能力, 训练集大小为 60000, 测试集大小为 10000。
- EuroSAT 数据集 [16]: 卫星影像数据, 格式为 JPG, 每张图像大小为 $(3 \times 64 \times 64)$, 由于显卡资源有限, 选取其中的工业区类别, 训练集大小为 2500, 测试集大小为 2000。相比于上述两个数据集, 工业区的分布和空间结构更加复杂, 且训练数据量很少。

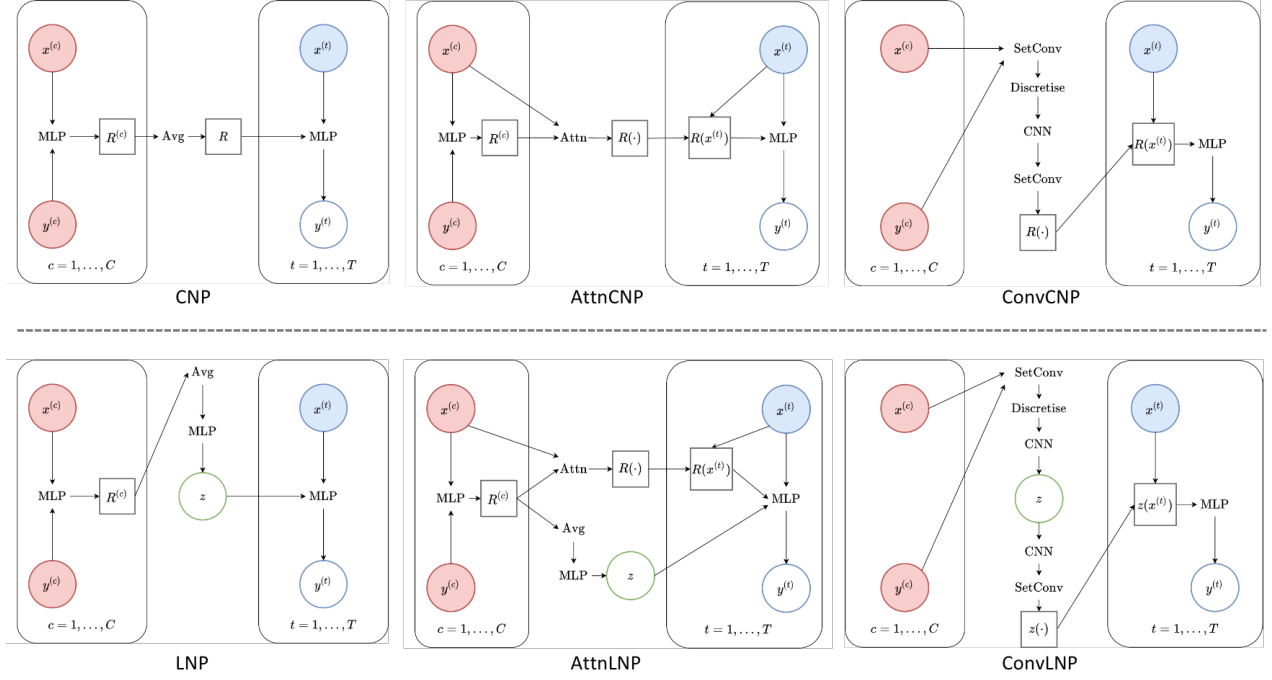


图 4. 神经过程家族 (NPF) [8]

模型设计. 考虑到目前没有生成数据的需求, 且训练计算资源相当有限, 于是选择了 CNP 家族成员中的卷积条件神经过程 ConvCNP, 其对二维图像的空间自相关性处理效果相对较好, 模型结构如图4右上部分所示。

本项目使用 PyTorch 框架构建模型, 方便进行 GPU 加速, 模型中编码器部分使用了 ResNet 结构 [17], 构建了 5 个 Block 总计 10 层卷积, 卷积核大小为 9, 聚合器使用和 AttnCNP 相同的架构, 解码器部分使用 4 层 128 神经元的 MLP。模型的输入为从图像上随机采样的背景点和测试点位置, 输出为测试点的预测值和不确定度。

模型训练. ConvCNP 作为条件神经过程的一种, 训练过程可以划分为 5 步:

- 从训练数据集 \mathcal{D} 中采样一个子任务数据集 $\{\mathcal{D}_i\}_{i=1}^{N_{tasks}}$
- 将子任务数据集随机划分为背景点与测试点 $\mathcal{D} = \mathcal{C} \cup \mathcal{T}$
- 以背景点为条件输入卷积条件神经过程获取测试点的预测 $p_{\theta}(y_{\mathcal{T}}|x_{\mathcal{T}}; \mathcal{C})$

- 计算损失函数对数似然 $\mathcal{L} = \log p_{\theta}(y_{\mathcal{T}}|x_{\mathcal{T}}; \mathcal{C})$, 其衡量测试点预测的表现情况。实际计算中使用的是负对数似然和。
- 计算损失函数的梯度 $\nabla_{\theta} \mathcal{L}$, 反向传播优化调节神经网络各参数

本项目使用相同的模型结构对三个数据集分别展开训练, 训练参数如下:

- 学习率: 0.001, 学习率衰减率: 0.1, 优化器: SGD
- 损失函数: 负对数似然和
- 最大训练轮数: 50, 训练批次: 32
- 10% 训练集作为验证集, 每轮训练结束后在验证集上计算损失函数, 选择损失函数最小的模型参数作为最终模型参数

实验环境. 本项目的软硬件环境如下:

- 软件环境: 操作系统为 Windows 11, Python 版本为 3.8, CUDA 版本为 12.1, pytorch 版本为 2.0.1, 使用 Anaconda 进行环境管理
- 硬件环境: CPU 为 Intel i5-10400, GPU 为 NVIDIA GeForce GTX 1650, 显存为 4GB

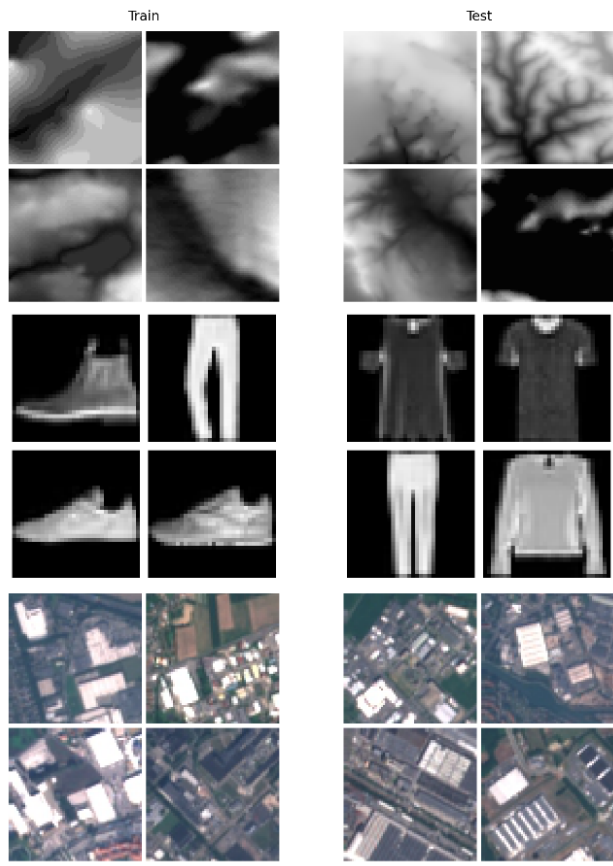


图 5. 数据集示例

4. 实验结果

训练过程. 三个数据集的训练过程如下所示，横轴为训练轮数 (Epoch)，纵轴为损失函数的负对数似然和。

TfaSR 数据集训练过程见图6，验证集损失函数在训练过程中出现了大量上升的情况，说明模型在验证集上的表现并不稳定，可能存在过拟合的情况，模型对于 DEM 地形数据的泛化能力不理想。最终最佳模型为第 40 轮训练的模型。

FashionMNIST 数据集训练过程见图7，训练集和验证集的损失函数都在不断下降，模型相比在 DEM 上的实验学习情况更好。最终最佳模型为第 26 轮训练的模型。

EuroSAT 数据集训练过程见图8，训练集和验证集的损失函数都在不断下降，且两者的表现基本一致，模型在 EuroSAT 数据集上的表现较好，这也

可能是由于该实验中数据量较少。最终最佳模型为第 47 轮训练的模型。

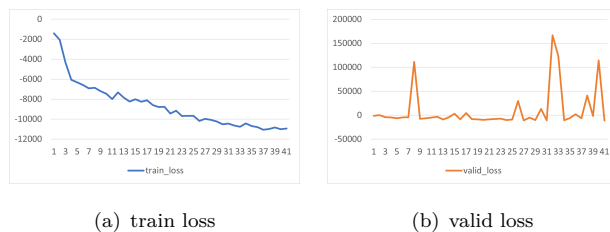


图 6. TfaSR 训练过程 (单轮需要 24min 完成)

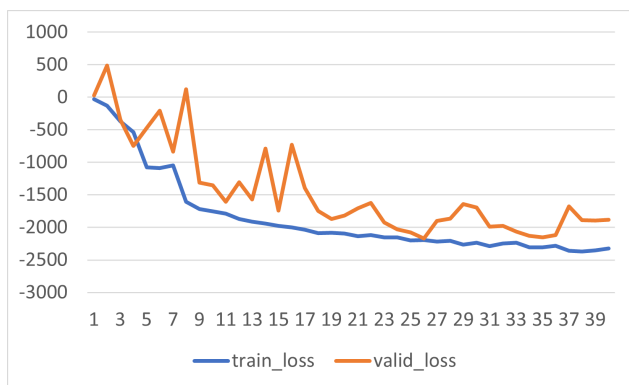


图 7. FashionMNIST 训练过程 (单轮需要 20min 完成)

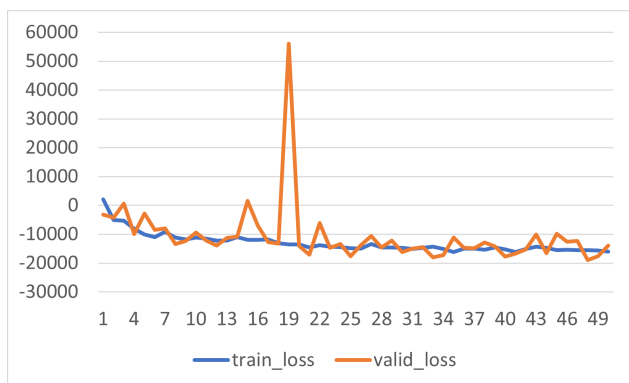


图 8. EuroSAT 训练过程 (单轮需要 4min 完成)

预测结果. 本项目使用模型对测试集进行预测，从测试图像上选取 1/8 数量的背景点，预测整张图像的插值重建情况，并将预测结果与最邻近插值、双线性插值、三次插值对比。具体情况可见图示，第 1 行为输入的背景点，第 2 行为预测均值，第 3 行为预测不确定度，第 4 行为最邻近插值，第 5 行为双线性插值，第 6 行为三次插值，第 7 行为真实图像。

TfaSR 数据集预测结果见图9，模型预测结果与基线相比效果一般，且相对更加平滑，很难恢复出 DEM 的精细地形特征，说明模型对于 DEM 数据的泛化能力不理想。另外观察不确定性图可以发现，模型对于地形特征的不确定性较小，这与实际情况不太符合，地形变化剧烈的位置不确定度预测没有较大变化，模型对数据的学习情况不理想。

FashionMNIST 数据集预测结果见图10，模型预测效果相当好，远超基线方法。进一步分析不确定性图可以看出模型训练得到的卷积感受野，说明模型从数据本身中学习到相对核大小，而且衣物的边缘不确定性相对较大，具有很强的解释性，能够指导新的采样位置。另外模型的预测也有较强的平滑效果，对纹理细节的恢复效果不是很好。

EuroSAT 数据集预测结果见图11，模型预测结果比基线效果好，但是相比真实图像差距还较大，呈现出方块形的图斑特征，这可能是由于训练数据中工业区的建筑物分布通常以几何规则矩形居多。而不确定性图中，模型对建筑物区域的不确定较大，这与实际情况相符合，说明模型仍然在少量数据中学习得到部分空间分布特征。

5. 讨论分析

本项目对卷积条件神经网络进行了实验测试，可见模型对于 DEM 地形预测效果不理想，对于 FashionMNIST 预测效果较好，对于 EuroSAT 预测效果一般。以下探讨模型的不足之处，并提出改进方向。

模型不足。 模型对数据的编码效果受限于编码器结构和训练数据量。对于 DEM 数据的实验表明，地形特征的变化范围较大，精细结构较多，当前模型难以学习到地形特征的分布规律。而对于 FashionMNIST 数据集的测试结果表明，模型对于背景和主体差别明显的样本学习效果较好。而对于 EuroSAT 数据集的测试结果表明，模型对于建筑物的学习效果一般，这可能是由于 EuroSAT 数据集的训练数据量较小，且建筑物的分布规律较为复杂，模型难以学习到建筑物的分布规律。

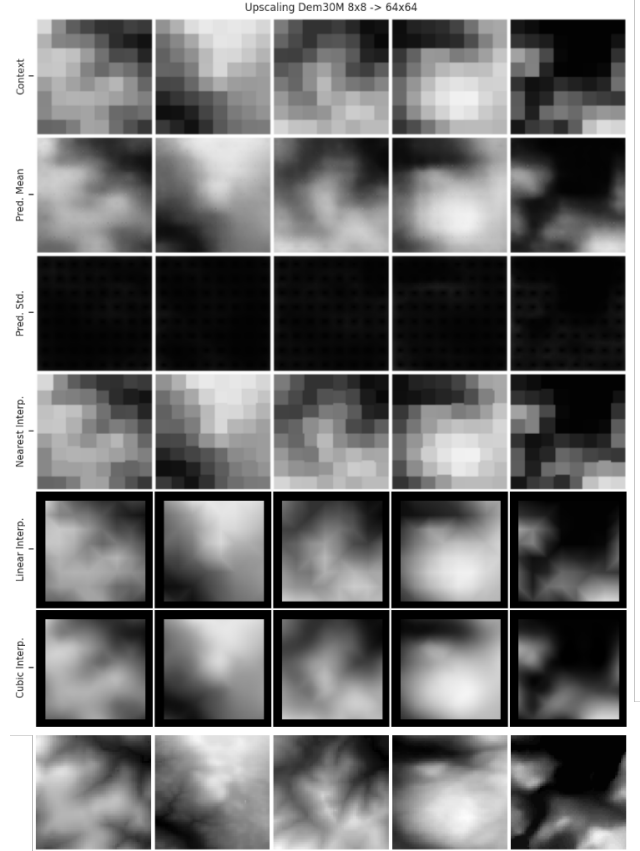


图 9. TfaSR 数据集预测情况

模型改进。 本项目使用的模型结构较为简单，仅使用卷积编码器和 MPL 解码器建模空间预测均值和不确定度，模型的学习能力受限。下一步考虑使用带有自注意力机制的卷积，如 ViT 作为编码器，以提升模型的学习能力。另外，本项目使用的训练数据也相对有限，各种超参数的优化还有上升空间，未来应当使用更大算力开展更多实验。

6. 未来展望

随着 2012 年以来的人工智能新浪潮崛起，深度学习技术的高速发展，以神经网络模型 (Neural Network, NN) 为代表的相关方法在计算机视觉、自然语言处理等领域取得了巨大的成功 [6]，其可以通过大量小批次的数据迭代训练，使用梯度下降进行参数调优，以高精度逼近有标签数据集，自适应特征提取效果好、可拓展性强 [18]。但是神经网络通常以黑箱形式存在，缺乏对不确定性的量化能力和可

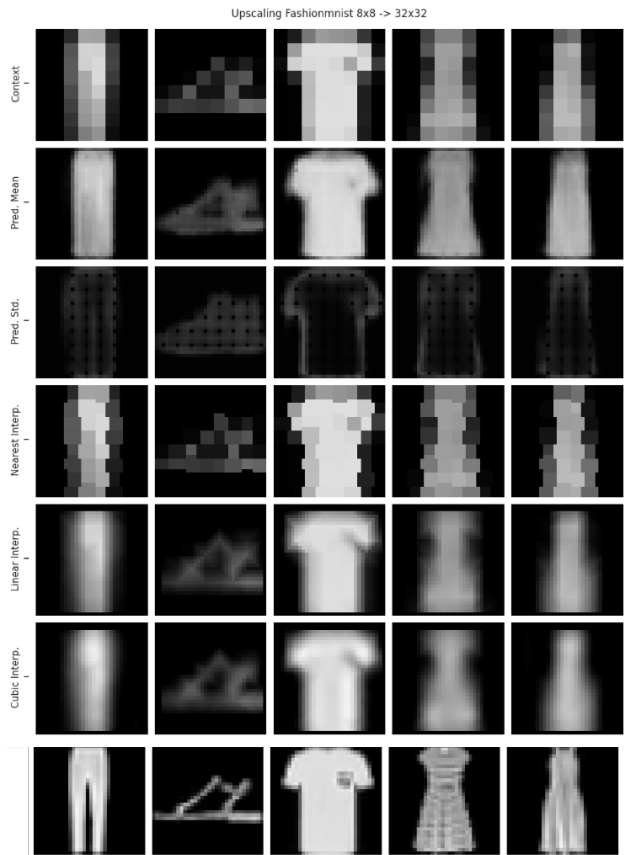


图 10. FashionMNIST 数据集预测情况

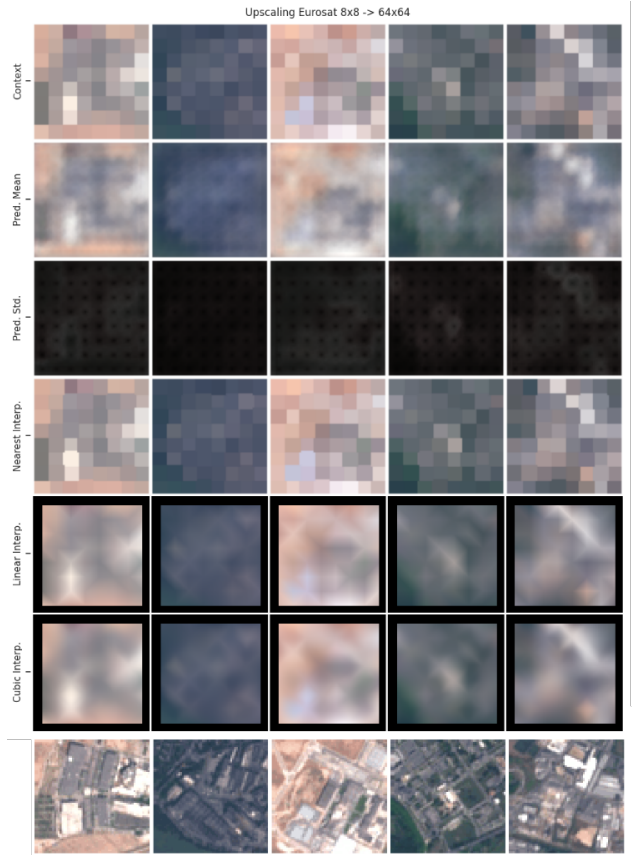


图 11. EuroSAT 数据集预测情况

解释性，并且作为一种参数量很大的模型，在普遍具有异质性的地学数据上迁移和泛化效果不佳，这两方面使得神经网络模型在地统计领域的应用受到限制 [19]。

高斯过程和神经网络各自具有优势，倘若能将两者结合起来，可为地统计领域带来新方法的机遇和新思路的拓展。

对空间预测插值和不确定度衡量的需求使得地统计从一开始就与高斯过程模型展开漫长的合作，诞生出一系列优秀的克里格方法。然而，在当前数据量与日俱增和深度学习日新月异的情形下，研究者们呼唤着新方法的出现，高斯过程和神经网络的结合在笔者看来是一个非常有力潜力的赛道，而神经网络是该赛道上的有力竞争者。

在未来的工作中，笔者计划进一步调研和梳理高斯过程与神经网络融合的方法，尤其以神经过程为主，在不同的任务场景下测试神经过程，与条件

变分自编码器（CVAE）、深度高斯过程（DGP）等方法进行比较，并希望能够多应用深度学习领域内优秀算法的先进思想到地统计学中，使地统计高斯过程的适用性和拓展性进一步增强。

致谢

本项目得到王选计算机研究所穆亚东教授和工学院濮国梁副教授的指导，感谢两位老师的帮助，使得本项目选择高斯过程和神经网络的交叉调研成为可能。本项目中神经过程章节所涉及的部分文献源自 GeoSOT 实验室研究生魏新光的推荐，在此感谢其给予的支持。最后特别感谢两位课程助教谭镇涛与沈天一，每次沟通都给予及时的解答和帮助，在笔者因为身体生病导致提交延后也给予了耐心的关心和照顾，再次感谢两位助教的善意。

参考文献

- [1] G. Matheron, “Principles of geostatistics,” *Economic geology*, vol. 58, no. 8, pp. 1246–1266, 1963. 1
- [2] E. Schulz, M. Speekenbrink, and A. Krause, “A tutorial on gaussian process regression: Modelling, exploring, and exploiting functions,” *Journal of Mathematical Psychology*, vol. 85, pp. 1–16, 2018. 1
- [3] M. A. Oliver, R. Webster et al., “Basic steps in geostatistics: the variogram and kriging,” Springer, Tech. Rep., 2015. 1
- [4] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*. MIT press Cambridge, MA, 2006, vol. 2, no. 3. 1
- [5] GISGeography. Kriging interpolation - the prediction is strong in this one. [Online]. Available: <https://gisgeography.com/kriging-interpolation-prediction/> 1
- [6] G. Hinton, Y. LeCun, and Y. Bengio, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015. 1, 5
- [7] S. Jha, D. Gong, X. Wang, R. E. Turner, and L. Yao, “The neural process family: Survey, applications and perspectives,” 2022. 1, 2
- [8] Y. Dubois, J. Gordon, and A. Y. Foong, “Neural process family,” <http://yanndubs.github.io/Neural-Process-Family/>, September 2020. 1, 2, 3
- [9] M. Garnelo, D. Rosenbaum, C. Maddison, T. Ramalho, D. Saxton, M. Shanahan, Y. W. Teh, D. Rezende, and S. A. Eslami, “Conditional neural processes,” in *International conference on machine learning*. PMLR, 2018, pp. 1704–1713. 1, 2
- [10] M. Garnelo, J. Schwarz, D. Rosenbaum, F. Viola, D. J. Rezende, S. Eslami, and Y. W. Teh, “Neural processes,” *arXiv preprint arXiv:1807.01622*, 2018. 2
- [11] K. Märtens. Neural processes as distributions over functions. [Online]. Available: <https://kasparmartens.rbind.io/post/np/> 2
- [12] H. Kim, A. Mnih, J. Schwarz, M. Garnelo, A. Eslami, D. Rosenbaum, O. Vinyals, and Y. W. Teh, “Attentive neural processes,” *arXiv preprint arXiv:1901.05761*, 2019. 2
- [13] J. Gordon, W. P. Bruinsma, A. Y. Foong, J. Requeima, Y. Dubois, and R. E. Turner, “Convolutional conditional neural processes,” *arXiv preprint arXiv:1910.13556*, 2019. 2
- [14] Y. Zhang, W. Yu, and D. Zhu, “Terrain feature-aware deep learning network for digital elevation model superresolution,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 189, pp. 143–162, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0924271622001332> 2
- [15] H. Xiao, K. Rasul, and R. Vollgraf. (2017) Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms. 2
- [16] P. Helber, B. Bischke, A. Dengel, and D. Borth, “Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2019. 2
- [17] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” 2015. 3
- [18] Y. Bengio, I. Goodfellow, and A. Courville, *Deep learning*. MIT press Cambridge, MA, USA, 2017, vol. 1. 5
- [19] J. S. Dramsch, “70 years of machine learning in geoscience in review,” *Advances in geophysics*, vol. 61, pp. 1–55, 2020. 6