arXiv:2303.08227v1 [cs.LG] 14 Mar 2023

用于增材制造的深度神经网络霍尔效应推力器设计

Korolev K.V *俄罗斯莫斯科学生，
korolev.konstantin.v@gmail.com

霍尔效应推力器是最通用、最受欢迎的太空电推进系统之一。面向行星际任务的行业趋势推动了此类推进系统设计开发的进步。众所周知，正确确定霍尔效应推力器放电通道的尺寸会对其性能产生重大影响。由于这类推进系统的完整物理模型尚未针对快速计算和设计迭代进行优化，大多数推力器仍采用所谓的尺度定律进行设计。然而，本文着重介绍一种较为新颖的方法，在文献中，这种方法的出现频率远低于常规的尺度设计方法。借助深度机器学习，可以建立性能预测模型，从而以远低于从头设计所需的计算功率，轻松获得满足所需特性的霍尔推力器设计；同时，该方法也比常规尺度方法更加灵活。

I. 符号说明

U_d = 放电电压P = 放电功率T =推力$\dot{m}_a$ = 质量流量I_{sp} = 比冲η_m = 质量利用效率η_a = 阳极效率
j = P/v [功率密度]v = 放电通道体积h, d, L = 通用几何参数
C_* = 尺度系数组g = 自由落体加速度M = 离子质量

II. 引言

THE 深度学习的应用领域极为广泛，但本研究聚焦于霍尔效应推力器的设计。霍尔效应推力器（HET）是一种相当简单的直流等离子体加速装置，但由于其物理过程复杂且具有非线性，我们目前还没有完整的解析性能模型。尽管工业界有许多方法能够以很高的效率设计这类系统，但这些方法需要投入数百万的研究经费和大量时间。神经网络设计方法或许能够解决这一问题，再辅以少量硬件迭代调整[1]。

按比例缩放的推力器通常性能良好，但由于诸多原因，这种方法不够灵活：首先，由于所有初始实验数值都存在较大偏差，精度可能不够高；其次，要设计具有不同功率密度或能效的推力器也很困难 I_{sp} 。

另一方面，神经网络设计方法的精度优势仅适用于数据集涵盖的领域[1],；但它能够同时建立多个放电参数与几何参数之间的关系，因此这一局限很容易得到弥补。因此，这种新颖的方法与比例关系相结合，有望成为霍尔效应推力器的终极设计工具。

请注意，这两种模型都未纳入阴极效率和性能，中性气体推力分量也未纳入。以往文献中的大多数相关性都是基于假设或物理定律建立的[2], 本文中，

*创始人，Pure EP

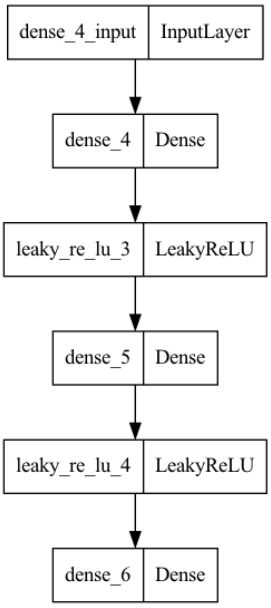


图 1 生成对抗网络生成器架构

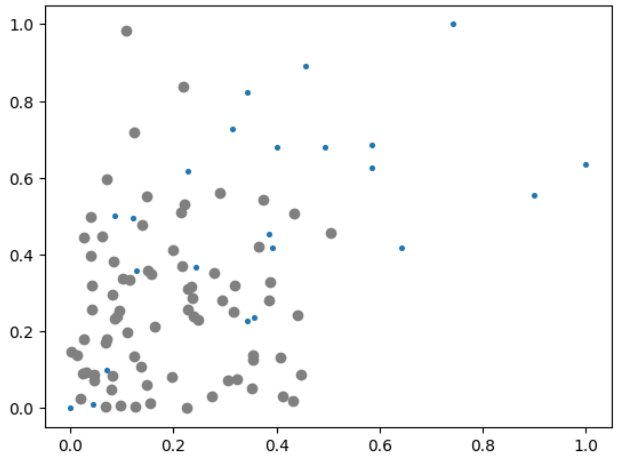


图 2 生成数据样本（d,T），灰色表示伪造数据，蓝色表示真实数据

提出了一种基于特征生成、GAN数据集扩充和机器学习特征选择的新方法。

A. 使用生成对抗网络扩充数据集

如前所述，现有数据不足以训练神经网络或大多数机器学习算法，因此我建议使用生成对抗网络生成更多相似的数据点。生成模型由两个不同的模型进行训练——生成器和判别器。生成器学习如何生成新的数据点，而判别器会将这些数据点判定为与真实数据集相似。当然，模型必须足够精确，避免对数据过拟合或产生新的未知相关性，这一点显而易见。我们通过平均绝对百分比误差（平均绝对百分比误差）和物理边界条件对模型进行了检验。在确定最有前景的架构后，该模型能够生成平均绝对百分比误差为 4.7% 的伪造数据点。我们需要测量平均绝对百分比误差，以确保数据点位于与原始数据集相同的范围内，因为本研究关注的是亚千瓦推力器。模型生成新数据点后，我们检查了它们是否符合缩放数值的物理边界（例如，推力不能超过 2，效率不能超过 1.4，依此类推；数据依据原始数据集进行了缩放，以保持质量），结果仅有 0.02% 的数据点被判定为异常值。生成对抗网络架构和数据集样本如下所示。

III. 一般关系

由于我们只会使用低功率霍尔推力器的数据集，因此可以忽略非线性方程和关系的推导，在此采用传统方法。下面定义阳极的一些参数：

$$\alpha = \frac{\dot{m}\beta}{\dot{m}_a}, \tag{1}$$

其中，α 是 β 推力器参数中的阳极参数。之所以这样选择，是因为这样一来，阴极和其他损耗就不会纳入模型。该方法的一个关键区别在于只拟合最佳且最合适的数据，从而消除尺度定律中的部分方差。不过，机器学习方法需要大量信息，而这些信息并没有足够的数据量可供使用。因此，可以进行一些简化并作出一些假设。首先，如前所述，我们不将中和器效率纳入模型。其次，该模型仅在数据集所定义的特定领域内成立；许多参数（例如阳极功率和 I_{sp}）仍采用半经验建模方法。我们要预测的是机器学习算法的输出：比冲、推力、效率、最佳质量流率和功率密度。输入的函数仅取决于功率和电压范围。就此而言

关于这个主题，我们来介绍用于电流推进器缩放的半经验方程。 .

$$h = C_h d \tag{2}$$

$$\dot{m}_a = C_m h d \tag{3}$$

$$P_d = C_p U_d d^2 \tag{4}$$

$$T = C_t \dot{m}_a \sqrt{U_d} \tag{5}$$

$$I_{spa} = \frac{T}{\dot{m}_a g} \tag{6}$$

$$\eta_a = \frac{T}{2\dot{m}_a P_d} \tag{7}$$

其中， C_x 是通过解析建模得到的缩放系数，它使方程呈线性。通常，该模型具有95%预测带，但如前所述，正是这种线性给当前电流推进器设计带来了问题（质量大、功率密度不变、平均性能）。原始数据集仅包含24条记录。参考文献为

推力器	功率，W	U_d，V	d，mm	h，mm	L，毫米	$\dot{m}_a$，毫克/秒	T, N	I_{spa}，秒
SPT-20	52.4	180	15.0	5.0	32.0	0.47	3.9	839
SPT-25	134	180	20.0	5.0	10	0.59	5.5	948
Music-si	140	288	18	2	6.5	0.44	4.2	850
HET-100	174	300	23.5	5.5	14.5	0.50	6.8	1386
KHT-40	187	325	31.0	9.0	25.5	0.69	10.3	1519
KHT-50	193	250	42.0	8.0	25.0	0.88	11.6	1339
HEPS-200	195	250	42.5	8.5	25.0	0.88	11.2	1300
BHT-200	200	250	21.0	5.6	11.2	0.94	12.8	1390
KM-32	215	250	32.0	7.0	16.0	1.00	12.2	1244
...								
HEPS-500	482	300	49.5	15.5	25.0	1.67	25.9	1587
UAH-78AM	520	260	78.0	20	40	2	30	1450
BHT-600	615	300	56.0	16.0	32	2.60	39.1	1530
SPT-70	660	300	56.0	14.0	25.0	2.56	40.0	1593
MaSMi60	700	250	60	9.42	19	2.56	30	1300
MaSMiDm	1000	500	67	10.5	21	3	53	1940
SPT-100	1350	300	85.0	15.0	25.0	5.14	81.6	1540

如下[3][4][5][6][7]下一节将讨论所使用的神经网络架构。

IV. 数据驱动的HET设计

神经网络是一类常用于人工智能领域的机器学习算法。它们是能够通过训练识别大型数据集中的模式的数学模型。生成对抗网络（GAN）的生成器架构已在前文展示。本节将重点介绍全连接网络，这类网络最常用于此类任务。HETFit 代码利用元学习算法树结构帕岑估计器，为用户选择的每个数据输入动态生成全连接神经网络（FcNN）的架构。该代码采用 OPTUNA 提供的先进实现。动态建议的架构包含 2 至 6 层，每层有 4 至 128 个节点

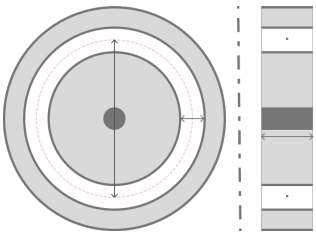


图 3 霍尔效应推力器几何形状

每层节点采用 SELU、双曲正切或 ReLU 激活函数，并使用最优优化器。代码的用户界面如下：1. 指定工作环境 2. 加载或生成数据 3. 调整架构 4. 训练并获得稳健缩放模型

A. 全连接神经网络

所有全连接神经网络均使用 PyTorch 实现，因为它是用于实验的最强大的机器学习/人工智能库。生成网络架构后，所有网络的训练循环都相似，因为它们使用梯度下降算法：损失函数：

$$L(w,b) \equiv \frac{1}{2n}\sum_{x}\|y(x)-a\|^2 \tag{8}$$

这里使用的是均方误差（MSE）误差函数，这是全连接神经网络中最常用的误差函数。接下来，我们以这种方式迭代并更新权重，持续指定数量的训练轮次。训练轮次循环：

- 获取预测值：$\hat{y}$- 计算

损失：L(w, b) - 执行反

向传播 - 更新优化器

由于数据中的偏差很大，电推进数据集极其复杂。得益于数据科学和机器学习的进步，现在可以对其进行处理。

这样，我们在 P<1000 W 输入功率和 200-500 V 范围内构建了投资回报率领域的数据集。遗憾的是，这类模型的局限之一是无法超出实际数据库的范围，同时又不牺牲性能和准确性。

B. 物理信息神经网络

为了处理未缩放的数据，我们引入了物理信息神经网络（PINN），利用方程 2-7 生成 C_x 系数。诚然，前文已指出，这种方法无法生成性能更优的霍尔效应推力器；但由于我们在与 Lee 等人相同的范围内构建了更大的数据集， [7] 因此有必要确认我们的数据集仍与原始数据集具有相同的质量。使用上述物理信息神经网络，可以拟合系数；结果显示，少数数值仅有轻微偏差，差异为百分之几，尚可接受。

C. 机器学习方法说明

我们已经讨论过 HETFit 代码的工作原理及其可生成的结果，下一节将对此作概述。但在此我想提醒大家，这项工作具有很强的实验性，因此对机器学习方法应始终持审慎态度。霍尔效应推力器中的某些等离子体放电物理机制尚未得到充分理解，因此，数据驱动的方法在特定区间的预测中可能会出现误差。关于我在这项工作中开发的设计工具，有几点说明：该工具面向机器学习经验很少甚至没有经验的人，帮助他们快速分析设计，或创建用于仿真的基准设计。该工具甚至可用于处理一般的表格数据，因为它对输入数据几乎没有任何限制。

D. 两个输入变量的预测

推力器的一项主要特性是效率。在这项工作中，我研究了多个输入值对 η_t的依赖关系。结果以预测矩阵可视化的形式呈现如下。图3同时考虑了所有先前的结果；再次说明，如果没有机器学习，这会困难得多。

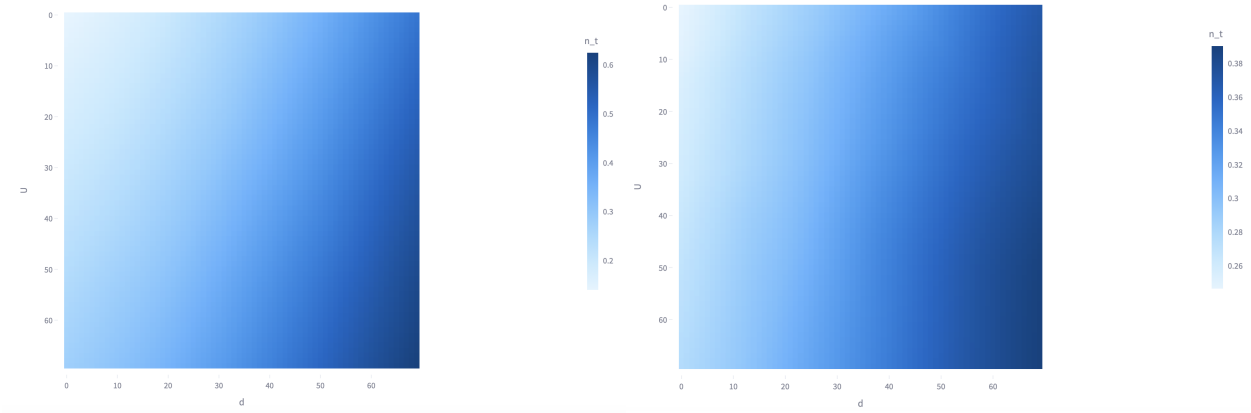


图 4 U，$d \rightarrow \eta_t$ 预测结果

图 5 U，$d \rightarrow \eta_t$ 预测结果

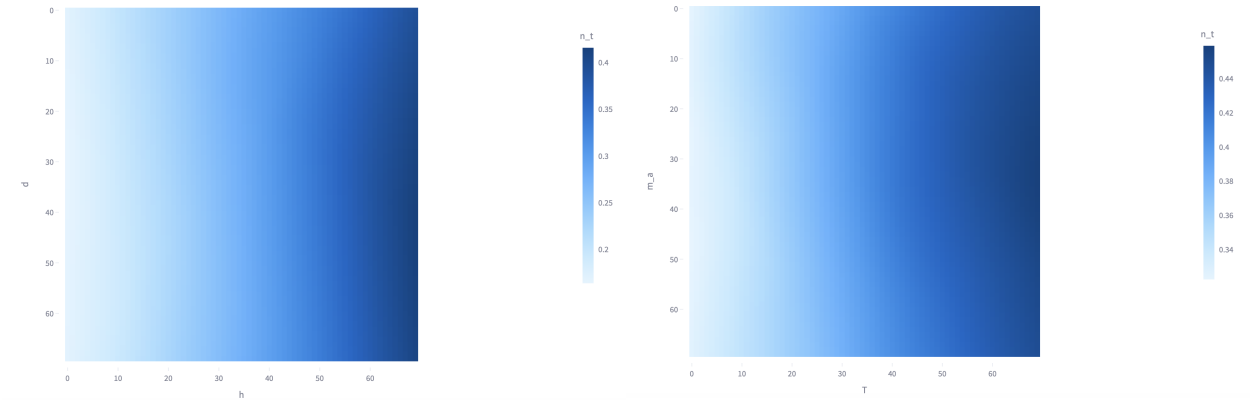


图 6 d，$h \rightarrow \eta_t$ 预测结果

图 7 $m_{\dot{a}}$，$T \rightarrow \eta_t$ 预测结果

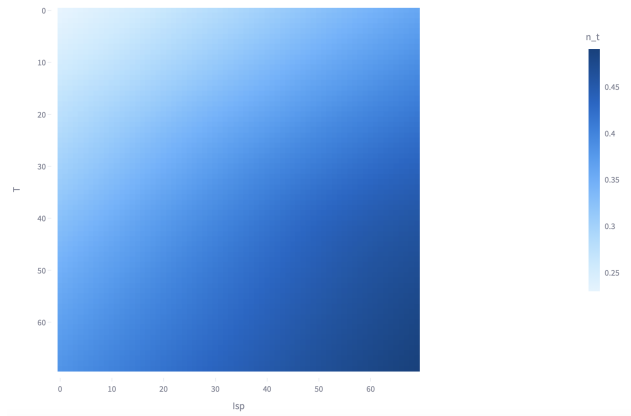


图 8 I_{sp}，$T \rightarrow \eta_t$ 预测结果

V. 结果讨论

让我们比较一下半经验方法[7], 论文中的方法[1], 以及我们的方法所作出的预测。值得一提的是，当前的方法最容易从头开始重新设计。

A. 神经网络架构生成算法

与50次迭代的情况一样，前文讨论的元学习模型能够在几秒内生成得分为0.9+ 的架构。HETFit支持登录neptune.ai环境，以全面控制仿真。示例试运行如下所示。

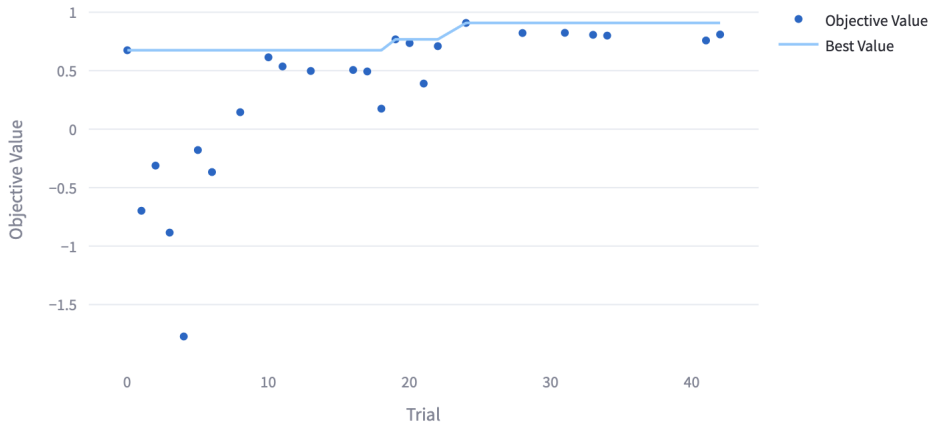


图9 TPE算法架构优化

B. 功率密度与磁通量的依赖关系

目前，这两种模型除了通用物理关系外，都不支持将磁通量纳入考量；不过，我们̃计划在后续论文中更新模型。目前，机器学习方法尚未明确功率与 B 之间的关系，但可以计算z轴上的磁场分布，磁屏蔽推力器的计算结果如下：

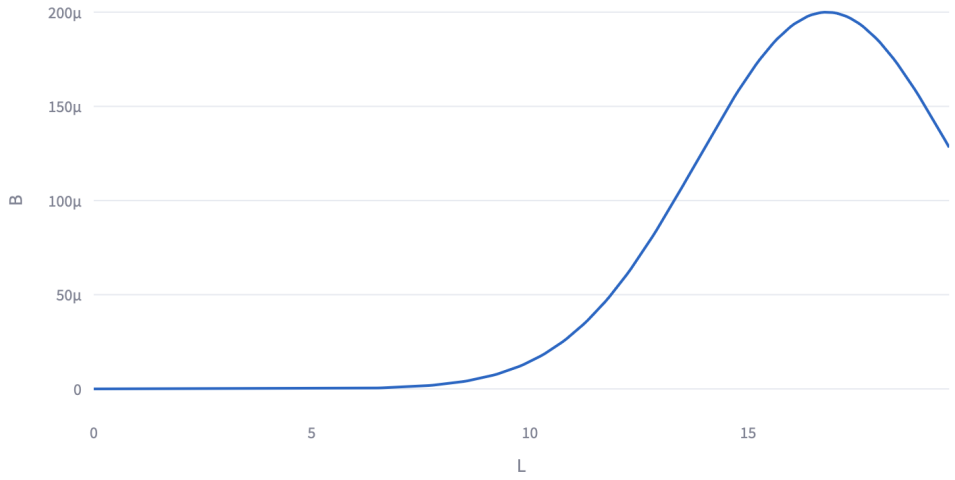


图10 dB/dz分布

C. T 对 d、P 的依赖关系

下图描述了推力与通道直径和宽度的关系，其中色调图表示推力。这种依赖关系广为人知，其周围有几条 95%预测带 [7]

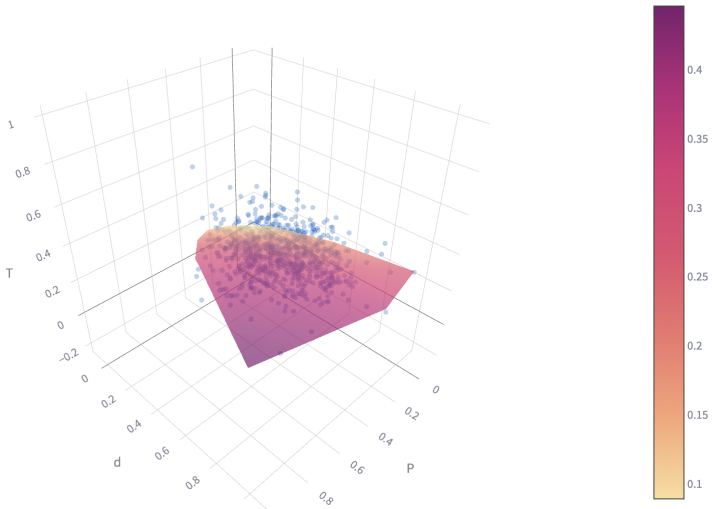


图 11 Ton d,P 的依赖关系

D. T 对 P,U 的依赖关系

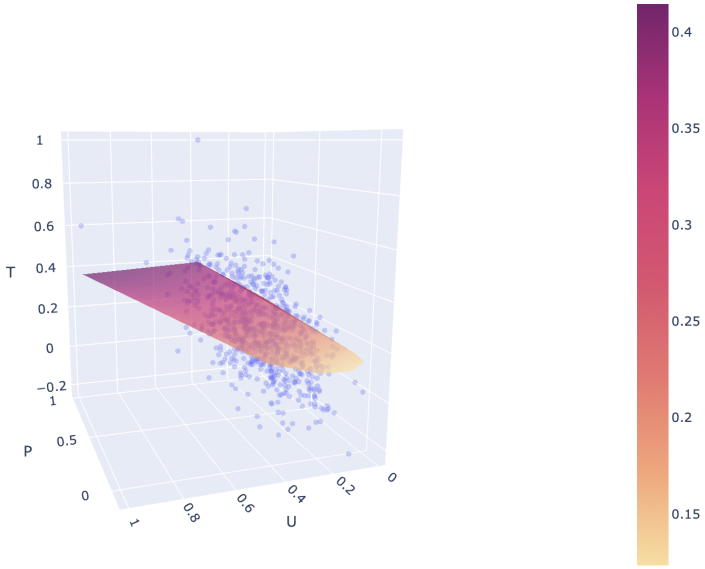


图 12 T 对 P,U 的依赖关系

E. T 对 m_a,P 的依赖关系

与[2] 相比，该模型考虑的参数比线性关系更多。因此，与半经验线性关系相比，这种方法在特定领域内的精度更高。

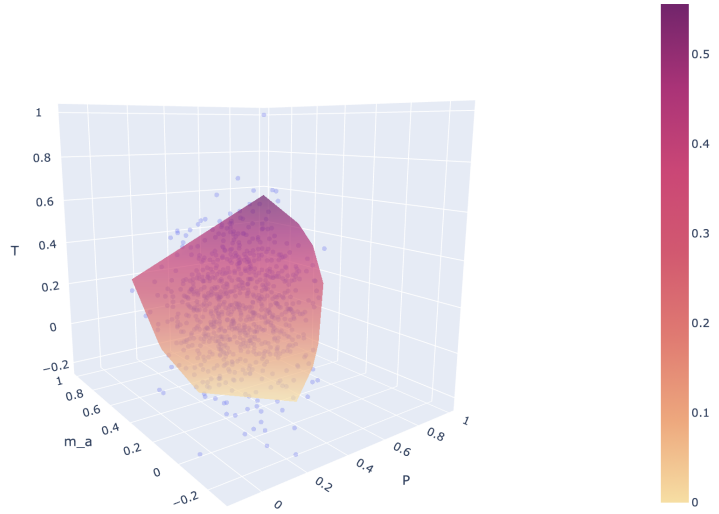


图13. T 对 m_a、P 的依赖关系

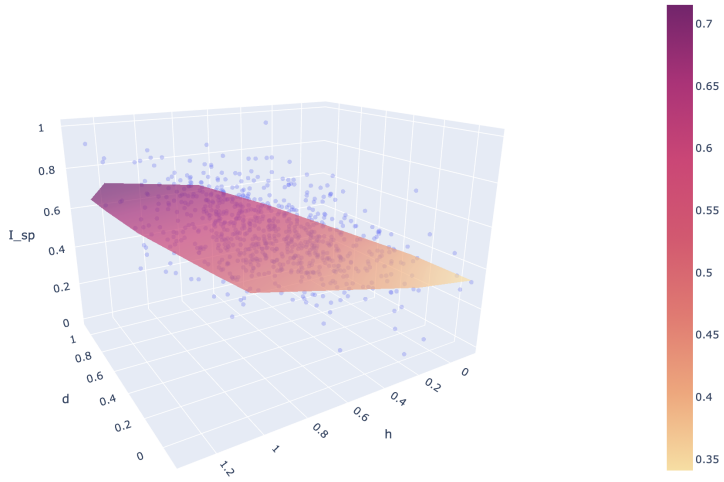


图14. I_{sp} 对 d、h 的依赖关系

F. I_{sp} 对 d、h 的依赖关系

到目前为止，我们已经生成了许多模型，但借助机器学习，我们可以构建一个模型，同时处理所有参数，因此这些图往往是该模型推理结果的三维投影。

G. 预训练模型在霍尔效应推力器通道增材制造中的应用

上述模型用于预测通道的几何形状，随后对该通道进行了仿真。作为对比，第二个通道则通过常用的尺度定律计算得出。两者的初始条件如下：

结果表明，机器学习设计的几何形状能够产生更高密度的离子，从而更高效地产生推力。HETFit 代码给出的 HET 参数估计值偏低，以补偿 HET 模型中未纳入的变量。实验已证明，这种估计方式是有效的，因为 SEM 对推力的预测始终高于真实性能。Lee 等人。 [7]

初始条件	数值
$n_{e,0}$	$1e13[m^{-3}]$
ϵ_0	4 [伏特]
V	300 [伏特]
T	293.15 [开尔文]
P_{abs}	0.5 [托]
$\mu_e N_n$	$1e25$ [$1/(V \cdot m \cdot s)$]
dt	1e-8 [s]
Body	Ar

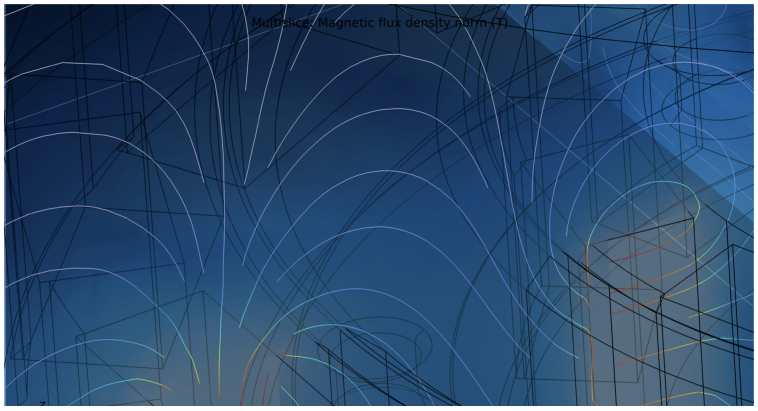


图15 磁通密度分布，磁屏蔽配置

H. 代码说明

主要概念：- 每次观测/设计会话称为一个环境，目前环境可以是 RCI 或 SCI（真实或缩放接口）

- 大多数运行参数都在此对象初始化时指定，包括通过生成对抗网络生成新样本 $\vec{}$- 内置特征生成（以10为底的功率对数、效率、 B等）- 针对每种情况进行顶级特征选择。（Boruta算法）- 使用所选模型编译环境，可以是任意 torch 模型或 scikit-learn 模型 - 训练 - 绘图、推理、保存、导出为 JIT/ONNX 格式、测量性能

I. COMSOL HET仿真

这些仿真是在 COMSOL 的等离子体物理接口中进行的，该接口能够根据初始条件和几何形状准确计算电子密度、温度和能量分布函数。以下是两个通道的对比。

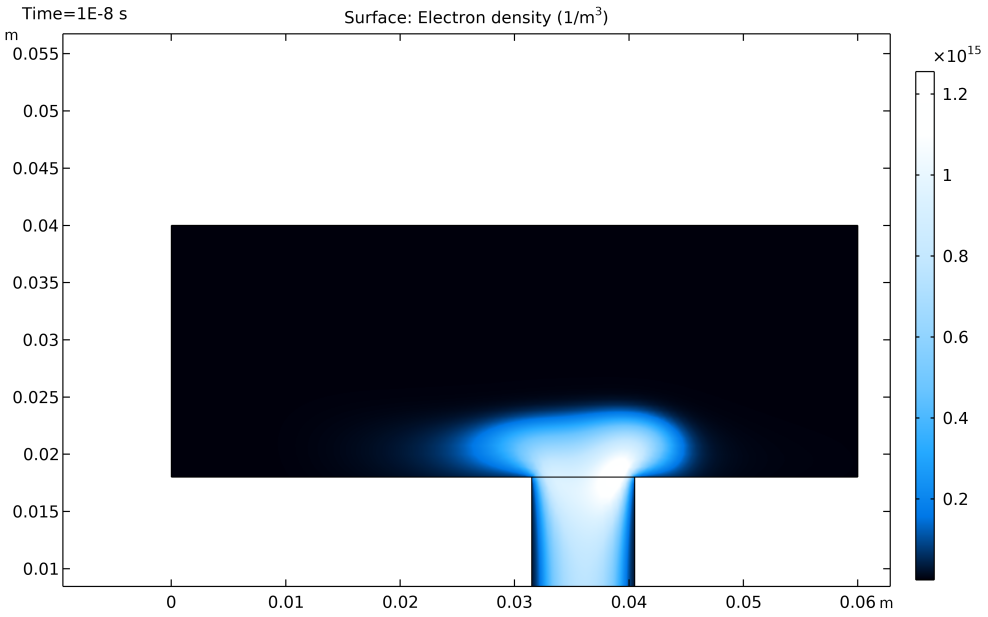


图 16 线性 SEM 几何结构下的电子密度

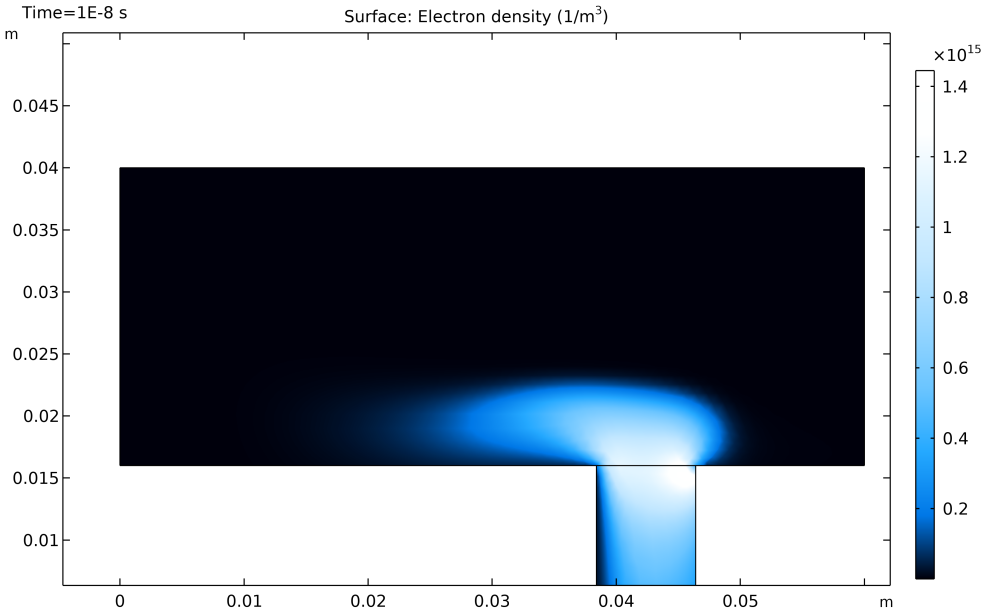


图 17 HETfit 几何形状下的电子密度

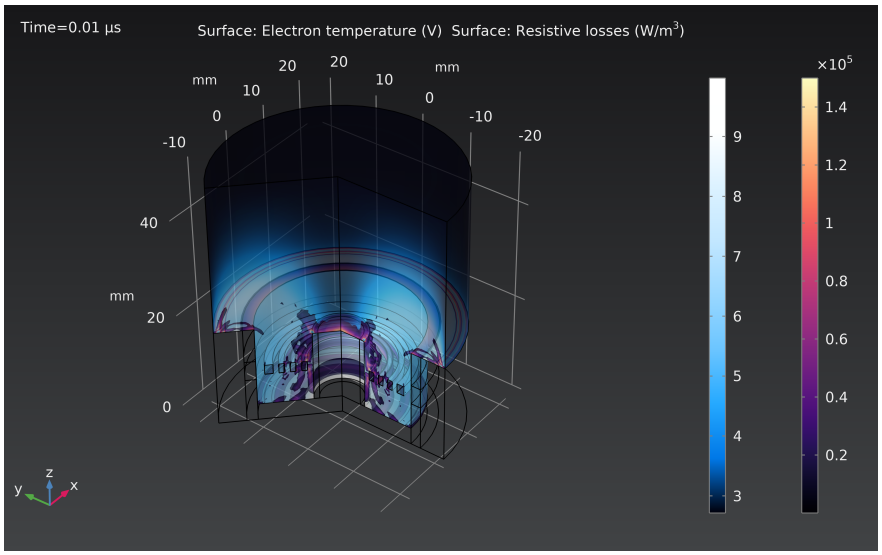


图 18 COMSOL 对所设计推力器启动过程的仿真



图 19 制造的通道

VI. 结论

总之，本文建立并提出了另一种尺度定律模型。HETFit 代码为开源代码，任何人都可以免费使用。本文还通过增材制造打印了通道，以证明其可制造性。希望这项工作有助于开发更现代的比例关系，因为现有的比例关系远非完美。

本文采用的方法最初用于 Plyashkov、Shagayda、Kravchenko、Ratnikov 和 Lovtsov [1]的研究，相较于 SEM 方法具有以下优势：能够在给定范围内更准确地预测性能，并能考虑实验数据。我认为，随着输入数据增多，机器学习方法在推力器设计中的应用将更加广泛。

本研究中的代码也可用于其他表格实验数据，因为大多数情况和任务往往相同：特征选择和模型优化。

参考文献

[1] Plyashkov, Y. V., Shagayda, A. A., Kravchenko, D. A., Ratnikov, P. D., 以及 Lovtsov, A. S.，“使用神经网络对霍尔效应推力器进行缩放”，2022年10月25日。网址：http://arxiv.org/abs/2206.04440，编号：arXiv:2206.04440。[2] Shagayda, A. A.，以及 Gorshkov, O. A.，“霍尔推力器缩放定律”，第29卷，第2期，2013年3月，第466–474页。https: //doi.org/10.2514/1.B34650，网址：https://arc.aiaa.org/doi/10.2514/1.B34650。[3] Beal, B. E.、Gallimore, A. D.、Haas, J. M. 和 Hargus, W. A.，“霍尔推力器集群羽流中的等离子体特性”，第20卷，第6期，2004年11月，第985–991页。https://doi.org/10.2514/1.3765，网址：https://arc.aiaa.org/doi/10.2514/1.3765。[4] Belikov, M.、Gorshkov, O.、Muravlev, V.、Rizakhanov, R.、Shagayda, A. 和 Snnirev, A.，“高性能低功率霍尔推力器”，第*37*届联合推进会议与展览会，美国航空航天学会，2001年7月8日。https://doi.org/10.2514/6.2001-3780，网址：https://arc.aiaa.org/doi/10.2514/6.2001-3780。[5] Kronhaus, I.、Kapulkin, A.、Balabanov, V.、Rubanovich, M.、Guelman, M. 和 Natan, B.，“同轴磁隔离纵向阳极霍尔推力器的放电特性表征”，第29卷，第4期，2013年7月，第938–949页。https://doi.org/10.2514/1. B34754，网址：https://arc.aiaa.org/doi/10.2514/1.B34754。[6] Misuri, T. 和 Andrenucci, M.，“霍尔电推进器缩放方法：改进与评估”，第*44*届*AIAA/ASME/SAE/ASEE*联合推进会议与展览会，美国航空航天学会，2008年7月21日。https://doi.org/10.2514/6.2008-4806，网址：http://arc.aiaa.org/doi/abs/10.2514/6.2008-4806。[7] Lee, E.、Kim, Y.、Lee, H.、Kim, H.、Doh, G.、Lee, D. 和 Choe, W.，“亚千瓦级霍尔效应推力器的缩放方法”，第35卷，第6期，2019年11月，第1073–1079页。https://doi.org/10.2514/1.B37424，网址：https://arc.aiaa.org/doi/10.2514/1.B37424。