arXiv:2602.00982v1 [cs.CV] 1 Feb 2026

简约导航，深度对齐：Mouse vs. AI 2025 获胜方案

Phu-Hoa Pham*
VNU-HCM理工大学 胡志明市，越南
23122030@student.hcmus.edu.vn

Chi-Nguyen Tran*University of Science，VNU-HCM 胡志明市，越南
23122044@student.hcmus.edu.vn

Dao Sy Duy Minh
VNU-HCM理工大学 胡志明市，越南
23122041@student.hcmus.edu.vn

Nguyen Lam Phu Quy
University of Science，VNU-HCM，胡志明市，越南
23122048@student.hcmus.edu.vn

Huynh Trung Kiet University of Science，VNU-HCM，胡志明市，越南
23122039@student.hcmus.edu.vn

摘要

视觉鲁棒性与神经对齐仍是开发能够媲美生物视觉系统的人工智能体所面临的关键挑战。我们介绍了 HCMUS_TheFangs 团队在 NeurIPS 2025 Mouse vs. AI：鲁棒视觉觅食竞赛两个赛道上的获胜方案。在赛道 1（视觉鲁棒性）中，我们证明架构简洁性与针对性组件相结合能够带来更优的泛化能力，借助由门控线性单元和观测归一化增强的轻量级双层卷积神经网络，取得了 95.4% 的最终得分。在赛道 2（神经对齐）中，我们开发了一种具有 16 个卷积层和基于 GLU 的门控的深度类 ResNet 架构，以 1780 万参数实现了神经预测性能的第一名。我们对在 60K 至 1.14M 步之间训练的十个模型检查点进行系统分析，发现训练时长与性能呈非单调关系，约在 200K 步时取得最优结果。通过全面的消融研究与失败案例分析，我们阐释了为何更简单的架构在视觉鲁棒性上表现更优，而容量更大的深层模型则能实现更好的神经对齐。我们的结果挑战了视觉运动学习中关于模型复杂度的传统假设，并为开发鲁棒的、受生物学启发的视觉智能体提供了实用指导。

1 引言

在多变的视觉条件下实现鲁棒导航，是生物视觉的一项基本能力，而人工系统至今仍难以企及。深度强化学习虽然在受控环境中取得了显著成功，但部署后的智能体在遇到训练分布之外的视觉扰动时，往往会失效。这种脆弱性与小鼠等生物系统形成了鲜明对比：尽管环境变化显著，它们仍能保持稳定的导航性能，并且只需少量接触便能适应新的视觉条件 [9]。

*贡献相同。

NeurIPS 2025 的 The Mouse vs. AI：鲁棒视觉觅食竞赛通过两项互补的评测赛道，为研究这一鲁棒性差距提供了独特的基准。赛道 1（视觉鲁棒性）评估智能体在多种留出的视觉扰动下的表现，包括雾、光照变化及其他符合生态实际的变换。赛道 2（神经对齐）评估人工视觉表征对同一导航任务期间从小鼠视觉皮层记录的超过 19,000 个神经元的神经活动的预测能力。该双重评测框架能够同时考察行为鲁棒性和生物学合理性，贯通强化学习、计算机视觉与系统神经科学 [9]。

我们的初步研究遵循了传统观点，即复杂的视觉导航任务需要相应复杂的架构。我们实现并系统评估了几种最先进的方法，包括用于多尺度视觉处理的 InceptionNet [12] 、包含 24 个残差块的深度 IMPALA ResNet [2] ，以及基于 LSTM 的时序模型。然而，这些复杂架构始终表现出训练不稳定性、严重过拟合，以及对受扰动视觉输入的泛化能力较差等问题。InceptionNet 完全未能收敛，而深度 ResNet 虽然取得了较高的训练性能，但在视觉扰动条件下评估时性能下降了 35%。

这些失败促使我们从根本上重新审视研究方法。我们没有继续增加架构复杂度，而是研究经过谨慎选择的增强措施能否让更简单的设计实现更优的鲁棒性和神经对齐性能。这项研究最终形成了两种针对各自赛道进行优化的不同架构。针对赛道 1，我们开发了一种轻量级双层卷积神经网络，采用门控线性单元和观测归一化；尽管参数数量显著少于竞争方法，仍取得了 95.4% 的最终得分。针对赛道 2，我们设计了一种受深度 ResNet 启发的架构，包含 16 个卷积层并采用基于 GLU 的门控，以 1,780 万个参数实现了排名第一的神经对齐性能。

赛道 1 与赛道 2解决方案之间的差异揭示了行为表现与神经相似性之间关系的一项重要发现。视觉鲁棒性得益于架构简洁性，因为它能避免模型对训练特定模式过拟合；而神经对齐则需要足够的模型容量，以捕捉生物视觉皮层中丰富的层次化表征。我们还系统分析了覆盖不同训练时长（6 万至 114 万步）的 10 个模型检查点，进一步发现训练时间与性能之间并非单调关系：最优检查点出现在约 20 万步时，而非训练收敛时。

本文的主要贡献有四项。第一，我们证明，在视觉鲁棒性方面，采用针对性增强的简洁架构显著优于复杂的深度网络，这挑战了人们对视觉运动任务模型需求的普遍认知。第二，我们表明，深度更大、容量更高的架构能够实现更优的神经对齐，而简洁架构则在行为鲁棒性方面表现出色。第三，我们全面记录了多种失败的方法，包括 InceptionNet、深度 ResNet、LSTM 集成以及数据增强策略，为避免常见问题提供实用指导。第四，我们通过系统的消融研究量化了各个组件的贡献，并分析了训练时长、架构选择与泛化性能之间的关系。

2 相关工作

强化学习中的视觉鲁棒性。 在分布偏移下实现稳健的视觉感知，仍是强化学习面临的一项根本挑战。领域随机化 [13]通过让智能体在多样的视觉条件下进行训练来应对这一问题，而数据增强技术 [6]则通过人为扰动训练观测来提升泛化能力。然而，这两种方法都会增加训练的复杂性，若使用不当，甚至可能适得其反，损害性能。我们的实验验证了这一现象：全面的增强流程使我们的 ResNet 基线最终得分从 87.7% 降至 59.8%。近期研究探索了辅助任务、正则化技术和元学习等替代方法，但这些方法往往需要谨慎进行超参数调优，并消耗额外的计算资源。

视觉控制的卷积神经网络架构。 Mnih等人 [8] 提出的Nature CNN确立了三层卷积架构作为视觉强化学习的事实标准。后续工作探索了更深的架构，IMPALA [2] 引入了残差连接-

用于可扩展的分布式训练。InceptionNet [12] 表明，多尺度卷积处理能够提升图像分类效果，因而被应用于各类视觉领域。然而，将这些架构创新从监督学习迁移到强化学习并非易事；在强化学习场景中，更深的网络常常表现出训练不稳定性和过拟合。我们的工作提供了系统性证据，表明对于某些视觉控制任务，架构深度可能有害，尤其是当主要目标是对扰动的鲁棒性时。

注意力机制与门控。注意力机制改变了现代深度学习，使模型能够有选择地关注相关信息。Transformer 架构 [14]推广了用于序列建模的自注意力，而门控线性单元 [1] 则引入了乘法门控，以实现信息的选择性流动。包括 SwiGLU [11]在内的近期变体，在各种语言建模任务中展现出更优的性能。在视觉领域，空间 Softmax [7]为从卷积层中提取空间连贯特征提供了一种有效机制，使模型能够学习关注与任务相关的图像区域。我们的工作将这些门控机制应用于鲁棒视觉导航，并表明 GLU 模块能够持续提升行为表现和神经对齐。

神经行为学基准。神经科学与人工智能的融合由来已久，人们也日益认识到，生物智能能够为人工系统设计提供启发 [16]。近期的基准测试开始将神经数据与行为指标结合起来，从而能够直接比较人工与生物表征。Mouse vs. AI 竞赛进一步拓展了这一方法，通过在小鼠和智能体执行的相同任务上，同时评估行为鲁棒性和神经对齐。这种双重评估框架揭示，行为能力并不一定意味着生物学合理性，我们两个赛道的最优架构各不相同便是明证。了解小鼠视觉处理 [4] 有助于深入认识高效的导航策略，并为设计鲁棒的人工智能体提供参考。

3 竞赛概述与问题表述

3.1 任务描述与环境

Mouse vs. AI 竞赛设置了一项在自然风格的 3D Unity 环境中实现的视觉引导觅食任务。智能体从随机起始位置出发，朝视觉提示的目标球体移动，并通过第一人称自我中心摄像机观察环境。每个时间步的视觉观测由一幅 86×155 像素的灰度图像构成；选择这一分辨率是为了近似对应行为实验中小鼠的视敏度。动作空间包含三个连续控制维度：前后速度、横向（侧移）速度和角旋转速率。回合会在成功获取目标或达到固定超时时间后终止。奖励结构会对抵达目标给予正向强化，同时惩罚耗时过长，以激励高效导航。

该环境融入了多项设计元素，以确保生态效度并与小鼠实验保持一致。目标球体在不同回合中的位置各不相同，避免智能体记住固定轨迹。视觉干扰物和环境复杂性要求智能体真正理解视觉信息，而非仅进行简单的特征匹配。连续动作空间要求智能体进行平滑、协调的控制，而不是离散决策。这些特性共同确保成功的智能体必须具备稳健的视觉感知与视运动协调能力。

3.2 赛道 1：视觉鲁棒性评估

赛道 1 评估智能体对未见过的视觉扰动的泛化能力。这些扰动在改变低层视觉统计的同时保留了任务语义。评估方案分别考察在标准训练分布上测得的平均成功率（ASR），以及在一系列留出扰动下测得的修正成功率（MSR）。这些扰动包括不同浓度的雾、模拟一天中不同时段的光照变化，以及小鼠在自然环境中会遇到的其他生态真实视觉变换。最终得分将这些指标加权平均，在标准性能与鲁棒性之间取得平衡。具体而言，赛道 1 的得分为

计算如下：

$$\text{Final Score}_{\text{Track 1}} = \frac{1}{2} \cdot \text{ASR} + \frac{1}{2} \cdot \text{MSR} \tag{1}$$

3.3 赛道 2：神经对齐评估

赛道 2 引入了一种截然不同的评估标准，侧重于生物学合理性。该赛道不评估行为指标，而是考察人工智能体的内部表征在多大程度上能够预测小鼠在执行同一导航任务时记录到的视觉皮层神经活动。竞赛组织者提供了来自 19,000+ 个神经元的神经记录，涵盖初级视觉皮层（V1）和高级视觉区；这些记录是在小鼠执行觅食任务时通过双光子钙成像采集的。参赛者从训练好的智能体中提取特征表征，并将其提交至评估服务器，由服务器计算神经预测指标。

主要评估指标是线性读出 R^2，用于量化通过岭回归利用智能体特征预测神经反应的效果。对于给定层的特征 $\mathbf{F} \in \mathbb{R}^{N \times D}$ 和神经反应 $\mathbf{N} \in \mathbb{R}^{N \times K}$，其中 N 为样本数量，D 为特征维度，K 为神经元数量，评估拟合如下：

$$\hat{\mathbf{N}} = \mathbf{F}\mathbf{W} + \mathbf{b} \tag{2}$$

使用带交叉验证的岭回归来确定正则化强度。决定系数 R^2 用于量化预测质量。次要指标通过使用表征不相似性矩阵的相关性比较智能体表征空间与神经表征空间的几何结构，来评估表征相似性。总之，这些指标评估智能体是否学到了能够映射生物视觉处理的视觉表征。

3.4 赛道之间的关系

双赛道结构揭示了设计人工视觉系统时存在的根本张力。赛道 1 强调行为鲁棒性和泛化，倾向于采用抽象掉低层视觉统计的表征。赛道 2 强调生物对齐，这可能要求表征保留有关视觉刺激结构的信息，即使这些信息与行为无关。我们的研究表明，这些目标可能需要不同的架构选择：更简单的模型在行为鲁棒性方面表现出色，而更深的基于注意力的架构则能实现更优的神经对齐。

4 方法

4.1 赛道 1：轻量级架构与针对性增强

我们的赛道 1 方案优先考虑架构简洁性，同时融入了能够增强视觉鲁棒性的特定组件。该架构包含三个主要部分：用于提取视觉特征的双层卷积骨干网络、用于选择性特征筛选的门控线性单元模块，以及用于实现尺度不变性的观测归一化。这种设计理念源自系统性的实验；实验表明，复杂架构始终无法在训练条件之外实现泛化。

视觉编码器由两个卷积层组成，并采用激进的空间下采样。第一层对 $86 \times 155 \times 1$ 输入应用步幅为 4 的 8×8 卷积核，生成 16 个特征通道。这种激进的下采样会立即降低空间分辨率，同时提取边缘和颜色梯度等低级视觉特征。第二层应用步幅为 2 的 4×4 卷积核，将特征通道扩展至 32 个，同时进一步降低空间维度。两层均采用负斜率为 0.2 的 LeakyReLU 激活，使梯度能够通过非正激活值传播。生成的特征图会被展平，并通过全连接层投影至 256 维。视觉编码过程的数学表示如下：

$$\mathbf{h}_1 = \text{LeakyReLU}(\text{Conv2D}_{8\times8, s=4}(\mathbf{x}), \alpha = 0.2) \tag{3}$$

$$\mathbf{h}_2 = \text{LeakyReLU}(\text{Conv2D}_{4\times4, s=2}(\mathbf{h}_1), \alpha = 0.2) \tag{4}$$

$$\mathbf{z} = \text{FC}_{256}(\text{Flatten}(\mathbf{h}_2)) \tag{5}$$

其中，$\mathbf{x} \in \mathbb{R}^{86\times155\times1}$ 表示输入观测，$\mathbf{z} \in \mathbb{R}^{256}$ 表示编码后的特征向量。

门控线性单元模块对编码后的特征进行处理，以实现信息的选择性流动。GLU 采用两条并行路径处理编码后的特征：一条是使用 Swish 激活函数的特征变换路径，另一条是使用 Sigmoid 函数激活的门控路径。Swish 激活函数定义为 Swish(x) $= x \cdot \sigma(x)$，能够提供平滑的非线性变换，并具有良好的梯度特性。门控机制能够学习识别在视觉扰动下仍然可靠的特征，同时抑制对噪声敏感的成分。GLU 的变换表示为：

$$\mathbf{f} = \text{Swish}(\text{FC}_{\text{hidden}}(\mathbf{z})) \tag{6}$$

$$\mathbf{g} = \sigma(\text{FC}_{\text{hidden}}(\mathbf{z})) \tag{7}$$

$$\mathbf{h}_{\text{GLU}} = (\mathbf{f} \odot \mathbf{g}) \tag{8}$$

$$\mathbf{y} = \text{FC}_{256}(\mathbf{h}_{\text{GLU}}) \tag{9}$$

其中，$\odot$ 表示逐元素相乘，隐藏维度与输入维度相同，以保留特征容量。输出投影将维度恢复为 256，从而可按需使用残差连接。

观测归一化是我们 Track1 架构的最后一个组成部分。我们采用训练期间通过指数移动平均维护的运行统计量，并按以下方式计算归一化后的观测值：

$$\hat{\mathbf{x}} = \frac{\mathbf{x} - \boldsymbol{\mu}_{\text{running}}}{\boldsymbol{\sigma}_{\text{running}} + \epsilon} \tag{10}$$

其中，$\boldsymbol{\mu}_{\text{running}}$ 和 $\boldsymbol{\sigma}_{\text{running}}$ 分别是按通道计算的均值和标准差估计值，$\epsilon = 10^{-8}$ 用于防止除以零。这种归一化可实现对全局光照变化的不变性，而全局光照变化是评估方案中的主要视觉扰动来源。运行统计量会在训练期间逐步调整，以确保学习过程稳定。

4.2 赛道 2：采用 GLU 门控的深度残差网络架构

我们的赛道2解决方案采用了显著更深的架构，专门用于捕捉与生物视觉皮层相一致的层次化视觉表征。该架构包含 16 个卷积层，以残差结构组织，并采用基于 GLU 的门控机制，共计 1780 万个参数。更高的模型容量使其能够学习从低级边缘到高级物体表征的多尺度视觉特征，从而呼应哺乳动物视觉皮层的层次化组织。

视觉编码器首先采用激进的空间下采样层，应用步幅为 4 的 4×4 卷积核，从灰度输入生成 64 个特征通道。该第一层提取初级视觉特征，同时将空间分辨率降低为原来的 1/16。随后，架构经过四个通道容量逐步增加的阶段：$64 \to 128 \to 256 \to$512 个通道。每个阶段采用 2×2 步幅卷积进行空间下采样，随后是成对的步幅为 1 的 3×3 卷积层，并按照 ResNet 设计模式 [3]组织为残差块。残差连接有助于梯度流在深度网络中传播：

$$\mathbf{h}_0 = \text{Conv2D}^{64}_{4\times4,s=4}(\mathbf{x}) \tag{11}$$

$$\mathbf{h}_i = \mathbf{h}_{i-1} + \mathcal{F}(\mathbf{h}_{i-1}; \mathbf{W}_i) \quad \text{for } i = 1, \ldots, 15 \tag{12}$$

其中 $\mathcal{F}$ 表示一个采用 LeakyReLU 激活的两层残差块。

在卷积特征提取之后，我们采用多层 GLU 模块来实现自适应特征路由与选择。GLU 层应用学习得到的门控函数，以确定如何对特征进行加权和选择性过滤。对于给定的特征表示 $\mathbf{h}$，门控机制计算：

$$\mathbf{g} = \text{Softmax}(\mathbf{W}_g \mathbf{h}) \tag{13}$$

其中 $\mathbf{W}_g$ 为学习得到的门控权重。softmax 运算确保门控激活形成用于选择性特征加权的概率分布。我们的实证分析表明，只有第一个 GLU 层在训练过程中发生显著学习，后续层的参数更新极小，从而对卷积特征进行选择性过滤。最终表征被送入独立的策略头和价值头，分别用于动作选择和价值估计。

表 1：赛道 1 视觉鲁棒性结果。ASR 衡量标准条件下的成功率，MSR 衡量视觉扰动下的成功率。最终得分是加权结果，侧重于鲁棒泛化。

模型架构	ASR（%）	MSR（%）	最终得分（%）
IMPALA ResNet（24 层）	80.96	51.00	65.98
IMPALA ResNet（4 层）	91.40	84.00	87.70
+数据增强	72.60	47.00	59.80
InceptionNet		未能收敛	
SimpleCNN（我们的方案）	94.20	89.00	91.60
SimpleCNN + 门控线性单元	95.60	88.00	91.80
SimpleCNN + 门控线性单元 + 归一化	**96.80**	**94.00**	**95.40**

庞大的参数量（1780 万，而赛道 1 为 140 万）提供了捕捉丰富层次化表征的容量。尽管这种深度会对视觉鲁棒性不利（正如我们在赛道 1 消融实验中所证明的），神经对齐目标却能从增强的表征容量中受益，从而匹配视觉皮层区域中生物神经元多样的调谐特性。

4.3 训练流程和超参数

两种架构均使用近端策略优化（PPO）进行训练， [10], 这是当前最先进的同策略强化学习算法。PPO 通过裁剪代理目标，在样本效率与训练稳定性之间取得平衡，因此非常适用于连续控制任务。附录中的表 4 列出了两个赛道的完整训练配置。

对于赛道 1，我们采用两阶段训练策略：先训练卷积主干网络 1,400,000 步，然后添加门控线性单元模块，并从第一阶段的最佳检查点继续训练 350,000 步。这种分阶段方法使主干网络能够先学习稳健的视觉特征，再由门控线性单元模块进一步优化特征选择。对于赛道 2，我们进行了广泛实验，训练步数从 60,000 到 1,140,000 不等，并每隔 20,000 步保存一次检查点。系统性的检查点分析表明，最佳性能出现在约 200,000 步时，而非训练收敛时；延长训练有时反而会降低泛化能力。所有模型均使用 ML-Agents 工具包在 NVIDIA GPU 上训练 [5], ，训练时间根据训练步数而异：赛道 1 为 6–8 小时，赛道 2 为 12–24 小时。

5 实验结果

5.1 赛道 1：视觉鲁棒性结果

表 1 展示了我们在赛道 1 上的全面实验结果，将我们的轻量级架构与若干基线方法进行了比较。我们结合 SimpleCNN、门控线性单元和观测归一化的最终模型取得了 95.4% 的最终得分，在赛道 1 排行榜上稳居首位。相较于具有 4 层的 IMPALA ResNet 基线，这是一项显著提升；该基线尽管参数量明显更多，最终得分仅为 87.7%。与更深的架构相比，性能差距更为显著：24 层 ResNet 因严重过拟合，最终得分仅为 65.98%。

这些结果揭示了有关视觉鲁棒性架构选择的若干重要见解。首先，ASR 与 MSR 之间的性能差距直接反映了模型对训练条件的过拟合。深达 24 层的 ResNet 存在 30 个百分点的差距（80.96% 对比 51.00%），这表明模型记住了训练特定视觉模式，而非学习可泛化特征。相比之下，我们的 SimpleCNN 仅有 2.8 个百分点的差距（96.80% 对比 94.00%），展现出强大的泛化能力。其次，将深度从 4 层增加到 24 层会显著损害性能，这与“更深的网络能实现更好的视觉理解”这一传统观点相悖。第三，为 ResNet 添加数据增强后，性能反而从

表 2：赛道 2 神经对齐结果，展示不同训练检查点的性能。所有深度模型均采用 ResNet + 门控线性单元架构，包含 1780 万个参数。所示分数为竞赛排名分数（越低越好）。

模型 ID	训练步数	竞赛排名	分数	参数
368923	1,139,998	**1**	**0.1517**	1780万
368751	199,996	2	0.1507	1780万
368551	219,977	3	0.1497	1780万
368578	319,976	4	0.1496	1780万
368628	79,998	5	0.1495	1780万
368967	199,994	6	0.1470	1780万
368507	375,323	6	0.1470	17.8M
369015	519,976	6	0.1470	17.8M
368622	59,975	11	0.1467	17.8M
简单CNN（赛道 1）	500,027	13	0.1451	1.4M

87.70% 降至 59.80%，这表明简单的数据增强可能会造成分布不匹配，从而损害学习效果。

5.2 赛道 2：神经对齐结果

我们的赛道 2 提交在神经对齐指标上取得了第一名的性能，表明我们的深度残差网络架构所学到的表征与生物视觉处理高度吻合。表 2 汇总了我们系统训练时长分析中不同模型检查点的结果。该架构由组织在残差块中的 16 个卷积层构成，并采用基于门控线性单元的门控，所有检查点共计 1780 万个参数。有趣的是，神经对齐性能并未随训练时长单调提升，最优检查点出现在约 200,000 步附近。

系统的检查点分析揭示了若干值得注意的模式。模型 368923 经过 114 万步训练，取得了最高的神经对齐分数 0.1517，表明延长训练可以优化表征，使其更好地匹配生物处理过程。然而，训练时长与性能之间的关系高度非单调。模型 368751 仅训练了 20 万步，便以 0.1507 的分数取得第 2 名，尽管训练步数少了 5.7× ，性能差距仅为 0.66%。多个训练了 20 万步的模型（368751、368967、368551）均取得了强劲性能，表明这代表了最优训练时长。相反，训练了 52 万步的模型 369015 仅取得第 6 名，表现甚至不如仅训练了 8 万步的模型 368628。这种非单调性表明，即使衡量的是神经对齐而非行为表现，延长训练也可能导致过拟合或灾难性遗忘。

比较赛道 1 和赛道 2 的架构，可以看出行为鲁棒性与神经对齐所需条件的不同。赛道 1 中轻量级的 SimpleCNN 仅有 140 万个参数，得分为 0.1451，排名第 13。虽然其神经对齐表现尚可，但与针对赛道 2 优化的深度 ResNet 架构相比，差距显著。参数量相差 12.8×（140 万与 1,780 万），神经对齐得分约提升 4.6%。这表明，生物视觉皮层采用了丰富的层次化表征，需要相当大的模型容量才能近似；即便在导航任务中，更简单的表征已足以胜任行为层面的任务。

5.3 消融研究

为系统量化各个架构组件的贡献，我们对两条赛道都进行了全面的消融研究。表 3 展示了赛道 1 的结果，评估了观测归一化和 GLU 模块的影响。观测归一化带来了最大的单项提升，使最终得分从 91.6% 提高至 95.4%，提升了 3.8 个百分点。这一显著提升体现了输入预处理在实现

表 3：赛道 1 架构的消融研究，通过系统地移除组件来量化各组件的独立贡献。

配置	ASR (%)	MSR (%)	最终得分 (%)
完整模型（SimpleCNN +GLU + 归一化）	96.80	94.00	95.40
移除归一化	95.60	88.00	91.80
移除 GLU	94.20	89.00	91.60
同时移除两者	94.20	89.00	91.60

稳健视觉感知方面的关键作用。视觉扰动通常表现为全局光照变化或对比度变化，而归一化通过将输入重新缩放至标准化分布，直接应对这些问题。

与基础 SimpleCNN（91.6%）相比，GLU 模块使 SimpleCNN + GLU 的准确率达到 91.8%，提升幅度较为有限，仅为 0.2 个百分点。尽管这一提升看似微小，但在多次训练轮次和不同评估条件下都表现出一致性。门控机制能够选择性地抑制噪声敏感特征，从而提升鲁棒性，尤其是在严重扰动条件下。有趣的是，归一化与 GLU 的贡献近似具有可加性：两者结合时达到 95.4%，而未使用任一组件时为 91.6%，这正好体现了各自独立提升幅度的预期总和。

对于赛道 2，我们将深度 ResNet 架构（1780 万个参数）与赛道 1 中的轻量级 SimpleCNN（140 万个参数）进行比较，以评估模型容量对神经对齐的影响。如表 2 所示，该深度架构在各个检查点上均取得第 1—6 名，而 SimpleCNN 仅取得第 13 名，表明增加模型容量能够显著改善神经对齐。参数量增加 12.8$\times$ 倍后，模型能够在多个尺度上学习层次化特征：从早期层中的低级边缘检测器，到后期层中的高级语义表征，这与生物视觉通路的组织方式相呼应。残差块结构和 GLU 门控机制促进了梯度流与自适应特征路由，使该架构能够近似复现视觉皮层区域中观察到的多样化调谐特性。

5.4 失败方法分析

记录失败的方法，有助于深入了解开发稳健视觉智能体时常见的陷阱。我们系统地评估了几种复杂的架构和技术，但它们的表现最终不如简单的基线。InceptionNet 通过采用具有不同卷积核尺寸的并行分支实现多尺度卷积处理，但在训练过程中完全未能收敛。在超过 500,000 个训练步数后，模型的回合回报仍然存在较大波动，没有呈现持续改善的趋势。对学习到的行为进行分析后发现，当目标出现在附近时，智能体表现尚可，但面对远处目标时则完全失效，这表明多分支架构难以在强化学习环境中学习连贯的视觉表征。

具有 24 层的 Deep IMPALA ResNet 取得了良好的训练性能，ASR 达到 80.96%，但在受到扰动时表现灾难性地失效，MSR 仅为 51.00%。标准条件与受扰动条件之间相差 30 个百分点，这直接量化了模型对训练特定视觉模式的过拟合。对学习到的特征进行可视化后发现，深层会形成高度特异的特征检测器：它们对训练分布中的输入反应强烈，却无法良好地响应受扰动的观测。这表明，增加深度虽然能提升记忆能力，却未必能改善学习到的表征质量。

LSTM 的集成代表了引入时序建模的一次尝试，其依据是这样一种直觉：导航需要记住先前的观测，而时序一致性可以提升鲁棒性。我们同时实现了标准 LSTM 与 SRU-LSTM [15]（空间增强循环单元）变体。从随机初始化开始训练时，LSTM 模型表现出严重的不稳定性，回合回报剧烈波动且无法收敛。当从预训练的 SimpleCNN 检查点初始化时，奖励立即从大约 -20 跌至 -50，并在 100,000 步内未能恢复。即便在 LSTM 训练最终趋于稳定的情况下，89.9% 的最终性能仍低于前馈 SimpleCNN 基线的 91.6%。我们

假设 LSTM 对空间信息的顺序处理破坏了对视觉导航至关重要的二维空间结构，并且额外的循环参数会导致对特定训练轨迹的过拟合。

数据增强策略也带来了令人意外的负面结果。我们实现了一套全面的增强流程，包括天气效果（雨、雾、夜间条件）、颜色抖动（亮度、对比度和饱和度变化）以及随机遮挡。将这套流程应用于 IMPALA ResNet 后，性能从 87.7% 降至 59.8%，下降了 27.9 个百分点。这一反直觉的结果表明，若不加区分地应用数据增强，可能会引入分布不匹配，干扰模型学习。评估方案中的视觉扰动保持了语义一致性和逼真的外观，而过于激进的数据增强则可能产生不真实的外观，提供误导性的训练信号。这凸显了将数据增强策略与评估条件仔细匹配的重要性，而不是一概采用通用技术。

6 讨论

6.1 为什么简洁性能够实现视觉鲁棒性

我们的赛道 1 结果表明，架构简洁性与有针对性的增强相结合，在视觉鲁棒性方面显著优于复杂的深度网络。多个因素促成了这一反直觉的发现。首先，浅层网络的容量限制阻止其记忆训练特定模式，迫使模型学习能够迁移至不同条件的可泛化特征。拥有数百万参数的深度网络可能会形成高度专门化的特征检测器，虽然在训练中达到完美性能，却无法应对受扰动输入。我们的双层 SimpleCNN 表征容量有限，因此必须学习在视觉变化下仍保持稳定的特征。

其次，导航任务可能并不需要图像分类中深度架构所旨在实现的层次化多尺度处理。检测目标球体并避开障碍物，主要需要识别显著视觉特征（颜色、形状、相对位置），而非构建复杂的物体部件层次结构。浅层网络足以从觅食环境中相对简单的视觉场景提取这些特征。这种与任务相匹配的归纳偏置，即让架构约束符合问题结构，能够比灵活的深度架构实现更好的泛化；后者必须通过学习才能发现适当的表征。

第三，在强化学习场景中，训练稳定性至关重要，因为深度网络经常出现梯度病态、局部极小值和灾难性遗忘。我们的浅层架构在不同随机种子和超参数设置下都能可靠地完成训练，并始终收敛到高质量解。深度网络在不同训练轮次间表现出较高方差，有些初始化甚至完全无法学会任务。这种可靠性优势在汇总不同评估条件下的结果时，会直接转化为更好的性能。

6.2 神经对齐的模型容量

赛道 2 的结果表明，要匹配生物神经表征，需要采用不同的架构。尽管深度 ResNet 的参数量比我们的赛道 1 方案多 12.8×，但它仍实现了最高的 top-1 神经对齐精度。这表明，生物视觉皮层具有丰富的层次化表征，需要相当大的模型容量才能近似。具有渐进式通道扩展（$64 \to 128 \to 256 \to 512$）的 16 层层次化架构，能够学习从低级边缘到高级语义表征的多尺度视觉特征。残差连接促进梯度流和跨层特征复用，类似于生物视觉皮层中广泛的循环连接。GLU 门控机制可实现自适应特征路由，其中主 GLU 层能够学习根据特征的相关性对其进行选择性加权，这类似于生物视觉中的注意力机制：神经群体会表现出刺激依赖性选择性。

有趣的是，行为表现不一定能预测神经对齐。我们轻量级的 SimpleCNN 实现了更强的行为鲁棒性（95.4%，而大多数赛道 2 检查点为 89–91%），但在神经指标上的表现却明显落后（排名第 13，而其他模型排名第 1–6）。这种分离表明，任务最优表征与生物合理表征并不一致，至少对于这一特定导航任务而言是如此。生物视觉皮层的进化是为了支持多种行为需求，

其功能超越了简单导航，可能需要比完成任何单一任务严格所需的更丰富表征。

6.3 训练与性能之间的非单调关系

我们对赛道 2 中从 6 万到 114 万训练步数的 10 个检查点进行了系统分析，发现训练时长与行为性能和神经性能之间存在显著的非单调关系。训练至 20 万步的模型 368751 排名第 2，表现几乎与训练时长多出 5.7× 的顶尖模型不相上下。训练至 8 万步的模型 368628 优于训练至 52 万步的模型 369015，尽管前者的训练量少了 6.5× 。这些模式表明，即使衡量的是神经对齐而非行为指标，延长训练也可能损害泛化能力。
有几种机制可能解释这种非单调性。首先，持续训练可能导致模型对训练分布中的特定视觉刺激过拟合，从而降低其对评估所用的更广泛刺激集中的神经反应的对齐程度。其次，随着策略不断改进并接触不同的状态分布，灾难性遗忘可能会削弱先前学习到的特征。第三，行为优化（最大化奖励）与神经对齐（匹配生物表征）之间的关系可能并非单调；在任务特定优化占据主导之前，最佳对齐状态可能出现在训练的中间阶段。

这些发现具有重要的实践意义。与其训练至收敛后提交最终检查点，实践者不如评估训练不同阶段的多个检查点，并根据验证性能进行选择。在进行架构搜索或超参数优化时，与其在 1M 步后再寻找合适的检查点，在 200K 步时就找到它们所节省的计算资源（5× 减少）相当可观。

6.4 局限性与未来方向

我们的方法存在一些局限性，也由此指明了未来的研究方向。首先，任务特异性仍值得关注——研究结果未必能推广到更复杂的视觉导航场景，例如视觉场景更丰富、规划时域更长或扰动类型更多样的场景。在其他视觉控制任务上评估我们的架构原则，有助于确定其更广泛的适用性。其次，我们的神经对齐评估依赖线性读出和表征相似性指标，而这些方法只能捕捉生物视觉的某些方面。纳入动态、注意力或因果结构的更精细对齐指标，或可为生物学合理性提供更深入的见解。

第三，我们尚未充分探索各种可能的架构增强方向。采用自注意力机制的 Transformer 架构是一条值得探索的路径，有望兼顾注意力带来的鲁棒性优势与架构简洁性。视觉 Transformer在监督学习中已展现出良好前景，其在鲁棒视觉强化学习中的应用也值得研究。第四，我们的训练流程采用标准 PPO，未使用好奇心驱动探索、后见之明经验回放或元学习等复杂技术。这些方法或许能进一步提升鲁棒性和神经对齐水平。

最后，竞赛环境仅提供自我中心的视觉观察，而小鼠则运用结合视觉、本体感觉和前庭信息的多感官整合。扩展我们的架构以处理多模态感官输入，同时保持鲁棒性，是开发真正受生物启发的智能体的一项重要未来挑战。

7 结论

我们介绍了在 NeurIPS 2025 Mouse vs. AI：鲁棒视觉觅食竞赛两个赛道中的获胜方案，证明通过有针对性的增强，在保持架构简洁性的同时能够实现卓越的视觉鲁棒性，而深层层次架构则能实现最佳神经对齐。我们的赛道 1 解决方案由一个双层卷积神经网络、门控线性单元和观测归一化组成，尽管参数量比深层替代方案少 12.8× ，仍取得了 95.4% 的最终得分。我们的赛道2解决方案采用包含残差块和基于 GLU 的门控的 16 层 ResNet 架构，以 1780 万个参数实现了顶尖的神经预测性能。

通过全面实验和系统性的消融研究，我们记录了成功的组件和失败的方法。观测归一化对鲁棒性的提升最大（3.8 个百分点），而 GLU 机制也带来了稳定但较小的提升（0.2 个百分点）。包括 InceptionNet、深度残差网络、LSTM 变体和朴素数据增强在内的复杂架构均未能胜过简单基线，这凸显了选择适合任务的架构的重要性。我们对训练动态的分析揭示，训练时长与性能之间存在非单调关系；最优检查点约出现在 200K 步，而非训练收敛时。

这些发现挑战了关于视觉鲁棒性和生物对齐所需架构的传统观点。简单的架构有助于防止过拟合，从而提升视觉鲁棒性；而神经对齐则需要足够的模型容量，以捕捉层次化生物表征。行为上最优的表征与生物合理表征之间的分歧，揭示了这两类目标之间关系中仍有许多重要问题亟待探究。我们希望，对架构选择、训练流程和失效模式所做的全面记录，能为开发鲁棒且受生物学启发的视觉智能体的研究人员提供有价值的指导。

致谢

我们感谢 Mouse vs. AI 竞赛组织者，包括 Marius Schneider、Joe Canzano、Jing Peng、Yuchen Hou、Spencer LaVere Smith 和 Michael Beyeler，感谢他们创建了这一连接强化学习与神经科学的宝贵基准。我们衷心感谢 VNU-HCM理工大学提供计算资源。我们也感谢广大的 ML-Agents 社区开发出优秀的工具，使这项研究得以开展。

参考文献

[1] Yann N. Dauphin、Angela Fan、Michael Auli 和 David Grangier。使用门控卷积网络进行语言建模。载于 Doina Precup 和 Yee Whye Teh 编，《第*34*届国际机器学习会议论文集》，《机器学习研究论文集》第70卷，第933–941页。PMLR，2017年8月6日至11日。网址：https://proceedings. mlr.press/v70/dauphin17a.html。[2] Lasse Espeholt、Hubert Soyer、Remi Munos、Karen Simonyan、Volodymir Mnih、Tom Ward、Yotam Doron、Vlad Firoiu、Tim Harley、Iain Dunning、Shane Legg 和 Koray Kavukcuoglu。IMPALA：基于重要性加权的智能体-学习器架构实现可扩展的分布式深度强化学习，2018年。网址：https://arxiv.org/abs/1802.01561。[3] Kaiming He、Xiangyu Zhang、Shaoqing Ren 和 Jian Sun。用于图像识别的深度残差学习。载于*2016*年*IEEE*计算机视觉与模式识别会议（*CVPR*），第770–778页，2016年。doi：10.1109/CVPR.2016.90。[4] Andrew D. Huberman、Marla B. Feller 和 Barbara Chapman。视觉地图和感受野发育的潜在机制。《神经科学年度评论》，31：479–509，2008年。ISSN 0147-006X。doi：10.1146/annurev.neuro.31.060407.125533。[5] Arthur Juliani、Vincent-Pierre Berges、Ervin Teng、Andrew Cohen、Jonathan Harper、Chris Elion、Chris Goy、Yuan Gao、Hunter Henry、Marwan Mattar 和 Danny Lange。Unity：面向智能体的通用平台。*arXiv*预印本 *arXiv:1809.02627*，2018年。[6] Michael Laskin、Kimin Lee、Adam Stooke、Lerrel Pinto、Pieter Abbeel 和 Aravind Srinivas。使用增强数据进行强化学习，2020年。网址：https://arxiv.org/abs/2004. 14990。[7] Sergey Levine、Peter Pastor、Alex Krizhevsky、Julian Ibarz 和 Deirdre Quillen。利用深度学习和大规模数据采集学习机器人抓取中的手眼协调。《国际机器人研究杂志》，37(4-5)：421–436，2018年。doi：10.1177/0278364917710318。网址：https://doi.org/10.1177/0278364917710318。[8] Volodymyr Mnih、Koray Kavukcuoglu、David Silver、Andrei A. Rusu、Joel Veness、Marc G. Bellemare、Alex Graves、Martin Riedmiller、Andreas K. Fidjeland、Georg Ostrovski、Stig Pe- tersen、Charles Beattie、Amir Sadik、Ioannis Antonoglou、Helen King、Dharshan Kumaran、Daan

Wierstra、Shane Legg 和 Demis Hassabis。通过深度强化学习实现人类水平控制。《自然》，518(7540)：529–533，2015年。ISSN 1476-4687。doi：10.1038/nature14236。网址：https://doi.org/10.1038/nature14236。[9] Marius Schneider、Joe Canzano、Jing Peng、Yuchen Hou、Spencer LaVere Smith 和 Michael Beyeler。Mouse vs. AI：用于评估视觉鲁棒性和神经对齐的神经行为学基准，2025年。网址：https://arxiv.org/abs/2509.14446。[10] John Schulman、Filip Wolski、Prafulla Dhariwal、Alec Radford 和 Oleg Klimov。近端策略优化算法，2017年。网址：https://arxiv.org/abs/1707.06347。[11] Noam Shazeer。GLU变体改进Transformer，2020年。网址：https://arxiv.org/abs/ 2002.05202。[12] Christian Szegedy、Wei Liu、Yangqing Jia、Pierre Sermanet、Scott Reed、Dragomir Anguelov、Dumitru Erhan、Vincent Vanhoucke 和 Andrew Rabinovich。深入卷积网络。载于*2015*年*IEEE*计算机视觉与模式识别会议（*CVPR*），第1–9页，2015年。doi：10.1109/CVPR.2015.7298594。[13] Josh Tobin、Rachel Fong、Alex Ray、Jonas Schneider、Wojciech Zaremba 和 Pieter Abbeel。通过领域随机化将深度神经网络从仿真迁移到现实世界。载于*2017*年*IEEE/RSJ*智能机器人与系统国际会议（*IROS*），第23–30页，2017年。doi：10.1109/IROS.2017.8202133。[14] Ashish Vaswani、Noam Shazeer、Niki Parmar、Jakob Uszkoreit、Llion Jones、Aidan N Gomez、Ł ukasz Kaiser 和 Illia Polosukhin。注意力即一切。载于 I. Guyon、U. Von Luxburg、S. Bengio、H. Wallach、R. Fergus、S. Vishwanathan 和 R. Garnett 编，《神经信息处理系统进展》第30卷。Curran Associates, Inc.，2017年。网址：https://proceedings.neurips.cc/paper_files/paper/2017/file/ 3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf。[15] Fan Yang、Per Frivik、David Hoeller、Chen Wang、Cesar Cadena 和 Marco Hutter。通过端到端强化学习实现远距离无地图导航的空间增强型循环记忆，2025年。网址：https://arxiv.org/abs/2506.05997。[16] Anthony M. Zador。对纯粹学习的批判，以及人工神经网络能从动物大脑中学到什么。《自然·通讯》，10(1)：3770，2019年。ISSN 2041-1723。doi：10.1038/s41467-019-11786-6。网址：https://doi.org/10.1038/s41467-019-11786-6。

详细训练配置

本附录提供全面的训练配置和额外的实验细节，以便于复现。

A.1 超参数配置

表 4 展示了用于训练赛道 1 和赛道 2 模型的完整 PPO 超参数。

表 4：两个赛道的完整 PPO 训练超参数。赛道 1 使用更高的学习率和更大的缓冲区，而赛道 2 则需要较低的学习率，以稳定训练更深的架构。

超参数	赛道 1（SimpleCNN）	赛道 2（深度残差网络+门控线性单元）
近端策略优化算法设置		
学习率	9×10^{-5}	5×10^{-6} 至 9×10^{-6}
学习率调度	线性衰减至 0	线性衰减至 0
批次大小	128	128
缓冲区大小	4,096	1,024
训练轮数	3	3
熵系数（β）	0.005	0.005
裁剪参数（ϵ）	0.2	0.2
广义优势估计（GAE）系数（λ）	0.95	0.95
折扣因子（γ）	0.99	0.99
网络架构		
隐藏单元	256	256
层数（全连接）	1	1
视觉编码器	SimpleCNN（2 个卷积层）	深度残差网络（16 个卷积层）
总参数量	1.4M	17.8M
训练日程		
第一阶段步数	1,400,000	N/A
第二阶段步数	350,000	N/A
总训练步数	1,750,000	60K 至 1.14M
检查点间隔	20,000	20,000
摘要频率	1,000	1,000
环境设置		
时间跨度	64	64
每回合最大步数	变量	变量
并行环境数量	8-16	8-16
观测空间	$86 \times 155 \times 1$	$86 \times 155 \times 1$
动作空间	连续（3D）	连续（3D）

A.2 架构规格

A.2.1 赛道 1：SimpleCNN +GLU 架构

表 5 列出了我们赛道 1 架构按层划分的完整规格。

A.2.2 赛道 2：深度 ResNet +GLU 架构

表 6 总结了赛道 2 所采用的深层层次架构。

A.3 赛道 2 检查点性能分析

表 7 对全部十个赛道 2 模型检查点进行了全面的性能分析，并纳入了可用的行为指标。

表 5：赛道 1 SimpleCNN + 门控线性单元模型的详细架构规格。输入维度表示为（高度、宽度、通道）。各层之后的输出维度根据标准卷积公式计算。

层	Type	输出形状	卷积核/单元	参数
输入	-	(86,155, 1)	-	0
归一化	运行统计量	(86,155, 1)	-	0
卷积层1	二维卷积	(20,38, 16)	8×8，步幅 4	1,040
激活层1	带泄漏的ReLU	(20,38, 16)	$\alpha = 0.2$	0
卷积层2	二维卷积	(9，18， 32)	4×4，步幅 2	8,224
激活函数2	带泄漏的ReLU	(9，18， 32)	$\alpha = 0.2$	0
展平	-	(5、184)	-	0
FC1	线性	(256)	-	1,327,360
门控线性单元 特征_	线性	(256)	-	65,792
门控线性单元 Swish_	Swish	(256)	-	0
门控线性单元 门控_	线性	(256)	-	65,792
门控线性单元_Sigmoid 函数	Sigmoid 函数	(256)	-	0
门控线性单元_相乘	逐元素	(256)	-	0
门控线性单元_输出	线性	(256)	-	65,792
策略头	线性	(3)	-	771
价值头	线性	(1)	-	257
总计				**1,395,256**

检查点分析的关键观察：

• 训练时长的**最佳区间**似乎在约 200K 步（8 小时），此时可达到最佳性能的 99.34%，同时节省 5.7× 的计算资源。
• 模型 368628 仅训练了 8 万步（4 小时），便显著达到第 5 名，表现优于训练时间更长的模型6。5×
• 训练超过 20 万至 30 万步后，收益递减，有时性能还会下降（例如，模型 369015 在训练了 52 万步后）。
• 这种非单调关系表明，基于验证集神经对齐结果提前停止训练，可能显著提高计算效率。

A.4 计算资源与训练基础设施

所有实验均使用以下计算配置进行：

- **硬件：** NVIDIA GPU（RTX 3080/3090 或同等型号）
- **框架：** Unity ML-Agents Release 19 与 PyTorch 1.13+
- **训练环境：** 8–16 个并行 Unity 环境实例
- **总计算时间：**赛道 1 约需 150 GPU 小时， 200+ 赛道 2 GPU 小时（包括所有检查点）
- **检查点存储：**每 20K 步保存一次模型，每个检查点需要 ∼200MB

A.5 失败方法详情

表 8 总结了未能超越简单基线的架构的量化性能，为我们的设计决策提供了背景信息。

表 6：赛道 2 深度残差网络 + GLU模型的架构规范，共包含 16 个卷积层。该架构每 2 层采用一次残差连接，并逐步进行通道扩展（64→128→256→512）。

组件	层	输出形状	详细信息	参数
输入	-	(86, 155, 1)	灰度	0
初始卷积	二维卷积	(21,38,64)	4×4，步幅 4	1,088
阶段 1：残差块 1-2	2×(卷积+修正线性单元)	(21,38,64)	3×3，步幅 1	∼74K
阶段 1：残差块 3-4	2×(卷积+修正线性单元)	(21,38,64)	3×3，步幅 1	∼74K
下采样	二维卷积	(10,19,128)	2×2，步幅 2	32,896
第2阶段：残差块5-6	2×(卷积+修正线性单元)	(10,19,128)	3×3，步幅 1	∼295K
下采样	二维卷积	(5, 9, 256)	2×2，步幅 2	131,328
阶段 3：残差块 7-10	4×(卷积+修正线性单元)	(5,9, 256)	3×3，步幅 1	∼2.36M
下采样	二维卷积	(2, 4, 512)	2×2，步幅 2	524,800
阶段 4：残差块 11-12	2×(卷积+修正线性单元)	(2,4, 512)	3×3，步幅 1	∼4.72M
展平	-	(4096)	空间展平	0
GLU门控层	门控	(256)	自适应路由	∼1.05M
稠密投影	线性	（256）	特征投影	∼1.05M
策略头	线性	(3)	连续动作	771
价值头	线性	(1)	状态价值	257
总计	**16 卷积 + 门控线性单元**			**17,870,417**

表 7：赛道 2 的综合检查点分析，展示训练时长、神经对齐分数与模型性能之间的关系。所有检查点均采用相同的架构（深度残差网络 + 门控线性单元，1780 万个参数），但训练时长各不相同。

检查点 ID	训练步数	Rank	神经得分	训练时间（小时）	相对性能
368923	1,139,998	1	0.1517	24	基线（100%）
368751	199,996	2	0.1507	8	99.34%（快 5.7×）
368551	219,977	3	0.1497	9	98.68%（快 5.2×）
368578	319,976	4	0.1496	12	98.62%（快3.6×倍）
368628	79,998	5	0.1495	4	98.55%（快14.3×倍）
368967	199,994	6	0.1470	8	96.90%
368507	375,323	6	0.1470	14	96.90%
369015	519,976	6	0.1470	18	96.90%
368622	59,975	11	0.1467	3	96.70%（快 19.0× 倍）
359198	500,027	13	0.1451	16	95.65%（小型卷积神经网络）

表 8：失败的架构方法的性能总结。这些实验让我们决定采用更简单的架构，并进行有针对性的增强。

架构	参数	ASR (%)	MSR (%)	最终结果 (%)
IMPALA ResNet（24 层）	∼8M	80.96	51.00	65.98
IMPALA ResNet（4 层）	∼3M	91.40	84.00	87.70
ResNet-4 + 数据增强	∼3M	72.60	47.00	59.80
InceptionNet	∼5M		未能收敛	
SimpleCNN + LSTM	∼2.5M	92.80	87.00	89.90
SimpleCNN + SRU-LSTM	∼3M	91.50	86.50	89.00
SimpleCNN + 门控线性单元 + 归一化（我们的方法）	**1.4M**	**96.80**	**94.00**	**95.40**