

学术硕士研究生

学位论文

**猕猴V4双光子成像数据的深度学习模型研究**

**作 者：林泽瀚**

**导 师：吕海东**

**培养单位：心理学部**

**学 号：202121061031**

**学科专业：心理学**

**完成日期：2024年5月**

**北京师范大学学位论文原创性声明**

本人郑重声明： 所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

学位论文作者签名：　　　　　　　　　　日期： 年 月 日

**学位论文使用授权书**

学位论文作者完全了解北京师范大学有关保留和使用学位论文的规定，即：学校有权保留并向国家有关部门或机构送交论文的复印件和电子版，允许学位论文被查阅和借阅；学校可以公布学位论文的全部或部分内容，可以允许采用影印、缩印或其它复制手段保存、汇编学位论文。保密的学位论文在解密后适用于本授权书。

本人签名： 日期：

导师签名： 日期：

猕猴V4双光子成像数据的深度学习模型研究

摘 要(小三号黑体，居中)

（小四号宋体）段落按照“首行缩进”格式，每段开头左缩进2个汉字符，标点符号占一格，行距20磅。

关键词**：**腹侧通路, 深度学习, 人工神经网络，猕猴视皮层V4区, 双光子钙成像

英文题目(三号，全部大写字母，居中)

**ABSTRACT**(四号加粗，居中)

**KEY WORDS:** Ventral pathway, Deep learning, Artificial neural network, Area V4 in macaque, Two-photon calcium imaging

目 录（小二号，黑体，居中）

[第1章 研究背景 1](#_Toc151121245)

[1.1 视觉并行通路理论 1](#_Toc151121246)

[1.1.1 视觉并行通路理论的发展 1](#_Toc151121247)

[1.2 V4区的位置及主要功能区 2](#_Toc151121248)

[1.3 神经网络与视觉系统的联系 2](#_Toc151121249)

[1.4 科学问题与研究思路 2](#_Toc151121250)

[第2章 实验操作步骤 3](#_Toc151121251)

[2.1 实验材料 3](#_Toc151121252)

[2.2 动物手术 3](#_Toc151121253)

[2.3 麻醉猴内源光学成像 3](#_Toc151121254)

[2.3.1 内源光学成像原理 3](#_Toc151121255)

[2.3.2 内源信号采集 4](#_Toc151121256)

[2.3.3 视觉刺激-内源信号光学成像 4](#_Toc151121257)

[2.4 麻醉猴双光子成像 5](#_Toc151121258)

[2.4.1 双光子成像原理 5](#_Toc151121259)

[2.4.2 双光子信号采集 5](#_Toc151121260)

[2.4.3 视觉刺激-双光子成像 6](#_Toc151121261)

[第3章 数据分析方法 7](#_Toc151121262)

[3.1 OI数据分析-SVM map 7](#_Toc151121263)

[3.2 双光子成像数据处理 7](#_Toc151121264)

[3.3 深度学习模型构建 8](#_Toc151121265)

[3.3.1 神经元DCNN编码模型 8](#_Toc151121266)

[3.3.2 AlexNet预训练卷积层的选择 8](#_Toc151121267)

[3.3.3 单位点与多位点模型 8](#_Toc151121268)

[3.3.4 模型评估与训练 8](#_Toc151121269)

[3.3.5 偏好图像生成 8](#_Toc151121270)

[3.3.6 重复次数增量实验 8](#_Toc151121271)

[3.3.7 卷积核权重评估 8](#_Toc151121272)

[第4章 实验结果 10](#_Toc151121273)

[4.1 刺激图像的修改 10](#_Toc151121274)

[4.2 不同数据集的比较 10](#_Toc151121275)

[4.3 卷积核的设计 10](#_Toc151121276)

[第5章 讨论 11](#_Toc151121277)

[5.1 11](#_Toc151121278)

[参考文献 13](#_Toc151121279)

[附 录 14](#_Toc151121280)

[攻读学位期间取得的学术成果 15](#_Toc151121281)

[致 谢 16](#_Toc151121282)

图 目 录

**未找到图形项目表。**

表 目 录

未找到图形项目表。

# 研究背景

(黑体三号，居中，章序与标题之间空1个汉字符）

## 视觉并行通路理论

哺乳类动物的神经系统由无数精密相连的神经网络构成，其中大脑扮演着系统的指挥官角色，掌管着生命体的全部活动，并管理着与外部世界的信息交换。在漫长的演化旅程中，人类及其他非人类灵长动物的大脑架构演变成了一座错综复杂的神经迷宫。20世纪初，Brodmann基于大脑皮层的细胞构造，将其精心划分为43个区域，这一数字在后续研究中已经增至52个。这些划分不仅定义了大脑皮层的地图，而且与大脑的功能紧密相连，每个区域都承担着特定的角色。

对于人类以及其他高等灵长类动物来说，视觉系统是接受外界信息的主要通道。超过七成的感官信息是通过视觉得到的，这些信息涉及物体的亮度、形状、颜色、运动（方向与速度），以及立体视觉，即深度感知。相关的大脑区域主要集中于枕叶、颞叶和顶叶。过去数十年的科研揭示了视觉信息并行处理路径的理论，这一理论深刻总结了视觉信息在大脑各部位的传递流程，并对理解视觉处理的生物学机制提供了极为重要的理论支撑。

1982年，Mishkin和Ungerleider 根据行为学，电生理，形态学和脑损伤研究结果，定位了视觉皮层的并行通路：一条位于大脑腹侧，被称为腹侧视觉通路（简称腹侧通路，ventral visual pathway），主要包括纹状皮层（striate cortex，又称为初级视觉皮层V1），纹外皮层（extrastriate cortex 包括V2、V3、V4），延伸至下颞叶区域（TEO和TE），主要负责物体视觉（object vision），这些区域与腹侧边缘系统和额叶的链接可能是物体信息与情绪，动机等认知功能联合的结构基础；另一条通路位于大脑背侧，称为背侧视觉通路（dorsal visual pathway），从纹状皮层，纹前皮层至下顶叶区域，主要负责空间视觉（spatial vision），用于定位特定的物体，指导相关的行为。同年年末，根据过去十几年非人灵长类动物的电生理和形态学研究结果，Van Essen 和 Maunsell 对视觉皮层并行通路进行了更精细地描述：腹侧通路主要涉及脑区为V1，V4，VP和IT，背侧通路为V1，MT和MST，并区分了腹背侧通路主要加工的视觉信息，认为腹侧通路主要涉及加工颜色和形状信息，背侧主要涉及加工运动信息。

纵观这些研究，不同时间和方式的研究发现均支持以下理论即视皮层不同脑区接受信息的来源有所不同，传递信息的去向也不尽相同。正是这种神经信号处理通路的不同为处理不同的视觉信息提供了必备的神经基础，让视觉信息在视皮层中的分类和加工更加高校，同时也使视皮层中的信息交流更加准确。在本研究中，我们重点关注腹侧视觉通路中的V1和V4脑区，并尝试使用神经网络模型来拟合V1、V4神经元的反应。

## V4区的位置及主要功能区

### V4区的定位

在视觉科学领域，视皮层四区（V4区）最初被认定为颜色处理的关键脑区。早期的研究，如Semir Zeki等人的工作，通过损伤实验确定了位于猕猴脑内月状沟前和上颞叶沟之间的一个独立脑区，并将其命名为V4区。后续研究通过单细胞记录技术揭示，V4区的神经元对特定波长的视觉刺激表现出响应，表明这些神经元具有颜色选择性。更为有趣的是，Zeki发现这些具有颜色选择性的神经元按功能柱状结构分布。

随着研究的深入，科学家们在V4区发现了对朝向和形状有选择性反应的神经元。因此，位于视觉处理腹侧通路中的V4区，现在被普遍认为是处理物体颜色和形状信息的重要脑区。

关于V4区的确切定位和结构细节，科学界依然存在一些争议。然而，关于V4区的大致位置和分区已达成较广泛的共识。V4区主要位于脑的腹侧部位，处于月状沟（lunate）与颞上沟（STS）之间，后缘与V2区相连，前缘则延伸至内颞叶区域IT。V4区的一部分隐藏于月状沟和颞上沟的沟壑中，因此在大脑皮层表面并不容易观察到完整的V4区。

### V4的功能和分区

V4区位于视觉信息传递路径的中枢地带，扮演着将初级视觉皮层（V1区）的输入与内颞叶区域（IT区）的高级视觉处理连接起来的关键角色。在视觉信息从简单到复杂的转化过程中，V4不仅仅是一个信息的中转站，而是在这一连串复杂且细致的处理过程中，担任分工与整合的复杂任务。这种高效的神经机制为大脑的高级区域处理极其复杂的视觉刺激奠定了基础，使我们能在丰富多变的环境中形成接近实际情景的认知。

研究揭示了V4区内部神经元的多样性：不仅有对特定颜色和形状敏感的选择性神经元，还有对复杂图案和运动趋势作出反应的神经元。这种多样性让我们能迅速而准确地在自然环境中辨识物体。特别地，V4区的颜色选择性神经元机制支撑了我们在五彩斑斓的世界中维持颜色恒定性的能力。此外，V4区的深度信息处理能力，使我们能够感知物体的立体结构。

来自我们实验室的最新研究进一步拓宽了对V4区功能的认识，发现这一区域内还存在对颜色和运动方向敏感的选择性神经元。这表明，V4区不仅参与颜色、形状、深度、朝向的处理，还涉及运动方向的识别。V4区内不同种类神经元的独特属性共同构成了一个复杂且精细的视觉信息处理网络。

## 神经网络模型与视觉系统的联系

### 神经网络模型的起源

20世纪中叶，Hubel等人[1]在猫的初级视觉皮层（V1）中，发现了两种主要的细胞类型——简单细胞和复杂细胞。其中，简单细胞会对在某一个特殊空间位置的亮暗光栅有反应，每个细胞对于光栅的朝向有选择性，并且随着逐渐偏离偏好方向反应值降低（调谐曲线）。而对复杂细胞来说，没有那么多严格的反应特性，它们仍然拥有朝向选择性，但是与其他方向的反应差别不大。1980年，Fukushima 提出了Neocognitron模型[2]，是现代CNN的先驱。它包含两种主要的细胞类型。S细胞模拟的是简单细胞，重复出它的特性，二维的权重被分配到输入图片的每个位置。S细胞的平面具有视野的布局，所有细胞共享相同的偏好视觉特征，并且在一个层中存在多个平面。C细胞模拟的是复杂细胞，它得到将同一平面不同位置的简单细胞经过非线性运算的结果。一层的简单细胞和复杂细胞代表V1的基础运算，Neocognitron 仅仅是重复上面的过程。因此，第一层的复杂细胞作为第二层简单细胞的输入。到了19世纪90年代，许多类似的层级模型被提出，其中最突出的是HMAX，它对简单细胞集合的反应值进行使用简单的“max”操作确定复杂细胞的反应，在图像多变时表现优异。由于有图片对应的人类实验数据，模型的结果能够直接与人类快速视觉分类能力对比，类似的模型可以被用来拟合V4神经元对复杂形状刺激的反应[3]。

在2014年的一项研究中，研究人员记录了猕猴在观看复杂物体图片时的细胞活动，并对真实的V4和IT神经元的反应与人工神经网络的反应进行了回归分析。研究发现，那些在物体识别任务中表现更佳的神经网络，往往能够更准确地预测神经反应[4]，相似的结果也在视频分类任务中观察到[5] 。文章提到，作者构建的人工神经网络中的倒数第一层在预测IT神经元活动方面表现最优，而倒数第二层则更适合预测V4神经元的活动。这些发现在一定程度上表明，人工神经网络拥有与视觉系统腹侧通路类似的层次结构与处理特性。

### 神经元DCNN模型

在2019年，Dicarlo等研究人员[6] 通过向猕猴的V4区域植入一个由96个微型电极组成的阵列，成功采集了猴子神经元对图片刺激的放电反应数据。基于这些数据，他们设计了一个人工神经网络模型，这个模型不仅能够精准预测神经元的放电模式，还能够根据这些预测生成专门的控制图片，这些控制图片的目的是为了最大化特定神经元的放电率。

在合成这些控制图像时，研究团队采用了两种方法：一种是“拉伸”法（stretch control），该方法致力于最大程度地增加目标神经元的活动，尽管在此过程中其他神经元的活动也可能会有所提高。另一种是“独热群”法（one–hot–population control），这种方法在增强目标神经元响应的同时，还会刻意抑制周围其他神经元的活动。

虽然这些生成的刺激图像在外观上可能不够自然，也不直接关联到网络训练时使用的图像，但它们在实际应用中却能有效促使神经元活动超过常规水平。这一结果进一步验证了人工神经网络在捕捉视觉处理流程中基础构成方面的有效性，突显了其在模仿生物视觉系统方面的潜力。

前人的研究认为V4脑区是视觉物体识别中的一个关键环节，其神经元表现出对各种刺激特征（如朝向、颜色、弧度、空间频率（SF）、纹理以及亮度）的多维度选择性。然而，究竟哪些特征是V4神经元的主要处理对象，以及这些特征与V4脑区已知的功能结构（包括朝向、颜色、曲线和运动）之间的具体关联，目前仍然没有很明确的结论。

### 组里的相关工作

在前人研究的基础上[6]，我们采用了双光子钙信号成像技术，记录了522个分布于不同功能区的V4神经元对超过800张自然图像和人工刺激的响应。通过深度卷积神经网络（DCNN）对这些V4神经元的活动进行建模，我们对模型中的神经元感受野进行了深入分析。分析结果揭示出曲率、方向和颜色是构成神经元群体反应的三个主要维度，并且这些特征与神经元所处的功能区域高度一致。对神经元反应进行的降维分析进一步支持了这一发现。

该研究确认了颜色、方向和圆弧形状是V4神经元的关键反应特性，而这些特性恰好也是神经元所在位置的主要功能特征。这表明V4区域的功能架构反映了其主要处理的视觉信息特征。同时，我们的工作验证了DCNN模型在模拟V4神经元的钙信号成像结果方面的适用性，为未来深入探索V4区域的视觉特性奠定了坚实的基础。

## 科学问题与研究思路

在上一阶段的研究中，我们构建了一个深度学习模型，用于处理V4区双光子成像数据。该模型经过适量双光子数据训练后，能够准确拟合神经元活动，并生成针对特定神经元的优化刺激。本研究旨在基于现有模型进行深入探索，挖掘其潜在应用，并提高模型的拟合精度。

**问题1：DCNN神经元模型是否适用于双光记录的神经元反应？**

Dicarlo等研究人员[6] 在猕猴电生理学数据上已经取得了卓越的成果。但是，是否可以将DCNN神经元模型应用于双光子记录的神经元活动数据，并取得同样的效果呢？

双光子钙成像技术的一个显著优点是能够在不损伤细胞的情况下，进行长期和重复的实验，从而获取更加可靠的神经元活动数据。那么，双光子钙成像技术的先进性是否能够为我们提供更加丰富的结果？此外，我们亦需评估双光子钙成像中trial相关性对结果的具体影响程度。

**研究思路：**

在数据采集阶段，我们计划执行多次实验，从多个trials中收集数据，并将这些数据平均化以构建数据集。我们将采用随机抽样方法来评估增加trials数量对模型性能的影响。

**问题2：现有的模型有哪些问题，怎样提高？**

与Dicarlo[6]团队在猕猴电生理数据上的成果相比，我们观察到模型在双光子数据上的精度较低。这引发了关于模型当前局限性的思考，并促使我们审视实验方案的设计合理性。如何才能进一步提升模型的精度呢？

**研究思路：**

考虑到感受野的特性，我们将尝试改变刺激的呈现方式，比如裁剪、置灰等，以探索最适合模型训练的刺激形态。

对比同一组神经元在不同刺激集下的反应数据。通过分析两组数据的差异，我们希望能够更深入地理解模型的工作原理，引导后续实验的设计，从而探究提高模型精度的可能方案。

# 实验操作步骤

## 实验材料

在本次研究中，所有参与实验的动物均源于国内一流的猕猴繁育基地，并在北京师范大学的认知神经科学与学习国家重点实验室的专业动物设施内得到妥善饲养。在实验过程中，所有涉及动物的操作均严格遵守美国国立卫生研究院（NIH）制定的动物使用规范，并已获得北京师范大学实验动物管理和伦理委员会的正式审查与授权。本研究的数据收集涉及到3只动物，包括2只恒河猴和1只食蟹猴，他们的3个大脑半球均参与了研究，我们在所有的大脑半球上进行了内源信号光学成像以及双光子钙信号成像实验。

## 动物手术

动物手术主要包括术前准备、手术过程和术后维护。

动物先用氯胺酮（10mg/kg）在猴笼进行初始麻醉。之后转移到手术室，动物头部固定于立体定位仪上，在呼吸机和异氟烷（1.5%-2.5%）麻醉。

对动物进行开颅手术，开颅窗口直径为22mm，窗口中心根据研究脑区（V1或V4）有所调整，距中缝18-35mm，距后ridge 15-25mm。打开硬脑膜，根据lunate, STS 和IOS的脑沟位置确定视皮层V1, V2和V4区。我们先对暴露视皮层进行内源信号光学成像，之后依据光学成像所得的功能图，对目标功能区进行定点病毒注射。工具病毒为AAV9.Syn.GCaMP6S.WPRE.SV40 (CS1282, titer 3.34e13 (GC/ml), Addgene)，注射深度为500μm, 每点注射量为500nL。病毒注射后密封颅骨窗口，并在6周后再次手术植入一个密封光学窗口（直径13mm）。之后所有的内源信号光学成像和双光子钙信号成像都在这密封的光学窗口上进行，每次成像实验间隔7-10天。

术后会定期对小室 (chamber) 进行清理维护，确保皮层的状态保持正常。

## 麻醉猴内源光学成像

### 内源光学成像原理

内源性光学成像技术是通过监测大脑皮层血氧水平的动态变化来揭示群体神经元对视觉刺激的响应机制。这种响应体现为神经活动的增强，随之带来的是局部脑组织的耗氧量上升，从而促使含氧血红蛋白向去氧血红蛋白转换，这一过程影响了血液对光的吸收与反射特性。在实验设置中，我们选择了波长为630纳米的红光作为照射源，这是因为去氧血红蛋白对此波长的红光吸收率较高，而反射率较低。随后，利用高灵敏度的CCD成像系统来精确记录由此产生的反射光强度变化，从而实时追踪和分析脑内血氧变化，为理解视觉刺激与神经元活动间的相互作用提供了一种手段。

### 内源信号采集

在手术当天，为了进行内源性光学成像，我们采取了无成像窗的技术手段。显露出的皮层被玻璃片覆盖和平压，而玻璃片与皮层之间的空隙被琼脂填补，确保成像区域的稳定。随后，在成像窗植入之后，便可进行直接成像。在成像过程中，动物起初采用异氟烷吸入麻醉，随后转为静脉推注麻醉剂（第一例实验中是丙泊酚与舒芬太尼的混合，第二和第三例实验则使用了氯胺酮），同时维库溴铵的注射用以保持眼球位置的固定。动物的生命体征，包括心率、体温、血氧饱和度以及呼出气中二氧化碳水平被持续监测，确保麻醉水平的适宜，并据此调整麻醉剂的用量。

为了确保动物的瞳孔充分敞开，我们使用了0.5mg/ml的阿托品溶液，并通过精心选择的角膜接触镜来调整视线聚焦在57厘米远的屏幕上。利用632纳米波长的红光进行照射，皮层反射率的变化——即内源性血氧信号——被Optical Imaging公司的Imager 3001系统以4赫兹的频率捕获。图像的分辨率为540×654像素，这对应于一个15.5毫米×19毫米的成像区域。在每个实验周期中，视觉刺激持续3.5秒钟，血氧信号的记录从刺激前的0.5秒开始直到刺激结束，总共记录16帧图像，每次刺激之间的间隔时间大约为6秒。

### 视觉刺激-内源信号光学成像

在本次研究中，视觉刺激的产生由Cambridge Research Systems Ltd. 的ViSaGe系统控制，并通过位于动物正前方57厘米的21英寸CRT显示器（SONY CPD-G520型号）呈现给实验动物。这款显示器具备100Hz的刷新率，并且经过精确的gamma校正，以确保图像质量的准确性和稳定性。我们使用了动态的全屏正弦波光栅刺激来获得Orientation map和 Color map。这些光栅刺激设计有两个空间频率——0.25和1周期/度，并在8个不同的方向上移动（以45度为间隔，从0度至315度），所有刺激的时间频率均设为4周期/秒。无论是黑白光栅还是红绿光栅，它们的平均亮度都被设定为28.92坎德拉/平方米。

为了生成 Curvature map，我们采用了运动的全屏形状刺激，如圆形和三角形。这些形状刺激的直径均为2.5度视角，它们按照3度的步进间隔排列。刺激中的亮区亮度高达111.2坎德拉/平方米，其边缘宽度为0.2度，而背景亮度则保持在20.57坎德拉/平方米。在每一次试验中，这些形状刺激将在随机选定的一个方向上移动（从8个可能的方向中选择），运动速度统一为4度/秒。

通过这种方法，我们能够创造出一系列刺激，旨在精确测定和映射动物视觉皮层的反应，进而分析其对不同形状、颜色和运动方向的敏感度。

## 麻醉猴双光子成像

### 双光子成像原理

双光子激发技术是一种独特的成像方法，它依赖于两个低能量光子—通常是红外或近红外光—在极小体积内几乎同时与单个荧光分子相互作用。这一过程中，这两个光子共同提供足够的能量使荧光分子激发到较高能级，然后这个分子在返回基态时发射一个光子，产生荧光。这个精巧的物理现象使得科学家能够穿透更深的组织层，以实现对活体大脑内部结构更为精细的成像，进而在三维空间中探测细胞层面的活动。

双光子成像系统通过强度极高的飞秒激光脉冲、高数值孔径物镜以及敏感的光电倍增管（PMT）检测器相结合，实现了对生物组织的高分辨率成像。在此技术的帮助下，我们可以以精确的控制和优异的深度穿透能力，捕获立体图像，并洞察到不同深度的神经元活动。

在神经科学领域，钙离子的浓度变化被广泛地作为神经元活动的一个可靠指标。通过双光子钙成像，我们能够实时追踪神经元内的钙浓度变化。这项技术涉及先将细胞膜上的钙通道转染以发出绿色荧光，然后使用双光子显微镜发出的红外激发光线捕捉这些荧光信号。这样，我们可以通过观察荧光强度的变化，得到神经元活动强度的直观表示，从而在神经信号传递的复杂网络中，精准地描绘出活动图谱。

### 双光子信号采集

为了进行双光子显微成像实验，动物的准备和麻醉流程与内源信号光学成像的步骤保持一致。双光子显微镜使用的是Bruker Nano公司生产的Prairie Ultima IV型（In Vivo）双光子显微镜，而激光发射设备则是COHERENT品牌的Chameleon激光器。在这些成像会话中，激光的波长被设定在980纳米，功率范围介于30至45毫瓦。显微镜配备了Nikon品牌的16倍水浸物镜（数值孔径为0.8），使得对应的成像视野达到830微米乘以830微米的大小。

双光子扫描采取的是Galvo模式，以每秒1.3帧的速率进行，图像分辨率为512像素乘以512像素。为了同步视觉刺激与显微镜的成像过程，我们精确地将每次视觉刺激的开始与双光子扫描的下一帧启动时间对齐。

由于显微镜物镜的设计是垂直于样本的，因此在成像时，我们采用特制的旋转夹具将动物的头部旋转至水平位置，与显微镜镜头对齐，同时让动物的身体侧卧以维持其稳定。成像过程中，由于动物的呼吸和心跳，脑组织可能会发生轻微的位移。我们实时监控细胞的形态，并在必要时进行在线调整，以确保皮层在XY平面上的位移不超过50微米，在Z轴上的位移不超过150微米。

在本研究中，我们在3个半球的8个不同位置进行了详尽的双光子成像（如图Figure Sx所示）。成像深度范围在155至295微米之间，每个成像点只进行了一次特定深度的双光子成像。

### 视觉刺激-双光子成像

在本项研究中，视觉刺激是由ViSaGe系统产生，并通过位于试验动物正前方57厘米处的一台21英寸DELL E1913Sf LCD显示器进行展示。该显示器的刷新率设定为60Hz，经过精确的gamma校正，以保障视觉刺激的色彩和亮度准确传达，最大亮度调至80.82坎德拉/平方米。所有的视觉刺激都是在一个中性灰色背景上进行展示的，该背景亮度为40.41坎德拉/平方米。为了实验的需要，我们仅向动物的左眼展示视觉刺激，而右眼则保持闭合。

为了精确定位神经元的感受野，初始步骤中，我们通过手动移动一个圆形方波运动光栅（直径1到3度）并实时观察皮层的荧光响应来识别感受野的大致区域。确定感受野后，我们以此区域为中心，在一个5乘5的网格内，随机在其中一个格点上展示直径1到3度的视觉刺激。这些刺激包括三种形状——方波光栅、三角形和圆形，以及四个不同的运动方向（从45度至315度），且运动速度固定为每秒4度。每一刺激的呈现时间为1.5秒，在此期间会运动两次，其间隔为0.8秒。刺激展示完毕后，我们会立即对数据进行处理，以确定当前成像区域内细胞群的感受野位置和范围（采用高斯函数拟合，并以高斯函数高度的一半为阈值）。

在探究神经元对自然视觉刺激的响应方面，我们构建了一套包含800张彩色图片的自然客体刺激库，这些图片涵盖了450类不同物体，全部来源于www.freepngs.com网站。对于V4区域的5个位点，刺激库中还额外增加了50种人工形状刺激，所有刺激均未经亮度调整。在实验过程中，所有刺激的大小被统一调整到感受野直径的80%，并以随机顺序进行呈现。每个刺激的展示时间为2秒，每两次刺激之间间隔3.4秒。为了避免视觉适应现象，刺激的大小会在呈现前1秒内从感受野直径的80%逐渐增大至95%，然后在后1秒再缩小至80%。在一轮刺激展示完毕后，我们会重新测定感受野的位置，进而准备下一次的实验周期。通常情况下，一个实验会完成5个周期。在两次实验之后，我们会获得10个周期的数据，并将这些数据用于深度卷积神经网络（DCNN）模型的训练，以此来识别出各个神经元的偏好图像。在第三轮实验中，我们将一些神经元的偏好图像加入到刺激集中，以进一步验证模型的预测效果。

# 数据分析方法

## OI数据分析-SVM map

在我们的研究中，领域的性质与尺寸主要通过支持向量机（SVM）图来确定。SVM算法相较于传统的差异图像，更有效地抑制了血管引起的噪声和低频信号干扰，因此在提高信噪比和区分皮层不同功能区域方面表现出更加优越的性能（参考文献：Xiao et al., 2008）。所有的光学成像（OI）数据分析工作均在MATLAB 2017版本（The MathWorks, Natick, MA）中完成。

首先，我们计算每个试验周期中对每种刺激的内源性信号响应，即ΔR/R值。这里的ΔR/R定义为（第8到第16帧的平均反射率 - 第1到第4帧的平均反射率）除以第1到第4帧的平均反射率。接着，利用这些响应图，我们进一步生成SVM图。具体来说，通过SVM算法对两类刺激的响应图进行分类，每个像素点在分类中的权重反映了对应皮层区域在区分这两类刺激上的贡献程度。权重的大小指示了区分度的高低，而权重的正负值表示了该像素点对特定刺激的偏好。由此，我们可以将所有像素点的权重绘制成一幅图，揭示皮层的功能分布。用于此部分的Matlab SVM程序由Chih-Jen Lin提供（LIBLINEAR: A Library for Large Linear Classification, 2008; 可在https://www.csie.ntu.edu.tw/~cjlin/liblinear/获取）。

对于每一张生成的原始SVM图，我们首先应用高斯低通滤波（滤波器尺寸为5，标准差为1）进行平滑处理。然后，利用绿色血管图将成像区域划分为皮层区和背景区，其中背景区包括血管区和chamber外区域。随后，我们采用背景区灰度值分布的标准差对SVM图进行归一化处理（即计算z-score），并以±2倍标准差作为确定领域显著性的阈值（领域尺寸需大于50像素）。同时，我们还将各个双光子成像图像与血管图对齐，以此计算各细胞在各光学成像图中的z-score值。

## 双光子成像数据处理

我们的双光子成像数据分析同样在MATLAB 2017环境中完成。图像校准与细胞识别的方法继承自我们先前的研究。为了补偿成像过程中可能出现的皮层移动引起的图像模糊，我们将同一视场下的各个会话图像（包括跨日数据）都与一个统一的模板对齐，该模板是第一个会话前1000帧图像的平均值。这样的对齐通常能纠正小于20像素的位移。

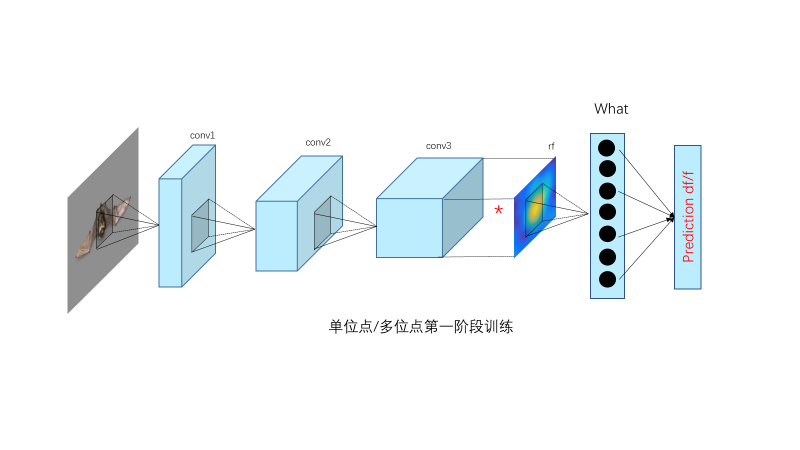
细胞的识别是基于其对视觉刺激的反应。具体操作中，我们首先将每个刺激的各次实验结果取平均，然后以刺激前两帧的平均图像作为基线，以刺激后第二、第三帧的平均图像作为反应图。通过从反应图中减去基线图，我们获得了用于细胞识别的差分图像。细胞是否被识别的标准是，在某个差分图像中至少有30个相连像素点的灰度值超过了整个图像灰度的平均值加上2.5倍的标准差。

一旦确定了细胞的位置，我们便计算该细胞对于各种视觉刺激的反应强度，即荧光变化率（ΔF/F0）。这里的ΔF表示的是刺激反应图中细胞灰度的变化值，F0则是基线图中细胞的灰度均值。如果一细胞对于所有实验刺激的平均反应强度均未超过0.3，那么这一细胞将不被纳入进一步分析中。细胞对各种视觉刺激的平均反应将被用来进行主成分分析（PCA）以及深度卷积神经网络（DCNN）模型的训练与验证。

## 深度学习模型构建

### 神经元DCNN编码模型

我们所采用的神经元编码模型是基于Bashivan等人在2019年提出的原型[6]。此模型结构包括AlexNet的部分预训练卷积层，一个代表感受野（RF）的“where”层，以及一个“what”层[7]，旨在构建从刺激图像到AlexNet特征空间的映射，最终映射到神经元的响应空间。与[6]中的原始模型相比，本研究中的刺激图像在送入模型之前并未进行鱼眼变换处理。模型的拟合过程分为两个阶段。在第一阶段，我们保持AlexNet的预训练卷积层参数不变，仅对“where”层和“what”层进行训练，直至模型达到收敛，这与文献[6]中的做法相似。在此基础上，我们引入了第二阶段的训练，允许AlexNet的卷积核参数参与调整，直到再次达到收敛。这一增加的微调阶段显著提高了编码模型的拟合度，EVE值提升了0.05至0.1。模型训练的损失函数由四个部分构成：第一部分是真实神经响应与预测响应间的均方误差​；第二部分是减少不必要边缘的空间滤波器拉普拉斯正则化项；第三部分是空间滤波器L2正则化项；第四部分是通道滤波器的L2正则化项。（公式待补）



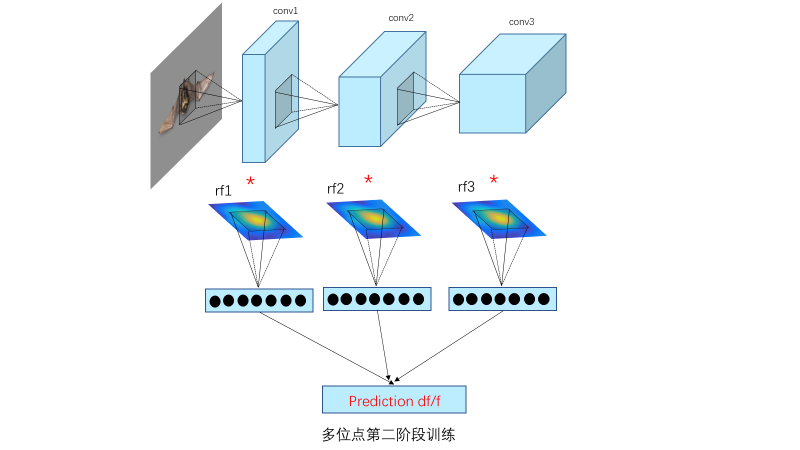


图1 神经元编码模型框架

### AlexNet预训练卷积层的选择

在本研究中，我们对V1和V4神经元的模型拟合进行了深入探讨，这涉及到为每个位点精选最适合的AlexNet卷积层。为了高效地挑选出对应于每个位点的最优卷积层，我们首先采用稀疏随机映射技术[8]来对AlexNet各卷积层的特征进行降维处理，确保降维后的维数符合Johnson-Lindenstrauss引理对样本量的要求。继而，我们利用降维后的特征在岭回归模型中预测神经元的响应，通过双重交叉验证方法为每个神经元选定最佳的正则化参数alpha。最终，我们根据各个位点所有神经元平均拟合精度的评估，挑选出表现最佳的卷积层，以此作为编码模型中AlexNet的特征输出层。在本研究中，V1区域前三个位点的最优AlexNet卷积层分别是第1、第2和第2层，值得注意的是，第三个位点位于V1和V2区域的交界。而V4区域的后五个位点，其最优卷积层均为第三层。这一筛选结果不仅映射了AlexNet与灵长类动物视觉腹侧通路之间的层级对应关系，而且与前人的研究结果[9]保持了一致性。

### 单位点与多位点模型

在本研究中，我们采用了单位点模型和多位点模型两种不同的方法来训练神经网络，以探究神经元在信息处理上的机制。

单位点模型专注于单个位点的神经元反应数据。在这种模型训练方法中，我们仅将同一位点的神经元反应数据输入模型进行训练。在此模型框架下，所有神经元共享相同的卷积层参数，而where和what层的参数则针对每个神经元独立优化。图4和图5展示的结果是基于这种单位点模型得出的。

为了更深入地比较V4区域内不同功能区神经元（通常分布于不同位点）在信息处理上的差异，我们设计并实施了多位点模型。在这个模型的训练过程中，所有神经元（包括V1和V4区的）被同时纳入训练。与单位点模型相同，不同位点对应的最优AlexNet卷积层也被保留。在多位点模型中，所有神经元共享卷积层参数，但where和what层的参数依然是针对每个神经元独立进行优化的。图6展示的结果是基于多位点模型得出的，我们发现多位点模型的总体预测精度略低于单位点模型（详见图Sx）。

总的来说，通过这两种模型的对比分析，我们可以更加深入地理解和揭示不同功能区神经元在信息处理上的独特机制和差异。

### 模型评估与训练

在本研究中，我们从800个自然视觉刺激及其对应的神经元响应中随机选取了720组作为训练集，而将余下的80个自然刺激以及额外的50个针对V4区域的人工刺激及其神经响应作为验证集。在模型的训练过程中，我们特别在验证集上对损失函数中的超参数进行了细致的网格搜索，其中的范围设定在[0.01, 0.1, 1]，和的范围同样设定在[0.01, 0.1, 1]。这一过程旨在寻找到能够实现最优拟合精度的超参数组合。

为了有效地进行模型优化，我们将学习率设定为1e-4，并选用了批处理大小为50的Adam优化器。为了防止模型过拟合，我们采取了早停策略：即如果在连续500次迭代中均方误差（MSE）的下降幅度持续小于5e-5，那么训练将被提前终止。

在量化模型拟合精度方面，我们采用了可解释方差的解释比例（EVE，Explainable Variance Explained）作为指标。EVE计算公式为：



其中，表示真实神经响应值与模型预测值之间的均方误差，是神经元在不同刺激下响应的方差，而则是神经元在同一刺激下响应的方差在所有刺激上的平均值。通过这种方式，我们能够更加精确地评估并解释模型的预测效果。

### 偏好图像生成

为了直观地展示每个神经元所编码的特征，我们受到文献[1]的启发，采用了梯度上升法在图像像素空间对训练好的编码模型进行优化，从而为每个神经元生成其偏好图像。这一过程的核心是对以下损失函数进行优化：



此处，L的第一项旨在最大化目标神经元的预测响应，而第二项则是总变异（total variation）正则项，其作用是为了避免生成图像中出现过多的高频噪声。在这里，我们设置了 = 500。图像像素空间初始状态为随机均匀分布的噪声，学习率设定为0.01。整个优化过程包括200次迭代。

通过这种方法，我们能够生成和提取每个神经元的偏好图像，这些图像反映了神经元对特定视觉特征的响应模式，有助于我们更深入地理解神经元的编码机制。

### 卷积核消融实验

在我们已经训练好的多位点模型中，我们采取了一种特殊的方法来评估第一卷积层各卷积核的重要性。具体来说，我们逐个对第一卷积层的每个卷积核进行“损伤”处理，即将单个卷积核的权重设置为0，然后重新评估编码模型在验证集上的表现。这一过程的关键在于观察当特定卷积核被移除后，模型的可解释方差的解释比例（EVE）会发生怎样的变化。

我们以“损伤”某个卷积核所导致的EVE损失与总EVE的比例作为衡量该卷积核在神经元编码模型中重要性的指标。这种方法允许我们量化每个卷积核对于模型整体性能的贡献，从而更深入地理解模型的工作机制以及各卷积核在编码过程中的作用。通过这种系统的分析，我们能够揭示多位点模型中不同卷积核对神经元响应编码的影响，进而为深入理解神经网络的工作原理提供新的见解。

​

。

# 实验结果

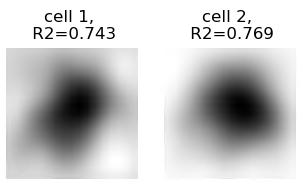
## 刺激图像的修改

### 目标

本节的目标是通过改进刺激图像的展示方式来增强模型的预测准确性。在模型训练的过程中，我们观察到每张图像首先通过一个预训练的AlexNet进行特征抽取。模型随后尝试挖掘这些特征与相应神经元反应之间的内在联系，这一步骤涉及到了对神经元感受野的定位，以及对各个通道权重的估计，这个过程我们称之为“What and Where”。

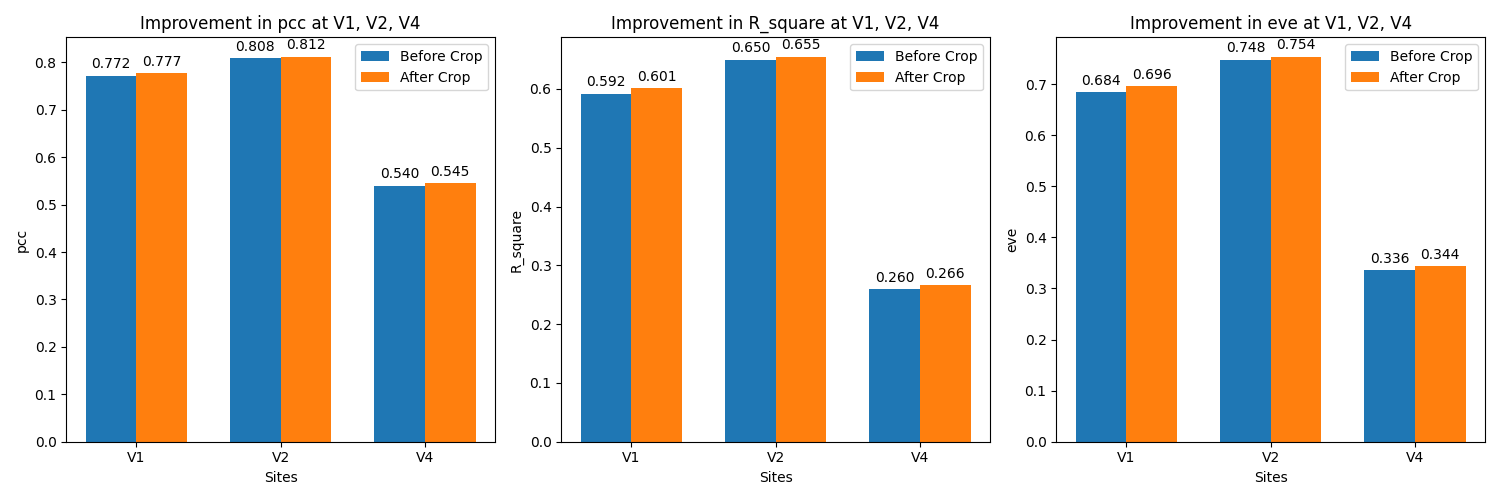
在训练过程中，模型致力于从720张训练图像中解码出神经元的感受野。但神经科学的研究表明，每个神经元的感受野是独特的。我们在实验中已经将图像定位于神经元感受野的中心，因此，仅有位于中心的像素会投射到神经元的感受野上。因此，对于一个300px的图像，我们考虑采取一种称为“crop”的策略，即裁剪出中心的圆形区域，并将圆外的像素设置为统一的灰色背景。

通过这种刺激图像的精简方法，我们能够显著减少解码的复杂性，进而有效地提高模型的预测准确度。为了进一步提升模型性能，我们采用了这种裁剪刺激图像的手段，并已经初步验证了其效果。



### 操作

在V1、V2、V4各取一个位点数据，将原有的刺激进行crop，比较crop前后神经元编码模型在两阶段训练下的拟合精度结果，如下图所示。





### 结论

在大部分数据集中，采用crop的方式能够有效地提升模型拟合精度。

## 不同数据集的比较

### 目标

在我们的数据集里，数据8,9是同一批神经元细胞在两组不同的图片刺激集下采集到的结果。我们希望通过比较这两组数据的拟合精度情况，从而指导模型的改进和提升。

### 基本结果比较

在对数据集8和9进行比较时，首先需要指出的是，两组数据所使用的刺激图片在分辨率上有所不同。具体来说，数据组8使用的刺激图片分辨率为300×300像素，而数据组9的图片分辨率为450×450像素。除此之外，两者在刺激内容上也存在显著差异：数据组9的刺激包含了背景元素，而数据组8的图片则以灰色背景为特征，正如图2所展示的。。



图3 数据8,9刺激图片对比

从统计学的视角进一步分析两类刺激集对113个神经元反应的影响，我们着手研究了它们在不同刺激集下的平均反应值和变异度。如图4所示，结果表明，当刺激图片含有背景时，神经元的反应强度通常较弱，且反应的变异度也较小，这两者的差异在统计上达到了显著水平(p < 0.005)。箱线图进一步证实了当刺激中包含背景元素时，神经元激发的反应强度普遍降低。

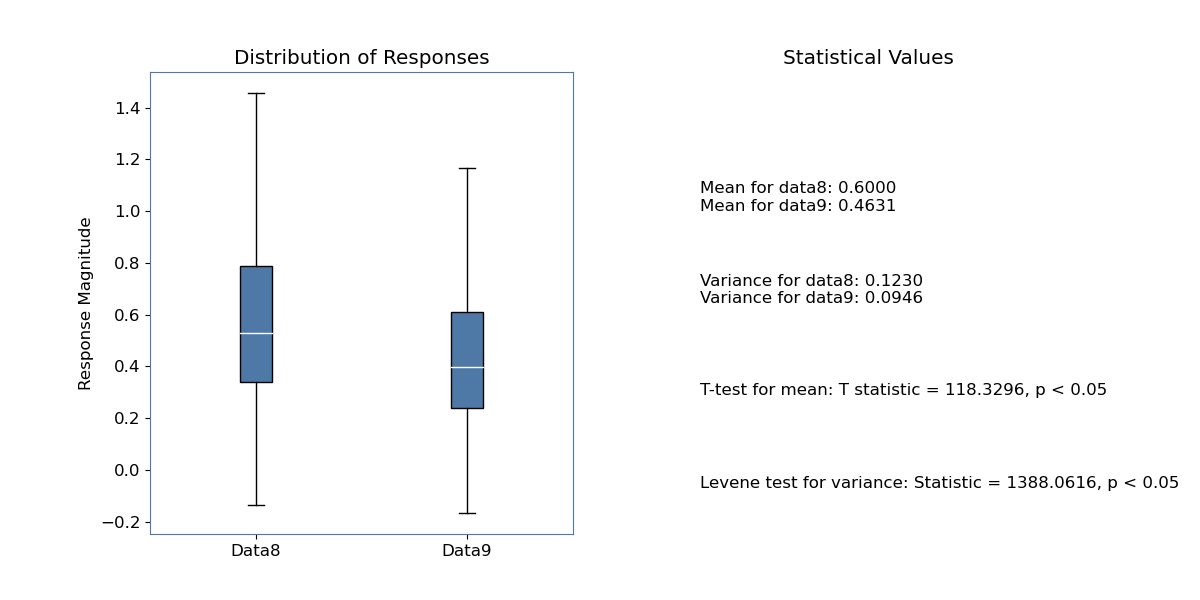
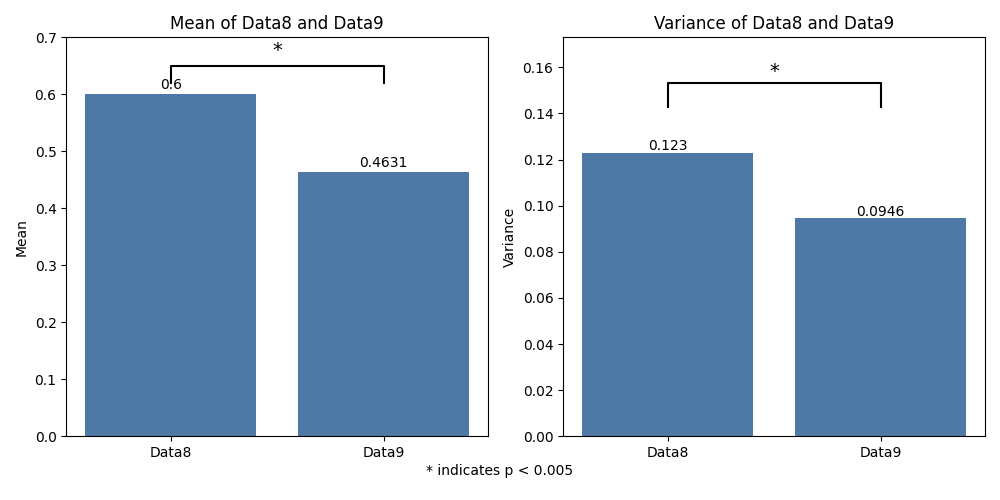


图5 数据8,9神经元反应统计分析

除了上述分析，我们还深入研究了数据组8和9中的试次相关性（trial correlation）。对每个神经元而言，我们计算了其在数据组8的刺激集下所有图片的试次相关性，并取其平均值；在数据组9中，我们采取了相同的方法。结果显示，与数据组8相比，数据组9中神经元的平均试次相关性显著降低。这表明，在多次实验中，神经元对于带有背景的图片刺激表现出较低的一致性反应。特别地，试次相关性低于0.2的神经元在数据组9中占据了不小的比例。这一趋势暗示，神经元对于包含背景元素的图片刺激反应的规律性较弱。为了直观展现这一差异，我们在图6中呈现了数据组8和9的神经元平均试次相关性的散点图。

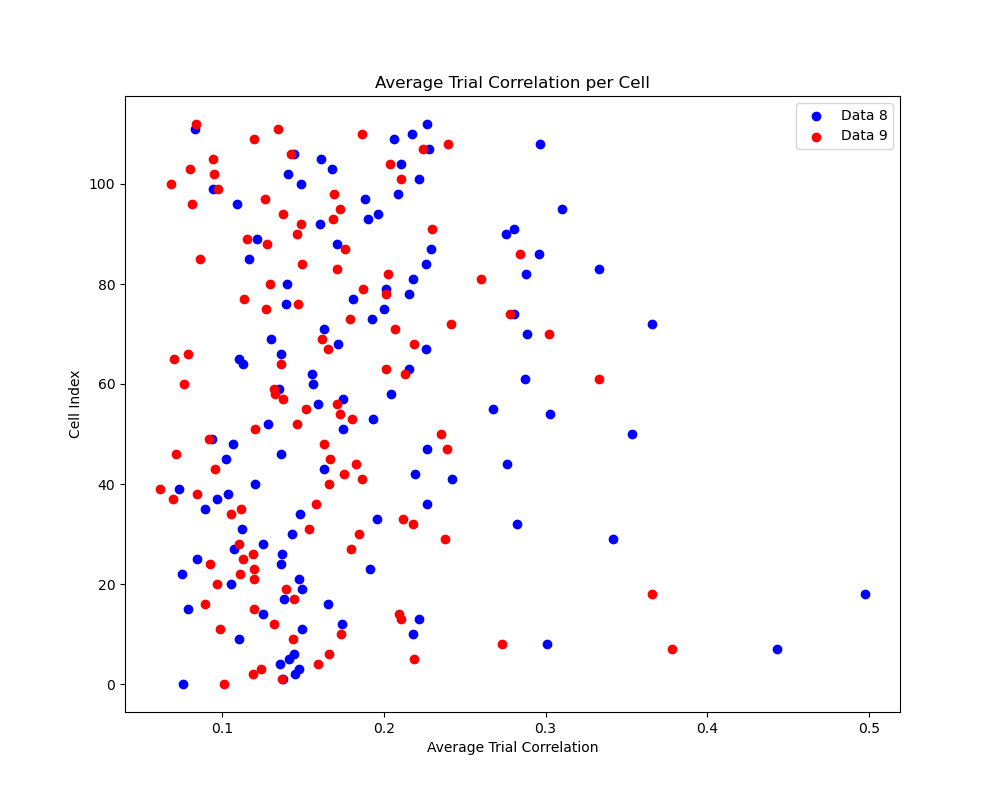


图7 数据8,9 trial correlation 对比

### 最优刺激比较

#### 单独训练

在我们模型里面，最大的优势就是能够通过生成偏好图片来可视化结果，从而判断模型的准确性，因此我们尝试分析使用数据8,9各自训练的模型得到的最优图片。

#### 交叉训练

#### 评价指标探索

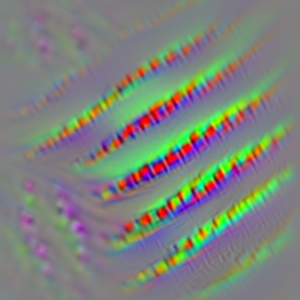
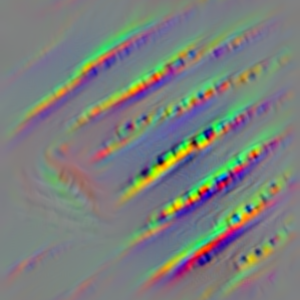
对于上述的偏好图片，我们希望能够有一种指标能够验证两张偏好图片的相似性，具体而言，我们希望有一种指标能够符合人眼的观感，越相似数值越大。一开始我们尝试了结构相似度(ssim)作为指标，结果如图所示。

图8 偏好图片进行SSIM计算

分析结果我们发现，对于某一个神经元的偏好刺激，即使给定另一个随机神经元的偏好刺激，ssim值也不会特别低。初步猜测是由于偏好图片的背景都是灰色导致两者的ssim会偏高。

为了排除背景因素的干扰，考虑到我们的偏好图片像素值为300\*300px，我们对偏好图片进行掩码操作，只对图片中心半径为125px的圆形区域进行ssim计算，掩码效果如图9所示。

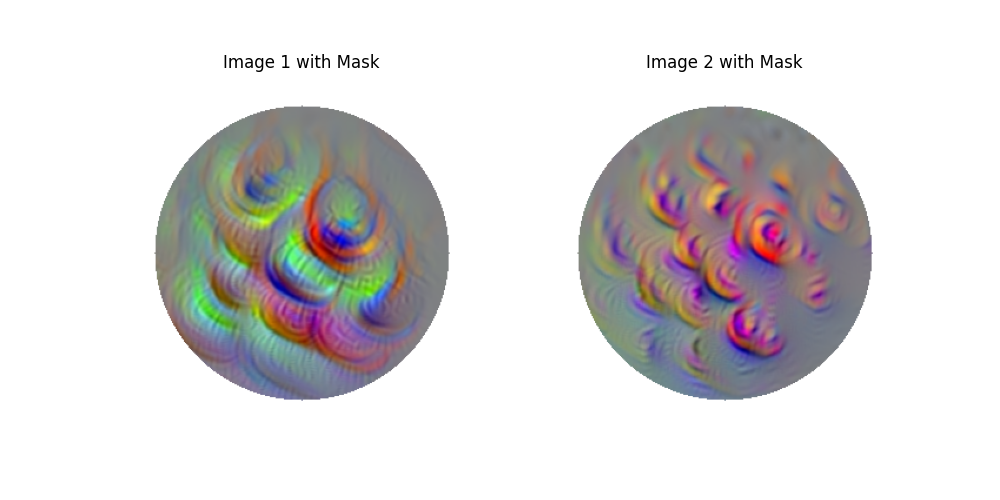


图10 通过掩码后进行比较

### 结论

# 讨论

## 

# 参考文献

1. 1. Hubel, D.H. and T.N.J.T.J.o.p. Wiesel, *Receptive fields, binocular interaction and functional architecture in the cat's visual cortex.* 1962. **160**(1): p. 106.
2. 2. Fukushima, K.J.B.c., *Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position.* 1980. **36**(4): p. 193-202.
3. 3. Cadieu, C., et al., *A model of V4 shape selectivity and invariance.* 2007. **98**(3): p. 1733-1750.
4. 4. Yamins, D.L., et al., *Performance-optimized hierarchical models predict neural responses in higher visual cortex.* 2014. **111**(23): p. 8619-8624.
5. 5. Tacchetti, A., L. Isik, and T.J.P.c.b. Poggio, *Invariant recognition drives neural representations of action sequences.* 2017. **13**(12): p. e1005859.
6. 6. Bashivan, P., K. Kar, and J.J. DiCarlo, *Neural population control via deep image synthesis.* Science, 2019. **364**(6439).
7. 7. Klindt, D., et al., *Neural system identification for large populations separating “what” and “where”.* Advances in neural information processing systems, 2017. **30**.
8. 8. Li, P., T.J. Hastie, and K.W. Church. *Very sparse random projections*. in *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*. 2006.
9. 9. Güçlü, U. and M.A.J.J.o.N. van Gerven, *Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream.* Journal of Neuroscience, 2015. **35**(27): p. 10005-10014.
10. 10. Schrimpf, M., et al., *Brain-Score: Which Artificial Neural Network for Object Recognition is most Brain-Like?* BioRxiv, 2018.
11. [10]

# 附 录

“附录”二字与题名间空一个汉字符位。宋体小四号（英文用Times New Roman体小四号），两端对齐书写，段落首行左缩进2个汉字符。行距20磅，段前段后0磅。

# 攻读学位期间取得的学术成果

**已发表（或正式接受）的学术论文：**（书写格式同参考文献）

**申请或已获得的专利：**（书写格式同参考文献，无专利时此项不必列出）

**参加的研究项目及获奖情况：**

# 致 谢

本篇论文的完成标志着我在北师大三年硕士研究生学习的顺利结束。这段学术之旅的每一步，都凝聚着诸多恩师与良朋的支持与鼓励，对此我心怀感激。

我首先要向我的恩师，吕海东教授，表达最深的敬意和感谢。吕老师不仅是我的学术导师，更是我心灵的引路人。在我刚踏入神经科学领域时的迷茫与徘徊中，是吕老师的悉心指导与不懈鼓励，点亮了我前行的灯塔。吕老师推荐的《视觉信息处理的脑机制》一书，极大地补充了我在视觉科学领域的知识空白，同时吕老师对科研的严谨态度与孜孜不倦的探索精神，更是我学术研究中的不竭动力。

对于实验室中的同仁，我同样怀有无尽的感激之情。每一位师兄师姐的悉心指导、每一次的共同探讨与协作，都让我在科研的道路上不断成长，开阔了视野。在实验的点滴中，我不仅提升了个人的实验技能，更对视觉科学有了更为深刻和全面的理解。

我还要特别感谢我的家人和朋友们。家人们无私的爱与支持，是我在追求学术梦想过程中最坚强的后盾。在每一次的迷惘与困惑中，家是我最温暖的避风港。父母亲的理解与支援，女朋友陈莹的鼓励与陪伴，都是我能够坚持下来的重要力量。同时，我也要感谢我的班主任郁曦老师，您的辛勤工作和教诲，让我们这个集体更加团结和谐，营造了一个有益于学术追求的优良环境。每一位朋友的关心与帮助，都为我的生活和研究增添了无限的色彩。

在此，我对所有给予我帮助与支持的人表示最诚挚的感谢和最崇高的敬意。这篇论文的每一个字节，都凝聚着大家的智慧与情谊。

2024年5月