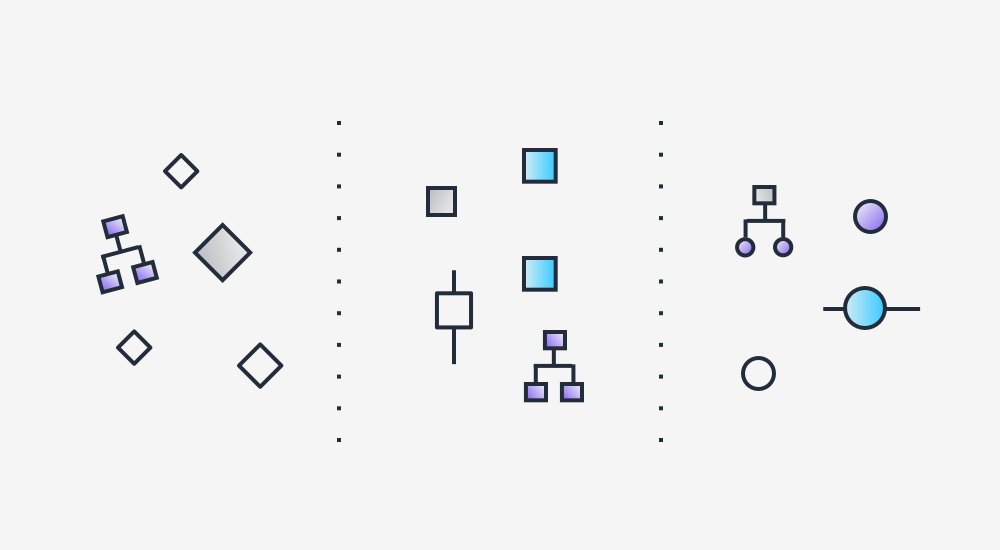
# 深层网络能学习不变量吗？



注：Alexey Potapov是奇点人工智能研究团队的一员。你可以在奇点人工智能研究实验室了解更多关于他们工作的信息。要直接与我们的团队和社区聊天，请访问奇点人工智能社区论坛。

人工通用智能（AGI）应该能够看到。特别是，它应该能够识别对象，并学会从尽可能少的示例中识别新的对象类。

这意味着它应该泛化。那会怎样？

假设我们有一个模式x，它可以通过某种转换T进行转换，并且转换参数w的所有值的转换模式T（x | w）应该与原始模式一致。目前还不清楚这种转变。也就是说，要么学习变换本身，要么学习不变识别模型。

深神经网络（DNNs）在计算机视觉领域是非常成功的。但是他们能在训练集的区域之外学习不变量或推广吗？或者他们只能插值和记忆吗？我们的团队一直在试验这些可能性。

### 承认

#### 实验装置

识别模型应该不变的常见变换是空间变换。利用卷积神经网络（CNNs）对位移不变性进行硬编码。旋转不变的cnn有时也被使用，但并不太流行。实现不变识别的常用技术是利用原始图像的空间变换版本扩展训练集。

但这有助于识别新姿势的物体吗？

理想情况下，机器学习系统应该能够在训练集中看到的参数值范围之外进行外推。例如，如果我们训练一个识别模型，识别在范围内旋转的所有MNIST数字[–45ᵒ，45ᵒ]，它应该能够识别旋转90ᵒ的这些数字。

我们从一个简单的任务开始。我们扩展了MNIST训练集（去掉“6”和“9”位），除了“3”和“4”位在整个角度范围内旋转，而“3”和“4”位在[–45ᵒ，45ᵒ]范围内旋转。基本但有指导意义的实验是，看看DNN模型是否有可能在不明确引入旋转不变性功能的情况下，识别出旋转角度超出范围[–45ᵒ，45ᵒ]的数字“3”和“4”。

注：目前有一些模型能够学习空间变换（如空间变换），特别是使识别模型对它们不变性尽管这样的解比硬编码对具体变换的不变性更实用、更一般，但它仍然假定变换的类及其适当的参数化是已知的。我们对DNNs实现对先验未知变换不变性的能力感兴趣。

一个更相关的工作的例子是“流形结构的非局部估计”，它精确地考虑了超出训练集区域的外推任务，并将这种外推（例如旋转流形结构）转移到新的图像类。不幸的是，作者只研究了“轻微的旋转”，应用于新图像已经产生了不完美的结果。因此，我们考虑通用DNN模型。

#### CNN公司

我们用基本的CNN网络做了第一个实验。我们的基线网络包含两个卷积层和两个密集层，softmax作为输出层（代码可用）。我们考虑了以下精度测试：

* 整个测试集的精确度，旋转角度与训练时使用的范围相同：Ptest=0.989。
* 旋转角度为[45ᵒ，315ᵒ]：Pout=0.212的数字3和4（测试集图像）的精度。
* 旋转角度为[135ᵒ，225ᵒ]：Pinv=0.003的数字3和4（测试集图像）的精度。

可以看出，该模型在训练集范围外的泛化能力很差，在训练集范围外的角度范围内，其精度几乎降为零（远远低于随机猜测），即使它是在其他数字的所有旋转角度上进行训练。

批处理规范化（BN）等技术提高了神经网络的泛化能力。我们尝试了BN和dropout以及它们的组合。最佳模型达到Ptest=0.993，Pout=0.257，Pinv=0.041。总体上精度较高，但仍远低于不变性检验的随机猜测。

我们可以假设网络不够深，更深的特征可能能够更好地推断流形结构。然而，在我们的实验中，在最佳模型中加入更多的层与两个卷积层相比，Ptest、Pout略有改善，Pinv也略有下降。

因此，传统的cnn无法在没有附加手段的情况下推广旋转的概念（不仅是外推到从未见过的旋转角度，而是将遇到的角度的识别能力从一个类转移到另一个类）。

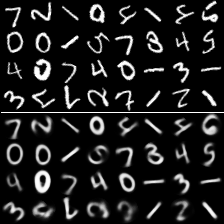
#### CapsNets公司

胶囊网络（CapsNets）是专为捕捉部分-整体关系而设计的，考虑到物体及其部分的姿势。我们在这里考虑胶囊网络，因为它们不显式地操作空间位置和转换，而是利用神经元层之间推理时间路由的一般机制，这依赖于胶囊的矢量输出。

首先，我们考虑了动态路由的CapsNet。在论文的基础上对其进行了初步的实现。

对于CapsNet的实现，我们得到了Ptest=0.991，Pout=0.249，Pinv=0.009，这并不优于带BN的CNNs（不变性较差）。

这个CapsNet模型包含一个用于正则化的解码器，还可以用来查看模式的最高级别表示中包含哪些信息。下面是训练解码器输出的一个例子（输入图像在上半部分，解码图像在下半部分）。



对于遇到的早期方向的图像，重建结果是合理的。然而，如果我们观察以下数字3和4在新方向上的重建结果，它们是完全错误的，并且类似于其他数字（因为它们被认为是这样的）。



我们还研究了使用它的性能。

得到以下分数：Ptest=0.976，Pout=0.218，Pinv=0.090。有趣的是，这个CapsNet在不变性测试中显示了最好的分数，虽然它在未扩展测试集上的准确度不太高，但它低于随机猜测的水平（所以它没有推广，只是没有过多的拟合）。

详细介绍了我们的实验。

### 一代

#### 实验装置

生成模型在AGI中非常重要。特别是，AGI系统应该具有能够以不同方式（例如，从不同的角度观察，但不仅仅是从不同的角度观察）渲染新对象的图像的想象力。我们在这里不会详细讨论，只是问生成网络是否能够学习执行“心理旋转”。

在给定旋转角度信息的情况下，我们训练模型来构造输入图像的旋转版本。由于我们对生成模型感兴趣，我们在实验中使用了自动编码器。编码器将输入图像转换成一些潜在的代码，然后将其与编码的旋转角连接起来，然后将其直接输入解码器，然后要求解码器重建正确旋转的图像。我们发现，表示模型输入角度a的最佳方法是通过一对值：cos（a）和sin（a）。

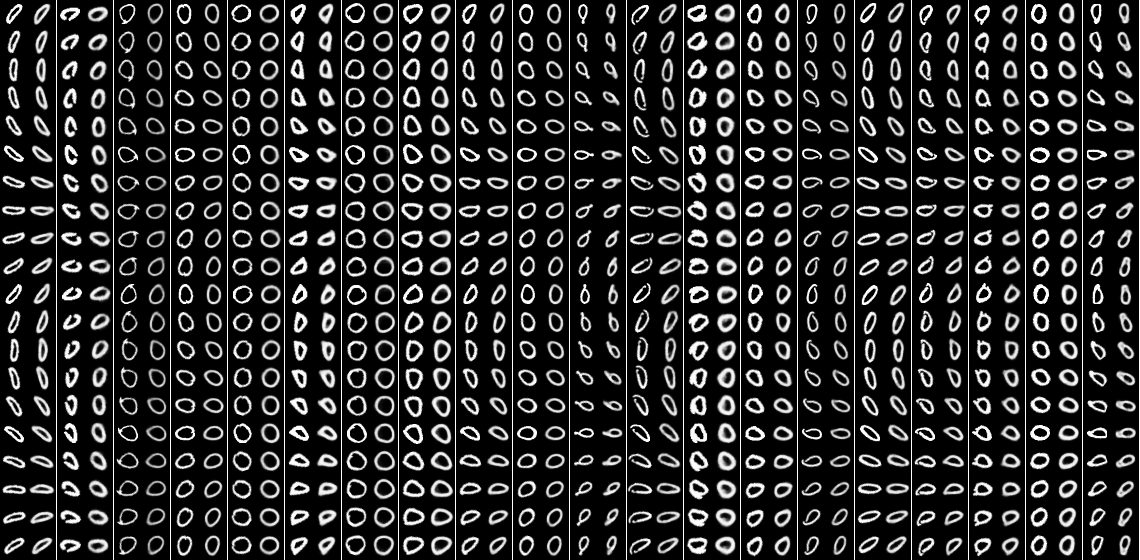
模型是在MNIST数据集上训练的。输入MNIST训练集的原始数字图像，输出旋转图像。对于每个训练样本，旋转角度从[–180ᵒ，180ᵒ]中随机选择，除了4和9，其范围为[–45ᵒ，45ᵒ]。应该注意的是，由于旋转角度是提供给模型的，所以混淆6和9没有问题。然后，评估来自测试集的图像的旋转。

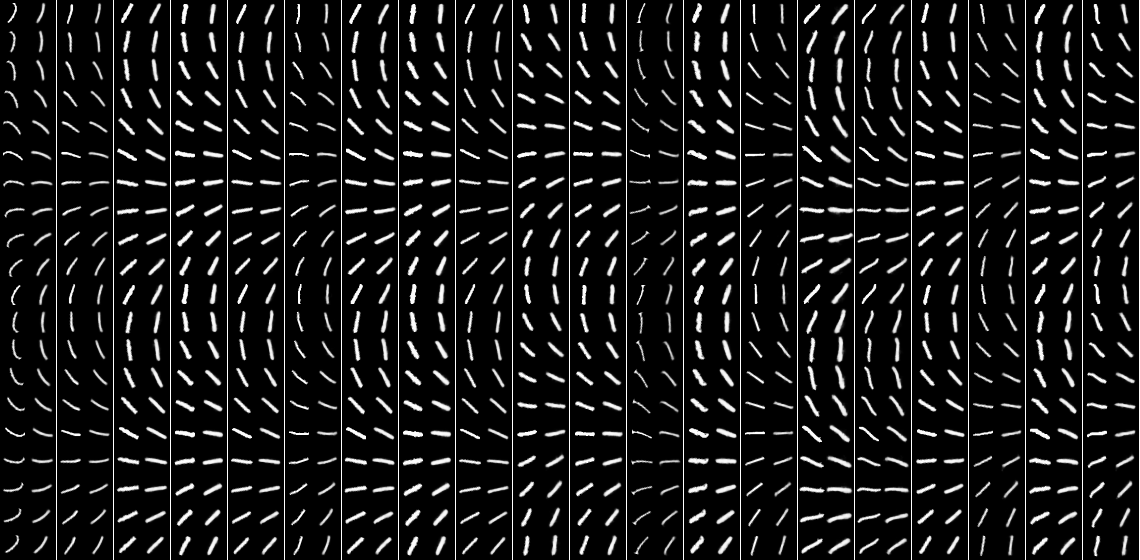
#### 旋转试验

我们对由两个卷积层（核=4，步幅=2）组成的编码器（64和128个特征映射）、1024个单位的中密层和10个单位的后密层组成的解码器进行了实验，解码器由相同数量的层和单位组成，但恢复了联系。整个潜在代码包含两个接收cos（a）和sin（a）的变量以及10个由编码器计算的变量。我们考虑了香草（非正则）自编码器的重建损失训练和两次更新（重建损失和对抗损失）训练。前者的重建结果稍差，但作为生成模型更有用。可以找到实现。

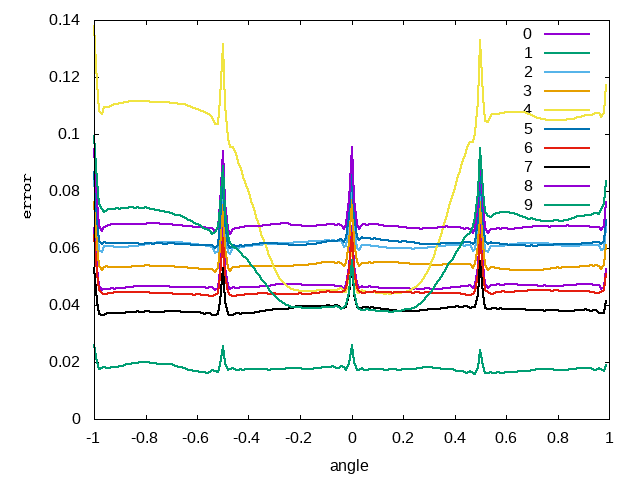
为了重建旋转图像，我们对自动编码器进行了端到端的训练，希望编码器能够学习到有用的潜在表征。

经过训练的模型能够几乎完美地重建在训练集范围内以角度旋转的已知数字的新图像（左栏包含预期的重建结果 - ；而右栏包含实际的重建结果）：

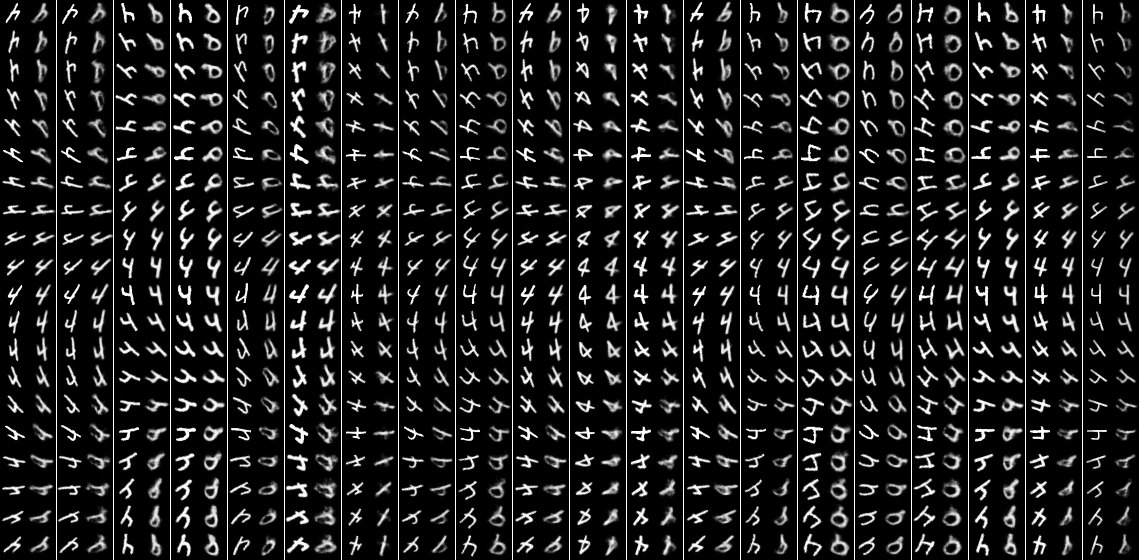


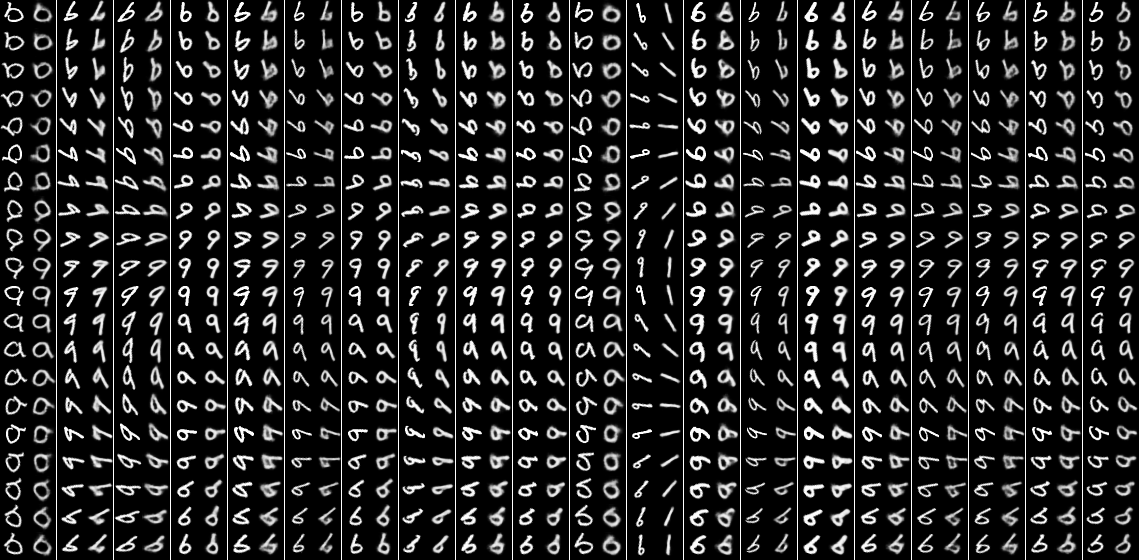


但是，数字“9”的角度超出范围[-45ᵒ，45ᵒ]时的重建误差似乎很大，而数字“4”的重建误差则要大得多。应忽略0ᵒ和±90ᵒ上的峰值，因为它们与图像在0ᵒ和±90ᵒ不同角度下的双线性模糊有关。



考虑以下不同角度的“4”和“9”的重建（旋转）图像示例。



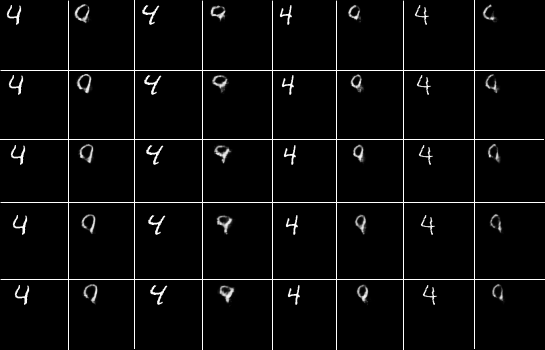


虽然在某些情况下，将“4”旋转到大角度的结果并不太差，但它远不是完美的，而且往往类似于其他数字。“9”的情况稍微好一些，因为网络学会了在所有方向绘制“6”。该模型似乎成功地将“6”和“9”的潜在码之间的联系外推到了训练集中所给出的角度范围之外，但它并没有推广旋转过程。在某些情况下，还可以看到模型将“9”转换为“0”、“1”或“7”。

因此，该模型不仅不能将旋转外推到新的角度，而且不能应用它所获得的将一些数字旋转到其他数字的能力。因此，我们的结论是，虽然进行了一些推广，但该模型主要记住了不同角度旋转的不同数字的外观。

#### 移位试验

我们做了一个类似的实验，但不是旋转而是移位。由于使用了卷积译码器，因此学习如何重建图像的移位版本应该没有问题。但是，结果与旋转相同。



这个结果很容易理解。潜在的编码神经元与层的每个神经元都有单独的连接，这被解释为一组特征映射，然后对其应用（转置）卷积。显然，通向这些特征地图每个区域的连接的权重是独立学习的。

* 如果模型学会了在一个地方激活适当的特征来绘制一个数字，那么它将无法帮助模型在另一个地方绘制相同的数字，因为相应的连接将不会被训练。
* 如果译码器输入一个数字潜码和一个新的移位，它将在适当的位置激活特征映射的神经元，但它们将对应于具有相似潜码的数字特征的混合，模型知道如何在这个位置绘制它们。

因此，该模型不会将一个数字的图像移动到一个新的位置，而是将其从其他数字的图像（在特征层次上融合）中组合起来。当然，如果我们尝试用移位重建一个数字，而这个移位根本不存在于训练集中，那么模型将无法用这个移位呈现任何内容。

更详细地描述了使用自动编码器的实验。

### 识别奇异网络的唯一神经网络设计

正如我们所看到的，神经网络无法学习（泛化/外推）区分环境中的不变量或生成环境中的变换，尽管它们不只是记忆模式和显示某种形式的泛化。

尽管传统的神经网络是非常强大的模型，可以用来建立许多奇点网络节点，但要想创建AGI，很可能需要创建能够真正泛化的新的神经模型，或者将DNN节点连接到一种符号归纳节点，这将有助于克服彼此的限制。

这些实验的结果为能够产生AGI的神经网络提供了新的可能性。我们将在以后的文章中介绍一些关于这些计划的早期工作。

### 你怎么能参与进来？

虽然我们的人工智能研究实验室可以让您进入我们的人工智能计划，但我们还没有完成！在上，您可以直接与我们的人工智能团队，以及来自世界各地的开发人员和研究人员聊天。这是你直接影响人工智能未来的机会。

[**SingularityNET Community Forum** Join our community today to help shape the future of AI.community.singularitynet.io](https://community.singularitynet.io/)

### 敬请期待！

在接下来的几周里，我们将有大量令人兴奋的内容发布在我们的出版物上。在接下来的几周里，我们将看到：

* 团队介绍- 我们将直接向您和社区介绍我们人工智能团队的每个新成员。
* 访谈——我们的团队已经吸引了人工智能和机器学习领域的几位领先思想家。我们将做独家采访，让他们讨论他们的工作将如何影响奇点网和其他。
* 研究论文——我们的大部分工作将涉及生物信息学、语言学习、元学习等领域的开创性研究。奇点网社区将首先获得我们的突破。
* 项目更新-我们正在努力使我们的幕后工作100%透明和互动。期待看到我们所有研究领域的最新进展。单击下面的链接了解更多信息：

[**Launching the SingularityNET AI Research Lab** Bringing radical transparency to the groundbreaking AI initiatives that will seed our network.medium.com](https://medium.com/singularitynet-ai-research-lab/launching-the-singularitynet-ai-research-lab-27daf6501e5)

