

摘 要

单视角三维物体形状重建 (Single-view 3D Reconstruction) 是三维视觉领域一直以来的核心问题之一。由于拟合非线性方程和学习模式的有效性, 深度神经网络被期望在重建三维非线性形状的任务中表现出色。本课题探究了深度神经网络在三维重建任务中的有效性和内在机制, 主要的工作和创新如下:

1. 针对基于自编码器结构的深度神经网络多次训练结果差异较大的问题, 提出采用多任务训练对编码器进行初始化, 以及引入残差层增加解码器网络深度这两种方法优化网络架构。在 ShapeNet 公开数据集上进行测试, 取得了优于基于检索的非学习方法 Oracle Nearest Neighbor 的结果。
2. 为了衡量三维数据集的聚类程度, 提出了基于 affinity propagation 和 silhouette score 的度量标准, 该方法得到的量化结果与数据集降维后的可视化结果吻合。
3. 提出了影响三维重建任务中深度神经网络内在机制的主要因素是数据集的本质特征: 当训练集中的三维形状集有相比于图片集更高的聚类程度的结构时, 在这个数据集上训练的深度神经网络更有可能会执行识别机制而非重建机制, 并定性和定量的证明了这一强相关性。

最后, 本文分析了上述结论对三维重建任务中的数据收集和神经网络训练的指导意义。

关键词: 三维重建, 点云, 深度学习, 数据挖掘

ABSTRACT

Single-view 3D reconstruction is one of the elementary tasks in the field of 3D vision. Due to the effectiveness in learning and approximating non-linear function, deep neural networks are expected to perform well on the task of reconstructing 3D nonlinear shapes. This paper investigates the effectiveness and internal mechanism of deep neural network in 3D reconstruction task. The main contribution and innovation are summarized as follows:

1. For the problem that the variance of multiple training results of deep neural network based on autoencoder structure is too large, it proposes to leverage multi-task training to initiate the encoder and introduce residual layer to deepen the decoder in order to optimize network architecture. These methods are trained and tested on ShapeNet public dataset and they outperform the non-learning method based on retrieve, namely Oracle Nearest Neighbor.
2. Define novel way to measure clustering coefficient of 3D reconstruction dataset based on affinity propagation and silhouette score, the quantitative results of this metric is corresponding to qualitative results of visualization of dataset in low dimension.
3. It claims that the bias of internal mechanism of network is mainly affected by the intrinsic properties of dataset: when the training set of 3D shapes has a more clustered structure than images, the deep neural networks trained on this dataset become more likely to perform recognition than reconstruction, and it proves the strong correlation between the deep neural networks and dataset property both qualitatively and quantitatively.

Finally, this paper analyzes the significance of the above conclusions for data collection and neural network training.

KEY WORDS: 3D Reconstruction, Point Cloud, Deep Learning, Data mining

摘 要	1
ABSTRACT	2
目 录	3
第一章 前言	1
1.1. 数学公式	1
1.1.1. 简单的数学公式	1
1.1.2. 带自动编号的公式	1
1.1.3. 带等号对齐的公式	1
1.2. 伪代码	1
1.3. 插入图片	1
1.4. 引用论文	1
第二章 研究内容	4
2.1. 本章小结	4
参考文献	5
附录 A L ^A T _E X 实验	6
附录 B MATLAB 实验	7
致 谢	8

第一章 前言

1.1 数学公式

1.1.1 简单的数学公式

卷积（convolution）在图像分析的线性方法中是一种重要的运算。卷积是一个积分，反映一个函数 $f(t)$ 在另一个函数 $h(t)$ 移动时所重叠的量。函数 f 和 h 在有限域 $[0, t]$ 上的 1D 卷积 $f * h$ 由下式给出：

$$(f * h)(t) \equiv \int_0^t f(\tau)h(t - \tau)d\tau$$

1.1.2 带自动编号的公式

这里可以限定在 $[0, t]$ 区间，原因是我们假设负坐标部分的值是 0。为了准确起见，我们还可以将卷积积分的上限扩展为 $(-\infty, \infty)$ ：

$$(f * h)(t) \equiv \int_{-\infty}^{\infty} f(\tau)h(t - \tau)d\tau = \int_{-\infty}^{\infty} f(t - \tau)h(\tau)d\tau \quad (1.1)$$

1.1.3 带等号对齐的公式

卷积可以推广到更高维。令 2D 函数 f 和 h 的卷积 g 记为 $f * h$ ，则有：

$$\begin{aligned} (f * h)(x, y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(a, b)h(x - a, y - b)dadb \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x - a, y - b)h(a, b)dadb \end{aligned} \quad (1.2)$$

1.2 伪代码

在写论文的时候我们通常要写伪代码，伪代码里面有时甚至还要包含数学公式（如根号一类的）。伪代码会自动找一个比较好的位置插入图片。

1.3 插入图片

在使用该命令的时候，图片会自动找一个他觉得比较好的位置插入图片，我们就不需要担心前面改了文字之后后面的格式乱掉。

1.4 引用论文

使得论文符合要求^{[1][2]}。

算法 1 用归并排序求逆序数

输入: *Array* 数组, *n* 数组大小

输出: 逆序数

```

1: function MergerSort(Array, left, right)
2:   result  $\leftarrow$  0
3:   if left < right then
4:     middle  $\leftarrow$  (left + right)/2
5:     result  $\leftarrow$  result + MergerSort(Array, left, middle)
6:     result  $\leftarrow$  result + MergerSort(Array, middle, right)
7:     result  $\leftarrow$  result + Merger(Array, left, middle, right)
8:   end if
9:   return result
10: end function
11:
12: function Merger(Array, left, middle, right)
13:   i  $\leftarrow$  left
14:   j  $\leftarrow$  middle
15:   k  $\leftarrow$  0
16:   result  $\leftarrow$  0
17:   while i < middle and j < right do
18:     if Array[i] < Array[j] then
19:       B[k + +]  $\leftarrow$  Array[i + +]
20:     else
21:       B[k + +]  $\leftarrow$  Array[j + +]
22:       result  $\leftarrow$  result + (middle - i)
23:     end if
24:   end while
25:   while i < middle do
26:     B[k + +]  $\leftarrow$  Array[i + +]
27:   end while
28:   while j < right do
29:     B[k + +]  $\leftarrow$  Array[j + +]
30:   end while
31:   for i = 0  $\rightarrow$  k - 1 do
32:     Array[left + i]  $\leftarrow$  B[i]
33:   end for
34:   return result
35: end function

```

<i>Form.</i>	<i>Phonetic Value.</i>	<i>Name.</i>	<i>Form.</i>	<i>Phonetic Value.</i>	<i>Name.</i>
A a	papa, father	Alpha	N v	now	Nu
B β	bed	Beta	Ξ ξ	wax	Xi
Γ γ	go or sing (10)	Gamma	Ο ο	obey	Omicron
Δ δ	do	Delta	Π π	pet	Pi
Ε ε	met	Epsilon	Ρ ρ	run	Rho
Ζ ζ	adze	Zeta	Σ σ s	sit	Sigma
Η η	prey	Eta	Τ τ	tell	Tau
Θ θ	thin	Theta	Υ υ	French u, German ü	Upsilon
Ι ι	pin, machine	Iota	Φ φ	graphic	Phi
Κ κ	kill	Kappa	Χ χ	German buch	Chi
Λ λ	land	Lambda	Ψ ψ	hips	Psi
Μ μ	men	Mu	Ω ω	tone	Omega

图 1-1 图片的一个简单应用场景

<i>Form.</i>	<i>Phonetic Value.</i>	<i>Name.</i>	<i>Form.</i>	<i>Phonetic Value.</i>	<i>Name.</i>	<i>Form.</i>	<i>Phonetic Value.</i>	<i>Name.</i>	<i>Form.</i>	<i>Phonetic Value.</i>	<i>Name.</i>
A a	papa, father	Alpha	N v	now	Nu	A a	papa, father	Alpha	N v	now	Nu
B β	bed	Beta	Ξ ξ	wax	Xi	B β	bed	Beta	Ξ ξ	wax	Xi
Γ γ	go or sing (10)	Gamma	Ο ο	obey	Omicron	Γ γ	go or sing (10)	Gamma	Ο ο	obey	Omicron
Δ δ	do	Delta	Π π	pet	Pi	Δ δ	do	Delta	Π π	pet	Pi
Ε ε	met	Epsilon	Ρ ρ	run	Rho	Ε ε	met	Epsilon	Ρ ρ	run	Rho
Ζ ζ	adze	Zeta	Σ σ s	sit	Sigma	Ζ ζ	adze	Zeta	Σ σ s	sit	Sigma
Η η	prey	Eta	Τ τ	tell	Tau	Η η	prey	Eta	Τ τ	tell	Tau
Θ θ	thin	Theta	Υ υ	French u, German ü	Upsilon	Θ θ	thin	Theta	Υ υ	French u, German ü	Upsilon
Ι ι	pin, machine	Iota	Φ φ	graphic	Phi	Ι ι	pin, machine	Iota	Φ φ	graphic	Phi
Κ κ	kill	Kappa	Χ χ	German buch	Chi	Κ κ	kill	Kappa	Χ χ	German buch	Chi
Λ λ	land	Lambda	Ψ ψ	hips	Psi	Λ λ	land	Lambda	Ψ ψ	hips	Psi
Μ μ	men	Mu	Ω ω	tone	Omega	Μ μ	men	Mu	Ω ω	tone	Omega

(a) the first subfigure

(b) the second subfigure

图 1-2 子图应用场景

第二章 研究内容

2.1 本章小结

参考文献

- [1] Yao D, Zhao P, Yu C, et al. Sparse Online Relative Similarity Learning. 2015 IEEE International Conference on Data Mining (ICDM). IEEE, 2015. 529–538.
- [2] hbsjzzxgssb. 东南大学学位论文封面, 2007.

附录 A L^AT_EX 实验

技术实验结果在这里写

附录 B MATLAB 实验

技术实验结果在这里写

致 谢

这次的毕业论文设计总结是在我的指导老师 xxx 老师亲切关怀和悉心指导下完成的。从毕业设计选题到设计完成，x 老师给予了我耐心指导与细心关怀，有了莫老师耐心指导与细心关怀我才不会在设计的过程中迷失方向，失去前进动力。x 老师有严肃的科学态度，严谨的治学精神和精益求精的工作作风，这些都是我所需要学习的，感谢 x 老师给予了我这样一个学习机会，谢谢！

感谢与我并肩作战的舍友与同学们，感谢关心我支持我的朋友们，感谢学校领导、老师们，感谢你们给予我的帮助与关怀；感谢肇庆学院，特别感谢计算机科学与软件学院四年来为我提供的良好学习环境，谢谢！