基于强化学习与形状签名的排样学习模型

Learn to pack with reinforcement learning and shape signature

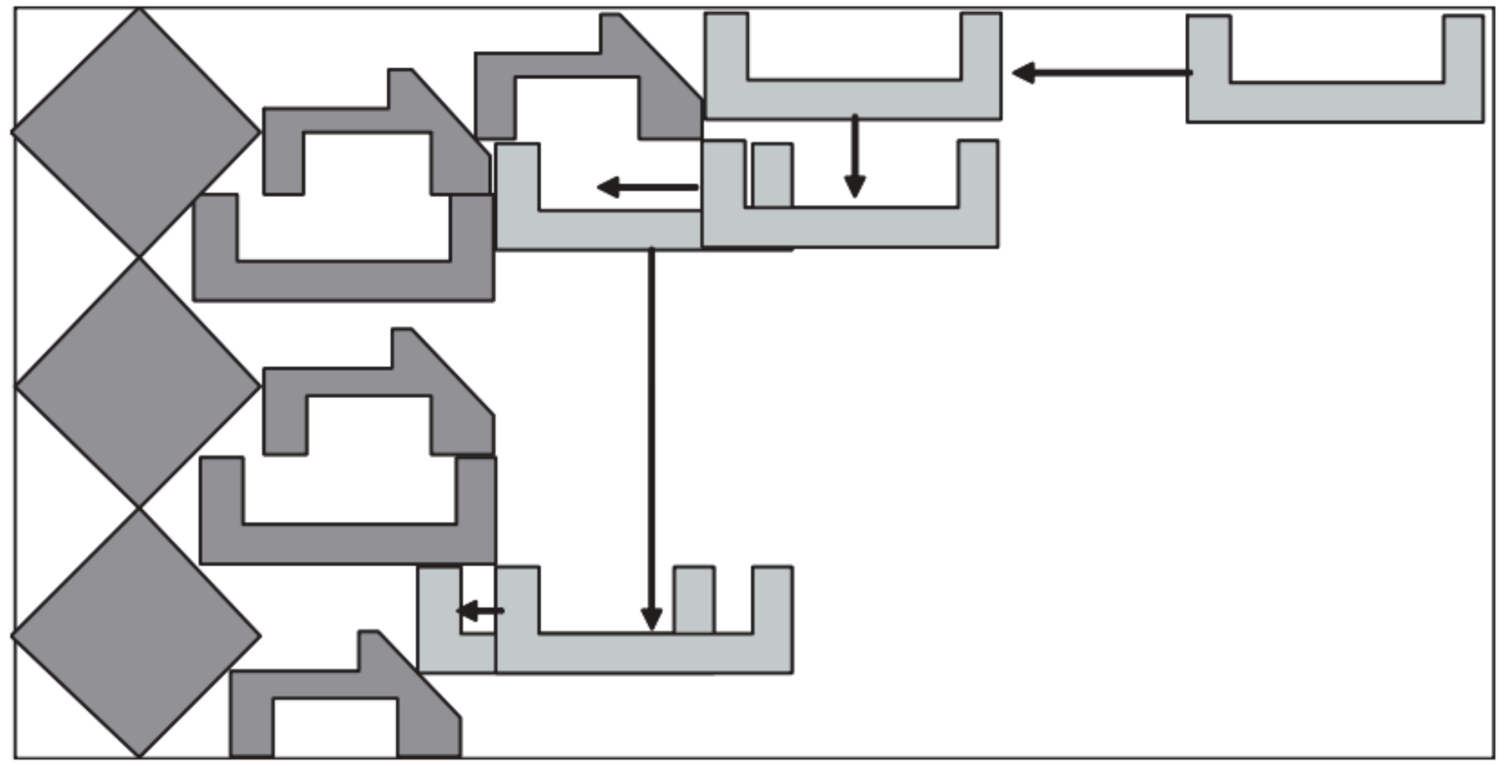
王子路, 羊山

2020.6.28

研究简介

Introduction to our research

同濟大學 TONGJI UNIVERSITY



左底部启发式排样

通过启发式标准加入形状，比如按照面积降序、高度降序，将形状逐个加入容器中，选择最左底部位置[4][5]

二维排样问题 [2D irregular packing problem]

一种常见的组合优化问题，一般通过左底部算法获得初始解（如上），再通过基于布局优化寻找更优的解，优化过程受初始解影响，该问题为NP-Hard[1][2][3]

研究问题

初始加入序列直接影响后续优化，但是启发式算法的效果视数据集而定，且无法获得适应数据集的序列

我们做了什么？

1. 建立了面向二维排样的学习模型

通过图像签名（Shape Signature）[10]解决了图形数据嵌入问题，基于强化学习解决了数据集的扩展性问题

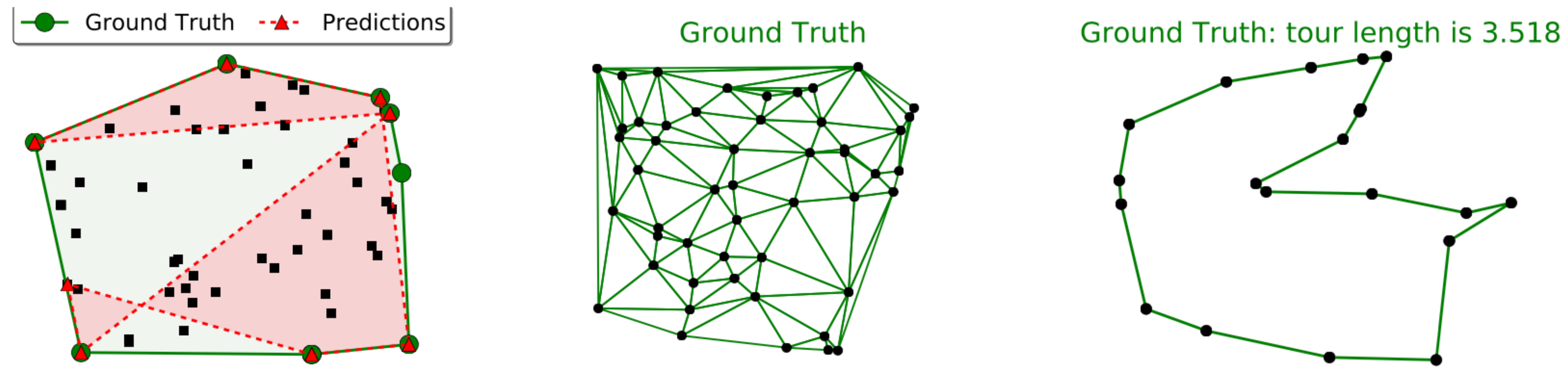
2. 序列预测的实验结果优于启发式标准

我们通过上述模型与序列学习网络Pointer-Network[7]，在生成数据集及部分训练集上，都获得了优于最好的启发式标准5%-15%的排样高度

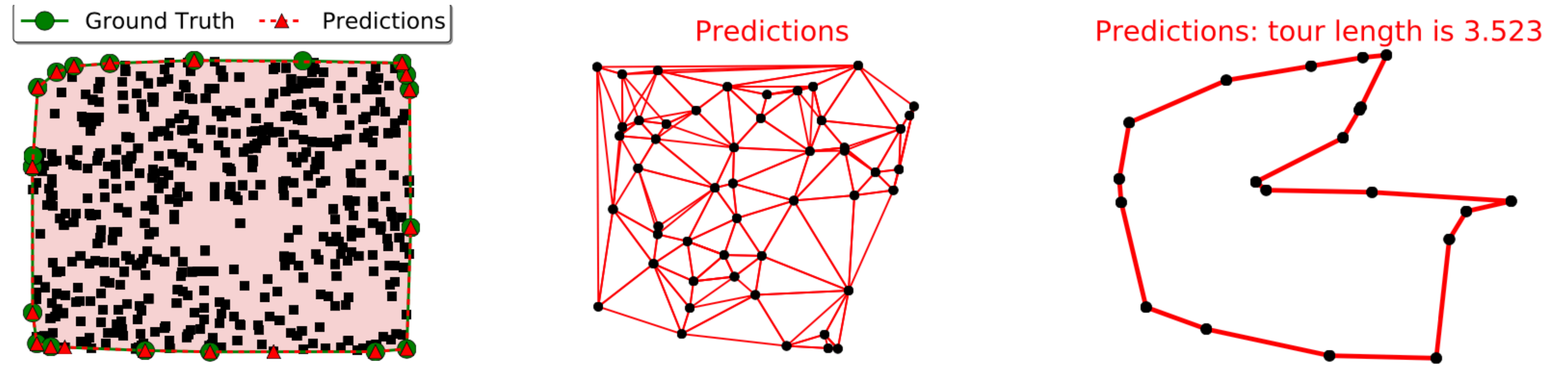
在深度学习与二维不规则排样的结合进行了第一次探索

组合优化与深度学习

Combinatorial optimization and deep learning



(a) LSTM, m=50, n=50　(b) Truth, n=50　(c) Truth, n=20



(d) Ptr-Net, m=5-50, n=500　(e) Ptr-Net , m=50, n=50　(f) Ptr-Net , m=5-20, n=20

Algorithm 1 Training Procedure

1: Training set S, number of training steps T, batch size B.

2: Initialize Pointer Net params θ.

3: Initialize baseline value according to heuristic algorithm.

4: **for** t = *1 to T* **do**

5: Select a batch of sample s_i for $i \in \{1, \cdots, B\}$.

6: Sample solution o_i based on $p_\theta(\cdot|s_i)$ for $i \in \{1, \cdots, B\}$.

7: Let $g_\theta = \frac{1}{B}\sum_{i=1}^{B}[(SA(o_i|s_i) - b(s_i))\nabla_\theta log p_{\boldsymbol{\theta}}(o_i|s_i)]$.

8: Update $\theta = ADAM(\theta, g_\theta)$.

9: Update baseline $b(s_i) = b(s_i) + \alpha(SA(o_i|s_i) - b(s_i))$ for $i \in \{1, \cdots, B\}$.

10: **end for**

11: return pointer net parameters θ.

序列学习网络

现阶段，Pointer Network是较为成熟的序列学习算法，已经应用在凸包寻找、旅行商问题等序列为基础的组合优化问题，同样也可以用在排样问题上[7][8][9]

强化学习

强化学习算法能够解决组合优化学习的训练集不足的问题，比如基于Policy Gradient的强化学习模型可以持续学习持续优化[11][12][13]

Learn to pack with reinforcement learning and pointer network

思路来源

The original idea

同濟大學 TONGJI UNIVERSITY

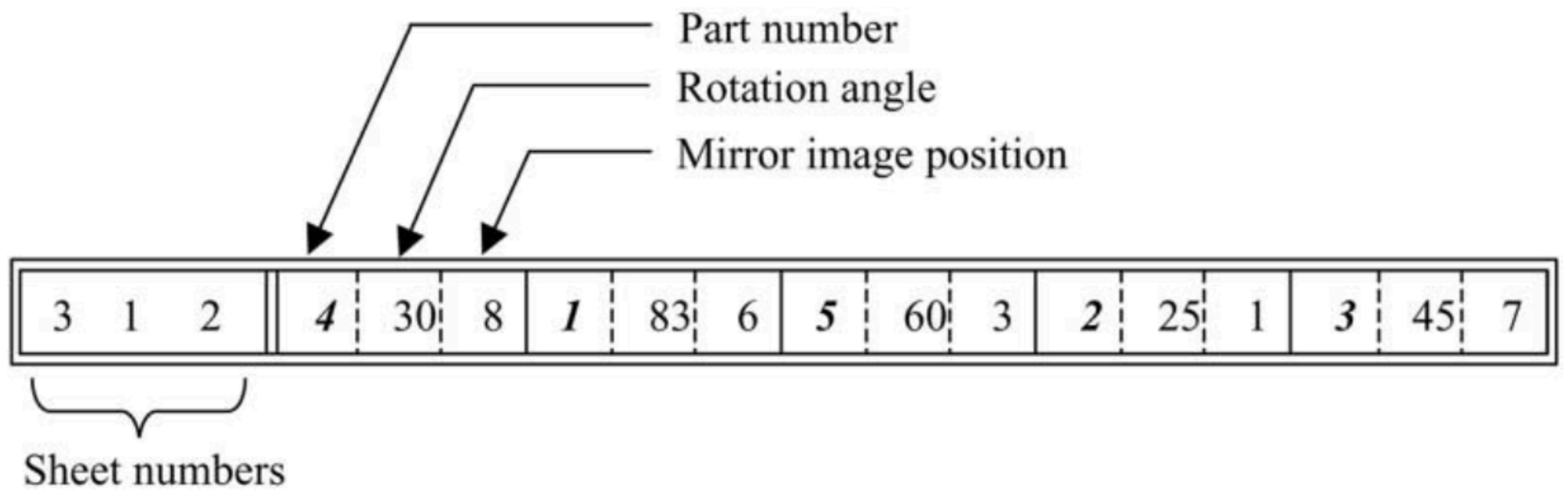
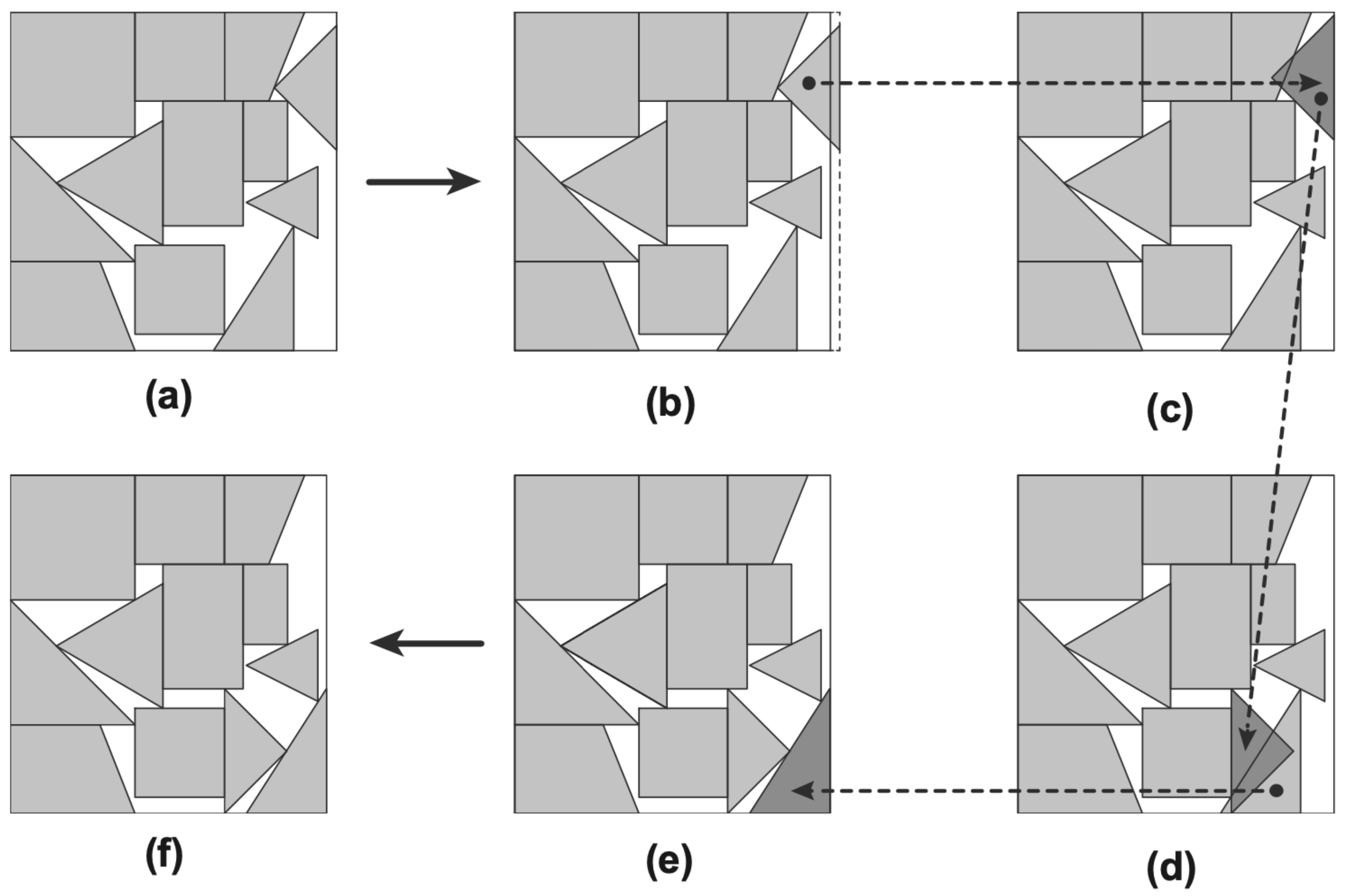


Fig. 5. A typical genetic string for three sheets and five parts.

通过遗传算法优化加入序列[6]



通过优化算法优化布局

Learn to pack with reinforcement learning and pointer network

解决思路

1. 能否直接预测合理序列?

一般来说通过遗传算法获得优化后的序列，可以通过Pointer Network序列网络，来学习在确定输入下，应该如何排序。

2. 能否直接预测形状布局?

考虑到直接预测形状布局涉及多个影响因素，可行性有待考证。

具体落实

1. 直接预测序列有技术基础可行性，且网络较为简单，可以与启发式序列和优化序列分别比较

2. 将形状嵌入一维向量后即可进行训练，理论上预测的序列将逊于遗传算法获得的优化序列，但是会优于启发式序列

方案概述

Algorithm outline

The *centroid distance* function is expressed by the distance of the boundary points from the centroid (x_c, y_c) of the shape

$$r(t) = ([x(t) - x_c]^2 + [y(t) - y_c]^2)^{1/2}$$

$r(t)$ is invariant to translation. Computation of $r(t)$ is low. Figure 1 shows the centroid distance signatures of the an apple shape (referred to as the apple in this paper).

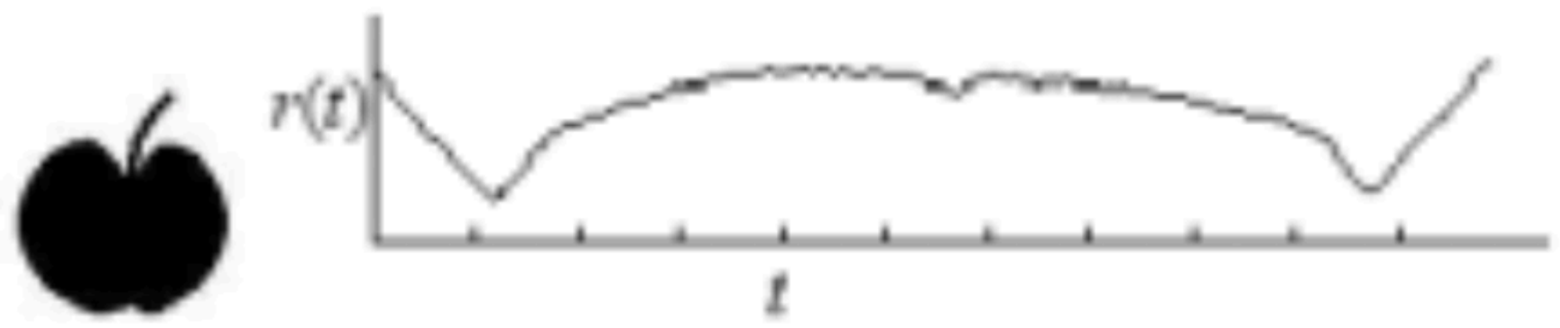


Figure 1. The centroid distance signatures of an apple.

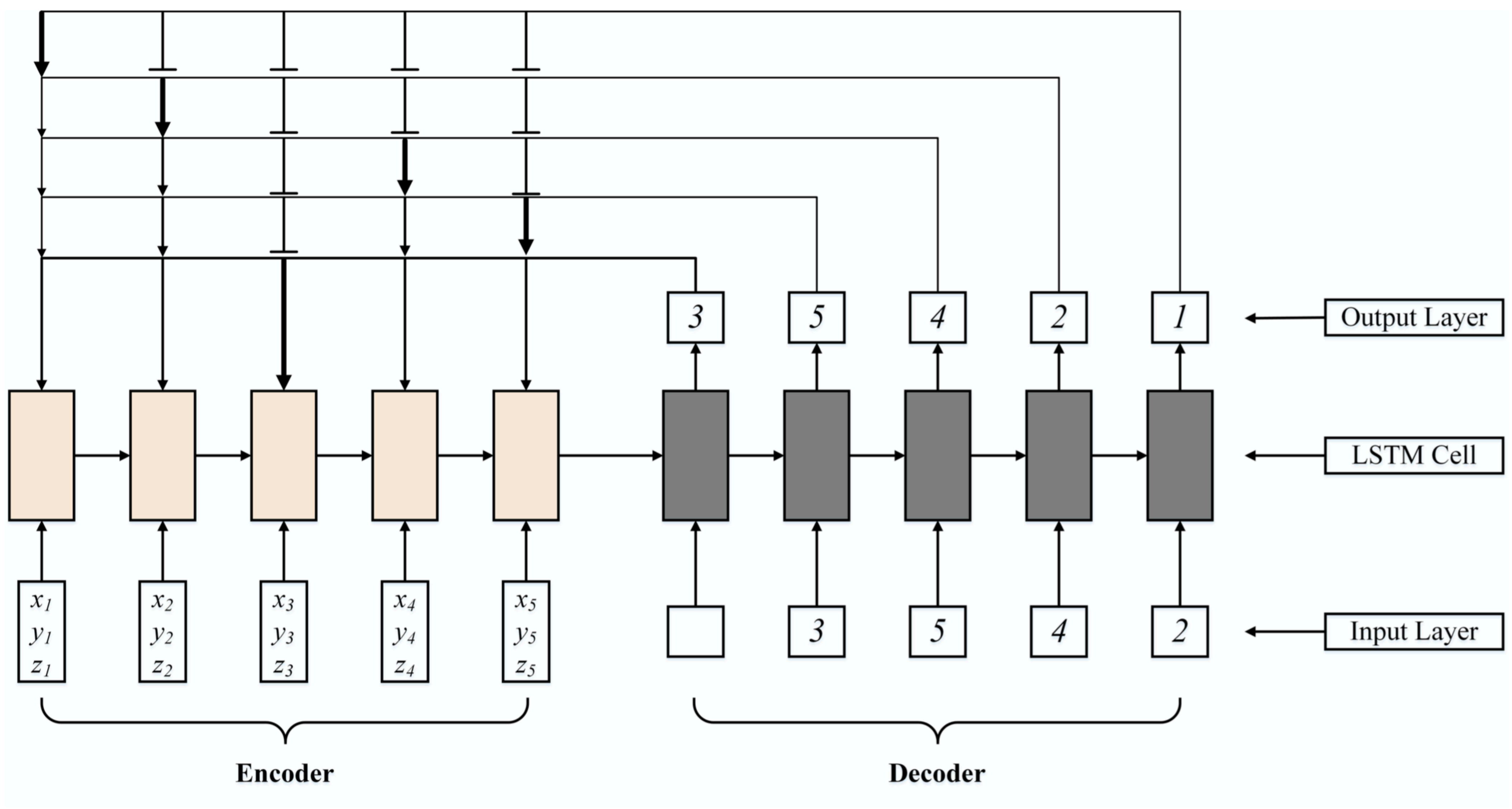


Figure 1: Architecture of the neural network

嵌入一维向量

形状签名（Shape Signature）可以将形状比较好的嵌入一个一维向量，同时不造成过多损失，我们对基准数据集实现了基本2%以内损失压缩[10]

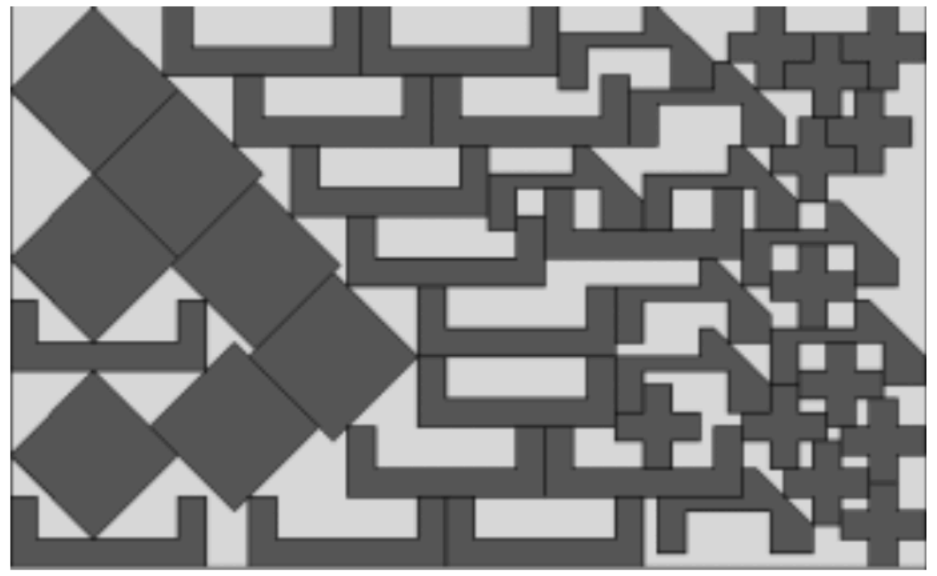
序列网络+强化学习

将通过形状签名嵌入的一维向量的形状，输入由LSTM组成的序列网络Pointer Network，基于Policy Gradient，实现对序列的持续学习[13]

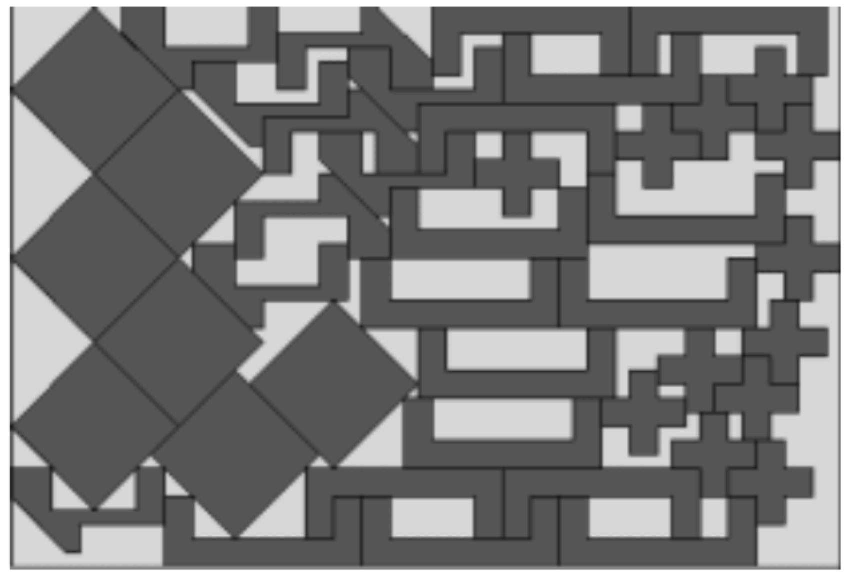
Learn to pack with reinforcement learning and pointer network

实验数据与训练过程

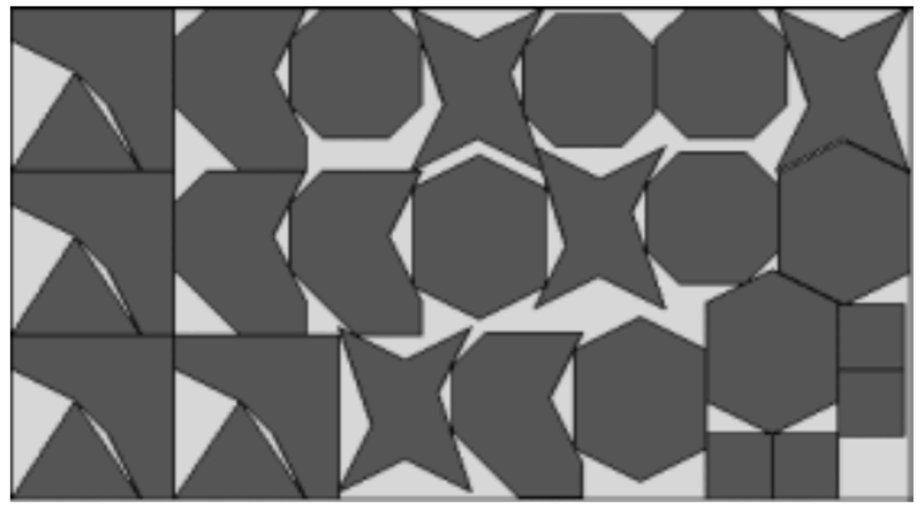
Experimental data and training process



(a)



(b)



(c)

形状与数据集生成

我们参考了几个常作为EURO的部分数据集，生成了2379个训练样本与817个验证样本，每个样本有12个形状，全部嵌入范围[-1,1]的一维向量

Algorithm 1 Training Procedure

Training set S, number of training steps T, batch size B.

Initialize Pointer Net params θ.
Initialize baseline value according to heuristic algorithm.

for t = *1 to T* **do**
    Select a batch of sample s_i for $i \in \{1, \cdots, B\}$.
    Sample solution o_i based on $p_\theta(\cdot|s_i)$ for $i \in \{1, \cdots, B\}$.
    Let $g_\theta = \frac{1}{B}\sum_{i=1}^{B}[(SA(o_i|s_i) - b(s_i))\nabla_\theta log p_{\boldsymbol{\theta}}(o_i|s_i)]$.

    Update $\theta = ADAM(\theta, g_\theta)$.
    Update baseline $b(s_i) = b(s_i) + \alpha(SA(o_i|s_i) - b(s_i))$ for $i \in \{1, \cdots, B\}$.
end for
return pointer net parameters θ.

训练过程

采用Policy Gradient策略，Batch Size = 16, Optimizer = Adam, Initial Learning Rate = 1e-2, 每30轮按照Decay Factory=0.96衰减，训练238轮，优化预测网络

实验结果与分析

Experimental results and analysis

训练集合

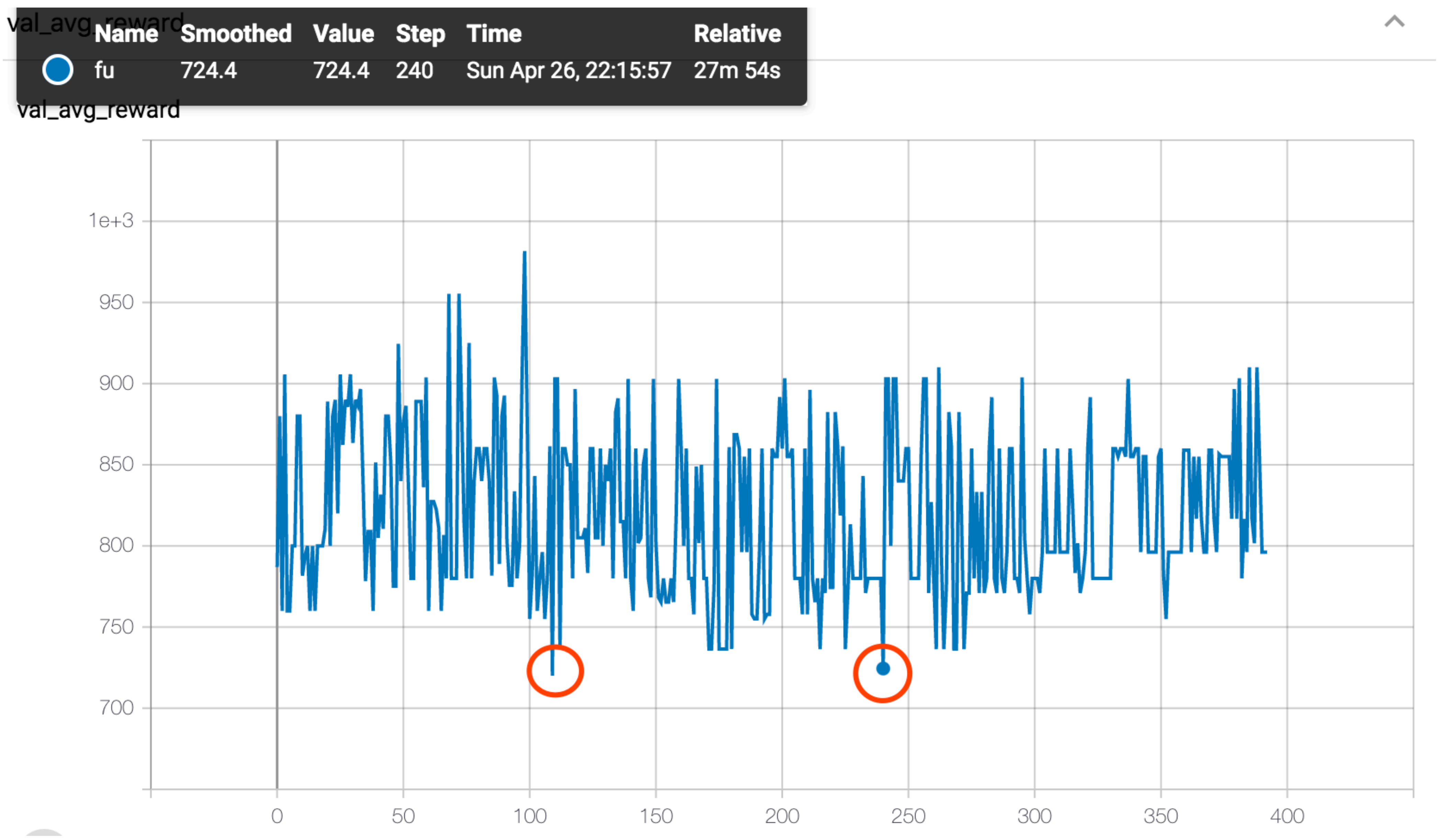
测试集合

训练集和数据集反馈均持续下降

1. 可以看到在三百轮以内，训练集合和测试集都是持续下降，即预测效果提升，未发现过拟合情况。
2. 随机选择验证样本，各种启发式标准获得高度高于平均期望高度，即能够预测出高于启发式标准的结果。
3. 部分训练集会使得验证集长时间陷入局部最优，具体的实验结果比较依赖训练集本身。
4. 一定程度上证明了神经网络可以像人一样，根据形状的情况，预测一个合理的初始序列。

实验结果与分析

Experimental results and analysis



具体数据集测试

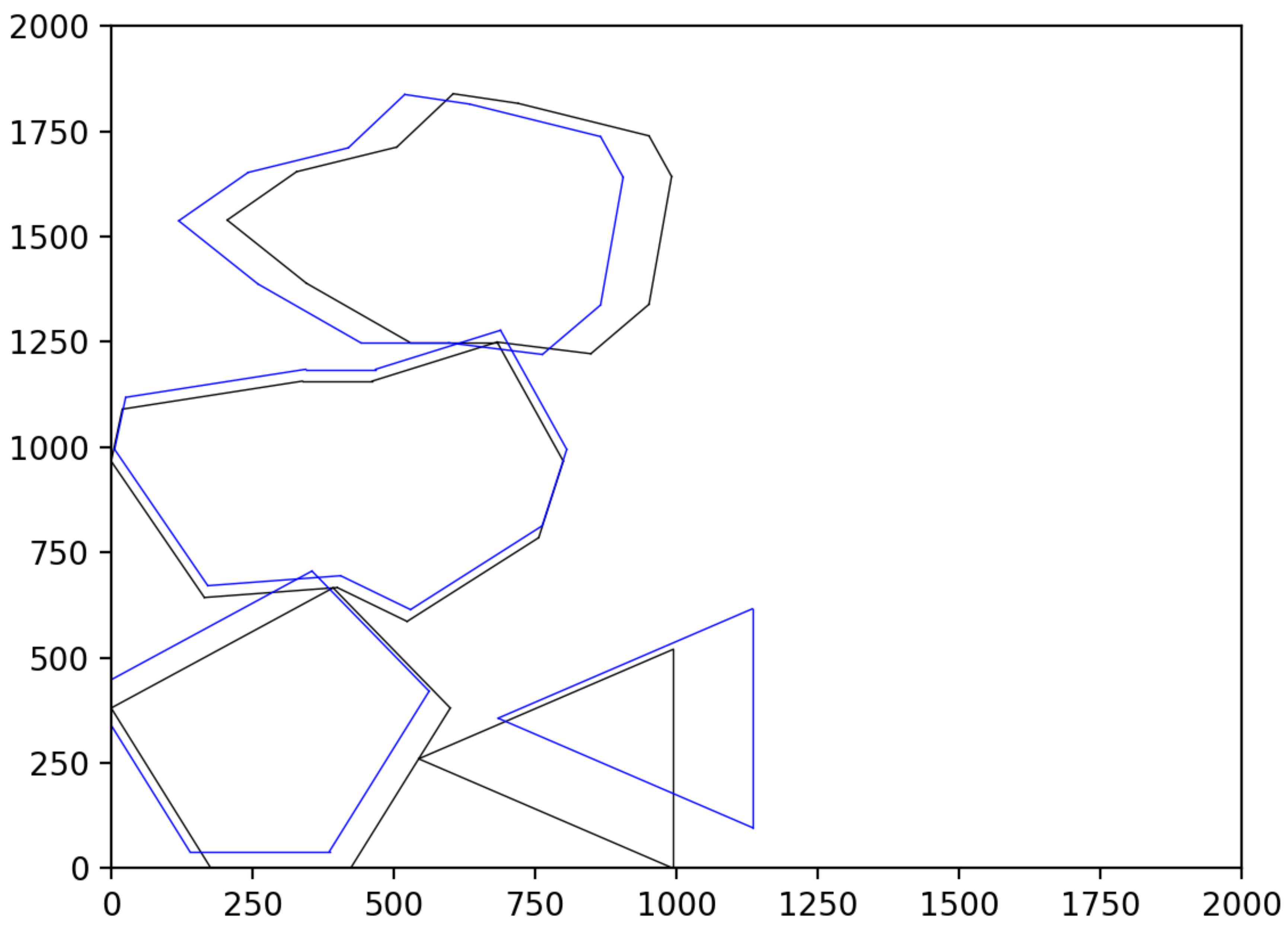
在数据集Fu上，不同轮次的训练模型的预测结果有较大的波动，但是有接近一半的结果是优于启发式排样的，并且有个别的排样结果比较接近GA/SA优化后的结果

综合数据集情况

在有限的训练次数内，我们暂时没有发现进入过度拟合的情况，预测结果的排样高度在持续下降，考虑到排样问题本身比较复杂，我们未来可以通过增加网络复杂度和训练轮数，来增加网络的学习能力

实验结论与展望

Experimental conclusion and prospect



根据序列预测位置[测试]

输入嵌入向量的多个形状，输出左底部排样的结果，对LSTM网络进行训练后，我们可以获得如上结果，即LSTM可以基于形状输入对位置进行映射

结论分析

1. 可以通过序列网络学习排样序列

排样序列作为序列问题，只要对形状进行嵌入，就可以通过强化学习和序列网络进行学习，而且未来提升空间不小

2. 预测模型对单个数据预测结果不太稳定

预测结果整体下降，对单个数据集而言可降可升，但是可以通过增加网络复杂度、提升学习能力来解决该问题

未来展望

1. 增加数据集推广性[当前目标]

序列预测效果比较依赖输入数据集的规律性和输出与输入数据的相关性，需要研发出相对通用的学习模型

2. 混合预测模型

根据对直接预测相对位置的实验结果分析[原理：位置映射]，可能可以尝试搭建预测排样优化结果的混合模型

参考文献

Reference

文献综述

排样问题的几何原理、概述以及分类

[1] Bennell J A, Oliveira J F. The geometry of nesting problems: A tutorial[J]. European Journal of Operational Research, 2008, 184(2): 397-415.

[2] Bennell J A, Oliveira J F. A tutorial in irregular shape packing problems[J]. Journal of the Operational Research Society, 2009, 60(1): S93-S105.

[3] Wäscher G, Haußner H, Schumann H. An improved typology of cutting and packing problems[J]. European journal of operational research, 2007, 183(3): 1109-1130.

启发式排样算法

采用左底部算法或者是重心最低算法获取初始解

[4] Liu H, He Y. Algorithm for 2D irregular-shaped nesting problem based on the NFP algorithm and lowest-gravity-center principle[J]. Journal of Zhejiang University-Science A, 2006, 7(4): 570-576.

[5] Dowsland K A, Vaid S, Dowsland W B. An algorithm for polygon placement using a bottom-left strategy[J]. European Journal of Operational Research, 2002, 141(2): 371-381.

序列优化算法

一般通过遗传算法或者上模拟退火算法进行序列优化

[6] Babu A R, Babu N R. A generic approach for nesting of 2-D parts in 2-D sheets using genetic and heuristic algorithms[J]. Computer-Aided Design, 2001, 33(12): 879-891.

参考文献

Reference

同濟大學 TONGJI UNIVERSITY

组合优化+深度学习

序列到序列的学习、路径规划、图模型嵌入与学习

[7] Vinyals O, Fortunato M, Jaitly N. Pointer networks[C]//Advances in neural information processing systems. 2015: 2692-2700.

[8] Khalil E, Dai H, Zhang Y, et al. Learning combinatorial optimization algorithms over graphs[C]//Advances in Neural Information Processing Systems. 2017: 6348-6358.

[9] Kool W, Van Hoof H, Welling M. Attention, learn to solve routing problems![J]. arXiv preprint arXiv:1803.08475, 2018.

形状嵌入算法

将形状嵌入到一维向量中

[10] Zhang D, Lu G. Review of shape representation and description techniques[J]. Pattern recognition, 2004, 37(1): 1-19.

强化学习+组合优化

通过强化学习解决三维装箱、路径规划问题等问题

[11] Bello I, Pham H, Le Q V, et al. Neural combinatorial optimization with reinforcement learning[J]. arXiv preprint arXiv:1611.09940, 2016.

[12] Nazari M, Oroojlooy A, Snyder L, et al. Reinforcement learning for solving the vehicle routing problem[C]//Advances in Neural Information Processing Systems. 2018: 9839-9849.

[13] Hu H, Zhang X, Yan X, et al. Solving a new 3d bin packing problem with deep reinforcement learning method[J]. arXiv preprint arXiv:1708.05930, 2017.