

基于迁移学习的泥水平衡盾构隧道开挖面稳定性预测方法研究

马志刚¹, 严靖轲²

¹上海隧道工程有限公司, 上海

²同济大学土木工程学院, 上海

收稿日期: 2026年2月11日; 录用日期: 2026年3月4日; 发布日期: 2026年3月13日

摘要

为解决传统开挖面稳定性评估方法效率低、难以应对突发情况的难题, 以排浆密度、进浆密度为依据, 提出了一种基于迁移学习的开挖面稳定性预测方法。相比于理论计算与数值模拟, 所提出的基于Attn-LSTM的智能预测模型能够从监测数据中学习不同地层条件下掘进参数与泥浆密度间的关系, 且在复杂地层条件下也具备较高的预测精度。利用中俄东线天然气管道工程监测数据的验证结果表明, 该方法对排浆密度与进浆密度的预测值与真实值相比仅有0.1858%的平均绝对百分比误差。最后, 基于迁移学习方法, 在沪通铁路吴淞口长江隧道站前IV标工程监测数据上验证了提出方法的有效性。

关键词

开挖面稳定性, 迁移学习, 神经网络, 泥浆密度

Research on Transfer Learning Based Prediction Method for Face Stability of Slurry Pressure Balance Shield Tunnel

Zhigang Ma¹, Jingke Yan²

¹Shanghai Tunnel Engineering Co., Ltd., Shanghai

²College of Civil Engineering, Tongji University, Shanghai

Received: February 11, 2026; accepted: March 4, 2026; published: March 13, 2026

Abstract

To address the inefficiency of traditional methods for evaluating face stability and their difficulty in handling unexpected situations, a prediction method for face stability based on transfer learning is

proposed, utilizing return slurry density and feed slurry density. Compared with theoretical calculations and numerical simulations, the proposed intelligent prediction model based on Attn-LSTM can learn the relationship between tunneling parameters and slurry density under different geological conditions from monitoring data, and exhibits high prediction accuracy even in complex strata. Validation using monitoring data from the China-Russia Eastern Natural Gas Pipeline Project shows that the average absolute percentage error between the predicted and actual values of discharged slurry density and injected slurry density is only 0.1858%. Finally, the effectiveness of the proposed method is verified on monitoring data from Section IV of the Wusongkou Yangtze River Tunnel station of the Shanghai-Nantong Railway Project using a transfer learning approach.

Keywords

Face Stability, Transfer Learning, Neural Network, Slurry Density

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着复杂地层中隧道建设的需求不断增加,泥水平衡盾构机(SPB)因为较好的地层适应性,被广泛用于复杂地层中隧道的建设[1]。泥水平衡盾构通过带压泥浆来平衡开挖面前方的水土压力。泥水平衡盾构推进过程中,泥浆在泥水舱内泥浆压力和开挖面前方水压力差的作用下会沿土体颗粒间孔隙向地层渗透形成泥膜。泥膜封闭渗流通道同时也为支护压力提供了有效的作用面。盾构推进时,呈现泥膜被刀盘切削后又因泥浆渗透迅速形成的动态循环[2]。由此可见,泥水平衡盾构隧道开挖面稳定性是一个极其复杂的问题。如果泥水支护压力过小,开挖面前方土体会发生主动失稳,造成地表沉降;如果泥水支护压力过大,开挖面前方土体会发生被动失稳导致地面隆起或地表冒泥冒浆。近年来因开挖面失稳造成地层劈裂冒浆、掌子面塌方江海水倒灌等重大工程事故屡有发生。如南京长江纬三路过江通道建设过程中发生的开挖面失稳冒浆险情,荷兰 Heinenroord 第二隧道发生的塌方进水事故等。

传统的开挖面稳定性评估方法包括理论分析与数值模拟[3]-[5],这两种方法都难以反映泥浆与地层间的相互作用。理论分析和数值模拟针对的是确定的工况,无法对盾构推进过程中开挖面的稳定性进行实时分析。数值模拟方法相当耗时,难以对突发情况做出指导。隧道开挖过程中能采集到大量的传感器数据,且隧道工程中的问题往往需要考虑多方面的因素。基于机器学习、深度学习方法通过学习数据间的联系进行预测,不需要任何假设且训练完成以后运行速度极快,已应用于包括盾构隧道掘进参数预测[6][7]、最大沉降预测[8]、地层条件识别[9]等方面。深度学习更适用于解决高维的复杂问题,在经过大量数据的训练后其表现比机器学习方法更好,因此深度学习更受关注[10]。

开挖面稳定性评估作为隧道工程的一大难题,其错误评估会给施工带来极大的困扰甚至人员伤亡。但是,目前开挖面稳定性的研究集中于理论分析与数值模拟,并不能提前对开挖面稳定性进行感知。面对施工过程中的突发情况,这些方法也无法及时给出应对策略。

针对上述问题,本文提出了一种基于注意力机制与 LSTM 的深度学习神经网络模型。该模型通过注意力机制与 LSTM 学习数据不同特征维度(即不同盾构掘进参数)和时间上的联系,且可在实际应用中通过迁移学习进行微调以适用新工程。随后,本文通过对模型进行修改建立了另外三个模型: Attn-Attn-LSTM、

Attn-GRU-Attn-LSTM、LSTM-Attn-LSTM, 以讨论不同神经网络层在预测过程中的作用。最后, 在沪通铁路吴淞口长江隧道站前IV标工程验证了模型的有效性。

2. 基于 Attn-LSTM 的智能预测模型

2.1. 长短时记忆神经网络(LSTM)

人脑中除了注意力能够帮助处理信息外, 记忆力在信息处理中也能发挥很大作用。人脑通常会记住和忘记一些事情。基于此, 科学家们提出了长短时记忆神经网络(LSTM) [10]。在时间序列预测上面, 通过“遗忘”与“记住”一些信息, LSTM对时间效应的把控也有更加精准。本文研究的时间序列问题时间跨度大, 因此, 借助 LSTM“遗忘”掉与当前时间步无关的历史数据, 可以更好地提高预测的精度。

LSTM 通过遗忘门、输入门、输出门进行信息的控制与传递。当一个信息传到 LSTM 单元时, 遗忘门、输入门、输出门会对其进行处理, 分别判断该时刻应该保留多少历史信息、输入多少当前信息以及有多少信息输出。以第 t 时间步的输入 \mathbf{x}_t 为例, 遗忘门、输入门、输出门的计算方式分别为:

$$\mathbf{f}_t = \sigma(\mathbf{W}_f [\mathbf{h}_{t-1}; \mathbf{x}_t] + \mathbf{b}_f) \quad (1)$$

$$\mathbf{i}_t = \sigma(\mathbf{W}_i [\mathbf{h}_{t-1}; \mathbf{x}_t] + \mathbf{b}_i) \quad (2)$$

$$\mathbf{o}_t = \sigma(\mathbf{W}_o [\mathbf{h}_{t-1}; \mathbf{x}_t] + \mathbf{b}_o) \quad (3)$$

其中, $\sigma()$ 为 Logistic 函数, \mathbf{h}_{t-1} 为上一时刻的隐藏状态; \mathbf{W} , \mathbf{U} , \mathbf{b} 为可学习的参数; $[\mathbf{h}_{t-1}; \mathbf{x}_t]$ 表示矩阵拼接。LSTM 单元中还有候选状态、记忆单元以及隐藏状态, 在确定了遗忘门、输入门、输出门后, 三者可分别由式(4)~(6)计算:

$$\tilde{\mathbf{c}}_t = \tanh(\mathbf{W}_c [\mathbf{h}_{t-1}; \mathbf{x}_t] + \mathbf{b}_c) \quad (4)$$

$$\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \tilde{\mathbf{c}}_t \quad (5)$$

$$\mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{c}_t) \quad (6)$$

从式(1)~(6)中可以看出, 不同时间步中 LSTM 神经元的信息由 \mathbf{h}_t , \mathbf{c}_t 传递。同时, 不同时刻的 \mathbf{h}_t 也是每个时间步的输出。

2.2. 注意力机制

动物的各个感官系统接收了大量的信息, 但是这些信息受到的关注程度是不同的。动物的注意力能够让其将只关注最有价值的部分信息[11]。受此影响, 深度学习领域的学者们提出了注意力机制(attention mechanism), 用于选择关键信息, 提高神经网络效率[12]。通俗来讲, 注意力机制将一些权重赋予不同向量, 从而实现信息的选择。

例如对于输入向量 $[\mathbf{x}_1, \dots, \mathbf{x}_N]$, 为了从中对关键信息施加较高的注意力, 需要引入一个“提示”, 也就是与任务相关的查询向量 \mathbf{q} 。通过打分函数可以计算每个输入向量与 \mathbf{q} 之间的相关性。那么, 每个向量的注意力权重可以由式(7)计算:

$$\alpha_n = \frac{\exp(s(\mathbf{x}_n, \mathbf{q}))}{\sum_{j=1}^N \exp(s(\mathbf{x}_j, \mathbf{q}))} \quad (7)$$

$$s(\mathbf{x}, \mathbf{q}) = \mathbf{v}^T \tanh(\mathbf{W}\mathbf{x} + \mathbf{U}\mathbf{q}) \quad (8)$$

其中 α_n 称为注意力分布, $s()$ 为加性注意力打分函数。当计算完所有向量的注意力权重后, 对所有向量

计算加权平均得到最终输出。

2.3. 模型架构

Hübner 等人的研究表明将注意力机制分为两阶段建模能得到较好地结果[13]。因此本文采用的模型架构如图 1 所示。该模型由编码器和解码器两部分组成，两部分均基于注意力机制与 LSTM(后文简称该模型为 Attn-LSTM)。编码器中，我们参考 Qin 等人提出的注意力机制[14]，结合 LSTM 对输入序列的不同时间步、不同维度的特征予以权重，实现关键信息的选择。解码器采用类似的方式进行，不过解码器采用了时间注意力机制进行编码器隐藏状态的选择。

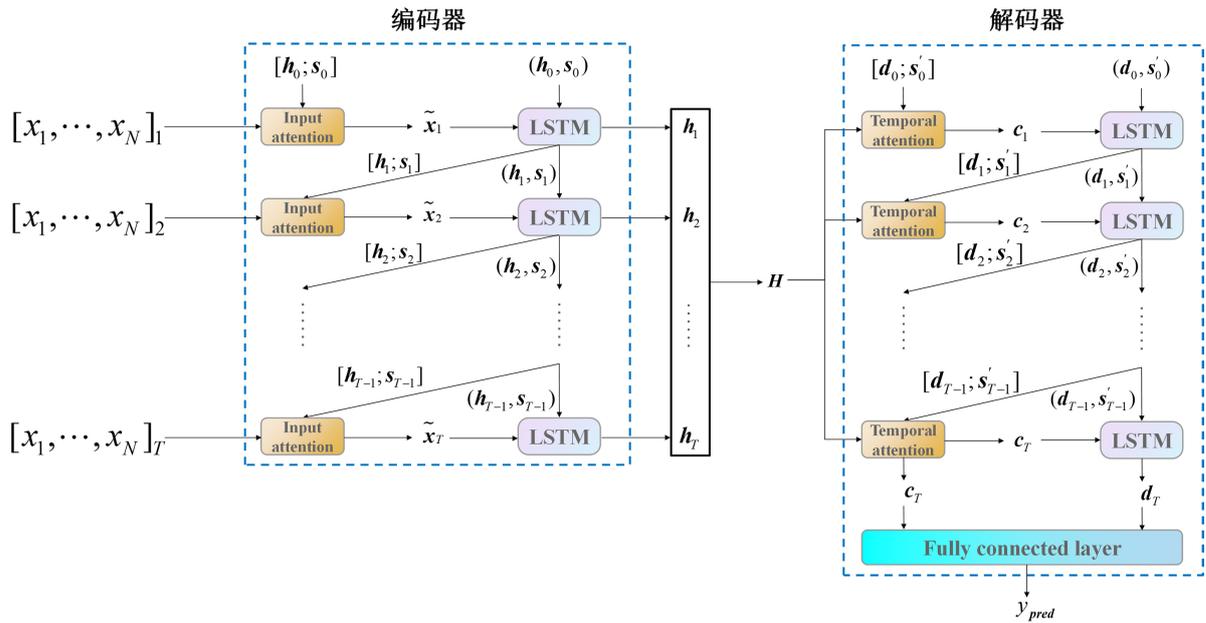


Figure 1. Deep neural network architecture based on Attn-LSTM
图 1. 基于 Attn-LSTM 的深度神经网络架构

2.4. 模型评价体系

本文选取回归问题中常用的评价指标均方误差(MSE)、均方根误差(RMSE)以及平均相对百分比误差(MAPE)对最终预测结果进行评价。三个指标的值越小，代表着预测值与真实值越接近，模型精度越高。MSE、RMSE 以及 MAPE 可以分别由下式计算：

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_{true}^i - y_{pred}^i)^2 \quad (9)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{true}^i - y_{pred}^i)^2} \quad (10)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_{true}^i - y_{pred}^i}{y_{true}^i} \right| \times 100\% \quad (11)$$

其中 n 为预测值的个数， y_{true}^i 为真实值， y_{pred}^i 为预测值。三者均用于衡量预测值与真实值之间的差距，但各有不同：

MSE 表示真实值与预测值之间距离的平方，其常作为模型训练中的损失函数，不难看出，二者之间

的差距越大, 损失函数对其的“惩罚”比 RMSE 和 MAPE 更多。

RMSE 将量纲进行了统一, 反映了真实值与预测值之间的平均绝对误差。

MAPE 表示绝对误差的百分比, 能直观的反映模型的预测精度。

3. 工程实例

3.1. 工程概况

中俄东线天然气管道工程(永清 - 上海)在江苏省南通市与常熟市之间将相继穿越新江海河、常熟港海轮锚地、主航道、专用航道、白茆河及其之间区域, 穿越长度约为 10.23 km, 穿越水面宽度约为 7.5 km。隧道直径为 7.60 m。该工程目前已完成推进, 有丰富的训练数据。

沪通铁路吴淞口长江隧道站前IV标工程, 正线迄止里程为 DK140+700~DK144+210, 线路长度 2.905 km, 位于上海市宝山区境内, 包括 1 号工作井、明挖暗埋段、陆域段隧道, 总长约 2.28 km, 隧道直径为 10.69 m。该工程目前正在进行, 可用于迁移学习和模型验证。

在施工过程中, 为了确保盾构机顺利运行, 盾构机内部存在大量的传感器以一定频率实时记录盾构掘进参数。记录的参数包括刀盘转速、推进压力、排浆流量、注浆压力等多个参数。在开挖过程中, 驾驶员将根据记录的历史数据对盾构机掘进参数进行调整, 以确保盾构机的正常行驶。

实际工程中以采样频率为单位进行预测并无太大意义, 不同工程的采样频率也不一致, 本文对原始数据进行了 30 s 的均值重采样。需要注意的是, 重采样后的数据仍存在大量的停机数据以及一定比例的异常数据。在数据驱动的机器学习/深度学习方法中, 这些数据会严重降低模型的性能, 因此, 本文在记录的数据中删除了停机数据并从中提取了有效掘进循环构成数据集。

3.2. 输入参数选取

在建立开挖面稳定性智能预测模型之前, 需要确实如何判断开挖面是否失稳。虽然学者们针对开挖面稳定性问题做了很多理论研究。但是在预测方面, 用什么参数预测? 怎样通过预测的参数判断稳定性? 目前还没有理论与工程实例。

结合盾构机的实际开挖过程, 我们发现当开挖面失稳时, 前方土体会产生塌陷。一方面, 这些土体落入盾构机内部, 会影响泥水舱中泥浆的密度, 造成进浆密度与排浆密度的出现较大差异。另一方面, 一些土渣与传感器产生碰撞也会造成测量值突变。因此, 本文将排浆密度与进浆密度确定为开挖面稳定性预测指标。

如 2.1 节中所述, 记录的数据中包含了 936 个掘进参数。但是, 同时将除输出参数之外的所有参数作为特征输入数据驱动模型, 往往会导致“维数灾难”。为了解决上述问题, 本文通过文献调研初选了 14 个参数, 并采用皮尔逊相关系数计算不同特征间的相关系数, 其计算公式与结果分别如式(15)和图 2 示。

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (12)$$

其中 x_t , y_t 表示不同序列的第 t 个值; \bar{x} 和 \bar{y} 表示不同序列的均值; n 表示序列长度。

需要注意的是, 虽然地勘报告中记录了不同桩号下样本的压缩模量、粘聚力以及内摩擦角, 但这样记录的数据离散性大, 且隧道的同一截面内存在多组性质不同的样本。同时, 随着隧道的开挖, 盾构机对地层不断产生扰动, 开挖前测得的数据不一定符合实际工程情况。本文的依托工程在盾构机掘进过程中实时记录了贯入度。该参数具有实时性, 能够准确地反映地层在开挖时的性质, 因此, 本文将采用贯入度这一参数作为反应地层性质的输入参数。

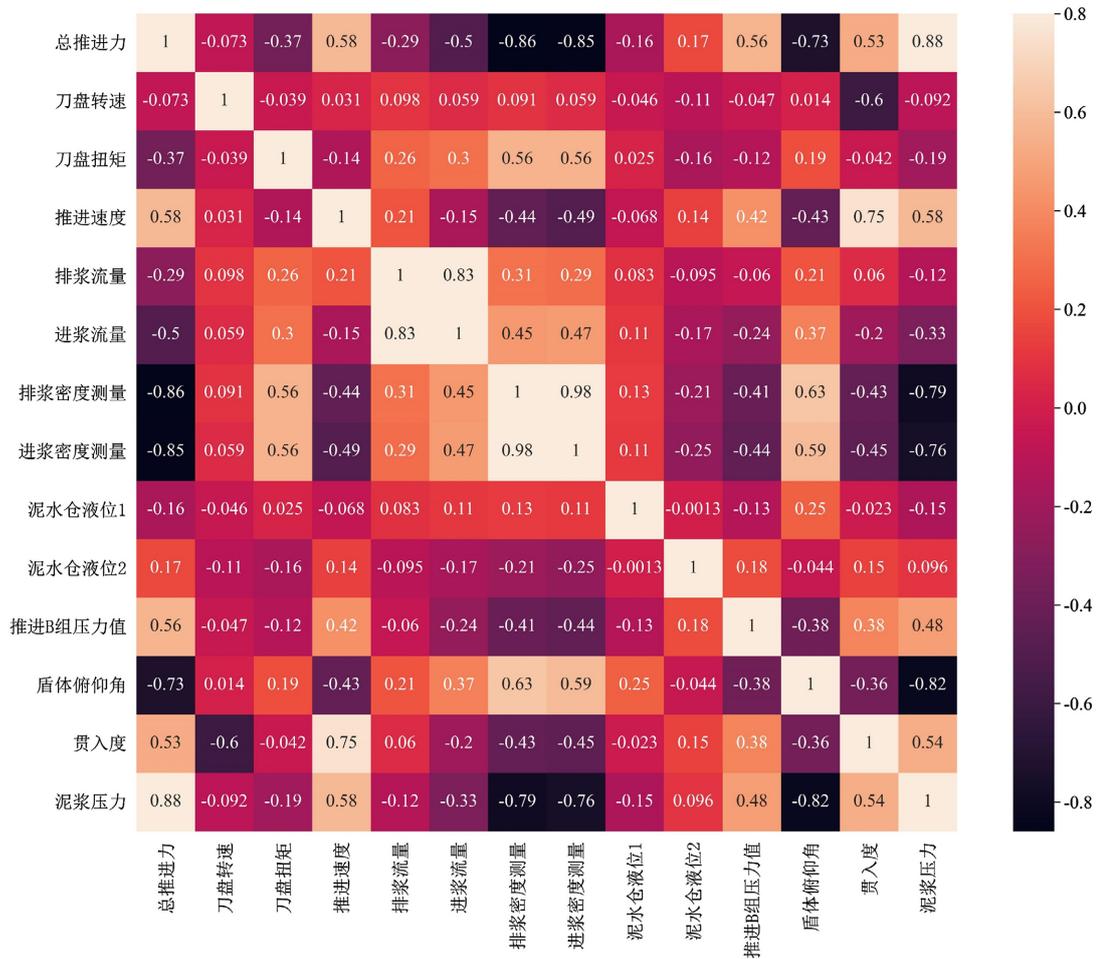


Figure 2. Pearson correlation coefficient heatmap between selected parameters
图 2. 所选参数间的皮尔逊相关系数热力图

在所选的 14 个参数中，总推进力、泥浆压力两个参数与排浆密度、进浆密度的皮尔逊相关系数远大于 0.6。参考 Liu 等提出的阈值[9]，本文排除了总推进力与泥浆压力两个参数以避免过拟合。最后选择的输入参数、输出参数以及二者相关统计指标如表 1 所示。表 1 中的参数组成如下：

Table 1. Input and output parameters with their statistical characteristic indices

表 1. 输入与输出参数及相关统计指标

掘进参数	单位	最大值	最小值	平均值	标准差
刀盘转速	rpm	6.00	0	0.25	0.42
刀盘扭矩	kN·m	9957.21	-225.28	263.73	457.84
推进速度	mm/min	129.36	0	13.93	24.09
排浆流量	m ³ /h	2000.00	0	425.61	551.73
进浆流量	m ³ /h	2000.00	0	444.63	546.02
排浆密度测量	t/m ³	1.50	1.00	1.25	0.18
进浆密度测量	t/m ³	1.50	1.00	1.27	0.18

续表

泥水舱液位 1	m	92.89	-89.37	2.14	15.11
泥水舱液位 2	m	2.29	-2.00	-0.03	-0.69
推进 B 组压力值	bar	400.00	0	129.17	68.09
盾体俯仰角	mm/m	5.00	-5.00	-0.34	0.99
贯入度	mm	145.18	0	15.50	27.12

直接反映开挖面稳定性的盾构掘进参数：刀盘转速、刀盘扭矩、推进速度、推进 B 组压力值。

间接反映开挖面稳定性和地层性质的泥浆相关参数：泥水仓液位、进排浆流量、进排浆密度。

盾构机姿态相关的盾构掘进参数：盾体俯仰角反应地层性质的参数：贯入度。

综上，输入模型的 10 个参数能够全面地反映盾构机掘进过程中开挖面稳定的相关情况。

4. 模型训练与预测

4.1. 数据集划分

同 2.1 节所述，本文参考 Zhang 等[8]的方法对数据集进行清洗。之后，以 8:1:1 的比例划分为训练集、验证集和测试集，分别用于模型的训练、超参数的选择以及模型性能的评估。由于输入参数的数量级有较大的差异(表 1)，直接输入可能导致模型训练过程中损失函数收敛较慢甚至无法收敛。因此，本文在特征(即不同输入参数)维度对数据进行归一化：

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (13)$$

其中 x_{\min} 和 x_{\max} 分别代表某一输入参数所有数据的最小值与最大值。

4.2. 模型训练策略

本文采用的模型超参数优化方法为随机搜索算法。根据进行的文献调研[15][16]，将超参数限制在一定范围内。此时的所有超参数形成了一个搜索空间(表 2)，本文在该搜索空间中进行一定次数的随机搜索，实现最优超参数的选择。

Table 2. Hyperparameter searching space

表 2. 超参数搜索空间

超参数类型	范围
时间窗口长度	[20, 40, 60]
初始学习率	[0.0001, 0.0002, 0.001]
批量大小	[32, 64, 128]
编码器隐藏单元数	[16, 32, 64, 128]
解码器隐藏单元数	[16, 32, 64, 128]

搜索过程中，以模型在验证集上的 MSE 为评价指标进行超参数的更新，最后保留的超参数即为最优超参数。另外，在训练过程中，引入了学习率衰减，即，在训练过程迭代一定次数后将学习率 $\times 0.9$ ，使得模型的更新在后期没有太大波动从而更加接近最优解[17]。

为了对比结构性能, 本节还通过对 Attn-LSTM 模型结构的更改建立了另外三个神经网络模型: 编码器不使用注意力机制的 LSTM-Attn-LSTM 模型、编码器只使用注意力机制的 Attn-Attn-LSTM 模型以及将编码器中的 LSTM 替换为 GRU 的 Attn-GRU-Attn-LSTM 模型。他们的超参数搜索空间均为表 2 所示。需要注意的是, 所有模型都采用了 Adam 优化器, 最终得到的最佳组合如表 3 所示。

Table 3. The optimal combination of hyperparameters
表 3. 超参数最佳组合

模型类型	超参数类型	值
Attn-LSTM	时间窗口长度	40
	初始学习率	0.0001
	批量大小	64
	编码器隐藏单元数	64
	解码器隐藏单元数	128
LSTM-Attn-LSTM	时间窗口长度	40
	初始学习率	0.0001
	批量大小	64
	编码器隐藏单元数	16
	解码器隐藏单元数	32
Attn-Attn-LSTM	时间窗口长度	40
	初始学习率	0.0001
	批量大小	64
	编码器隐藏单元数	-
	解码器隐藏单元数	32
Attn-GRU-Attn-LSTM	时间窗口长度	40
	初始学习率	0.0001
	批量大小	64
	编码器隐藏单元数	64
	解码器隐藏单元数	128

需要注意的是, 该训练策略仅在中俄东线天然气管道工程数据集上用于智能预测模型的训练。在实际应用(沪通铁路吴淞口长江隧道站前IV标工程)中, 应对训练完成的模型, 使用迁移学习方法进行微调。

4.3. 模型预测结果分析

图 3 展示了 Attn-LSTM 模型的排浆密度与进浆密度预测结果, 从图中可以看出 Attn-LSTM 模型能够精准地预测二者的变化趋势。在时间序列图中真实值出现波动甚至剧烈波动的时候, Attn-LSTM 模型仍能精确预测。频数分布直方图表示了 0.01 t/m^3 的间隔内预测值与真实值的差异。不难看出, 二者的频数分布直方图基本重合, 表现出极高的精度。排浆密度真实值和预测值的散点图基本围绕着最佳预测线在极小区域内分布。但是进浆密度真实值与预测值的散点图却出现了较大波动, 结合该参数时间序列图与频数分布直方图分析, 可能的原因是该参数突变的地方较多, 导致模型在无法完美捕捉所有突变处的变化趋势。尽管如此, 该模型在训练集、测试集上仍表现出了较高的预测精度。

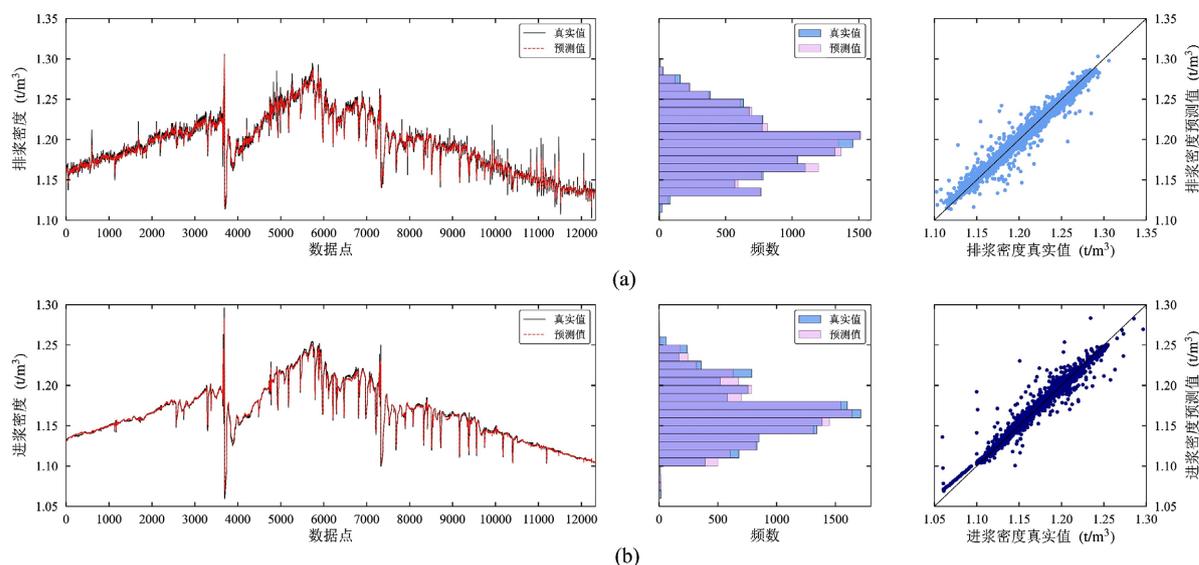


Figure 3. Comparison between true values and values predicted by Attn-LSTM in the test set

图 3. 测试集中真实值与 Attn-LSTM 预测值对比: (a) 排浆密度; (b) 进浆密度

为了进一步验证该模型的性能, 还对 4.2 节提出的另外三个模型在同样数据集上进行训练与预测。为了更直观的提现四个模型性能的优劣, 将他们在训练集与测试集上的评价指标汇总到表 4、表 5 中。

Table 4. Evaluation indicators of different models on train set

表 4. 不同模型在训练集上的评价指标

模型	RMSE	MAPE
Attn-LSTM	1.7223×10^{-2}	0.8430%
LSTM-Attn-LSTM	4.0221×10^{-2}	2.4248%
Attn-GRU-Attn-LSTM	2.5389×10^{-2}	1.3419%
Attn-Attn-LSTM	2.0802×10^{-2}	1.0112%

Table 5. Evaluation indicators of different models on test set

表 5. 不同模型在测试集上的评价指标

模型	RMSE	MAPE
Attn-LSTM	3.9194×10^{-3}	0.1858%
LSTM-Attn-LSTM	2.1464×10^{-2}	1.6729%
Attn-GRU-Attn-LSTM	6.8545×10^{-3}	0.4235%
Attn-Attn-LSTM	6.3790×10^{-3}	0.4230%

4.4. 模型应用

将该模型用于沪通铁路吴淞口长江隧道站前 IV 标工程以检验所得最优智能预测模型(Attn-LSTM)的泛化性能。需要注意的是, 4.2 节中的模型训练过程中并未使用该工程的数据。为了更全面地考察模型的泛化性能, 分别使用来自新数据集 0 环、50 环、100 环、200 环的数据进行微调, 微调后的误差如图 4 所示。可以看到, 随着来自新数据集的数据增加, 预测精度逐渐提高, 这是因为智能模型通过掘进参数直

接和间接地学习到了地层信息和施工工法信息。从定量数据来看, 数据的加入将进排浆密度预测值的 RMSE 从 0.131 g/cm^3 降低到了 0.074 g/cm^3 , MAPE 从 10.28% 降低到了 5.62%, 精度提升了约 50%, 有效验证了模型的适用性。

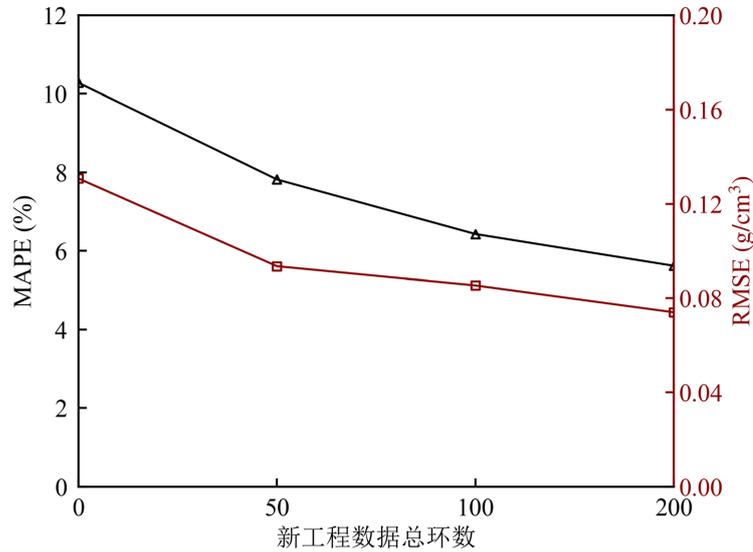


Figure 4. Application effectiveness of predictive models in new engineering projects
图 4. 预测模型在新工程中的应用效果

仅使用新工程数据按 4.2 节的策略对模型进行重新训练, 得到的结果与微调后的结果总结到了表 6 中。由于新工程数据量小, 模型在小样本数据中无法学习到完整的进排浆密度与掘进参数之间的联系, 因此模型性能相对较差。另外, 新工程的地表存在既有建筑, 需要根据地表沉降调整运行策略, 受人为因素的影响大, 而中俄东线输气工程属于海底隧道, 地表沉降对其影响较小, 主要考虑开挖面稳定性, 人为因素影响小, 因此模型在新工程上的表现较差。从表 6 中可以看出, 尽管未经迁移学习的模型预测精度略低于重新训练的模型, 但随着新数据的加入, 模型在迁移学习过程中学到了更多不同掘进参数与进排浆密度的联系, 模型的预测精度逐渐提高。

Table 6. Comparative analysis of model transfer performance
表 6. 模型迁移性能对比

模型	RMSE	MAPE
仅新工程数据 400 环	0.116 g/cm^3	9.37%
新工程数据 0 环微调	0.131 g/cm^3	10.28%
新工程数据 50 环微调	0.094 g/cm^3	7.82%
新工程数据 100 环微调	0.085 g/cm^3	6.43%
新工程数据 200 环微调	0.074 g/cm^3	5.62%

5. 结论

本文针对开挖面稳定性预测问题, 分析了反映开挖面稳定性的掘进参数, 构建了基于 Attn-LSTM 的深度神经网络智能预测模型, 并在中俄东线天然气管道工程以及沪通铁路吴淞口长江隧道站前 IV 标工程

中得到了验证, 主要结论如下:

1) 采用的模型在测试集中的 MAPE 达到了 0.1858%, 相比于 Attn-Attn-LSTM、LSTM-Attn-LSTM、Attn-GRU-Attn-LSTM 三种模型分别下降了 56.08%、88.89%、56.13%。这说明本文采用的模型对进浆密度、排浆密度的预测优于另外三种模型。

2) 从训练集真实值曲线与预测值曲线的变化趋势来看, Attn-LSTM 对时间效应的把握优于其他模型。LSTM-Attn-LSTM 模型虽然精度在四个模型中最差, 但是对时间效应的把握优于 Attn-GRU-Attn-LSTM 和 Attn-Attn-LSTM 模型起。

3) 本文建立的 Attn-LSTM 模型有良好的适用性, 尽管在完全未知的新工程中模型预测结果会有较大误差, 但随着相关数据的加入, 模型的预测误差会大幅提升, 说明该智能模型有良好的学习性能。

参考文献

- [1] 冯爱军. 中国城市轨道交通 2021 年数据统计与发展分析[J]. 隧道建设(中英文), 2022, 42(2): 336-341.
- [2] 尹鑫晟. 泥水盾构成膜规律及开挖面稳定性[D]: [博士学位论文]. 杭州: 浙江大学, 2017.
- [3] Anagnostou, G. and Kovári, K. (1994) The Face Stability of Slurry-Shield-Driven Tunnels. *Tunnelling and Underground Space Technology*, **9**, 165-174. [https://doi.org/10.1016/0886-7798\(94\)90028-0](https://doi.org/10.1016/0886-7798(94)90028-0)
- [4] Mollon, G., Dias, D. and Soubra, A. (2009) Probabilistic Analysis and Design of Circular Tunnels against Face Stability. *International Journal of Geomechanics*, **9**, 237-249. [https://doi.org/10.1061/\(asce\)1532-3641\(2009\)9:6\(237\)](https://doi.org/10.1061/(asce)1532-3641(2009)9:6(237))
- [5] Zeng, S., Lü, X. and Huang, M. (2019) Discrete Element Modeling of Static Liquefaction of Shield Tunnel Face in Saturated Sand. *Acta Geotechnica*, **14**, 1643-1652. <https://doi.org/10.1007/s11440-019-00806-w>
- [6] 高昆, 于思溟, 许维青, 等. 基于 Attention-ResNet-LSTM 混合神经网络的盾构掘进速度预测新方法[J]. 隧道建设(中英文), 2023, 43(4): 592-601.
- [7] 郑永光, 张娜, 刘扬扬, 等. 基于多模态控制策略的 TBM 掘进参数预测模型研究[J]. 隧道建设(中英文), 2023, 43(4): 583-591.
- [8] Zhang, P., Wu, H., Chen, R., Dai, T., Meng, F. and Wang, H. (2020) A Critical Evaluation of Machine Learning and Deep Learning in Shield-Ground Interaction Prediction. *Tunnelling and Underground Space Technology*, **106**, Article ID: 103593. <https://doi.org/10.1016/j.tust.2020.103593>
- [9] Liu, Z., Li, L., Fang, X., Qi, W., Shen, J., Zhou, H., et al. (2021) Hard-Rock Tunnel Lithology Prediction with TBM Construction Big Data Using a Global-Attention-Mechanism-Based LSTM Network. *Automation in Construction*, **125**, Article ID: 103647. <https://doi.org/10.1016/j.autcon.2021.103647>
- [10] Greff, K., Srivastava, R.K., Koutnik, J., Steunebrink, B.R. and Schmidhuber, J. (2017) LSTM: A Search Space Odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, **28**, 2222-2232. <https://doi.org/10.1109/tnnls.2016.2582924>
- [11] Zhang, A., Lipton, Z.C., Li, M., et al. (2021) Dive into Deep Learning. arXiv: 2106.11342.
- [12] 邱锡鹏. 神经网络与深度学习[M]. 北京: 机械工业出版社, 2020.
- [13] Hübner, R., Steinhauser, M. and Lehle, C. (2010) A Dual-Stage Two-Phase Model of Selective Attention. *Psychological Review*, **117**, 759-784. <https://doi.org/10.1037/a0019471>
- [14] Qin, Y., Song, D., Chen, H., Cheng, W., Jiang, G. and Cottrell, G.W. (2017) A Dual-Stage Attention-Based Recurrent Neural Network for Time Series Prediction. *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, Melbourne, 19-25 August 2017, 2627-2633. <https://doi.org/10.24963/ijcai.2017/366>
- [15] 李佩禅. 基于 CNN-LSTM 算法的盾构掘进姿态智能预测[J]. 四川建筑, 2024, 44(5): 266-268.
- [16] 高修强, 彭达, 王国光, 等. 考虑盾构机参数主动控制的隧道掘进地表沉降智能预测方法[J]. 北京交通大学学报, 2024, 48(3): 120-129.
- [17] 姜山, 刘石磊, 张红兴, 等. 基于深度学习的盾构掘进速度预测算法[J]. 计算机仿真, 2025, 42(6): 213-219.