

如何捕捉长时间序列量价数据的规律

华泰研究

2024年3月14日 | 中国内地

深度研究

人工智能系列之 75: patch 思想用于长时间序列量价选股模型

随着高频数据的普及和算力的发展，量化投资中使用的长时间序列数据长度正逐渐扩展。传统 GRU 模型在处理长序列数据时可能存在信息遗忘、难以捕捉周期性和异质性规律等问题。本研究引入 patch 的思想，按照交易日将股票的长时间序列量价数据划分为多个 patch，设计 PatchModel1 和 PatchModel2 两个模型，并在两个选股场景下进行测试。结果表明，patch 模型具有增量信息，模型融合后相比 GRU 均有提升。使用两个场景下的合成因子对前期报告的全频段融合因子加以改进，回测表现有所提高。

传统 GRU 模型在处理长序列数据时可能存在一些“盲区”

作为一种经典的时间序列深度学习模型，GRU 在量化投资中有着广泛应用。然而，GRU 在处理长时间序列量价数据存在一些缺陷。首先，当序列非常长的时候，GRU 会遇到梯度消失、信息遗忘的问题。其次，高频量价数据具有一定的周期性，GRU 难以捕捉这种周期性的规律。此外，长时间序列量价数据的日内和日间信息传递具有异质性，但参数共享的设计造成 GRU 只能一视同仁地处理所有时间点的数据。

模型引入 patch 设计能够有效缓解 GRU 的不足

Patch 的思想可概括为对数据进行分块处理，并将每一块作为一个整体传入模型，在时间序列预测和计算机视觉等领域均有应用。本研究按照交易日将股票的长时间序列量价数据划分为多个 patch，使模型可以有效缓解信息遗忘的问题，并引入以日为周期的先验知识，差异化地分析日内和日间信息传递。本文设计了 PatchModel1 和 PatchModel2 两个模型。PatchModel1 使用 GRU 处理日内的时序数据，再通过注意力机制构建日间的联系；PatchModel2 将日内时点信息拆解为不同的特征，再使用 GRU 来挖掘日间的时序规律。

Patch 模型相较于基准 GRU 模型具有增量信息

本研究在两个场景下测试 patch 模型的表现。在 15 分钟频量价数据序列中，patch 模型回测表现优于 GRU，且模型间预测值相关性不高，等权合成因子的表现进一步提升。样本空间为全 A 股，GRU 与 patch 模型合成后，2017/1/4~2024/2/29 的回测期内周度 RankIC 均值从 8.86% 提升到 9.58%，分 10 层 TOP 组合年化超额收益率从 21.15% 提升到 24.65%。在 30 分钟频量价特征序列中，patch 模型回测表现略弱于 GRU，但仍能提供增量信息。GRU 与 patch 模型合成后，周度 RankIC 均值从 8.27% 提升到 8.62%，分 10 层 TOP 组合年化超额收益率从 20.42% 提升到 21.64%。

改进全频段融合因子，回测表现有所增强

使用上述两个实验场景中的合成因子，对前期报告的全频段融合因子加以改进。全频段融合因子在 2017/1/4~2024/2/29 的回测期内周度 RankIC 均值从 10.42% 提升到 11.33%，分 10 层 TOP 组合年化超额收益率从 32.61% 提升到 34.40%。基于全频段融合因子 2.0 版本构建指数增强组合。在周双边换手率分别控制为 30%、40% 和 50% 的情况下，2017/1/4~2024/2/29 回测期内中证 500 增强组合年化超额收益率为 18.93%、18.57% 和 18.43%，信息比率为 3.27、3.12 和 3.00；中证 1000 增强组合年化超额收益率为 29.25%、30.92% 和 28.94%，信息比率为 4.35、4.48 和 4.12。

风险提示：借助高频因子、人工智能构建的选股策略是历史经验的总结，存在失效的可能。深度学习的可解释性较弱，使用需谨慎。

研究员

SAC No. S0570516010001
SFC No. BPY421

林晓明

linxiaoming@htsc.com
+(86) 755 8208 0134

研究员

SAC No. S0570520080004
SFC No. BRB318

何康, PhD

hekang@htsc.com
+(86) 21 2897 2039

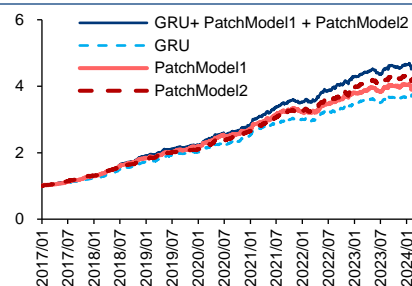
联系人

SAC No. S0570123070272

卢炯

lujiong@htsc.com
+(86) 755 8249 2388

15 分钟频量价模型相对净值



资料来源：Wind，华泰数据中台，华泰研究

正文目录

研究导读3

Patch 思想4

 Patch 的概念与优点4

 Patch 的相关研究4

实验设计6

 方法6

 PatchModel16

 PatchModel27

 实验场景7

 15 分钟频量价数据序列7

 30 分钟频量价特征序列8

结果10

 15 分钟频量价数据序列10

 30 分钟频量价特征序列11

全频段融合因子的改进13

 全频段融合因子测试13

 指数增强策略14

 中证 500 增强14

 中证 1000 增强15

总结17

 风险提示17

 参考文献18

研究导读

随着高频数据的普及和算力的发展，量化投资中使用的时间序列数据长度正逐渐扩展。以华泰金工的研究为例，2020年6月报告《AlphaNet：因子挖掘神经网络》中，使用过去30个交易日的日频量价数据作为模型输入，时间序列的长度为30；2023年5月报告《神经网络多频率因子挖掘模型》中，使用过去20个交易日的15分钟K线数据作为模型输入，时间序列的长度达到320。

更长的时间序列量价数据蕴含着更加丰富的信息，也带来了新的问题：传统时间序列模型可能难以充分挖掘长序列中的规律。例如，门控循环单元（GRU）是一种经典的时间序列深度学习模型，在量化投资中有着广泛应用。然而，GRU模型在处理长序列数据时可能存在一些“盲区”：（1）尽管GRU中使用了重置门和更新门来控制信息的流动，但仍会遇到梯度消失、信息遗忘的问题，当序列非常长的时候尤为明显；（2）高频量价数据具有一定的周期性，比如早盘和尾盘的交易量一般比盘中更大，GRU难以捕捉这种周期性的规律；（3）长时间序列量价数据的信息传递具有异质性，特别是从收盘到次日开盘，隔夜信息很可能与日内信息有较大差异，但参数共享的设计造成GRU只能一视同仁地处理所有时间点的信息。如何使用合适的模型来捕捉长时间序列数据的规律？

Patch思想可能是值得借鉴的解决方案，通过对时间序列数据进行分块处理，能够有效兼顾局部和全局信息。本研究将长时间序列量价数据按照交易日划分为多个patch，对日内和日间的时间序列展开差异化建模，以改善GRU的表现。本文设计了两个模型PatchModel1和PatchModel2，并在15分钟频量价数据序列和30分钟频量价特征序列两个实验场景下进行测试。结果表明：patch模型接近或优于基准GRU模型，而且信息互补，模型融合后相比GRU均有提升。

图表1：15分钟频量价模型的IC值分析和分层回测结果汇总

	RankIC 均值	RankIC 标准差	IC_IR	IC>0 占比	TOP 组合年化超额收益率	TOP 组合信息比率	TOP 组合胜率	TOP 组合换手率
GRU	8.86%	7.47%	1.19	88.76%	21.15%	4.06	82.56%	120.05%
PatchModel1	8.43%	7.18%	1.17	87.90%	21.95%	4.30	81.40%	111.90%
PatchModel2	8.78%	6.72%	1.31	92.22%	23.52%	4.62	90.70%	115.28%
GRU+PatchModel1+PatchModel2	9.58%	7.54%	1.27	92.22%	24.65%	4.59	87.21%	110.15%

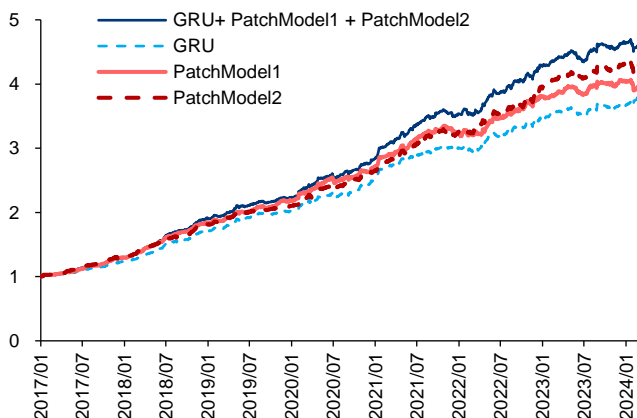
资料来源：Wind，华泰数据中台，华泰研究

图表2：30分钟频量价模型的IC值分析和分层回测结果汇总

	RankIC 均值	RankIC 标准差	IC_IR	IC>0 占比	TOP 组合年化超额收益率	TOP 组合信息比率	TOP 组合胜率	TOP 组合换手率
GRU	8.27%	6.38%	1.30	90.78%	20.42%	3.86	86.05%	106.86%
PatchModel1	7.83%	6.41%	1.22	87.32%	19.72%	3.91	81.40%	96.42%
PatchModel2	8.03%	6.82%	1.18	86.17%	18.03%	4.56	89.54%	119.15%
GRU+PatchModel1+PatchModel2	8.62%	6.64%	1.30	89.05%	21.64%	4.69	86.05%	104.95%

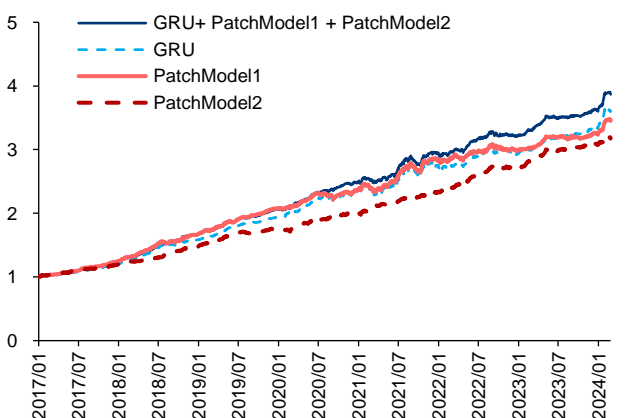
资料来源：Wind，华泰数据中台，华泰研究

图表3：15分钟频量价模型相对净值（分10层TOP层）



资料来源：Wind，华泰数据中台，华泰研究

图表4：30分钟频量价模型相对净值（分10层TOP层）



资料来源：Wind，华泰数据中台，华泰研究

Patch 思想

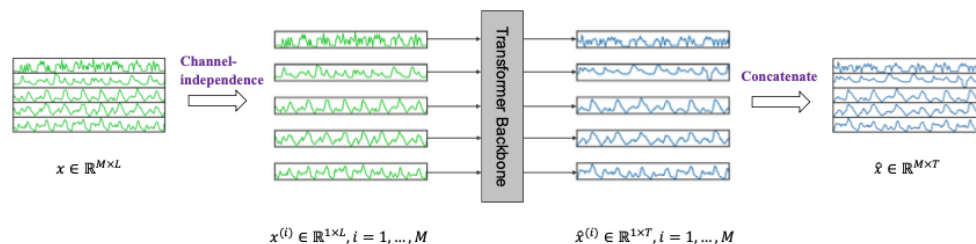
Patch 的概念与优点

Patch 的思想可以概括为对时间序列数据进行分块处理，并将每一块作为一个整体传入模型。对数据进行 patch 处理具备以下优点：（1）模型保留了局部信息，在时间序列中模型可以捕捉某一时间点前后一段时间的信息，而不只是这一时间点的信息；（2）在给定相同时间窗口下，对模型进行 patch 处理减少了算力以及内存的占用，提升了模型的运行效率；（3）模型可以捕捉更长时间窗口的信息，由于时间序列通常携带大量的时间冗余信息，因此在过往的研究中通常调低采样频率或设计稀疏连接的方法来忽略部分数据点，而进行 patch 处理后的数据可以在保留全部数据点的基础上避免冗余信息的影响。

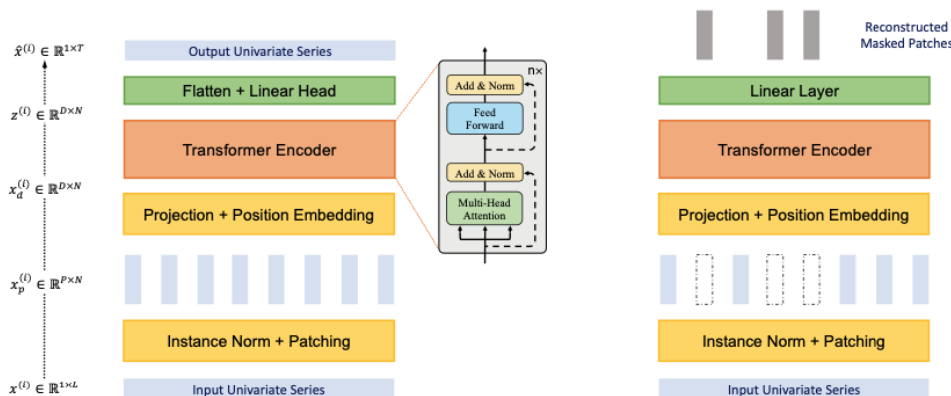
Patch 的相关研究

近年来许多学者将 patch 的思想应用在不同的深度学习领域。Nie 等（2022）提出 patch 时间序列 Transformer 模型（PatchTST），并应用于多变量时间序列预测和自监督表征学习中。PatchTST 模型包括 patch 和通道独立两个核心部分：patch 将时间序列分解为多个子序列，作为 Transformer 的输入 token；通道独立意味着每个输入 token 只包含来自单个变量的信息。论文比较了 PatchTST 与多个 Transformer 类模型在不同领域数据集上的表现，发现 PatchTST 展现了更好的预测能力。论文还尝试了自监督学表征学习，模型性能也有提升。

图表5：PatchTST 模型



(a) PatchTST Model Overview



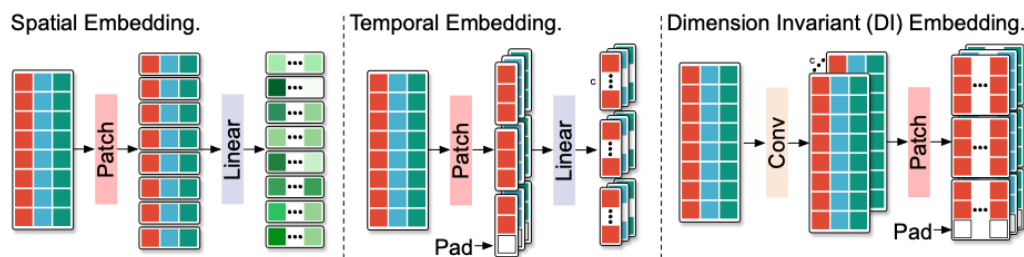
(b) Transformer Backbone (Supervised)

(c) Transformer Backbone (Self-supervised)

资料来源：A time series is worth 64 words: Long-term forecasting with transformers，华泰研究

Zhang 等（2023）提出多尺度 Transformer 金字塔网络（MTPNet），用于有效捕捉在多个不受约束的尺度上的时间依赖性。模型使用了维度不变编码（Dimension Invariant Embedding）的方法。以往的编码方式包括两种，在空间上分 patch 或在时间上分 patch。该研究认为应该同时考虑空间和时间两个维度，先通过卷积对原来的时间序列做处理，实现空间维度信息的融合，然后在时间维度上进行分 patch 的操作。这种方式使得每个 patch 的编码同时保留了空间和时间维度的信息。

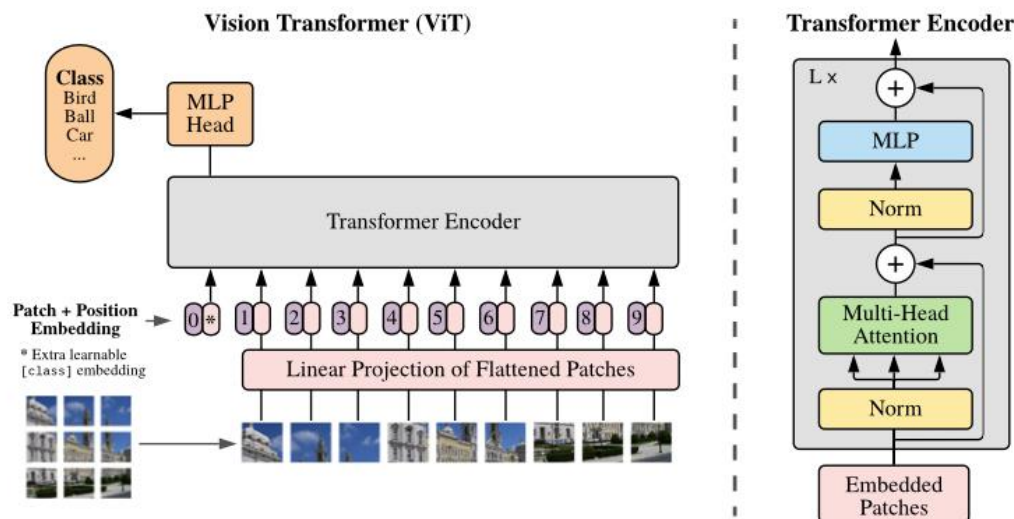
图表6：空间、时间、维度不变编码方法



资料来源：Multi-scale Transformer Pyramid Networks for Multivariate Time Series Forecasting，华泰研究

此外，在计算机视觉领域，Dosovitskiy 等（2020）基于 patch 提出视觉 Transformer 模型（ViT），并得到了比卷积神经网络更出色的性能。应用于大型数据集上时，ViT 模型可以捕捉图片中局部信息，并减少需要的计算资源。

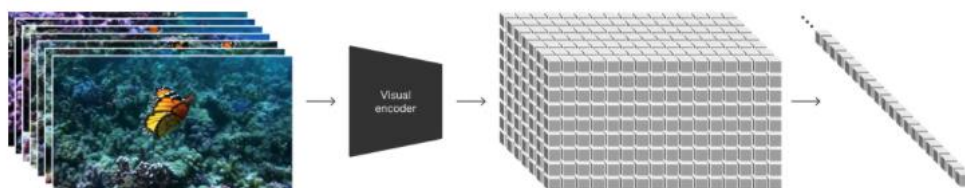
图表7：ViT 模型



资料来源：An image is worth 16x16 words: Transformers for image recognition at scale，华泰研究

OpenAI 最新成果 Sora 模型能够将文本描述转换为相应的视频内容，在技术和效果上都有巨大突破，核心之一是 spacetime patch，先将视频压缩到低维空间，再分解为时空 patch 序列，作为 Transformer 的输入 token，这使得 Sora 能够适应不同分辨率的视频。

图表8：Sora 模型的 spacetime patch



资料来源：OpenAI，华泰研究

实验设计

方法

借鉴上述思想，股票的长时间序列数据也可划分为多个 patch，很自然的方法便是按照交易日来分割，即每个 patch 为同一交易日的日内数据，不同 patch 属于不同交易日。这种做法可以有效缓解 GRU 的不足：（1）GRU 在分析长时间序列数据时会出现信息遗忘的问题，而 patch 能够处理更长时间窗口的信息；（2）GRU 难以捕捉高频量价数据的周期性规律，patch 划分方法可引入以日为周期的先验知识；（3）GRU 只能同质化地处理不同时刻的信息传递，patch 能够差异化地分析日内和日间信息传递。

将股票分割为 patch 之后，还会涉及两个问题：（1）patch 内的数据如何建模；（2）patch 间的数据如何建模。一种思路是，patch 内数据具有时序规律，可使用 GRU 进行建模，之后再通过注意力或拼接等方式构建 patch 间的联系，由此我们设计出第一个模型 PatchModel1。还有一种思路是，patch 内数据可拆解为不同的特征（例如早盘数据是一个特征，尾盘数据是另一个特征），再使用 GRU 来挖掘 patch 间的时序规律，基于这种思路我们设计出第二个模型 PatchModel2。

PatchModel1

PatchModel1 使用 GRU 捕捉 patch 内的时序信息，再通过注意力机制构建 patch 间的联系，网络结构如下图。

输入维度为 $t \times f$ 的数据（其中 t 代表序列长度， f 代表特征数量），首先将输入数据划分为多个 patch，每个 patch 属于同一天，得到维度为 $m \times n \times f$ 的向量（其中 m 代表天数， n 代表每一天内时间序列的长度）。例如，序列长度为 320 的 15 分钟频开、高、低、收、vwap、成交量数据，可转换为 $20 \times 16 \times 6$ 的向量，即过去 20 个交易日，每个交易日有 16 个 15 分钟区间，每个区间有 6 个特征。

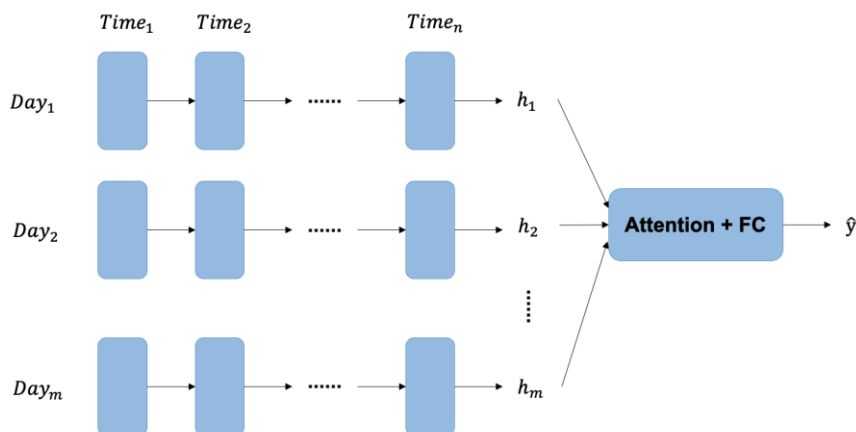
接着，使用 GRU 模型提取每个 patch 内部的时间序列信息，取最后一个时间步的输出，得到 m 个维度为 h 的向量，其中 h 代表隐含层维度。考虑到每天的序列都是从开盘到收盘，具有一定的共性，本研究使用相同参数的 GRU 对不同 patch 进行建模。最后，使用注意力机制对不同 patch 的输出进行加权，再接入全连接层，得到预测值。计算注意力的方法为：

$$a_i = \tanh(wh_i + b)$$

$$a_i = \frac{\exp(a_i)}{\sum_{j=1}^m \exp(a_j)}$$

$$h = \sum_{i=1}^m a_i h_i$$

图表9：PatchModel1 网络结构



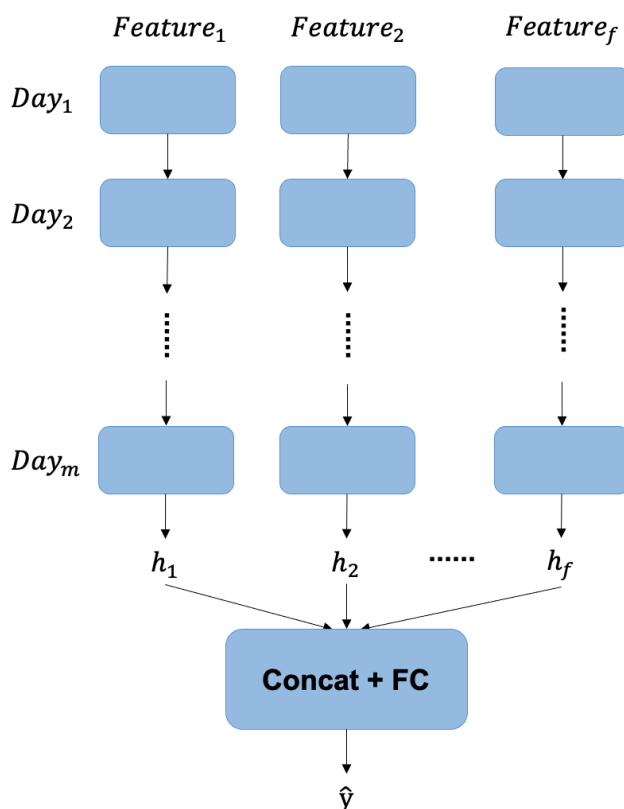
资料来源：华泰研究

PatchModel2

PatchModel2 使用 GRU 来挖掘 patch 间的时序规律，网络结构如下图。

同 PatchModel1 一样，输入维度为 $t \times f$ 的数据，通过 patch 的划分可变换得到维度为 $m \times n \times f$ 的向量。接着，借鉴 PatchTST 论文中通道独立的设计，对于每个原始特征，使用 GRU 提取 patch 间的时间序列信息，其中以每个 patch 作为时间步，patch 内更细颗粒度的时间作为新的特征。考虑到不同通道的时间序列规律不一定相似，此模型使用不同参数的 GRU。取最后一个时间步的输出，得到 f 个维度为 h 的隐含层。最后，将不同原始特征的隐含层拼接起来，再接入全连接层，得到预测值。

图表10：PatchModel2 网络结构



资料来源：华泰研究

实验场景

15 分钟频量价数据序列

华泰金工报告《神经网络多频率因子挖掘模型》（2023.5.11）中设计了基于 GRU 的 15 分钟频量价数据模型，使用个股过去 20 个交易日的 15 分钟频开、高、低、收、vwap、成交量数据来预测未来 10 个交易日的收益率。报告还尝试加入注意力机制来更好地记忆长序列信息，但相比原始 GRU 模型并无优势。本文将在此场景下测试 patch 模型的有效性。

图表11：模型及数据细节

特征和标签	特征 X: 个股过去 20 个交易日的 15 分钟频开、高、低、收、vwap、成交量数据，序列长度为 $20 \times 16 = 320$ 。
	标签 y: 个股未来 10 个交易日 ($T+1 \sim T+11$) 的收益率。
	特征预处理: 每个特征先进行时间序列标准化，即将特征时间序列的每个值除以该序列的均值。再对每个特征进行截面 z-score 标准化。
	标签预处理: 对标签进行截面 z-score 标准化。
	样本内训练数据从 2013 年开始，每 5 个交易日采样一次，训练集和验证集依时间先后按照 4:1 的比例划分。
损失函数: 预测值与标签之间 IC 的相反数。	
模型训练参数	batch_size: 5000。
	训练最大迭代轮数: 100，早停轮数: 10。
	学习率: 0.005，优化器: Adam。
	模型每隔半年重新训练一次。

资料来源：华泰研究

图表12：网络结构对比

GRU 基准模型	GRU: 输入维度 320×6 ，输出维度 30，层数为 1 层。
	BN: 对 GRU 的输出进行批标准化。
	FC: 全连接层，输入维度 30，输出维度 1。
PatchModel1	GRU: 输入维度 16×6 ，输出维度 30，层数为 1 层，数量 20 个。
	Attention: 计算每个 GRU 输出隐含层的得分，再加权求和，输出维度 30。
	BN: 对 Attention 的输出进行批标准化。
	FC: 全连接层，输入维度 30，输出维度 1。
PatchModel2	GRU: 输入维度 20×16 ，输出维度 30，层数为 1 层，数量 6 个。
	Concat: 将多个 GRU 的输出拼接，输出维度 180。
	BN: 对 Concat 的输出进行批标准化。
FC: 全连接层，输入维度 180，输出维度 1。	

资料来源：华泰研究

30 分钟频量价特征序列

华泰金工报告《基于全频段量价特征的选股模型》(2023.12.8) 中使用分钟频、逐笔成交和逐笔委托数据，构建多个日频化因子。本研究将这些因子提升到 30 分钟频，使用个股过去 40 个交易日的 30 分钟频量价特征作为深度学习模型的输入，预测未来 10 个交易日的收益率，并测试此场景下 patch 模型的有效性。值得一提的是，量价特征使用 30 分钟频的原因是与前个实验 15 分钟频量价数据保持差异，以便论证模型的鲁棒性。

图表13：模型及数据细节

特征和标签	特征 X: 个股过去 40 个交易日的 30 分钟频量价特征，序列长度为 $40 \times 8 = 320$ 。
	标签 y: 个股未来 10 个交易日 ($T+1 \sim T+11$) 的收益率。
	特征预处理: 对每个因子进行中位数去极值、行业市值中性化、截面 z-score 标准化、缺失值填充。
	标签预处理: 对标签进行截面 z-score 标准化。
	样本内训练数据从 2013 年开始，每 5 个交易日采样一次，训练集和验证集依时间先后按照 4:1 的比例划分。
损失函数: 预测值与标签之间 IC 的相反数。	
模型训练参数	batch_size: 5000。
	训练最大迭代轮数: 100，早停轮数: 10。
	学习率: 0.005，优化器: Adam。
	模型每隔半年重新训练一次。

资料来源：华泰研究

图表14：网络结构对比

GRU 基准模型	GRU：输入维度 320x22，输出维度 30，层数为 1 层。
	BN：对 GRU 的输出进行批标准化。
	FC：全连接层，输入维度 30，输出维度 1。
PatchModel1	GRU：输入维度 8x22，输出维度 30，层数为 1 层，数量 40 个。
	Attention：计算每个 GRU 输出隐含层的得分，再加权求和，输出维度 30。
	BN：对 Attention 的输出进行批标准化。
	FC：全连接层，输入维度 30，输出维度 1。
PatchModel2	GRU：输入维度 40x8，输出维度 30，层数为 1 层，数量 22 个。
	Concat：将多个 GRU 的输出拼接，输出维度 660。
	BN：对 Concat 的输出进行批标准化。
	FC：全连接层，输入维度 660，输出维度 1。

资料来源：华泰研究

结果

我们使用单因子测试的方法，对以上模型进行测试。为了减轻随机性干扰，本文的深度学习模型都用不同随机数种子训练三次，将三次的模型等权集成，作为最终的因子信号。

单因子测试方法如下：

1. 股票池：全 A 股，剔除 ST 股票，剔除每个截面期下一交易日停牌、涨停的股票。
2. 回测区间：2017/1/4~2024/2/29。
3. 调仓周期：周频，不计交易费用。
4. 因子预处理：因子去极值、行业市值中性化、标准化。
5. 测试方法：IC 值分析，因子分 10 层测试，因子间相关性分析。

15 分钟频量价数据序列

测试结果如下所示。与 GRU 基准模型相比，PatchModel1 和 PatchModel2 整体表现更优，尽管 RankIC 均值略低，但 TOP 组合年化超额收益率和信息比率明显更高，TOP 组合换手率也有所下降。

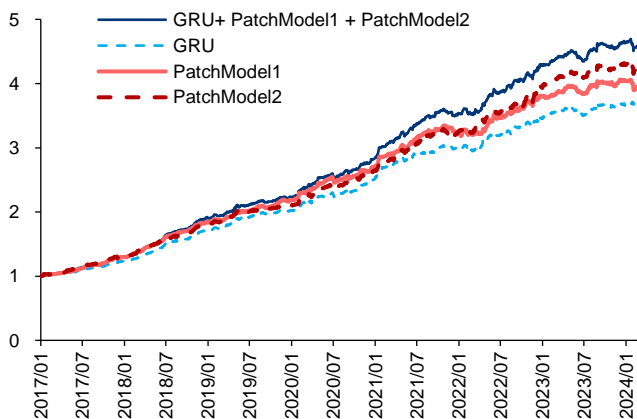
模型间预测值的相关性不高，GRU、PatchModel1 和 PatchModel2 三个模型间预测值的相关系数在 0.5~0.7 之间。进一步将模型进行等权合成，合成因子的回测表现强于单因子。其中，三模型合成因子的 TOP 组合年化超额收益率先其他因子，达到 24.65%。

图表 15：15 分钟频量价模型的 IC 值分析和分层回测结果汇总

	RankIC 均值	RankIC 标准差	IC_IR	IC>0 占比	TOP 组合年化超额收益率	TOP 组合信息比率	TOP 组合胜率	TOP 组合换手率
GRU	8.86%	7.47%	1.19	88.76%	21.15%	4.06	82.56%	120.05%
PatchModel1	8.43%	7.18%	1.17	87.90%	21.95%	4.30	81.40%	111.90%
PatchModel2	8.78%	6.72%	1.31	92.22%	23.52%	4.62	90.70%	115.28%
GRU+PatchModel1	9.13%	7.57%	1.21	90.78%	22.32%	4.26	79.07%	112.41%
GRU+PatchModel2	9.58%	7.45%	1.29	92.51%	24.33%	4.57	88.37%	113.12%
GRU+PatchModel1+PatchModel2	9.58%	7.54%	1.27	92.22%	24.65%	4.59	87.21%	110.15%

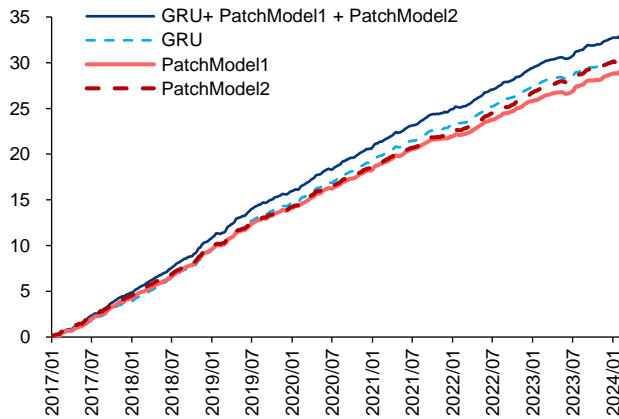
资料来源：Wind，华泰数据中台，华泰研究

图表 16：15 分钟频量价模型相对净值



资料来源：Wind，华泰数据中台，华泰研究

图表 17：15 分钟频量价模型累计 RankIC



资料来源：Wind，华泰数据中台，华泰研究

图表 18：15 分钟频量价模型间预测值相关性

	GRU	PatchModel1	PatchModel2	GRU+PatchModel1+PatchModel2
GRU		0.69	0.58	0.84
PatchModel1	0.69		0.63	0.86
PatchModel2	0.58	0.63		0.81
GRU+PatchModel1+PatchModel2	0.84	0.86	0.81	

资料来源：Wind，华泰数据中台，华泰研究

15 分钟频 patch 模型相对 GRU 模型的增量信息是否仅来源于对日间信息的捕捉，改用日频量价数据模型能否囊括这些信息？本研究训练了一个基于过去 40 个交易日的日频量价数据来预测未来 10 日收益率的日频 GRU 模型。首先进行相关性分析，15 分钟频 GRU 与日频 GRU 模型的相关性为 0.53，而 PatchModel1、PatchModel2 与日频 GRU 模型的相关性均更低，三模型合成因子与日频 GRU 模型的相关性为 0.55，仅提高 0.02。接着使用 15 分钟频量价模型信号对日频 GRU 信号做回归，再对残差进行回测，PatchModel1、PatchModel2 以及三模型合成因子的残差在多数指标上都好于 GRU 残差。此外，还可使用 patch 模型信号对日频 GRU 和 15 分钟频 GRU 信号同时做回归，残差超额同样显著。综上，patch 模型通过网络结构设计更好地融合日内和日间信息，并非 15 分钟频 GRU 和日频 GRU 的简单加和。

图表 19：15 分钟频量价模型与日频量价模型的相关性

	15 分钟频 GRU	PatchModel1	PatchModel2	GRU+PatchModel1+PatchModel2
日频 GRU	0.53	0.50	0.46	0.55

资料来源：Wind，华泰数据中台，华泰研究

图表 20：15 分钟频量价模型残差的 IC 值分析和分层回测结果汇总

	RankIC 均值	RankIC 标准差	IC_IR	IC>0 占比	TOP 组合年化超额收益率	TOP 组合信息比率	TOP 组合胜率	TOP 组合换手率
GRU 残差	3.92%	5.90%	0.66	76.66%	4.88%	1.04	66.28%	132.63%
PatchModel1 残差	3.57%	5.61%	0.64	75.79%	6.32%	1.45	66.28%	121.76%
PatchModel2 残差	4.21%	5.48%	0.77	78.67%	7.96%	1.77	70.93%	128.18%
GRU+PatchModel1+PatchModel2 残差	4.57%	6.03%	0.76	78.10%	8.79%	1.98	73.26%	124.73%

资料来源：Wind，华泰数据中台，华泰研究

30 分钟频量价特征序列

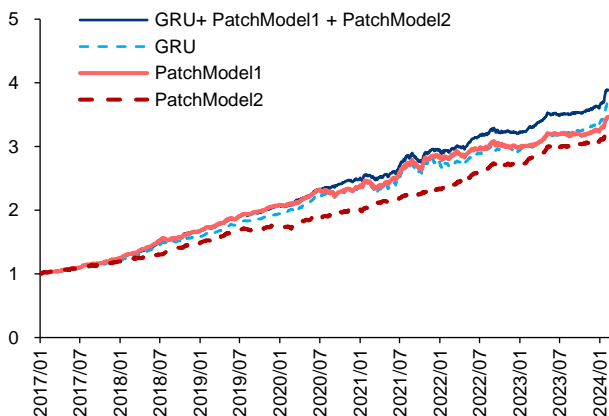
测试结果如下所示。在此实验场景下，PatchModel1、PatchModel2 的 RankIC 均值和 TOP 组合年化超额收益率表现略弱于 GRU 基准模型。不过，单模型间预测值的相关系数在 0.6~0.8 之间，patch 模型仍能提供增量信息，与 GRU 等权重合成后的因子表现强于单因子。其中，三模型合成因子表现较好，RankIC 均值为 8.62%，TOP 组合年化超额收益率为 21.64%。

图表 21：30 分钟频量价模型的 IC 值分析和分层回测结果汇总

	RankIC 均值	RankIC 标准差	IC_IR	IC>0 占比	TOP 组合年化超额收益率	TOP 组合信息比率	TOP 组合胜率	TOP 组合换手率
GRU	8.27%	6.38%	1.30	90.78%	20.42%	3.86	86.05%	106.86%
PatchModel1	7.83%	6.41%	1.22	87.32%	19.72%	3.91	81.40%	96.42%
PatchModel2	8.03%	6.82%	1.18	86.17%	18.03%	4.56	89.54%	119.15%
GRU+PatchModel1	8.36%	6.52%	1.28	88.76%	21.62%	4.07	84.88%	99.69%
GRU+PatchModel2	8.64%	6.65%	1.30	89.91%	21.88%	5.16	88.37%	111.80%
GRU+PatchModel1+PatchModel2	8.62%	6.64%	1.30	89.05%	21.64%	4.69	86.05%	104.95%

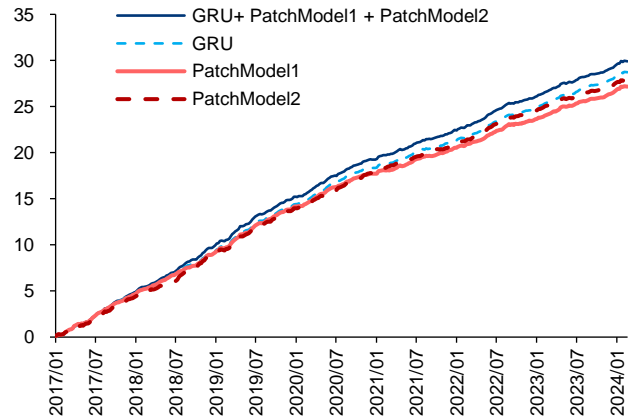
资料来源：Wind，华泰数据中台，华泰研究

图表 22：30 分钟频量价模型相对净值



资料来源：Wind，华泰数据中台，华泰研究

图表 23：30 分钟频量价模型累计 RankIC



资料来源：Wind，华泰数据中台，华泰研究

图表24： 30 分钟频量价模型间预测值相关性

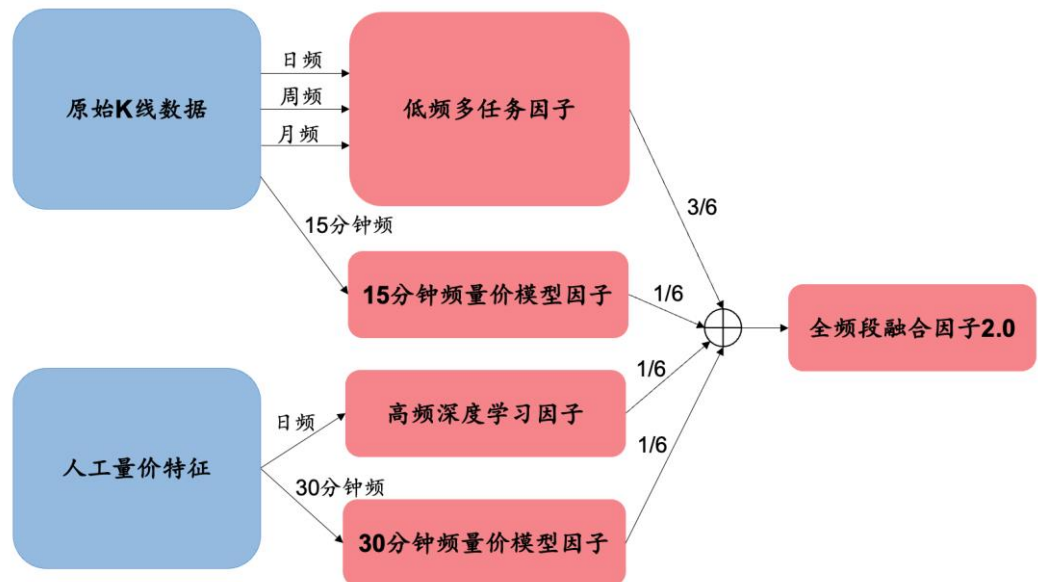
	GRU	PatchModel1	PatchModel2	GRU+PatchModel1+PatchModel2
GRU		0.75	0.66	0.88
PatchModel1	0.75		0.66	0.88
PatchModel2	0.66	0.66		0.84
GRU+PatchModel1+PatchModel2	0.88	0.88	0.84	

资料来源：Wind，华泰数据中台，华泰研究

全频段融合因子的改进

华泰金工报告《基于全频段量价特征的选股模型》（2023.12.8）中基于高频因子和低频量价数据，使用深度学习训练得到高频深度学习因子和低频多任务因子，再将两者按照 1:3 比例合成得到全频段融合因子。我们用上述两个实验场景中的合成因子对全频段融合因子加以改进。15 分钟频量价模型因子、30 分钟频量价模型因子、高频深度学习因子、低频多任务因子按照 1:1:1:3 比例进行合成，得到全频段融合因子 2.0 版本。

图表25：全频段融合因子 2.0 构建方法



资料来源：华泰研究

全频段融合因子测试

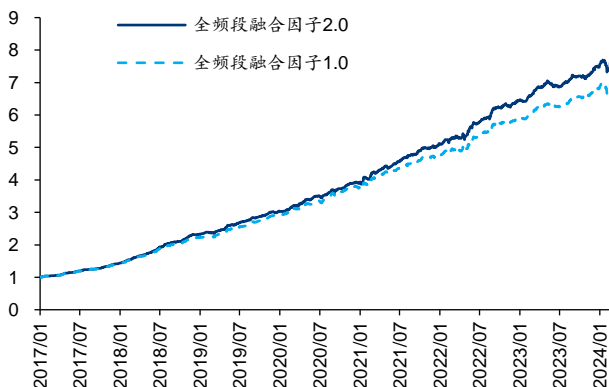
使用单因子测试的方法，对全频段融合因子 1.0 和 2.0 版本进行测试。全频段融合因子 2.0 在 RankIC 均值、IC_IR、TOP 组合年化超额收益率、TOP 组合信息比率、TOP 组合胜率等多项指标上表现更加突出。

图表26：全频段融合因子 IC 值分析和分层回测结果汇总

	RankIC 均值	RankIC 标准差	IC_IR	IC>0 占比	TOP 组合年化超额收益率	TOP 组合信息比率	TOP 组合胜率	TOP 组合换手率
全频段融合因子 1.0	10.42%	7.23%	1.44	92.22%	32.61%	6.12	89.54%	90.50%
全频段融合因子 2.0	11.33%	7.55%	1.50	91.93%	34.40%	6.38	93.02%	92.32%

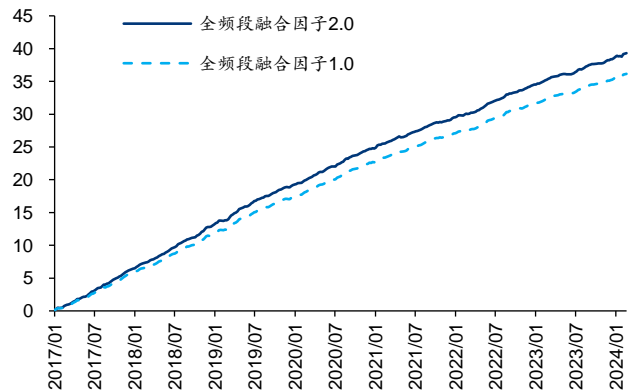
资料来源：Wind，华泰数据中台，华泰研究

图表27：全频段融合因子相对净值



资料来源：Wind，华泰数据中台，华泰研究

图表28：全频段融合因子累计 RankIC



资料来源：Wind，华泰数据中台，华泰研究

指数增强策略

使用全频段融合因子，构建中证 500 和中证 1000 指数增强组合，构建方法如下。

图表29：指数增强组合构建细节

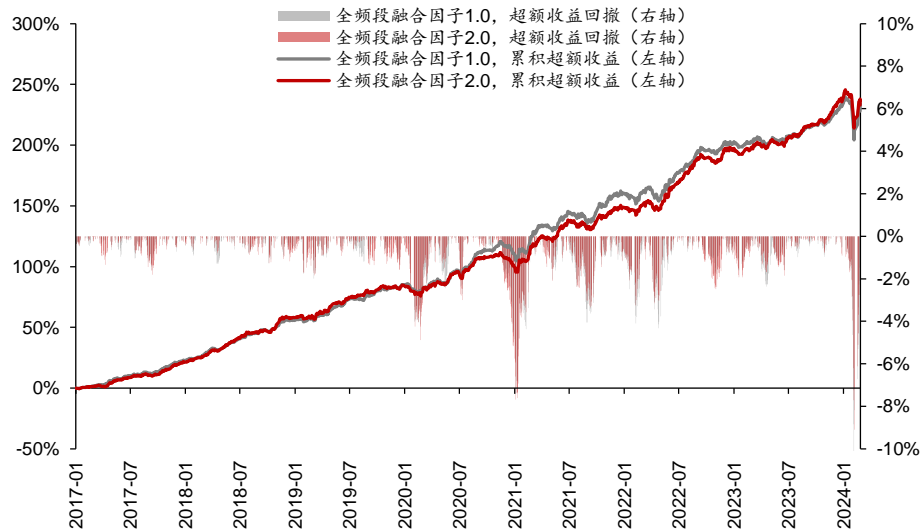
优化目标	最大化预期收益
成分股权重约束	不低于 80%
个股权重偏离上限	0.8%
风格因子约束	行业暴露<0.02，市值暴露<0.2
换手率约束	周双边换手率上限分别为 30%、40%和 50%
调仓频率和交易成本	周频调仓，调仓当日以 vwap 价格成交，交易成本双边千分之四

资料来源：华泰研究

中证 500 增强

中证 500 增强组合回测结果如下。两种版本全频段融合因子构建的中证 500 增强策略，超额收益接近，但 2.0 版本的跟踪误差和超额收益最大回撤更低，因此信息比率和 Calmar 比率普遍更高。

图表30：中证 500 增强组合累计超额收益（换手率 30%）



资料来源：Wind，华泰数据中台，华泰研究

图表31：中证 500 增强组合回测绩效

	年化收益率	年化波动率	夏普比率	最大回撤	年化超额收	年化跟踪误	超额收益最	信息比率	Calmar 比率	相对基准月	调仓双边换
					益率	差	大回撤				
全频段融合 换手率=30%	15.28%	20.94%	0.73	30.55%	18.66%	5.94%	10.37%	3.14	1.80	80.00%	32.65%
因子 1.0 换手率=40%	15.46%	21.06%	0.73	30.28%	18.86%	6.11%	10.45%	3.09	1.80	78.82%	42.10%
换手率=50%	15.33%	21.18%	0.72	31.27%	18.76%	6.20%	10.67%	3.03	1.76	82.35%	51.56%
全频段融合 换手率=30%	15.55%	20.84%	0.75	29.64%	18.93%	5.79%	9.10%	3.27	2.08	80.00%	32.83%
因子 2.0 换手率=40%	15.18%	21.01%	0.72	29.84%	18.57%	5.96%	9.57%	3.12	1.94	77.65%	42.46%
换手率=50%	15.02%	21.15%	0.71	31.36%	18.43%	6.15%	9.92%	3.00	1.86	76.47%	52.00%
中证 500	-2.98%	20.75%	-0.14	41.69%							

资料来源：Wind，华泰数据中台，华泰研究

图表32：中证 500 增强组合逐年收益率

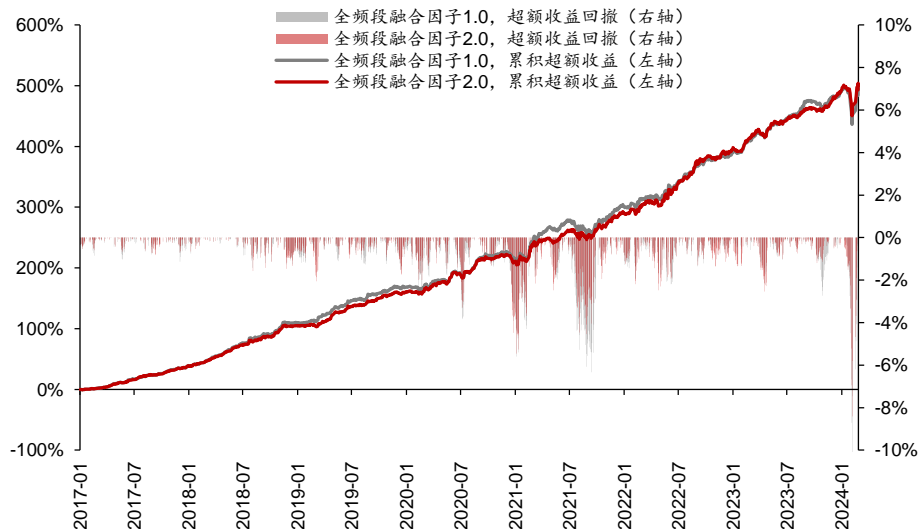
	2017 年收益率	2018 年收益率	2019 年收益率	2020 年收益率	2021 年收益率	2022 年收益率	2023 年收益率	2024 年收益率
全频段融合因子 换手率=30%	20.47%	-15.53%	50.24%	37.86%	43.30%	-7.74%	2.04%	-2.54%
1.0 换手率=40%	21.53%	-12.91%	51.23%	33.75%	44.87%	-10.33%	3.19%	-2.24%
换手率=50%	21.42%	-12.49%	49.75%	34.89%	46.51%	-11.53%	2.72%	-2.60%
全频段融合因子 换手率=30%	19.22%	-13.47%	47.87%	31.64%	43.80%	-5.06%	4.95%	-1.92%
2.0 换手率=40%	20.03%	-12.67%	46.83%	34.41%	42.52%	-6.76%	2.88%	-2.44%
换手率=50%	20.34%	-11.72%	44.51%	36.93%	45.70%	-9.66%	0.82%	-2.03%
中证 500	-2.25%	-33.32%	26.38%	20.87%	15.58%	-20.31%	-7.42%	-1.52%

资料来源：Wind，华泰数据中台，华泰研究

中证 1000 增强

中证 1000 增强组合回测结果如下。全频段融合因子 2.0 版本同样在跟踪误差、超额收益最大回撤、信息比率和 Calmar 比率上体现出优势。

图表33：中证 1000 增强组合累计超额收益（换手率 30%）



资料来源：Wind，华泰数据中台，华泰研究

图表34：中证 1000 增强组合回测绩效

	年化收益率	年化波动率	夏普比率	最大回撤	年化超额收益率	年化跟踪误差	超额收益最大回撤	信息比率	Calmar 比率	相对基准月胜率	调仓双边换手率
全频段融合 换手率=30%	19.74%	23.65%	0.83	34.74%	28.97%	7.03%	10.17%	4.12	2.85	85.88%	33.38%
因子 1.0 换手率=40%	20.42%	23.83%	0.86	34.46%	29.73%	7.26%	11.19%	4.09	2.66	85.88%	42.91%
换手率=50%	19.81%	23.89%	0.83	34.90%	29.08%	7.38%	11.99%	3.94	2.43	84.71%	52.22%
全频段融合 换手率=30%	20.02%	23.51%	0.85	33.24%	29.25%	6.73%	8.41%	4.35	3.48	87.06%	33.99%
因子 2.0 换手率=40%	21.54%	23.65%	0.91	32.97%	30.92%	6.90%	9.36%	4.48	3.30	88.24%	43.33%
换手率=50%	19.68%	23.78%	0.83	34.25%	28.94%	7.03%	10.23%	4.12	2.83	87.06%	52.66%
中证 1000	-7.23%	22.95%	-0.31	52.14%							

资料来源：Wind，华泰数据中台，华泰研究

图表35： 中证 1000 增强组合逐年收益率

	2017 年收益率	2018 年收益率	2019 年收益率	2020 年收益率	2021 年收益率	2022 年收益率	2023 年收益率	2024 年收益率
全频度融合因子 换手率=30%	12.05%	-4.36%	61.06%	40.80%	53.71%	-4.68%	13.31%	-9.74%
1.0 换手率=40%	10.35%	-3.41%	64.77%	42.11%	55.89%	-5.35%	14.47%	-9.94%
换手率=50%	9.71%	-1.56%	63.52%	41.23%	53.66%	-4.85%	12.35%	-10.67%
全频度融合因子 换手率=30%	12.01%	-6.44%	59.69%	42.84%	51.80%	-0.55%	11.68%	-8.37%
2.0 换手率=40%	11.69%	-1.63%	62.11%	47.27%	52.02%	-2.94%	14.17%	-8.83%
换手率=50%	10.98%	-0.04%	61.55%	44.11%	48.73%	-6.33%	11.56%	-9.69%
中证 1000	-19.06%	-36.87%	25.67%	19.39%	20.52%	-21.58%	-6.28%	-9.22%

资料来源：Wind，华泰数据中台，华泰研究

总结

随着高频数据的普及和算力的发展，量化投资中使用的时间序列数据长度正逐渐扩展。传统 GRU 模型在处理长序列数据时可能存在信息遗忘、难以捕捉周期性和异质性规律等问题。本研究引入 patch 的思想，按照交易日将股票的长时间序列量价数据划分为多个 patch，设计 PatchModel1 和 PatchModel2 两个模型，并在两个选股场景下进行测试。结果表明，patch 模型具有增量信息，模型融合后相比 GRU 均有提升。使用两个场景下的合成因子对前期报告的全频段融合因子加以改进，回测表现有所提高。

传统 GRU 模型在处理长序列数据时可能存在一些“盲区”。作为一种经典的时间序列深度学习模型，GRU 在量化投资中有着广泛应用。然而，GRU 在处理长时间序列量价数据存在一些缺陷。首先，当序列非常长的时候，GRU 会遇到梯度消失、信息遗忘的问题。其次，高频量价数据具有一定的周期性，GRU 难以捕捉这种周期性的规律。此外，长时间序列量价数据的日内和日间信息传递具有异质性，但参数共享的设计造成 GRU 只能一视同仁地处理所有时间点的数据。

模型引入 patch 设计能够有效缓解 GRU 的不足。Patch 的思想可概括为对数据进行分块处理，并将每一块作为一个整体传入模型，在时间序列预测和计算机视觉等领域均有应用。本研究按照交易日将股票的长时间序列量价数据划分为多个 patch，使模型可以有效缓解信息遗忘的问题，并引入以日为周期的先验知识，差异化地分析日内和日间信息传递。本文设计了 PatchModel1 和 PatchModel2 两个模型。PatchModel1 使用 GRU 处理日内的时序数据，再通过注意力机制构建日间的联系；PatchModel2 将日内时点信息拆解为不同的特征，再使用 GRU 来挖掘日间的时序规律。

Patch 模型相较于基准 GRU 模型具有增量信息。本研究在两个场景下测试 patch 模型的表现。在 15 分钟频量价数据序列中，patch 模型回测表现优于 GRU，且模型间预测值相关性不高，等权合成因子的表现进一步提升。样本空间为全 A 股，GRU 与 patch 模型合成后，2017/1/4~2024/2/29 的回测期内周度 RankIC 均值从 8.86% 提升到 9.58%，分 10 层 TOP 组合年化超额收益率从 21.15% 提升到 24.65%。在 30 分钟频量价特征序列中，patch 模型回测表现略弱于 GRU，但仍能提供增量信息。GRU 与 patch 模型合成后，周度 RankIC 均值从 8.27% 提升到 8.62%，分 10 层 TOP 组合年化超额收益率从 20.42% 提升到 21.64%。

改进全频段融合因子，回测表现有所增强。使用上述两个实验场景中的合成因子，对前期报告的全频段融合因子加以改进。全频段融合因子在 2017/1/4~2024/2/29 的回测期内周度 RankIC 均值从 10.42% 提升到 11.33%，分 10 层 TOP 组合年化超额收益率从 32.61% 提升到 34.40%。基于全频段融合因子 2.0 版本构建指数增强组合。在周双边换手率分别控制为 30%、40% 和 50% 的情况下，2017/1/4~2024/2/29 回测期内中证 500 增强组合年化超额收益率为 18.93%、18.57% 和 18.43%，信息比率为 3.27、3.12 和 3.00；中证 1000 增强组合年化超额收益率为 29.25%、30.92% 和 28.94%，信息比率为 4.35、4.48 和 4.12。

风险提示

借助高频因子、人工智能构建的选股策略是历史经验的总结，存在失效的可能。深度学习的可解释性较弱，使用需谨慎。

参考文献

- [1] Nie Y, Nguyen N H, Sinthong P, et al. A time series is worth 64 words: Long-term forecasting with transformers[J]. arXiv preprint arXiv:2211.14730, 2022.
- [2] Zhang Y, Wu R, Dasalu S M, et al. Multi-scale transformer pyramid networks for multivariate time series forecasting[J]. IEEE Access, 2024.
- [3] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[J]. arXiv preprint arXiv:2010.11929, 2020.
- [4] Dehghani M, Mustafa B, Djolonga J, et al. Patch n'pack: Navit, a vision transformer for any aspect ratio and resolution[J]. Advances in Neural Information Processing Systems, 2024, 36.
- [5] Arnab A, Dehghani M, Heigold G, et al. Vivit: A video vision transformer[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2021: 6836-6846.

免责声明

分析师声明

本人，林晓明、何康，兹证明本报告所表达的观点准确地反映了分析师对标的证券或发行人的个人意见；彼以往、现在或未来并无就其研究报告所提供的具体建议或所表达的意见直接或间接收取任何报酬。

一般声明及披露

本报告由华泰证券股份有限公司（已具备中国证监会批准的证券投资咨询业务资格，以下简称“本公司”）制作。本报告所载资料是仅供接收人的严格保密资料。本报告仅供本公司及其客户和其关联机构使用。本公司不因接收人收到本报告而视其为客户。

本报告基于本公司认为可靠的、已公开的信息编制，但本公司及其关联机构（以下统称为“华泰”）对该等信息的准确性及完整性不作任何保证。

本报告所载的意见、评估及预测仅反映报告发布当日的观点和判断。在不同时期，华泰可能会发出与本报告所载意见、评估及预测不一致的研究报告。同时，本报告所指的证券或投资标的的价格、价值及投资收入可能会波动。以往表现并不能指引未来，未来回报并不能得到保证，并存在损失本金的可能。华泰不保证本报告所含信息保持在最新状态。华泰对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

本公司不是 FINRA 的注册会员，其研究分析师亦没有注册为 FINRA 的研究分析师/不具有 FINRA 分析师的注册资格。

华泰力求报告内容客观、公正，但本报告所载的观点、结论和建议仅供参考，不构成购买或出售所述证券的要约或招揽。该等观点、建议并未考虑到个别投资者的具体投资目的、财务状况以及特定需求，在任何时候均不构成对客户私人投资建议。投资者应当充分考虑自身特定状况，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素。对依据或者使用本报告所造成的一切后果，华泰及作者均不承担任何法律责任。任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。

除非另行说明，本报告中所引用的关于业绩的数据代表过往表现，过往的业绩表现不应作为日后回报的预示。华泰不承诺也不保证任何预示的回报会得以实现，分析中所做的预测可能是基于相应的假设，任何假设的变化可能会显著影响所预测的回报。

华泰及作者在自身所知情的范围内，与本报告所指的证券或投资标的不存在法律禁止的利害关系。在法律许可的情况下，华泰可能会持有报告中提到的公司所发行的证券头寸并进行交易，为该公司提供投资银行、财务顾问或者金融产品等相关服务或向该公司招揽业务。

华泰的销售人员、交易人员或其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。华泰没有将此意见及建议向报告所有接收者进行更新的义务。华泰的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。投资者应当考虑到华泰及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突。投资者请勿将本报告视为投资或其他决定的唯一信赖依据。有关该方面的具体披露请参照本报告尾部。

本报告并非意图发送、发布给在当地法律或监管规则下不允许向其发送、发布的机构或人员，也并非意图发送、发布给因可得到、使用本报告的行为而使华泰违反或受制于当地法律或监管规则的机构或人员。

本报告版权仅为本公司所有。未经本公司书面许可，任何机构或个人不得以翻版、复制、发表、引用或再次分发他人（无论整份或部分）等任何形式侵犯本公司版权。如征得本公司同意进行引用、刊发的，需在允许的范围内使用，并需在使用前获取独立的法律意见，以确定该引用、刊发符合当地适用法规的要求，同时注明出处为“华泰证券研究所”，且不得对本报告进行任何有悖原意的引用、删节和修改。本公司保留追究相关责任的权利。所有本报告中使用的商标、服务标记及标记均为本公司的商标、服务标记及标记。

中国香港

本报告由华泰证券股份有限公司制作，在香港由华泰金融控股（香港）有限公司向符合《证券及期货条例》及其附属法律规定的机构投资者和专业投资者的客户进行分发。华泰金融控股（香港）有限公司受香港证券及期货事务监察委员会监管，是华泰国际金融控股有限公司的全资子公司，后者为华泰证券股份有限公司的全资子公司。在香港获得本报告的人员若有任何有关本报告的问题，请与华泰金融控股（香港）有限公司联系。

香港-重要监管披露

- 华泰金融控股（香港）有限公司的雇员或其关联人士没有担任本报告中提及的公司或发行人的高级人员。
- 有关重要的披露信息，请参华泰金融控股（香港）有限公司的网页 https://www.htsc.com.hk/stock_disclosure 其他信息请参见下方 “美国-重要监管披露”。

美国

在美国本报告由华泰证券（美国）有限公司向符合美国监管规定的机构投资者进行发表与分发。华泰证券（美国）有限公司是美国注册经纪商和美国金融业监管局（FINRA）的注册会员。对于其在美国分发的研究报告，华泰证券（美国）有限公司根据《1934 年证券交易法》（修订版）第 15a-6 条规定以及美国证券交易委员会人员解释，对本研究报告内容负责。华泰证券（美国）有限公司联营公司的分析师不具有美国金融监管（FINRA）分析师的注册资格，可能不属于华泰证券（美国）有限公司的关联人员，因此可能不受 FINRA 关于分析师与标的公司沟通、公开露面和所持交易证券的限制。华泰证券（美国）有限公司是华泰国际金融控股有限公司的全资子公司，后者为华泰证券股份有限公司的全资子公司。任何直接从华泰证券（美国）有限公司收到此报告并希望就本报告所述任何证券进行交易的人士，应通过华泰证券（美国）有限公司进行交易。

美国-重要监管披露

- 分析师林晓明、何康本人及相关人士并不担任本报告所提及的标的证券或发行人的高级人员、董事或顾问。分析师及相关人士与本报告所提及的标的证券或发行人并无任何相关财务利益。本披露中所提及的“相关人士”包括 FINRA 定义下分析师的家庭成员。分析师根据华泰证券的整体收入和盈利能力获得薪酬，包括源自公司投资银行业务的收入。
- 华泰证券股份有限公司、其子公司和/或其联营公司，及/或不时会以自身或代理形式向客户出售及购买华泰证券研究所覆盖公司的证券/衍生工具，包括股票及债券（包括衍生品）华泰证券研究所覆盖公司的证券/衍生工具，包括股票及债券（包括衍生品）。
- 华泰证券股份有限公司、其子公司和/或其联营公司，及/或其高级管理层、董事和雇员可能会持有本报告中所提到的任何证券（或任何相关投资）头寸，并可能不时进行增持或减持该证券（或投资）。因此，投资者应该意识到可能存在利益冲突。

评级说明

投资评级基于分析师对报告发布日后 6 至 12 个月内行业或公司回报潜力（含此期间的股息回报）相对基准表现的预期（A 股市场基准为沪深 300 指数，香港市场基准为恒生指数，美国市场基准为标普 500 指数），具体如下：

行业评级

增持：预计行业股票指数超越基准

中性：预计行业股票指数基本与基准持平

减持：预计行业股票指数明显弱于基准

公司评级

买入：预计股价超越基准 15% 以上

增持：预计股价超越基准 5%~15%

持有：预计股价相对基准波动在-15%~5%之间

卖出：预计股价弱于基准 15% 以上

暂停评级：已暂停评级、目标价及预测，以遵守适用法规及/或公司政策

无评级：股票不在常规研究覆盖范围内。投资者不应期待华泰提供该等证券及/或公司相关的持续或补充信息

法律实体披露

中国：华泰证券股份有限公司具有中国证监会核准的“证券投资咨询”业务资格，经营许可证编号为：91320000704041011J

香港：华泰金融控股（香港）有限公司具有香港证监会核准的“就证券提供意见”业务资格，经营许可证编号为：AOK809

美国：华泰证券（美国）有限公司为美国金融业监管局（FINRA）成员，具有在美国开展经纪交易商业业务的资格，经营业务许可编号为：CRD#:298809/SEC#:8-70231

华泰证券股份有限公司

南京

南京市建邺区江东中路228号华泰证券广场1号楼/邮政编码：210019

电话：86 25 83389999/传真：86 25 83387521

电子邮件：ht-rd@htsc.com

深圳

深圳市福田区益田路5999号基金大厦10楼/邮政编码：518017

电话：86 755 82493932/传真：86 755 82492062

电子邮件：ht-rd@htsc.com

北京

北京市西城区太平桥大街丰盛胡同28号太平洋保险大厦A座18层/

邮政编码：100032

电话：86 10 63211166/传真：86 10 63211275

电子邮件：ht-rd@htsc.com

上海

上海市浦东新区东方路18号保利广场E栋23楼/邮政编码：200120

电话：86 21 28972098/传真：86 21 28972068

电子邮件：ht-rd@htsc.com

华泰金融控股（香港）有限公司

香港中环皇后大道中99号中环中心58楼5808-12室

电话：+852-3658-6000/传真：+852-2169-0770

电子邮件：research@htsc.com

http://www.htsc.com.hk

华泰证券（美国）有限公司

美国纽约公园大道280号21楼东（纽约10017）

电话：+212-763-8160/传真：+917-725-9702

电子邮件：Huatai@htsc-us.com

http://www.htsc-us.com

©版权所有2024年华泰证券股份有限公司