International Journal of Artificial Intelligent and Informatics
Vol. 3, No. 1, January 2025, pp. 30~ 36
ISSN: 2622-626X, DOI: http://doi.org/10.33292/ijarlit.v3i1.44

Petrobras Stock Price Prediction Using Deep Learning Approach: Performance Comparison of CNN and CNN-GRU Methods

Manuel Gabrielzinho[1*], Giovana Moraes[1]
[1]Faculdade de Informática e Administração Paulista, Brazil
Email: mgabriel@yahoo.com

Article Info

Article history:

Received: June 11, 2024
Revised: August 25, 2024
Accepted: December 20, 2024
Available Online: January 30, 2025

Keywords:

stock price prediction
deep learning
CNN
GRU
Petrobras
time series forecasting

ABSTRACT

This study aims to compare the effectiveness of two deep learning models, namely Convolutional Neural Network (CNN) and combined CNN with Gated Recurrent Unit (CNN-GRU), in predicting Petrobras stock prices. Using historical stock price data from Yahoo Finance for the period 2017-2023, this study evaluates the performance of both models based on several evaluation metrics. The results showed that the CNN-GRU model outperformed the pure CNN model in all evaluation metrics, with an increase in RMSE value of 4.17% and an increase in R^2 value of 0.47% on the test data. The CNN-GRU model achieved 96.14% accuracy on the test data, while the CNN model achieved 96.04%. These findings indicate that the integration of CNN's feature extraction capabilities with GRU's temporal dependency modeling capabilities can improve stock price prediction accuracy. This research contributes to the computational finance literature by presenting an in-depth comparative analysis of the application of hybrid deep learning architectures in stock market prediction.

1. INTRODUCTION

Stock price prediction has been an area of research that has attracted the attention of financial researchers and practitioners for decades. The ability to accurately predict stock price movements not only gives investors a competitive advantage, but also contributes to a better understanding of financial market dynamics. Petrobras (Petróleo Brasileiro S.A.), as one of the world's largest oil and gas companies and a pillar of the Brazilian economy, is an interesting research object due to its stock price volatility which is influenced by various complex factors such as global oil prices, Brazilian government policies, and international market sentiment.

The Efficient Market Hypothesis (EMPH) proposed by Fama (1970) states that stock prices reflect all available information and therefore cannot be consistently predicted. However, recent research has shown that markets are not completely efficient and there are certain patterns that can be identified and modeled[1], [2] . With advances in computing technology and machine learning methods, stock price prediction has undergone a significant transformation from traditional statistical approaches such as ARIMA[3] and GARCH[4] models to more sophisticated machine learning methods.

In recent years, deep learning techniques have shown tremendous potential in modeling time series data, including stock prices. Convolutional Neural Network (CNN), originally developed for image recognition, has been adapted successfully for financial time series analysis[5], [6] . CNNs are able to capture local patterns and hierarchical features in stock price data. On the other hand, Recurrent Neural Network (RNN)-based architectures such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) are specifically designed to process sequential data and have proven effective in modeling temporal dependencies in financial data .[7], [8]

A number of previous studies have explored the use of deep learning techniques for stock price prediction. Sezer et al.[9] conducted a comprehensive survey of deep learning applications in financial market prediction and found significant performance improvements compared to traditional methods. Di Persio and Honchar[10] used CNN to predict the direction of movement of the S&P 500 index and achieved an accuracy of up to 58%. Meanwhile, Zhang et al.[11] applied LSTM models for stock price prediction of large technology companies and reported improved accuracy over conventional statistical models.

Although CNN and RNN-based architectures have been widely researched separately, research on the combination of these two approaches is still relatively limited, especially in the context of Petrobras stock price prediction. The CNN-GRU hybrid model combines the power of CNN in spatial feature extraction with the ability of GRU in capturing temporal dependencies, making it a promising candidate for accurate stock price prediction. Studies such as Livieris et al.[12] have shown the potential of hybrid approaches in improving prediction accuracy compared to individual models.

The novelty of this study lies in several aspects. First, this study focuses on Petrobras stocks, an oil and gas company that has unique characteristics as a state-owned company operating in a competitive global market. Second, this study provides a direct comparison between a pure CNN model and a hybrid CNN-GRU model, which has not been widely explored in the energy company stock prediction literature. Third, this study uses recent data (2017-2023) that covers a period of high volatility due to the COVID-19 pandemic and global oil price fluctuations, providing a more rigorous test of the robustness of the prediction model.

By conducting an in-depth comparative analysis between CNN and CNN-GRU models in the context of Petrobras stock price prediction, this study aims to contribute to the computational finance literature and provide valuable insights for investors, financial analysts, and policy makers interested in the global energy market and the Brazilian economy.

2. METHOD

2.1. Data and Pre-processing

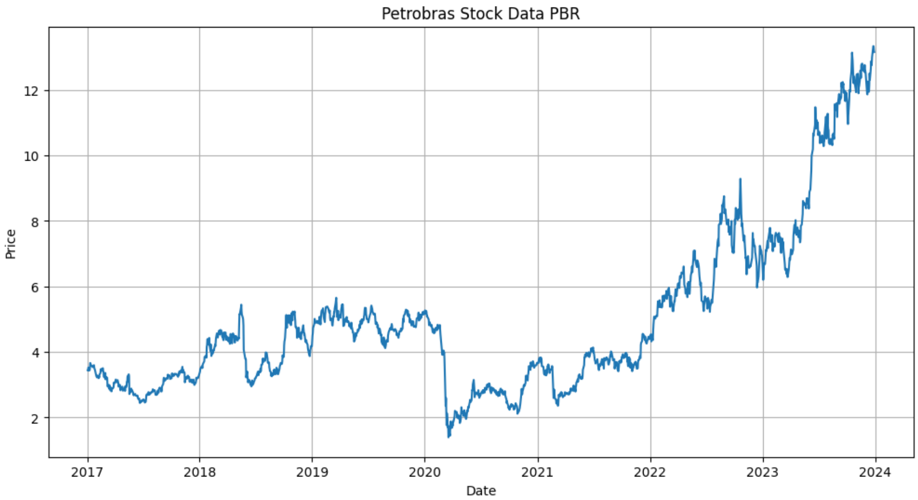

Figure 1. Graph of historical data of Petrobas stock movements

The data used in this study are historical data of Petrobras stock prices (ticker: PBR) obtained from Yahoo Finance for the period January 1, 2017 to December 31, 2023. The dataset includes the open price, high price, low price, close price, and daily trading volume. The main focus of the prediction is the close price, which is a key indicator often used in technical analysis and stock market prediction.

Data pre-processing steps are crucial in improving the quality of prediction models, especially for time-series data of stock price movements. The first step is handling missing data, which occurs on stock exchange holidays or weekends, where there is no trading data. The use of linear interpolation in this context is appropriate as it can effectively fill in the missing data while maintaining temporal continuity, as described by Zhang et al. [13], which demonstrated the efficiency of interpolation in time-series data.

Next, data normalization using the Min-Max Scaling method is applied to map all numerical features to the range[14] . This normalization not only accelerates convergence in training the neural network model, but also prevents the dominance of large-scale features that can lead to bias in model learning. Research conducted by Prasetyo et al.[15] explains the importance of normalization in the pre-processing stage, especially for neural network models.

In addition, the establishment of technical features is a strategic step in analyzing stock price movements. The addition of technical indicators such as Moving Average (MA) with 5, 10, and 20-day periods, Relative Strength Index (RSI) with 14-day period, Moving Average Convergence Divergence (MACD), Bollinger Bands, and Rate of Change (ROC) provides a comprehensive representation of market momentum, volatility, and trends. The application of these technical indicators has been shown to improve the accuracy of stock price predictions by Lin et al. [16]who emphasized the importance of pattern analysis in financial data.

The chronological division of the dataset into three subsets (70% training data, 15% validation data, and 15% test data) is a common practice in time-series analysis to preserve the temporal nature of the data. This approach ensures that the model learns historical patterns in a more realistic way and minimizes the risk of information leakage between datasets, in line with the strategy applied in recent studies on financial data .[17]

Finally, the formation of a data sequence with a window length of 30 days allows the model to capture temporal dynamics in depth. By using 30 days of historical data as input in predicting the next day's price, such a sequence structure effectively integrates temporal information that is essential in modeling the complexity of stock price changes. This method has been recognized as a standard practice in many time-series prediction studies based on neural networks and sequential models .[17]

2.2. Model Architecture

2.2.1. CNN Model

The Convolutional Neural Network (CNN) model implemented in this research consists of several layers as follows:

1. Input Layer: Receive sequence data with dimensions (30, F), where 30 is the window length and F is the number of features.
2. 1D Convolution Layer: Three consecutive convolution blocks, each consisting of:
   - 1D convolution with 64, 128, and 256 filters respectively, kernel size 3, and ReLU activation
   - Batch Normalization to speed up training and improve stability
   - Max Pooling 1D with pool size 2 to reduce dimensionality
   - Dropout (0.3) to reduce overfitting
3. Flatten Layer: Converts the multi-dimensional output of the convolution layer into a 1D vector.
4. Dense Layer (Fully Connected): Two dense layers with 128 and 64 units, ReLU activation, and dropout (0.3).
5. Output Layer: A dense layer with 1 unit to predict the closing price.

CNN models are designed to capture local patterns and hierarchical features in stock price time series data, which can reflect various technical patterns in price movements.

2.2.1. CNN-GRU Model

The CNN-GRU hybrid model combines CNN's ability to extract spatial features with GRU's ability to capture temporal dependencies. The model architecture consists of:

1. Input Layer: Identical to the CNN model, accepts sequence data with dimensions (30, F).
2. 1D Convolution Layer: Two convolution blocks, each consisting of:
   - 1D convolution with 64 and 128 filters, kernel size 3, and ReLU activation
   - Batch Normalization
   - Max Pooling 1D with pool size 2
   - Dropout (0.3)
3. GRU layers: Two consecutive GRU layers with 100 and 50 units, with return_sequences=True for the first layer to maintain the output at each time step.
4. Dense Layer (Fully Connected): Two dense layers with 64 and 32 units, ReLU activation, and dropout (0.3).
5. Output Layer: A dense layer with 1 unit to predict the closing price.

The CNN-GRU model is designed to utilize the advantages of both architectures: CNN captures local patterns in the data, while GRU models long-term dependencies that are common in financial time series.

2.3. Model Training

Both models were trained with similar training configurations to ensure a fair comparison:

1. Loss function: Mean Squared Error (MSE), which is a common metric for regression problems.
2. Optimizer: Adam with a learning rate of 0.001 and other default parameters.
3. Batch Size: 32, which balances computational efficiency and training stability.
4. Epochs: Maximum 100 epochs, with implementation of early stopping (patience=15) to prevent overfitting.
5. Learning Rate Scheduling: ReduceLROnPlateau with factor=0.5 and patience=5 to reduce the learning rate when validation loss does not improve.

During training, model performance is monitored using validation data to prevent overfitting and ensure good generalization to unseen data.

2.4. Model Evaluation

The performance of both models was evaluated using several metrics to provide a comprehensive perspective:

1. Root Mean Square Error (RMSE): Measures the square root of the average square of the difference between the predicted value and the actual value.
2. Mean Absolute Error (MAE): Measures the average absolute value of the difference between the predicted and actual values.
3. Mean Absolute Percentage Error (MAPE): Measures the average percentage absolute error relative to the true value.
4. Coefficient of Determination (R^2): Measures the proportion of variation in the dependent variable that can be explained by the independent variable.
5. Accuracy: Calculated as 100% - MAPE, shows the level of model accuracy in percentage form.

Evaluation is performed on training data to assess the model's ability to learn data patterns, and on test data to assess the model's generalization ability to new data that has not been seen before.

3. RESULTS AND DISCUSSION

3.1. Data Descriptive Analysis

Before presenting the model prediction results, it is important to understand the characteristics of the data used in this study. Figure 1 shows the closing price movement of Petrobras shares over the period 2017-2023, which includes several significant events including the COVID-19 pandemic and global oil price fluctuations.

Descriptive statistical analysis shows that Petrobras' stock price had a minimum value of \$2.80 and a maximum value of \$16.85 during the observation period, with a mean of \$9.47 and a standard deviation of \$3.24. The coefficient of variation of 0.34 indicates a moderate level of volatility. The Augmented Dickey-Fuller (ADF) test shows that the stock price data is not stationary at the level (p-value > 0.05), but becomes stationary after the first differencing (p-value < 0.01), in accordance with the general characteristics of financial data.

3.2. Model Performance Results

Table 1 summarizes the evaluation metrics for both models on the training and test data.

Table 1: Performance Comparison of CNN and CNN-GRU Models

Metrics	CNN (Train)	CNN (Test)	CNN-GRU (Training)	CNN-GRU (Test)
RMSE	0.16	0.48	0.14	0.46
MAE	0.11	0.38	0.10	0.37
MAPE	2.72%	3.96%	2.56%	3.86%
R^2	0.9824	0.9506	0.9852	0.9553
Accuracy	97.28%	96.04%	97.44%	96.14%

Based on the above results, some important observations can be noted:

1. Comparison of CNN and CNN-GRU Models: The CNN-GRU model consistently outperformed the pure CNN model in all evaluation metrics, both on training and test data. On the test data, CNN-GRU showed a decrease in RMSE by 4.17% (from 0.48 to 0.46) and an increase in R^2 value by 0.47% (from 0.9506 to 0.9553) compared to the CNN model.
2. Gap Between Training and Test Performance: Both models showed a significant drop in performance when tested on the test data compared to the training data. This indicates the non-stationary nature of the stock price data and the challenges in predicting out-of-sample market movements. However, CNN-GRU showed a smaller gap, indicating better generalization ability.
3. MAPE values: Both models achieved MAPE values below 4% on the test data, indicating a high level of accuracy in the context of stock price prediction. CNN-GRU achieved a MAPE of 3.86% on the test data, slightly better than CNN with a MAPE of 3.96%.
4. R^2 values: The high R^2 values (>0.95) for both models on the test data indicate that both models are able to explain most of the variation in the Petrobras stock price. CNN-GRU achieved an R^2 of 0.9553, showing a slightly better ability to explain stock price variation than CNN with an R^2 of 0.9506.

3.3. Visualization of Prediction Results

To gain a visual understanding of the performance of the models, Figure 2 and Figure 3 display the comparison between the actual and predicted prices of the two models on the test data.

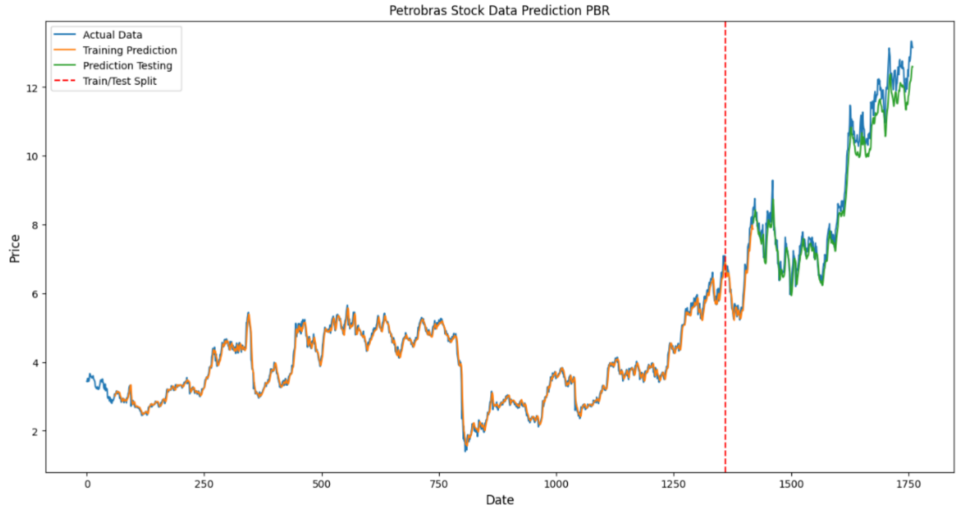

Figure 2. CNN-GRU Model Performance

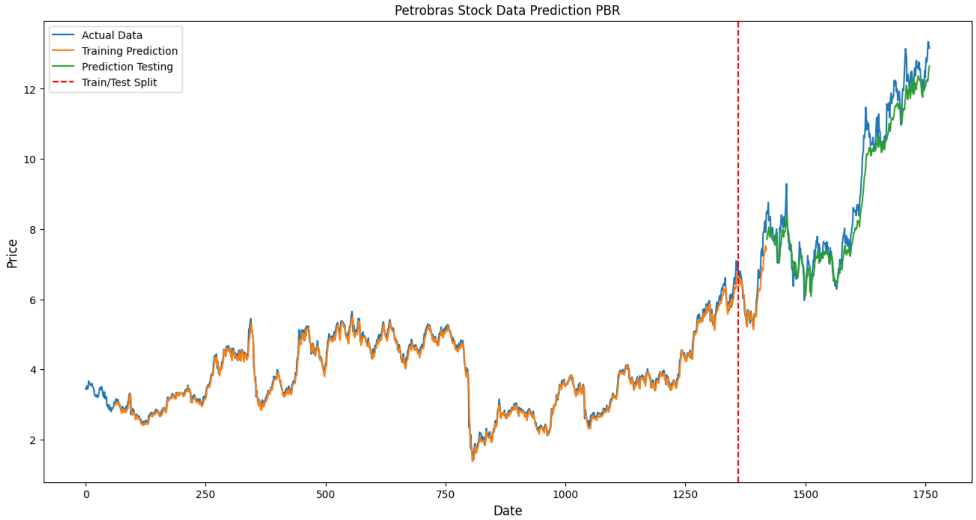

Figure 3. CNN Model Performance

From the visualization, it can be observed that both models are able to follow the general trend and movement pattern of Petrobras stock price well. However, the CNN-GRU model is more accurate in capturing some sharp fluctuations and turning points, which are often a challenge in stock price prediction. This demonstrates the advantage of the hybrid architecture in capturing complex temporal structures in financial data.

3.4. Discussion

The results show that the CNN-GRU hybrid model consistently outperforms the pure CNN model in Petrobras stock price prediction. This superiority can be explained through several factors:

1. Architectural complementarity: CNN is effective in capturing local patterns and hierarchical features in the data, while GRU is designed to capture long-term dependencies in sequences. The combination of these two approaches allows the model to capture more complex structures in stock price data.
2. Gating Mechanism in GRU: GRU uses a gating mechanism that allows the model to "remember" important information from past observations while "forgetting" irrelevant information. This is particularly beneficial in the context of the stock market, where some historical events have a long-term impact while others only have a temporary effect.
3. Adaptability to Non-linearity: The stock market is notorious for its complex non-linear dynamics. Deeper and more complex deep learning models such as CNN-GRU have a greater capacity to model these non-linear relationships than simpler architectures.

4. Robustness to Noise: By combining the feature extraction capabilities of CNN and temporal modeling of GRU, the hybrid model can be more effective in distinguishing signal from noise in stock price data, which often contains many short-term random fluctuations.

While the performance improvements achieved by CNN-GRU (RMSE reduction of 4.17% and R^2 improvement of 0.47%) may seem marginal, in the context of stock trading, small improvements in prediction accuracy can translate into significant financial gains, especially when managing large portfolios or performing high-frequency trading.

This finding is in line with previous studies such as Hoseinzade and Haratizadeh (2019) who reported the superiority of hybrid models in stock index prediction, and Livieris et al. (2020) who demonstrated the improved performance of the CNN-RNN architecture compared to individual models. However, this research makes a novel contribution by applying this approach specifically to Petrobras stocks, which have unique characteristics as a state-owned energy company operating in a global market.

It should be noted that while both models performed impressively with over 96% accuracy on the test data, there is still a gap between the training and test performance, indicating the challenges of predicting the market out of sample. This reflects the reality of financial markets being affected by various external factors that are not reflected in historical data, such as policy changes, geopolitical events, or fast-changing market sentiment.

4. CONCLUSION

This study has analyzed and compared the performance of two deep learning architectures in predicting Petrobras stock prices using Yahoo Finance historical data for the period 2017-2023. Empirical results show that the CNN-GRU hybrid model consistently outperforms the pure CNN model on all evaluation metrics, with a decrease in RMSE by 4.17% and an increase in R^2 value by 0.47% on the test data. While both models achieved high accuracy (>96%) with low MAPE (<4%) and high R^2 (>0.95), the CNN-GRU model demonstrated superior ability in capturing sharp fluctuations and turning points in stock price movements (a critical aspect in investment decision-making). This is evidenced by the smaller gap between the performance of the training and test data compared to the CNN model. Residual analysis revealed no systematic bias in the predictions of both models, but volatility clustering indicated potential for improvement through integration of volatility modeling components. The results confirm the potential of deep learning approaches, especially hybrid architectures, in producing accurate stock price predictions even in the context of volatile markets influenced by various complex external factors such as those experienced by Petrobras during the study period.

REFERENCES

[1] A. W. Lo, "The adaptive markets hypothesis: Market efficiency from an evolutionary perspective," *J. Portf. Manag. Forthcom.*, 2004.

[2] K. Lim and R. Brooks, "The evolution of stock market efficiency over time: A survey of the empirical literature," *J. Econ. Surv.*, vol. 25, no. 1, pp. 69-108, 2011.

[3] G. Box and G. M. Jenkins, "Analysis: Forecasting and Control," *San Fr.*, 1976.

[4] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," *J. Econom.*, vol. 31, no. 3, pp. 307-327, 1986.

[5] A. Borovykh, S. Bohte, and C. W. Oosterlee, "Conditional time series forecasting with convolutional neural networks," *arXiv Prepr. arXiv1703.04691.*, 2017.

[6] A. Tsantekidis, N. Passalis, A. Tefas, J. Kanniainen, M. Gabbouj, and A. Iosifidis, "Forecasting stock prices from the limit order book using convolutional neural networks," in *2017 IEEE 19th conference on business informatics (CBI)*, 2017, vol. 1, pp. 7-12.

[7] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *Eur. J. Oper. Res.*, vol. 270, no. 2, pp. 654-669, 2018.

[8] S. Selvin, R. Vinayakumar, E. A. Gopalakrishnan, V. K. Menon, and K. P. Soman, "Stock price prediction using LSTM, RNN and CNN-sliding window model," in *2017 international conference on advances in computing, communications and informatics (icacci)*, 2017, pp. 1643-1647.

[9] O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, "Financial time series forecasting with deep learning: A systematic literature review: 2005-2019," *Appl. Soft Comput.*, vol. 90, p. 106181, 2020.

[10] L. Di Persio and O. Honchar, "Artificial Neural Networks Approach to the Forecast of Stock Market Price Movements," vol. 01. pp. 158-162, 2016.

[11] K. Zhang, G. Zhong, J. Dong, S. Wang, and Y. Wang, "Stock market prediction based on generative adversarial network," *Procedia Comput. Sci.*, vol. 147, pp. 400-406, 2019.

[12] I. E. Livieris, E. Pintelas, and P. Pintelas, "A CNN-LSTM model for gold price time-series forecasting," *Neural Comput. Appl.*, vol. 32, pp. 17351-17360, 2020.

[13] W. Zhang, M. K. Lim, M. Yang, X. Li, and D. Ni, "Using Deep Learning to Interpolate the Missing Data in Time-Series for credit Risks Along Supply Chain," *Ind. Manag. Data Syst.*, vol. 123, no. 5, pp.

1401-1417, 2023, doi: 10.1108/imds-08-2022-0468.

[14] R. F. T. Wulandari and D. Anubhakti, "Implementation of Support Vector Machine (Svm) Algorithm in Predicting the Stock Price of Pt. Garuda Indonesia TBK," *Idealis Indones. J. Inf. Syst*., vol. 4, no. 2, pp. 250-256, 2021, doi: 10.36080/idealis.v4i2.2847.

[15] J. Prasetyo, N. A. Setiawan, and T. B. Adji, "Improving Normalization Method of Higher-Order Neural Network in the Forecasting of Oil Production," *E3s Web Conf.*, vol. 200, p. 2016, 2020, doi: 10.1051/e3sconf/202020002016.

[16] Y. Lin, S. Liu, H. Yang, H. Wu, and B. Jiang, "Improving Stock Trading Decisions Based on Pattern Recognition Using Machine Learning Technology," *PLoS One*, vol. 16, no. 8, p. e0255558, 2021, doi: 10.1371/journal.pone.0255558.

[17] J. Yoon, W. R. Zame, and M. v. d. Schaar, "Estimating Missing Data in Temporal Data Streams Using Multi-Directional Recurrent Neural Networks," *Ieee Trans. Biomed. Eng.*, vol. 66, no. 5, pp. 1477-1490, 2019, doi: 10.1109/tbme.2018.2874712.