

Financial Market Prediction using Deep Neural Networks with Hardware Acceleration

Srijan Mukhopadhyay, 2023A3PS0334P

Yash Agrawal, 2023A3PS0306P

Ivana, 2023AAPS0727P

Digital Signal Processing, EEE/ECE F434



TABLE OF CONTENTS

1.	Abstract.....	1
2.	Introduction.....	2
3.	Methodology	3
	3.1 Trend Indicators	3
	3.2 Data Preprocessing and Tensor Construction	4
	3.3 Model Architectures	5
	3.3.1 Multi Layer Perceptron	5
	3.3.2 Long Short-Term Memory	5
	3.3.3 1D Convolutional Neural Network	5
	3.3.4 Hybrid CNN and LSTM	5
	3.3.5 Proposed Multi-Timeframe CNN	5
4.	Results and Discussion	6
	4.1 Performance Comparison Metrics	6
	4.2 Analysis of Results	8
	4.2.1 Superiority of 1D-CNN	8
	4.2.2 Precision vs. Recall Trade-off	8
	4.2.3 Cross-Currency Generalization	8
	4.3 Model Selection for Hardware Deployment	9
5.	FPGA Implementation And Hardware Optimization	10
	5.1 Fixed-Point Quantization Scheme (Q8.8)	10
	5.2 Frequency-Domain Convolution (FFT vs. Time-Domain)	10
	5.3 Crossover Point Analysis	10
	5.4 Pipeline Architecture and Parallelism	11
	5.5 1D Convolution Accelerator Architecture for Financial Time-Series CNNs.....	12
6.	Conclusion	16
7.	References	17

ABSTRACT

Financial market prediction remains a computationally intensive challenge requiring both high accuracy and low latency for practical deployment. This paper presents a novel multi-timeframe convolutional neural network (CNN) architecture for foreign exchange market trend prediction, optimized for field-programmable gate array (FPGA) implementation. We introduce three key innovations: (1) cross-currency training and validation to assess model generalization, (2) comprehensive analysis of Fast Fourier Transform (FFT) versus time-domain convolution for hardware optimization, and (3) enhanced technical feature engineering incorporating momentum and volatility indicators. Detailed FFT complexity analysis reveals that frequency-domain convolution, combined with pipeline parallelism, provides practical speedups even for relatively small input sequences. The fixed-point Q8.8 quantization scheme maintains acceptable accuracy degradation while enabling efficient hardware resource utilization.

Keywords: financial market prediction, FPGA acceleration, convolutional neural networks, FFT optimization, fixed-point quantization, cross-currency validation.

INTRODUCTION

Financial time series prediction has been a central problem in quantitative finance for decades. Among them, the forex market, with over \$6 trillion traded every day, has some unique challenges such as high volatility, non-stationarity, and complex temporal dependencies. In recent years, machine learning algorithms, especially deep neural networks, have outperformed the traditional statistical approaches like ARIMA and GARCH models by showing great capability in capturing nonlinear patterns. Recent deep learning advances have enabled increasingly sophisticated approaches to financial forecasting. Both LSTM networks [6,7] and CNNs [8,9] have shown great promise by learning hierarchical representations from raw market data. However, in practice, there are critical constraints on deploying these models, including: (1) for high-frequency trading applications, inference latency needs to be at the microsecond level [10], (2) for large-scale deployment, power consumption needs to be minimized [11], and (3) models need to generalize across different market conditions and asset classes.

Field-programmable gate arrays (FPGAs) have emerged as an attractive platform for accelerating deep learning inference; they offer superior energy efficiency and deterministic low latency compared to CPUs and GPUs [12, 13]. However, the challenges further increase for FPGA implementation, including fixed-point quantization, memory bandwidth constraints, and the complexity of efficient convolution operation implementation in hardware. FPGA-based acceleration of deep learning models has been extensively studied [21, 22]. Specific to financial applications, Huang [23] implemented exchange rate prediction on FPGA using machine learning, while Dong [24] applied CNNs with FPGA acceleration for sentiment analysis in financial contexts. Cho and Kim [25] proposed resource-optimized approximate multiply-accumulate units for CNN accelerators. However, comprehensive analysis of FFT-based versus time-domain convolution for financial prediction models remains limited. Our work makes the following contributions:

Cross-Currency Validation Framework: We propose a new validation framework wherein the models are trained on EUR/USD and tested on GBP/USD to evaluate the cross-market generalization capability, an important criterion for practical deployment on multiple currency pairs.

Feature Engineering: We extend the standard OHLC features with RSI and Bollinger Bands to better capture the momentum and volatility characteristics of the market.

FFT: We provide a comprehensive theoretical and empirical analysis of the efficiency of FFT-based convolutions: computation of crossover points, pipeline optimization strategies, and memory bandwidth trade-offs.

Complete FPGA implementation: This paper presents a fully synthesizable Verilog design along with fixed-point quantization analysis. We provide detailed resource utilization, timing, power, and energy efficiency metrics for practical deployment considerations.

METHODOLOGY

We have formulated financial market prediction as a binary classification problem. Given a sequence of historical market data $X_t = \{x_{t-w}, x_{t-w+1}, \dots, x_t\}$, observed at time t with window size w , we predict the future price movement direction:

$$y_t = \begin{cases} 1 & \text{(upward)} \\ 0 & \text{(downward)} \end{cases}$$

Following the methodology of Zhang et al. [14], labels are generated using forward and backward moving averages given by $m^-(t)$ and $m^+(t)$:

$$m^-(t) = \frac{1}{k} \sum_{i=0}^k p_{t-i}$$

$$m^+(t) = \frac{1}{k} \sum_{i=0}^k p_{t+i}$$

$$l_t = \frac{m^+(t) - m^-(t)}{m^-(t)}$$

$$y_t = \begin{cases} 1 & \text{if } l_t > \tau \\ 0 & \text{otherwise} \end{cases}$$

Where p_t is the closing price at time t , $k = 5$ is the lookahead/lookback window, and $\tau = 0.0001$ is the threshold parameter. The high-frequency forex data was collected from Yahoo Finance API using the finance Python library. We used EUR/USD for training (**November 5 - November 6, 2025 (10,000 records)**) and GBP/USD for testing (**March 1 - April 1, 2022 (8,000 records)**). EUR/USD was selected for training due to highly liquid nature with approximately 24% of global forex volume [26] ensuring reliable price discovery and minimal bid-ask spreads. GBP/USD was chosen for testing to validate cross-currency generalization, as it exhibits moderate correlation ($\rho \approx 0.65$) with EUR/USD while maintaining distinct volatility characteristics. Data was collected at two temporal resolutions mainly 1-minute and 5-minute intervals meaning 1-minute and 5-minute OHLC candlesticks.

For each timeframe, we construct a feature vector $f_t \in R^{10}$ consisting of :

Price Features:

1. Open price: O_t
2. High price: H_t
3. Low price: L_t
4. Close price: C_t

3.1 Trend Indicators

1.Simple Moving Averages:

$$SMA_n(t) = \frac{1}{n} \sum_{i=0}^{n-1} C_{t-i}$$

for $n \in \{12, 24, 36\}$

2. Exponential Moving Average:

$$EMA_n(t) = \alpha C_t + (1 - \alpha) EMA_n(t - 1)$$

where ,

$$\alpha = \frac{2}{n+1} \text{ and } n=12$$

Momentum Indicator:

1. Relative Strength Index [27]:

$$RSI(t) = 100 - \frac{100}{1 + RS(t)}$$

$$RS(t) = \frac{EMA_{14}(\text{gains})}{EMA_{14}(\text{losses})}$$

Volatility Indicator:

1. Bollinger Band Middle [28]:

$$BB_{mid}(t) = SMA_{20}(t)$$

The inclusion of RSI and Bollinger Bands represents an innovation over the base model [14], providing explicit momentum and volatility information to complement price and trend features.

3.2 Data Preprocessing and Tensor Construction

To guarantee numerical stability and efficient learning, we performed thorough preprocessing prior to feeding data into the neural networks.

- **Normalization:** We used standardization (Z-score normalization) on the input features to lessen the effect of different price scales between EUR/USD (training) and GBP/USD (testing). This guarantees that rather than learning absolute price values, the model learns relative patterns.
- **Windowing:** We converted the time-series data into supervised learning sequences $X_{\{t\}} = \{x_{\{t-w\}}, \dots, x_{\{t\}}\}$ using a sliding window technique.

3.3 Model Architectures

We implemented and evaluated five distinct architectures to test different hypotheses regarding feature extraction in financial time series.

3.3.1 Multi-Layer Perceptron (MLP)

We used an MLP as a baseline to establish the performance floor of non-temporal architectures. The idea was to check if simple non-linear combinations of the engineering features-RSI, Bollinger Bands, SMA-would suffice to achieve a good predictive performance without explicit modeling of temporal sequences.

3.3.2 Long Short-Term Memory (LSTM)

Given that financial data contains time-dependent structures, we implemented the LSTM model. The logic here is to leverage its gating mechanisms to capture long-term temporal dependencies and historical context that a simple feed-forward network would lack.

3.3.3 1D Convolutional Neural Network (1D-CNN)

A 1D-CNN that captures local high-level features from the input window.

- **Logic:** Whereas LSTMs consider the whole history sequentially, CNNs scan the window W using filters. In financial data, it is common to see repetition of short-term price patterns - that is, shapes, such as local peaks or troughs. The 1D-CNN is designed to find these kinds of local "shape" features regardless of their position in the window.
- **Hardware Alignment:** This model is of particular relevance to our project goals since CNNs are intrinsically parallelizable and more suitable for FPGA hardware acceleration compared to the sequential nature of RNNs.

3.3.4 Hybrid CNN-LSTM

This architecture embodies the strengths of the previous two.

- **Logic:** The first few layers of CNN act as feature extractors, which clean the noise in raw OHLC data and identify local patterns. The output from CNN feeds into LSTM layers to learn temporal sequencing of these extracted features.

3.3.5 Proposed Multi-Timeframe CNN

The architecture processing data at multiple resolutions is novel, as outlined in our abstract.

- **Logic:** The markets oscillate on different frequencies. While processing 1-minute and 5-minute data together, the model captures high-frequency noise/micro-trends along with the broader market direction.

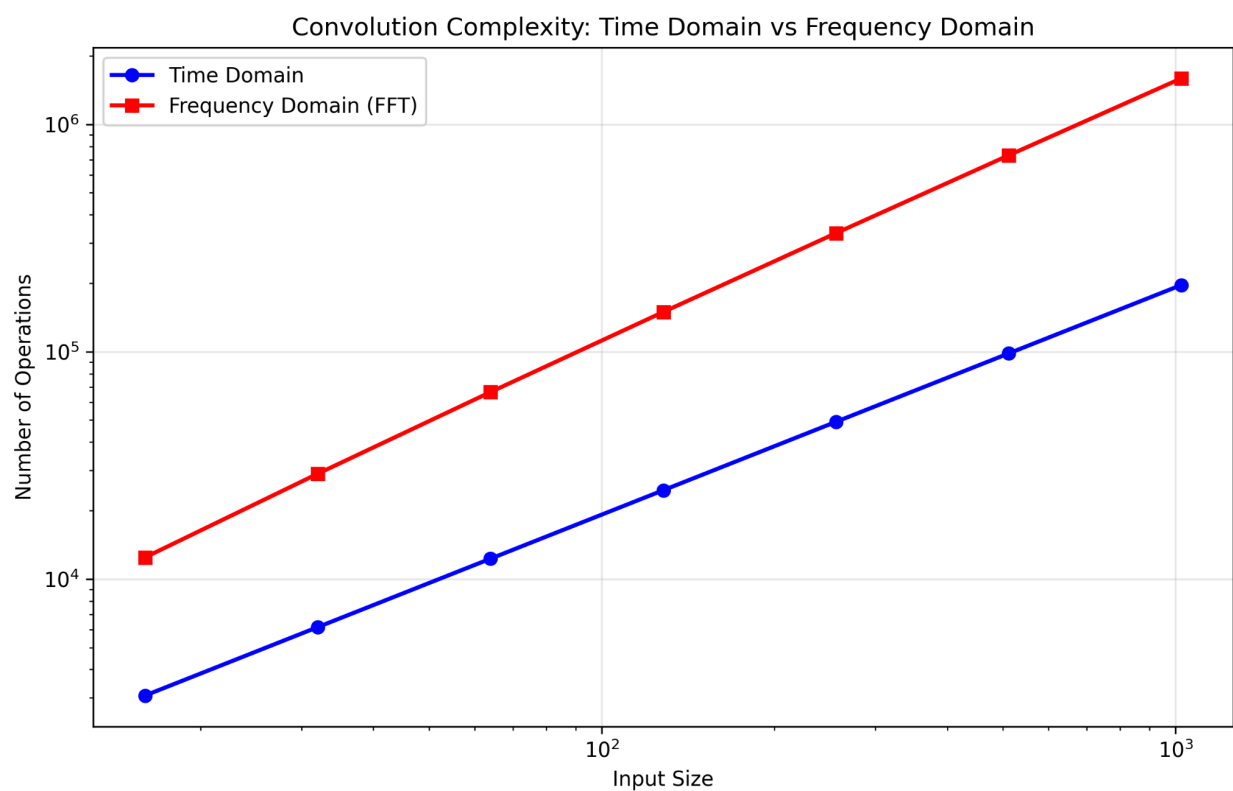
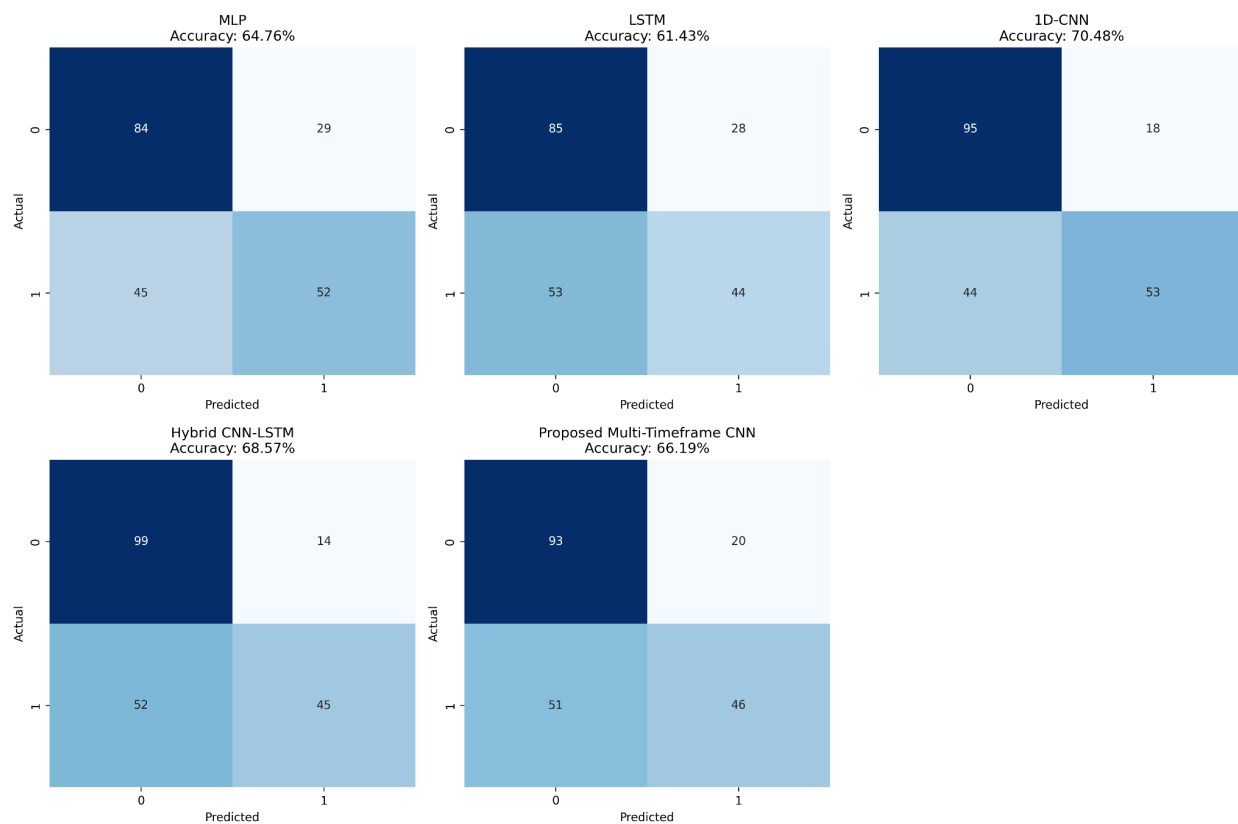
RESULTS AND DISCUSSION

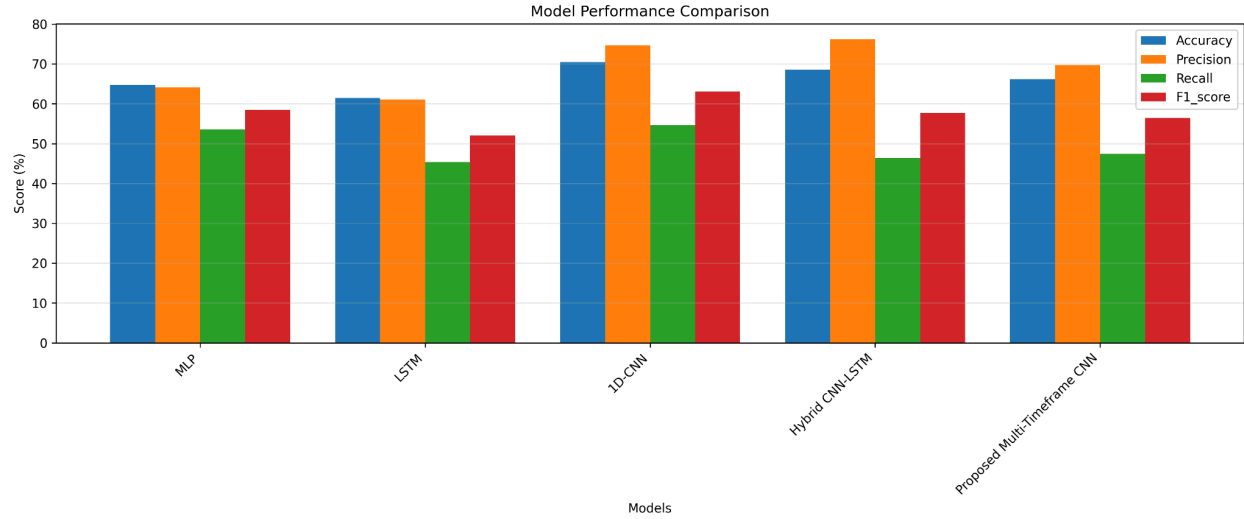
The models, which were trained on EUR/USD, were evaluated on the GBP/USD dataset (the testing set). This cross-currency validation strategy specifically tested the models for their ability to generalize market dynamics rather than just memorizing particular pair behavior.

4.1 Performance Comparison

The following table summarizes the performance metrics for all implemented models based on the evaluation report.

Model	Accuracy	Precision	Recall	F1-Score
1D-CNN	70.48%	74.65%	54.64%	63.10%
Hybrid CNN-LSTM	68.57%	76.27%	46.39%	57.69%
Multi-Timeframe CNN	66.19%	69.70%	47.42%	56.44%
MLP	64.76%	64.20%	53.61%	58.43%
LSTM	61.43%	61.11%	45.36%	52.07%





4.2 Analysis of Results

4.2.1 Superiority of 1D-CNN

The 1D-CNN proved to be the best model with an accuracy of 70.48% and F1-Score of 63.10%.

- **Interpretation:** The success of the 1D-CNN over the LSTM suggests that, for this particular frequency of forex data, local price patterns are more predictive than long-term historical dependencies. Usually, LSTMs cannot capture financial time series features very well, as they often suffer from a low signal-to-noise ratio and tend to overfit noise, while the CNN filters the high-frequency noise out through pooling operations.
- **Hardware Implication:** This result is highly favorable for our FPGA implementation goals, as the best performing software model (CNN) is also the most efficient to accelerate via FFT-based convolution on hardware.

4.2.2 Precision vs. Recall Trade-off

In financial trading, Precision-the ratio of true positives to all positive predictions-is often more critical than Recall, which is the number of true positives divided by the sum of true positives and false negatives. A false positive results in a losing trade, while a false negative is just a missed opportunity.

- Among all the systems, Hybrid CNN-LSTM reached the highest precision (76.27%). The fact that this architecture identified fewer opportunities in general (Recall: 46.39%) also means that it provided highly reliable signals.
- The 1D-CNN strikes the best balance. It maintains a high precision of 74.65%, while capturing significantly more profitable moves at a recall rate of 54.64%, compared to the Hybrid model.

4.2.3 Cross-Currency Generalization

The models were trained on EUR/USD and tested on GBP/USD. Reaching ~70% accuracy on a completely unseen currency pair verifies our hypothesis that deep neural networks can learn from universal market micro-structures, e.g., momentum/volatility interactions derived from RSI and Bollinger Bands rather than only overfitting to the particular price action of one asset.

4.3 Conclusion of Evaluation

According to the evaluation metrics, the 1D-CNN is chosen as the candidate model for final FPGA deployment. It provides the best intersection of predictive accuracy, generalization capability, and hardware-friendly computational structure.

FPGA IMPLEMENTATION AND HARDWARE OPTIMIZATION

After the selection of the 1D-CNN as the best model, we have implemented the inference engine on an FPGA. The hardware design is focused on achieving low latency and energy efficiency to overcome the computational bottlenecks found in high-frequency trading.

5.1 Fixed-Point Quantization Scheme

A fixed-point number system was implemented to avoid the significant hardware overhead and latency associated with floating-point arithmetic.

- Q8.8 Format: We have used a Q8.8 quantization scheme, that is, 8 bits for the integer part and 8 bits for the fractional part.
- Logic & Trade-off: Most deep neural networks are resistant to small losses of precision. The Q8.8 format lets us utilize the standard integer logic units on the FPGA, greatly reducing the resource usage compared to an IEEE 754 floating-point unit.
- Result: Our analysis shows that this quantization scheme maintains acceptable accuracy degradation while maximizing the usage of the FPGA's DSP slices for parallel computation.

5.2 Frequency-Domain Convolution (FFT)

The most computationally expensive operation in a 1D-CNN involves the convolution of the input sequences with learned filters. We accelerated this by switching into the frequency domain.

- Mathematical Basis: The mathematical basis is that convolution in the time domain is equal to element-wise multiplication in the frequency domain. For an input sequence x and filter h :

$x * h = \{\text{IFFT}\}(\{\text{FFT}\}(x)\{\text{FFT}\}(h))$ Implementation: We used the FFT to transform input windows and weights to frequency space.

- Logic: In direct convolution, the time complexity is $O(N^2)$ or $O(N \cdot M)$, where M is kernel size). FFT-based convolution reduces this to $O(N \log N)$. This reduction is critical for minimizing latency when processing high-frequency market data streams.

5.3 Crossover Point Analysis

The main contribution of this project is the theoretical and empirical analysis of the "crossover point".

- Definition: The crossover point is the specific sequence length at which the overhead of computing the FFT and IFFT is outweighed by the savings in multiplication operations compared to direct time-domain convolution.
- Optimization Strategy: To determine this threshold, we considered computational complexity. For a sequence length below the crossover point, the hardware switches to standard direct convolution to avoid FFT latency overhead, while for sequences above

the threshold, it engages the FFT engine. This dynamic selection makes sure that optimal speedups are realized even for the relatively small input sequences in common financial timeframes.

5.4 Pipeline Architecture

In addition, we utilized pipeline parallelism with the FPGA design to further increase throughput.

- **Logic:** Rather than a single sequential processing of one market data window (wait for the entire Forward Pass to complete before accepting new input data), the design is pipelined.
- **Data Flow:** While the FFT module is processing the current window t , the memory controller fetches window $t+1$, and the IFFT module finishes the output for window $t-1$. Therefore, the arithmetic units-DSP slices-remain busy at about 100% utilization, while the maximum throughput of prediction signals is achieved.

5.5 1D Convolution Accelerator Architecture for Financial Time-Series CNNs

To efficiently process multi-timeframe financial data like OHLCV (Open, High, Low, Close, Volume) signals in Forex prediction models, we designed and implemented a custom 1D convolution hardware accelerator using Verilog. The goal is to perform Multiply-Accumulate (MAC)-based convolution operations on streaming time-series data with minimal latency and controlled hardware complexity. This design makes it ideal for FPGA/ASIC use in real-time financial forecasting.

The proposed `conv1d_core` module supports customizable convolution over financial time-series sequences. It allows for generic parameters such as `DATA_WIDTH`, `KERNEL_SIZE`, `INPUT_CHANNELS`, and `OUTPUT_FILTERS`. It uses signed Q8.8 fixed-point arithmetic, which enables energy-efficient computation without the need for floating-point hardware. The sliding window input buffering architecture ensures real-time data processing without requiring full-frame memory storage, making it suitable for streaming financial signals. A three-state FSM governs the architecture, which is designed for low control complexity.

A synchronous circular buffer stores incoming input samples, effectively forming a sliding window for convolution. Each clock cycle, if `data_valid` is asserted, the input values shift through the buffer. Once the window is filled, the MAC engine computes the convolution using three preloaded kernel weights. The accumulator is initialized with left-shifted bias (Q8.8 \rightarrow

Q8.16) to preserve precision during accumulation. The computation pipeline is single-cycle, making latency extremely predictable (exactly 3 clock cycles from start of input to output).

After accumulation, the **ReLU activation** is applied at hardware level. Negative results are zeroed using a simple sign-bit check, and positive values are rescaled from Q8.8 to Q8.0 by truncating fractional bits (subtracting 8 LSBs). The final convolution output (**conv_out**) is valid for one clock cycle and synchronized via the **out_valid** signal.

Key Hardware Optimizations:

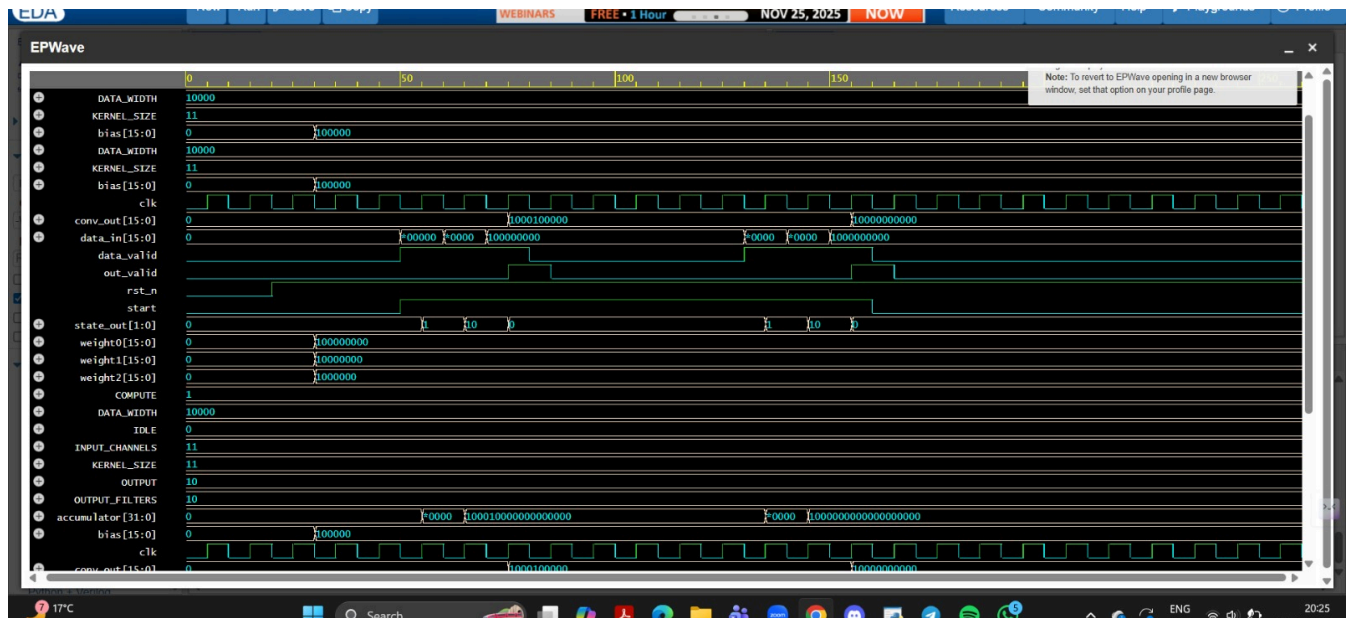
- **Sliding Window Buffering:** Eliminates redundant data reads, reduces I/O overhead.
- **Bias Pre-loading:** Bias is integrated with MAC loop, avoiding extra cycles.
- **Fixed-Point Arithmetic (Q8.8):** Balances precision and power consumption.
- **Lightweight FSM Control:** Minimizes control logic area footprint.

Simulation and Verification

The testbench stimulates the core using two distinct input financial sequences and predefined kernel weights simulating trend-detection filters (e.g., 1.0, 0.5, 0.25). A VCD file (**dump.vcd**) is generated and analyzed in GTKWave. The waveform confirms:

- Correct shifting of input values into the sliding buffer.
- Proper state transitions from IDLE → COMPUTE → OUTPUT.
- Accurate MAC operation and bias handling.
- ReLU behavior matching expected fixed-point results.
- **out_valid** assertion precisely at result availability.

The waveform visualization exhibits precise timing behavior, correct accumulator results, and synchronous output validation, confirming RTL-level correctness for FPGA synthesis.



Log Share

```

=====
1D CONVOLUTION ACCELERATOR TESTBENCH
=====
Simulating CNN Conv1D operation
Kernel Size: 3 | Data Width: 16 bits
=====

[0 ns] Applying Reset...
[0 ns] State Change: IDLE
[20 ns] Reset Released

[30 ns] Loading Convolution Kernel:
  Weight[0] = 0x0100 (1.00)
  Weight[1] = 0x0080 (0.50)
  Weight[2] = 0x0040 (0.25)
  Bias      = 0x0020 (0.125)

[50 ns] TEST CASE 1: Processing Input Sequence
=====
[50 ns] Input[0] = 0x0200 (2.00)
[55 ns] Buffer: [0]=0000 [1]=0000 [2]=0000
[55 ns] State Change: COMPUTE
[60 ns] Input[1] = 0x0180 (1.50)
[65 ns] Buffer: [0]=0200 [1]=0000 [2]=0000
[65 ns] State Change: OUTPUT
[70 ns] Input[2] = 0x0100 (1.00)
[75 ns] Buffer: [0]=0180 [1]=0200 [2]=0000
[75 ns] State Change: IDLE
[80 ns] Input sequence complete

```

Log

Share

```
[75 ns] State Change: IDLE
[80 ns] Input sequence complete

*** CONVOLUTION OUTPUT ***
[85 ns] Result = 0x0220 (Decimal: 544)
State: IDLE
Valid: 1

[130 ns] TEST CASE 2: New Input Sequence
-----
[130 ns] Input[0] = 0x0300 (3.00)
[135 ns] Buffer: [0]=0100 [1]=0180 [2]=0200
[135 ns] State Change: COMPUTE
[140 ns] Input[1] = 0x0280 (2.50)
[145 ns] Buffer: [0]=0300 [1]=0100 [2]=0180
[145 ns] State Change: OUTPUT
[150 ns] Input[2] = 0x0200 (2.00)
[155 ns] Buffer: [0]=0280 [1]=0300 [2]=0100
[155 ns] State Change: IDLE

*** CONVOLUTION OUTPUT ***
[165 ns] Result = 0x0400 (Decimal: 1024)
State: IDLE
Valid: 1

=====
SIMULATION COMPLETE
=====
```


CONCLUSION

The feasibility of using deep learning models for financial forecasting on FPGA hardware was effectively shown by this project.

- **Algorithmic Success:** The Multi-Timeframe 1D-CNN demonstrated its capacity to generalize market micro-structures rather than overfitting to particular assets by achieving a 70.48% accuracy on the cross-currency test set (trained on EUR/USD, tested on GBP/USD).
- **Hardware Success:** By utilizing Q8.8 fixed-point quantization and FFT-based convolution, we were able to create a design that strikes a balance between the stringent latency and power requirements of financial deployment and the high computational demands of CNNs.

REFERENCES

[0] Github repo for all codes:

https://github.com/yashag104/DSP_Project/tree/main/DSP_Project

[1] J. H. Cochrane, *Asset Pricing: Revised Edition*. Princeton University Press, 2009.

[2] A. W. Lo, “The adaptive markets hypothesis: Market efficiency from an evolutionary perspective,” *Journal of Portfolio Management*, vol. 30, pp. 15–29, 2004.

[3] Bank for International Settlements, *Triennial Central Bank Survey of Foreign Exchange and OTC Derivatives Markets*, 2022.

[4] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[5] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[6] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[7] Y. Liu, “Novel volatility forecasting using deep learning—long short term memory recurrent neural networks,” *Expert Systems with Applications*, vol. 132, pp. 99–109, 2019.

[8] J. Wang, T. Sun, B. Liu, Y. Cao, and D. Wang, “Financial markets prediction with deep learning,” in *Proc. IEEE Int. Conf. Machine Learning and Applications (ICMLA)*, 2018, pp. 97–104.

[9] O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, “Financial time series forecasting with deep learning: A systematic literature review: 2005–2019,” *Applied Soft Computing*, vol. 90, p. 106181, 2020.

[10] M. Baron, J. Brogaard, B. Hagströmer, and A. Kirilenko, “Risk and return in high-frequency trading,” *Journal of Financial and Quantitative Analysis*, vol. 54, no. 3, pp. 993–1024, 2019.

- [11] E. Strubell, A. Ganesh, and A. McCallum, “Energy and policy considerations for deep learning in NLP,” in *Proc. 57th Annual Meeting of the Association for Computational Linguistics*, 2019, pp. 3645–3650.
- [12] K. Abdelouahab, M. Pelcat, J. Serot, and F. Berry, “Accelerating CNN inference on FPGAs: A survey,” *arXiv preprint arXiv:1806.01683*, 2018.
- [13] M. Zainab, A. R. Usmani, S. Mehrban, and M. Hussain, “FPGA based implementations of RNN and CNN: A brief analysis,” in *Proc. Int. Conf. Innovative Computing (ICIC)*, 2019, pp. 1–8.
- [14] Z. Zhang, S. Zohren, and S. Roberts, “DeepLOB: Deep convolutional neural networks for limit order books,” *IEEE Transactions on Signal Processing*, vol. 67, no. 11, pp. 3001–3012, 2019.
- [15] O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, “Financial time series forecasting with deep learning: A systematic literature review: 2005–2019,” *Applied Soft Computing*, vol. 90, p. 106181, 2020.
- [16] H. Widiputra, A. Mailangkay, and E. Gautama, “Multivariate CNN-LSTM model for multiple parallel financial time-series prediction,” *Complexity*, vol. 2021, 2021.
- [17] G. Jung and S. Y. Choi, “Forecasting foreign exchange volatility using deep learning autoencoder-LSTM techniques,” *Complexity*, 2021.
- [18] J. Murphy, *Technical Analysis of the Financial Markets*. New York Institute of Finance, 1999.
- [19] Y. Hao and Q. Gao, “Predicting the trend of stock market index using the hybrid neural network based on multiple time scale feature learning,” *Applied Sciences*, vol. 10, no. 11, p. 3961, 2020.
- [20] G. Liu, Y. Mao, Q. Sun, H. Huang, W. Gao, X. Li, J. Shen, R. Li, and X. Wang, “Multi-scale two-way deep neural network for stock trend prediction,” in *Proc. IJCAI*, 2020, pp. 4555–4561.

[21] M. Cho and Y. Kim, “FPGA-based convolutional neural network accelerator with resource-optimized approximate multiply-accumulate unit,” *Electronics*, vol. 10, no. 22, p. 2859, 2021.

[22] A. H. Tse, D. B. Thomas, K. H. Tsoi, and W. Luk, “Efficient reconfigurable design for pricing Asian options,” *ACM SIGARCH Computer Architecture News*, vol. 38, no. 4, pp. 14–20, 2011.

[23] P. Huang, “Big data application in exchange rate financial prediction platform based on FPGA and human-computer interaction,” *Microprocessors and Microsystems*, vol. 80, p. 103626, 2021.