



D23-047

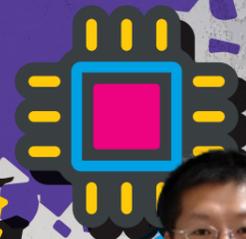
## 應用於植入式神經傳感系統之高效支持 向量機學習加速器

An Energy-Efficient SVM Learning Accelerator for  
Implantable Devices of Neural Sensing System

隊伍名稱 | 滴西欵斯喵喵喵

DCS Meow Meow Meow

隊長 | 謝伊妍 / 臺灣大學電子工程學研究所



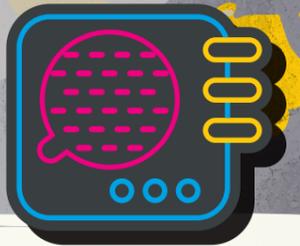
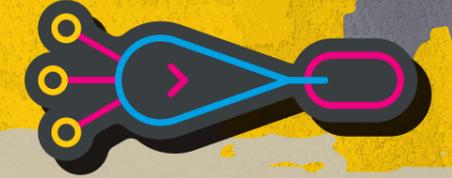
指導教授

楊家驥 | 臺灣大學電機工程學系

美國加州大學洛杉磯分校電機博士，現為臺灣大學電機工程學系教授。實驗室致力於開發低功耗之客製化晶片以提升資料處理速度與能量效率。

研究領域

AI晶片設計、基頻通訊積體電路、生醫訊號處理晶片設計



### 作品摘要

隨著神經傳感技術的進步，偵測並辨識腦波相關特徵，在神經疾病調控和腦機介面中扮演了重要的角色。透過腦波模式識別，感知腦部動態進行神經疾病控制及腦機介面已被應用於現實場景。

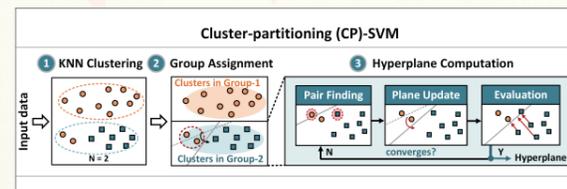
其中，支持向量機 (support vector machine, SVM) 在神經信號處理中被廣泛應用，可以有效地分類和識別腦部信號，並能在訓練有限訓練樣本的情況下，實現優於一般神經網路的性能。然而其傳統的軟體實現方式，存在較大的延遲和能耗問題。因此本作品開發一支持SVM學習加速器，目標為提高嵌入式神經傳感系統的能效和處理速度，同時具備高度擴展性，替需要高準確度、高能效的神經傳感系統，帶來一嶄新的解決方案。

此設計具備以下設計特點：

1. 支持訓練和推論過程，具備晶片上模型調適 (on-chip model adaptation) 的功能，針對高度時變性的神經信號能維持模型的高準確度。
2. 在演算法最佳化上，引入叢集分割支持向量機 (cluster-partitioning SVM) 加速整體訓練與推論流程，相較於傳統SVM，在訓練收斂速度上提升數十倍，並大幅減少推論時間，整體演算法流程如圖一所示。
3. 在硬體面積與延遲最佳化上，透過處理單元最佳化與核心計算器架構探索，達到最低的面積時間乘積。此外也引入稀疏感知最佳化，避免冗餘計算降低整體延遲。並提出跨叢集分類器，在運算遮罩步驟前進行排序，減少約一半的硬體實現面積。同時提出鍊狀資料交換器，僅將相鄰的PE相連，大幅減少資料交換器的面積。最後採用了以二為基底的核心取代高斯函數以避免乘法運算，進而大幅減少處理單元的硬體實現面積。

本作品採用40nm CMOS工藝製造，核心面積為1.4mm<sup>2</sup>，整合了2.1M邏輯閘。本作品的研究成果實現了極高的能量效率與面積效率，與過往文獻的最佳成果相比提升了數十倍。

本設計開發之低能耗SVM學習加速器，對於神經感測和腦機介面領域具有重要意義，它能夠實現更快速、更精確的神經狀態檢測，為患者提供一條嶄新的道路，也能與電刺激器或其他致動器偕同整合，大幅提升未來治療的效率和效果。本作品之應用前景非常廣泛，將對許多領域產生重要的影響，包括腦機介面技術、神經運動控制、人機互動等，同時也期望在未來醫學研究領域，為人類健康帶來更多福祉。



圖一 叢集分割支持向量機演算法流程。

### Abstract

Neural sensing technology advancements have been crucial for detecting and recognizing brainwave-related features in neurological disorders and brain-machine interfaces (BCI). Support vector machines (SVMs) are commonly used in neural signal processing to classify and identify brain signals, outperforming general neural networks with limited training samples. However, traditional implementations of SVMs suffer from long hardware latency and power limitations. To address this, we developed an energy-efficient SVM learning accelerator for implantable neural sensing systems. It features:

1. On-chip model adaptation supporting both training and inference to maintain high accuracy for time-varying neural signals.
2. Introduction of the cluster-partitioning (CP)-SVM algorithm optimization, leading to faster training convergence and a significant reduction in inference time compared to traditional SVMs.
3. Hardware optimization achieved through processing unit and core calculator architecture exploration, employing sparsity-aware skipping to reduce latency, a cross-cluster sorter to decrease hardware area significantly, a chained data exchanger to minimize the data exchanger's area, and adopting a base-2 kernel transformation to replace Gaussian kernel functions, resulting in a notable reduction in processing unit area.

The proposed SVM learning processor architecture (Fig. 2) comprises a scheduler array, a processing element (PE) array, and a classifier. The scheduler array hierarchically distributes clustering computations among multiple PEs, with each scheduler assigned to process data for one cluster. The PE array calculates the scoring function for inference or the evaluation error for training, utilizing an L1-norm calculator and CORDICs to compute the SVM kernel function. The classifier includes a cluster assigner for partitioning and assigning inputs to PEs, and an evaluator for classification decisions based on assigned labels.

The proposed processor integrates 2.1M logic gates within a 1.4mm<sup>2</sup> core area. Compared to a conventional SVM kernel configuration, it achieves much higher area and energy efficiency for inference and training. This work presents an energy-efficient SVM learning accelerator for brain pattern recognition, implementing CP-SVM on silicon and enabling fast learning with a massive electrode array, showcasing a promising, efficient brain pattern recognition kernel for neural sensing systems.

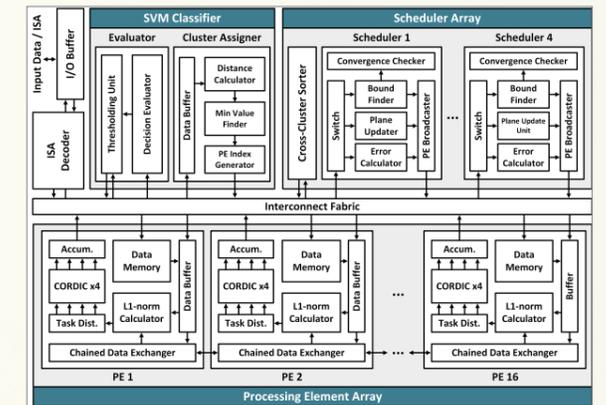


Fig.2 System architecture.