Hippocampus-Inspired Artificial Neural Network Enables Robust Classification under Sparsity: Structured versus Brute-Force Robustness

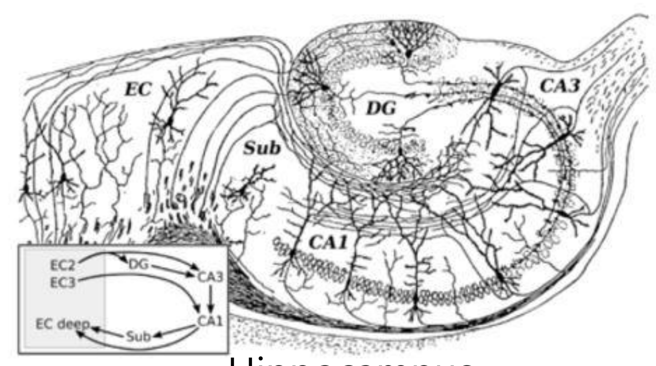
Sang-Yoon Kim and Woochang Lim
Institute for Computational Neuroscience and Department of Science Education, Daegu National University of Education, Daegu 42411, S. Korea

Motivation and Central Question

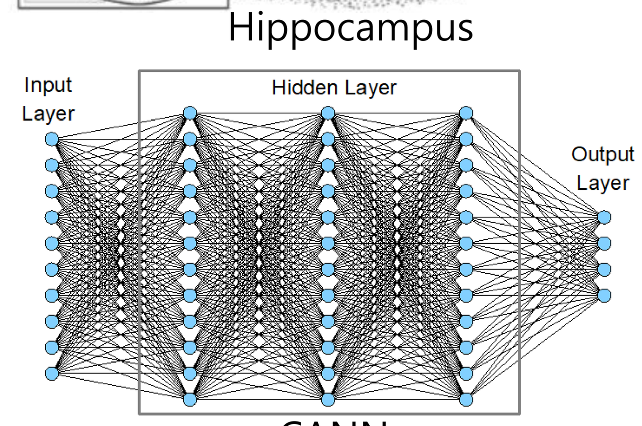
Motivation: Why Study Structural Sparsity?

- **Biological Hippocampus**
  (Evolved over hundreds of millions of years across vertebrate species for efficient memory processing)
  Sparse and Structured Architecture:
  - Competitive sparsification in DG
  - Recurrent stabilization in CA3
  - Integrative refining in CA1
- **Conventional ANN (CANN)**
  - Feedforward multilayer perceptron (MLP)
  - Robustness typically relies on dense connectivity

Central Question:
Do hippocampal circuit principles yield robust classification as connectivity becomes sparse?



Hippocampus

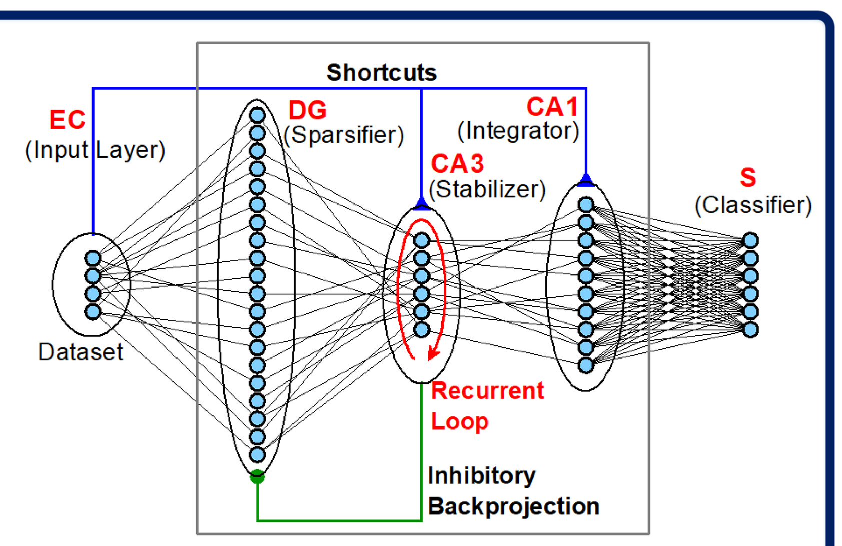


CANN

Architectures: OurANN vs CANN

- **OurANN: Hybrid Classifier Integrating:**
  (a) Hippocampus-Inspired Modular Architecture with Cooperative Functional Modules
  (b) Engineering Machine Learning

  3 Functional Modules: DG, CA3, CA1
  - EC (input layer)→DG→CA3→CA1→S (output classifier)
  - DG: competitive sparsification
  - CA3: recurrent stabilization (temporal recurrence)
  - CA1: integrative refinement
  - Shortcut connections: EC→CA3, EC→CA1: direct input supply
  - Inhibitory backprojection: CA3→DG: enhanced sparsification
  - Population sizes: DG : CA3 : CA1 = 100 : 30 : 50 (anatomical ratio)



OurANN (Hippocampus-Inspired ANN)

- **CANN (Conventional ANN: Baseline)**
  3 Hidden Layers
  - Input → Hidden → Hidden → Hidden → Output
  - Same unit counts: 100 : 30 : 50 → action selection device (gearbox in an auto)

Computational Setup and Sparsity Control

- **Input Dataset**
  - MNIST handwritten digits (0-9)
  - 28 × 28 grayscale images → 784-D flattened vector input to EC
  - 10-class classification
- **Training Protocol**
  - Supervised Machine Learning
  - Identical training schedule for both models
  - Multiple random realizations
- **Structural Sparsity**
  - p_c: inter-layer connection probability
  - Sparse projections instantiated directly with probability p_c
  - Sweep: 1.0 → 0.5 → 0.1 → 0.05 → 0.01
  - Final readout layer kept fully dense (OurANN: CA1 → S, CANN: Hidden → Output)
- **Layer-Size Scaling**
  - η: layer-size scaling factor
  - Multiplies all module/hidden sizes in both models
- **Baseline Condition**
  - Layer-size scaling: η=1
  - Results shown first for unscaled networks
- **Sparsity Regimes**
  Connection Probability Sweep
  - p_c: inter-layer connection probability
  - Evaluated at: 1.0 → 0.5 → 0.1 → 0.05 → 0.01
- **Defined Regimes**
  - Dense & Moderate: p_c =1.0 ~ 0.1
  - Sparse: p_c = 0.1 ~ 0.05
  - Extremely Sparse: p_c = 0.05 ~ 0.01
- **Analysis Strategy**
  - Compare OurANN vs CANN within each regime
  - Track how robustness and performance changes as p_c decreases
  - Identify where architectural differences emerge

Robustness Metrics

- **Local Metrics**
  - Training & Test Accuracy
  - Efficiency = Accuracy / p_c (p_c : representative proxy for overall system cost)
  - Trade-off Score = Accuracy × $(1-p_c)$ (balances performance and wiring cost)
- **Global Metrics: Degradation Rate δ**
  - Plot: Accuracy vs. $\ln(p_c)$
  - Fit: $A(p_c)=a+b\ \ln(p_c)$
  - Degradation rate δ = $|b|$
  - Meaning: speed of accuracy collapse as sparsity increases
- **Robustness Index R**
  - Defined as: R = ("mean accuracy over p_c") × "survival factor"
  - Survival factor = $1-\delta$ (persistence across sparsity range)

Results I: Global Robustness across Sparsity Regimes

Structural Robustness Emerges Under Increasing Sparsity
Main Message: OurANN degrades more slowly and maintains higher robustness than CANN when connectivity becomes sparse.

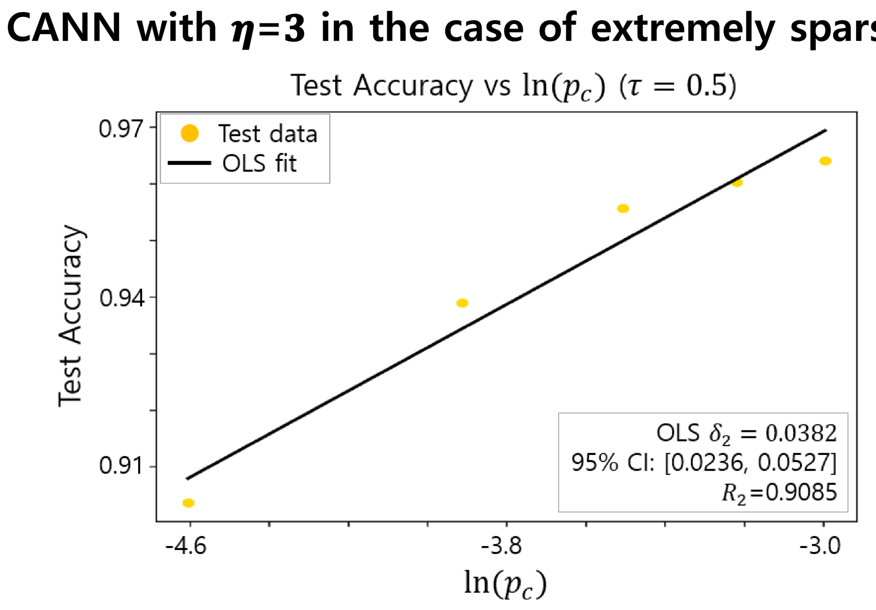
Test-based Degradation Rate (δ) and Robustness Index (R); values shown as: (δ,R) :

Model	Dense/Moderate (1.0~0.1)	Sparse (0.1~0.05)	Extremely Sparse (0.05~0.01)
OurANN	(0.0101, 0.9631)	(0.0302, 0.9175)	(0.0431, 0.8703)
CANN	(0.0105, 0.9637)	(0.0415, 0.9034)	(0.3466, 0.5023)

Interpretation
Dense / Moderate: The two models are essentially indistinguishable.
Sparse: OurANN begins to degrade more slowly than CANN.
Extremely Sparse: OurANN maintains structural robustness, while CANN collapses rapidly.

Results II: Limits of Brute-Force Scaling for CANN

Brute-Force Scaling for CANN with η=3 in the case of extremely sparsity



Test-based global metrics (δ,R):
(0.0382, 0.9085) for η=3 vs (0.3466, 0.5023) for η=1 [cf., OurANN with η=1: (0.0431, 0.8703)]
Brute-force scaling (η=3) yields degradation rates and robustness indices comparable to OurANN at η=1, but only at the cost of substantially increased parameter counts (parameter redundancy).

Structural Robustness in OurANN vs. Brute-Force Scaling Robustness in CANN

Conclusions

- **Main Findings:**
  - **Dense / Moderate (1.0 ~ 0.1)**: OurANN ≈ CANN
  - **Sparse (0.1 ~ 0.05)**: **OurANN degrades more slowly than CANN**
  - **Extremely Sparse (0.05 ~ 0.01)**: **OurANN maintains robustness; CANN collapses**
  - **CANN recovers only via brute-force scaling** (η=3)
- **Key Interpretation:**
  Emergent Structural Robustness vs Brute-Force Scaling Robustness
  - Hippocampus-inspired architecture yields **structural robustness** under sparsity.
  - **Structural robustness** emerges from cooperative dynamics among functional modules DG, CA3, and CA1, supported by shortcuts and inhibitory backprojection.
  - CANN relies on **brute-force scaling via parameter expansion.**
- **Future Works:**
  (1) Structural Robustness under Dataset Complexity: MNIST-Family → CIFAR-10/100 → ImageNet
  (2) Dynamical Memory System:
  Extending OurANN beyond Classifier to A Dynamical Memory System