Attractor network ensemble

- One of the main trends of research in Machine Learning, along with "deep learning", is the study of "networks of networks".
- One central question of interest is comparing a net of nets with overall N neurons and K degree with only one single net with the same size (constant $N \times K$) and check their different performances in terms of storage/load, computational complexity, etc.
- We introduced an ensemble of attractor networks using a simple divide-and-conquer strategy

$$N \times K \equiv N \times K_b \times n$$

- where $K_b = K/n$ for n network modules.

Computation with Neural Networks

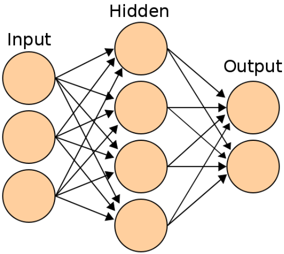
- **Distributed (parallel)**=Memory+Processing vs. Von Neumann Architecture (CPU+RAM).
- **Robust and tolerant:** Deal with noisy data, is an associative device.
- **Dynamics:** Learning (synaptic), Retrieval (recursive update). Memories have an attraction basin.

Applications

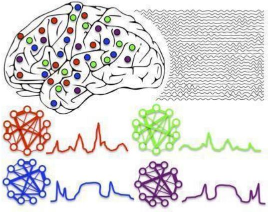
- Neuroscience modeling:
  - Memory (short x long)
  - Motor behavior
  - Sensory discrimination...
- Engineering:
  - Pattern classification, (retrieval, generalization...)
  - Data mining
  - Shape recognition
  - Heuristic algorithms

Neural network models

- From simple tasks (classify): **Perceptron**



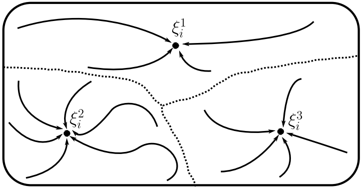
- To complex ones (learn): **Associative Memory.**

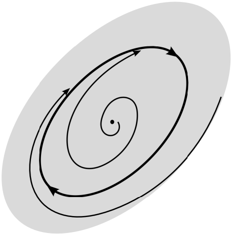


Attractor Neural Networks

An Attractor Neural Network (ANN)

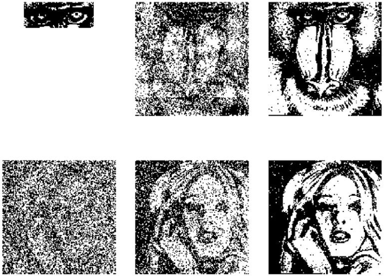
- is an arrangement of nodes (neurons) often recurrently connected
- its time dynamics relaxes to a stable memory state (attractor)
- memories have a basin of attraction





Attractor networks as associative memory devices

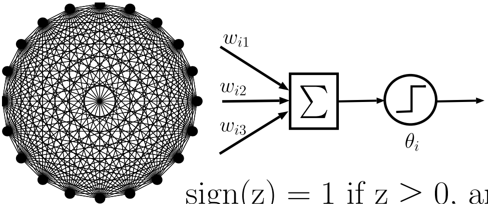
- Two concepts or stimuli are associated when the experience of one leads to the effects of another.
- Associative memory is content addressable.
- Attractor networks are capable of performing pattern denoising and pattern completion



The Hopfield network

The basic problem is:

- To store a set of P patterns ξ_i.
- When presented a new pattern ζ_i, respond by producing whichever one of the stored patters most closely resembles ζ_i.
- Hebbian learning rule: $W_{ij}^{\mu} = W_{ij}^{\mu-1} + \xi_i^{\mu}\xi_j^{\mu}$

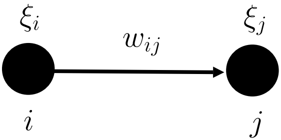


$$\sigma_i := \text{sign}\left(\Sigma_j\, w_{ij}\sigma_j - \theta_i\right)$$

$$\text{sign(z)} = 1 \text{ if z} \geq 0, \text{ and sign(z)} = -1 \text{ if z} < 0$$

Training phase:
A set of patterns ξ_i^μ, $\mu = 1, ..., P$, is presented to the network during a training phase, and the connection strengths w_{ij} are adjusted.

$$\textit{Training pattern:} \quad \vec{\xi^\mu} = \{\xi_1, \xi_2, \dots\}$$



$$w_{ij} = \xi_i \xi_j \qquad \textit{Hebb rule}$$

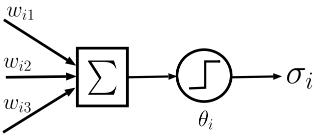
$$W_{ij}^\mu = W_{ij}^{\mu-1} + \xi_i^\mu \xi_j^\mu$$

Neural threshold updating:

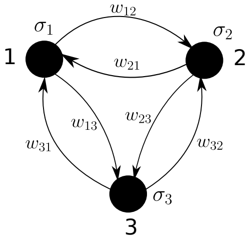
$$\sigma_i^0 \equiv \xi_i^\mu$$

$$\sigma_i^{t+1} = \mathrm{sign}(h_i^t - \theta),$$

$$h_i^t \equiv \sum_j J_{ij}\sigma_j^t, \quad J_{ij} \equiv C_{ij}W_{ij}$$



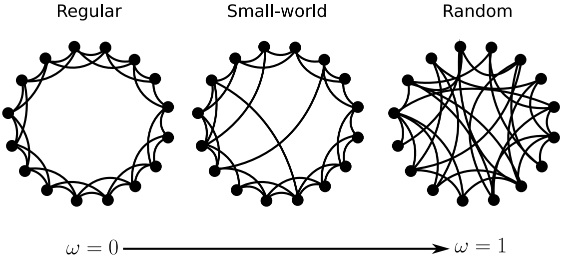
State of the network: $\vec{\sigma} = \{\sigma_1, \sigma_2, \sigma_3\}$



Information measures (Quality, Quantity):

$$m \equiv \frac{1}{N}\sum_i^N \xi_i\sigma_i, \text{ (overlap)}, \quad \alpha = P/K \text{ (pattern load)}$$

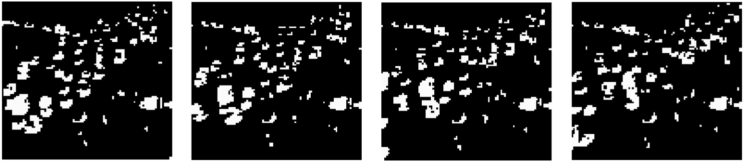
A “prototype” complex network (architecture)



Small-world network: large clustering coefficient, small path length.

- Topology matrix: $\mathbf{C} = \{C_{ij}\}$.
- $C_{ij} \in \{0, 1\}$ splits in local K_L and random links K_R.
- Network parameters:
  - *Connectivity ratio:* $\gamma = K/N$.
  - *Randomness ratio:* $\omega = K_R/K$.

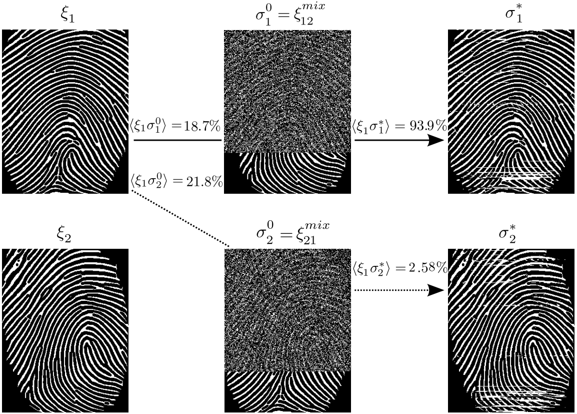
- Learn and retrieve sequences of sparse-coding correlated patterns: automotive video.
    - Content-based traffic video retrieval using a holistic method retrieving a complete video from a (possible noisy) query frame.
- **We had to use a densely connected network ($N = 8544$, $\gamma \sim 0.7$) in order to learn a reasonable amount of frames.**



González, M., Dominguez D., and Sánchez A. Learning sequences of sparse correlated patterns using small-world attractor neural networks: An application to traffic videos. Neurocomputing 74.14 (2011): 2361-2367.

Storing fingerprints patterns

- Forensic applications dealing with a **small set of fingerprints.**
  - The case of separating latent overlapped fingerprints, where one can extract the intended fingerprint from an overlapped impression in a crime scene, according the number of long-range shortcuts.



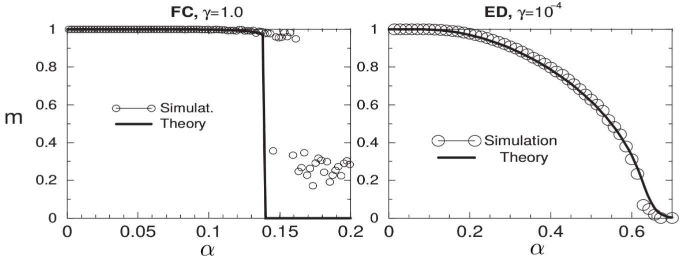
González, M., Dominguez D., Sánchez A. and Rodríguez FB. Retrieval of noisy fingerprint patterns using metric attractor networks. International journal of neural systems 24.07 (2014).

Increasing storage capacity with an ensembled model

- Attractor networks are rich models in neuroscience and engineering applications.
- Their dynamical properties are useful in a variety of applications such as denoising and pattern completion.
- But its limitation in terms of storage capacity are well known.
- How can we increase the storage while keeping the computational cost: using an attractor network ensemble.

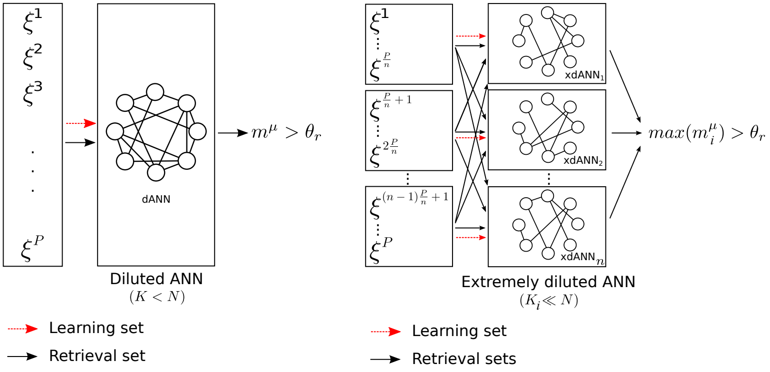
Increasing storage capacity (random unbiased patterns)

- Full connectivity:
  - Costly wiring and heavy computation due to extensive feedback.
- Diluted networks:
  - Wiring saving and light computation.
  - Increased storage capacity in terms of $\alpha = P/K$.



- **We can take advantage of the continuous transition in the diluted case to build modules in an ensemble.**
- If we allow some noise while discriminating retrieval from non retrieval.

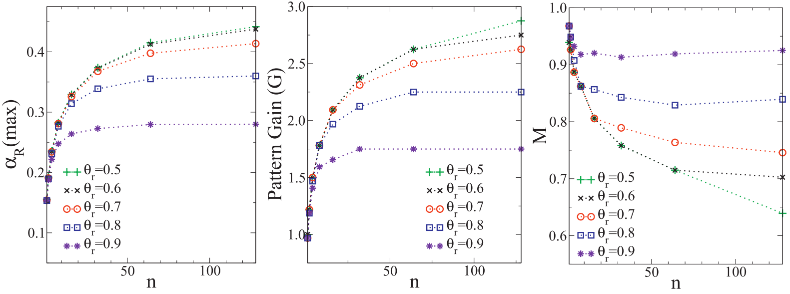
Ensemble of extremely diluted ANN (xdANN)



- **We keep the computational cost constant** for the single dANN ($\boldsymbol{K}$), and the ensemble of xdANN ($\boldsymbol{K_i}$) with $\boldsymbol{n}$-components, in terms of number of connections: $K = n \times K_i$.

Ensemble Performance

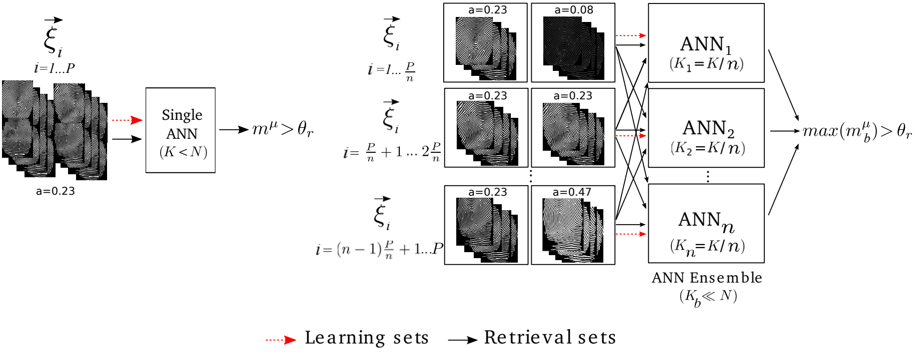
- The larger the number of components is, the larger the pattern (load) gain of the ensemble system.
- We triple the number of stored patterns while discriminating between pattern subsets assigned to each component in the ensemble.



González, Mario, et al. "Increase attractor capacity using an ensembled neural network." Expert Systems with Applications 71 (2017): 206-215.

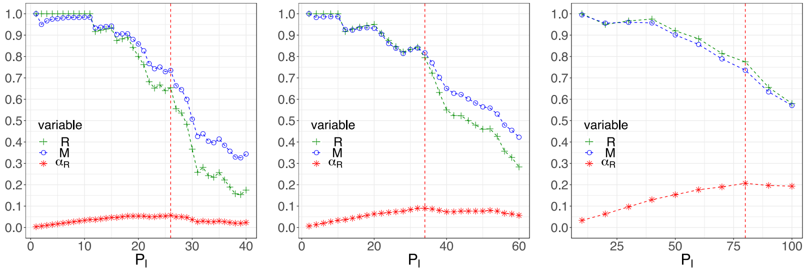
Ensemble of ANN for fingerprints retrieval

1. **Deal with real patterns**, i.e. spatially structured and activity biased (e.g. **fingerprints**).
2. Test the retrieval capacity of the EANN model for such patterns.
3. Compare with a single attractor of **equal connectivity**.
4. **Specialize the network modules** for patterns subsets with different features (i.e. pattern activity).



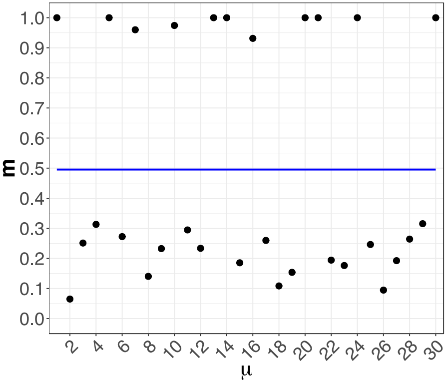
EANN fingerprint retrieval performance I

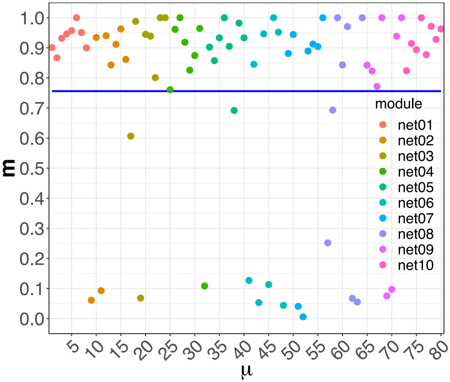
- Fingerprints dataset with $a \sim 0.23, \theta = 0.656$. Systems with $N = 89420, K = 300, n = \{1, 2, 10\}$.



$$R = \frac{P_r}{P_l}, \quad M = \langle m^\mu \rangle = 1/P_l \sum_{\mu=1}^{P_l} m^\mu, \quad \alpha_R = \frac{P_r}{K_b \times n}, \quad G = \frac{P_r^e}{P_r^s}$$

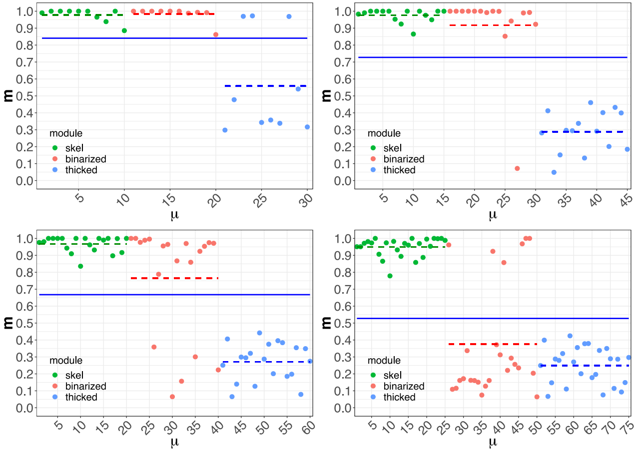
- **Fingerprint retrieval examples for $n = \{1, 10\}$.**





EANN modules specialization (ensembled voting)

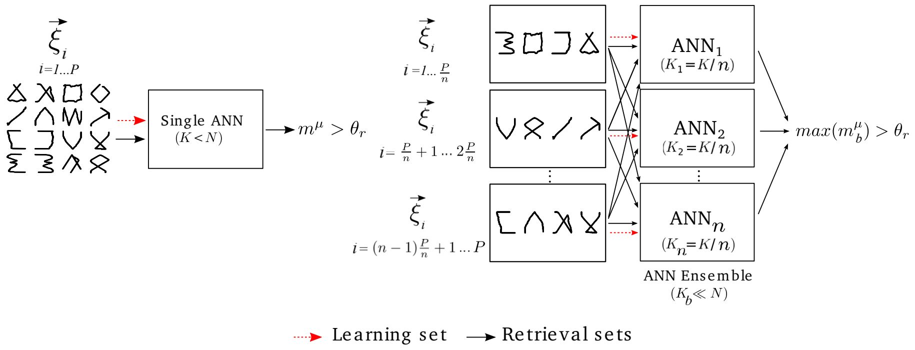
- Module 1: **skeletonized fingerprints** $a \sim 0.0844,\ \theta = 1.4951$), module 2: **normal binarized** ($a \sim 0.2258,\ \theta = 0.656$), and module 3: **thickened fingerprints** ($a \sim 0.4660,\ \theta = 0.0681$).



EANN retrieval and modules specialization

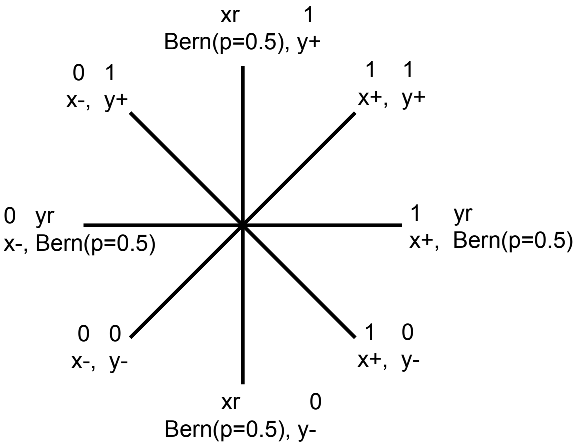
- With the same connectivity cost the EANN system **more than triple the storage capacity** of the single network.
- This gain is higher that the one found for random patterns, indicating that the modularization helps to reduce the cross-talk noise for highly correlated patterns.
- Dealing with real structured patterns, the EANN system will be more robust than a single attractor, retrieving a larger number of such correlated patterns.
- The modularity specialization can be useful, in particular we found that the more sparse the code (lower activity) the larger the number of patterns retrieved by the module.
- The modularization could perform ensembled voting.

1. **Deal with offline 2D gesture patterns**, i.e. spatially/temporal structured and activity biased.
2. Importance of the neural coding.
3. Test **gesture invariants**.
4. **Specialize the network modules**.

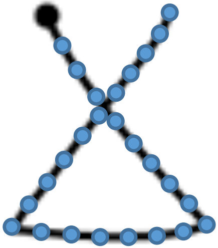


Gesture encoding (binarization) I

Codification



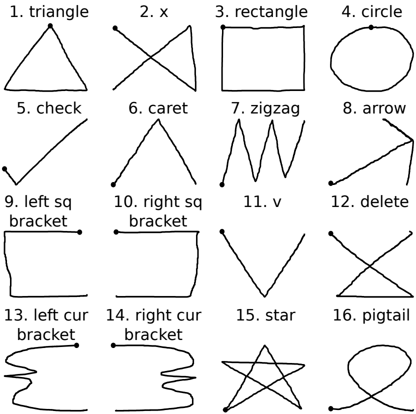
Gesture

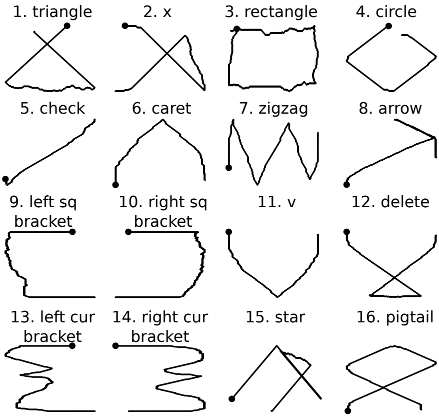


Binary chain

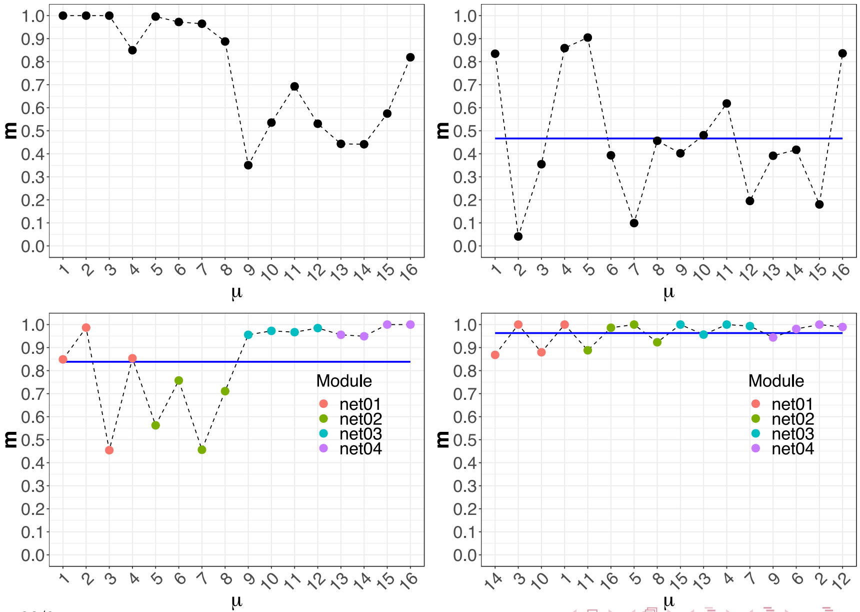
101010101010101010
0yr0yr0yr0yr0yr0yr0yr
1111111111111111111

Gesture encoding (binarization) II

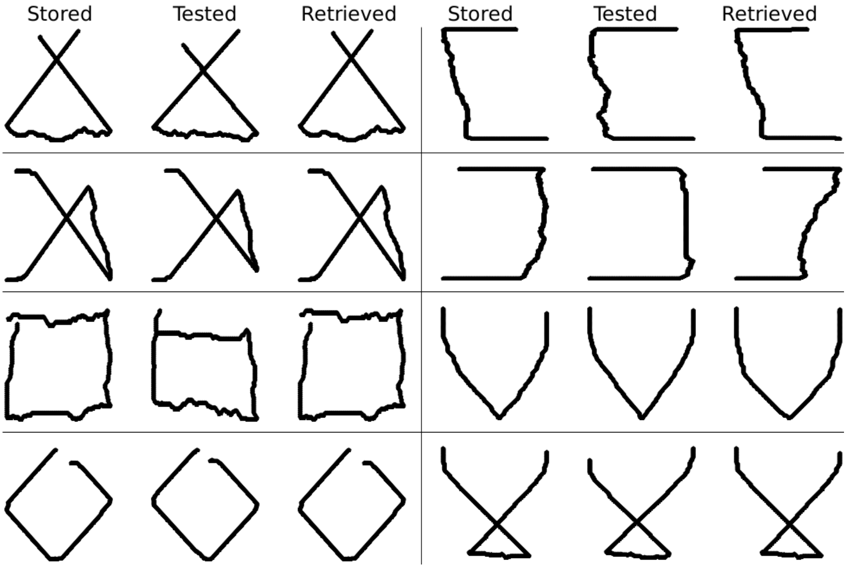




Gesture retrieval overlaps $n = \{1, 4\}, K = 200$

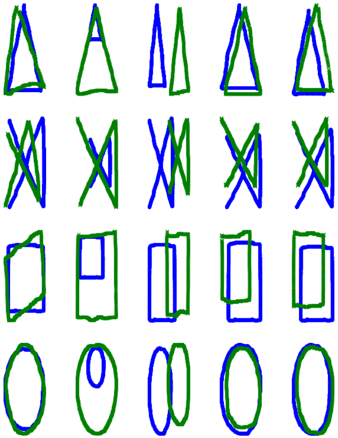


Examples of retrieved gestures ($min(m^0) = 0.6$)



EANN is robust to (some) gesture invariants

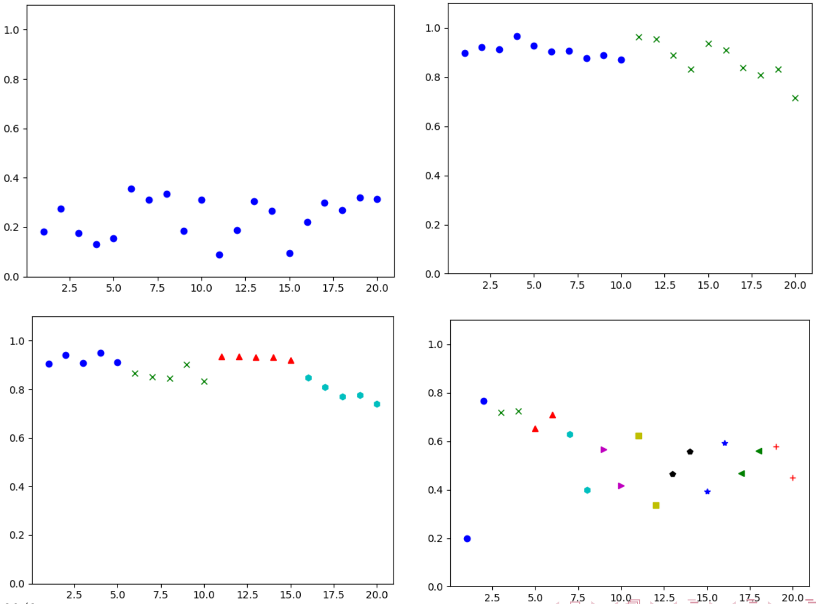
👎Orientation.
👍(Slight) Rotation.
👍Translation, 👍scale, 👍multi-stroke.



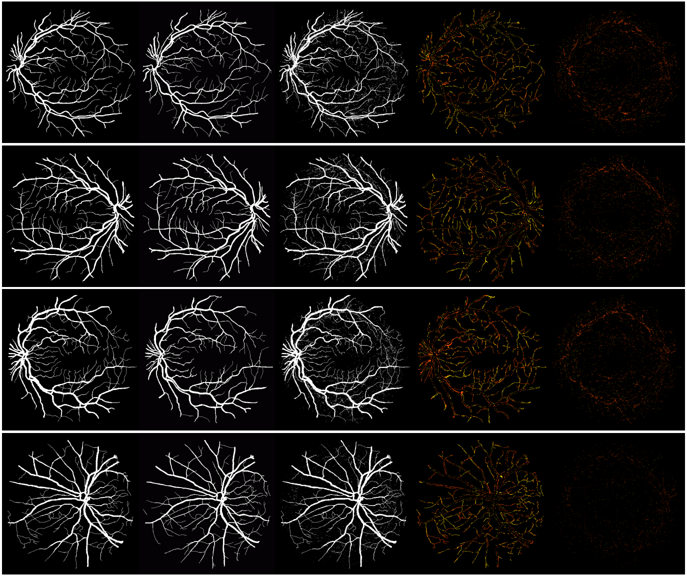
EANN for gesture retrieval

- The EANN is robust to the gesture initial noise, promising robustness for invariants.
- EANN model opens the possibility of a combinatorial optimization of the input of patterns to the ensemble modules.
- The modularity specialization can be useful, to deal with gesture invariants.
- Heuristic optimization can be used to select the input order.
- Include dynamics of the gesture.
- Test 3D gestures.

(DRIVE) Retrieval of retinal images ($n = \{1, 2, 4, 10\}$)



Examples of retrieved retinal images



Conclusions and future work

- We tested applications to real patterns: biometrics and 2D Gestures.
- The systems increase the storage capacity of the single network (higher than for random patterns).
- The importance of the pattern encoding was highlighted.
- The modules specialization can perform retrieval of patterns with different features, or perform ensembled voting.
- The modules can be non-uniform (connectivity, patterns subsets, etc.).
- The **optimization of the input** to modules assignment will be explored:
  - Heuristic optimization.
- Explore Behavioral biometrics:
  - Signature recognition.

Thank you!

The algorithms we take for granted have been explored to their limits before, so we can take advantage of them today.

Resources

- González, Mario, et al. "Increase attractor capacity using an ensembled neural network." Expert Systems with Applications 71 (2017): 206-215.
- González, Mario, et al. "Capacity and retrieval of a modular set of diluted attractor networks with respect to the global number of neurons." International Work-Conference on Artificial Neural Networks. Springer, Cham, 2017.
- Ensemble of Attractor Networks for 2D Gesture Retrieval. International Work-Conference on Artificial Neural Networks 2019.
- Fingerprint Retrieval using a Specialized Ensemble of Attractor Networks. International Work-Conference on Artificial Neural Networks 2019.