Modular Neural Networks

CSE 537 DATA MINING

Professor Anita Wasilewska

References

- http://www.sciencedirect.com/science/article/pii/S0925231213001604
- http://www.kovan.ceng.metu.edu.tr/~emre/literature/LOMNN/Modular_Neural_Networks_A_Survey.pdf
- http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.7.6594&rep=rep1&type=pdf
- http://www.teco.edu/~albrecht/neuro/report.pdf
- https://en.wikipedia.org/wiki/Modular_neural_network
- A Parallel and Modular Multi-Sieving Neural Network Architecture for Constructive Learning, by Lu et al.
- A modular neural network architecture with concept, by Yi Dinga, Qi Fenga, Tianjiang Wanga, Xian Fub
- https://www3.cs.stonybrook.edu/~cse634/

Overview

- History of Neural networks
- Motivation behind Neural networks
- Biological neural networks
- Modular neural networks
- Motivation behind Modular Neural Networks
- A paper study on a parallel and fault tolerant MNN
- Benefits

History of Neural Networks

- In 1943, neurophysiologist Warren McCulloch and mathematician Walter Pitts modelled a simple neural network using electrical circuits

- Nathanial Rochester from the IBM research laboratories simulated a hypothectical network and failed in his first attempt.

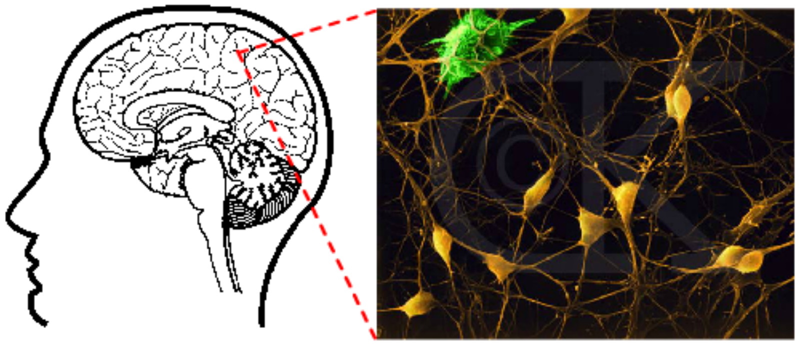
History of Neural Networks

- In 1959, Bernard Widrow and Marcian Hoff of Stanford developed models called "ADALINE" and "MADALINE.“

- ADALINE was developed to recognize binary patterns so that if it was reading streaming bits from a phone line, it could predict the next bit

- MADALINE was the first neural network applied to a real world problem, using an adaptive filter that eliminates echoes on phone lines.
- While the system is **as ancient as** air traffic control systems, like air traffic control systems, it is **still in commercial use** !!

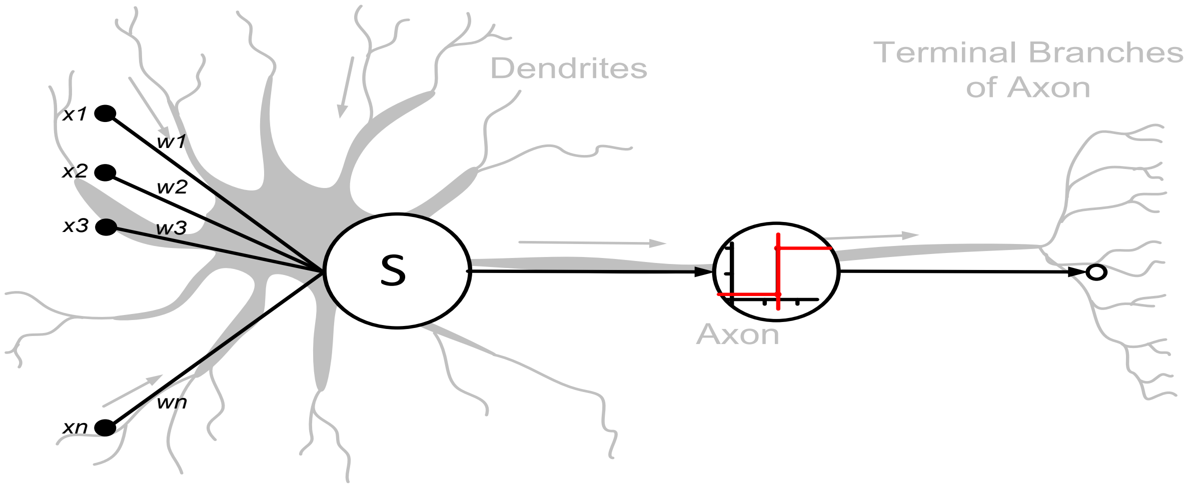
Biological Neural Network



- Each of the yellow blobs in the picture above are neuronal cell bodies: **Soma**
- The lines are the input and output channels: **dendrites** and **axons** which connect them
- Each **neuron** receives **electrochemical inputs** from other neurons at the dendrites

Biological Neural Network

- If the **sum** of these electrical inputs is sufficiently powerful to **activa**te the neuron, it **fires** or transmits these signals to other neurons
- It is important to note that a **neuron fires only** if the total signal received at the cell body **exceeds** a certain level
- Each neuron performs a weighted sum of its inputs, and **fires** a binary signal if the **total input** exceeds a certain level
- **This is the model on which artificial neural networks are based**

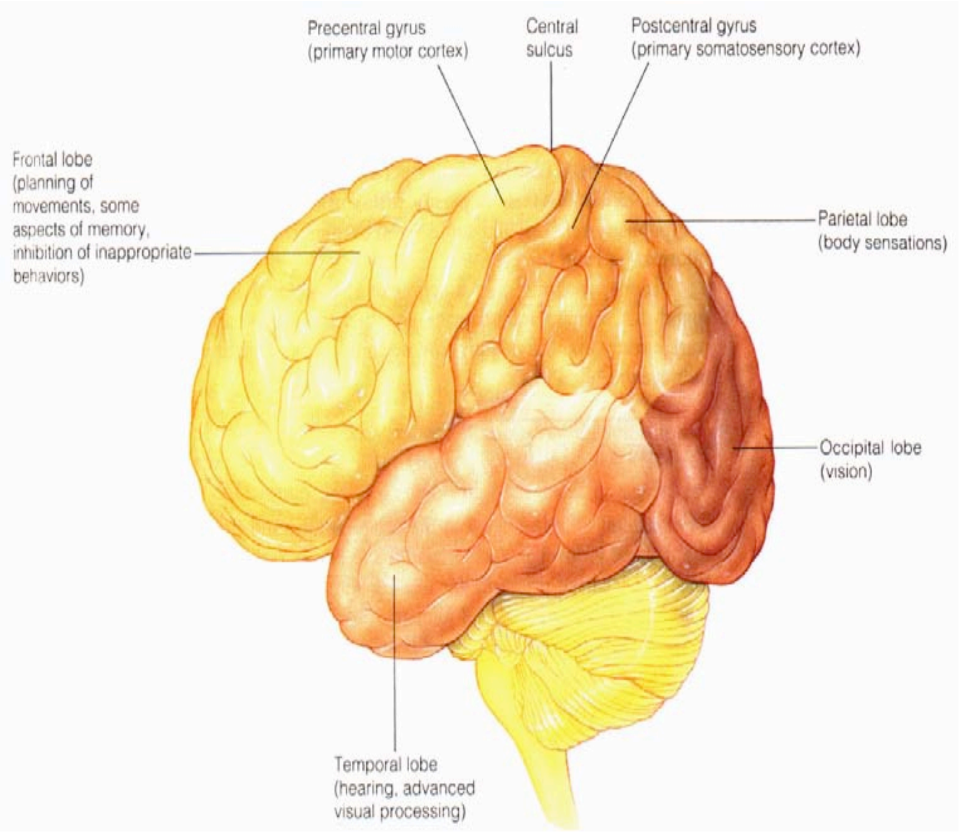


Modularity

- Biological modularity is the first idea which **motivated** many Modular NN designs

- Brain is **modular** on different spatial scales
- On the smallest scale, synapses are clustered on dendrites

- On the largest scale, the brain is **composed** of several anatomically and functionally distinct areas



Modularity in Neural Networks

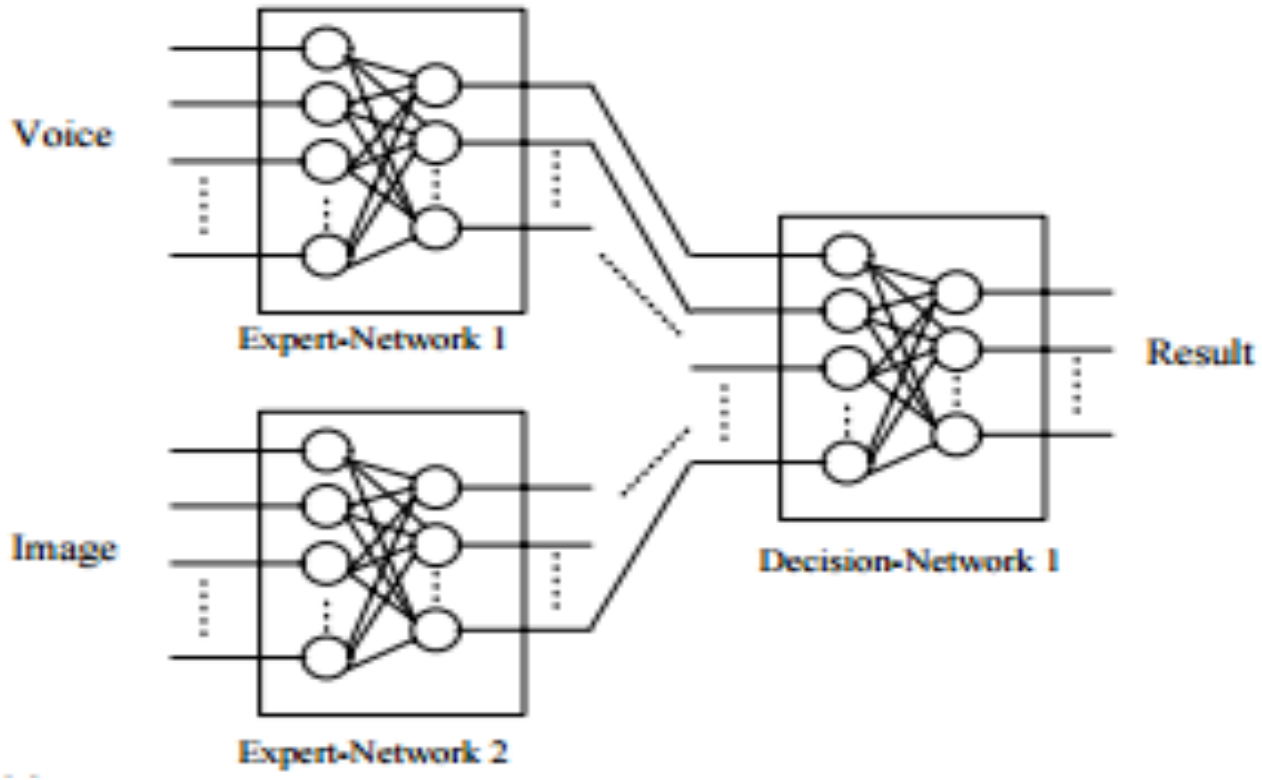
- The most used artificial neural networks have a **monolithic** structure and perform well on a small input space
- The **complexity** increases and the performance decreases rapidly with a growing input **dimension**

- **Different** models of NN combined into a single system form

modular neural networks

- Each **single** network is made into **a module** that can be freely **intermixed** with modules of other types in that system.

Modular Neural Networks



Modular Neural Networks

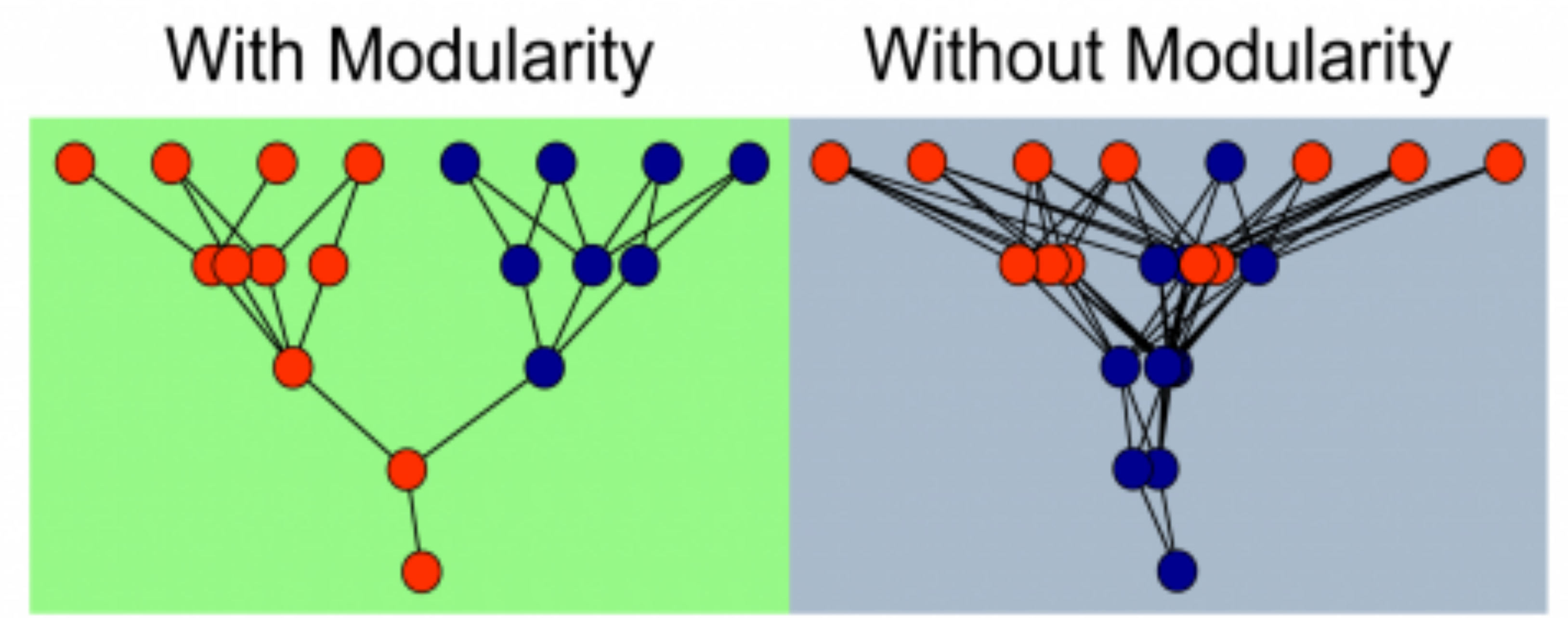
- **Modular** (multiple) NN are used for strongly separated architecture
- **Each** of the networks works independently on its own domain
- The single networks are **built** and **trained** for its domain and for their specific task
- The **final decision** is made on the **results** of the individual networks

Modular Neural Networks

- The **decision system** can by **implemented** by a logical
  majority vote function,
  another neural network or
  a rule based expert system may be employed

- The **individual network** is trained on its domain only
- The **output** by a single network is according to its **specific input**
-

Modular Neural Networks



One of the networks is trained to **identify a person by voice** while the other network is trained to **identify the person by vision**

Modular Neural Networks

- Modular neural network architecture builds a **bigger network** by using modules as **building blocks**

- The **architecture** of a single module is simpler and the sub-networks are **smaller** than a monolithic network.

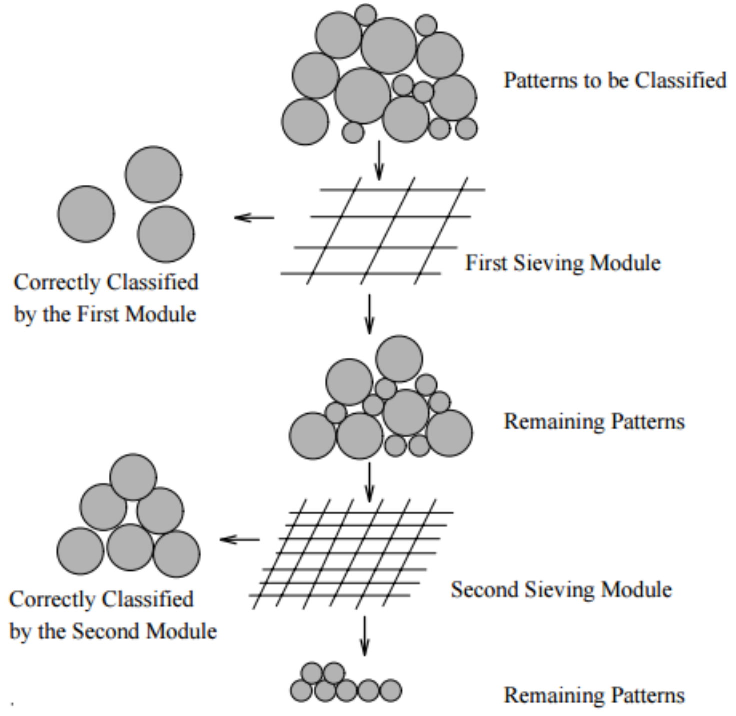
- Due to the structural modifications the **task** the module has to learn is in general easier than the **whole task** of the network

Modular Neural Networks

- The modules are independent to a certain level which **allows** the system to work in parallel

- For this modular approach it is always necessary to have a **control system** to enable the modules to work together in a useful way

Modular Multi-Sieving Neural Network Architecture



Modular Multi-Sieving Neural Network Architecture

- The **patterns** are classified by this **algorithm** on different levels
-
- In the **first level** , a very **rough sieve**, some patterns may be
- recognized correctly while others will not
- The correctly classified samples are **taken out** of the training set

- The **next level** , a **less rough sieve**, is only trained on the
- remaining ones
- **After** the training of this level the correctly recognized patterns
- are **removed** from the training set.

Modular Multi-Sieving Neural Network Architecture

- The remaining **patterns** form the training set for the **next** level
- This process is **repeated** until all patterns are classified correctly
-
- **Each level** of learning , **each sieve generates** a neural network
- with the **ability** to recognize a subset of the original training set
-
- These networks called **sieving modules** face a simpler **recognition** task than the whole problem

Motivations for Modular Neural Network

- **Biological motivations :**
    - **Modularity, Functional specialization concept, Fault tolerance, scalability**
- **Psychological motivations**
    - **Learning in stages,**
    - **Decomposing tasks**
        - A way to cope NP-completeness, some papers used MNN to solve travelling salesman problem
- **Hardware motivations**
    - Hardware reaching theoretical limits, need for speed and less memory structures
- **Computational motivations**
    - improves the speed of learning by reducing the effect of conflicting training information (or crosstalk). Crosstalk degrades the ability of the network to perform correctly for a group of patterns.

MNN Design stages

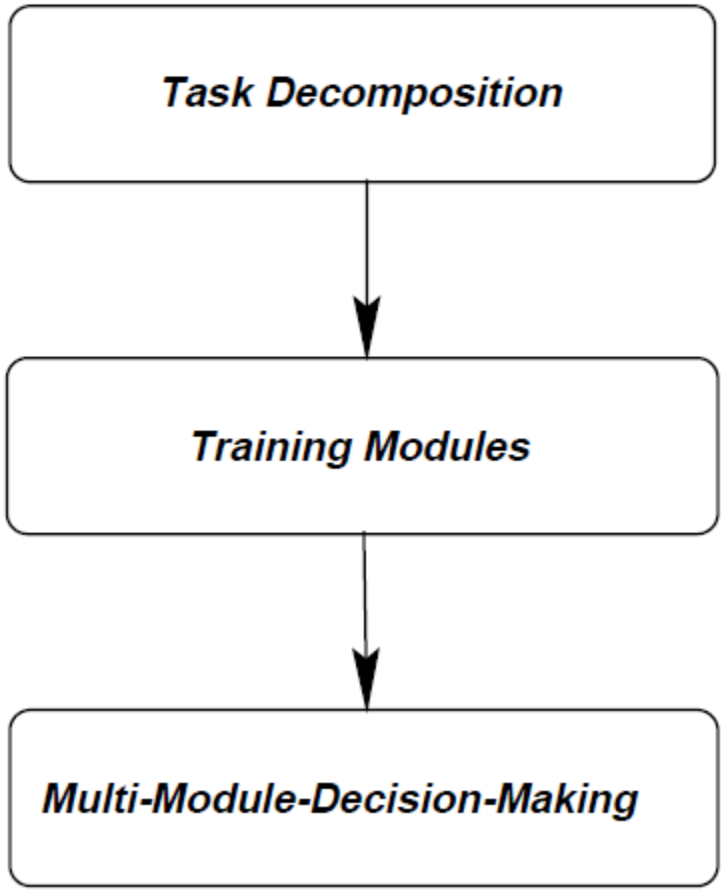
- **Efficiency**

  -multimodule **decision-making** strategy has to take part in order to integrate the different local decisions (at the modular level) into a global one

- **Reasonable balance**
- between sub-tasks simplification and decision-making efficiency

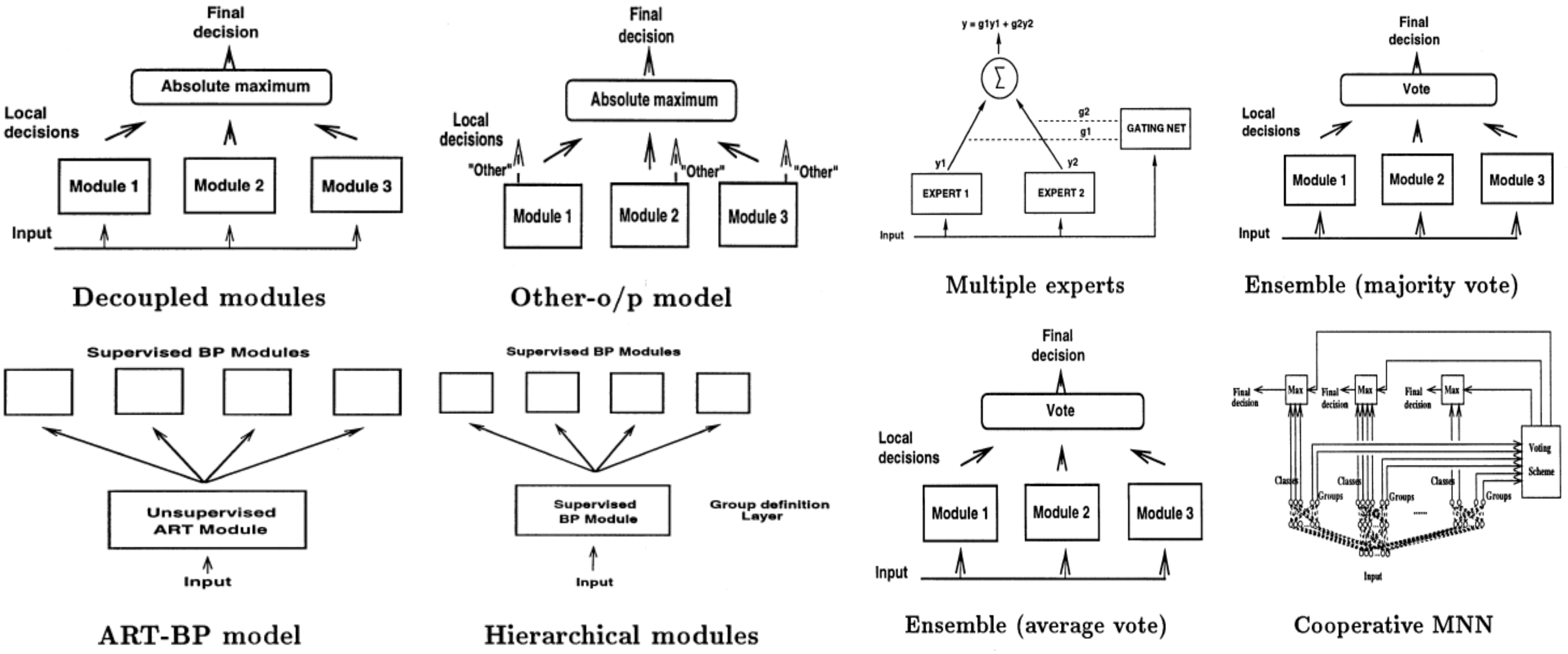
  - sub-tasks as simple as possible

  - give the multimodule decision-making strategy enough information



The general three main stages for designing MNN.

MNN Architectures



Eight of the most common MNN architectures.

Research Paper

A modular neural network architecture with concept

Yi Ding, Qi Feng, Tianjiang Wang, Xian Fu

Journal of Neurocomputing

Volume 125, February, 2014
Pages 3-6

Elsevier Science Publishers B. V. Amsterdam, The Netherlands

- Aim: Develop a parallel and fault tolerant MNN.
- Motivation: Human Nervous system.

Evolution of the Model

- Is the network **homogenous** or **heterogeneous:**
  Are all modules of **one type** or are there **different types** of modules used
- What **type of architecture** is used for the modules:

  MLP, LNN, SOM, ART
- How are the modules **interconnected:**

  Are only feedforward connections used, are recurrent connections allowed, are connections only made from one layer to next
- What is the **general network structure:**

  has the network a single layer or n layers

Evolution of the Model

- How are the **inputs connected** to the network:

  Are the inputs connected to all modules overlapping or

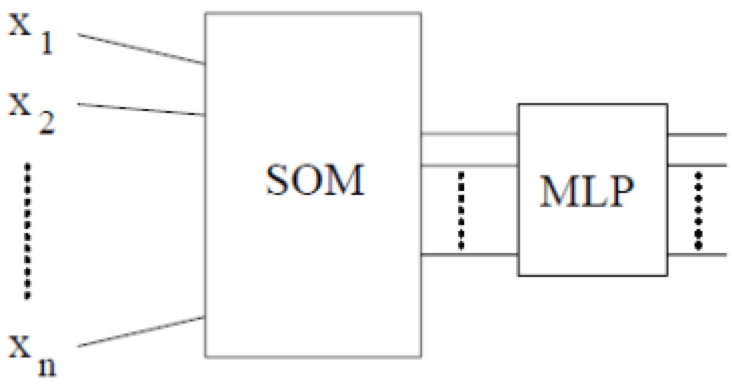
  only to one module non-overlapping

- What **training algorithm** is used for the modules

  Is a supervised or unsupervised learning method used

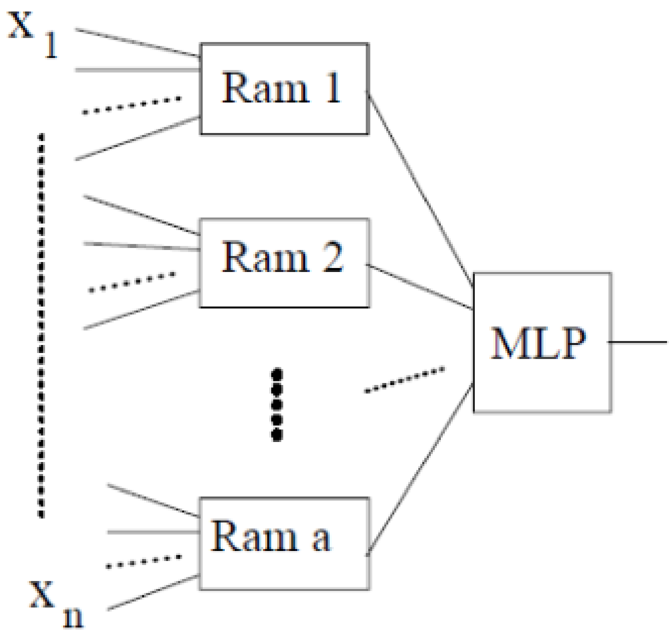
- How is **learning organized** for the whole network:
- Does the **network learn** in stages?
- Is noise used during the training?

Evolution of the Model



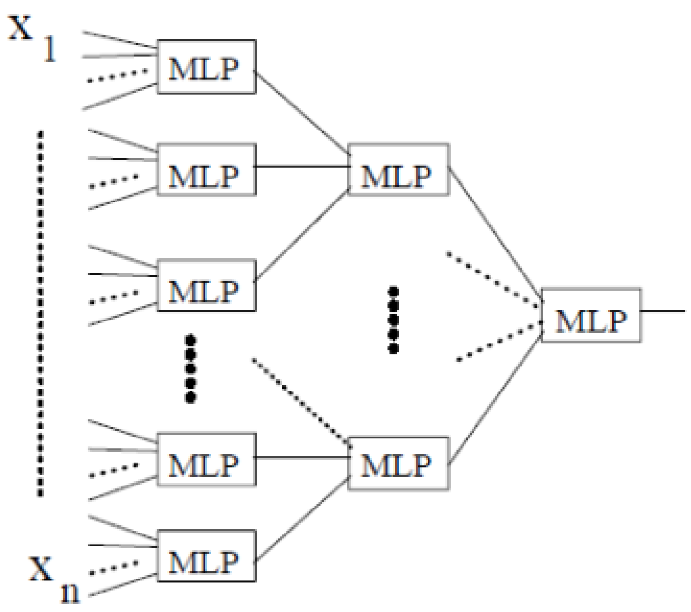
- A **heterogeneous architecture** consisting of two modules
- One is a self-organizing map SOM and is used to reduce the input dimension the other one is a multilayer Perceptron that works on the reduced input space

Evolution of the Model



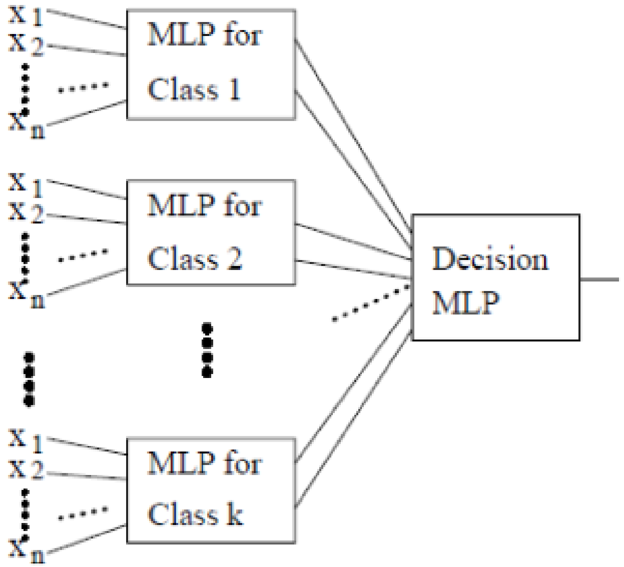
- MLP on the output of logical adaptive nodes is one idea for
- another heterogeneous structure

Evolution of the Model



- A homogeneous pyramidal structure
- uses small MLPs instead of logical nodes
- This should provide fast training of the network as well as improved generalization ability

Evolution of the Model



- The input and output dimensions of the problem remain
- the same
- Each subnetwork is only **concerned** with a single class

The Final Architecture:

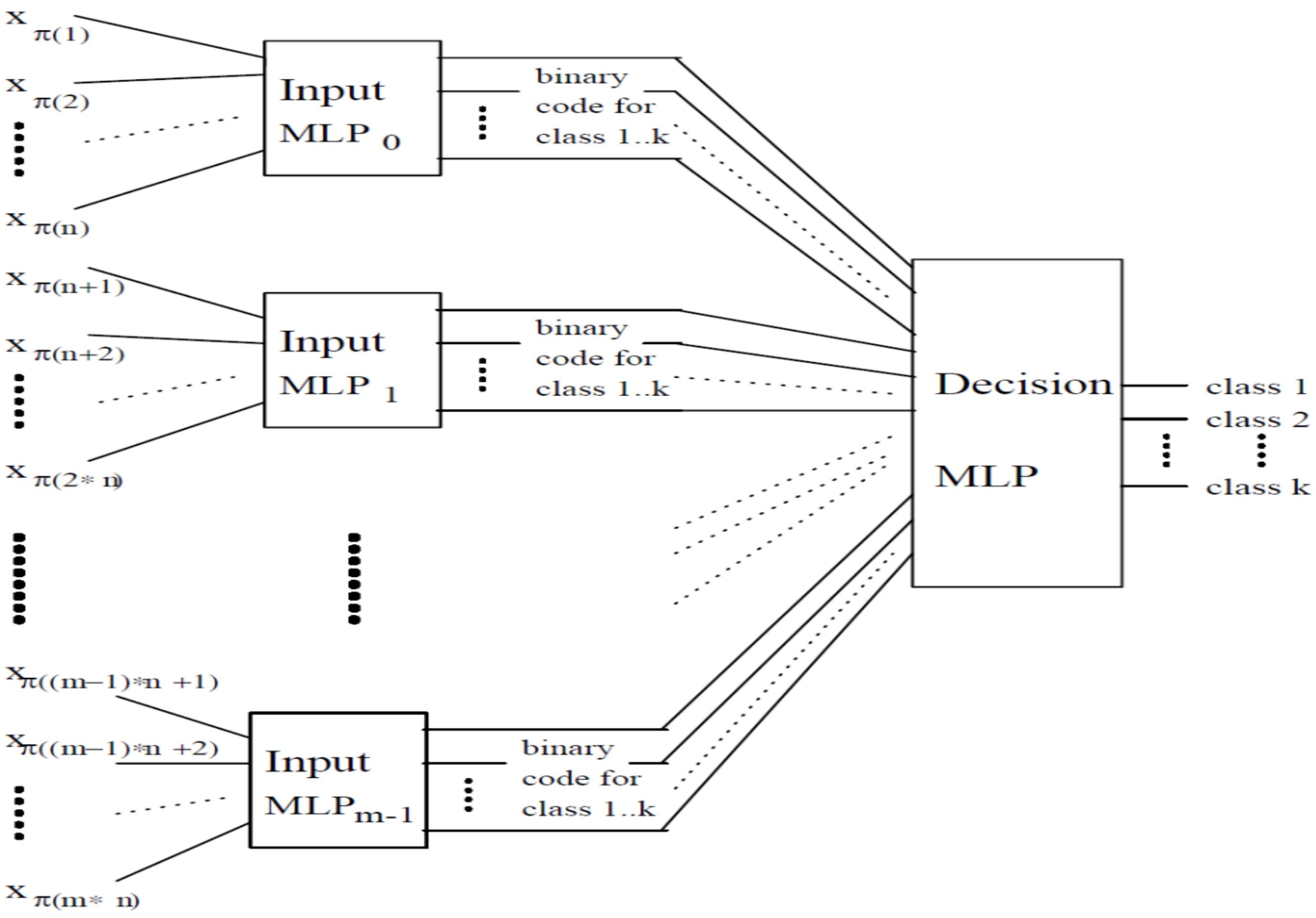


Figure 5.2: The Proposed Modular Neural Network Architecture.

Characteristics

- All sub-networks are MLPs
- the number of inputs and the number of outputs of the module is determined by the system
- The internal structure such as the number of hidden layers and the number of neurons in each hidden layer can be chosen independent of the overall architecture

Definition: A Module

A module is a multilayer feedforward neural network defined by a 3-tuple:

$$\mathcal{M} = (a, b, \mathcal{H})$$

Where a is the number of inputs of the module, b is the number of output nodes, and $\mathcal{H}$ is a list containing the numbers of neurons in each of the hidden layers.

Example:

A multilayer Perceptron module $\mathcal{M}$ with eight inputs, twelve neurons in the first hidden layers, ten neurons in the second hidden layer, and four outputs is described as: $\mathcal{M} = (8, 4, [12, 10])$.

Definition: A Modular Neural Network

A modular neural network is a set of interconnected modules defined by a 7-tuple.

$$N = (l, k, m, r, \pi, \mathcal{I}, \mathcal{D})$$

Where l is the number of inputs, k the number of classes, m the number of modules in the input layer, r is the type of the intermediate representation ($r \in \{small, large\}$), π is the permutation function, $\mathcal{I}$ is the input layer module, and $\mathcal{D}$ is the decision module.

Algorithm for MNN

The Training Algorithm:

- Stage 1 – Training the Input Layer:

  1. Select the training sets TS_i from the original training set TS, for all $i = 0 \ldots m-1$.
  2. Train all modules MLP$_i$ on TS_i using the BP algorithm.

- Stage 2 – Training the Decision Network

  1. Calculate the response r of the first layer for each input vector j.
     $r^j = \Phi((x_1^j, x_2^j, \ldots, x_l^j))$.
  2. Build the training set for the decision network
     $TS_d = \{(r^j; d_{BIT}^j) | j = 1, \ldots, t\}$
  3. Train the decision network on the set TS_d using the BP algorithm

Generalization

- The proposed architecture combines two methods of **generalization**:

- **One method** is built in to the MLP
- Each of the networks has the ability to generalize on its input space
- This type of **generalization** is common to connectionist systems.

- The **second method** of generalization is due to the architecture of the proposed network

  It is a way of **generalizing** according to the similarity of **input patterns**

Training

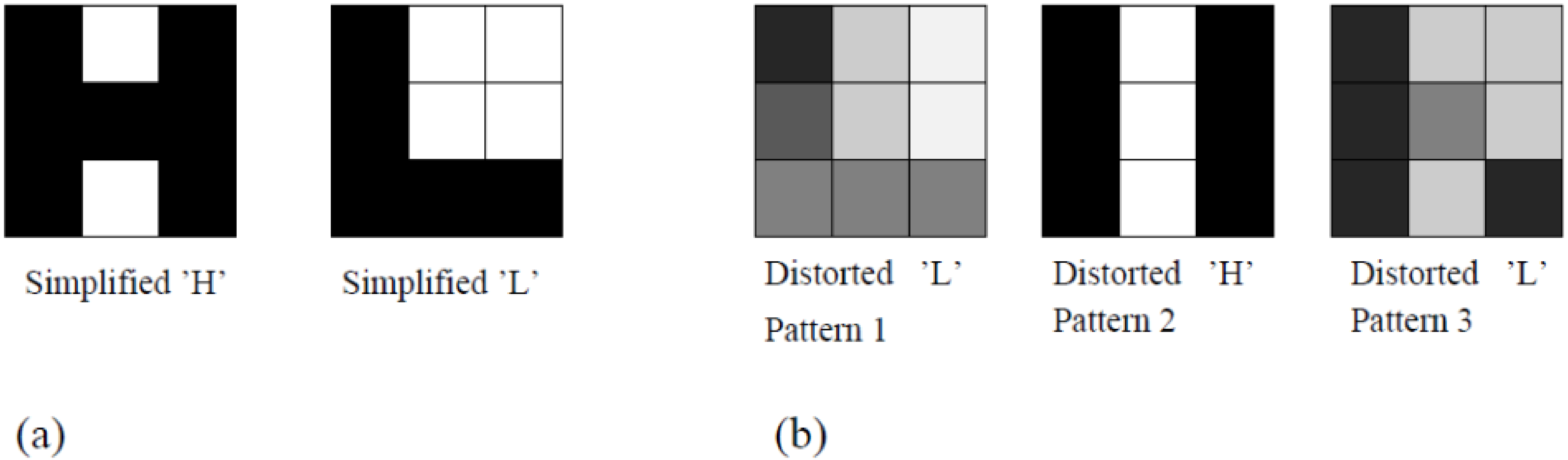
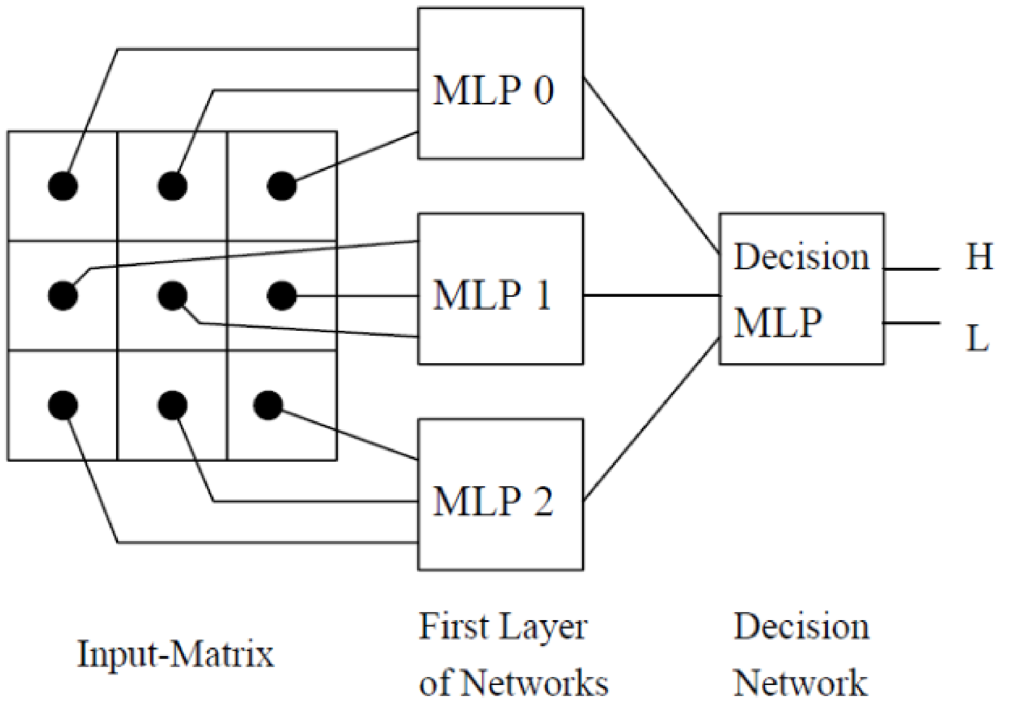


Figure 5.5: (a) The Training Set.
(b) The Test Set.

Example Architecture



Training

- Input :
- 9 values between 0 and 1 (black = 1, white = 0)

- The network needs to be trained to recognize the simplified letters ‘H’ and ‘L’

- Output: ‘H’ = 0 ; ‘L’ = 1

Training Subsets

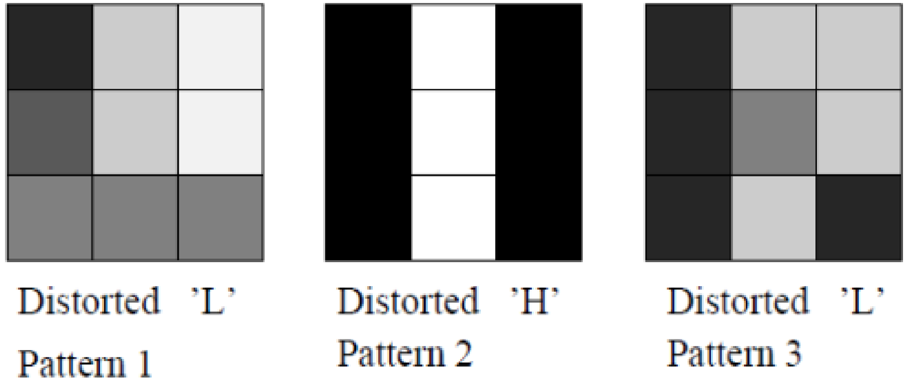
Simplified 'H' Simplified 'L'

MLP_0	MLP_1	MLP_2
(1,0,1;0)	(1,1,1;0)	(1,0,1;0)
(1,0,0;1)	(1,0,0;1)	(1,1,1;1)

Training Set for decision Network

$$(\Phi(1,0,1,1,1,1,1,0,1);1,0) = (0,0,0;1,0)$$
$$(\Phi(1,0,0,1,0,0,1,1,1);0,1) = (1,1,1;0,1)$$

Use Test Set



- The 1st character tests generalization within the input modules
- Second tests the generalization on the number of correct sub-patterns
- Third character is a combination of both

Global Decision Making (Classification)

$$r_1 = \Psi(\Phi(0.9, 0.2, 0.1, 0.7, 0.2, 0.1, 0.5, 0.5, 0.5))$$
$$= \Psi(0.95, 0.86, 0.70) = (0.04, 0.96) \quad \Rightarrow \quad 'L'$$

$$r_2 = \Psi(\Phi(1.0, 0.0, 1.0, 1.0, 0.0, 1.0, 1.0, 0.0, 1.0))$$
$$= \Psi(0, 0.49, 0) = (0.91, 0.09) \quad \Rightarrow \quad 'H'$$

$$r_3 = \Psi(\Phi(0.9, 0.2, 0.2, 0.9, 0.5, 0.2, 0.9, 0.2, 0.9))$$
$$= \Psi(0.92, 0.65, 0.09) = (0.15, 0.89) \quad \Rightarrow \quad 'L'$$

Benefits of Modular Neural Network

- **Efficiency**

The possible connections increases at a daunting rate as **nodes** are added to the network

Since computation time depends on the **number** of nodes and their **connection**s, any increase here will have drastic consequences in the processing time

Speedup of 10 times achieved

Benefits of Modular Neural Network

- **Training**
- A large neural network attempting to **model** multiple parameters can suffer from interference as new data can dramatically alter existing connections or just serve to confuse.

- With some foresight into the **subtasks** to be solved, each neural network can be **tailored** for its task

- Provide **unique** training algorithm and training data for each sub-network

Implemented much more quickly

Benefits of Modular Neural Network

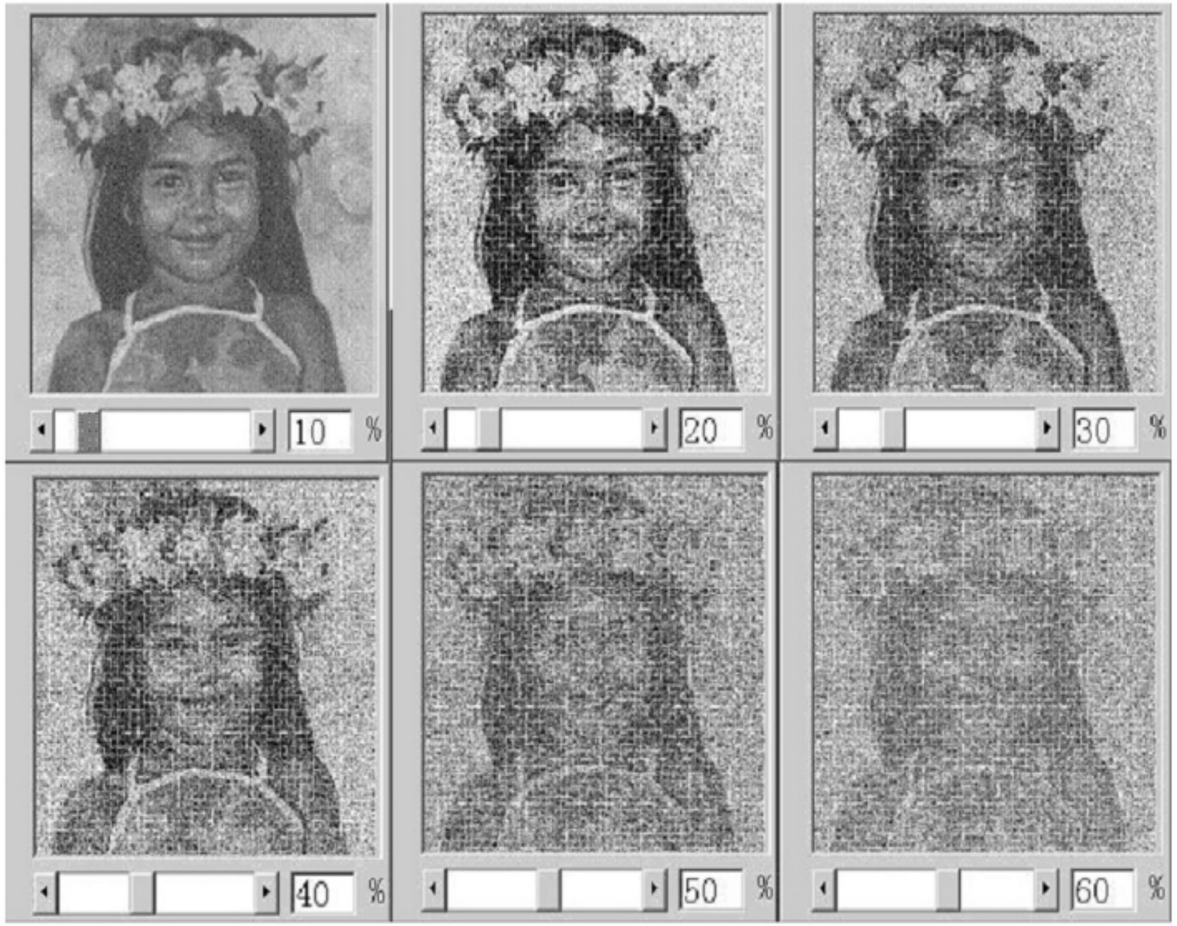
- **Robustness**
- Regardless of whether a large neural network is **biological** or **artificial**, it remains largely susceptible to interference at and failure in any one of its nodes

- By compartmentalizing **subtasks**, failure and interference are much more readily diagnosed and their **effects** on other sub-networks are **eliminated** as each one is independent of the other.

Benefits of Modular Neural Network

Reference :

A modular neural network architecture with concept, by Yi Dinga, Qi Fenga, Tianjiang Wanga, Xian Fu



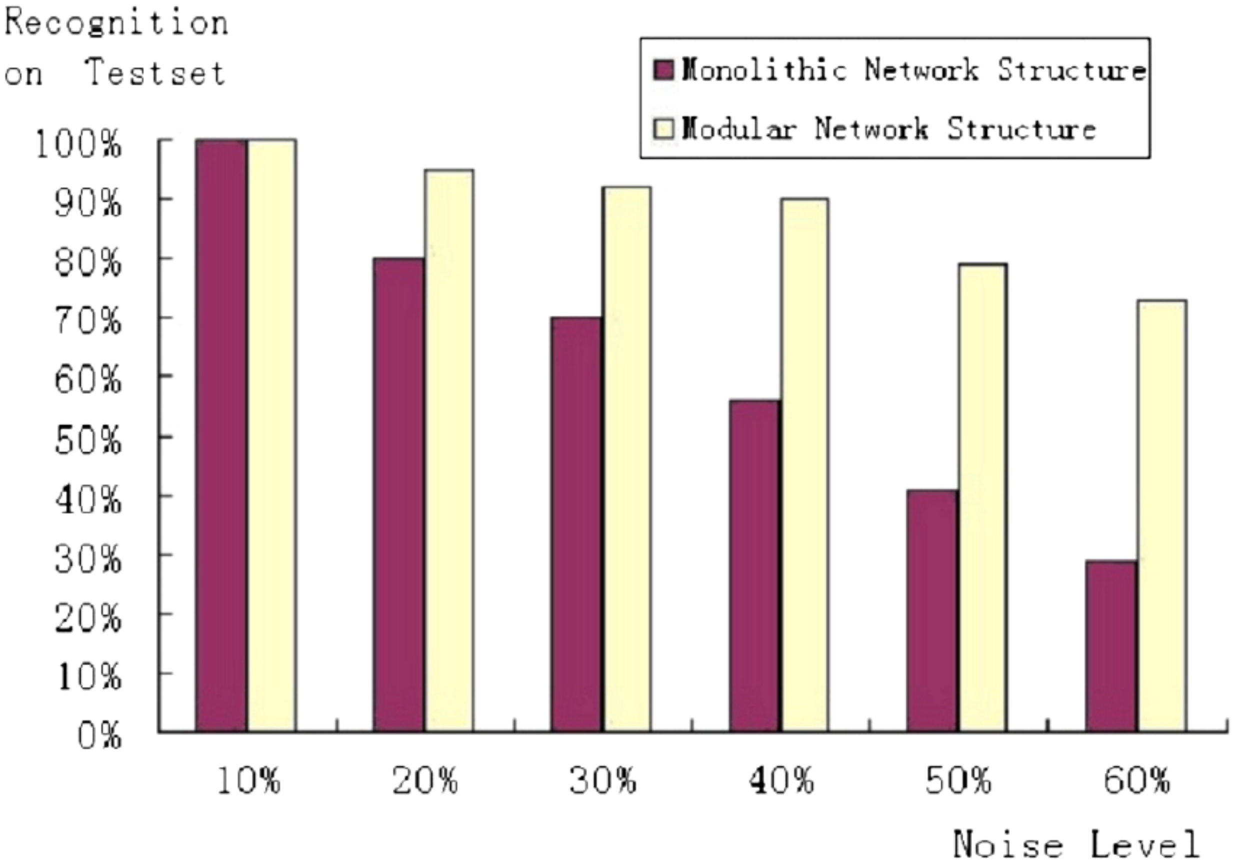


Fig. 3. The performance on noisy inputs.

Summary

- Modular Neural Networks, its connection to NN
- Motivations
- Benefits of MNN
- MNN Architecture
- Example of developing a MNN