

# Behaviour-Based Clustering of Neural Networks

**María José Castro-Bleda**

*Universidad Politécnica de Valencia, Spain*

**Slavador España-Boquera**

*Universidad Politécnica de Valencia, Spain*

**Francisco Zamora-Martínez**

*Universidad Politécnica de Valencia, Spain*

## INTRODUCTION

The field of off-line optical character recognition (OCR) has been a topic of intensive research for many years (Bozinovic, 1989; Bunke, 2003; Plamondon, 2000; Toselli, 2004). One of the first steps in the classical architecture of a text recognizer is preprocessing, where noise reduction and normalization take place. Many systems do not require a binarization step, so the images are maintained in gray-level quality. Document enhancement not only influences the overall performance of OCR systems, but it can also significantly improve document readability for human readers. In many cases, the noise of document images is heterogeneous, and a technique fitted for one type of noise may not be valid for the overall set of documents. One possible solution to this problem is to use several filters or techniques and to provide a classifier to select the appropriate one.

Neural networks have been used for document enhancement (see (Egmont-Petersen, 2002) for a review of image processing with neural networks). One advantage of neural network filters for image enhancement and denoising is that a different neural filter can be automatically trained for each type of noise.

This work proposes the clustering of neural network filters to avoid having to label training data and to reduce the number of filters needed by the enhancement system. An agglomerative hierarchical clustering algorithm of supervised classifiers is proposed to do this. The technique has been applied to filter out the background noise from an office (coffee stains and footprints on documents, folded sheets with degraded printed text, etc.).

## BACKGROUND

Multilayer Perceptrons (MLPs) have been used in previous works for image restoration: the input to the MLP is the pixels in a moving window, and the output is the restored value of the current pixel (Egmont-Petersen, 2000; Hidalgo, 2005; Stubberud, 1995; Suzuki, 2003). We have also used neural network filters to estimate the gray level of one pixel at a time (Hidalgo, 2005): the input to the MLP consisted of a square of pixels that was centered at the pixel to be cleaned, and there were four output units to gain resolution (see Figure 1). Given a set of noisy images and their corresponding clean counterparts, a neural network was trained. With the trained network, the entire image was cleaned by scanning all the pixels with the MLP. The MLP, therefore, functions like a nonlinear convolution kernel. The universal approximation property of a MLP guarantees the capability of the neural network to approximate any continuous mapping (Bishop, 1996).

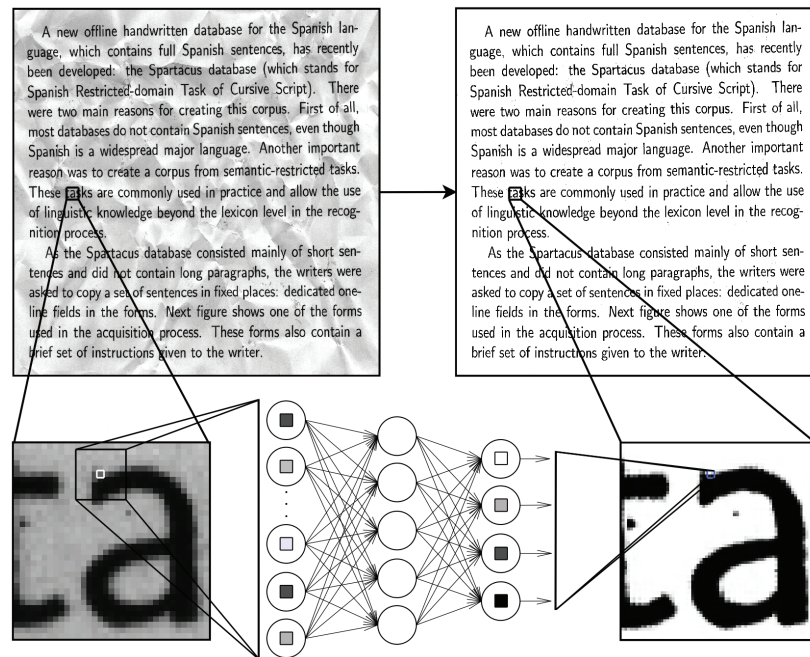
This approach clearly outperforms other classic spatial filters for reducing or eliminating noise from images (the mean filter, the median filter, and the closing/opening filter (Gonzalez, 1993)) when applied to enhance and clean a homogeneous background noise (Hidalgo, 2005).

## BEHAVIOUR-BASED CLUSTERING OF NEURAL NETWORKS

### Agglomerative Hierarchical Clustering

Agglomerative hierarchical clustering is considered to be a more convenient approach than other clustering

Figure 1. An example of document enhancement with an artificial neural network. A cleaned image (right) is obtained by scanning the entire noisy image (left) with the neural network.



algorithms, mainly because it makes very few assumptions about the data (Jain, 1999; Mollineda, 2000). Instead of looking for a single partition (based on finding a local minimum), this clustering algorithm constructs a hierarchical structure by iteratively merging clusters according to certain dissimilarity measure, starting from singletons until no further merging is possible (one general cluster). The hierarchical clustering process can be illustrated with a tree that is called dendrogram, which shows how the samples are merged and the degree of dissimilarity of each union (see Figure 2). The dendrogram can be easily broken at a given level to obtain clusters of the desired cardinality or with a specific dissimilarity measure. A general hierarchical clustering algorithm can be informally described as follows:

1. Initialization:  $M$  singletons as  $M$  clusters.
2. Compute the dissimilarity distances between every pair of clusters.
3. Iterative process:

- a) Determine the closest pair of clusters  $i$  and  $j$ .
  - b) Merge the two closest clusters into a new cluster  $i+j$ .
  - c) Update the dissimilarity distances from the new cluster  $i+j$  to all the other clusters.
  - d) If more than one cluster remains, go to step a).
4. Select the number  $N$  of clusters for a given criterion.

## Behaviour-Based Clustering of Supervised Classifiers

When the points of the set to be clustered are supervised classifiers, both a dissimilarity distance and the way to merge two classifiers must be defined (see Figure 2):

1. The dissimilarity distance between two clusters can be based on the behaviour of the classifiers with respect to a validation dataset. The more similar the output of two classifiers is, the closer they are.
2. To merge the closest pair of clusters, a new classifier is trained with the associated training data

3 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: [www.igi-global.com/chapter/behaviour-based-clustering-neural-networks/10253](http://www.igi-global.com/chapter/behaviour-based-clustering-neural-networks/10253)

## Related Content

---

### Detecting Synchronization Signal Jamming Attacks for Cybersecurity in Cyber-Physical Energy Grid Systems

Danda B. Rawat and Brycent A. Chatfield (2017). *Security Solutions and Applied Cryptography in Smart Grid Communications* (pp. 68-78).

[www.irma-international.org/chapter/detecting-synchronization-signal-jamming-attacks-for-cybersecurity-in-cyber-physical-energy-grid-systems/172671](http://www.irma-international.org/chapter/detecting-synchronization-signal-jamming-attacks-for-cybersecurity-in-cyber-physical-energy-grid-systems/172671)

### Cognitive Parameter Based Agent Selection and Negotiation Process for B2C E-Commerce

Bireshwar Dass Mazumdar and R. B. Mishra (2011). *Intelligent, Adaptive and Reasoning Technologies: New Developments and Applications* (pp. 181-203).

[www.irma-international.org/chapter/cognitive-parameter-based-agent-selection/54431](http://www.irma-international.org/chapter/cognitive-parameter-based-agent-selection/54431)

### IoT Analytics and ERP Interoperability in Automotive SCM: ANN-Fuzzy Logic Technique for Designing Decision Support Systems

Paul Jayender and Goutam Kumar Kundu (2022). *International Journal of Fuzzy System Applications* (pp. 1-19).

[www.irma-international.org/article/iot-analytics-and-erp-interoperability-in-automotive-scm/306282](http://www.irma-international.org/article/iot-analytics-and-erp-interoperability-in-automotive-scm/306282)

### A Study on Machine Learning-Based Water Quality Assessment and Wastewater Treatment

Satakshi Singh, Suryanshi Mishra, Tinku Singha and Shobha Thakur (2023). *Artificial Intelligence Applications in Water Treatment and Water Resource Management* (pp. 146-162).

[www.irma-international.org/chapter/a-study-on-machine-learning-based-water-quality-assessment-and-wastewater-treatment/329350](http://www.irma-international.org/chapter/a-study-on-machine-learning-based-water-quality-assessment-and-wastewater-treatment/329350)

### Representing Non-Rigid Objects with Neural Networks

José García-Rodríguez, Francisco Flórez-Revuelta and Juan Manuel García-Chamizo (2009). *Encyclopedia of Artificial Intelligence* (pp. 1363-1369).

[www.irma-international.org/chapter/representing-non-rigid-objects-neural/10417](http://www.irma-international.org/chapter/representing-non-rigid-objects-neural/10417)