

# Generación de Características y Reconocimiento Estadístico de Patrones

Luciano Lorenti, Lucía Violini, Javier Giacomantone, María José Abásolo,  
Oscar Bria, Marcelo Naiouf, Armando De Giusti

**Instituto de Investigación en Informática LIDI (III-LIDI)**  
**Facultad de Informática – UNLP**

{llorenti, lviolini, jog, mjabasolo, obria, mnaiouf, degiusti}@lidi.info.unlp.edu.ar

## CONTEXTO

Esta línea de investigación y desarrollo (I/D) forma parte del proyecto “Cómputo Paralelo de Altas Prestaciones (HPC). Fundamentos y Evaluación de Rendimiento en HPC. Aplicaciones a Sistemas Inteligentes, Simulación y Tratamiento de Imágenes – 11/F017”, del Instituto de Investigación en Informática LIDI acreditado por la UNLP en el marco del Programa de Incentivos.

## RESUMEN

Este trabajo describe una línea de I/D y los resultados esperados de la misma. El objetivo principal es analizar, desarrollar y evaluar modelos y métodos computacionales. A partir de los resultados obtenidos y los métodos propuestos el segundo objetivo es la transferencia de los mismos y el estudio de las técnicas de enseñanza-aprendizaje más adecuadas para los temas abordados.

Los dos ejes principales de investigación son, generación de características y reconocimiento estadístico de patrones. Se analizan métodos de generación de características a partir de señales en general y de imágenes digitales en particular, considerando el método de aprendizaje supervisado subyacente y su poder de discriminación. El segundo eje se centra en estudiar, desarrollar y evaluar métodos computacionales de reconocimiento estadístico de patrones, en particular métodos de clasificación supervi-

sada, no supervisada y reducción de dimensión. Todos los sistemas de reconocimiento de patrones diseñados o analizados responden a un modelo integral, en los cuales el énfasis está dado en el análisis científico de cada etapa.

**Palabras Clave:** Reconocimiento de Patrones. Análisis de Imágenes. Aprendizaje Automático. Generación de Características.

## 1. INTRODUCCION

En general todo tipo de fenómeno físico o químico del cual podamos directa o indirectamente extraer descriptores viables, es potencialmente tratable por un sistema automático de reconocimiento de patrones. Algunos de ellos tienen protocolos precisos para su implementación producto de años de investigación básica. Es necesario por lo tanto diferenciar claramente las hipótesis planteadas para proponer modelos o métodos. Es de particular interés en esta línea de I/D el estudio de la adecuada generación de características a partir de imágenes digitales. El trabajo actual lo podemos clasificar en cuatro tópicos principales bien diferenciados correspondientes a sub-disciplinas dentro del área. La primera sub-disciplina es clasificación supervisada donde el énfasis de nuestro trabajo se centra en el estudio de métodos de clasificación basados en núcleos dispersos, en particular máquinas de soporte vectorial. La segunda es clasificación no supervisada donde la principal línea

de trabajo son las técnicas de agrupamiento. Actualmente con énfasis en detección de valores atípicos, métodos basados en teoría espectral de grafos y su aplicación a imágenes de rango. El tercer tópico de fundamental importancia es el de reducción de dimensión en particular selección de características. El cuarto tópico que involucra a los tres anteriores puede ser denominado como de transferencia y técnicas de enseñanza-aprendizaje específicas para los temas de esta línea de investigación.

## 2. LINEAS DE INVESTIGACION y DESARROLLO

- Selección y extracción de características.
- Máquinas de soporte vectorial. Kernels y algoritmos de optimización.
- Clasificación no supervisada. Técnicas de agrupamiento (clustering).
- Detección de anomalías.
- Modelos y métodos computacionales en Análisis de Imágenes y Reconocimiento de Patrones.
- Métodos de estimación de parámetros para clasificadores Bayesianos.
- Métricas y pseudométricas.
- Criterios de evaluación de desempeño en sistemas de clasificación automática.
- Criterios y algoritmos para combinación de clasificadores.
- Clasificación de series temporales y clasificación contextual.
- Análisis de Señales Digitales.

## 3. RESULTADOS OBTENIDOS /ESPERADOS

- Desarrollar modelos y optimizar algoritmos particulares de clasificación supervisada y no supervisada.
- Evaluación de los métodos de análisis de desempeño y su aplicación sobre los clasificadores y conjuntos de datos propuestos.

- Construcción de una mesa multi-táctil basada en visión por computador para su uso en educación especial [29].
- Se estudiaron y propusieron métodos para detección en series temporales de fMRI [23][32].
- Se realizó la evaluación de rendimiento en sistemas de reconocimiento de patrones supervisados y de clasificación binaria [16].
- Se desarrollaron métodos de segmentación de imágenes de rango [20][30][33].
- Desarrollo de técnicas de selección en espacios multi-dimensionales [31].

## 4. FORMACION DE RECURSOS HUMANOS

La formación de recursos humanos en esta línea I/D se puede clasificar como indirecta o directa. En forma directa, actualmente, hay dos investigadores realizando su doctorado y cuatro alumnos han concluido tesis. Indirectamente como área permanente de consulta y transferencia para investigadores de líneas de I/D estrechamente relacionadas.

## 6. BIBLIOGRAFIA

1. Batagelj V, Bock H, Ferligoj A. "Data Science and Classification". Springer, (2006).
2. Fukunaga K. "Introduction to Statistical Pattern Recognition". Second Edition. Academic Press, (1990).
3. Devijer P, Kittler, J. "Pattern Recognition: theory and applications". Springer, (1986).
4. Anke Meyer-Base. "Pattern Recognition for Medical Imaging". Academic Press, (2004).
5. Kim H.Y., Giacomantone J. O., Cho, Z. H. Robust Anisotropic Diffusion to Produce Enhanced Statistical Parametric Map, Computer Vision and Image Understanding, v.99, pp.435-452 (2005).
6. Kim H.Y., Giacomantone J. O., A New Technique to Obtain Clear Statistical Para-

- metric Map by Applying Anisotropic Diffusion to fMRI, IEEE, International Conference on Image Processing. Proceedings, Genova, Italy, v.3, pp.724-727 (2005).
7. Corte C, Vapnik V, Support vector networks. Machine Learning v.20, pp.273-297 (1995).
  8. Vapnik, V. The Nature of Statistical Learning Theory. N. Y. Springer (1995).
  9. Li S., Fevens L., Krzyzak A., Li S. Automatic Clinical Image Segmentation Using Pathological Modelling, PCA and SVM, MLDN, LNAI 3587 pp.314-324, (2005).
  10. Lei Z., Yang Y., Wu Z.. Ensembles of Support Vector Machine for Text-Independent Speaker Recognition, IJCSNS v.6 n.5A pp. 163-167, (2006).
  11. Boekhorst R., Abnizova I., Wernich L. Discrimination of regulatory DNA by SVM on the basis of over and under-represented motifs, ESANN pp. 481-486 (2008).
  12. Vossen Anselm. Support Vector Machines in High Energy Physics, CERN, Geneva, Switzerland, pp.23-33 (2005).
  13. Wang L., Chang M., Feng J. Parallel and Sequential Support Vector Machines for Multi-label Classification, International Journal of Information Technology, v.11 n.9 pp. 11-18, (2005).
  14. Rüping S., SVM kernels for time series analysis, G1-Workshop-Woche Lernen-Lehren-Wissen-Adaptivitet, pp.43-50 (2001).
  15. Yang K., Shahabi C. A pca-based kernel for kernel pca on multivariate time series, IEEE Intern. Conf. on Data Mining (2005).
  16. Giacomantone J., De Giusti A., ROC performance evaluation of RADSPM technique, XIV Congreso Argentino de Ciencias de la Computación (CACIC), Chilecito (2008).
  17. Chaovalitwongse W., Pardalos P. On the Time Series Support Vector Machine using Dynamic Time Warping Kernel for Brain Activity Classification, Cybernetics and Systems Analysis v.44 pp.125-138 (2008).
  18. Von Luxburg U. A Tutorial on Spectral Clustering. Statistics and Computing, 17(4), (2007).
  19. Shi J., Malik J. Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(8), 888-905, (2000).
  20. Lorenti L., Giacomantone J. Segmentación espectral de imágenes utilizando cámaras de tiempo de vuelo. XI Workshop Computación Gráfica, Imágenes y Visualización. pp. 430-439. Mar del Plata, Argentina, (2013).
  21. Yang, P.; Huang, B. An Outlier Detection Algorithm Based on Spectral Clustering. Pacific-Asia Workshop on Computational Intelligence and Industrial Application, vol.1, pp. 507-510, (2008).
  22. Han Y., Feng X., Baciú G. Variational and PCA based natural image segmentation. Pattern Recognition 46, pp. 1971-1984 (2013).
  23. Giacomantone J., Tarutina T. Diffuse Outlier Detection Technique for Functional Magnetic Resonance Imaging. Computer Science and Technology Series. XVI Argentine Congress of Computer Science Selected Papers. pp. 255-265 (2011).
  24. Penne et al. Robust real time 3-D respiratory motion detection using time of flight cameras. International Journal of Computer Assisted Radiology and Surgery v.3 pp.427-431. (2008).
  25. Kollors, A., Penne, J., Hornegger, J., Barker A. Gesture Recognition with a Time-of-Flight camera. International Journal of Intelligent Systems Technologies and Applications, v. pp.334-343, (2008).
  26. Li J., Wang Y. Visual tracing and learning using speeded up robust features. Pattern Recognition Letters 33, pp. 2094-2101 (2012).
  27. Prato M., Zanni L. A practical use of regularization for supervised learning with kernel methods. Pattern Recognition Letters 34, pp. 610-618 (2013).
  28. Antunez E., Marfil R., Bandera A. Combining boundary and región features inside the combinatorial pyramid for topology preserving perceptual image segmentation. Pattern Recognition Letters 33, pp. 2245-2253 (2012).
  29. Cristina Manresa-Yee, Ramón Mas, Gabriel Moyá, María J. Abásolo, Javier Giacomanto-

- ne. Interactive multi-sensory environment to control stereotypy behaviours. *Computer Science & Technology*, pp. 121-128, (2012).
30. Lorenti L., Violini L., Giacomantone J. Selección sub-óptima del espectro asociado a la matriz de afinidad. IV Workshop Procesamiento de Señales y Sistemas de Tiempo Real. pp. 1417-1425. Mar del Plata, Argentina, (2013).
31. Giacomantone J., De Giusti A. Detección de áreas de interés bajo la hipótesis de relación espacial de voxels activados en fMRI. XX Congreso Argentino de Ciencias de la Computación. San Justo. Argentina. (2014).
32. Lorenti L., Giacomantone J., De Giusti A. Segmentación de imágenes de tiempo de vuelo vía clustering espectral co-regularizado. XX Congreso Argentino de Ciencias de la Computación. San Justo. Argentina. (2014).