

# Using visual and semantic features for anti-spam filters

**Francesco Gargiulo**, Antonio Penta, Antonio Picariello, Carlo Sansone

Università di Napoli Federico II  
Dipartimento di Informatica e Sistemistica  
via Claudio 21 — 80125 Napoli (Italy)  
`{francesco.grg,a.penta,picus,carlosan}@unina.it`

NIPS 2007–21st Neural Information Processing Systems Conference  
December 7–8, 2007 — Whistler, Canada

# Introduction

- Unsolicited Commercial Emails, commonly known as *spam*, are becoming a serious problem for email accounts of single users, small companies and large institutions.
- Previous approaches:
  - ▶ *Text Classification*
  - ▶ *Image Analysis and Processing*
  - ▶ *Combination of both*
- We propose a novel anti-spam system which uses a combination of:
  - ▶ *Image-Based features*
  - ▶ *Semantic Analysis techniques*in the email, in order to determine whether a message is spam.

# Methodologies

## Text Level

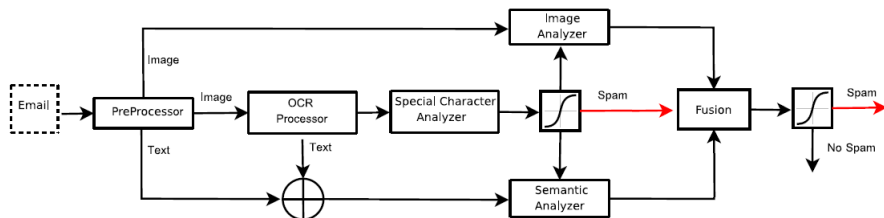
We apply Semantic Analysis processing with NLP techniques:

- **Part of Speech**
- **Latent Semantic Analysis**

## Image Level

We consider images that contain embedded text or not and we use :

- **Visual Features**
- **OCR-based Features**



## Preliminary Results

Database from the mail server of University of Naples “Federico II” composed by:

- 400 (text + attached images) spam e-mails
- 300 (text + attached images) no-spam e-mails

We used a *Decision Tree Algorithm* to obtain the results shown in the tables below.

- **Semantic Analysis Results** Accuracy =  $\frac{TruePositive + TrueNegative}{TotalNumberOfEmails}$

Features	Accuracy Mean	Accuracy Deviation
Adjectives	85.04%	36.31%
Nouns	94.56%	5.90%
Nouns + Adjectives	92.59%	21.71%

- **Image Analysis Results**

Features	Accuracy Mean	Accuracy Deviation
Visual	82.54%	1.31%
OCR-based	87.82%	4.14%
Visual + OCR-based	93.09%	4.85%

- **Combined Results with Behavior-Knowledge Space (BKS):**

- ▶ **Accuracy means = 92.80%, Accuracy Deviation = 13.46%**