

# Curriculum for a Graduate Course on the Integration of Vision and Language

## Prerequisite:

Introduction to Language Processing, Introduction to Digital Image Processing, Probability Theory, Introduction to Optimization, Introduction to Linear Algebra

## Objectives:

- Reviewing the basic of language and image foundations;
- Mastering the algorithms for joint processing;
- Application of aforementioned algorithm in image annotation and language understanding;
- Question and answering of multimedia data.

## Topics:

- Revision of basics of language: language foundations, word and sentence representations and their relation to syntax and semantics, basic language understanding tasks such as semantic role labeling, coreference resolution, temporal and spatial processing, and machine translation (1.5 weeks)
- Revision of basics of image: image foundation, spatial and frequency representation, feature extractions, colour representations, arithmetic and statistical operations on image (1.5 weeks)
- Supervised, semi-supervised and unsupervised learning: probabilistic modeling and HMM, deep learning, neural network, SVM, PCA/LDA/ICA (5 weeks)
- Structured learning and prediction (2 weeks)
- Optimization techniques (2 weeks)
- Applications in multimedia search (1 week)

**Note:** All the examples and illustrations will be in line with the joint language and image processing.

## Assessment:

HW/Assignments	25%
Course Project	40%
Final Exam	35%