

# Speaker and Language Recognition Techniques

Sandro Cumani

Politecnico di Torino

# Introduzione

- Riconoscimento vocale
  - Riconoscimento del parlato
  - Riconoscimento del parlatore
  - Riconoscimento della lingua
- Numerosi ambiti applicativi
  - Trascrizione automatica
  - Autenticazione
  - Ambito forense
- Disponibilità di risorse assenti in passato

## Sommario

- Modelli acustici e fonetici
- Metodi di classificazione generativi
- Metodi di classificazione discriminativi

# Modelli fonetici

- Stato dell'arte: Decoder fonetici basati su reti neurali
- Migliorare la decodifica: addestramento basato su GPU
  - Notevole riduzione dei tempi di addestramento
  - Possibilità di utilizzare dataset di dimensioni maggiori
  - Elevato impatto dal punto di vista applicativo

# Modelli generativi

- Stato dell'arte basato su i-vectors
  - Vettore di dimensioni ridotte che caratterizza una frase
  - Richiede elevate risorse computazionali
  
- Semplificazione del processo di estrazione di i-vectors
  - Basato su approssimazioni variazionali di distribuzioni
  - Permette un'estrazione più rapida con un notevole risparmio dal punto di vista della memoria
  - Mantiene l'accuratezza del sistema classico
  - Applicabile in campo mobile

# Modelli discriminativi per il riconoscimento della lingua

- Stato dell'arte basato su Support Vector Machines
- Ottimizzazione algoritmica
  - Analisi di diversi algoritmi di classificazione applicati a problemi di riconoscimento della lingua
  - Identificazione delle soluzioni ottimali per l'addestramento dei modelli
- Caratterizzazione di frasi attraverso vettori di ridotte dimensioni

# Modelli discriminativi per il riconoscimento del parlatore

- Stato dell'arte basato su Support Vector Machines
  - Notevoli criticità
- Definizione di un nuovo framework per l'addestramento discriminativo
  - Risolve le criticità dei sistemi allo stato dell'arte
  - Permette di ottenere risultati nettamente superiori rispetto ai modelli precedenti
  - Permette la costruzione di sistemi le cui performance eguagliano e, in alcuni casi, superano quelle dei sistemi generativi

# Conclusioni

Tecniche innovative sia dal punto di vista algoritmico che dal punto di vista computazionale

Ricadute in termini di applicazioni reali

Integrate in applicazioni di verifica e identificazione in ambiti aziendali