

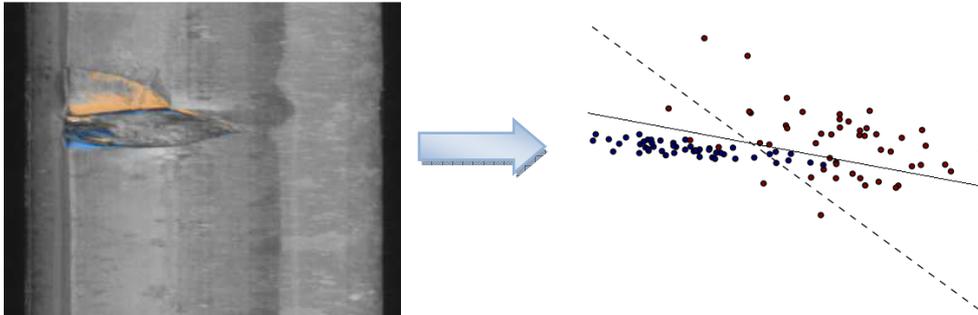
Proposta per tesi di Laurea Magistrale area INFORMATICA-MATEMATICA

Data mining and machine learning (estrazione di informazione e apprendimento automatico).

CLASSIFICAZIONE SEMI-SUPERVISIONATA DI DIFETTI INDUSTRIALI (DA ANALISI DI IMMAGINI)

Nessuna produzione industriale può essere dichiarata esente da imperfezioni. La classificazione di difetti nelle componenti industriali è un importante ambito di automazione, poiché permette di automatizzare e velocizzare la produzione, e migliorare la qualità del prodotto. L'obiettivo di questa tesi è studiare e sviluppare un sistema automatico *non supervisionato* per la classificazione di difetti superficiali in prodotti industriali.

I difetti superficiali dei prodotti industriali, una volta individuati, possono essere rappresentati con dei descrittori ("feature"), che vengono estratti dalle misure fornite dai sensori ottici (immagini). La classificazione ha il compito di decidere a quale classe di appartenenza debba essere assegnato ciascun descrittore. Può essere di tipo supervisionato, nel caso in cui l'insieme delle classi sia definito a priori, oppure *non supervisionato*, nel caso sia il classificatore ad individuare e discriminare le classi naturali presenti nei dati. Esempi di classificatori non supervisionati sono k-means, HMM, PCA, SVD. Nel caso in cui pochi esempi di classi vengono fornite al sistema, che successivamente provvede ad aggiornare la sua conoscenza mediante tecniche adattive, si parla di addestramento *semi-supervisionato*.



E' a disposizione dello studente una banca dati con i maggiori difetti industriali per l'applicazione in oggetto. La complessità e la durata dello studio e dell'implementazione potranno essere regolate in base al background dello studente.

Organizzazione della tesi

- Ricerca bibliografica della letteratura riguardante la classificazione semi-supervisionata, contestualizzata per i difetti industriali.
- Analisi di una o due tecniche adattive (es: online boosting, online clustering, ecc.) o tecniche non convenzionali (manifolds, ecc.).
- Scelta di un classificatore (generativo) e suo addestramento con pochi campioni di input.
- Addestramento incrementale con campioni randomizzati e tecniche conservative di apprendimento.
- Test su modelli sintetici e su immagini reali.

Requisiti

- Conoscenza di elaborazione d'immagini di base.
- Conoscenze matematiche buone/avanzate, in particolare algebra lineare, metodi numerici, statistica e probabilità.
- Dimestichezza con la programmazione in linguaggi di alto livello (MATLAB o C++/C#).

Complessità

- Alta.

Tempi previsti di sviluppo: 9 mesi

- Possibilità di pubblicare i risultati.

Riferimenti Danieli Automation**Alessandro ARDESI**

BU Manager
Instrumentation&Robotics

Phone +39 0432.518535 (3535)

Fax +39 0432.518011

E-mail a.ardesi@dca.it

WebSite www.dca.it

Riferimenti UniUD**Lauro SNIDARO**

Assistant Professor
Computer Science

Phone +39 0432 558444

E-mail lauro.snidaro@uniud.it

WebSite www.dimi.uniud.it/snidaro