UNIVERSITÀ CATTOLICA DEL SACRO CUORE
INTERFACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
E SCIENZE BANCARIE, FINANZIARIE E ASSICURATIVE

CORSO DI LAUREA IN APPLIED DATA SCIENCE FOR BANKING
AND FINANCE

TESI DI LAUREA MAGISTRALE

Image classification for Inspection Machines: a comprehensive review and application study

Relatore:
Prof. Enrico Barbierato
Corelatore:
Prof. Alfredo Marzocchi

Candidata:
Elisa Salvi

Matricola:
5005374

ANNO ACCADEMICO 2021/2022

Contents

Introduction

The healthcare industry has made substantial progress during the past several decades embracing more innovations and cutting-edge technologies. The research of Computer Vision, imaging processing and pattern recognition enables more efficient diagnosis, treatment, health care operations, efficient medical records management and more, especially during emergencies. Machine Learning algorithms can handle complex data sets with applications in many fields like radiology, medical robotics, medical imaging, drug discovery, medical transcription and so on. The expectation for the near future is to see more and more healthcare companies and institutions experiment with technologies at the frontiers of knowledge to improve their services.

The progress of Computer Vision for healthcare applications is documented in a collection of articles called the "Journal of Healthcare Engineering" [1]. This journal enhances the development of these research fields and aims to bring together researchers, engineers, mathematicians and programmers in order to understand the problems usually encountered in these contexts through rapid online publications. This journal deals with several application engineering areas like cardiovascular, clinical, rehabilitation, neural, respiratory systems and so on. Moreover, the papers include topics such as biomedical imaging, image processing and sensors for bioinstrumentation.

Medical image technology is highly valuable to clinical analysis and treatment because it can provide important information about internal organs. Moreover, it can help doctors understand the patient's health condition and treat various diseases converting scanned images into interactive 3D models.

Computer Vision does not merely analyze medical images, but it also offers accurate data about phenomena that are hard to measure with traditional methods. A significant example is the case of the "Orlando Health Winnie Palmer Hospital for Women and Babies" [2]. One of the main causes of mortality in childbirth is postpartum hemorrhaging; this hospital is developing image analysis algorithms to precisely measure the blood loss during childbirth. In this way, the doctor estimation of the amount of blood lost during the delivery is improved in accuracy, minimizing human errors and improving treatment procedures for women.

Computer Vision is a branch of Artificial Intelligence (AI). The goal of Computer Vision is to develop digital systems to process, analyze and interpret visual data. Artificial Intelligence is a field of computer science that deals with brain-inspired software architectures for fast data-driven decision making and embraces Machine Learning, Computer Vision, natural language understanding, computational logic and processing. AI aims at building efficient algorithms for the recognition and the classification of objects, situations, and events. The challenge consists of under-

standing how the human brain processes information and apply the same methods to train specific software, making it capable of accurate and robust results. The incipit of "Artificial Intelligence" as a new discipline is defined in 1956, when a famous conference was held at Dartmouth College in Hanover, New Hampshire, America, and the organizers became the founders of AI research. This conference was proposed in August 1955 by J. McCarthy, M. L. Minsky, N. Rochester and C. E. Shannon. These researchers wrote a paper called "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence" [3] that suggested the examination of topics such as Neural Networks, computability theory, natural language processing and recognition. This was a highly stimulating meeting; relevant successes in the field of AI would be achieved by the very participants or their students. The participants were all prominent scientists of the time: the mathematician John McCarthy, the researcher in mathematics and neurology from Harvard Marvin Minsky, the director of information research at IBM Nathaniel Rochester, the mathematician and researcher in information theory Claude Shannon. The event was conducted as a brainstorming session where researchers could debate openly about the possibility to make machines behave like humans.

The scientific community agrees on the year 1956 as the first time the term "Artificial Intelligence" was coined, although substantial contributions to AI ware made in previous years. AI had a strategic role during the Second World War. The idea that human intelligence could be simulated through the use of machines was born in the 1940s with the great contribution of the mathematician Alan Turing. The British scientist is considered one of the fathers of modern computing and his work had important applications especially in the field of cryptography. Turing collaborated with the British government to decipher the German military transmissions and develop the code behind the Enigma device. Moreover, he formulated the concepts of calculability, computability, and the Turing machine, that is a theoretical model of a machine capable of executing any kind of computable sequence. In 1950, Turing developed the "Turing Test", also known as the "Imitation Game", whose aim was to assess the presence of human intelligence in a machine. This work conducted by Turing, who died in 1954, played a key role in the Dartmouth conference.

The entrance of Computer Vision was marked by human visual system research conducted in various universities during the 1960s. The studies about how neurons react to different kind of stimuli yielded optimistic results. Scientists understood that human vision is characterized by a hierarchical analysis and detection of feature, like shapes, texture and edges. Furthermore, they took inspiration from how the human brain processes visual information to develop a Computer Vision system capable of achieving a complete understanding of an image. Many algorithms that are still relevant today were already described, such as edge detection, motion analysis and so on. The hype of these new projects was so high that researchers experienced difficulties to live up the public's expectations. The technical background failed to keep up with the complexity of the problems. This difficult period was called "AI winter" and scientist realized that their challenges would be solved at best in the next decades. The following years were characterized by more accurate mathematical researchers who laid the foundations of Computer Vision algorithms.

Computer Vision algorithms can handle both black and white and color images

and operate in many different areas; some examples of Computer Vision tasks are:

- image classification analyses an image and determines if it belongs to a certain class. This task often faces some challenges such as find the right classification for ambiguous images. This technique is used in the field of skin cancer detection to automate the identification process of cancerous tissue and reduce the chances of human error. The identification and the diagnosis of moles from melanomas are often difficult to the naked eye. Therefore, several algorithms are trained with a vast database of images consisting of both healthy and cancerous tissue.

- Object detection can identify particular entities within an image or a video. A concrete example is automatization of the damage detection on an assembly line, instead of training personnel for the inspection process that is way more expensive. Quality control is crucial to ensure the highest level of customer approval and limit possible complications after the sale of the product. In the medical field, this method is applied in x-ray radiology, ultrasound, endoscopy and many more procedures in order to help medical professionals identifying any issues or abnormalities in patients' internal organs.

- Image segmentation divides the image into sections and analyzes the delimited areas individually. This method is useful to highlight the pixels of a tumorous section of a tissue in a medical report. Tumors spread quickly to other parts of the body, so making early detection is especially crucial. It can reduce the possibility of human error and save a patient's life by identifying even the slightest difference with a high level of certainty.

- Action and emotion recognition concerns the identification of specific actions and sentiments showed in an image or a video. This technique can be used to assess the quality of a surgery by measuring the level of activity, analyzing and detecting medical professional movements.

- Object tracking detects and monitors an object in real-time video feeds. This technique is used, for example, for autonomous vehicles to detect, classify and track other cars in motion to avoid collisions. This is also crucial for to develop monitoring systems in order to identify any risky situations for the workers or the environment.

- Image editing modifies an image, for example to obscure sensitive data for privacy issues.

- Content-based image retrieval searches images from large data collections, relying only on the content of the images.

The Covid pandemic presented a tough challenge for the global healthcare system and has required worldwide researchers to elaborate and implement new strategies. Computer Vision has provided a significant contribution against this biological threat [4]. The previous listed tasks can be applied to the diagnosis, control, treatment, and prevention of Covid-19 infections. Several tools are implemented to

monitor and prevent the spread of the pandemic, for example masked face detection software, radiography images analysis, remote patients monitoring and occupancy detection to ensure social distance. These techniques are adopted more and more by healthcare facilities to improve service quality and increase efficiency. In January 2021 the first open-source network, called "COVID-Net", was designed to handle Covid-19 hospitalization data, like age, sex, ethnicity and underlying health conditions of Covid patients, and also allowing to weekly update hospitalization rates.

However, the implementation of Computer Vision algorithms deals with some critical issues. Firstly, in real life the datasets are not sufficiently large to train the algorithms properly. Consequently, "Data Augmentation" technique, explained in chapter 1.1, are often used in order to develop algorithms robust to image transformations, like non-optimal brightness conditions, partial coverage of the subject, scale variations, blurring and so on. Moreover, the process of image feature extraction can be negatively influenced by inconspicuous features.

Despite these difficulties machine vision systems have numerous applications. In the industrial and manufacturing sectors, Computer Vision has an important role thanks to the possibility of being integrated directly on production lines and in factory environments. This technology is exploited for many areas of 4.0 industry, like monitoring and predicting the maintenance of industrial assets or classifying anomalies on the surface of mechanical component [5].

This dissertation aims to review of the most used techniques for image preprocessing, like filtering, edge detection and Hough trasform, and feature extraction, such as RGB levels, histogram of oriented gradients, scale invariant feature and texture, in order to achieve an accurate classification of a database composed of visual imagery. Different techniques for the classification will be presented, like Decision Tree, Random Forest, Support Vector Machine, K-Nearest Neighbors and Convolutional Neural Network. Is important to analyze various algorithms because it is impossible to determine the best way to classify an image a priori, this depends on various features that characterize the data. At the end, an application example about the classification of pharmaceutical ampoule caps will be described as a practical demonstration of the explained theoretical background presented.

DA FINIRE