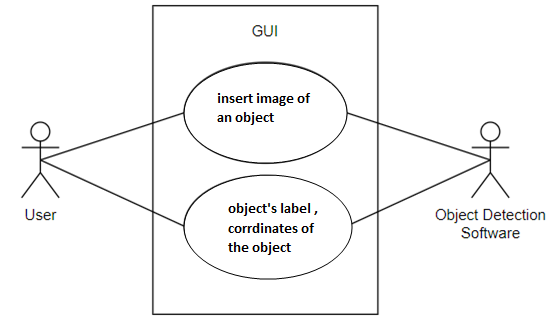
**AI documentation**

1. ***Project Idea:***

* A desktop application that can identifies some objects from an image, using Artificial Neural Networks based on classification model that classifies data set that covers objects which include pens, phone, cup, keyboard, notebook. However, identification problem can be broken down into multiple tasks, starting with object detection at order level then the object classification illustrated below:
* A computer on a table

  Description automatically generated with low confidence

1. ***Main functionalities:***



* The software will be used to identify objects based on artificial neural network classification, which includes the following: Phone, Pen, Notebook, Cup, and Keyboard.
* The user will be prompted to insert an image of an object they would like to identify, and by using Object Detection the software will be outputting the object’s label, and coordinates of the object for the inserted image. Classifying the objects in the image.

***(3) Similar applications in market:***

Object detection is **a computer vision technique for locating instances of objects in images or videos**

***Autonomous Driving.*** Self-driving cars depend on object detection to recognize pedestrians, traffic signs, other vehicles, and more. For example, Tesla’s Autopilot AI heavily utilizes object detection to perceive environmental and surrounding threats such as oncoming vehicles or obstacles. 

***Animal detection in Agriculture***. Object detection is used in agriculture for tasks such as counting, animal monitoring, and evaluation of the quality of agricultural products. Damaged produce can be detected while it is in processing using machine learning algorithms 

***People detection in Security***. A wide range of security applications in video surveillance are based on object detection, for example, to detect people in restricted or dangerous areas, suicide prevention, or automating inspection tasks in remote locations with computer vision.

***Medical feature detection in healthcare***. Object detection has allowed for many breakthroughs in the medical community. Because medical diagnostics rely heavily on the study of images, scans, and photographs, object detection involving CT and MRI scans has become extremely useful for diagnosing diseases, for example with ML algorithms for tumour detection. 

***Vehicle detection with AI in Transportation.*** Object recognition is used to detect and count vehicles for traffic analysis or to detect cars that stop in dangerous areas, for example, on crossroads or highways. 

***Object detection in Retail***. Strategically placed people counting systems throughout multiple retail stores are used to gather information about how customers spend their time and customer footfall. AI-based customer analysis to detect and track customers with cameras helps to gain an understanding of customer interaction and customer experience, optimize the store layout, and make operations more efficient. A popular use case is the detection of queues to reduce waiting time in retail stores. 

***(4) An initial literature review of Academic publications (papers) relevant to the idea:***

***Abstract***

***Paper1***

As one of the important tasks in computer vision, target detection has become an important research hotspot in the past 20 years and has been widely used. It aims to quickly and accurately identify and locate a large number of objects of predefined categories in a given image. According to the model training method, the algorithms can be divided into two types: single-stage detection algorithm and two-stage detection algorithm. In this paper, the representative algorithms of each stage are introduced in detail. Then the public and special datasets commonly used in target detection are introduced, and various representative algorithms are analysed and compared in this field. Finally, the potential challenges for target detection are prospected.

**INTRODUCTION**

Object detection is a basic research direction in the fields of computer vision, deep learning, artificial intelligence, etc. It is an important prerequisite for more complex computer vision tasks, such as target tracking, event detection, behaviour analysis, and scene semantic understanding. It aims to locate the target of interest from the image, accurately determine the category and give the bounding box of each target. It has been widely used in vehicle automatic driving, video and image retrieval, intelligent video surveillance [1], medical image analysis [2], industrial inspection [3] and other fields. Traditional detection algorithms on manually extracting features mainly include six steps: pre-processing, window sliding, feature extraction, feature selection, feature classification and post-processing and generally for specific recognition tasks. Its disadvantages mainly include small data size, poor portability, no pertinence, high time complexity, window redundancy, no robustness for diversity changes, and good performance only in specific simple environments. In 2012, AlexNet image classification model based on convolutional neural network (CNN) was proposed by Krizhevsjy[4]and others.

 In the image classification competition of the image dataset ImageNet [5], they won the competition with a huge advantage of 11% accuracy over the second place using traditional algorithms. Many scholars have begun to apply deep convolutional neural networks to target detection tasks, and have proposed many excellent algorithms. It can be roughly divided into two categories: the single-stage detection algorithm based on region proposal and the two-stage detection algorithm based on regression.

For more information about the paper

***resources:***

* shorturl.at/aAC18

***Paper 2***

***Introduction to Artificial Neural Network (ANN) Methods:***

What They Are and How to Use Them\*. Jure Zupan1),

Department of Chemistry, University Rovira i Virgili, Tarragona, Spain Basic concepts of ANNs together with three most widely used ANN learning strategies (error back-propagation, Kohonen, and counter propagation) are explained and discussed.

In order to show how the explained methods can be applied to chemical problems, one simple example, the classification and the prediction of the origin of different olive oil samples, each represented by eight fatty acid concentrations, is worked out in detail. Introduction In the last few decades, as the chemists have got accustomed to the use of computers and consequently to the implementation of different complex statistical methods, they are trying to explore multi-variate correlations between the output and input variables more and more in detail. With the increasing accuracy and precision of analytical measuring methods it become clear that all effects that are of interest cannot be described by simple uni-variate and even not by the linear multivariate correlations precise, a set of methods, that have recently found very intensive use among chemists are the artificial neural networks (or ANNs for short). \_\_\_\_\_\_\_\_\_\_ \*) The lecture presented at the VI-th COMETT Italian School on Chemometrics, Alghero, Sardinia, Italy, 26-30-st September 1994. 1) On leave from the National Institute of Chemistry, Ljubljana, Slovenia Jure Zupan, Introduction to ANNs Acta Chimica Slovenica 41/3/1994, pp. 327-352 2 Therefore, the analytical chemists are always eager to try all new methods that are available to solve such problems. One of the methods, or to say more Due to the fact that this is not one, but several different methods featuring a wide variety of different architectures learning strategies and applications, we shall first start with explaining the overall strategy, goals, implications, advantages and disadvantages, and only after explaining that, we shall discuss the fundamental aspects of different approaches to these methods and how they can be put to use in analytical chemistry. The ANNs are difficult to describe with a simple definition. Maybe the closest description would be a comparison with a black box having multiple input and multiple output which operates using a large number of mostly parallel connected simple arithmetic units. The most important thing to remember about all ANN methods is that they work best if they are dealing with non-linear dependence between the inputs and outputs (Figure 1). ANNs can be employed to describe or to find linear relationship as well, but the final result might often be worse than that if using another simpler standard statistical technique. Due to the fact that at the beginning of experiments we often do not know whether the responses are related to the inputs in a linear on in a nonlinear way, a good advice is to try always some standard statistical technique for interpreting the data parallel to the use of ANNs. Black box Input variables Non

- ***resources:***

shorturl.at/eEI15

***Paper 3***

***Abstract***—An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. ANNs, like people, learn by example. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons. This is true of ANNs as well. This paper gives overview of Artificial Neural Network, working & training of ANN. It also explain the application and advantages of ANN

***Introduction*** The study of the human brain is thousands of years old. With the advent of modern electronics, it was only natural to try to harness this thinking process. The first step toward artificial neural networks came in 1943 when Warren McCulloch, a neurophysiologist, and a young mathematician, Walter Pitts, wrote a paper on how neurons might work. They modeled a simple neural network with electrical circuits. Neural networks,

with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained neural network can be thought of as an "expert" in the category of information it has been given to analyze.

***Other advantages include:***

1. Adaptive learning: An ability to learn how to do tasks based on the data given for training or initial experience. 2. Self-Organization: An ANN can create its own organization or representation of the information it receives during learning time. 3. Real Time Operation: ANN computations may be carried out in parallel, and special hardware devices are being designed and manufactured which take advantage

***resources:***

* shorturl.at/eEI15

***Paper 4***

***Abstract-*** The essence of the study is to analyze an algorithm which will provide a robust and computationally light method, which might be suitable to implement in the real-time industrial application such as object detection and recognition. For industrial applications, the primary step in automatic detection and classification of an object is to find the object automatically from an image using features related to its shape. This chore is a very complex one. Therefore, to hit the target Histogram of oriented gradient (HOG) algorithm is selected to extract the image features. Average Magnitude Difference Function AMDF is employed to correct the alignment defect. Finally, Artificial Neural Network (ANN) was employed to detect the type of object in the image efficiently. None the less, a database was generated. The database consists of images of real industrial products which are of different shapes and sizes, captured under different lightning conditions. The outcome of the experiment conducted on the database recorded 98.10% success.

***Introduction*** For industrial applications, the primary step in automatic detection and classification of an object is to find the object automatically from an image using features related to its shape. Therefore, shape detection plays an important role and has an intensive usage in various applications [1] such as in robotics; object classification is needed to recognize a certain object in a cluttered scene [2]. In the industries detecting objects with computers is difficult. This is as a result of some Complications such as Initially, products from industries vary in many forms such as type, size and Shape. Hence, using object detection system in the industries is only possible if the algorithm is robust enough to eliminate these variations. Subsequently, the industrial operation involves moving objects from one place to the other; this might result in objects being placed in different poses. The methods used needs to distinguish the objects with different orientations accurately. Thirdly, the illumination in the environment varies due to, shadows of other machine or humans. Therefore, the algorithms are required to function efficiently under different illumination condition.

Finally, partial occlusion can be a huge problem to the performance of the algorithms. Partial occlusion is inevitable in the industries since many products are processed at a time. To deal with this situation the algorithm needs to recognize an object even if it is partially occluded.

The actual idea behind this work is to utilize an algorithm that is computationally light, but it also shows high performance in detecting and classifying an object in an image. In this study, a comprehensive analysis of the algorithm was carried out using a more complex database and the emphasis will be on the industrial application.

On the one hand, HOG is utilized to extract image feature. However, HOG in not entirely invariant for rotation, therefore AMDF is employed to detect the image pose, which is then removed. Then, the features are passed to the MLP for training. While, on the other hand, after MLP is trained and then used to classify already aligned HOG features obtained from an input image.

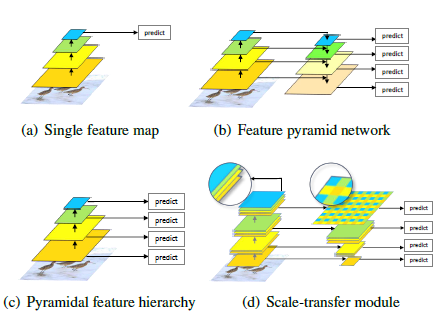
***resources:***

* shorturl.at/hruxC

***Paper 5***

***Scale-Transferrable Object Detection***

***Abstract-*** Scale problem lies in the heart of object detection. In this work, we develop a novel Scale-Transferrable Detection Network (STDN) for detecting multi-scale objects in im-ages. In contrast to previous methods that simply combine object predictions from multiple feature maps from differ-ent network depths, the proposed network is equipped with embedded super-resolution layers (named as scale-transfer layer/module in this work) to explicitly explore the inter-scale consistency nature across multiple detection scales. Scale-transfer module naturally ﬁts the base network with little computational cost. This module is further integrated with a dense convolutional network (DenseNet) to yield a one-stage object detector. We evaluate our proposed archi-tecture on PASCAL VOC 2007 and MS COCO benchmark tasks and STDN obtains signiﬁcant improvements over the comparable state-of-the-art detection models.

***Introduction*** Scale problem lies in the heart of object detection. In order to detect objects of different scales, a basic strategy is to use image pyramids [1] to obtain features at different scales. However, this will greatly increase memory and computational complexity, which will reduce the real-time performance of object detectors.

In recent years, convolutional neural networks (CNN) [18] have achieved great success in computer vision tasks, such as image classiﬁcation [17], semantic segmentation [23], and object detection [10]. The hand-engineered features are replaced with features computed by convolutional neural networks, which greatly improves the performance of object detectors. Faster R-CNN [25] uses convolutional feature maps computed by one layer to predict candidate re-gion proposals with different scales and aspect ratios (Figure 1(a)). Because the receptive ﬁeld of each layer in CNN is ﬁxed, there exists inconsistency between the ﬁxed receptive ﬁeld and the objects at different scales in natural im-ages. This may compromise object detection performance. SSD [22] and MS-CNN [3] use feature maps from different layers within CNN to predict objects at different scales (See Figure 1(c)). Shallow feature maps have small receptive ﬁelds that are used to detect small objects, and deep feature maps have large receptive ﬁelds that are used to detect large objects. Nevertheless, shallow features have less semantic information, which may impair the performance of small object detection. FPN [20], ZIP [19] and DSSD [7] integrate semantic information on feature maps at all scales. As shown in Figure 1(b), a top-down architecture combines high-level semantic feature maps with low-level feature maps to yield more semantic feature maps at all scales. However, in order to improve detection performance, feature pyramids must be carefully constructed….

***resources:***

* shorturl.at/ryP23

***(5) Dataset used:***

We captured some real-life images of the following objects, to train the model. {pen, notebook, cup, keyboard, phone}

***Drive link:***

- shorturl.at/lBJ03

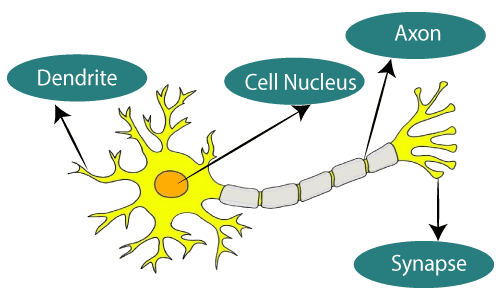
***(6)Details of the algorithm(s)/approach(es)***

***The Algorithms we will use:***

* Artificial neural networks (ANN)
* Convolution neural networks (CNN)

***What are Artificial neural networks?***

* The term "**Artificial Neural Network**" is derived from Biological neural networks that develop the structure of a human brain.
* Similar to the human brain that has neurons interconnected to one another, artificial neural networks also have neurons that are interconnected to one another in various layers of the networks. These neurons are known as nodes.

Diagram, schematic

Description automatically generated

Biological Neural Network. Artificial Neural Network

***Characteristics of ANNs:***

1. Self-organizing and self-learning ability: It can change and adjust its structure by interacting with the environment.

1. Strong promotion ability

1. Highly parallel: A large number of similar or independent operations can be performed simultaneously.

1. Strong information synthesis ability: It can process both quantitative and qualitative information, and can coordinate multiple input information fusion and multimedia technologies.

***Why we use ANN to solve Object detection problem?***

* ANNs have the ability to learn and model non-linear and complex relationships, which is really important because in Object Detection problem, the relationships between inputs and outputs are non-linear as well as complex.

***How we will use ANN To solve Object detection problem? :***

* The inputted image is first divided into nearly two thousand region sections.

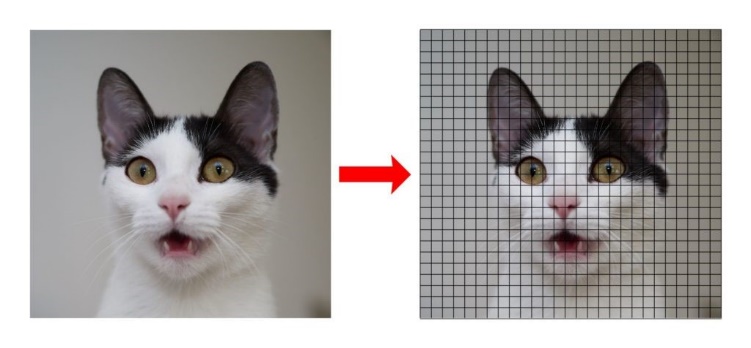
* Then a neural network is applied for each region, respectively.

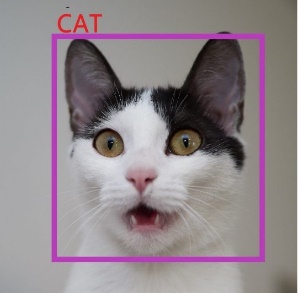
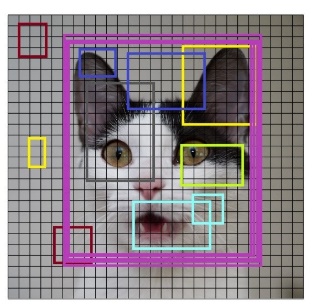
* The size of the regions is calculated, and the correct region is inserted into the neural network.

* It classifies and creates bounding boxes individually.

***What is Convolution neural networks ?  :***

* Convolutional Neural Network, also known as CNN or ConvNet, is a class of neural networks that specializes in processing data that has a grid-like topology, such as an image. A digital image is a binary representation of visual data. It contains a series of pixels arranged in a grid-like fashion that contains pixel values to denote how bright and what colour each pixel should be.



   
***Characteristics of CNNs:***

*•* Convolutional neural network is a class of deep learning methods which has become dominant in various computer vision tasks and is attracting interest across a variety of domains, including radiology.

• Convolutional neural network is composed of multiple building blocks, such as convolution layers, pooling layers, and fully connected layers, and is designed to automatically and adaptively learn spatial hierarchies of features through a backpropagation algorithm.

• Familiarity with the concepts and advantages, as well as limitations, of convolutional neural network is essential to leverage its potential to improve radiologist performance and, eventually, patient care.

***Block digrams***

Diagram

Description automatically generated

Diagram

Description automatically generated

***(7) Development Platform:***

# We used:

* VsCode as IDE
* Python Tensor flow to implement the algorithm
* Keras to build models
* Tkinter to build the GUI