**4.4 Machine Learning algorithms for image classification**

**Reading material**

|  |
| --- |
| Machine Learning Algorithms for Image Classification  So far you learned what machine learning is and the key steps in devising an ML solution. In this session you will have a look at the general concepts and algorithms for image classification. |
| **General Concepts**  As a broad subfield of artificial intelligence, machine learning is concerned with algorithms and techniques that allow computers to “learn” by example. The major focus of machine learning is to extract information from data automatically by computational and statistical methods. It allows us **to give our data a voice**. Machine learning is now being routinely used to work with large volumes of data in a variety of formats such as image, video, sensor, health records, etc.  When machine learning is used for classification, empirical models are built to classify the data into different categories, aiding in the more accurate analysis and visualization of the data. Applications of classification include facial recognition, credit scoring, and cancer detection. When it is used for clustering, or unsupervised classification, it aids in finding the natural groupings and patterns in data.  Applications of clustering include medical imaging, object recognition, and pattern mining. Object recognition is a process for identifying a specific object in a digital image or video. Object recognition algorithms rely on matching, learning, or pattern recognition algorithms using appearance-based or feature-based techniques. |
| **What is an image?**   * For the purposes of computer science, an image is an array of pixels. * A pixel (picture element) is the smallest addressable unit in an image and is represented by a single number or a group of numbers and its position in the array. * Since a pixel is just a number or a group of numbers and an image is an array of pixels then an image is just an array of numbers * Depending on the type of the pixels an image can be a two-dimensional or three-dimensional array     Source: <https://medium.com/analytics-vidhya/cnn-series-part-1-how-do-computers-see-images-32462a0b33ca>  Source: https://www.ibm.com/docs/bg/gddm?topic=ivu-what-is-image |
| **What is image classification?**   * Image classification is the task of assigning a label or class to an entire image. * A model takes an image as input and outputs probability that the image contains an object that belongs to one of the classes in the label set. * The main aim of any image classification-based system is to assign semantic labels to captured images and consequently.   Source: https://huggingface.co/tasks/image-classification |
| * Image classification can be used for image search   + Image search is when you use text to find images that contain the object you described in the text * Image classification can be used for automatic image tagging   + Automatic image tagging adding textual or categorical description to an image   + For example, when a user uploads an image containing an orange cat you may want to add descriptions such as orange, cat, Garfield * The state-of-the-art image classification systems use deep learning. For example, in the challenge known as ImageNet Large Scale Visual Recognition Challenge (ILSVRC) deep learning methods perform far better than the classical ML methods in classifying 1000 classes of objects. |
| * **What is deep learning?**     Source: https://www.flickr.com/photos/eatyourgreens/2635712607/ |
| **Deep Learning**   * Deep learning is a subset of machine learning that uses an Artificial Neural Network (ANN) with three or more layers as its classifier. For now, we will not define what ANN is or what it means to have three or more layers. These concepts are discussed in-depth in Module 7. * Why deep learning?   + One of the biggest advantages of deep learning as compared to the rest of machine learning is its ability to handle unstructured data well   + In the rest of machine learning, a human expert curates the set of features to be used for each machine learning task. With deep learning it is not necessary to do this. Deep learning classifiers can identify relevant features on their own.     Source: <https://labelyourdata.com/articles/machine-learning-and-training-data>  Source: <https://www.ibm.com/topics/deep-learning> |
| Is it still hard to grasp what machine learning and deep learning is? Let's try to explain it again with a video.  <https://www.youtube.com/watch?v=q6kJ71tEYqM> (time duration 7:48) |
| **Traditional Image Classification Approaches**  Before the beginning of use of deep learning for image classification, image classification was highly reliant on some standard descriptors (image features).  Descriptors are small “interesting”, descriptive or informative patches in images. Several image processing algorithms, such as edge detection, corner detection or threshold segmentation may be involved in this step. As many features as practicable are extracted from images and these features form a definition (known as a bag-of-words) of each object class. At the deployment stage, these definitions are searched for in other images. If a significant number of features from one bag-of-words are in another image, the image is classified as containing that specific object (i.e. chair, horse, etc.). |
| The difficulty with this traditional approach is that it is necessary to choose which features are important in each given image. As the number of classes to classify increases, feature extraction becomes more and more cumbersome. It is up to the computer vision engineer’s judgment and a long trial and error process to decide which features best describe different classes of objects. Moreover, each feature definition requires dealing with a plethora of parameters, all of which must be fine-tuned by the computer vision engineer.  Source: https://arxiv.org/abs/1910.13796 |
| **Traditional Image Classification**  The face image below shows a traditional descriptor known as SURF descriptor. The SURF descriptor describes blobs in the image that can be used to classify the image. In the given image you can see there are detected features around the eyes, the nose, the mouth, the chin and the jaw. Aren’t these face features what we use to recognize faces?    Source: <https://www.flickr.com/photos/daniel-sikar/49800412293> |
| **Traditional ML-based Image Classification Vs Deep Learning Image Classification**  The following image shows the differences between the traditional ML workflow and the deep learning workflow for image classification. In the case of deep learning, the image is simply given as input to the model while in traditional ML a human makes several decisions on parameters used to extract features and/or how to combine them. In the case of SURF, it is unlikely you will have good results for your application domain if you don’t experiment with the SURF feature extraction parameters.  There are different parameters such as threshold, number of octaves, number of layers which must be decided by a human expert based on experimentation or to get good results. A human does not need to be involved at any level for feature extraction in the case of deep learning. |
| **Image Classification in Remote Sensing**  The traditional approaches for remote sensing image analysis are based on low-level and mid-level feature extraction and representation. These techniques have shown good performance by using different feature combinations and machine learning approaches. These earlier approaches have used small scale image dataset. The recent trends for remote sensing image analysis are shifted to the use of machine learning models which help to process a huge amount of dataset in accurate and rapid way.  Satellite image classification is a multilevel process that starts from extracting features from images to classifying them into categories. Image classification is a step-by-step process that starts with designing a scheme for classification of desired images. |
| After that, the images are preprocessed which includes image enhancement, and scaling. Thereafter, the algorithm is applied on the images to get the desired classification. Corrective actions are made after applying algorithms, which is also called post processing.  In the final phase the accuracy of this classification is assessed. It is possible to apply both supervised and unsupervised machine learning methods for image classification in remote sensing.  Source: Remote Sensing Image Classification: A Comprehensive Review and Applications by Maryam Mehmood et al (March,2022), https://doi.org/10.1155/2022/5880959 |
| **Application Domains of ML in Remote Sensing Data**    Source: https://artemisat2.ulpgc.es/wp-content/uploads/2019/11/MachineLearningRemoteSensing-GustauCamps-Valls.pdf |
| **Machine Learning Tasks in Remote Sensing**  Machine learning has been used for various tasks in remote sensing.  As you discussed in Modules 1 – 3 of this course, a remote-sensing dataset may contain data from visible and invisible bands. An ML algorithm is expected to take some of or all of the available bands or extracted features and output classifications for each pixel or regions in the input area.  Generally, there are two approaches to applying ML to remote sensing. The first one is pixel-based classification and the second one is object-based classification. In pixel-based classification, the spectral band features of a single pixel are used to assign it a label. Object-based, however, uses both spatial and spectral features to output classifications for regions in the input. |
| **Semantic segmentation**  In semantic segmentation task, each pixel is assigned to one pre-defined class and pixels of the same class are grouped together to one semantic segment.    Source: https://uni-bonn.sciebo.de/s/kixZbCxTErCjW2Y |
| **Object-based Image Classification** Object-based classification uses both spectral and spatial information for classification. The process involves categorization of pixels based on their spectral characteristics, shape, texture and spatial relationship with the surrounding pixels. Object-based classification methods were developed relatively recently compared to traditional pixel-based classification techniques. While pixel-based classification is based solely on the spectral information in each pixel, object-based classification is based on information from a set of similar pixels called objects or image objects. Object-based classification is a two-step process, first the image is segmented or broken into discrete objects or features and then each object is classified. For example, in object-based classification, we might be interested in classifying buildings and football fields from remote sensing images. |
| **General Steps to Process EO Data**  In Modules 1 – 3, you studied that EO data contains multiple bands. In section 4.2 of this module, we discussed the key steps in ML. In this section, we will discuss general steps that can be used to select and process EO data. Please note these are general steps and some of them may not be relevant to some tasks.   1. Select best available image according to pre-defined thresholds 2. Select best features (channels, spatial) that describe the problem (classification, retrieval) 3. Remove noise and distortions due to clouds, acquisition (sun glint) or transmission (vertical stripes) 4. Use band operations to create band ratios, and indices through linear/non-linear combinations of existing bands 5. Pass the input (along with the extracted features) and use an ML algorithm to assign semantic classes to objects (pixels, patches, regions) in the scene   All the steps above are in some way related to feature extraction from EO data. Let’s see why feature extraction is important for EO data. |
| **Importance of Feature Extraction from EO Data**  Extracting features from remote sensing images is essential to:   * Make image processing algorithms more robust (to noise) * Visualize data characteristics * Understand the underlying physical relations   Specific feature extraction techniques will be discussed in subsequent modules. We will now discuss how the different types of Machine Learning can be applied to EO data at a high level. |
| **Supervised Learning for Image Classification in Remote Sensing**  Supervised learning is based on labeled training data consisting of a set of training samples (input-output pairs), it generally has two types: ***classification*** and***regression.***   * **Classification** isused when the output variables are categorical (i.e., with 2 or more classes). It helps to map the data from a certain dimensional feature space (number of bands) to one-dimensional space of the classes. The different classes are represented as integers. |
| * For example, if you need to separate two grass classes (dry and green) from information contained in two bands (refer to figure below), you can pass the band information as input to a machine learning model and produce either a green grass (class 1) or a dry grass (class 2) as output.       Source: Prof. Gabriele Cavallaro, School of Engineering and Natural Sciences, University of Iceland, Lecture 2 from the Course REI506M “Applications of Data Science in Remote Sensing” (Fall 2022), p.44 |
| Supervised classification has been applied to assign semantic labels to areas in a remote sensing dataset. |
| * **Regression** is defined as a statistical method used to establish the relationship between a dependent variable and one or more independent variables. In machine learning, Regression is a supervised learning technique since the dataset contains input features and the corresponding labels as well. Regression is used when the number of possible values for the dependent (target) variable are not countable. For example, if the target variable is a rational number there are infinitely many possibilities for what the target variable can be. In remote sensing regression can be used to estimate continuous quantitative metrics such as crop yield and building energy needs. Generally, regression has been applied to predict temperatures, stock prices, and credit scores. In the case of stock price prediction, for example, the target variable stock price can assume any positive value. Thus, we must apply regression methods to be able to output stock price that can assume any value in the valid range.   Regression vs. Classification  Source : <https://www.simplilearn.com/regression-vs-classification-in-machine-learning-article> |
| This part was a little bit more complex, wasn’t it? We know that it’s getting more abstract. To help you understand the concepts of regression and classification, please watch the following video:  <https://www.youtube.com/watch?v=TJveOYsK6MY> (time duration 2.47) |
| **Unsupervised learning**  In this case, data without labels is given. The classification algorithm is expected to find “hidden” patterns. Unsupervised learning methods generate clusters without semantic meaning. It is the task of the human to make sense of these clusters. Its classical frameworks include **clustering and dimensionality reduction.** Dimensionality reduction refers to techniques that reduce the number of input variables in a dataset.    Source: Remote Sensing Image Classification: A Comprehensive Review and Applications by Maryam Mehmood et al (March,20220), <https://doi.org/10.1155/2022/5880959> |
| Please watch this video to learn more on reinforcement learning:  <https://www.youtube.com/watch?v=2xATEwcRpy8> (time duration 2:27) |
| **Deep Learning in Remote Sensing**   * Many deep learning approaches have found applications in remote sensing * Some of the most used deep learning models in remote sensing are convolutional neural networks, * One popular application of deep learning in remote sensing is target recognition. Target recognition is the identification of objects such as ships, cars, and planes that exist in an image. In the context of remote sensing, target recognition is very challenging. * Target recognition in remote sensing is hard because   + The objects appear much smaller in an RS video   + The objects appear in a complex neighboring environment * Deep learning methods can be used for pixel wise classification or regional classification. Moreover, deep learning methods can work with spatial features, spectral features or both. |

**Exercise materials and tasks**

**Quiz questions**

|  |
| --- |
|  |

As recap of this session, please complete below quiz on machine learning for image classification subject:

1. Which of the following is not an example of image classification?
   1. **Using ML to identify the edges of cats in an image**
   2. Using ML to Recognizing faces from images
   3. Using ML to Identify different bird species from their images
   4. **Extracting features (descriptors) from images**
2. Which of the following is true about the traditional image classification approaches and deep learning-based image classification approaches?
   1. **Traditional ML image classification approaches require heavy involvement of a human expert to get good performance**
   2. **Deep Learning-based image classification can work well without explicit human specified features**
   3. Even if deep learning image classification can work well without human intervention the state-of-the-art image classification systems use traditional ML image classification methods
   4. For a deep learning image classification approach the input images must not be pre-processed in any way.
3. Which of the following requires a regression approach?
   1. Classifying images of domestic animals
   2. Recognizing a speaker from a sound file
   3. **Determining the price of a house based on its location, and number of bathrooms**
   4. Showing a customized advertisement to a user on a website
4. Which of the following is not true about Deep Learning?
   1. Supervised deep learning approaches can extract good features themselves without the help of a human
   2. Deep learning approaches can work well on unstructured data
   3. **An Artificial Neural Network that contains two layers of three neurons each can be considered deep learning**
   4. **There are some deep learning approaches that do not use Artificial Neural Networks**
5. If you are given some satellite images and are asked to group parts of each image into three groups, what kind of learning would you have to use?
   1. **Supervised**
   2. **Unsupervised**
   3. Reinforcement