The amount of data (of varying quality) which is being generated every day grows tremendously in a majority of real-life domains, including, among others, medical imaging, text categorisation, computational biology, genomics and banking. Although it may appear quite beneficial at the first glance—more data could mean more possibilities of extracting and revealing useful underlying knowledge—handling extremely large and difficult data sets became a pivotal issue and it is one of the largest challenges faced by various research communities, especially in the era of *big data*. This big data revolution has notably affected many fields of science and engineering, including statistics, machine learning, parallel computing, and computer systems as a whole. Within our project, we will primarily focus on processing computer vision data, such as images and video sequences, but the developed methods will be applicable to other domains as well.

In a classic approach towards image analysis, two substantial phases may be distinguished, namely feature extraction and classification of the extracted feature vectors. The first phase consists in extracting some useful features that would make it possible to represent an image using a vector of numbers describing not the image as such, but an object that appears in the image. For example, in the case of a facial image, among such features we could have the length of the nose with relation to the distance between the eyes or the roundness of the chin. These features must be manually defined and carefully selected by the engineers and programmers, who develop specific algorithms to analyse the objects of a certain category. Subsequently, the extracted feature vectors are classified to a certain group of objects (e.g. for a facial image it may be decided whether a person is wearing glasses). For this purpose, the *classifiers* are used, which learn the classification rules based on sufficiently large and representative training sets (in the aforementioned example, it could mean hundreds of images presenting people wearing glasses against a similar number of individuals without the glasses). *Support vector machine* (SVM) is one of the most popular classifiers, which has been successfully applied to a variety of pattern recognition problems, being a natural choice when designing new solutions that require data classification.

Nowadays, we may witness a revolution in computer vision and machine learning, which is concerned with the emergence of *deep learning* algorithms. Actually, this term embraces a number of techniques that exploit multi-layer neural networks which are capable of learning both the features, as well as the classification rules. In the case of images and video sequences, deep convolutional neural networks have been developed – the initial bottom layers perform the feature extraction, while the top layers classify the extracted feature vectors. This makes it possible to avoid the tedious process of designing hand-crafted features that must be extracted from the visual data. On the other hand, it is non-trivial to elaborate appropriate architecture of deep convolutional networks and this problem is usually solved relying on intuition and experience of the research engineers.

SVMs are considered as better classifiers than artificial neural networks, hence potentially they could be employed to classify the features extracted by a CNN. However, such approach has not been deployed in practice due to some limitations of the SVM training. They are concerned with high time and memory complexity, as well as with the difficulties in tuning the parameters (including the kernel function) and selecting the useful features and samples from the training set. These problems are particularly important in training deep networks, which actually demand huge amounts of data to be learned correctly. Our project is aimed at enduring SVMs against these problems relying on evolutionary algorithms, which allow for finding the best (or almost the best) solution using the mechanisms resembling biological evolution. As a result, this will make the SVMs applicable to learn from difficult data, including feature vectors extracted with CNNs – our goal is to improve the latter and decrease the influence of their architecture on the retrieved scores.

Achieving the project goals will make it possible to couple the advantages of deep convolutional neural networks with the high classification capacities of SVMs. Thanks to adequate exploitation of evolutionary algorithms, large amounts of data will no longer form an obstacle in applying SVMs, but on the contrary—they will be advantageous in obtaining higher classification results. During the project, we will validate the introduced solutions in a number of computer vision scenarios, including face and gesture recognition, emotion analysis from facial videos, as well as medical image analysis for cancer diagnostics.