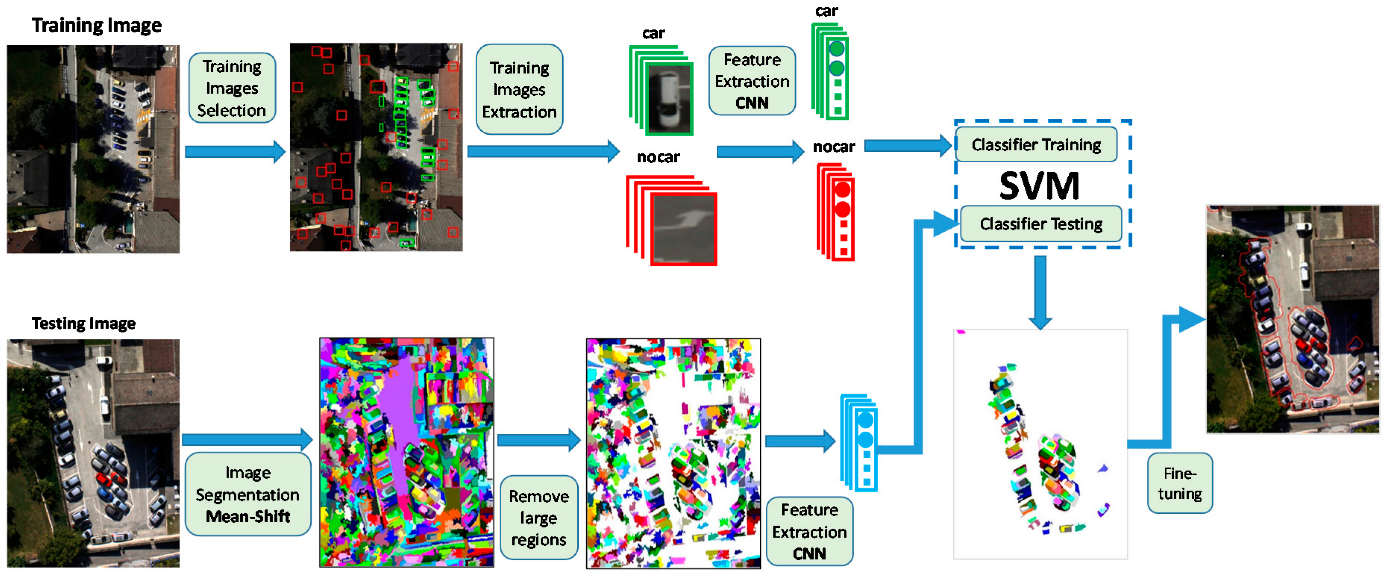
L'apprentissage profond prend beaucoup plus de temps que les classificateurs tels que XGBoost, RF, SVM, etc. dans ce cas, cette approche consiste à extraire (**Sélection de variables**) avec une technique comme ( CNN, The feature importances from Random forest) puis à ajouter un classificateur.

J’ai trouvé quelques travaux, propose la même technique.



<https://towardsdatascience.com/cnn-application-on-structured-data-automated-feature-extraction-8f2cd28d9a7e>

**Article : A Feature Selection Method for Multi-Label Text Based on Feature Importance**

*Currently, there are some researches on feature selection for multi-label data, which can be divided into three categories: wrapped methods, embedded methods, and filter methods. In wrapped methods, several different feature subsets are constructed in advance, the pros and cons of the feature subsets are then evaluated by the predictive precision of the classification algorithm, and the final feature subset is determined based on the evaluation [9–13]. In embedded methods, the feature selection process is integrated into the classification model training process, that is, features with high contribution to model training are selected to construct the feature subset in the process of classification model construction [14–16]. In filter methods, the classification contribution of each feature is calculated as its feature importance, and features with higher importance are selected to construct the feature subset [17–20]. Based on this selected feature subset, the training of the classification model is performed. Therefore, this type of feature selection methods has low computational complexity and high operational efficiency, and is very suitable for text data [21].*

**Feature ranking for multi-label classification using predictive clustering trees**

*In this work, we present a feature ranking method for multilabel data. The method is motivated by the the practically relevant multilabel applications, such as semantic annotation of images and videos, functional genomics, music and text categorization etc. We propose a feature ranking method based on random forests.*