Title: Explainable Artificial Intelligence for Developing Smart Cities Solutions

Paper link: https://www.mdpi.com/2624-6511/3/4/65

Traditional Artificial Intelligence (AI) technologies used in developing smart cities solutions rely more on best representative training datasets and features engineering and less on the available domain expertise. This makes the outcome of solutions less explainable. In this work, we propose a new solution to the problem of flood monitoring using Semantic Web technologies, which integrates deep learning (DL) and semantic rules (designed with close consultation with experts) to build a hybrid classifier. This hybrid classifier has on average 11% improvement in image classification performance.

1. Introduction

In 2017 there were more than 250 smart cities projects in 178 cities worldwide, and Machine Learning and Deep Learning techniques were used to solve various smart cities problems.

Over the years, numerous ML algorithms have been applied for image classification tasks. The performance of intelligent systems is matching or even bettering human intelligence. Most ML models are data-driven, and are developed by applying iterative training, evaluation and fine-tuning based on datasets. These models are highly non-transparent, and appear as a black box to the end-users. Human understanding, interpretation and explanation on decision making is crucial on complex tasks, such as many smart city applications, as the fundamental tenant of democratic smart cities is for the policymakers to explain decisions made by smart solutions powering public services in their cities.

The loss of control over interpretability of decision making is becoming a serious concern for high impact problems, and end-users expect explanations to assure confidence and trust over the system. Explainable

AI enhances the trust on decision making, as has been the case for medical education, research and clinical decision making.

This paper is from our experience of developing a flood monitoring application for a consortium of European cities using image classification with deep learning. We have built a novel image classification approach based on DL (Convolutional Neural Network: CNN) with an IoT-enabled camera to monitor gullies and drainages. However, the model lacked transparency in terms of how objects were related to each other and how the model was classifying images into different classes. To address the 'explainability' deficit of CNN models, we propose to use Semantic Web technologies, in particular ontologies and reasoners. These technologies can be used to capture the relationships between concepts and entities in a particular domain for better reasoning and explanation.

In this research, we propose a hybrid model that combines both machine learning and semantic rules set out to analyse the inclusion of expert knowledge in image classification. This model can be used more transparently and effectively in the context of real-time applications.

Section 2 surveys the literature on the use of ML and Semantic Web technologies in building smart cities solutions, and Section 3 provides context to this work.

2. Literature Review

Smart Cities integrate critical infrastructure and services through sensor and IoT devices, and use the rational method to design and plan any decision making more systematically. Human experts often make the decision based on data coming to any application.

In smart cities, data from sensors and IoT based applications is stored and processed to support decision making. This data is combined with qualitative and quantitative data for decision making to meet some objectives.

The use of semantic Web technology in combination with data analysis has been used for expert-based recommendation systems in the smart city. The semantic techniques give the flexibility to apply human experts' knowledge and control over the prediction model.

For companies and service providers, explaining how their system works and why it works well can enhance the trust of end-users in the system. For society, considering the possible impact of AI in terms of increased inequality and unethical behaviours is important.

Flood monitoring is a major concern in most of the cities around the world. While improved access to rainfall data, water level reading, satellite imagery and improved forecasting accuracy has been applied, real-time monitoring to support decision making is still a challenge.

Object analysis has been used for analysing image-based decision-making applications, scene recognition, and retrieval of a specified image from the library. In multi-object scenarios, multi-class label approaches are applied to classify such images, and ontology-based approaches are also used to retrieve a specified image from the library.

Machine learning algorithms have been widely used for image classification, and explainability has been applied in recent years to allow control and understanding of a machine learning model. Stream data mining algorithms can incrementally adapt to the non-stationary changes and data.

Adding explainability to a model allows to verify the system, understand the weakness and detect the biases of the model, and to modify the model with new insights. A visualisation tool could be applied to understand the structure of the model.

To design an explainable model, two major technical challenges need to be highlighted: (1) accurately extract features from noisy and sparse data into the model; and (2) generate easy and understandable explanation from the multilevel model structure.

The explainable system links human context-based reasoning with facts to construct contextual explanatory models, and applies human capabilities and understanding to make any decision or action. It is used in multi-domain operations to strengthen confidence, knowledge representation and reasoning.

A twin system consisting of a black-box built from the machine-learning method, and a white-box built from human knowledge has been used for the interpretation of chronic renal disease. A hybrid human-in-the-loop approach has been applied where machine learning is improved using extracted domain expert knowledge.

A combination of logic-based approach and probabilistic machine learning approach is required to build context-adaptive systems. A user-centric explainable decision support system was applied, where the system linked the human reasoning process with intelligent explainable techniques.

Convolution Neural network (CNN) models had been applied in image classification tasks. Multiple CNN models could be designed based on the task, and the output of each model was combined.

CNN achieved remarkably higher accuracy on many image analysis applications, but the network was heavily depended on the number of data that were used for the training. Different image augmentation techniques were applied, such as geometric transformations, feature space augmentation, colour space augmentation and random erasing.

Semantic technologies have been used to integrate multiple heterogeneous IoT devices for data monitoring real-time events and reasoning to support intelligent systems in smart city environments.

Ontologies play a key role in semantic web applications, and allow researchers to share and reuse domain knowledge. The ontology-based semantic approach has improved the interoperability between the applications.

Several projects have used semantics and ontologies within smart cities to add value to data collected from sensor and social data streams, to combine sensor and social data streams with machine learning techniques and to facilitate interoperability and information exchange.

Over the past few years, a variety of ontologies have been developed for use within smart cities environments for IoT, sensors, actuators and sensor observations, with the SSN ontology being one of the most commonly extended and adapted ones. However, the application of deep learning and semantic web in disaster response has been limited.

3. Flood Monitoring in Smart Cities

Flood monitoring has been identified as one of the major issues in smart cities. Real-time capturing of drainage and gully images using a smart camera and analysing and classifying the image can detect the potential flooding threat.

3.1. Major Objects and Their Significance

To learn about the major objects causing drainage and gully blockages, a workshop was organised with five experts. Four major objects were identified as the most common objects on monitoring drainage and gully blockage.

Leaves are one of the most prevalent problems when it comes to blockages in drainage systems.

Figure 2. Sample image for object: Leaves.

Silt is a solid, dust-like sediment that water, ice and wind transport and deposit. If not sufficiently cleaned regularly, it can cause problems with drainage and gully blockage.

Figure 3. Sample image for object: Mud.

Plastic and bottles were identified as a major risk to drainage system due to their capability of blocking the drainage and restricting the water flow into the sewage system.

3.2. Convolutional Neural Network for Object Coverage Detection

In this flood monitoring application, CNN models are used to detect object coverage level within drainage and gully images. This results in efficient image classification.

3.3. Semantics for Flood Monitoring

Semantic techniques enable understanding the characteristics of objects and the context of these objects, with the use of explicit formal rules. Expert knowledge adds the explainability of the system for decision making, and is articulated with semantic representations and formulation of semantic rules.

We propose a hybrid image classification model that uses machine learning and semantic techniques to classify drainage and gully images into a class label. The classification method consists of three computational steps.

4.1. Object Coverage Detection

Drainage and gully may get blocked with litter materials such as leaves, mud, plastics and bottles. To strengthen the classification decision, the level of coverages of the detected object within an image is used.

4.2. Semantic Representation and Rule Base Formulation

Semantic rules for image classification are defined based on expert knowledge captured during the workshop. Five experts were used to classify single images and the majority count approach was applied for knowledge extraction.

4.3. Inferencing and Image Classification

Inferencing is applied to classify images based on object coverage and semantic rules.

5. Methodology

In this section, we present a hybrid image classification model that uses machine learning and semantic rules for object coverage detection.

5.1. Data Construction

To prepare data sets for object detector, images were collected from publicly available image sources and manually pre-processed to remove unwanted, noisy and blurred images. Then, the image dataset for each object type was prepared.

5.2. Image Augmentation

To build an effective CNN model with higher accuracy, lower training and validation loss, a larger training dataset is required. Image augmentation has proven to be a powerful technique to enhance model performance.

Image augmentation techniques such as geometric transformation, random erasing, colour space transformations and feature space augmentation have been applied to increase the training image dataset. The ImageDataGenerator class from Keras library has been used as an alternative for image augmentation.

5.3. Image Annotation and Coverage Level

The object count method is not viewed as a feasible option for object coverage detection, as small objects appearing in a group and water and mud cannot be counted in discrete numbers. Additionally, the object count method does not appropriately address the coverage area proportion.

To overcome the complexities of image annotation, boxes were created by covering the objects in the image. The boxes covered some proportion of the total area within the image.

Figure 8. Screenshot of "Image Annotation".

The next challenge was to categorise the annotated image into different coverage levels. The coverage area percentage was applied to calculate the coverage level, and the annotated images were categorised into four coverage class levels.

5.4. Coverage Detector Implementation

The convolution layer is the first layer in a CNN model and uses a convolution operation on the input image to extract features. The number of convolution layers was adjusted based on the model training accuracy.

Pooling layers are down-sampling layers applied after the convolution layer to minimise the spatial resolution of the feature maps.

Four CNN models are designed by altering the number of layers, and a Softmax layer is used to calculate the final probabilities of each class. ReLU function was applied as the activation for the CNN models.

Regularisation modifies the model's learning parameters to improve the performance of the model during models training. L2 (lambda) regularisation parameter value is set as 0.001.

Dropout: random removal of some of the hidden nodes during the training. Image Augmentation: high number of training image datasets generated using data augmentation.

5.5. Semantic Representation

A semantic representation of drainage and gully blockages was created after the analysis of an individual expert's view on object identification, classification and reasoning. A semantic rule-base was created based on the experts' reasoning in image classification into corresponding class labels.

5.6. Rule-Based Formulation

Experts have highlighted that a drainage system is classified as "Fully Blocked" when there are many objects that cover most of the image portion with the sign of severe restriction of water flow through the drainage system, or as "Partially Blocked" when there are fewer objects that block it. Mutually exclusive semantic rules are defined to classify image instances based on object coverage detection using experts knowledge of image classification.

6. Experimental Design and Result Analysis

The simulation was performed on a machine with Intel(R) Core(TM) i7-8750 HCPU @2.20 GHz processor and 15.5 GB (usable) of RAM running on Windows-10 64-bit operating system.

6.1. Object Coverage Detection Training

The model's training accuracy, training loss, validation accuracy and validation loss are analysed iteratively for each object coverage detectors. The model's training accuracy improves significantly, by up to 60 iterations, following which the accuracy and loss performance appears to stabilise.

6.2. Analysis of Semantic Rules Implementation

The inference engine generates class labels for test images and compares them with the class labels defined by experts for those test images. The matched rules are applied based on the object coverage detection and hence the classification decision.

The object coverage detector detects Mud level one, Leaf level zero, Plastic $ Bottle Level zero and Water Level zero and selects the rule: Image(?p1, ?c1) ... Fully_Blocked(?p1)

The object coverage detector detects that the third image has no objects (mud, leaves, plastic bottle and water) and selects the rule No_Blockage(?p1).

The CNN model detects the coverage level and selects the appropriate rule from the rule base for image classification.

6.3. Hybrid Class Performance Analysis

In general, image classifier models' accuracies have been evaluated in terms of correctly classifying the test image into corresponding class labels. In this proposed hybrid image classifier, the model performance was analysed in two stages.

The accuracy of object coverage level detection is crucial for the implementation of semantic rules for image classification.

Table 5. Combined confusion matrix of coverage detectors.

The confusion matrix showed that the object coverage detectors do not have uniform accuracy. Level Zero and level Three had higher accuracy than level One and level Two, because they had more coverage of the representative object during the model training.

The proposed hybrid image classifier has been compared with a machine learning-based classifier based on deep learning on the accuracy of correctly classifying Fully Blocked (FB) images as FB, Partially Blocked (PB) images as PB and No Blockage (NB) images as NB.

The hybrid image classifier improved performance compared to that of machine learning-based classifier. It had better accuracy on all the true positive classification, and its overall accuracy was 69.23%, which is an improvement of about 2% accuracy compared to the machine learning-based classifier.

In the hybrid model, the accuracy of a classifier depends on the accuracy of the object coverage detectors and the implementation of the semantic rules.

In the literature, explainable AI presents some challenges, such as being used as a selective decision making that focuses on explanations and background knowledge, a large amount of information and using case-specific decision making. We faced other challenges as well, such as creating a domain-specific ontology from scratch.

7. Conclusions and Future Work

In this work, we have made a case for Explainable AI with a hybrid image classification model that combines ontological representation of the domain including rules captured with the help of domain experts and a DL-based classifier. This hybrid model is applied in a real-world use case involving flood monitoring application. The accuracy of our proposed hybrid image classification model was improved in comparison to the machine-based image classifier.

The hybrid image classifier gives the flexibility to incorporate experts' knowledge in the classification process, which improves the classification accuracy and also shows which rules need to be revised.

Future work will enhance the accuracy of both the object coverage detectors and hybrid classifier, by adding a higher number of application-focused images for training, and by revisiting rules that have been identified as the main contributors to low accuracy.