

# Using Deep Learning for IoT-enabled Camera : A Use Case of Flood Monitoring

Bhupesh Kumar Mishra  
University of Bradford, UK  
b.mishra@bradford.ac.uk

Dhaval Kumar Thakker  
University of Bradford, UK  
d.thakker@bradford.ac.uk

Suvodeep Mazumdar  
University of Sheffield, UK  
s.mazumdar@sheffield.ac.uk

Daniel Neagu  
University of Bradford, UK  
d.neagu@bradford.ac.uk

Sydney Simpson  
Bradford City Council, UK  
sydney.simpson@bradford.gov.uk

**Abstract**—In recent years, deep learning has been increasingly used for several applications such as object analysis, feature extraction and image classification. This paper explores the use of deep learning in a flood monitoring application in the context of an EC-funded project, Smart Cities and Open Data REuse (SCORE). IoT sensors for detecting blocked gullies and drainages are notoriously hard to build, hence we propose a novel technique to utilise deep learning for building an IoT-enabled smart camera to address this need. In our work, we apply deep learning to classify drain blockage images to develop an effective image classification model for different severities of blockages. Using this model, an image can be analysed and classified in number of classes depending upon the context of the image. In building such model, we explored the use of filtering in terms of segmentation as one of the approaches to increase the accuracy of classification by concentrating only into the area of interest within the image. Segmentation is applied in data pre-processing stage in our application before the training. We used crowdsourced publicly available images to train and test our model. Our model with segmentation showed an improvement in the classification accuracy.

**Index Terms**—Image Classification, Image Segmentation, Deep Learning, DCNN, IoT Sensor, Drain Blockage

## I. INTRODUCTION

Image classification is a part of image analysis where the images are labelled with a predefined class set. Deep Convolutional Neural Networks (DCNN) with many layers and several nodes in each layer have been widely applied for image classification [1-5]. The use of DCNN has been increasing in recent years for image classification tasks in many applications because of the unavailability of a large number of training data sets [2]. DCNN has been used extensively for object recognition in images with higher accuracy [6]. Deep learning-based methods delivered higher accuracy performance in many application areas as DCNN can extract characteristic features to recognise objects and hence categorise images into different classes. In deep learning, the number of training data sets play a central role for image classification tasks [7].

With the wider availability of large training data sets and implementation of GPUs, deep learning has been shown to

outperform some of the conventional methods of classification for image classification in many cases [1]. To achieve higher accuracy, image classification with deep learning has been trained with large number of images [8, 9]. In spite of the promising results of applying deep learning technique for image classification, image classifier trained with limited training data set introduces challenges. Without having enough training data, the accuracy of learning phase of the classification model can suffer with the problem of underfitting, ultimately leading to a degraded classification model [5]. Apart from the limited availability of data sets for training, the wider context, number and types of distinct objects within the images plays a crucial role in image classification [2, 8, 9]. Often, the same image can be classified into different classes based on the context and objects. We encountered the issue of limited image datasets, and identification of context in the use case of flood monitoring as part of the EC Interreg project called SCORE<sup>1</sup>. The project involves nine EU cities, and 3 EU universities to explore the use of IoT and big data technologies for solving smart cities challenges in water, environment, and mobility themes.

As part of the SCORE project, our research focuses on addressing challenges relevant to water theme, in particular, flood monitoring. The level of blockage of gullies and drainage is one of the many important parameters for local authorities and city councils to monitor in real-time. IoT sensors for detecting blocked gullies and drainages are notoriously hard to build. Hence as a replacement, we implemented a DCNN classifier for image classification with limited training image data set collected from Google images to classify drain blockage incidents. The main goal of this paper is to present a scalable approach to classify the images by exploring the context of images by applying image segmentation at data pre-processing stages. We applied image segmentation to localise the area of interest in the images before the images are utilised for the training and validation phases. The experimental results showed that the image classification using image segmentation at data pre-processing stage has higher accuracy in compari-

<sup>1</sup><https://northsearegion.eu/score/>

son of the classification accuracy of the image classification without image segmentation.

The rest of the paper is organised with section 2 reviewing image classification approaches. Section 3 describes the concept of deep learning, image segmentation and image classification whereas section 4 describes the experimental design, data generation and result analysis. We conclude and describe the future work in section 5.

## II. LITERATURE REVIEW

A number of methods have been applied for image classification tasks such as k-nearest-neighbours, maximum likelihood minimum distance, and logistic regression [10]. Most of these methods suffer from the “curse of dimensionality” [11] as there occurs a number of dimensions within the image data set. To deal with high dimensionality, methods such as Transformation [11] and Band selection [12] have been applied. In general, reduction of the dimension of the image data set is performed at initial stage to improve the classification performance. Reduction in dimensionality to fit the input image data into small-scale feature extraction has been explored using techniques such as principal component analysis (PCA) [13], balanced local discriminant embedding (BLDE) [14], pairwise constraint discriminant analysis and nonnegative sparse divergence (PCDA-NSD) [15]. Also, approaches such as pixel-based, object-based analysis have also been implemented for image classifications problem [16] over many years.

In recent years, intelligent system approaches such as neural network models have also been used to improve performance on image classification tasks [1, 2, 14, 17]. Deep Belief Network (DBN) has been used for spectral-spatial classification of hyperspectral data [14] by applying a hybrid framework with principal component analysis hierarchical learning-based feature extraction, and logistic regression. The DBN model has used single layer restricted Boltzmann machine and multilayer deep network-based models to learn the shallow and deep features of hyperspectral data, correspondingly. In the DBN model, linear regression was applied to classify images based on the extracted features. Residual learning framework was used to the training of networks [18] to increase training efficiency by explicitly fitting stacked layers to the desired underlying mapping instead of direct mapping

The accuracy of image classification tasks depends on the volume of data available for training and validation. Some models [8, 9] have used large number of images to train the classification model. While typical supervised models need large volumes of training data set to learn their model parameters. In the real world, specific problem-based image classifiers have been trained with only a limited number of sample images. Deep learning has been used with reasonably higher accuracy for image classification where large image data sets were not available for training [4, 19]. Simple image classification tasks had been solved with relatively

higher accuracy with small datasets by augmenting with label-preserving transformations.

Wide range of image classification approaches had been applied in many applications. Those models were mostly been implemented on the stored image classification. However, in real-time such as monitoring flooding where the camera acts as a sensor, the classification need to be applied on the streaming images. Captured image in such cases can be used as a sensor data which can be useful in many real-time monitoring problems. Deep learning is widely used image classification approach for image classification because of it has large number of nodes. Implementation of deep learning for image classification gives higher accuracy.

## III. DEEP LEARNING AND IMAGE SEGMENTATION- OUR APPROACH

### A. Deep Convolutional Neural Network (DCNN)

Neural Network (NN) is mostly designed with relatively fewer numbers of layers and nodes in each layer. Due to low number of layers and nodes, NN often has lower accuracy on image classification task. To overcome this limitation of NN, Deep Convolutional Neural Network (DCNN) with multiple layers of neurons have been used for image classification. Each neuron in DCNN extracts a different level of non-linear characteristic features from the image. DCNN was applied to classify image data by using large volumes (1.2 million high-resolution images) of training data in the ImageNet LSVRC-2010 contest into 1000 different classes [9]. Not only with large image data set, DCNN has also been applied with smaller numbers (few thousands) of training image data for classification [1]. DCNN was found to be a useful tool for image classification tasks.

### B. Contextual Image Analysis

To achieve higher accuracy, integration of additional information within an image needs to be analysed. Humans implicitly use structural knowledge to analyse an image during manual classification processes by considering contextual information along with the information about the shape of objects and spatial relations between the image regions [20]. Traditionally pixel-based classification methods have been used for image analysis in many applications while the integration of the concept of object-based analysis into the image classification process showed increment in the accuracy [16, 20]. Context analysis of an image varies with the application and classification tasks. Objects, environment and focused area of an image is crucial for feature extraction and hence the classification task. The importance of context analysis of images is illustrated in following image sets. Images Fig. 1 are the positive examples of the drainage blockages whereas images in Fig. 2 are the negative examples of the drainage blockages. From Fig. 1 and 2, it is observed that images of the drainage blockage and no drain blockages are analogous. To classify such images, context analysis is required to define their actual classes.

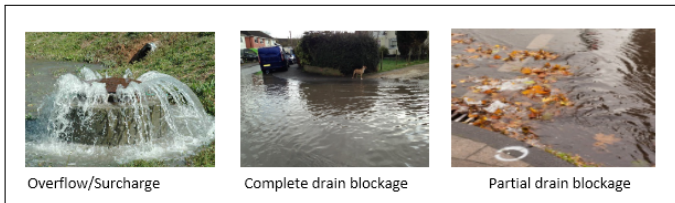


Fig. 1. Positive examples of drainage blockage



Fig. 2. Negative examples of drainage blockage

In general, images are captured focusing on the major area of concern at the centre. In other words, the objects and environment information have lesser importance as we move from the centre of an image. In fact, in supervised image capturing, monitoring and gathering, this is highly characteristic that occurred in most of the images. Considering this property, centre-focused image analysis for training can potentially be an effective approach for an image classification model.

### C. Image Segmentation

Segmentation of an image into a given number of non-overlapping regions provides freedom to select the chosen area of interest for the analysis. Segmentation helps in grouping the target objects or even focusing on single region target [16]. Segmenting an image into focused area of interest is one way to integrate the context information. For the focused region image analysis, degrees of freedom must be reduced to fulfil the assumed requirements [11]. With the case study on image classification of drain blockage, we analysed that most of the images have additional objects within images apart from drainage blockage features. With the limited low number of training image data set, identifying the individual context within the images is complex and may suffer poor performance on achieving classification accuracy.

Mostly, the nature of photography or capturing image of an event reflects that the centre of image has been focused as the main event. Considering this, we applied segmentation to locate central area of the image that reduce the impact of objects at the edges of the image, to analyse the images. In general, the representative characteristic features of an event are highly focused to the central region. The segmentation process with defined window size captures the focused area of interest leaving other objects in the images at the edges. In our model, pixel-wise segmentation has been used to crop the image edges. We applied edge removal from each side on images at data pre-processing stage by cropping 10% pixels

from all four sides of the image. Fig. 3.a is an example of an image collected in its original shape whereas Fig. 3.b is the image after segmentation. In the resulting image, after segmentation, it is observed that there is no change in the main focused context (Drainage blockage in this context) however impact of the objects and environments at the image edges have been reduced. Observing these two images it is clearly visible that some of the extra objects or context are removed or minimised applying segmentation.



Fig. 3. a: Image without segmentation b: Image after segmentation

## IV. EXPERIMENTAL DESIGN

### A. Data set

We used publicly available image as our data source for our case study data, retrieved using Google Image Search. The rationale for using this data source is to use images potentially uploaded by members of the public implying flooding and drain blockage events, as could be expected in realistic scenarios. Images of drainage blockage has been downloaded varying in types of drainage, colours, backgrounds, objects, blockage levels. The initially downloaded images data set was noisy containing irrelevant, blurred and out of context images. The data set was cleaned manually by visual inspection. After removal of noisy images from the downloaded data set, 800 images were prepared for the DCNN model. The images were classified categorically into three classes: fully blocked, partially blocked and no blockage as shown in Fig. 1. The images were further grouped into training data, validation data and test data in proportion 60%, 25% and 15% respectively.

### B. Deep Learning Model

We propose a deep learning-based image classifier as shown in Fig. 4 where unclassified image data are classified into classified image. The unclassified data can be classified into either in blockage class or partial blockage class or no blockage. Building block of the image classifier is shown in Fig. 5. Image segmentation is applied at data pre-processing where images are cropped with 10% edge removal. At the second stage DCNN parameters are defined such as type of activation functions, number of classes, batch size, error rate, number iterations and others. Model is trained and validated with sample image data set after the models' parameters are defined. After training and validation, the model is tested with test images. For the improvement in classification accuracy, the

model is trained and validated iteratively by adjusting model's parameter such as error rate, batch size, number of iterations.

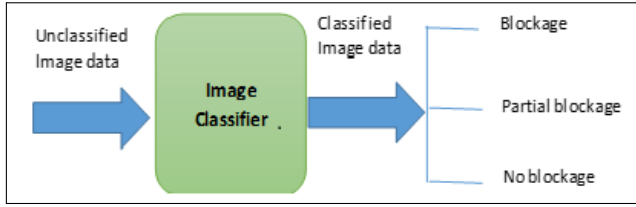


Fig. 4. Block diagram of Image Classifier

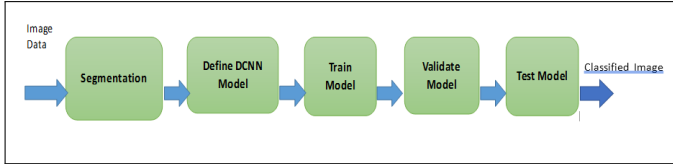


Fig. 5. Building blocks of Image Classifier

### C. Assumptions and Limitations

We apply following assumptions for the proposed image classification model. We also list some of the potential limitations of the model which we plan to explore further as future work.

#### Assumptions

- i. For model training and validation, classification of images into different class levels was based on human interpretation of the individual image.
- ii. We assumed that the all the gathered images had been captured as centrally focused image.
- iii. We have assumed that images obtained from the public domain via Google Image search are representative of what citizens would provide to local authorities. At a later stage of the project, we will adopt a similar approach to citizen provided data via our project partners.

#### Limitations

- i. The classifier can only classify drainage blockage images. Also, it can classify images into only three class levels though classifying images into a greater number of class levels gives higher level of interpretation of image analysis. We believe at this early stage of our project, these are the three primary factors local authorities are most interested in.
- ii. The classifier is limited to classify images only with jpeg format. We will explore the use of other formats in future work.

### D. Experimentation and Results

We analyse the efficiency of the proposed deep learning image classification with pre-processed segmented image.

TABLE I  
SIMULATION SETUP PARAMETERS

Model	VGG16
Class Mode	Categorical
Loss	categorical_crossentropy
Optimizer	optimizer_rmsprop(lr = 2e-5)
Steps per epoch	100
Epochs	100
Batch size	10
Validation steps	100

TABLE II  
RESULT SUMMARIES OF TRAINING, VALIDATION AND TEST OF THE MODEL.

Parameters	Accuracy without Segmentation	Accuracy with Segment.	Loss without Segmentation	Loss with Segmentation
Training	0.55	0.76	0.9	0.4
Validation	0.45	0.7	1.1	0.85
Test	0.6	0.66		

Simulations are run on an Intel(R) Xenon(R) CPU E5 -2620 v3 @ 2.40 GHz processor with 15.8 GB of RAM, 64-bit Operating System, X64-based processor with Windows 10 Enterprise. We used library functions included in the Keras core library with R programming language for the model simulation. We used VGG16, consisting 16 convolutional layers, in our classification model. Table 1 shows the simulations setup parameter. These parameters are fixed after applying different values such as batch sizes ranging from 5-15, steps per epochs and validation steps ranging from 50-100 during training and validation phases.

### E. Results Analysis

The classifier was trained and validated for image data sets without segmentation and with segmentation. Accuracy and loss for training and validation of classification model have been plotted for both the simulation tasks as shown in Fig. 6.a and 6.b. From the plots, we observed that there are improvements in the accuracy and loss on the proposed DCNN based image classification applying image segmentation at data pre-processing stage. The accuracy levels are increased by 21% and 25% for training and validation of the model respectively as listed in Table 2. There is a significant improvement in the loss value from 0.9 (model without segmentation) to 0.4 (model with segmentation) at training since lower the value of loss implies better is the classification model. Also, there is an improvement on loss value for validation of the model. These improvements are decisive on classifying images into their corresponding classes.

## V. CONCLUSIONS AND FUTURE WORKS

In our work as part of a flood monitoring use case, we applied DCNN for image classification task that classifies drainage images into either of three blockage classes based on its' severity: fully blocked, partially blocked or no blockage. A

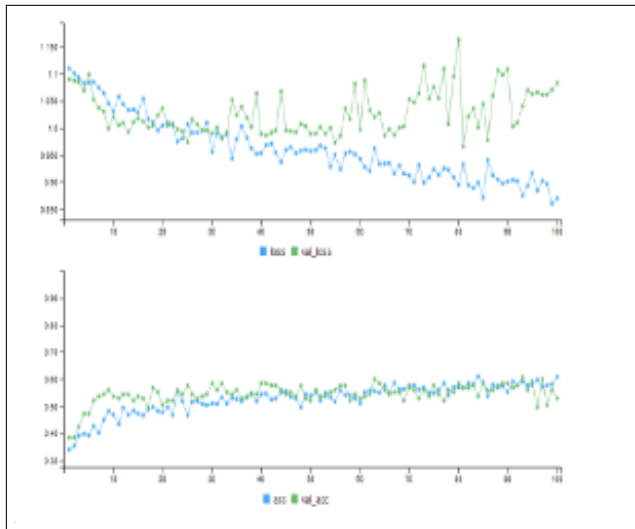


Fig. 6. Accuracy and loss plot on training for model training and validation with images without segmentation.

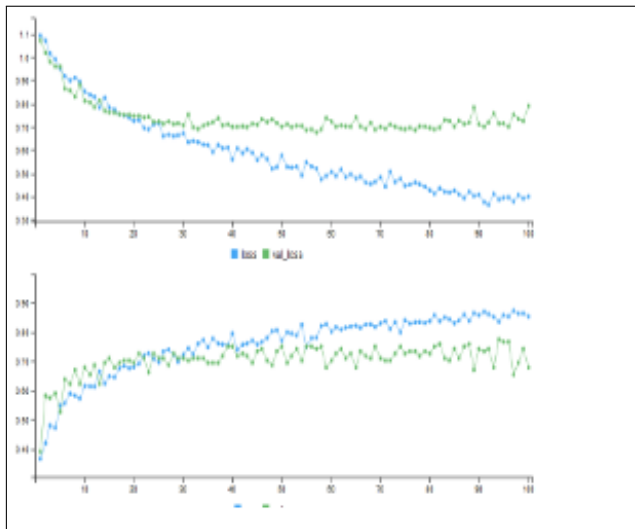


Fig. 7. Accuracy and loss plot on training for model training and validation with images with segmentation.

finer analysis of the level of blockage will be explored further – it was deemed that at this early stage of the project, the three identified classes would serve as a sufficient starting point, particularly as a proof of concept. We applied image segmentation at data pre-processing phase that localised the context of the images to improve the accuracy in image classification. The localisation of the area of interest in the image was applied to all images using segmentation. Pixel-wise 10% edges from all sides were removed from all the images at data pre-processing stage. These reduced image data were used for the training and validation of model to focus into central context of the images.

The experimental results indicated that the image classification using image segmentation at data pre-processing stage has higher accuracy in comparison of the classification accuracy

of the image classification without image segmentation. Our model was trained with a low number of publicly-available drain blockage images. As part of future work, more images with diversity in context, environment, types of objects within the images will be applied. The context and identification of types of objects within the image will help to make more accurate classification of the images. The improved classification model will be utilised in an IoT-enabled camera and will act as a real-time monitoring sensor.

#### ACKNOWLEDGMENT

The research presented in this paper is funded by the European Commission Interreg project Smart Cities and Open Data REuse (SCORE).

#### REFERENCES

- [1] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep Learning-Based Classification of Hyperspectral Data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, pp. 2094-2107, 2014.
- [2] H. Lee and H. Kwon, "Going Deeper With Contextual CNN for Hyperspectral Image Classification," *IEEE Transactions on Image Processing*, vol. 26, no. 10, pp. 4843-4855, 2017.
- [3] P. Liu, H. Zhang, and K. B. Eom, "Active Deep Learning for Classification of Hyperspectral Images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 2, pp. 712-724, 2017.
- [4] W. Hu, Y. Huang, L. Wei, F. Zhang, and H. Li, "Deep Convolutional Neural Networks for Hyperspectral Image Classification," *Journal of Sensors*, p. 12, 2015, Art. no. 258619.
- [5] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, "Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 10, pp. 6232-6251, 2016.
- [6] Y. Bengio, A. Courville, and P. Vincent, "Representation Learning: A Review and New Perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798-1828, 2013.
- [7] G. E. Hinton and R. R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks," *Science*, vol. 313, no. 5786, p. 504, 2006.
- [8] J. Deng, W. Dong, R. Socher, L. Li, L. Kai, and F.-F. Li, "ImageNet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248-255.
- [9] I. S. A. Krizhevsky, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Proceedings of the Advances in Neural Information Processing Systems 25 (NIPS '12)*, pp. 1097-1105, 2012.
- [10] G. M. Foody and A. Mathur, "A relative evaluation of multiclass image classification by support vector machines," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 6, pp. 1335-1343, 2004.
- [11] L. M. Bruce, C. H. Koger, and L. Jiang, "Dimensionality reduction of hyperspectral data using discrete wavelet transform feature extraction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 10, pp. 2331-2338, 2002.
- [12] F. Samadzadegan, H. Hasani, and T. Schenk, "Simultaneous feature selection and SVM parameter determination in classification of hyperspectral imagery using Ant Colony Optimization," *Canadian Journal of Remote Sensing*, vol. 38, no. 2, pp. 139-156, 2012/01/01 2012.
- [13] Y. Su, Z. Jiu-Fen, Z. Jiu-Ling, and L. Qing-Zhen, "A method for image classification based on Kernel PCA," in *2008 International Conference on Machine Learning and Cybernetics*, 2008, vol. 2, pp. 718-722.
- [14] Y. Chen, X. Zhao, and X. Jia, "Spectral-Spatial Classification of Hyperspectral Data Based on Deep Belief Network," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 2381-2392, 2015.
- [15] X. Wang, Y. Kong, Y. Gao, and Y. Cheng, "Dimensionality Reduction for Hyperspectral Data Based on Pairwise Constraint Discriminative Analysis and Nonnegative Sparse Divergence," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 4, pp. 1552-1562, 2017.

- [16] T. Blaschke, "Object based image analysis for remote sensing," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, no. 1, pp. 2-16, 2010/01/01/ 2010.
- [17] G. Giacinto and F. Roli, "Design of effective neural network ensembles for image classification purposes," *Image and Vision Computing*, vol. 19, no. 9, pp. 699-707, 2001/08/01/ 2001.
- [18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770-778.
- [19] P. Zhong, Z. Gong, S. Li, and C. Schönlieb, "Learning to Diversify Deep Belief Networks for Hyperspectral Image Classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 6, pp. 3516-3530, 2017.
- [20] T. Blaschke, "Object-based contextual image classification built on image segmentation," in *IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, 2003, 2003, pp. 113-119.