Automatic Architectural Drawing Labelling Using Deep Convolutional Neural Network

Seyed Masoud Sajjadian[1], Mina Jafari[2] and Xin Chen[2]

[1] School of Engineering and the Built Environment, Edinburgh Napier University, Edinburgh, UK

[2] Intelligent Modelling and Analysis Group, School of Computer Science, University of Nottingham, UK

Abstract

Architectural designers and technologists are able to make an assessment on buildability, thermal and hygrothermal performance of design details. To process drawings, human vision segments, classifies and distinguishes the drawing objects on the basis of their knowledge. With the rapid advancement of Artificial Intelligence methods, vast opportunities become available for performing tasks that used to require human intelligence or assistance by humans. Image processing and analysis is one of these tasks that consists of the manipulation of images using algorithms. There are various applications in different fields, and the use of it is increasing exponentially. This paper explores the use of image processing in identifying building materials in order to check compliance with building regulations and identify anomalies. In this paper, an encoder-decoder based deep convolutional neural network (DRU-net) for image segmentation is applied on architectural images to segment various materials including insulations, bricks and concrete in the conceptual development phase. An experimental analysis is performed on numerous detail drawings and an evaluation is made by mathematical models.

Keywords: Image Processing, Architectural Detail Drawings, Deep Learning.

1. Introduction

Building regulations are an essential part of the UK government approach to protect health and well-being of users, conserve fuel and preserve environment. Building Control Bodies (BCBs) have the authority to ensure requirements are met in each construction project. However, variety of construction methods and complexity of drawings, unpredictable changes and pressure of time may cause an increase in human errors. No substantial studies exist (as of today) to adequately demonstrate the compliance level to regulations typically achieved in each project. However, numerous studies reported significant failures in complying with building regulations not only in the UK but also in other countries like Canada and Australia. They also reported complexity of regulations, lack of consistency and knowledge of detail drawings, adequate monitoring and time pressure, etc. are causes of non-compliance (1).

Architecture and construction have significantly benefited from a wide range of computer programs from modelling to management and analysis. However, the industry still suffers from the lack of substantial digitization like other sectors and is among the least digitized (2). The industry still relies on paper to manage most of it processes that includes design drawings, procurement, daily progress reports, etc. The lack of a common and integrated platform in projects and the loosely coupled systems of tools where each component has limited knowledge of the other are known to be the main reasons for the industry's insignificant productivity in digitization compared to other industries (3).

The rapid development of Artificial Intelligence and deep learning as a subfield of it could help the construction industry in various aspects. Generally, deep learning is inspired by the human brain and learns from large amounts of data. It consists of networks capable of learning and the related algorithms perform a task repeatedly to improve the outcome similarly to how humans learn from an experience (4). The networks work on the basis of a collection of nodes (artificial neurons) that model the neurons in a human brain as shown in Figure 1. There has been numerous advantages to this method in fault tolerance, ability to learn very complex issues and development potential (5), industrial engineering inspection processes, quality control and detection of contaminants, in building performance to determine comfort level (6) (7) and in medical sciences (8).

Image segmentation is a process of dividing an image into some categories of objects. Currently most of the segmentation methods depend on thresholding algorithms (9). Gunay et al. (10) conducted a study on detecting occupants' presence in offices using image processing algorithms and correctly interpreted 95% of the occupied period and 93% of the unoccupied period. Similar study is conducted by Benezeth et al. and they reported an accuracy of 97% (11) Further studies by Amin et al. on potential accuracy of people counting systems on low resolution images resulted an error of within 3% in eighteen experiments (12). Because of such reliability, in this paper, an image segmentation method based on deep learning is used to segment architectural images. In the following the method is described in detail.

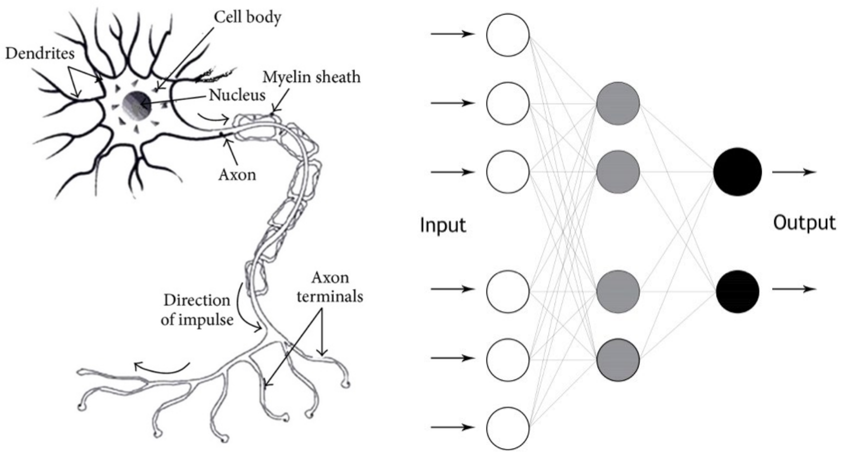


Figure 1: Structural of biological neuron, image on the left from (13)

2. Methods

In order to prove compliance with building codes it is vitally important to make a clear demonstration between the specification of the used materials and the relevant requirement from a building code (e.g how the design detail claim the required U-Value). The framework for the programme is adjusted as follows and shown in Figure 2:

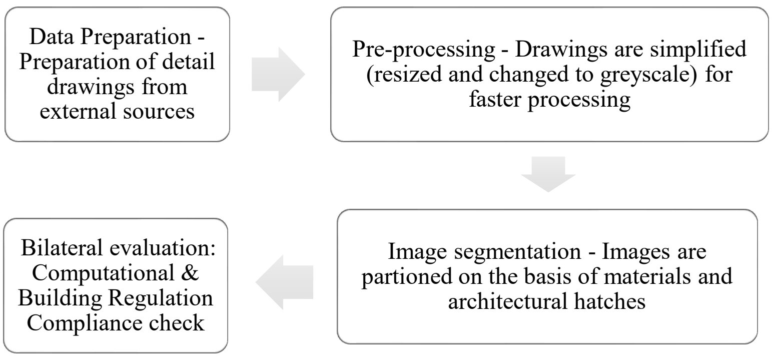


Figure 2: the overall scheme of our proposed method.

2.1 Data preparation:

Numerous detail drawings from Robust Detailing Limited, Accredited Construction Details (ACDs) and Building Research Establishment (BRE) in compliance with the energy efficiency requirements (Part L) of the Building Regulations were chosen to train the model. Each drawing is given a scale in order to determine the construction layer thicknesses. Each material in the drawing is given a colour that works as a reference code for material thermal specification. The colour can be mapped for a separate drawing by the model.

2.2. Image segmentation:

The goal of image segmentation is to label each pixel of an image. The output of the model is a high resolution image (the same size as input image) in which each pixel is classified to a particular class. Figure 3 presents some image examples and their corresponding ground truth labels.

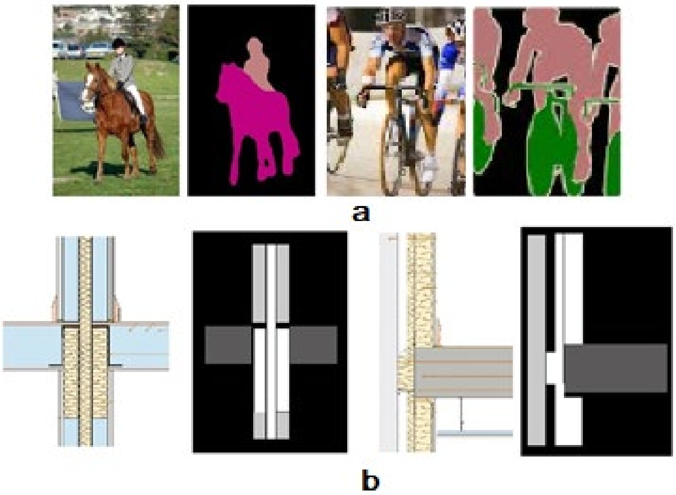


Figure 3: The images and the corresponding ground truth labels. a) natural images (14), b) architectural images used in this study.

The performance of computer vision systems has been significantly improved in a wide range of applications by the idea of deep convolutional neural network (DCNN) proposed by LeCun (15). DRU-net (16) is chosen to be used as the DCNN model in this paper. DRU-net is an encoder-decoder DCNN model for image segmentation. This network includes two paths, encoder and decoder. Encoder is used to extract the context and meaningful features from the input images, while the decoder enables an accurate localization by transforming these features into a segmentation map corresponding to the input image. The structure of the encoder consists of a stack of convolutional and max pooling layers, and the decoder includes convolutional and transposed convolutional layers. Figure 4 presents this network. For more details about the network, the readers are encouraged to read (16).

The input to this model is the architectural images and their corresponding ground truth labels, and the output is the segmentation results produced by the DRU-net as shown in Figure 4.

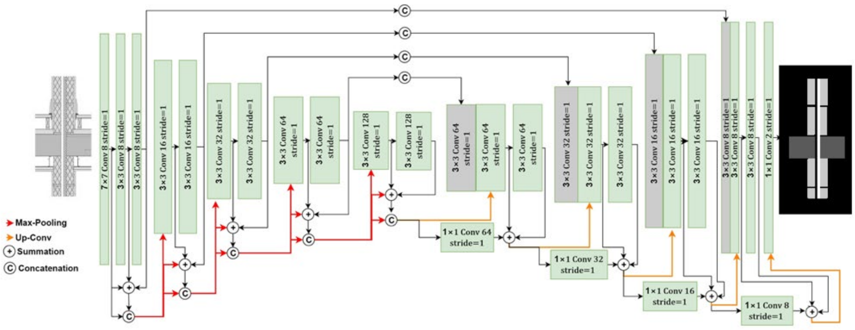


Figure 4: The structure of DRU-net (16).

2.3 Evaluation:

The accuracy of the proposed method needs to be evaluated to show how precise the proposed method is. The accuracy in image segmentation is measured based on the agreement between the estimated segmentation output and the ground truth segmentation mask. In this paper, the segmentation performance was assessed by using Precision, Recall, and Dice. These measures are often used to quantify the performance of image segmentation methods and reveal how similar the segmentation output of the method and the ground truth labels are. These measures are defined as follows:

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$Dice = \frac{2TP}{2TP + FP + FN} \tag{3}$$

where true positive (TP) represents pixels that are correctly predicted according to the target mask, whereas false positive (FP) represents pixels falsely segmented as foreground, and false negative (FN) shows pixels falsely detected as background. The higher value of these measures shows the higher performance of the methods. In other words, a higher value shows a higher agreement between the segmentation output of the model and the ground truth segmentation mask.

2.4 U-Value Measurement

U-value measurement is for a building element made of uniform and parallel layers. The heat flow is straight from inside to outside through such an element. To measure the U-value, a sum of the thermal resistances of each layer is required, however all building components have non-uniformities (layers that are not parallel or plane) which means that the heat does not travel straight through them. For the purpose of U-Values, in numerical methods, construction is always considered as uniform in one direction meaning that three dimensional effects do not influence the overall U-Value significantly (17). BS EN ISO6946 suggested calculation as follows (18) where R is material resistivity, d is the material thickness and K is thermal conductivity coefficient (K-Value):

$$R = \frac{d}{k} \tag{4}$$

$$U\ Value = \frac{1}{R1+R2+R3\ldots+Rn} \tag{5}$$

Furthermore, Approved document Part L1A sets out the concurrent notional dwelling specification as shown in Table 1, even though better fabric performance is likely to be required to achieve Target Emission rate (TER) and Target Fabric Energy Efficiency (TFEE) (19).

Table 1: Concurrent notional dwelling specification

Roof	0.13 W/m^2.K
Wall	0.18 W/m^2.K
Floor	0.13 W/m^2.K
Party wall	0.20 W/m^2.K
Windows	1.4 W/m^2.K

By deriving data from the drawings, calculations take place for a section on each drawings. U-Value calculator follows to the numerical method (BR 443) and make calculations based on conductivity rates and layer thickness. The method has the capacity to calculate Psi (Ψ) value too which is a measurement of heat loss (W/m.K) across a given junction between the external wall and another element. The material specification should be given manually to the developed program due to diversity of materials thermal characteristic range especially insulations (figure 5 demonstrates the thermal conductivity rates for most commonly used insulation materials), the algorithm only detects the material on the basis of architectural hatch, understand the scale and runs the calculations.

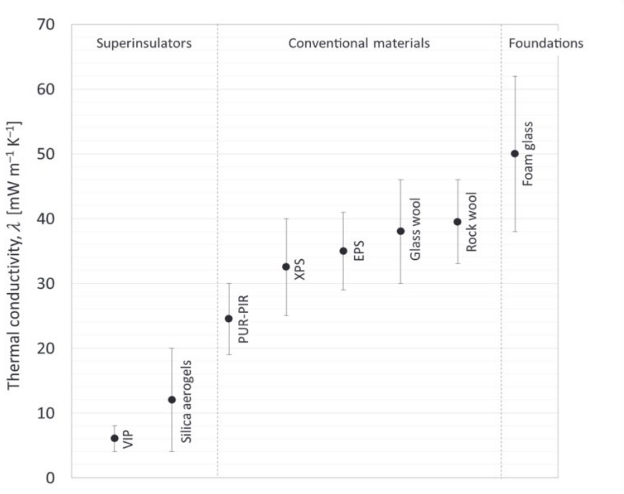


Fig. 5: Thermal conductivity of insulation materials. The dots represent average values and the bars indicate the available range. Image from (20)

3. Experimental results:

We trained the DRU-net on the pre-processed images and their corresponding ground truth labels of 2D architectural images. The datasets, training parameter settings and the experimental results are reported in this section.

3.1 Materials and datasets:

The model is trained to recognise standard hatching styles for construction materials. For the initial phases, insulation, bricks and concrete were used in the recognition. To train the model, we prepared 12 architectural images and their corresponding ground truth labels. The images were annotated manually by an expert. For all experiments, the dataset was randomly divided to 70% for training, 10% for validation and 20% for testing. Because the amount of data available is limited, 6-fold cross validation is used to reduce the sensitivity of the performance estimation to the partitioning of the data.

The deep-learning model run on a graphics-processing unit NVIDIA GeForce GTX 1080Ti. The code was implemented in Python based on Tensorflow 2.0, and the code is available at GitHub (https://github.com/MinaJf/DRU-net).

The model was trained for 100 epochs. The learning rate was set to 10^{-3}, and decayed by multiplying 0.8 for every 10 epochs. The initial feature channel number for the encoder-decoder architecture was 16, and the number of layers for both the encoder and decoder was 5.

3.2 Pre-processing:

The first step is to convert the image to grayscale image. In this process, a coloured two-dimensional image is transformed into a grayscale version of the image. Coloured images introducing unnecessary information which increase the amount of training time. Therefore, the main reason of this image conversion is that grayscale image simplifies the model training and reduces the computational time.

Besides, all the images were resized to 256 × 256 pixels. Another important pre-processing step is image normalization that rescales the pixel intensity values. In this work, all resized images were normalized to the range of -1 and 1 by using zero-mean normalization.

4. Results

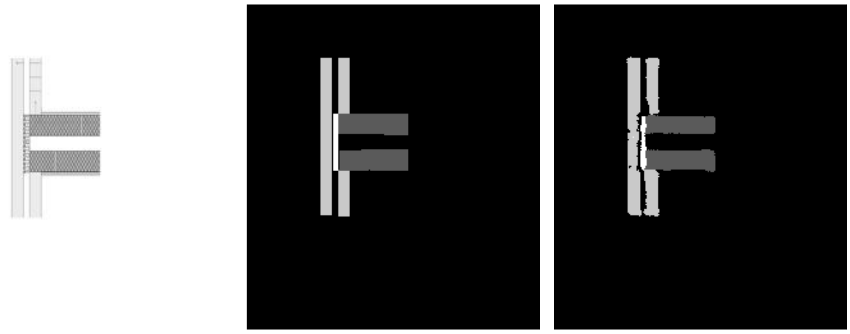
In this section, DRU-net is compared with U-net (21) and Residual U-net (RU-net) (22) in terms of accuracy. U-net and RU-net are two DCNN models that are chosen here to compare their performance with DRU-net. Dice, Precision, and Recall for U-net, RU-net, and DRU-net are reported in Table 2. It can be seen from Table 2 that DRU-net outperformed U-net and RU-net (a higher value indicates a better performance). U-net and RU-net could not segment some of the materials resulting in zeros for Dice, precision and recall measures, while DRU-net segmented all the classes.

Figure 6 shows one example of architectural images with its corresponding ground truth label, and the segmentation output of DRU-net. DRU-net produced a similar visual result to the ground truth label.

Table 2: Segmentation results based on U-net, RU-net, and DRU-net. The results are reported based on dice, Precision, and Recall for all testing images.

Method	Dice			Recall			Precision		
	C	B	I	C	B	I	C	B	I
U-net	0.634	0.545	0.178	0.755	0.587	0.294	0.560	0.530	0.225
RU-net	0.601	**0.668**	0.208	0.698	**0.835**	0.328	0.561	0.590	0.224
DRU-net	**0.754**	0.654	**0.227**	**0.863**	0.789	**0.444**	**0.678**	**0.620**	**0.335**

C: Concrete B: Brick I: Insulation



(a) (b) (c)

Figure 6: Visual results of one test example. a) Original image, b) ground truth label, c) the segmented output of the proposed method.

The level of accuracy achieved means the layer thickness for Brick and Concrete is detected correctly with very limited failure on some pixels. However, for insulation, the algorithm failed to achieve a high level of accuracy as the number of pixels for insulation layers is lower than the other two materials that makes the learning process more difficult for insulation segmentation.

Conclusion

Architectural detail drawings are of very different configurations and therefore human evaluations are prone to errors. In this study, a DCNN model is utilized to learn how to segment the most commonly used materials in construction on the basis of standard drawings and run the required calculations to check compliance with Part L1A of the UK building regulations. This article brings out a concept where image segmentation using deep learning can be considered in examining detail drawings. The framework development can address industry-wide problems in standard compliance, reduce human errors and deliver faster drawings check. The framework requires consistent and systematic review and feedback to improve the accuracy.

Our future research focus on improving the accuracy of the algorithm used in this study by balancing the focus of model learning on materials with lower pixels.

References

1. *Compliance with building regulations in England and Wales.* Baiche, Bousmaha, Walliman, Nicholas and Ogden, Raymond. 2006, Structural Survey , Vol. 24, pp. 279-299.

2. McKinsey & Company. *Imagining construction's digital future.* 2016.

3. *Paving the way to digital innovation: megaprojects, institutions and agency.* Papadonikolaki, Eleni, Morgan, Bethan and Davies, Andrew. s.l. : 7 th International Megaprojects Workshop: Theory meets Practice - Artificial Intelligence and Megaprojects: New processes and value, 2018.

4. *Study on image processing using deep learning techniques.* Rao Karanam, Santoshachandra, Srinivas, Y and Vamshi Krishna, M. 2020. Materials Today .

5. *Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes.* Tu, Jack V. 11, 1996, Clinical Epidemiology, Vol. 49.

6. *An Expandable, Contextualized and Data-Driven Indoor Thermal Comfort Model.* Sajjadian, Seyed Masoud, Jafari, Mina and Pekaslan, Direnc. 4, s.l. : Energy and Built Environment, 2020, Vol. 1.

7. *An Artificial Intelligence Method for Comfort Level Prediction.* Sajjadian, SM, Jafari, Mina and Siebers, Peer-Olaf. s.l. : International Conference on Sustainability in Energy and Buildings, 2018.

8. *Industrial Image Processing Using Fuzzy-Logic.* Gheorghe Amza, Catalin and Cicic, Dumitru Titi. s.l. : Elsevier, 2015, Vol. 100, pp. 492-498.

9. *Iterative thresholding algorithms.* Fornasier, Massimo and Rauhut, Holger. 2, 2008, Applied and Computation Harmonice Analysis, Vol. 25, pp. 187-208.

10. *Detecting Occupant's presence in office spaces: a case study.* Gunay, H.B, et al. Hamilton : s.n., 2016. eSIM.

11. *Towards a sensor for detecting human presence and characterizing activity.* Benezeth, Y, et al. 2-3, 2011, Energy and Buildings, Vol. 43, pp. 305-314.

12. *Occupancy detection in the office by analyzing surveillance videos and its application to building energy conservation.* Amin, I.J, et al. 6, 2008, Measurement, Vol. 41, pp. 589-599.

13. *A Review for Detecting Gene-Gene Interactions Using MachineLearning Methods in Genetic Epidemiology.* Lee Koo, ching, et al. 2013, BioMed Research International.

14. *Fully convolutional networks for semantic segmentation.* Long, Jonathan , Shelhamer, Evan and Darrell, Trevor. 2015. Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 3431-3440.

15. *Gradient-based learning applied to document recognition.* LeCun, Yann , et al. 1998, Proceedings of the IEEE, pp. 2278-2324.

16. *DRU-Net: An Efficient Deep Convolutional Neural Network for Medical Image Segmentation.* Jafari, Mina, et al. 2020. 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI). pp. 1144-1148.

17. Anderson, Brian. *Conventions for U-value calculations.* Garston : BRE Scotland, 2006.

18. *BS EN ISO 6946:2007 building components and building elements - Thermal resistance and thermal transmittance- Calculation method.* London : British Standards Institution, 2007.

19. Ministry of Hosuing, Communitites & Local Government. *Approved Document L1A; Conservation of fuel and power.* Newcastle Upon Tyne : NBS, part of RIBA Enterprises Ltd, 2018.

20. *A review an evaluation of thermal insulation materials and methods for thermal energy storage systems.* Villasmil, Willy, Fischer, Ludgar and Worlitschek, Jorg. 2019, Renewable and Sustainable Energy Reviews, Vol. 103, pp. 71-84.

21. *U-net: Convolutional networks for biomedical image segmentation.* Ronneberger, Olaf , Fischer, Philipp and Brox, Thomas. 2015. International Conference on Medical image computing and computer-assisted intervention. pp. 234-241.

22. *Deep residual learning for image recognition.* He, Kaiming , et al. 2016. Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770-778.