An Approach for Brick Wall Quantity Take-Off by U-Net Method Based on Deep Learning

Hasan Basri BASAGA[1*]
Kemal HACIEFENDIOĞLU[2]

ABSTRACT

This study presents a deep learning-based method for the quantity take-off in the construction industry. In this context, the brick wall quantity calculation was performed automatically over two-dimensional (2D) pictures by the U-Net method. 280 photos were first taken in the field at different distances and angles. 1960 images were, then, obtained by augmentation to increase the training accuracy. Pixel calculation of the automatically masked area in the images was made for wall estimation. The wall area was calculated by comparing this pixel value with that of the reference brick surface area. The method was tested on four sample photos including different wall images. A parametric study was carried out to reduce the errors. In the study, it has been shown that the proposed method is suitable for brick quantity calculation. In addition, it was concluded that the photo should be taken as close as possible, and more than one brick should be taken as a reference in close-up photos to increase the accuracy.

Keywords: Deep learning, fully convolutional network, U-Net, brick wall, quantity take-off.

1. INTRODUCTION

The advancement of technology has brought digitalization in every field. Especially with the introduction of web-based applications into business life, topics such as automation and big data processing have become even more important. The engineering and construction (AEC) sector closely follows these developments and integrates new ideas inside every day. Since the time factor is very important in the sector, the studies related to fast data analysis, instant job follow-up, instant evaluation of the job, fast communication and fast solutions, etc. have attracted more attention in the sector. In this context, Building Information Modeling (BIM)

technology, which is a three-dimensional (3D) information sharing process that can be used jointly by those involved in the design, construction, and maintenance of different projects, has also accelerated the sector. In addition, the development of artificial intelligence and the creation of solutions for various business areas have been important steps for the construction industry.

Computer vision (CV), a subset of Artificial Intelligence (AI), has revolutionized the analysis of visual data, encompassing images and videos, with the aim of emulating human visual comprehension. This review delves into the progression of CV techniques, tracing their evolution from conventional methods to the transformative influence of deep learning, while emphasizing their application domains and implications for the construction and asset management phases.

Historically, CV relied on manual efforts to design rules-based detectors and handcrafted feature descriptors to detect and classify objects within images [1]. These techniques encompassed feature detectors such as edge detection, corner detection, and blob detection, and feature descriptors including Scale Invariant Feature Transform (SIFT) [2] and Histogram of Oriented Gradients (HOG) [3]. These methods employed fixed-size rectangular regions sliding across images to apply descriptors and classifiers like Support Vector Machines (SVM) [4], though they proved inflexible and demanded expert engineering to design effective descriptors [5].

Advancements in machine learning, particularly deep learning, have ushered in a new era for CV. Deep learning enables the creation of Convolutional Neural Networks (CNNs), intricate networks acting as feature extractors that learn from input data, rendering them versatile for various classification tasks [6]. Deep learning fosters "end-to-end" learning, where CNNs autonomously learn and extract features from input images, culminating in enhanced accuracy and robustness [7].

Conventional CV methods and deep learning often synergize to maximize performance. In certain scenarios, techniques like sliding windows and SIFT are employed to identify regions of interest, followed by deep learning models for efficient processing. Techniques like Principle Component Analysis (PCA) are adopted to minimize feature dimensions and prevent model overfitting. The dynamic interplay between traditional and deep learning-based methods enhances outcomes across domains, such as image processing, object classification, object detection, pose estimation, and 3D reconstruction.

CV technology has gained significant attention within the construction industry due to its potential to revolutionize various aspects of project management, safety, quality control, and overall efficiency. Over the past decade, numerous studies have explored the application of CV methods in construction-related tasks. This literature review aims to provide an overview of the key contributions in this field, focusing on safety management, progress monitoring, productivity tracking, and quality control.

In safety management area, Suman and Paneru [8] presented a comprehensive review of CV applications in construction, with a specific emphasis on safety management. They highlighted the current state, opportunities, and challenges in this area, shedding light on the potential of CV to enhance safety practices. Meanwhile, Brian et al. [9] critically examined the integration of CV technologies for safety science and management in construction. Their review underscored the importance of these technologies and proposed future research

directions to advance safety practices. Moreover, Wu et al. [10] introduced a novel conceptual framework that combined CV and ontology techniques to improve on-site safety management in construction. Their study showcased the potential of semantic reasoning and visual data for enhancing safety protocols. Fang et al. [11] extended this discussion by presenting a framework that integrated deep learning and CV to ensure safety assurance on construction sites. Their work provided insights into the practical implementation of digital technologies to enhance safety performance.

In progress monitoring and productivity tracking, Xu et al. [12] critically reviewed state-of-the-art CV techniques employed for progress monitoring and productivity tracking in construction projects. By analyzing recent advancements, they highlighted the potential of these methods to provide real-time insights into project development. Notably, the potential of CV technology extends beyond individual construction phases. Xu et al. [13] conducted a comprehensive review encompassing the construction, operation, and maintenance phases of civil assets. Their critical assessment underscored the versatility of CV techniques in addressing challenges throughout the lifecycle of civil infrastructure.

Conclusively, these research papers collectively highlight the evolving prospect of CV applications in construction. From safety management to quality control, progress monitoring, and even lifecycle management, the integration of CV technologies promises transformative benefits. However, challenges and opportunities remain, suggesting a fertile ground for further research and innovation in this rapidly advancing field.

This study attempts the quantity calculation needed in all areas of the construction industry automatically by using the deep learning method. Since there is no study in the literature in which quantity calculations are made using deep learning, studies on deep learning applications in civil engineering are only summarized below. Since research on deep learning was carried out in various sub-branches of civil engineering, a limited number of studies are included here for information purposes only. Yang et al. [14], Fan et al. [15], and Wang et al. [16] developed deep learning-based methods for counting rebars in stock using images. These three studies are the closest ones to the subject considered in this work. Quantity calculation studies have been generally made on numerical data. For example, Akanbi et al. [17] developed deep learning models to predict the amount (in tonnes) of scrap and waste materials that could be recovered from buildings before demolition.

Many studies have also been carried out within the scope of occupational safety. These studies generally aimed to detect personal protective equipment or to detect situations that are not suitable for occupational health and safety in the field. Some examples are the determination of whether the helmet is worn [18,19], the joint control of helmet and vest [20], the suitability of the body ergonomic posture of the employees [21,22], and checking if there is a safety guardrail [23].

Another issue emphasized in the AEC industry is crack detection. To detect cracks in concrete surfaces, Pan et al. [24] developed a deep learning-based spatial channel hierarchical network. Yang et al. [25] proposed a new transfer learning method to detect concrete cracks in superstructures. Kang et al. [26] developed a hybrid method producing a pixel-based solution to measure crack thickness and length. In addition to method development studies, crack detection was also carried out in different structures and materials. For example,

structural cracks [27], cracks in buildings [28], road cracks [29,30], shield tunneling [31], cracks inside steel box girders of bridges [32], concrete highway bridges [33], etc.

The deep learning method has been used to produce solutions in different areas of the construction industry. 3D models classification of Building Information Modelling [34,35], construction cost prediction [36,37], building energy system behaviour prediction [38], long term electricity and heating load prediction [39,40], landslide detection [41], determination of earthquake-induced ground failure effects [42] are some examples. There are also literature reviews in which the research work on deep learning and machine learning in the construction sector is given [43,44].

2. AIM OF THE RESEARCH

The aim of this study is to present a method that uses deep learning and image processing techniques to automate the brick wall quantity take-off process from photographs. The proposed method offers several advantages. It eliminates the need for manual measurement, reduces the potential for human error, and increases the speed and accuracy of the take-off process. Furthermore, it can be used for progress control, as it enables automatic tracking of the construction process by comparing the actual construction progress with the planned progress based on the estimated brick quantity. The proposed method can also be integrated with Building Information Modeling (BIM) processes, which can improve the overall efficiency and accuracy of the construction project. Therefore, the objective of this study is to demonstrate the feasibility and effectiveness of using deep learning and image processing techniques for automating the brick wall quantity take-off process and its potential for integration with BIM processes and progress control.

3. U-NET ARCHITECTURE

The fully convolutional network [45] is an architecture for semantic segmentation. In the segmentation process, the image is divided into multiple parts to simplify the representation of the image. ResNet, SegNet, and U-Net are the architectures that are commonly used for fully convolutional networks. In this study, U-Net [46] is used. The architecture of U-Net is built based on a fully convolutional network for pixel-based classification, which is not an image (object) classification. The main difference between U-Net and traditional convolutional neural networks is that the former has the segmented image output, the latter has the image class output. U-Net is like a U-shaped model and has two parts. The first part is the encoder and the second part is the decoder. In the encoder part, by decreasing the dimension, it is progressed downwards. There are two convolution layers which are the max-pooling and dropout layers in each level.

The original U-Net architecture is designed for images with the size of 572×572×3. However, as the input size of the network increases, much more processing power is needed, so it was preferred to train a smaller network model. Input image which is 128×128 pixels was decreased by half at each level, as a result, it becomes 8×8 pixels at the fifth level. On the contrary, the decoder part continues by increasing the image dimensions from the fifth level to upwards. A segmented image with the same size as the input was obtained at the output. The ReLU activation function was used in the convolution layers and the kernel size in the max-pooling layers was taken by 2×2 pixels. The network is shown in Fig. 1 [46].
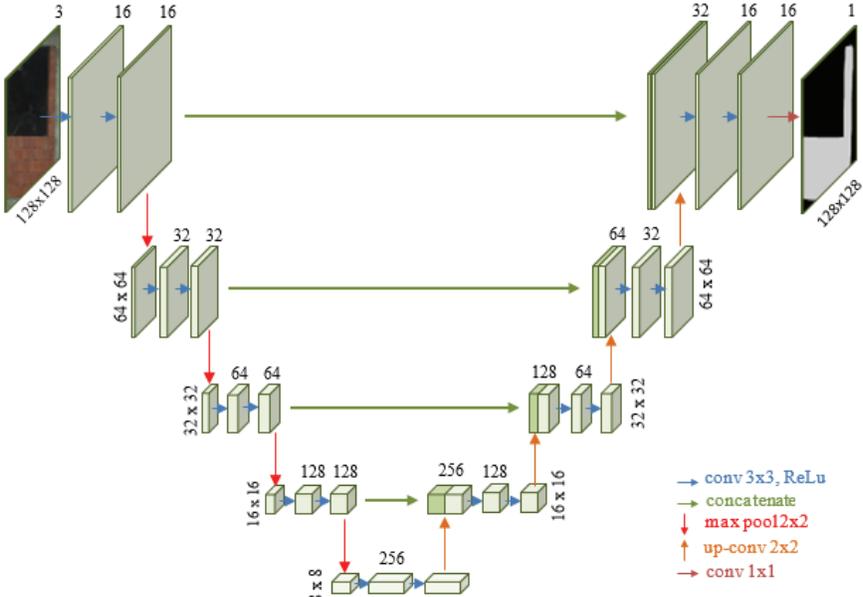
Fig 1 - U-Net architecture [46]

4. MATERIALS AND METHODS

This study, in which brick wall quantity take-offs are made with the U-Net method based on deep learning, consists of two stages:

1. Identifying the area view of the brick wall on the photo with the U-Net method
2. Calculation of the identified wall area.

The methodology of the study is given in Fig. 2.

4.1. Data Sources

The photos used in the training for automatic brick wall detection were taken by the authors with the help of a mobile phone. Different buildings were used so that the data set is not created by the photos of the walls made by the same master. To represent any condition, the photos were taken from different distances, different angles, indoors, outdoors, sunny areas, shaded areas, etc. In this way, a total of 280 images with 3024×3024 pixels were obtained. Some examples of wall photos are shown in Fig. 3.

4.2. Mask the Images

The training data prepared for the U-Net method was divided into two groups as the original image and the masked image. Masked images were obtained by painting the region to be

determined in the original images with a single color and the other parts with a different color. In the 280 photos used in the study, the areas with walls are masked as white, and the other areas are masked as black. Examples of masked images are given in Fig. 4.
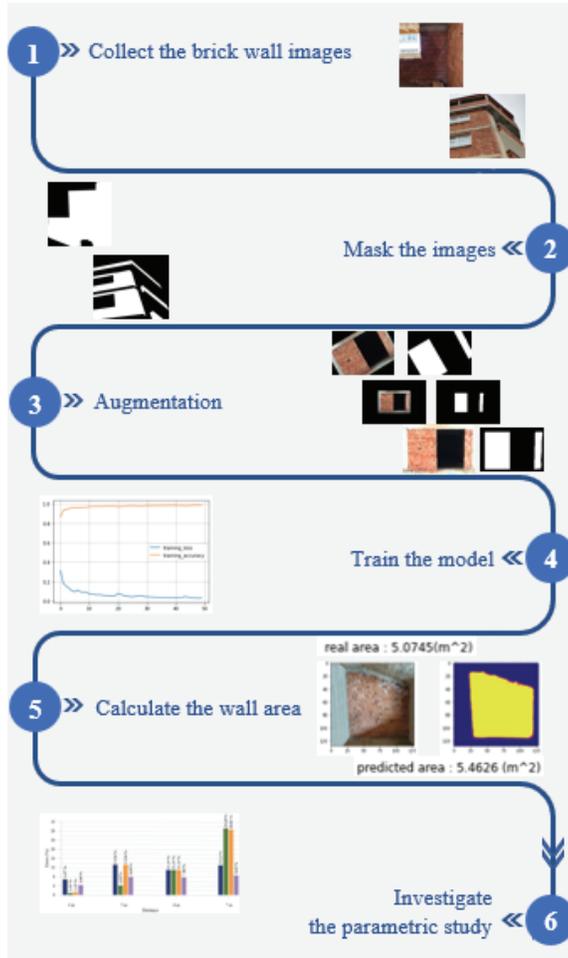
Fig. 2 - The methodology of the study

Fig. 3 - Examples of brick wall images

a. Original images

b. Masked form of the original images

Fig. 4 - Masked images

4.3. Augmentation

In the absence of a suitable dataset, the method of increasing the data is generally used for better training. Certain image enhancement techniques have been used to increase the amount of training data in the analysis. This is because the masking process takes too long. Masking a photo takes between 1-5 minutes, depending on the amount of detail. It took approximately 12 hours to mask the 280 images used in the study. The data has been duplicated because more image masking will require more effort and time. For this purpose, brightness (dark), brightness (light), rotation (30°), flip (vertical), rotation (-30°), and zoom (out) were performed for each photo. The same operations were performed on masked images. Thus, a total of 1960 images were obtained by adding 1680 (280×6) more images to the training set. Table 1 shows the operations for two images.

4.4. Settings

The study used Google Colab Notebook with virtual GPU, which is a free cloud service. This system develops deep learning applications using popular libraries such as Keras, TensorFlow, PyTorch, and OpenCV. Colab, up to Tesla K80 with 12 GB of GDDR5 VRAM, Intel Xeon Processor with two cores @ 2.20 GHz and 13 GB RAM, provided 12 hours of continuous execution time.

4.5. Accuracy Assessment

The IoU score is a standard performance measure for the object category segmentation problem. Given a set of images, the IoU measure gives the similarity between the predicted region and the ground-truth region for an object presented in the set of images and is defined by the following [47]:

$$IoU = \frac{TP}{TP+FP+FN} \tag{1}$$

where, TP, FP, and FN denote the true positive, false positive, and false negatives, respectively. If the prediction is completely correct, then IoU = 1. The lower the IoU, the worse the prediction result.

Table 1 - The augmented form of the images

Orginal	Brightness (dark)	Brightness (light)	Flip (Vertical)	Rotation (30^0)	Rotation (-30^0)	Zoom (out)
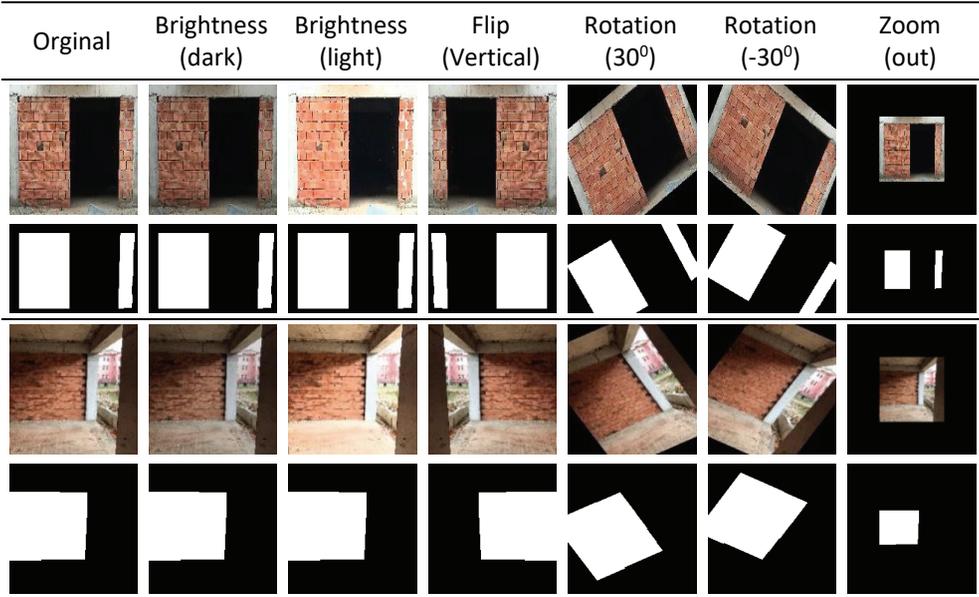
5. RESULTS

5.1. Model Training

The data set was trained using the U-Net architecture. 80% of the data set (1960 images) was reserved as training data and 20% as validation data. The number of epochs was set as 50 and the batch size was 16 since the running of the training data in the corresponding dataset of the model created for brick wall detection was considered as one epoch. The loss function and accuracy values are given in Table 2.

Table 2 - Loss assessment and accuracy results of U-Net architecture

Loss Function	Accuracy Value
0.0338	0.9846

The results of the loss evaluation and accuracy of the 128×128 U-Net architecture are shown in Fig. 5. In these graphics, blue curves indicate the training accuracy, and orange ones show the training loss rate. The best validation value was obtained at step 40.

The average IoU value of the training data set was calculated to be 0.95. This value also supports the results given in Table 2 and Fig. 5. According to the results obtained, the training data and epoch value were sufficient to identify the brick wall and high success was achieved. A comparison of masked and predicted images of some photos included in the training data is given in Fig. 6. They are selected from images with different IoU values.
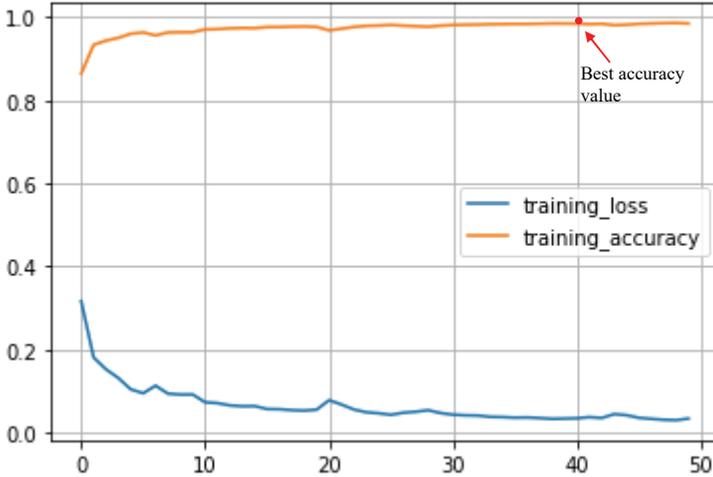
Fig. 5 - Loss assessment and accuracy results of U-Net architecture
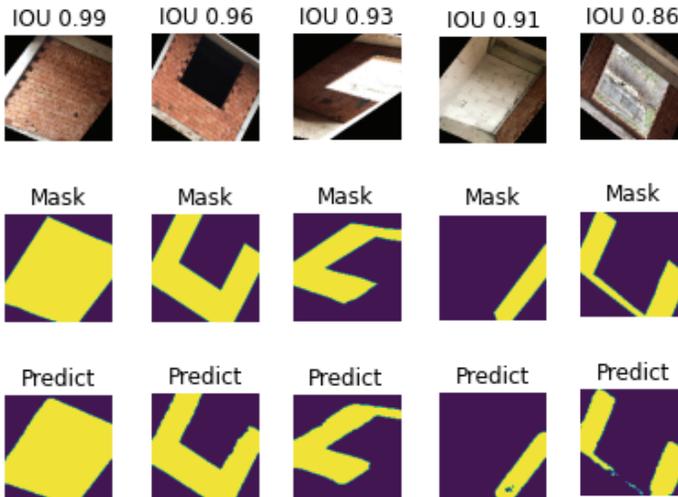
Fig. 6 - Comparison of masked and predicted images
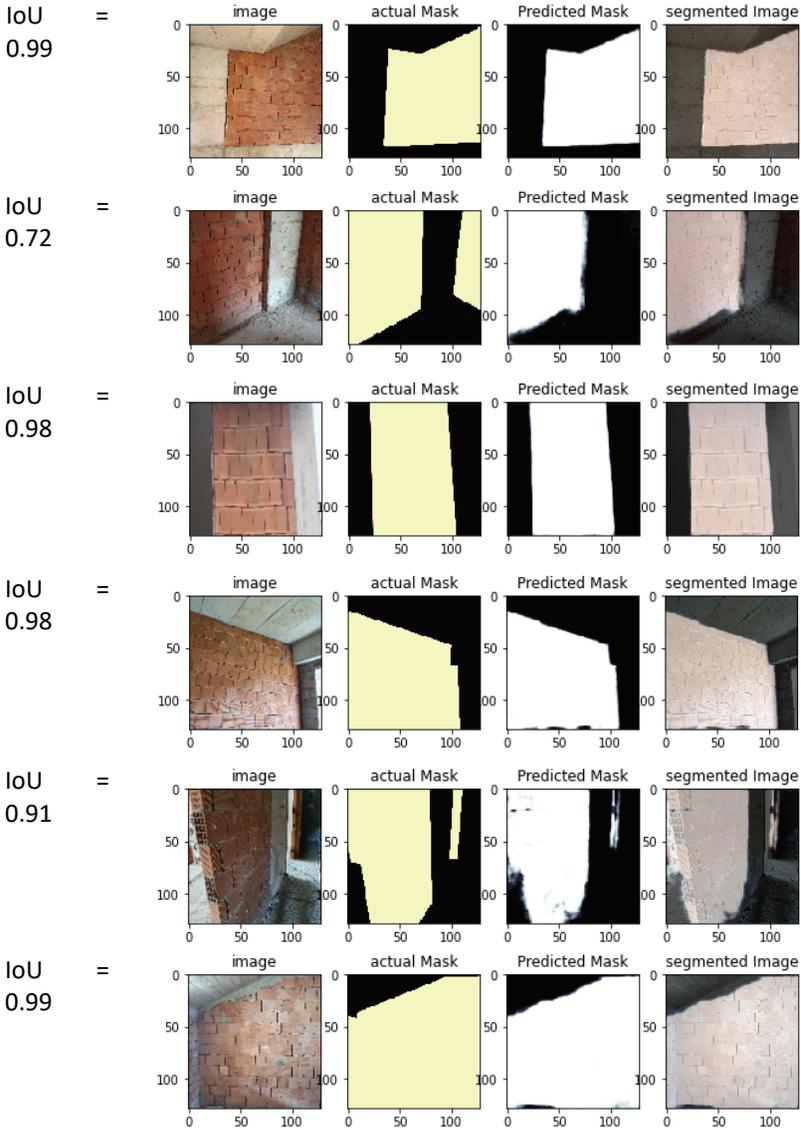
Fig. 7 - The segmentation results of the images

To test the success of the trained model, images taken from a different building, which are not included in the training dataset, were used. Fig. 7 shows the masked and predicted states of these images and the projection of the estimated area on the original image. While choosing the photos, the ones with complex wall shapes and details were preferred. According to the results, it is observed that most images have a high IoU value. Among the factors that reduce the value of the IoU (in other words, the success of identifying the wall) can be counted as light reflections, the presence of regions in the dark area, the details not

being seen very clearly. Success at this stage is very important, as identifying the correct location of the wall directly affects the finding of the wall area. Here, it is necessary to consider the factors affecting the success while taking the photo of the wall to be calculated.

5.2. The Brick Wall Area Calculation

Using the trained data set, the location of the brick walls can be detected and masked on the photo. This area can be calculated in pixel size. However, a reference field on the image is necessary to convert the pixel-sized area to the actual field value since the distance to the wall or the angle of shooting changes the pixel area of the wall on the photo. In this case, a transformation is not possible to calculate the actual field value. In this study, a method for this transformation is presented. The brick area was taken as the reference area on the photo of the wall whose area will be calculated. Since the actual dimensions of the brick are known, the transformation was performed with the following formula:

$$Brick\ wall\ predicted\ area = \frac{Brick\ wall\ predicted\ pixel\ area}{Brick\ pixel\ area} \times Brick\ real\ area \qquad (2)$$

Calculations occur automatically by following the steps below:

1. Any brick is masked on the picture on which the brick wall area calculation will be made
2. The trained model is given the original image, the masked image, and the actual area of the brick
3. The estimated mask image corresponding to the original image from the model is obtained
4. The mask pixel area is calculated on the estimated mask image
5. Brick mask pixel area is calculated
6. With the help of Eq. (2), the estimated wall area is calculated

Four examples were chosen to test the calculations described above: interior wall, outer wall, trapezoidal wall, and wall with a gap. Photos were taken close enough to see the wall with the mobile phone held vertically. No mechanism (such as a tripod) was used to ensure the verticality of the phone while taking the photo. The purpose of this is to measure the wall area estimation performance from photos taken directly from the field. The photos for which the wall area will be calculated and their dimensions are given in Fig. 8. stabilizer, a fully upright position may not be achieved. This may also affect the results. These challenges will be discussed further in the next sections.

After the photos were trained in the U-net model, calculations were made   and then the results are given in Fig. 9. Here, the image in the first column shows the original photo of the wall, the image in the second column shows the masked brick area, and the image in the last column shows the estimated area. A comparison of the results is given in Table 3. The walls are built with 19×19×13.5 cm bricks, and the 19×19 cm surface of the brick forms the wall area.

a. Interior wall

b. Outer wall

c. Trapezoidal wall

d. Wall with a gap
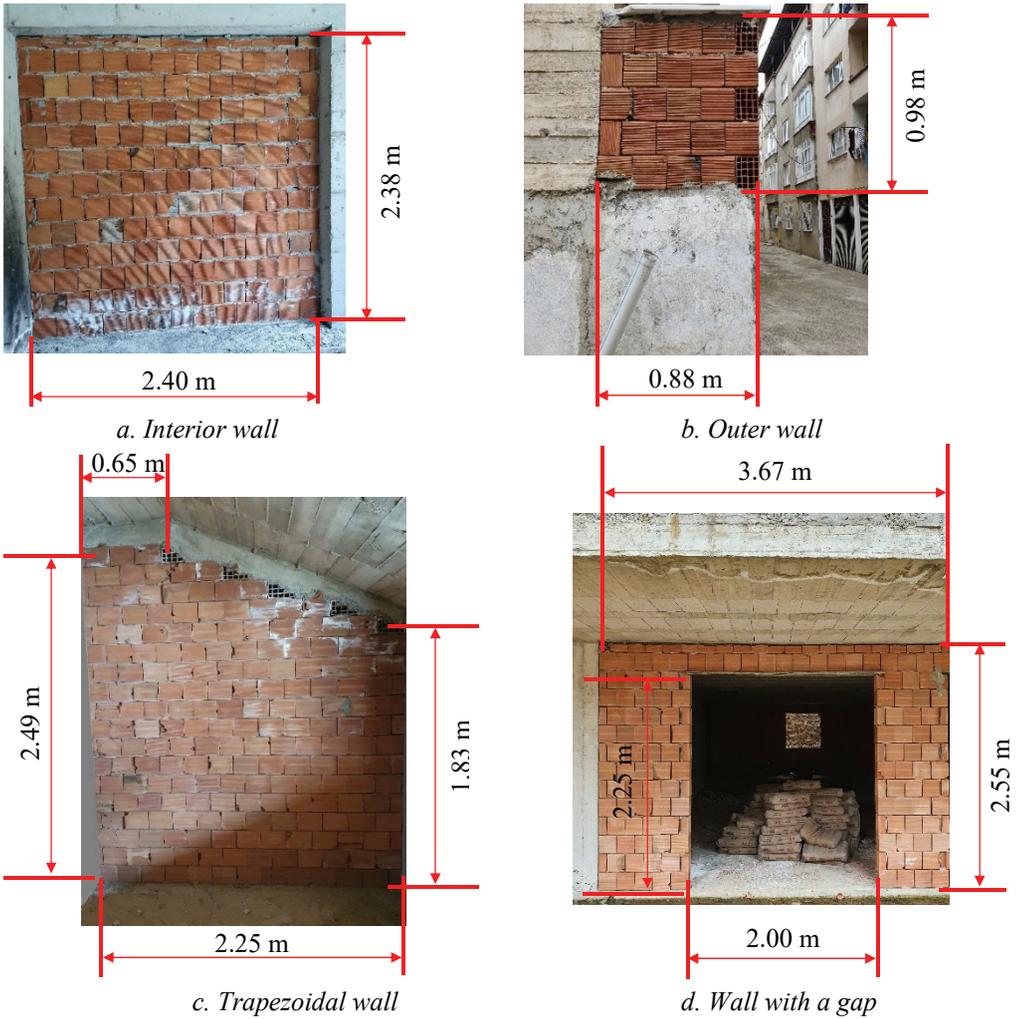
Fig.8 - The dimensions of the test images

As can be seen from Table 3, the error rate in the selected samples is calculated between 3.36 and 9.40. All of the IoU values, that show the success of detecting the wall in the photos, are 0.90 and above. As mentioned earlier, the first step in making a correct prediction is to correctly identify the wall. Mortar particles or dirt on the wall can also affect the masked area, as they can block the detection of the wall. To eliminate this situation and similar ones, it is necessary to increase the variety of photos used in training. This will provide realistic masking. According to the data in the table, there is no direct relationship between IoU and absolute error. This is due to deviations that can be made when masking the reference brick. Since the brick is an element within the wall, it is difficult to define its boundaries precisely. In addition, external factors such as shadows and the mixing of colors are also effective in determining the boundaries of the brick. Further, since the photos were taken without any

Table 3 - Comparison of the real and predicted areas of brick walls

	Real area (m²)	Predicted area (m²)	Absolute Error (%)	IoU
Interior wall	5.7120	5.1753	9.40	0.94
Outer wall	0.8624	0.8334	3.36	0.90
Trapezoidal wall	5.0745	5.4626	7.65	0.98
Wall with a gap	4.8585	4.5711	5.92	0.97
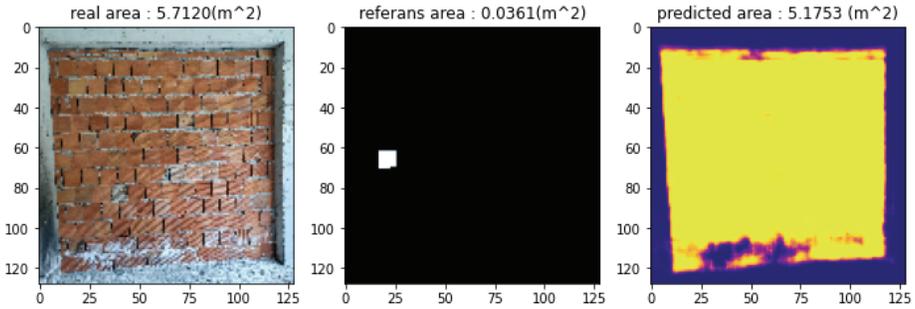
a. Interior wall
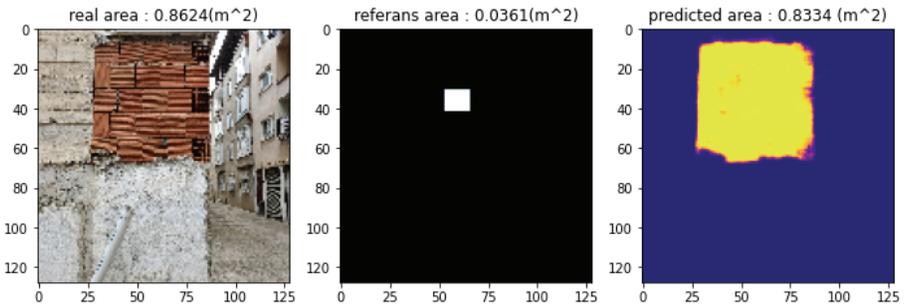
b. Outer wall
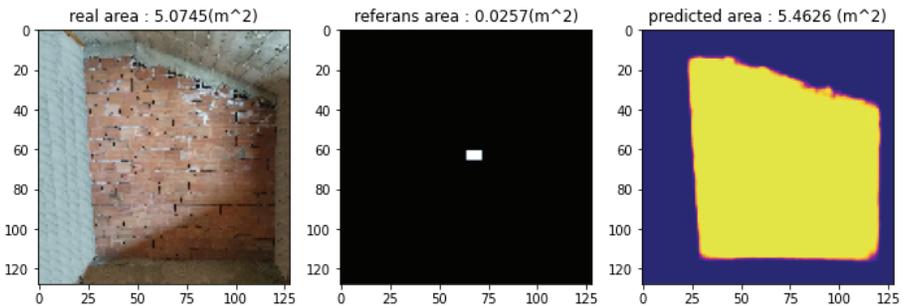
c. Trapezoidal wall
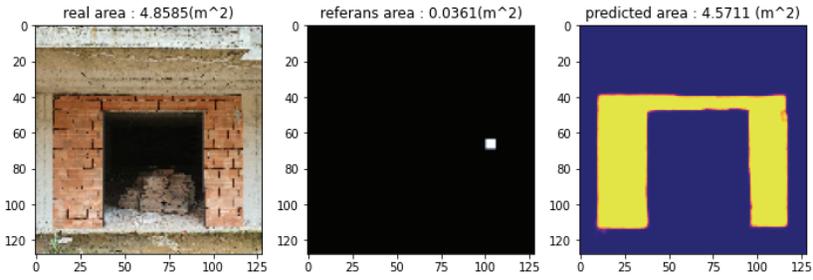
d. Wall with a gap

Fig 9 - The predicted masks and calculated areas by deep learning

5.3. Examining the Effect of Distance on the Wall Area Calculation

The photos, for which the wall area is calculated, were obtained by taking photos from the closest distance to see the wall. To determine the distance effect, photos were taken from a distance of 4, 5, 6, and 7 m from the wall, and the results were compared. Fig. 10 shows the IoU values of the walls and the estimated images. In Fig. 11, the obtained wall areas and errors (%) are shown. As can be seen from the figures, the highest IoU value and the lowest error rate value are obtained from the photographs taken from a distance of 4 m. As the
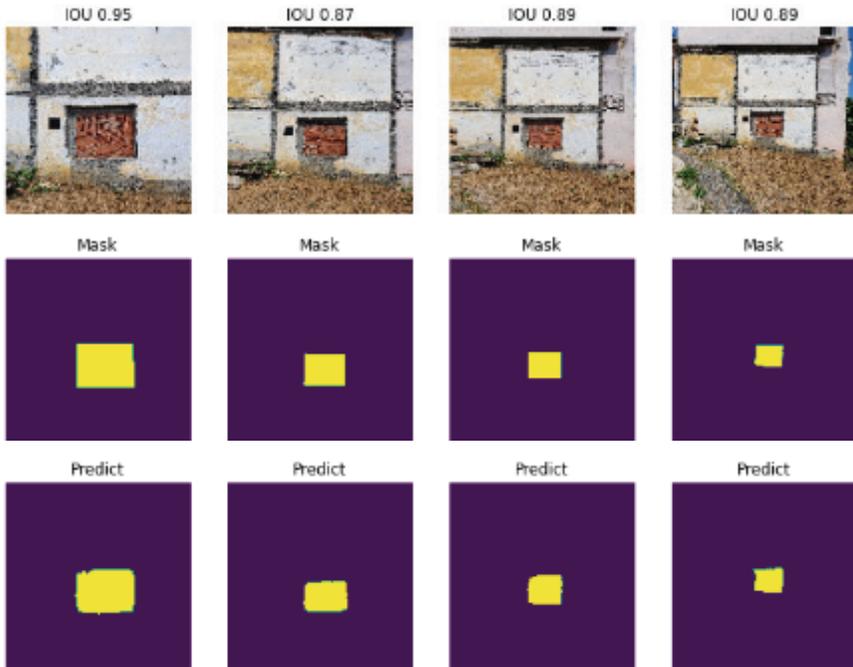
Fig. 10 - IoU values of the walls and the estimated images

distance increases, the loss of details, the more obvious shadows or sun glare, the increase of details entering the photo frame, and the presence of brick-like images cause the IoU to decrease. Also, as the distance increases   it becomes more difficult to mark the reference brick. Therefore, the increase in the distance causes the error rate to increase as well.

5.4. Examining the Effect of Reference Area on the Wall Area Calculation

As stated before, the exact marking of the reference area is of great importance in the correct calculation of the wall area. In particular, as the shooting distance increases, it becomes more difficult to distinguish colors and define brick boundaries. Parametric studies were carried out to reduce the error rate in long-shot images. Accordingly, the wall areas were calculated by considering two bricks, three bricks, and an area other than the brick, and different photographing distances (4, 5, 6, and 7 m). A4 size (21×29.7 cm) black paper was pasted next to the brick wall to determine the reference area outside the brick. The aim here is to examine the results that would be obtained if the reference field is visible at any distance. Table 4 gives reference markings according to the distances. The results obtained are shown in Fig. 12.

In case the photographing distance is 4 m, the error rate for the reference area of two and three bricks decreases to 1%. As the distance increases, the error rates generally increase. If the reference area is an A4 paper, the error rates are approximately close to each other as the photographing distance increases. According to these results, taking a photo of the wall as close as possible and considering more than one brick instead of a single brick as the reference area revealed that the minimum error rate will be obtained in the wall area calculation. If an area other than the brick is taken as a reference, the error rate does not change appreciably   as the distance increases, since this area could be masked more clearly. However, due to the mixing of the color of the paper with the shade, it becomes more difficult to set the paper boundaries as the distance increases, especially during masking. For this reason, color research for different weather conditions on the reference surface will reduce the error rate even more.
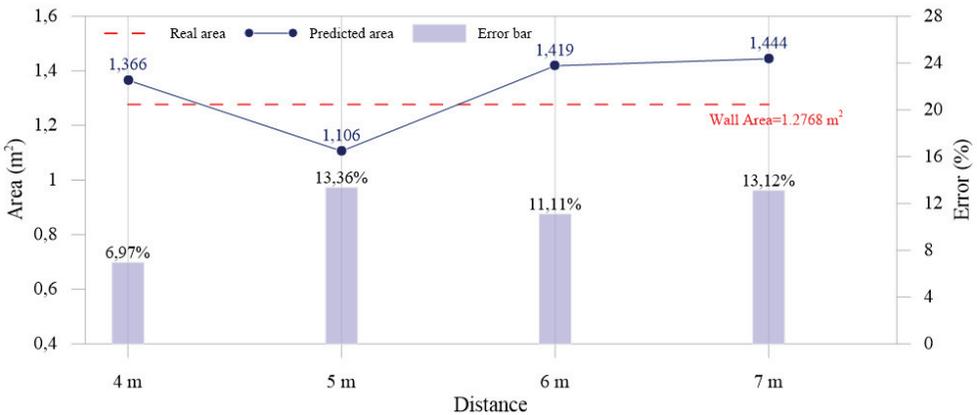
Fig. 11 - The calculated wall areas and errors (%)

15

Table 4 - The reference markings according to distances

Distance	Original	Ref: 1 brick	Ref: 2 bricks	Ref: 3 bricks	Ref: A4 paper
4 m					
5 m					
6 m					
7 m					
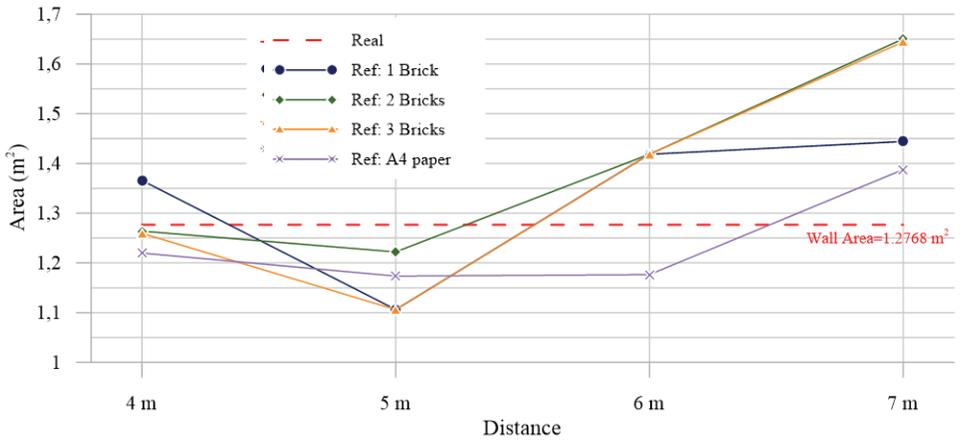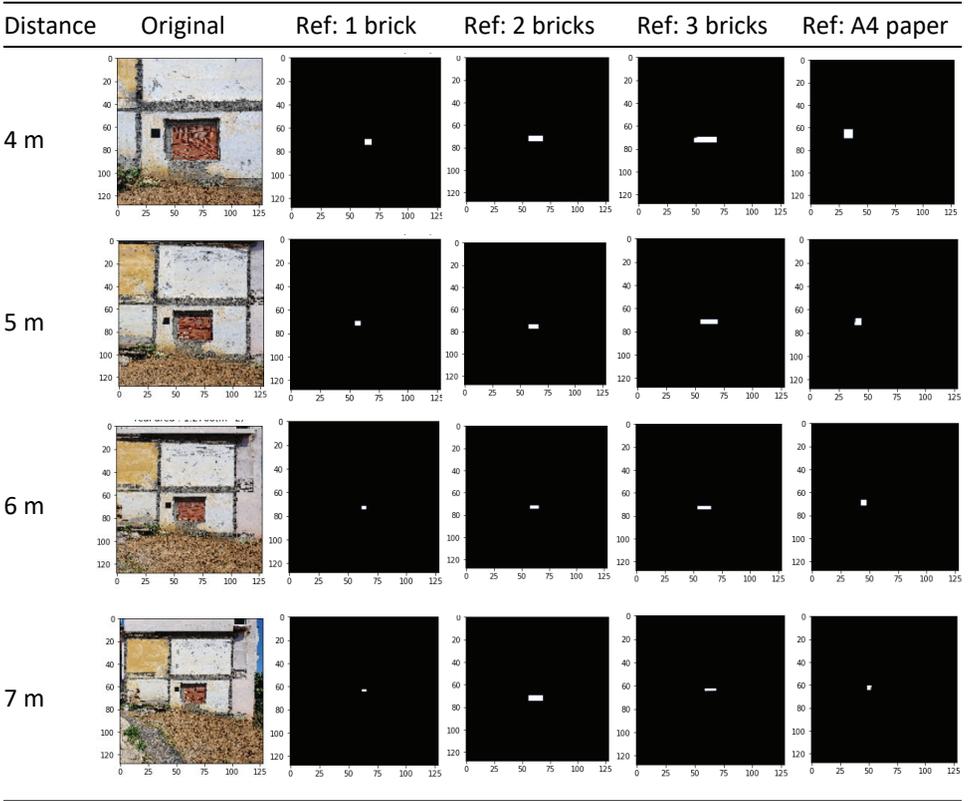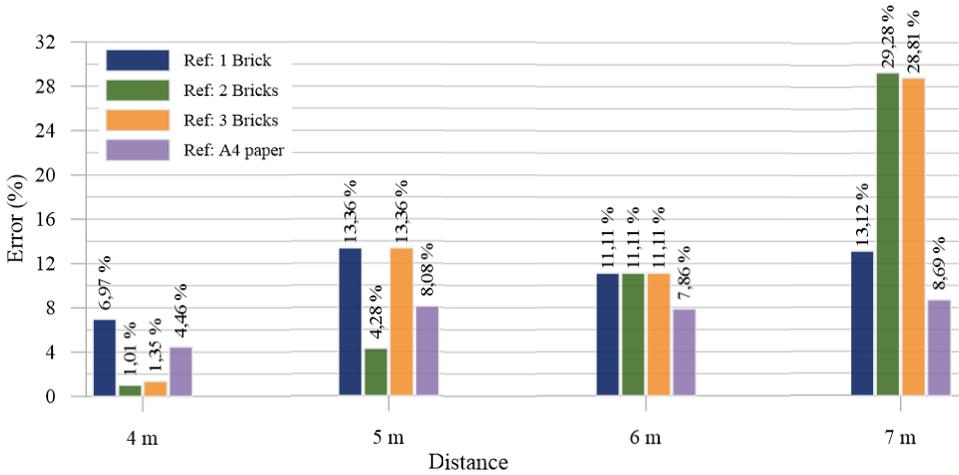
a. The wall areas

b. The errors

Fig. 12 - The wall areas and calculation errors according to distances and reference point

6. DISCUSSIONS

This study aimed to calculate the wall area automatically from the photograph by using the deep learning method. In this context, first of all, it is necessary to automatically determine the wall from the photograph. For this purpose, the photographs obtained by the authors were masked and trained with the U-Net segmentation method. The results showed that the dataset was sufficient for training, with high accuracy (98.46% accuracy and 0.95 mean Io U). Accordingly, the wall area can be determined automatically with high accuracy from the photograph. In order to find the real area of the wall determined in the photo, a reference area is needed on the photo. In the study, bricks of known dimensions were selected for this reference. The proposed method for wall area calculation was tested on four different walls; interior rectangle wall, outer rectangle wall, trapezoidal wall, and wall with a gap . The results (mean error: 6.58%) showed that this method can be used to calculate the wall area automatically.

In the study, a parametric study is also conducted to examine the effect of different photo shooting distances and the number of reference bricks on the accuracy of the wall area calculation. The results showed that taking photos from a closer distance can significantly reduce the error rate. For instance, when the photo shooting distance was 4m, the error rate was 6.97%, which is significantly lower than the error rates obtained from photos taken from 5m, 6m, and 7m distances.

Furthermore, in the study, the effect of using different numbers of bricks as a reference area has also been investigated. It was observed that taking more bricks as a reference area can significantly reduce the error rate. For instance, using two bricks as a reference area resulted in an error rate as low as 1.01% when the photo shooting distance was 4m. However, using

three bricks as a reference area did not always lead to a lower error rate. In some cases, it resulted in a higher error rate compared to using only one brick as a reference area.

Additionally, in the study, the effect of using an A4 size paper affixed to the wall as a reference area. has been examined, too. The results showed that using an A4 size paper can also lead to a lower error rate compared to using only one brick as a reference area. However, it is important to note that the error rate obtained with this method was not as low as the error rate obtained when using two bricks as a reference area. Overall, the results of the parametric study suggest that taking photos from a closer distance and using more bricks or an A4 size paper as a reference area can significantly improve the accuracy of the wall area calculation. These findings can contribute to the body of knowledge by providing valuable insights into the factors that can affect the accuracy of wall area calculations and the ways to improve them.

The proposed method can be integrated to BIM processes, which can provide valuable information for construction projects. By using the proposed method, the wall area can be automatically calculated and added to the BIM model, which can reduce the time and cost of manual quantity take-off. Additionally, the proposed method can be used for real-time data collection on construction sites. By using a mobile device, the workers can take photos of the walls, and the proposed method can automatically calculate the wall area. This can provide real-time information for project managers, which can help them make informed decisions and monitor the progress of the project.

Finally, the proposed method can be used in conjunction with digital twin technology. By creating a digital twin of the construction site, the proposed method can be used to automatically calculate the wall area in the virtual environment. This can provide valuable insights for architects and engineers, and can help them optimize the design and construction processes. Overall, the proposed method has the potential to improve the efficiency and accuracy of construction projects, and can be integrated with various digital technologies to provide valuable insights for construction professionals.

7. CONCLUSIONS

This study presented a novel method for automatically estimating brick wall quantity from photographs using deep learning. The U-Net architecture was employed to train a model to accurately mask the location of the wall on the photo, followed by estimating the actual area of the wall based on a reference area of known size. Our results showed that the model achieved an accuracy of 0.9846 with an IoU of 0.95, and the maximum error rate was 9.40% for the tested photos. Overall, these findings can be used to inform best practices for brick wall quantity take-off automation and provide guidance for improving the accuracy of construction estimation processes.

In the study, parametric studies were carried out to show the effect of the reference area and the shooting photo distance on the calculation of the wall area. The results showed that taking photographs at a closer distance and using multiple bricks as a reference area can significantly reduce the error rate in wall area calculations. These results showed that the main limitation of the current method is that the distance between the camera and the wall can affect the accuracy of the wall area calculation. Accordingly, future studies should be on developing

methods that can eliminate the effect of camera distance in wall area calculations. In addition, it is recommended to investigate the integration of the proposed method to   Building Information Modeling (BIM) processes and real-time data collection for more efficient use of this system.

In summary, the study demonstrates the potential of deep learning in automating brick wall quantity take-off from photos, and provides valuable insights for further research and development in this area.

Declaration of Competing Interest

The author declares that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1]   Huang, T.S. Computer Vision: Evolution And Promise. In 19th CERN School of Computing, CERN, Geneva; 1996; pp. 21–25.

[2]   Lowe, D.G. Distinctive Image Features from Scale-Invariant Keypoints. Int J Comput Vis, 2004, 60, 91–110.

[3]   Dalal, N.; Triggs, B. Histograms of Oriented Gradients for Human Detection. Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005, 2005, I, 886–893.

[4]   Cristianini, N.; Shawe-Taylor, J. An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods; Cambridge University Press, 2000.

[5]   O'Mahony, N.; Campbell, S.; Carvalho, A.; Harapanahalli, S.; Hernandez, G.V.; Krpalkova, L.; Riordan, D.; Walsh, J. Deep Learning vs. Traditional Computer Vision. Advances in Intelligent Systems and Computing, 2020, 943, 128–144.

[6]   Nanni, L.; Ghidoni, S.; Brahnam, S. Handcrafted vs. Non-Handcrafted Features for Computer Vision Classification. Pattern Recognit., 2017, 71, 158–172.

[7]   Chan, T.H.; Jia, K.; Gao, S.; Lu, J.; Zeng, Z.; Ma, Y. PCANet: A Simple Deep Learning Baseline for Image Classification? IEEE Transactions on Image Processing, 2015, 24, 5017–5032.

[8]   Paneru, S.; Jeelani, I. Computer Vision Applications in Construction: Current State, Opportunities & Challenges. Autom Constr, 2021, 132.

[9]   Guo, B.H.W.; Zou, Y.; Fang, Y.; Goh, Y.M.; Zou, P.X.W. Computer Vision Technologies for Safety Science and Management in Construction: A Critical Review and Future Research Directions. Saf Sci, 2021, 135, 105130.

[10] Wu, H.; Zhong, B.; Li, H.; Love, P.; Pan, X.; Zhao, N. Combining Computer Vision with Semantic Reasoning for On-Site Safety Management in Construction. Journal of Building Engineering, 2021, 42.

[11] Fang, W.; Ding, L.; Love, P.E.D.; Luo, H.; Li, H.; Peña-Mora, F.; Zhong, B.; Zhou, C. Computer Vision Applications in Construction Safety Assurance. Autom Constr, 2020, 110.

[12] Xu, S.; Wang, J.; Shou, W.; Ngo, T.; Sadick, A.M.; Wang, X. Computer Vision Techniques in Construction: A Critical Review. Archives of Computational Methods in Engineering, 2021, 28, 3383–3397.

[13] Xu, S.; Wang, J.; Wang, X.; Shou, W. Computer Vision Techniques in Construction, Operation and Maintenance Phases of Civil Assets: A Critical Review. Proceedings of the 36th International Symposium on Automation and Robotics in Construction, ISARC 2019, 2019, 672–679.

[14] Li, Y.; Lu, Y.; Chen, J. A Deep Learning Approach for Real-Time Rebar Counting on the Construction Site Based on YOLOv3 Detector. Autom Constr, 2021, 124, 103602.

[15] Fan, Z.; Lu, J.; Qiu, B.; Jiang, T.; An, K.; Josephraj, A.N.; Wei, C. Automated Steel Bar Counting and Center Localization with Convolutional Neural Networks. 2019.

[16] Wang, H.; Polden, J.; Jirgens, J.; Yu, Z.; Pan, Z. Automatic Rebar Counting Using Image Processing and Machine Learning. In 2019 IEEE 9th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER); IEEE, 2019; pp. 900–904.

[17] Akanbi, L.A.; Oyedele, A.O.; Oyedele, L.O.; Salami, R.O. Deep Learning Model for Demolition Waste Prediction in a Circular Economy. J Clean Prod, 2020, 274, 122843.

[18] Fang, Q.; Li, H.; Luo, X.; Ding, L.; Luo, H.; Rose, T.M.; An, W. Detecting Non-Hardhat-Use by a Deep Learning Method from Far-Field Surveillance Videos. Autom Constr, 2018, 85, 1–9.

[19] Wu, J.; Cai, N.; Chen, W.; Wang, H.; Wang, G. Automatic Detection of Hardhats Worn by Construction Personnel: A Deep Learning Approach and Benchmark Dataset. Autom Constr, 2019, 106, 102894.

[20] Nath, N.D.; Behzadan, A.H.; Paal, S.G. Deep Learning for Site Safety: Real-Time Detection of Personal Protective Equipment. Autom Constr, 2020, 112, 103085.

[21] Yu, Y.; Li, H.; Yang, X.; Kong, L.; Luo, X.; Wong, A.Y.L. An Automatic and Non-Invasive Physical Fatigue Assessment Method for Construction Workers. Autom Constr, 2019, 103, 1–12.

[22] Yang, K.; Ahn, C.R.; Kim, H. Deep Learning-Based Classification of Work-Related Physical Load Levels in Construction. Advanced Engineering Informatics, 2020, 45, 101104.

[23] Kolar, Z.; Chen, H.; Luo, X. Transfer Learning and Deep Convolutional Neural Networks for Safety Guardrail Detection in 2D Images. Autom Constr, 2018, 89, 58–70.

[24] Pan, Y.; Zhang, G.; Zhang, L. A Spatial-Channel Hierarchical Deep Learning Network for Pixel-Level Automated Crack Detection. Autom Constr, 2020, 119, 103357.

[25] Yang, Q.; Shi, W.; Chen, J.; Lin, W. Deep Convolution Neural Network-Based Transfer Learning Method for Civil Infrastructure Crack Detection. 2020.

[26] Kang, D.; Benipal, S.S.; Gopal, D.L.; Cha, Y.-J. Hybrid Pixel-Level Concrete Crack Segmentation and Quantification across Complex Backgrounds Using Deep Learning. Autom Constr, 2020, 118, 103291.

[27] Yang, C.; Chen, J.; Li, Z.; Huang, Y. Structural Crack Detection and Recognition Based on Deep Learning. Applied Sciences, 2021, 11, 2868.

[28] Zheng, M.; Lei, Z.; Zhang, K. Intelligent Detection of Building Cracks Based on Deep Learning. Image Vis Comput, 2020, 103, 103987.

[29] Zhou, S.; Song, W. Deep Learning-Based Roadway Crack Classification Using Laser-Scanned Range Images: A Comparative Study on Hyperparameter Selection. Autom Constr, 2020, 114, 103171.

[30] Hacıefendioğlu, K.; Başağa, H.B. Concrete Road Crack Detection Using Deep Learning-Based Faster R-CNN Method. Iranian Journal of Science and Technology, Transactions of Civil Engineering, 2021.

[31] Zhou, C.; Xu, H.; Ding, L.; Wei, L.; Zhou, Y. Dynamic Prediction for Attitude and Position in Shield Tunneling: A Deep Learning Method. 2019.

[32] Xu, Y.; Bao, Y.; Chen, J.; Zuo, W.; Li, H. Surface Fatigue Crack Identification in Steel Box Girder of Bridges by a Deep Fusion Convolutional Neural Network Based on Consumer-Grade Camera Images. Struct Health Monit, 2019, 18, 653–674.

[33] Zhang, C.; Chang, C.; Jamshidi, M. Concrete Bridge Surface Damage Detection Using a Single-stage Detector. Computer-Aided Civil and Infrastructure Engineering, 2020, 35, 389–409.

[34] Wang, L.; Zhao, Z.; Xu, N. Deep Belief Network Based 3D Models Classification in Building Information Modeling. International Journal of Online and Biomedical Engineering (iJOE), 2015, 11, 57.

[35] Wang, L.; Zhao, Z.; Wu, X. A Deep Learning Approach to the Classification of 3D Models under BIM Environment. International Journal of Control and Automation, 2016, 9, 179–188.

[36] Rafiei, M.H.; Adeli, H. A Novel Machine Learning Model for Estimation of Sale Prices of Real Estate Units. J Constr Eng Manag, 2016, 142, 04015066.

[37] Rafiei, M.H.; Adeli, H. Novel Machine-Learning Model for Estimating Construction Costs Considering Economic Variables and Indexes. J Constr Eng Manag, 2018, 144, 04018106.

[38] Mocanu, E.; Nguyen, P.H.; Gibescu, M.; Kling, W.L. Deep Learning for Estimating Building Energy Consumption. Sustainable Energy, Grids and Networks, 2016, 6, 91–99.

[39] Rahman, A.; Smith, A.D. Predicting Heating Demand and Sizing a Stratified Thermal Storage Tank Using Deep Learning Algorithms. Appl Energy, 2018, 228, 108–121.

[40] Rahman, A.; Srikumar, V.; Smith, A.D. Predicting Electricity Consumption for Commercial and Residential Buildings Using Deep Recurrent Neural Networks. Appl Energy, 2018, 212, 372–385.

[41] Hacıefendioğlu, K.; Demir, G.; Başağa, H.B. Landslide Detection Using Visualization Techniques for Deep Convolutional Neural Network Models. Natural Hazards, 2021.

[42] Hacıefendioğlu, K.; Başağa, H.B.; Demir, G. Automatic Detection of Earthquake-Induced Ground Failure Effects through Faster R-CNN Deep Learning-Based Object Detection Using Satellite Images. Natural Hazards, 2021, 105, 383–403.

[43] Akinosho, T.D.; Oyedele, L.O.; Bilal, M.; Ajayi, A.O.; Delgado, M.D.; Akinade, O.O.; Ahmed, A.A. Deep Learning in the Construction Industry: A Review of Present Status and Future Innovations. Journal of Building Engineering, 2020, 32, 101827.

[44] Xu, Y.; Zhou, Y.; Sekula, P.; Ding, L. Machine Learning in Construction: From Shallow to Deep Learning. Developments in the Built Environment, 2021, 6, 100045.

[45] Long, J.; Shelhamer, E.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. 2014.

[46] Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. 2015.

[47] Rahman, M.A.; Wang, Y. Optimizing Intersection-Over-Union in Deep Neural Networks for Image Segmentation. In; 2016; pp. 234–244.