






Chapter 36

RGB and Thermal Image Analysis for Marble Crack Detection with Deep Learning



Eleni Vrochidou , George K. Sidiropoulos , Athanasios G. Ouzounis ,
Ioannis Tsimperidis , Ilias T. Sarafis, Vassilis Kalpakis, Andreas Stamkos,
and George A. Papakostas 

1 Introduction

Marble stands out for its impressive aesthetic value, adding elegance and luxury anywhere it is being utilized, such as in constructions, interior decoration, statuary, and ornaments. By nature, marble is a unique material. Its physicochemical and mechanical properties make it a rather delicate material. Marble has a high porosity, which allows liquids to penetrate the stone and cause stains. Therefore, its internal structure can be influenced by humidity or decay factors that tend to penetrate materials with pores and deteriorate their durability. Moreover, micro-cracks on marble slabs may increase their porosity. Crack detection and treatment are therefore essential, especially for slabs being exposed to weathering, since temperature expansions may grow the already existing cracks [1]. Currently, marble crack detection is mainly performed by manual visual inspection of marble slabs by experienced workers. However, defect detection by the naked eye, especially when referring to micro-cracks on textured marble surfaces, may be inconsistent and error-prone. Machine vision-based automatic inspection can save time and improve the quality control of marble on the production line, providing a more constant, quicker, and cost-effective alternative [2]. Advances in machine learning (ML) and deep learning (DL) are

E. Vrochidou · G. K. Sidiropoulos · A. G. Ouzounis · I. Tsimperidis · G. A. Papakostas (✉)
MLV Research Group, Department of Computer Science, International Hellenic University,
65404 Kavala, Greece
e-mail: gpapak@cs.ihu.gr

I. T. Sarafis
Department of Chemistry, International Hellenic University, 65404 Kavala, Greece

V. Kalpakis · A. Stamkos
Intermek S.A., 64100 Kavala, Greece

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023
A. Yadav et al. (eds.), *Proceedings of International Conference on Paradigms of
Communication, Computing and Data Analytics*, Algorithms for Intelligent Systems,
https://doi.org/10.1007/978-981-99-4626-6_36

significantly impacting the field of machine vision, providing efficient algorithms for marble defects classification [3, 4], localization, and segmentation [5].

Automatic marble crack detection was first introduced in 1977 [6]; however, to date is not yet sufficiently investigated [7]. Moreover, DL applications for semantic segmentation of marble cracks are scarce in the literature. A DL segmentation approach was proposed in [8]. RGB images were used to train five different convolutional neural networks (CNNs). Results indicated ResNet-50 as the optimal architecture, reporting a mean Intersection over Union (mIoU) of 67.2%. More recently, in [5], a total of 112 DL segmentation model combinations were investigated for marble crack detection in color images. The combination of feature pyramid network (FPN) with SE-ResNet family feature extraction backbone, resulted in 71.35% mIoU. All previous works on marble crack segmentation to date are using color imaging. Thermal imaging has been extensively applied to surface crack detection problems, claiming to be able to better distinguish cracks on materials, compared to RGB [9]; yet, thermal imaging has never been applied for marble crack segmentation.

To this end, this work for the first time applies thermal imaging for marble crack segmentation and comparatively evaluates the results by using different DL models on pairs of thermal and color images. The objective of this study is to evaluate the performance of various image modalities, such as RGB images and thermal images, with the aim of investigating the impact of both image modality and deep learning network architecture on segmentation results. To accomplish this, we will test the performance of different deep learning models on each image modality separately. More specifically, a comparative evaluation of 112 DL segmentation models takes place, combining four semantic segmentation models with 28 feature extraction backbone networks. Experimental results are valued both based on the segmentation model (model-based evaluation) and on the feature extraction method (backbone-based evaluation).

The rest of the paper is structured as follows. Materials and methods are analyzed in Sect. 2. Experimental results are presented in Sect. 3. Discussion of the results and future research directions are provided in Sect. 4. Finally, Sect. 5 concludes the paper.

2 Materials and Methods

In this section, the used dataset, the proposed methodology, and the selected DL segmentation models are presented.

2.1 Dataset

The dataset analyzed during the current study was generated in four successive steps. In the first step, 38 marble tiles with cracks were handpicked by a domain expert

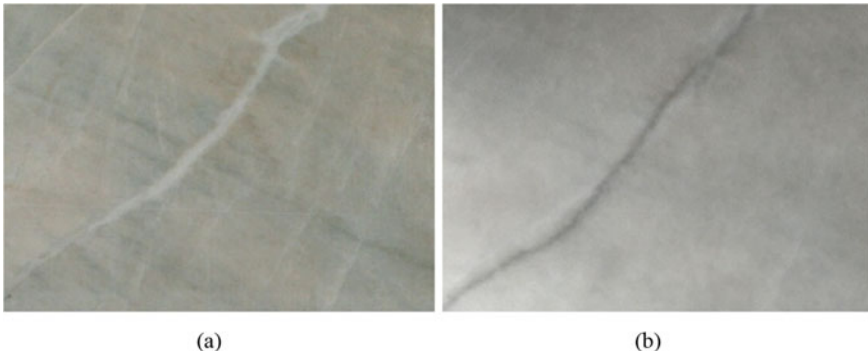


Fig. 1 Marble crack pair of images: **a** RGB image and **b** thermal image

from the production line of the marble quarrying company Solakis [10] in Drama, Greece. Four sizes of tiles were used: 20×40 cm, 30×60 cm, 40×40 cm, and 50×60 cm with hairline cracks of up to 2 mm wide. In the second step, all tiles were photographed at a steady distance of 90 cm using an MVLMF0824M-5MP lens mounted on an MV-CA050-10GM/GC digital camera. This method was complemented by an automatic screening machine with a diffusion box designed and implemented by Intermek A.B.E.E. in Kavala, Greece [11]. The process resulted in the high-resolution RGB images of the tiles used in this work. The thermal images were obtained in the third step in the laboratories of the International Hellenic University [12], in Kavala Campus, Greece. The tiles were first heated with an infrared source and then scanned with the thermal heat-sensitive 206×156 Seek Compact XR camera [13], focused on the cracked areas. In the final step, the RGB and thermal images were first paired, and then, the visible cracks were manually annotated by a domain expert using the LabelMe annotation tool [14].

The original dataset comprised 24 pairs of thermal and RGB images.¹ Figure 1 illustrates a pair of images, referring to an RGB image and the corresponding thermal image of the same marble crack.

Data augmentation was then applied to the original dataset. Random rotation between 0 and 90° , horizontal flip with 50% chance, and vertical Flip with 50% chance were selected, resulting in a total dataset of 244 images for each image category. Five-fold cross-validation was applied to the final dataset to increase the confidence of the model's performance.

2.2 Proposed Methodology

Figure 2 illustrates the basic steps of the proposed methodology. In the first step of the proposed methodology, all original images, RGB and thermal were subjected

¹ <https://github.com/MachineLearningVisionRG/mcs2-dataset>.

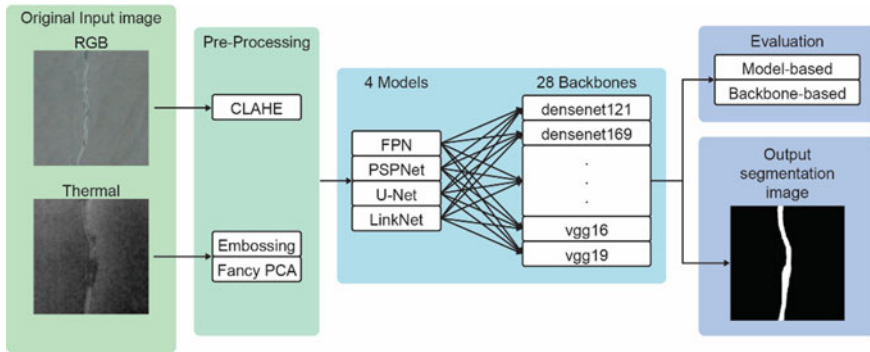


Fig. 2 Proposed methodology

to basic pre-processing. Contrast-limited adaptive histogram equalization (CLAHE) [15] was applied to all RGB images to reduce noise amplification. Thermal images were subjected to image embossing to raise crack patterns against the background. Moreover, principal components analysis color augmentation (Fancy PCA) was also applied to thermal images [16]. The next steps of the proposed approach are based on the methodology presented in [5]. Four semantic segmentation models are combined with 28 feature extraction networks. Results include the output segmentation image and the numerical results in terms of four well-known segmentation metrics, which are evaluated in two different ways: model-based evaluation and backbone-evaluation. The process is repeated for RGB and thermal images of the same cracks (pairs). It should be noted here, that in this work, the dataset is an original in-house dataset of thermal and RGB images of the same marble slabs toward a fair comparative evaluation. Therefore, a comparative study of thermal versus color imaging takes place, by simultaneously highlighting the optimal DL combination that better suits the marble crack segmentation problem for each of the two image categories.

2.3 DL Segmentation Models

The DL segmentation models in this work are combinations of four DL networks and 28 feature extraction backbone networks. The aim is to investigate the most efficient architecture for marble crack detection, for RGB images as well as for thermal images, and compare the results.

DL networks were selected based on their popularity, capabilities, and state-of-the-art reported performances. Therefore, the four selected deep convolutional neural networks models are the following: feature pyramid network (FPN) [17], LinkNet [18], pyramid scene parsing network (PSPNet) [19], and U-Net [20].

Feature extraction backbone networks were selected based on the most efficient families of networks as reported in the literature. Seven well-known families were selected, resulting in 28 total backbones: DenseNet (densenet121, densenet169, and densenet201), EfficientNet (efficientnetb0, efficientnetb1, efficientnetb2, efficientnetb3, and efficientnetb4), Inception (inceptionresnetv2 and inceptionv3), MobileNet (mobilenet and mobilenetv2), ResNet (resnet18, resnet34, resnet50, resnet101, resnet152, resnext50, and resnext101), SE-ResNet (seresnet101, seresnet152, seresnext50, and seresnext101), and VGG (vgg16 and vgg19).

3 Experimental Results

The proposed methodology was implemented in Python 3.9 employing TensorFlow and Keras. All experiments run on an Nvidia RTX 3090 GPU. Original thermal and RGB images were resized in 256×256 pixels size to be inserted as input to the DL models (except for PSPNet requiring 240×240 pixels size). For better convergence, all backbone networks were pretrained on ImageNet [21]. For all DL networks, 75% of their layers were frozen, while the last 25% of the model's layers were trainable.

In this work, the loss function (L) applied in all segmentation experiments, and it is calculated as the sum of focal (FL) and Dice loss (DL):

$$L = FL + DL \quad (1)$$

where the focal loss is

$$FL = -a_t(1 - p_t)^\gamma \log(p_t) \quad (2)$$

and Dice loss is calculated as follows:

$$DL = 1 - \frac{2y\hat{p} + 1}{y + \hat{p} + 1} \quad (3)$$

where in (2), $(1 - p_t)^\gamma$ is the modulator factor, γ is the focus factor, p_t the output of the activation function, and a_t a control weight, while in (3), y is the real, and \hat{p} the predicted value by the model.

The use of the latter unified loss is proven to better handle class imbalanced datasets in semantic segmentation problems and results in improved segmentation quality and a better balance between precision-recall [22]. Table 1 includes information regarding the same hyperparameters for all DL models.

The evaluation of the 112 DL segmentation models was conducted based on two different perspectives: based on the segmentation model and based on the feature extraction backbone network. In what follows, all experimental segmentation results are evaluated after fivefold cross-validation in terms of the following commonly used

Table 1 Hyperparameters of DL models

Hyperparameters	Setting
Activation	Sigmoid
Optimizer	Adam
Loss function	Focal loss, Dice loss
Learning rate	0.0005
Weight decay	1e-8 (0.00000008)
Epochs	50
Steps per epoch	20
Batch size	32
Early stopping	Min_delta = 0.05 Patience = 5

semantic segmentation metrics: Intersection over union (IoU), precision (P), recall (R), and F1-score.

3.1 Model-Based Evaluation

All values included in the tables of this section are the mean values of the performance results on the training set after fivefold cross-validation. Table 2 includes the segmentation results of all models, for both thermal and RGB images.

As it can be seen in Table 2, the best mIoU performance, 71.61%, is reported with the FPN model with RGB images. The second-best performance is reported with the same model with thermal images (68.07%). As a general notice, color and thermal images display similar performances, with small differences. In all cases, however, color images reported slightly better results compared to the corresponding thermal

Table 2 Model-based segmentation results (% mean values) for both thermal and color (RGB) images

Model	Image type	IoU	P	R	F1-score
FPN	Thermal	68.07	86.79	79.31	70.15
	RGB	71.61	93.96	77.01	73.40
U-Net	Thermal	57.24	72.73	79.36	59.24
	RGB	58.63	75.14	79.47	60.65
LinkNet	Thermal	45.89	54.89	87.99	48.24
	RGB	48.25	58.37	86.11	50.82
PSPNet	Thermal	63.15	85.91	74.53	65.35
	RGB	65.46	89.42	74.00	67.93

Best mean IoU for thermal and RGB are marked in bold

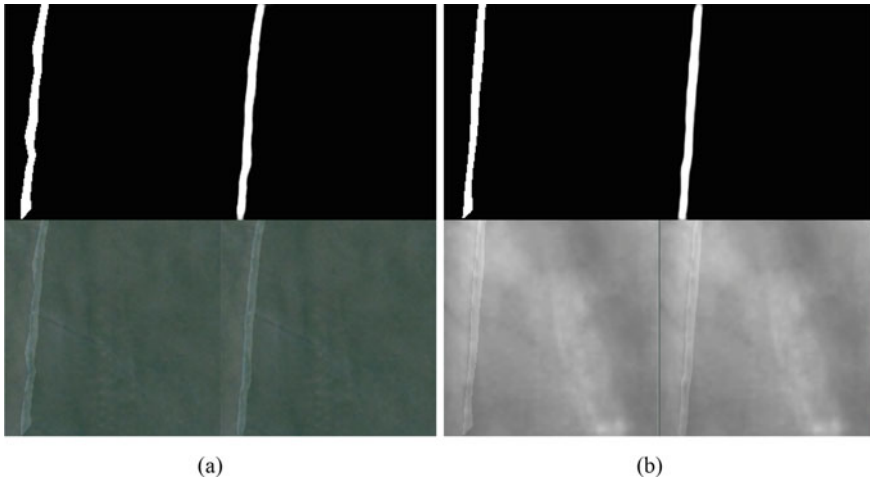


Fig. 3 Indicative results of FPN model to testing images (up-left = ground truth image, down-left = RGB/thermal input image, up-right = output segmentation, down-right = output segmentation mask applied to the input image): **a** from RGB input image (98.01% IoU with efficientnetb3 backbone), **b** from thermal input image (IoU 98.00% with seresnet152)

images. The latter can be attributed to the low resolution of the used thermal camera. Figure 3 displays indicative results of FPN with thermal and RGB images of the testing set reporting the mean mIoU and the used backbone in each depicted case. More specifically, Fig. 3a refers to an RGB input image, resulting in 98.01% IoU with efficientnetb3 backbone, while Fig. 3b refers to a thermal input image resulting in 98.00% IoU with seresnet152 backbone.

The best mIoU performances refer to a DL model combined with a certain backbone. The latter signifies the fact some feature extraction networks can help a model to result in a better segmentation result. In what follows, the evaluation of the results is presented from the perspective of the used backbone.

3.2 Backbone-Based Evaluation

Results are also evaluated from the backbone perspective, so as to highlight the contribution of each backbone to the DL segmentation models. Table 3 summarizes the segmentation results of all models, for both thermal and RGB images. All performance values are the mean values of the results on the training set after fivefold cross-validation.

The best-performing backbone family with thermal images is the EfficientNet (efficientnetb4) with 75.49% mIoU. However, the Inception family reported a higher average mIoU (73.55%), by considering the average of the results of both inception-resnetv2 and inceptionv3, compared to the averages of the rest backbone families.

Table 3 Backbone-based segmentation results (% mean values) for both thermal and color images

Backbone family	Name	Thermal images				RGB images			
		IoU	P	R	F1-score	IoU	P	R	F1-score
DenseNet	densenet121	70.64	89.42	78.76	73.35	74.79	94.53	78.83	77.56
	densenet169	71.14	91.43	77.43	73.71	72.42	95.31	76.13	74.57
	densenet201	67.31	84.60	80.35	70.11	72.64	95.42	76.19	74.86
EfficientNet	efficientnetb0	66.42	79.98	82.86	69.49	77.63	88.26	87.06	81.43
	efficientnetb1	64.47	75.15	86.81	68.02	78.23	85.82	90.09	82.46
	efficientnetb2	70.49	82.89	85.21	73.61	74.81	85.35	87.56	78.65
	efficientnetb3	70.00	81.59	85.54	73.44	75.92	82.88	90.92	80.14
	efficientnetb4	75.49	86.43	86.48	78.79	80.07	90.68	84.41	83.73
Inception	inceptionresnetv2	74.09	91.78	80.42	76.63	73.46	94.72	77.64	75.91
	inceptionv3	73.02	91.37	79.40	75.69	72.78	92.27	79.10	75.44
MobineNet	mobilenet	47.81	58.28	87.90	50.78	62.63	78.15	79.76	65.55
	mobilenetv2	48.23	67.53	76.40	50.36	54.43	86.00	66.98	56.18
ResNet	resnet18	67.01	87.27	77.46	69.40	69.52	94.92	73.67	71.58
	resnet34	63.85	83.79	76.52	66.22	71.24	93.80	76.44	73.54
	resnet50	66.41	90.63	73.72	68.30	68.15	96.01	71.29	70.08
	resnet101	60.63	85.66	73.18	62.09	62.35	98.28	63.67	63.42
	resnet152	63.71	91.05	70.34	65.22	60.64	95.94	63.69	61.62
	resnext50	70.15	92.49	76.40	72.18	70.59	97.00	72.94	72.50
	resnext101	69.38	92.69	75.13	71.13	55.57	85.60	68.11	57.28
SE-ResNet	seresnet18	68.31	87.32	77.92	70.80	69.94	94.78	74.22	71.97
	seresnet34	71.40	95.23	75.03	73.27	72.29	92.78	78.46	74.79
	seresnet50	62.85	74.31	85.63	65.91	79.22	91.16	86.34	82.74
	seresnet101	63.53	78.50	80.03	65.95	68.01	86.09	80.48	70.31
	seresnet152	69.41	88.00	79.58	71.72	61.88	77.01	79.78	64.39
	seresnext50	72.64	86.05	84.83	75.60	78.46	92.39	84.41	81.60
	seresnext101	74.80	90.58	82.41	77.57	65.21	80.86	81.75	67.97
VGG	vgg16	63.08	88.29	72.13	64.97	64.52	87.82	74.32	67.24
	vgg19	62.21	86.14	72.95	64.40	69.77	91.51	75.77	72.81

Best mean IoU for thermal and RGB are marked in bold

Considering the averages of families, the EfficientNet family which displays the higher mIoU performance comes third with 69.37% average mIoU, after Inception (73.55%) and DenseNet with 70.89%.

The best-performing backbone family with RGB images is again EfficientNet and efficientnetb4, with 80.07% mIoU. Moreover, considering the average performance of each family, the EfficientNet family ranks first (77.33%), followed by DenseNet (73.60%) and Inception (73.12%).

It could be therefore concluded, that in all cases, the three backbones standing out are those of EfficientNet, Inception, and DenseNet families. Experimental results verify the hypothesis that specific feature extraction networks can improve a model's segmentation performance.

4 Discussion

The proposed work experimentally demonstrated that deep learning could guarantee efficiency, quality, and reliability in visual inspection of marble slabs. However, further research needs to be carried out in order for the marble crack DL segmentation algorithms to be implemented in industrial settings.

The results of this work indicate similar performances for the DL models, for both thermal and RGB images. However, cracks can be slightly better detected in RGB images rather than in thermal. Infrared thermal imaging is widely used to identify cracks on the surface of materials and underneath them [9, 23, 24]. In most cases, thermal imaging has been proven most efficient for the detection of micro-cracks compared to color imaging [25]; however, there is no reported research on marble. Poor segmentation results of thermal images in this work, compared to those expected, can be attributed to the low resolution of the thermal camera. The thermal sensor used in this work was 206×156 pixels size, not referring to high-resolution thermal imaging. The latter can be clearly observed in the captured images, displaying a lot of noise and being of low quality compared to the corresponding RGB ones. Figure 4a shows a case of a thermal image with very strong dark areas on the marble slab and bright reflections, making the crack not clearly distinguishable, compared to the corresponding RGB image (Fig. 4b).

Therefore, marble crack detection based on deep learning will still be the main direction of future research, by additionally considering some important aspects. To

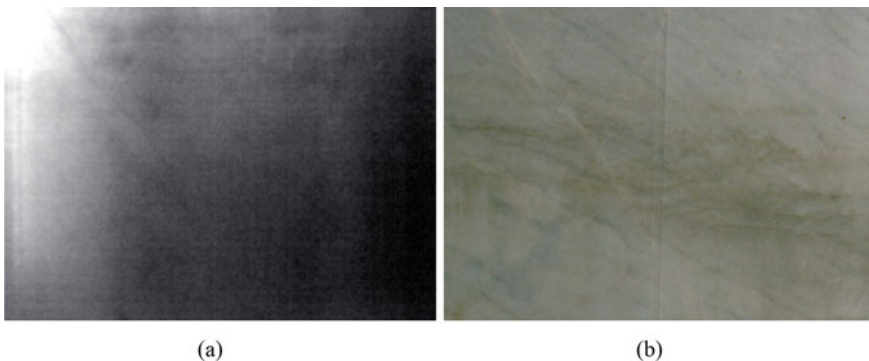


Fig. 4 Marble crack pair of images: **a** Thermal image and **b** RGB image. Result indicates that the model's performances were affected by the low quality and intense noise of thermal images

this end, future work first includes the use of a high-resolution thermal camera and re-capturing of thermal images. Moreover, the fusion of RGB and thermal images will also be investigated. In general, a fusion of technologies could lead to better segmentation results: micro-thermal sensors, ultrasonic waves, laser scanning thermography, etc., could also be integrated in order to achieve a full range of inspection of marble slabs.

In recent years, DL models have been established in visual inspection. Many DL models can be employed to enhance the segmentation performance; yet, high performance and fast implementations need to balance for real-time applications. Therefore, a future research direction is toward the investigation of robust DL model combinations that could accomplish both efficiency and accuracy. Next generation of computing technologies, characterized by quantum computing [26], is expected to provide fast computing capabilities for real-time visual inspection.

5 Conclusions

In this work, a performance evaluation of 112 DL segmentation architectures takes place, combining four DL models with 28 feature extraction networks, based on thermal and color imaging to detect cracks on marble slabs. Experimental results are evaluated based on two perspectives: based on the DL model and based on the backbone network. Results indicate that DL models can perform similarly on both thermal and RGB images, reporting FPN as the best-performing model, with 71.61 and 68.07% mIoU, for RGB and thermal images, respectively. Regarding the backbone network, results indicated as best-performing backbone family the EfficientNet with efficientnetb4, with 80.07 and 75.49% mIoU for RGB and thermal images, respectively.

Acknowledgements This research has been co-financed by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH—CREATE—INNOVATE (project code: T2EAK-00238).

References

1. Luque A, Ruiz-Agudo E, Cultrone G, Sebastián E, Siegesmund S (2011) Direct observation of microcrack development in marble caused by thermal weathering. *Environ Earth Sci* 62:1375–1386. <https://doi.org/10.1007/s12665-010-0624-1>
2. Ren Z, Fang F, Yan N, Wu Y (2022) State of the art in defect detection based on machine vision. *Int J Precis Eng Manuf Technol* 9:661–691. <https://doi.org/10.1007/s40684-021-00343-6>
3. Karaali İ, Eminağaoğlu M (2020) A convolutional neural network model for marble quality classification. *SN Appl Sci* 2:1733. <https://doi.org/10.1007/s42452-020-03520-5>
4. Ouzounis A, Sidiropoulos G, Papakostas G, Sarafis I, Stamkos A, Solakis G (2021) Interpretable deep learning for marble tiles sorting. In: *Proceedings of the 2nd international conference on*

- deep learning theory and applications. SCITEPRESS—Science and Technology Publications, pp 101–108. <https://doi.org/10.5220/0010517001010108>
5. Vrochidou E, Sidiropoulos GK, Ouzounis AG, Lampoglou A, Tsimperidis I, Papakostas GA, Sarafis IT, Kalpakis V, Stankos A (2022) Towards robotic marble resin application: crack detection on marble using deep learning. *Electronics* 11:3289. <https://doi.org/10.3390/electronics11203289>
 6. Lanzetta M, Tantussi G (1997) The quality control of natural materials: defect detection on Carrara marble with an artificial vision system. In: A.I.Te.M III, Proceedings of the 3rd conference of the italian association of mechanical technology. Fisciano Salerno, Italy, pp 449–456
 7. Sipko E, Kravchenko O, Karapetyan A, Plakasova Z, Gladka M (2020) The system recognizes surface defects of marble slabs based on segmentation methods. *Sci J Astana IT Univ* 1:50–59. <https://doi.org/10.37943/AITU.2020.1.63643>
 8. Akosman SA, Oktem M, Moral OT, Kilic V (2021) Deep learning-based semantic segmentation for crack detection on marbles. In: 2021 29th signal processing and communications applications conference (SIU). IEEE, pp 1–4. <https://doi.org/10.1109/SIU53274.2021.9477867>
 9. Yang J, Wang W, Lin G, Li Q, Sun Y, Sun Y (2019) Infrared thermal imaging-based crack detection using deep learning. *IEEE Access* 7:182060–182077. <https://doi.org/10.1109/ACCESS.2019.2958264>
 10. Solakis (2023) Solakis Marble Enterprises. <https://www.solakismarble.com/>. Last Accessed 04 Apr 2023
 11. Intermek (2023) Intermek. <https://www.intermek.gr/en/>. Last Accessed 04 Apr 2023
 12. IHU (2023) International Hellenic University. <https://www.ihu.gr/en/enhome>. Last Accessed 04 Apr 2023
 13. Seek Thermal: CompactXR (2023). <https://www.thermal.com/compact-series.html>. Last Accessed 04 Apr 2023
 14. Russell BC, Torralba A, Murphy KP, Freeman WT (2008) LabelMe: a database and web-based tool for image annotation. *Int J Comput Vis* 77:157–173. <https://doi.org/10.1007/s11263-007-0090-8>
 15. Zuiderveld K (1994) Contrast Limited adaptive histogram equalization. In: *graphics gems*. Elsevier, pp 474–485. <https://doi.org/10.1016/B978-0-12-336156-1.50061-6>
 16. An N, Xie J, Zheng X, Gao X (2015) Application of PCA in concrete infrared thermography detection. In: *Proceedings of the 2015 2nd international workshop on materials engineering and computer sciences*. Atlantis Press, Paris, France. <https://doi.org/10.2991/iwmcecs-15.2015.160>
 17. Lin T-Y, Dollar P, Girshick R, He K, Hariharan B, Belongie S (2017) Feature pyramid networks for object detection. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR). IEEE, pp 2117–2125. <https://doi.org/10.1109/CVPR.2017.106>
 18. Chaurasia A, Culurciello E (2017) LinkNet: exploiting encoder representations for efficient semantic segmentation. In: 2017 IEEE visual communications and image processing (VCIP). IEEE, pp 1–4. <https://doi.org/10.1109/VCIP.2017.8305148>
 19. Zhao H, Shi J, Qi X, Wang X, Jia J (2017) Pyramid scene parsing network. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR). IEEE, pp 6230–6239. <https://doi.org/10.1109/CVPR.2017.660>
 20. Ronneberger O, Fischer P, Brox T (2015) U-Net: convolutional networks for biomedical image segmentation. In: *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)*. pp 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
 21. Liu Y (2023) DeepCrack. <https://github.com/yhlleo/DeepCrack>. Last Accessed 04 Apr 2023
 22. Yeung M, Sala E, Schönlieb C-B, Rundo L (2022) Unified focal loss: generalising dice and cross entropy-based losses to handle class imbalanced medical image segmentation. *Comput Med Imaging Graph* 95:102026. <https://doi.org/10.1016/j.compmedimag.2021.102026>
 23. Cheng X, Cheng S (2022) Infrared thermographic fault detection using machine vision with convolutional neural network for blast furnace chute. *IEEE Trans Instrum Meas* 71:1–9. <https://doi.org/10.1109/TIM.2022.3218326>

24. Liu F, Liu J, Wang L (2022) Asphalt pavement crack detection based on convolutional neural network and infrared thermography. *IEEE Trans Intell Transp Syst* 23:22145–22155. <https://doi.org/10.1109/TITS.2022.3142393>
25. Chen C, Chandra S, Seo H (2022) Automatic pavement defect detection and classification using RGB-thermal images based on hierarchical residual attention network. *Sensors* 22:5781. <https://doi.org/10.3390/s22155781>
26. Tziridis K, Kalampokas T, Papakostas GA (2023) Quantum image analysis—status and perspectives. In: El-Alfy E-SM, George Bebis MZ (eds) *Intelligent image and video analytics*. 1st edn. CRC Press, Boca Raton. <https://doi.org/10.1201/9781003053262>