Vision-based Inspection System for Ornamental Stone Using a Weighted Hybrid Ensemble Classifier

Liliana Antão*, Rui Pinto*, Manuel J. Ferreira†, Tiago Pinto†, Gil Gonçalves*
*Research Center for Systems and Technologies
Faculty of Engineering, University of Porto
Rua Dr. Roberto Frias, 4200-465 Porto, Portugal
Email: { lpsantao, rpinto, gil }@fe.up.pt
†Neadvance – Machine Vision
Rua do Parque Poente, Lote 35 4705-002 Sequeira, Braga, Portugal
Email: { mferreira, tpinto }@neadvance.com

***Abstract*—With the increasing demands of customers in the ornamental stone industry, both in terms of the individual specifications of each product and in the delivery times, it is necessary to constantly adapt the manufacturing processes and their inherent complexity and, consequently, the automated systems that are essential to them. There is a strong movement of research in areas capable of generating non-destructive testing techniques applied to production systems in this sector. Currently, one of the main problems occurs during the ornamental stone slab polishing phase, where there is the need to monitor the polishing quality and diagnose possible defects in the surface of the slab. This can be used as feedback for self-correction and optimization of variables and process parameters in the polishing equipment. In this paper is proposed a monitoring system, based on machine vision techniques, used to detect defects in the surface of polished ornamental stone slabs. This approach is based on a weighted hybrid ensemble classifier, relying on image processing techniques and a Convolutional Neural Network. Results show that the ensemble classifier outperforms related classifiers.**

***Index Terms*—Ornamental Stone, Ensemble Classifier, Convolutional Neural Network, Machine Vision.**

I. Introduction

Regarding the natural and ornamental stone industrial sector, there is a positive response of ornamental stone manufacturing companies, especially in Portugal, in incorporating Industry 4.0 related practices and technologies in their production [1]. This allows them to enhance and achieve added value in this significant industrial sector since the transformation of rock minerals has enormous importance in the Portuguese market. In the stone extraction and transformation industry, typically, the first step marks the extraction of large blocks from quarries. There are several types of ornamental stone, such as marble, granite, limestone, among others. Then, the cutting and sawing process transforms these blocks into slabs. These slabs go through a polishing process to remove the cutting process's imperfections and restore brightness to the slab. Later, the slabs can suffer a final cut for smaller slabs or tiles, which will later be used in pieces, such as countertops and kitchen or bathroom tops. Finally, the process culminates in the packaging of the final product.

Ornamental stone manufactured goods, such as countertops and kitchen or bathroom tops, must satisfy specific aesthetic requirements, namely stone surface status or its reflecting properties, which can be accomplished by the polishing process. According to Bonifazi & Marinelli [2], there is not a defined way to know what constitutes a good polishing process. This greatly depends on the type of stone, its mineral composition, and textural attributes. Such judgment relies on human expertise. It can be influenced by several characteristics, such as mineral grains size and their relative arrangement, background color, presence of veins and plagues, cultural level of the human inspectors, and final destination of the stone manufactured good. Fig. 1 represents examples of ornamental stone manufactured goods.

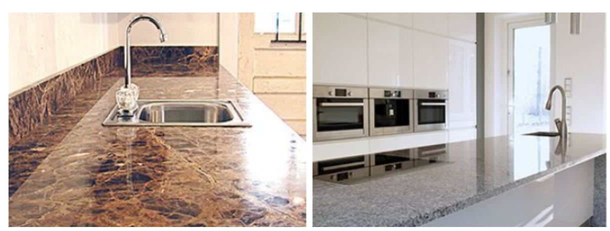

Fig. 1. Ornamental stone manufactured goods, such as kitchen/bathroom tops.

Mathielo & Bolonini [3] conducted a recent study in Brazil, where they concluded that 77% of companies in the ornamental stone sector perform analysis of the quality of the polished surfaces of stone via visual inspection, only 3% do this by measuring the brightness surface and 20% use both. This demonstrates the subjective character and the possible incompatibility between the analyzes of different human inspectors regarding the polishing quality of the finished surface. In this sense, it is necessary to adapt the stone manufacturing processes and their inherent complexity. Consequently, one needs to adapt the automated systems that are essential to them, namely automated methods to evaluate the efficiency of the polishing process in ornamental stone. Monitoring using sensors is a common method of recognition and adaptation to the environment by automatic systems, and it can be assumed

 ISBN: 978-1-61208-882-2

that computer vision is one of the most versatile technologies for this purpose [4].

Computer vision is based on the interpretation of images captured by non-invasive optical sensors, in order to extract useful information from a real scenario for analysis and/or process control. This technology allows identification and analysis of geometries, position detection, supervision, quality control, and measurement (speed, deformations, temperature, etc.). As a non-contact measuring system, it does not suffer considerable wear and can operate continuously even in hostile environments. For all these reasons, it is understandable that such a solution is of interest in a wide range of applications, such as closed-loop control processes [5].

In addition to computer vision, there is a research hype in areas capable of generating other types of Non-Destructive Testing (NDT) techniques for monitoring purposes, applied to the several stages in the ornamental stone production process. Montiel-Zafra *et al.* [6] present an impact-echo method to analyze the internal quality of ornamental stone blocks right after being extracted from the quarry.On the other hand, already on the polishing stage, Maria *et al.* [7] propose a wavelet technique to detect defects using images of ornamental limestone slabs.

Considering computer vision-related approaches for monitoring and defect detection, most of the literature work focuses on exploring Machine Learning (ML) classification techniques. This work intends to address the application of computer vision with ML for monitoring the quality of a stone slab polishing process by detecting defects on the stone's surface. The polishing defects may be seen as the presence of scratches and irregularities in the stone surface, which could be introduced by the polishing process itself or were already present in the stone, and the polishing process could not remove them. In this sense, we propose an approach, based on computational vision, to monitor complex defects and errors of high detection complexity during the ornamental limestone polishing process. This approach allows the verification and analysis of defects in polished ornamental limestone in a non-destructive way. This classification can later be used as feedback for self-correction for the regulation/optimization of the polishing variables and process parameters.

This paper is organized into four more sections. Section II provides the state of art about machine vision architectures and techniques for defect inspection in ornamental stone. Section III provides a detailed characterization of the proposed approach, a weighted hybrid ensemble classifier based on image processing and a Deep Learning model. Section IV discusses the experimental results achieved. Finally, Section V concludes the paper, stating final remarks about the work presented and provides orientations for future work.

II. Related Work

Tantussi & Lanzetta [8] provided a review about optical methods for stone surface inspection. Authors mention four main non-contact surface inspection methods, namely optical profilometry, glossmetry, riflectometry, and artificial vision. This section focuses on the literature review regarding artificial vision-based approaches for automated inspection of ornamental stone surface for defects.

It was mentioned before the difficulty to assess the quality of a stone surface due to the subjective nature of this classification, mostly by considering visible pictorial attributes and overall aesthetic features. Efforts have been made to propose methodologies and techniques able to quantify the aesthetic quality level of stone-based manufactured goods [9]. In this approach, Bonifazi *et al.* tried to assess the quality of the stone by evaluating the degree of polishing of the stone surfaces and the presence of defects, using image processing techniques. They also classify the main defects being: 1) Grooves; 2) Fissures and/or holes, and 3) Mineral inclusions. On the other hand, Yarahmadi *et al.* [10] present a new approach for quantifying the quality of stone products in quarries and processing plants.

In 2005, Lee *et al.* [11] propose an automated process for inspection of polished stone in order to detect process-induced defects, i.e., tooling marks induced by the polishing machine. The detected defects are characterized and used for adaptive control of the polishing process. The authors used the classical Circle Hough Transform (CHT) algorithm after collecting images of the surface of the stone since polishing tooling marks and scratches are distinguishable from natural flaws, characterized by their circular geometric form. So, this becomes a circle detection type of problem. The algorithm was tested using a range of images presenting defects encountered in the inspection of polished stone surfaces, obtaining good results. Issues presented were the optimization of the probability estimation, accuracy, and the relationship between speed and the proportion of edge pixels belonging to circular features to the total number present.

In 2012, Bianconi *et al.* [12] proposed an expert system for automatic classification of granite tiles through computer vision. The authors experimented with several classifiers (supervised ML) using a dataset of images. Classification takes into consideration both color and texture. Results show that the methods considered provide high granite classification accuracy, while Support Vector Machines (SVM) outperforms other methods. In 2013, Martínez *et al.* [13] proposed an automated classification approach, using ML techniques based on numeric variables obtained from 2D and 3D images captured by a linear 2D camera and a 3D laser scanner. Authors implemented both supervised and unsupervised ML techniques, such as SVM and Multilayer Perceptron neural networks (MLP), and cluster analysis and self-organizing maps. Results show that the error of automated classification was lower than for manual classification.

More recently, in 2018, Iglesias, Martínez & Taboada [14] proposed an automated inspection system for examining slate slabs based on capturing data with a 3D color camera and studying slate slab traits using computer vision algorithms. The authors tested the method on real slate slabs, which were previously classified by a human expert. Results show that the laboratory prototype system performed well, as the inspection

 ISBN: 978-1-61208-882-2

algorithms were able to accurately detect the same traits as the human expert, except for surface irregularities. Also, Ramil *et al.* [15] proposed a back-propagation Artificial Neural Network (ANN) in order to obtain the rapid and reliable identification of forming minerals in granitic rocks by means of RGB images. The results obtained, though preliminary, led to a high degree of correct identification of the forming minerals for three different granitic types.

III. PROPOSED APPROACH

To obtain a vision-based system that is able to correctly classify ornamental stone slabs as defective or not, we propose an approach based on a weighted hybrid ensemble classifier. Classification of defective slabs corresponds to the detection of defects on the surface of the slab. The overview of this approach is presented in Fig. 2. The ensemble is composed of two different classifiers, one based on traditional image processing techniques, such as adaptive filters applied according to structural and statistical methods, and another one based on a Deep Learning model, namely using a Convolutional Neural Network (CNN).

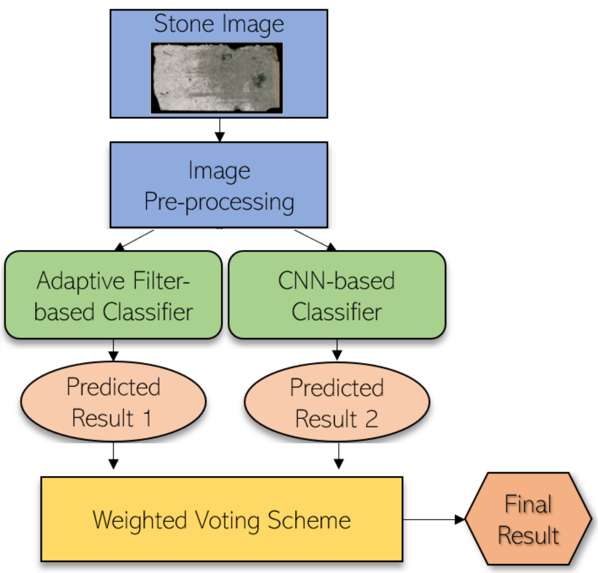

Fig. 2. Overview of the proposed hybrid classifier ensemble approach.

The inspection system starts by receiving an image of the ornamental stone's slab. This image is pre-processed in order to retrieve the biggest region of interest and feeding it as input to the classifiers. Then, each classifier obtains its predicted result separately for the input image. These predicted results are then summarized by applying a weighted voting scheme, obtaining the final result for that specific stone's surface. In this case, the result can be *OK* (no or non-significant defects) or *NOK* (significantly defective). Following Bonifazi *et al.* [9] guidelines, defect detection includes holes, fissures, decays, grooves, or deviated mineral inclusions. These are significant when occupying more than 0.05% of the region of interest of the stone surface.

A. Image Pre-processing

The first step to allow the system to correctly inspect the stone slab is provided by an image pre-processing module by converting the raw image of the stone slab into the desired input to the classifiers. Training/testing a model on raw images usually leads to bad classification performances [16], so this step helps to improve the classifier performance. Also, pre-processing techniques allow a faster training process. To allow the classification models to consider only the stone's surface, transitions between the slab and the image background or other elements present in the image are removed. Not only are those transitions not needed for this problem, but they may also cause a misdirection in the classification models since they introduce certain non-desired patterns, impacting the performance of the classifiers.

For this, it is performed segmentation of the tile to separate the stone slab surface from the image background. First, the original image is converted to grayscale, and then Gaussian filters are applied to eliminate noise while blurring the patterns and details inside the stone. Afterward, a linear threshold is applied, finally detecting the points that represent the outlines of the stone slab, using the same method proposed by [17]. To optimize memory usage, a simple approximation method can be used, which removes all redundant points and compresses the slab contours. With all the contours found, their features are evaluated, and the general external outlines of the slab stone are obtained. This step is important as one stone may appear split in more than one slab due to top to bottom fissures, and therefore return more than one external contour. In these cases, each part of the stone is evaluated separately to perceive which part needs correction, if any. This way, the pre-processing module separates the original stone image in two, allowing the classifiers to analyze the stone's defectiveness as distinct surfaces.

By using and pairing the external contour points obtained for each stone slab, it is possible to generate all rectangles that can be built using those pair of contour points. This enables to quickly reach the quadrilateral Region Of Interest (ROI), the biggest square or rectangle inside the stone slab, without any portion of the background. The ROI is then used as input to the classifiers, allowing to provide only stone surface information as input. After the initial rectangles are generated, the next set of rectangles are created by using the second point as the reference and so on.

An example of some possible rectangles is shown in Fig. 3 c). With all the rectangles generated, they are sorted by their area in descending order, being that the first rectangle in the list is the biggest one. Besides the area, additional criteria were added as sometimes the quadrilateral regions generated still included some background. Given this, a given rectangle was only selected if across its entire area there were no black dots. If any point in the perimeter of a rectangle is on a spot in which the respective pixel is 0, the respective rectangle is not valid. The rectangle that passes the criteria and has the biggest area is the desired one.

 ISBN: 978-1-61208-882-2

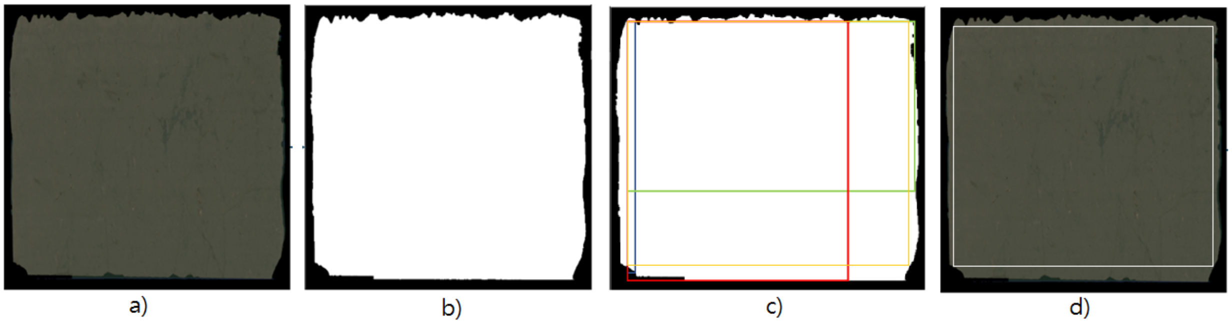

Fig. 3. Pre-processing steps: a) Original stone image, b) Image after segmentation, c) All possible rectangles; and d) Final rectangle chosen.

All the pre-processing steps performed and explained above can be seen in Fig. 3. The white square represented in Fig. 3 d) corresponds, approximately, to the biggest rectangle regarding the area, in the exemplified stone slab. The ROI is the region inside the limitation provided by the rectangle, used next as input for the classification models.

B. Weighted Hybrid Ensemble Classifier

To develop a system for inspection and detection of imperfections on the surface of the ornamental stone slabs, after obtaining the image in the desired format, the ensemble model is considered. A hybrid ensemble model is used to improve the system's performance and robustness. Compared to classifiers working individually, classifiers working together usually have the potential for a better performance.In this specific scenario, given the diversity in types of stones, this is particularly advantageous, as the limitations of one classification algorithm to a type of stone can be compensated by other classifiers.

The proposed ensemble model is composed of three main components: 1) Adaptive Filtering-based classifier that uses traditional image processing techniques with statistical and structural adaptations; 2) a Deep Learning model, particularly a Convolutional neural network (CNN-based classifier); 3) a decision-making scheme based on weighted voting. These three components are detailed in the following sections.

1) Adaptive Filtering-based classifier: Traditional filtering methods are utilized for a low cost, low level, and lightweight approach for inspection systems, as they enable to detect distinct features with no need for labeled data. Overall, these methods usually require a reference pattern without deformities, i.e., a threshold value that delimits the normal surface from the defective one. Typical filters are more suitable for images that are patterned and have periodic properties.

Considering the ornamental stone industry, where different types of stone can be extracted and used, applying a simple filtering method with static thresholds would result in poor classification performances. For instance, with marble slabs, the color and texture are typically light and homogeneous throughout its usable area, while for slate, although smooth, its texture does not follow a pattern, and its color is usually very dark. Granite, on the other hand, has a pattern of minerals along its surface, being highly variable in color and pattern (typically heterogeneous). Also, even considering the same type of stone, there may be found clear differences, according to the quarry. This difference in color and uniformity levels means that a given threshold defined as the best for one specific type of stone is not directly applicable to another. Thus, filtering methods alone are not suitable, and normally structural techniques are considered. By combining different algorithms, the defective regions can be distinguished from the stone's natural surface.

Given this, for the proposed classifier, the first step is to differentiate the basic pattern of each stone's ROI from possible defects. This step is important, as several images have innate patterns, not derived from anomalies, which could be wrongly perceived as a defect since they deviate from the normal pattern of the stone. For this a thresholding technique is applied to the image of the stone previously transformed into grayscale, using an adaptation of the Otsu method, called the valley-emphasis method [18]. In this method, the optimal threshold value is selected automatically using the images' gray-level histogram by applying the methods described in [18]. Therefore, the threshold is adapted according to the anomalies' characteristics to be isolated from each image.

After thresholding, the Sobel filter was used, followed by an oriented non-maximal suppression for edge detection. This is important to detect boundaries between the base pattern of the stone and anomalies. Then, morphological filters are applied to better identify and quantify the presence of defects on the stone surface. Each defect is identified, and its entity quantified in defect area and perimeter and the ratio of the surface area of the defect in respect to the ROI. This quantification allows checking if the detected anomaly complies with the established criteria mentioned before in Section III. If one or more defects detected pass these criteria, then the stone slab is classified as *NOK* (i.e., defective). In Fig. 4, some results from this classifier are presented. All these algorithms that comprise the adaptive filtering-based classifier are implemented in Python and OpenCV functions.

2) CNN-based classifier: As several images of stone slabs are available, a Deep Learning model is created and trained in order to classify the images as defective *NOK* or non-defective *OK*. For this, a CNN architecture is used since it is one of the best techniques for feature representation. It is

 ISBN: 978-1-61208-882-2

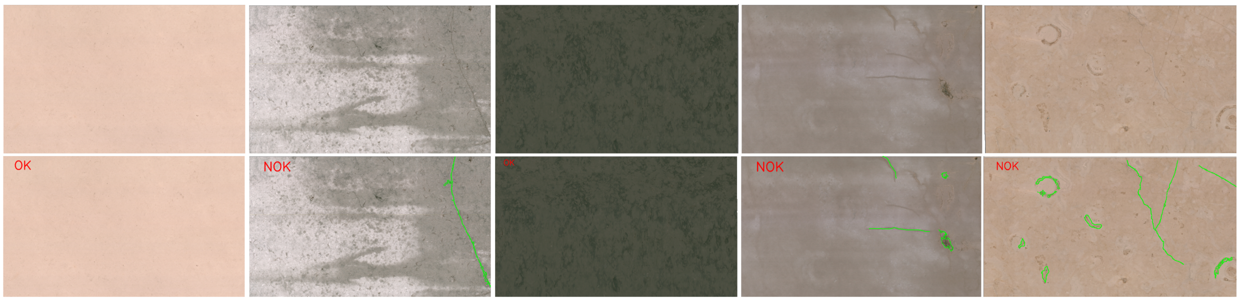

Fig. 4. Extracted ROIs of original stone images fed to the classifier (top), and corresponding classification and detection result (bottom).

easier to train than other ANN models, as they have many fewer parameters than fully connected networks with the same number of hidden units. To choose the final model, several different classifiers were considered with distinct combinations of architecture blocks and hyperparameters, randomly picked. The search space for the model was composed by four different encoders: DenseNet121, DenseNet201 [19], Inceptionv3 [20], resnet101v2 [21]. Combinations include average or max pooling, between 1 and 4 dense layers with a number of units in a Range(Start=128, End=2048, step=128). Finally, four different dropout values were tested: 0, 0.1, 0.2 or 0.3.

Given the search space defined above, each generated model with a combination of different parameters was trained three times in a dataset with a wide variety of defective and non-defective ornamental stones (more details on this dataset in Section IV). The data was divided into 70% of the images for training, 20% for validation, and the remainder used as test. The validation loss (categorial cross-entropy) was obtained from the average loss between the three runs and was used as the main metric for the model selection. The model selection was performed inside the training procedure and is divided into two stages: stage *i* that entails model elimination, where the models that performed badly were excluded; and stage *ii*, where the models resulted from the first stage are compared in performances on the test dataset.

For stage *i* of the model selection, the ten best models with the lowest average validation loss over the three trains were chosen for the final selection stage. Then, in stage *ii* of selection, those top ten models were evaluated regarding their F1-scores in the test dataset, where the final model selected had the highest score. The final model architecture achieved based on this model selection process had the following parameters: Densetnet121 as encoder, an average pooling method, and one dense layer with 1024 units and 0.1 of dropout. An Adam optimizer was used, and a learning rate of 0.0001 was selected. This final CNN was implemented in Python, using the Deep Learning framework Keras with Tensorflow backend.

3) Weighted Voting Scheme: A weighted voting ensemble is used as final decision-making for this system. Each model makes a prediction (votes) for each test instance, and the final output prediction is the one that receives more than half of the votes. In this case, only happening when both classifiers agree on the predicted result, which translates into a typical Majority voting algorithm. When the classifier models disagree, we increase the importance of one model above the other according to its performance.

This results in an approach where each model has a different significance, unlike majority voting. In this voting scheme, weights are determined according to the classification performances for each type of stone, defined by its initial characteristics, namely its heterogeneity or homogeneity. This criterion for weight setting is defined after the analysis of each classifier's performance, where it is clear that the stone surface's uniformity or lack of is the characteristic with the most impact. The scheme associates each trained classifier with a distinct weight according to its classification performance in the validation set for each type of stone. The final result for each input is done based on the highest weighted votes. If it is established that any classifier can make more confident predictions for a specific type of stone than the other, it is advisable to increase their weight to obtain more successful results.

IV. System Validation

In Portugal, the main mining district of ornamental limestones is the Maciço Calcário Estremenho (MCE). Limestones are fine to coarse-grained calciclastic sparitic rocks (rudstones and grainstones), i.e., formed by grains cemented by small amounts of translucid calcite [22]. There are several types of limestones, but the ones considered in this work are the *Cadoico Azul Mónica Silva* (CADOICO), *Salgueira Branco do Mar* (SBM) and *Salgueira Branco Real* (SBR).

In this sense, we have access to a dataset of limestone slabs' images collected right after a polishing process. The limestone slabs are first fed into polishing equipment in order to eliminate the marks resulting from the abrasive cutting processes. Fig. 5 represents a polishing equipment, namely the *StonePOLISH* model, produced by CEI [23]. These slabs may have different dimensions, with a maximum of 3.5 meters in length and 2.5 in height. After this process, the plate is digitized by a scanner that reproduces the image as the plate moves on the carpet, saving the images on the polisher's own computer. These images are captured in an RGB color scheme and saved in .jpg format, and occupy between 1MB to 3MB of disk space; most have a size of 2800 by 1500. Also, all images are collected in one dataset (CADOICO, SBM, and

 ISBN: 978-1-61208-882-2

SBR), in a total of 954 copies. In this case, the classification of limestone categories is done manually, but there are efforts to automate this process [24]. From 954 images, 707 were labeled as normal and 247 as defective.

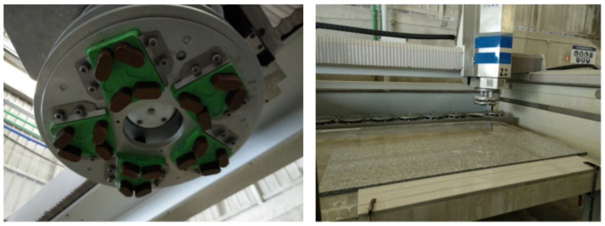

Fig. 5. Example of a polishing equipment.

Data augmentation was performed by applying rotations to the original images with 90º, 180º, and 270º, resulting in a total of 3816 images. This step was implemented to improve the CNN-based classifier performance, as data augmentation is widely recommended for image classification [25]. This augmented dataset was then split into train, validation, and test sets, with the same percentages mentioned before in Section III-B2. This division was needed for the CNN-based model, as only this classifier needs training and validation. For this training, an image input size of 400 was defined due to computational resources' limitations.

The proposed ensemble method was evaluated on the test set of the augmented limestone dataset mentioned before over standard performance metrics. From the three types of limestone mentioned, two levels of surface uniformity (calculated as detailed in Section III-A) were found: high homogeneity/uniformity for SBM and SBR (uniformity values bellow 9) and low uniformity/ high heterogeneity for CADOICO (uniformity values above 10). As the model's performance is highly impacted by this uniformity level, the performance metrics were retrieved for three variations of the dataset: (a) augmented homogeneous dataset (SBM and SBR augmented data); (b) augmented heterogeneous dataset (CADOICO augmented data); and (c) the complete augmented dataset (CADOICO, SBM, and SBR).

Also, to measure the performance improvement, the proposed model was compared with the isolated classifiers. A common metric like accuracy is suitable to assess the performance in defect detection but does not allows assessing if the classifier is good to distinct instances of *OK* and *NOK*. The ROC curve can be used for this, as it is a popular method for performance evaluation, plotting a graph with true positive rates over false positive rates. This way is possible to describe in a unique metric the trade-off between accurately classified positives and incorrectly classified negatives. The Area under the ROC curve (AUC) has been suggested as a robust classification performance metric, independent of the imbalance rate of the dataset, and can be used to compare performance between models completeness.

Given this, for validation of the classifier and comparison purposes, we have used AUC as the main evaluation metric.

TABLE I
AUC ON THE DIFFERENT AUGMENTED DATASETS WITH ADAPTIVE FILTERING, CNN-BASED, AND HYBRID WEIGHTED ENSEMBLE CLASSIFIER.

Datasets	(a) Homogeneous	(b) Heterogeneous	(c) Complete
	AUC	*AUC*	*AUC*
Adaptive Filtering-based	96.91%	84.84%	88.42%
CNN-based	93.3%	93.2%	95.8%
Hybrid Weighted Ensemble	96.96%	93.2%	96.04%

Table I shows the comparison of experimental results between each classifier alone (Adaptive filtering-based classifier and CNN-based classifier) and the proposed hybrid weighted ensemble method.

The AUC for an ideal and inaccurate model has values of 1 and 0.5, respectively. Given this, the average scores show that all the models perform accurately as well as exact and completeness. By comparing the AUC in the different variations of the dataset, it is clear that the adaptive filtering classifier works much better for homogeneous stones than for heterogeneous. This is mainly due to the classifier's difficulty with differentiating non-uniform patterns from defects, leading to false positives. Regarding the CNN-based model, although the AUC is worse than the filtering approach for the homogeneous stones, it is far more stable throughout all datasets, with less than 1% in AUC difference. When compared to the single models, the proposed ensemble model maintains the best performance from the isolated models, except in the total dataset where the AUC is increased.

The ensemble method's AUC in the homogeneous test stones outperforms the filtering model performance by a very small margin, indicating that the CNN-based model likely provided small betterment in classification when compared to the classification provided by the filtering approach alone. Conversely, in the heterogeneous dataset, the adaptive filtering model is not able to compensate for any CNN model's shortcomings, so the ensemble achieves the best possible result, the CNN's AUC. The ensemble model only shows significant improvements in the complete dataset. This is also where the applicability of this approach is more important. This increase in performance is derived from the compensation of the weak points of each isolated model.

As the weights given to each model's prediction derive from the stone's uniformity retrieved in pre-processing, more significance is given to the CNN in case of a low uniformity level and more to the filtering approach in case of a high uniformity value. This ensemble method also has the advantage of not utilizing extremely complex models that usually need more time in training and multiple parameter tuning, allowing for a lighter and simpler approach that achieves good performance while combining methods from different natures. By combining a supervised method with an adaptive filtering approach, the system is also more capable of correctly inspecting new types of stones or defects that may appear in the future when compared to a typical supervised approach.

V. Conclusions and Future Work

Considering the ornamental stone industrial sector, the automated detection of process-induced defects in stone slabs during the polishing phase is a common and important problem. Currently, most of the polishing quality inspection is performed manually by human experts. This inspection is very subjective since it is based on aesthetic properties. We propose an automated monitoring system based on machine vision to assist human operators with the quality of a polishing process. This monitoring system relies on a weighted hybrid ensemble classifier, which classifies polished ornamental stone slabs as *NOK* or *NOK*. Tests were performed to classify the polishing quality in limestone slabs by using a dataset of images collected after a polishing process. By inspection of results, it is possible to conclude that the proposed approach outperforms isolated classifiers in the same conditions.

However, there are some limitations in this study. For starters, the proposed solution may not be considered a truly online monitoring system since tests were performed in a dataset of images, where ideally, they should be performed using images collected in real-time during the polishing process of stone slabs. Secondly, the dataset considered is limited to limestone slabs from a specific Portuguese region, which reduces the flexibility of an automated monitoring approach. Thirdly, the proposed approach considers only the classification of the stone slabs. It is out of scope the automated classification feedback for self-correction of the polishing process. Finally, the stone slab classification in this study does not consider some important parameters used manually for quality inspection, such as the final destination of the stone manufactured good and the requirements from a specific client in that product.

As future work, we will deploy and validate the solution in real polishing equipment for real-time monitoring. Also, considering all the possible aesthetic requirements related to the stone manufactured good, we are considering an active or reinforcement learning approach to be considered in the ensemble approach.

Acknowledgment

This research was supported by the project PRODUTECH-SIF - Solutions for the Industry of the Future, financed by the Portuguese National program COMPETE 2020 and by Portuguese funds through FCT – Fundação para a Ciência e a Tecnologia (UID/EMS/00712/2013).

References

[1] A. da Silva and I. Almeida, "Towards industry 4.0— a case study in ornamental stone sector," *Resources Policy*, vol. 67, p. 101672, 2020.

[2] G. Bonifazi and S. Marinelli, "Ornamental stone finished product aesthetic inspection and characterization through a digital spectrophotometric approach," in *Machine Vision Applications in Industrial Inspection XI*, vol. 5011. International Society for Optics and Photonics, 2003, pp. 243–250.

[3] J. G. Mathielo and T. M. Bolonini, "Diagnóstico do processo de polimento de rochas ornamentais," 2016.

[4] P. Nerakae, P. Uangpairoj, and K. Chamniprasart, "Using machine vision for flexible automatic assembly system," *Procedia Computer Science*, vol. 96, pp. 428–435, 2016.

[5] X. Su and Q. Zhang, "Dynamic 3d shape measurement method: a review," *Optics and Lasers in Engineering*, vol. 48, pp. 191–204, 2010.

[6] V. Montiel-Zafra, F. Canadas-Quesada, M. Campos-Suñol, P. Vera-Candeas, and N. Ruiz-Reyes, "Monitoring the internal quality of ornamental stone using impact-echo testing," *Applied Acoustics*, vol. 155, pp. 180–189, 2019.

[7] M. C. Proença, M. Aniceto, P. N. Santos, and J. C. Freitas, "Automatic detection of defects in ornamental limestone using wavelets," *International Journal of Computer and Information Engineering*, vol. 10, no. 5, pp. 864–869, 2016.

[8] G. Tantussi and M. Lanzetta, "Analyses of stone surfaces by optical methods," in *AI Te. M 2007, 8th Conference of the Italian Association of Mechanical Technology*, 2007, pp. 100–128.

[9] G. Bonifazi, A. Gargiulo, S. Serranti, and C. Raspi, "Imaging-based logics for ornamental stone quality chart definition," in *Machine Vision Applications in Industrial Inspection XV*, vol. 6503. International Society for Optics and Photonics, 2007, p. 65030P.

[10] R. Yarahmadi, R. Bagherpour, S.-G. Taherian, and L. M. Sousa, "A new quality factor for the building stone industry: a case study of stone blocks, slabs, and tiles," *Bulletin of Engineering Geology and the Environment*, vol. 78, no. 1, pp. 533–542, 2019.

[11] J. Lee, M. L. Smith, L. N. Smith, and P. S. Midha, "Robust and efficient automated detection of tooling defects in polished stone," *Computers in industry*, vol. 56, no. 8-9, pp. 787–801, 2005.

[12] F. Bianconi, E. González, A. Fernández, and S. A. Saetta, "Automatic classification of granite tiles through colour and texture features," *Expert Systems with Applications*, vol. 39, no. 12, pp. 11 212–11 218, 2012.

[13] J. Martínez, M. López, J. M. Matías, and J. Taboada, "Classifying slate tile quality using automated learning techniques," *Mathematical and Computer Modelling*, vol. 57, no. 7-8, pp. 1716–1721, 2013.

[14] C. Iglesias, J. Martínez, and J. Taboada, "Automated vision system for quality inspection of slate slabs," *Computers in Industry*, vol. 99, pp. 119–129, 2018.

[15] A. Ramil, A. López, J. Pozo-Antonio, and T. Rivas, "A computer vision system for identification of granite-forming minerals based on rgb data and artificial neural networks," *Measurement*, vol. 117, pp. 90–95, 2018.

[16] K. K. Pal and K. Sudeep, "Preprocessing for image classification by convolutional neural networks," in *2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*. IEEE, 2016, pp. 1778–1781.

[17] S. Suzuki *et al.*, "Topological structural analysis of digitized binary images by border following," *Computer vision, graphics, and image processing*, vol. 30, no. 1, pp. 32–46, 1985.

[18] H.-F. Ng, "Automatic thresholding for defect detection," *Pattern recognition letters*, vol. 27, no. 14, pp. 1644–1649, 2006.

[19] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[20] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[22] J. M. Carvalho, J. V. Lisboa, A. Casal Moura, C. Carvalho, L. M. Sousa, and M. M. Leite, "Evaluation of the portuguese ornamental stone resources," in *Key Engineering Materials*, vol. 548. Trans Tech Publ, 2013, pp. 3–9.

[23] CEI, "Stone machinery - stonepolish," https://www.ceigroup.net/stonepolish, 2017.

[24] M. Tereso, L. Rato, and T. Gonçalves, "Automatic classification of ornamental stones using machine learning techniques a study applied to limestone," in *2020 15th Iberian Conference on Information Systems and Technologies (CISTI)*. IEEE, 2020, pp. 1–6.

[25] A. Mikołajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," in *2018 international interdisciplinary PhD workshop (IIPhDW)*. IEEE, 2018, pp. 117–122.