




RESEARCH ARTICLE

Granite classification using machine learning and edge computing [version 1; peer review: 1 approved, 2 approved with reservations]

Madhavi Karanam¹, Krishna Chythanya Nagaraju ², Gotham Sai P³, SaiKiran Manasa S³, Pranay Krishna G³

¹Professor, HoD, CSE,, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, Telangana, 500090, India

²Asst.Prof., CSE, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, Telangana, 500090, India

³UG STUDENT, CSE., Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, Telangana, 500090, India

V1 First published: 08 Nov 2022, 11:1276
<https://doi.org/10.12688/f1000research.124057.1>

Latest published: 08 Nov 2022, 11:1276
<https://doi.org/10.12688/f1000research.124057.1>

Abstract

Background: The outlook and the aura of any place are highly dependent on how a place is decorated and what materials are used in designing it. Granite is such a kind of rock which is vastly used for this purpose. Granite flooring and counters have a major influence on the interior d'ecor which is essential to set the mood and ambience of a house. A system is needed to help the end users differentiate between granites, which enhance the grandeur of their house and also check the frauds of different color granite being sent by the merchant as compared to what was selected by the end user. Several models have been developed for this cause using CNN and other image processing techniques. However, a solution for this purpose must be precise and computationally efficient.

Methods: For this purpose, researchers in this work developed a machine learning based granite classifier using Edge Computing and a website to help users in choosing which granite would go well with their d'ecor is also built. The developed system consists of a color sensor [TCS3200] integrated with an ESP8266 board. The data pertaining to RGB contrasts of different rocks is acquired by using the color sensor from a dealership. This data is used to train a Machine Learning algorithm to classify the rock into different granite types from a granite dealer and yield the category prediction.

Results: The proposed system yields a result of 94% accuracy when classified using Random Forest Algorithm.




Conclusion: Thus, this system provides an upper hand for the end users in differentiating between different types of granites.

Keywords

Granite, Machine Learning, Edge Computing, Random Forest, SVM, TCS3200, ESP8266

Open Peer Review

Approval Status   

	1	2	3
version 1 08 Nov 2022	 view	 view	 view

1. **Ratan Lal** , Northwest Missouri State University, Maryville, USA
2. **Blessing Olamide Taiwo** , Federal University of Technology Akure, Akure, Nigeria
3. **Zhi Yu** , Fuzhou University, Fuzhou, China

Any reports and responses or comments on the article can be found at the end of the article.



This article is included in the **Artificial Intelligence and Machine Learning** gateway.



This article is included in the **Computational Modelling and Numerical Aspects in Engineering** collection.

Corresponding author: Krishna Chythanya Nagaraju (nkrishna1397@grietcollege.com)

Author roles: **Karanam M:** Formal Analysis, Project Administration, Supervision, Validation, Writing – Review & Editing; **Nagaraju KC:** Conceptualization, Methodology, Resources, Supervision, Writing – Original Draft Preparation; **P GS:** Data Curation, Formal Analysis, Investigation, Resources, Software, Writing – Original Draft Preparation; **S SM:** Formal Analysis, Software, Validation, Visualization; **G PK:** Software, Validation, Visualization, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The author(s) declared that no grants were involved in supporting this work.

Copyright: © 2022 Karanam M *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Karanam M, Nagaraju KC, P GS *et al.* **Granite classification using machine learning and edge computing [version 1; peer review: 1 approved, 2 approved with reservations]** F1000Research 2022, 11:1276 <https://doi.org/10.12688/f1000research.124057.1>

First published: 08 Nov 2022, 11:1276 <https://doi.org/10.12688/f1000research.124057.1>

Introduction

Granite is a coarse-grained igneous rock. It is formed when magma is compressed due to the pressure underground. It is one of the most common rocks found on earth crust and constitutes a multimillion dollar granite industry. Granite owing to its beauty and composition, it is used in a wide range of applications which encompass house decor, flooring, counter tops etc. This flexibility of granite has instigated many people to cheat with the quality and type of granite. With the advancing technologies, like machine learning and Internet of Things (IOT) the verification and analysis of such samples has become an easy task. These technologies have made it possible to go about any particular task in an efficient and remote manner with pinpoint precision. This has led some researchers to explore the field and produce some workable solutions, models which recognize the type of granite using Convolutional Neural Network (CNN) on images of granite. These models classify the granites based on image patterns of different types of granite.¹

Models which use Computer Vision to classify rocks using the colors and textures of granites obtained from images. This model considers the color of granite and its texture as the distinguishing factor.² Models were developed which apply Machine Learning on data gathered using a spectrophotometer. These models gather data from a spectrophotometer and then apply Machine Learning to differentiate between granite types.³

The solutions currently developed all rely on heavy computing techniques and are not portable. A solution which can be executed from anywhere at any time needs to be developed to better cater the needs of end users. To meet the demands of end users and make the system more portable and efficient, a model with Machine Learning and Edge Computing is experimented in this work. The increasing standards of living and per capita income have also inspired people to try different styles and experience various luxuries, primarily the looks of their houses. This sudden inflow of income and need for granite has instigated some merchants to fraud the end users by charging for a higher quality rock and issuing a lower quality rock. With these frauds on the incline, the financial situations are on a decline. It's high time that we develop a method to verify the quality of the rock and terminate the frauds. The end users also often seek suggestions to choose from the plethora of granites that are available in the market. It is important now more than ever to devise a solution to corroborate the granite received in order to prevent the fraud efficiently and effectively without compromise on the portability. The integration of ML and Edge Computing is an effective way to do the required task much more easily.

This project utilizes the color variations of granites to classify them into various categories by collecting the color values from separate points on the rock and applying Machine Learning on them. The major objective of this work is to verify the rock delivered based on color.

Previous work

The homogeneity of pieces of granites being used based on color is very important for the aesthetic of place where granites are being used. In general humans carry out the classification process based on quality and look using human experience and exposure along with a bit of creativity. This opens doors for artificial intelligence based works to be carried out in this field in selection of a specific color granite.

Reference 1 work follows convolution neural network usage for granite classification. The researchers used transfer learning on MNSIT networks and CIFAR networks on a dataset of 1000 RGB images. Using Nearest Neighbor classifier with CIFAR they claim to achieve an accuracy of above 85% in classifying granites.

In the work² an expert system was developed based on computer vision concepts to classify granites. The authors claim that Support Vector Machine out performs than other methods they tried in classifying granites based on color and texture. The classification of granites using automatic artificial intelligence based systems would provide benefits of creating a repeatable procedure which is not possible in manual general process followed for decades in market. This further add to substantial upgrading in quality assessment procedures for companies. Also, lessening of losses owing to cancellation of order at customer site besides easy warehouse management can be achieved. The work was carried out using lot of 48 granites of twelve classes on total which were subdivided in to 64 samples per class and images were captured using a standard procedure in lab but finally the cropped image of "544×544 pixels" was used in experimentation. They used methods based on co occurrence matrix, Gabor features, chromatic features and Local Binary Patterns. The work made use of multiple classifiers including "Support Vector classifier", "Linear Classifier" "Naïve Bayes", "Nearest Neighbor" etc.

In some works sum and differences of histograms along with LVQ neural networks is used to classify granites.⁴ In the letter published by Ref. 5 the researchers claim to use "wavelet analysis" to classify the granites. Discrete wavelet transformation is used to generate sequences of signals in their work. The researchers made use of 30 image database having three classes of granites. The work claims to obtain 90% accuracy using LVQ neural networks.

With the work⁶ the researchers have proposed an “electro-mechanical system” that robotically categorizes marbles while on conveyor belt itself based on a “hierarchical clustering approach”. They made use of Programmable Logic Controller along with an auxiliary micro controller to bind between PLC and Matlab. Clustering based classification was implemented along with hierarchical classification based on cascaded features of color, texture and spectral. 83.6% accuracy as recorded in their work for correct classification of marbles. Reference 7 work makes use of transfer learning on AlexNet and VGGNet to develop a Convolution Neural Network based granite tiles classification using a dataset of 2000 RGB images classified into 25 classes. Out of experiments they observed that fine tuning VGGNet they could record an accuracy of almost 99.3%.

The authors of Ref. 8 have implemented their own CNN architecture and good bring out an accuracy of around 96.1% in marble quality classification. The authors of Ref. 9 have done extensive experimentation by examining fifteen varieties of convolution neural networks to sort dolomitic stone tiles. They observed that “DenseNet201” model could yield an accuracy of 83.24% while trained on a 489 digital image data set of granites classified into 3 classes.

Methods

The Model developed takes the color values of granite as input and then feeds it to a machine learning algorithm. The sensor relays data to a nodemcu, which analyses it and deploys a machine learning model to verify the granite being sampled and classifies it into a specific class of rock.

Machine learning

Machine learning is a part of Artificial Intelligence (AI) that trains a computing machine to learn from data and improve. The term machine learning depicts an automated process of pattern recognition and feature detection in any data. Machine learning also increased the comforts of living, personal assistants are now available which cater to a person’s personal needs and desires. Machine learning can be categorised as Supervised machine learning, Unsupervised machine learning, Semisupervised learning as shown in Figure 1 below.

Machine learning models require a large amount of data to be trained on for efficient prediction and increased accuracy. The data can be obtained from a variety of methods including real time data collection as done in this work. This data can be processed in a lot of ways such as normalization, removing data or filling the missing data with mean, median or mode and other ways. After processing the data, the Machine learning models are ready to be trained. The authors opted for classification model of ML to be implemented for this work as objective is to classify granites based on color. There are several Algorithms which analyze the data and predict/classify the output. Each of them works on a different principle and the ML model. Training is the part where the machine learns the data and analyses it. The Output and the efficiency majorly depend on this part. The data is trained to provide a specific output depending on the input provided. Evaluation of a model refers to the part where the model’s efficiency is tested. In this work Accuracy and Confusion matrix are considered to evaluate trained model. The overall machine learning process discussed can be depicted in figure as shown in Figure 2.

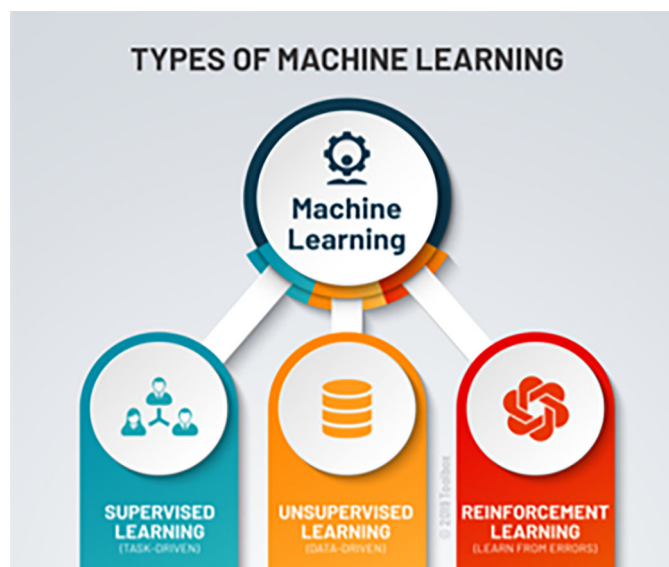


Figure 1. Types of machine learning.

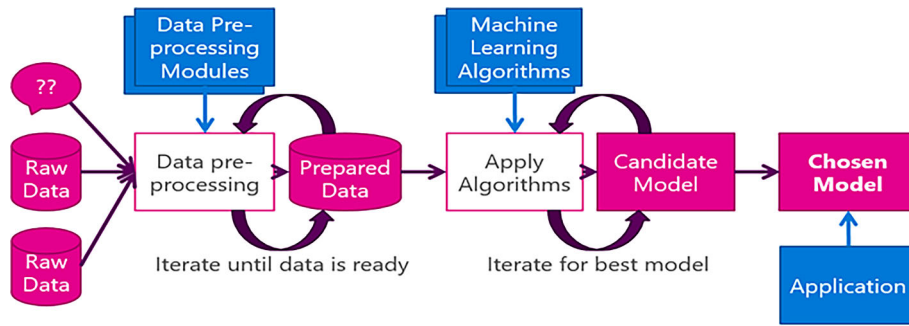


Figure 2. Machine learning process.

Edge computing refers to a distributed computing paradigm, in which the data is stored and computations are executed closer to the actual site in order to improve response time and carry the functions of any process faster. It’s more of a topology. The purpose of Edge computing is to move the entire computation of any process away from data centers and closer to the edge of the network in order to reduce the stress and load on the data center for efficient and quicker responses. It exploits the functions and specifications of smart objects like smart phones, controller boards etc., which have built-in memory to perform tasks and provide services instantaneously rather than accessing the data centers and retrieving information for every task.

Dataset

Dataset used is collected from a granite dealership with the consent of pertaining dealers in real time. This collected data is then used to train the developed IoT based machine learning model using edge computing. The Features of the dataset include RGB values of different types of granites and the last column giving the color of granite. The sample of data set collected is shown in **Figure 3** below.

The sensors were taken to a dealership and the RGB values pertaining to the different classes of granites are recorded. This data is then stored in a csv file for easier training of ML models.

Data collection

The ESP8266 and color sensor were integrated to build the experimental setup for the purpose of recording the RGB values of granites and for the collection of dataset. A dealer was contacted to get the permission for recording the values of granites and after taking necessary precautions and permission. The values of red granite, black granite and white granite were collected and tabulated as shown in **Figure 3** below. The data was collected by one of the authors of this work as can be found in image of **Figure 4** below. The dataset acquired accommodates 90 rows with 12 columns of rgb values describing 3 classes. A program was written in Arduino IDE to collect the sensor readings. The sensor gives “rgb” values

Values of Red(R)				Values for Green(G)				Values for Blue(B)				Class
302	413	302	302	413	302	301	413	301	301	413	302	white
302	413	302	302	413	302	302	413	302	302	413	301	white
302	413	302	302	413	301	301	413	301	302	413	301	white
301	413	301	302	413	302	301	413	302	302	413	302	white
302	412	302	301	412	301	301	412	301	301	413	302	white
302	414	302	302	414	303	302	414	302	301	413	302	white
302	413	302	302	413	302	302	413	302	302	414	302	white
302	413	301	302	413	302	302	413	301	302	413	302	white
302	413	302	302	413	301	302	413	301	302	412	302	white
301	414	301	302	414	302	302	414	303	302	414	302	white
302	413	301	302	414	302	302	413	302	302	413	302	white
302	413	301	302	413	301	302	413	301	302	413	302	white

Figure 3. The sample of data set used for in this experiment.



Figure 4. One of the authors collecting data using sensors.

at each point by adjusting the s0 and s1 pins. These RGB values are stored in an array. The sequence diagram of idea carried out can be seen in the figure below. The overall system architecture of the idea implemented can be seen as in **Figure 5** below. This work is implemented using Arduino IDE for edge computing, python language to perform machine learning and HTML, CSS to develop the web application. With the help of special libraries, machine learning is carried out on the Arduino platform. The Arduino (IDE) software makes the uploading of code to the NodeMCU much easier and quicker.

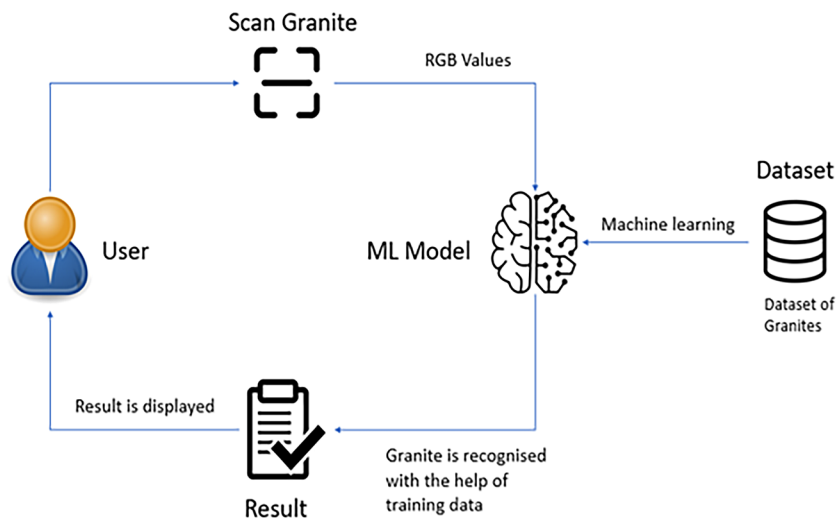


Figure 5. Architecture of the system implemented.

Python Language is used to implement the machine learning modules using Scikit Learn library which is then transferred to Arduino.

The Dataset built already is given to train machine learning models to learn. After the data is trained and models learnt the process, various algorithms are evaluated and compared to select the one with highest accuracy for classification i.e. Random forest. The Implementation of this project is done in 3 steps: 1. Data collection: collection of data from the granite dealership; 2. Model Selection: various ML algorithms were observed and a model is fixed; 3. Deployment: The model is deployed into a nodemcu for real time usage.

Various machine learning algorithms are used to classify any particular data and each one of them is dependent on a specific mathematical concept. These concepts are then applied to find a most suitable correlation with the given data or a function that best fits the data being analyzed. The algorithms that are compared for better output in this work are: Random Forests Algorithm¹⁰; Support Vector Machine Algorithm¹¹; K-Nearest Neighbors (KNN) Algorithm.¹² The algorithms are implemented in python and then converted to C for it to be executed in NodeMCU. NodeMCU is an open-source firmware, used for IoT applications on open source boards. Lua scripting language is used to implement the firmware on Espressif SDK for Esp8266. Esp8266 image is shown in figure below. A Wi-fi soc from espressif systems is available on the board. A dual in-line package provided by the prototyping hardware integrates a USB controller with an MCU and antenna laden board. UART, DAC and ADC interfaces are supported by the board and can be accessed through a specific set of pins among the 21 pins available on the board.

The deployment is done using python IDLE and Arduino IDE. Micromlgen is an open source library developed to bring machine learning to microcontrollers. It essentially converts a ML code into an optimized c code for the microcontrollers to execute it. This acts as an alternative to the Tensorflow package which is computationally complex and cannot be used on boards with smaller capacities and memories. Micromlgen supports some of the basic machine learning algorithms like SVM, Random Forests, Decision trees, Gaussian Naive Bayes, XGBoost among others. The obtained output code in C is then written in a header file which is to be included in the arduino sketch. Thus, ML is implemented in the NodeMCU effectively. The sensor is then connected to the NodeMCU and the model is loaded onto the board. The physical model gets ready and can be used to classify granites in real time.

TCS3200 color sensor contains a TAOS TCS3200 RGB sensor chip and 4 white LEDs. The image of TCS3200 color sensor is shown in the image in figure given below. The chip forms the most important module of the sensor, which is basically a color light-to-frequency converter. The sensor is arduino compatible, which makes it very useful and easy to handle. The corresponding frequency of various colors is given as output by sensing the presence of Red, Green and Blue colors. This module can be used in a variety of applications such as granite classification, color matching tests, color sorting robots, etc. The setup built for Edge Computing using ESP8266 and TCS3200 for carrying out the idea can be seen in the image of [Figure 6](#) below.

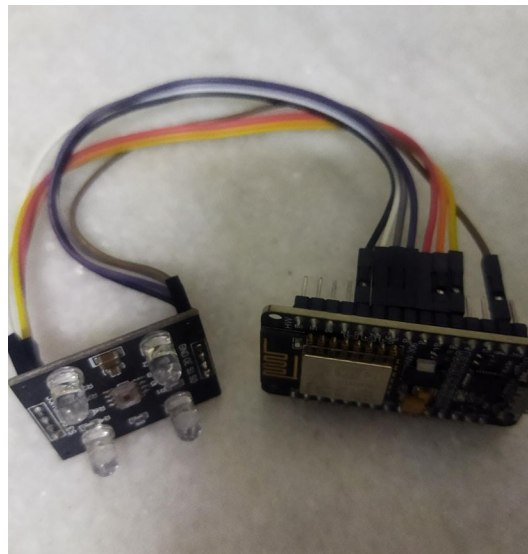


Figure 6. SetUp for Edge Computing using ESP8266 board and TCS3200 used for implementing this work.


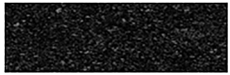

Test	Test	Sample Picture	Expected	Actual	Result Picture
1	Red Granite		Red Granite is classified	Recognized as Red Granite	Let me guess your You showed me red
2	Black Granite		Black Granite is	Recognized as Black	Let me guess your ol You showed me black
3	White Granite		White Granite is classified	Recognized as White Granite	Let me guess your ok You showed me white

Figure 7. Sample outputs for the test inputs given.

Table 1. Different algorithms tried and their accuracy.

Algorithm	Accuracy
Random forest	94%
SVM	78%
KNN	75%

The **Figure 7** above shows the sample results found during experimentation.

Conclusion

In this work the model built helps the users in choosing a granite that enhances the grandeur of their house and also reduces losses for return orders at merchant end. The model aims to help the end users in ascertaining the quality and type of granite they received. Improvising the age- old techniques of image classification, this model does the task with edge computing for increased efficiency and effectiveness. The researchers tested the dataset using several machine learning algorithms which have produced satisfactory results. Random Forest classifier gave 94% accuracy, SVM exhibited 78% accuracy and KNN could only achieve 75% accuracy. The values have been shown tabular form in **Table 1** above. Random Forest algorithm yielded the best results. The accuracy can be improved by collecting larger data.

Author contributions

Mrs. K. Madhavi was chief mentor for the work and supervised the whole research taken up with time to time inputs. She took care of reviewing the article and methodology.

Mr. Krishna Chythanya conceptualised the idea and assisted in Implementing the research with proper evaluations time to time and also writing the article.

Mr. P. Gowtham Sai was instrumental in collecting data as well as implementing the code.

Mr. S. Saikiran Manasa was contributing in algorithm implementation and evaluation.

Mr. G. Pranay Krishna was playing role during documentation and review.

Data availability

Figshare. dataset_G *raniteClassification.csv*. DOI: [10.6084/m9.figshare.20301006](https://doi.org/10.6084/m9.figshare.20301006).

This project contains the following underlying data:

- The data set has around 90 rows each consisting of 12 columns that stands of RGB values of Granite. Each of RGB is represented by 4 values respectively in a row. These values are recorded using TCS3200 color sensor on granite peices. The values were recorded by visiting couple of near by granite dealer shops in Hyderabad,

Telangana, India. The last column denotes the color of granite and this acts as a target variable in case of machine learning algorithms to predict the color of granite.

Source code

Zenodo. GRANITE CLASSIFICATION USING MACHINE LEARNING AND EDGE COMPUTING. DOI: [10.5281/zenodo.6835324](https://doi.org/10.5281/zenodo.6835324).

- The RAR file consists of three folders. Each folder has got code for each of the three machine learning algorithms implemented for Granite Classification using Edge Computing. The code has folders by name KNN, SVM and Random Forest.

References

1. Ferreira A, Giraldi G: **Convolutional neural network approaches to granite tiles classification**. *Expert Syst. Appl.* 2017; **84**: 1–11. [Publisher Full Text](#)
2. Bianconi F, González E, Fernández A, et al.: **Automatic classification of granite tiles through colour and texture features**. *Expert Syst. Appl.* 2012; **39**(12): 11212–11218. [Publisher Full Text](#)
3. Araújo M, Martínez J, Ordóñez C, et al.: **Identification of granite varieties from colour spectrum data**. *Sensors*. 2010; **10**(9): 8572–8584. [PubMed Abstract](#) | [Publisher Full Text](#)
4. Tomás-Balibrea L-M, et al.: **Automatic classification system of marble slabs in production line according to texture and color using artificial neural networks**. *International Conference on Computer Analysis of Images and Patterns*. Springer; 1999; pp. 167–174.
5. Luis-Delgado JD, Martínez-Alajarin J, Tomás-Balibrea LM: **Classification of marble surfaces using wavelets**. *Electron. Lett.* 2003; **39**(9): 714. [Publisher Full Text](#)
6. Alper Selver M, Akay O, Alim F, et al.: **An automated industrial conveyor belt system using image processing and hierarchical clustering for classifying marble slabs**. *Robot. Comput. Integr. Manuf.* 2011; **27**(1): 164–176. [Publisher Full Text](#)
7. Ather M, Khan B, Wang Z, et al.: **Automatic recognition and classification of granite tiles using convolutional neural networks (cnn)**. *Proceedings of the 2019 3rd International Conference on Advances in Artificial Intelligence*. 2019; pages 193–197.
8. Karaali I, Eminağaoğlu M: **A convolutional neural network model for marble quality classification**. *SN Appl. Sci.* 2020; **2**(10): 1–6. [Publisher Full Text](#)
9. Ouzounis AG, Sidiropoulos GK, Papakostas GA, et al.: **Interpretable deep learning for marble tiles sorting**. *DeLTA*. 2021; 101–108.
10. Ho TK: **Random decision forests**. *Proceedings of 3rd international conference on document analysis and recognition*. Vol. 1. IEEE; 1995; pp. 278–282.
11. Meyer D, Wien FT: **Support vector machines**. *R News*. 2001; **1**(3): 23–26.
12. Peterson LE: **K-nearest neighbor**. *Scholarpedia*. 2009; **4**(2): 1883. [Publisher Full Text](#)

Open Peer Review

Current Peer Review Status:   

Version 1

Reviewer Report 27 November 2023

<https://doi.org/10.5256/f1000research.136228.r225149>

© 2023 Yu Z. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Zhi Yu 

Zijin School of Geology and Mining, Fuzhou University, Fuzhou, China

1. The abstract requires revision, with the removal of specific terms such as background, methods, and results.
2. Clarify the term "d'ecor?" in the abstract.
3. Provide an introduction to the development of the machine learning model.
4. Introduce the data preprocessing steps and specify the ratio of training to testing datasets.
5. Include a comparative analysis of model performance.
6. Add sections on limitations and future scope.
7. In the introduction, highlight the limitations of prior studies and underscore the significance of this paper.
8. Enhance the quality of the figures presented in the paper.
9. Reorganize the content of the article for better flow.
10. Clarify the content of Figure 7, specifically the phrase "Lest me guess you?"

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

No

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Mining Engineering; Application of machine learning technique

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 27 November 2023

<https://doi.org/10.5256/f1000research.136228.r225152>

© 2023 Taiwo B. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The author(s) is/are employees of the US Government and therefore domestic copyright protection in USA does not apply to this work. The work may be protected under the copyright laws of other jurisdictions when used in those jurisdictions.



Blessing Olamide Taiwo

Federal University of Technology Akure, Akure, Ondo, Nigeria

Firstly, I want to thank the authors for putting this great work together and for their efforts.

The manuscript presents the development of machine learning based granite classifier using Edge Computing and a website to help users in choosing which granite would go well with their decor is also built. The authors considered color sensor [TCS3200] integrated with an ESP8266 board in the work.

I have the following comments and recommendation to improve the current state of the work.

1. Grammar and Presentation:

The work needs some rephrasing and corrections to improve the grammar and presentation. (Pg 4 last paragraph and other areas).

2. Literature Review:

The author provides some literature about previous work, but i recommend they include recent works relating to optimization algorithms, ensemble learning classifiers to provide more ground to users about the methodology used in this work.

3. Model Result validation and Data description:

I recommend that the author should provide the data in Table form and also provide the testing input dataset along with Figure 7.

The use of ROC as part of the result assessment should be included in the result section.

"**ROC curve** is a graphical representation of the performance of a binary classifier system as its discrimination threshold is varied. It is created by plotting the **true positive rate (TPR)** against the **false positive rate (FPR)** at various threshold settings."

Authors should provide the chart for both FPR and TPR for both training data only for testing data to enhance the result interpretation.

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Engineering application of Artificial intelligence

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 16 December 2022

<https://doi.org/10.5256/f1000research.136228.r157142>

© 2022 Lal R. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Ratan Lal 

Northwest Missouri State University, Maryville, MO, USA

The authors consider the problem of identifying the quality of Granite rock to help people and avoid fraud in Granite rocks. The broad approach is the following. First, a set of RGB colors for real images is collected through the color sensors. Then, the data has been used to train different machine learning models, such as Random Forest, Support Vector Machine, and K Nearest Neighbor. Upon training the model, the best among the selected models has been considered to deploy on NodeMCU, which is an open-source firmware used for IoT applications on open-source boards. The model is trained through Scikit Learn Library in python, and then it is converted into C using Micromlgen in order to deploy the ML model on the NodeMCU. Finally, the Random Forest model has 94% accuracy, and it is deployed on the NodeMCU.

Pros:

1. The idea of identifying the quality of Granite rock using IoT devices is interesting.
2. The paper is written well.

Major Comments:

1. I recommend the authors use the term "prediction" in place of "verification" as the machine learning model does not always give an accurate result.
2. It would be good if the authors could report the confusion matrix for the Random Forest model.
3. I recommend the authors remove methods:, results: heading and conclusion: paragraph from the abstract.

Minor comments:

1. On page 6, seurence -> sequence
2. On page 7, python IDLE -> python IDE

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

No

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Formal Verification, Cyber-Physical Systems, Robotics, Machine Learning, Internet of Things

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research