engineering proceedings

MDPI

Proceeding Paper

Evolutionary Image Augmentation with Genetic Algorithm for Enhancing CNN-Based Romblon Marble Pattern Recognition Model †

Marvin Rick G. Forcado [1,*] and Sivakumar Vengusamy [2]

[1] Center for Research in Artificial Intelligence and Information Technologies, Romblon State University, Odiongan 5505, Philippines
[2] Graduate School of Technology, Asia Pacific University of Technology and Innovation, Kuala Lumpur 57000, Malaysia; dr.sivakumar@apu.edu.my
* Correspondence: mforcado@rsu.edu.ph
† Presented at the 8th International Global Conference Series on ICT Integration in Technical Education & Smart Society, Aizuwakamatsu City, Japan, 20–26 January 2026.

Abstract

This study examines the effectiveness of Genetic Algorithm (GA)-based image augmentation in enhancing Convolutional Neural Network (CNN) performance for Romblon marble texture classification. Five CNN architectures, AlexNet, InceptionV3, VGG16, MobileNet, and ResNet50, were trained and evaluated on both raw and GA-augmented datasets. Unprocessed dataset achieved limited accuracy, with InceptionV3 performing best at 35.33%, followed closely by VGG16 at 33.78%. In contrast, GA-augmented data significantly boosted performance, with VGG16 achieving 94.68% accuracy, followed by MobileNet (92.68%) and InceptionV3 (92.46%). Entropy loss values consistently decreased across all models, indicating improved convergence and reduced overfitting. Although ROC-AUC scores remained close to 0.5, reflecting modest improvements in class separability, overall results confirm that evolutionary augmentation enriches dataset diversity and strengthens CNN learning capacity. MobileNet showed a solid balance between accuracy and computational economy, underscoring the possibility of GA-based augmentation as a workable option for real-world marble categorization, while VGG16 emerged as the most accurate of the studied architectures.

Keywords: convolutional neural network; data augmentation; marble classification; genetic algorithm

Check for updates

Academic Editors: Debopriyo Roy, George Fragulis and Peter Ilic

Published: 16 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Marble production has been a significant industry in the Philippines, particularly in the province of Romblon, which has long been recognized as the country's "Marble Capital" due to its vast deposits and high-quality stone. Since 2009, marble product exports have seen a consistent upward trend, with the Municipality of Romblon actively supporting exploration efforts to maximize the industry's potential [1]. According to a report by the United States Geological Survey (USGS), the Philippines produced approximately 224,786 cubic tons of marble between 2017 and 2018 [2]. Extracted marble from Romblon is meticulously processed into statues, monuments, souvenirs, and home décor, making it a key contributor to the province's economy and a valuable commodity in the global market [3]. The marble extraction process involves both manual and mechanized methods.

Workers inspect slabs using chisels, mattocks, and machinery to determine marble quality and classification [4]. However, this traditional approach heavily relies on human expertise, making it labor-intensive, time-consuming, and susceptible to misclassification errors. Selecting marble blocks that meet client specifications and quality standards is one of the most crucial stages in the manufacturing process, as misclassification can lead to significant economic losses [5,6].

Modern technologies like deep learning, pattern recognition, and convolutional neural networks (CNNs) are being investigated to automate and improve the marble categorization process in light of developments in image processing and artificial intelligence [7]. While image processing has been widely utilized in various industries, its application in marble mining and mineral classification remains relatively new [8]. CNNs, known for their ability to automatically learn discriminative features by stacking multiple layers, eliminate the need for handcrafted feature extraction, making them well-suited for marble classification tasks [9]. The integration of deep learning models into marble classification can significantly improve efficiency, enabling faster and more accurate identification of marble types, ultimately benefiting the industry by reducing errors and increasing productivity [10,11].

Numerous investigations into the application of CN models for marble classification have produced encouraging results [12,13]. However, to evaluate the performance of these models specifically for Romblon marble, a dataset consisting of high-resolution images of various marble types was collected from quarry sites in Romblon. To make model training and evaluation easier, the images were expertly labeled and divided into testing, validation, and training sets. In this study, various state-of-the-art CNN architectures were compared using Accuracy Score, F1 Score, and ROC-AUC Score to assess their effectiveness in classifying Romblon marble [6,14]. The findings from this research provide insights into the best-performing CNN models for automated marble classification, offering a potential solution to streamline the identification process and support the growth of the Philippines' marble industry. The rest of the paper is structured as follows. Section 2 describes DL methods for marble classification, whereas Sections 3 and 4 provide detailed descriptions of training and experimental evaluations. Conclusions are presented in the Section 5.

2. Deep Learning and Evolutionary Image Augmentation for Marble Classification

This section gives more information about the different CNN approaches for marble classification along with their usage.

2.1. VGG16 and VGG16

The consistent structure of VGGNet, created by Simonyan and Zisserman, made it a population architecture for image feature extraction, which is good at extracting visual information using 3×3 filter of convolution with 13 convolutions over the 16 layers. However, despite its strong performance, VGG16 has a high computational cost, requiring the handling of 138 million parameters, making it computationally expensive for large-scale applications. Several researchers have successfully applied VGGNet in marble classification tasks. Canayaz et al. [11] implemented VGGNet for marble classification, while Oktem et al. [15] used the VGG algorithm to classify 516 marble images, achieving an accuracy of 96.07% [16]. Additionally, Ather et al. [12] employed VGGNet in their study to automate the recognition and classification of granite tiles. However, despite its high classification accuracy, VGGNet faces certain limitations, such as long training times and a large number of parameters, which may hinder its efficiency in real-time applications.

2.2. ResNet50

Introduced in 2015, Residual Network (ResNet) revolutionized deep learning architecture by incorporating modular networks or networks within networks, enabling deeper models with improved training stability [17]. One of the key breakthroughs of ResNet, as highlighted by Zahisham et al. [18], is the residual connections allowing information to bypass certain layers, ensuring that each residual block retains portions of the previous activation map without modification [18]. This innovation effectively mitigates the problem of gradient, which is deleted, allowing the network to efficiently learn from it. Among the different ResNet variants, ResNet50 is one of the most widely used, comprising 50 weighted layers optimized for deep feature extraction [19]. Studies by Tourn et al. [20] and Sidiropoulos et al. [21] demonstrated that ResNet50 is a robust and reliable deep learning architecture and is good for high-precision image classification tasks [20,21].

2.3. InceptionV3

The Inception architecture, originally introduced by Sergey et al. in the research paper "Going Deeper with Convolutions," is an advanced version of GoogleNet designed to enhance computational efficiency and classification performance [22]. Over time, several refinements were introduced, including batch normalization, which led to the development of Inception-v2. Further improvements incorporated factorization techniques and an auxiliary classifier that helps propagate label information throughout the network, leading to the creation of Inception-v3. This version maintains high accuracy while improving computational efficiency, making it an excellent choice for deep learning-based image classification tasks.

2.4. MobileNetV2

MobileNetV2 is a lightweight DNN architecture optimized for efficient computation and real-time applications. Developed by Google, it improves upon its predecessor by using linear bottlenecks and inverted residual, dramatically lowering computation costs without sacrificing accuracy [23]. MobileNetV2 is particularly well-suited for edge devices, mobile applications, and low-power environments, making it an efficient alternative to traditional deep learning models. Compared to models like VGG16 and ResNet50, MobileNetV2 is more computationally efficient because of its ability to achieve high classification accuracy with fewer parameters. Depthwise separable convolutions are used in the design, which drastically lowers the number of parameters and floating-point operations (FLOPs) needed for picture classification. Due to its efficiency, MobileNetV2 is commonly used in real-world applications that require fast inference times while maintaining reliable classification performance. Studies have shown that MobileNetV2 can outperform heavier architectures in terms of computational efficiency while achieving comparable accuracy scores, making it an ideal candidate for resource-constrained environments. Its capability to deliver state-of-the-art performance with minimal computational overhead makes it a strong contender for marble classification tasks, where real-time processing is essential.

2.5. Genetic Algorithm (GA)

The Genetic Algorithm (GA) is an optimization method that draws inspiration from genetics and natural selection. In addition to optimization challenges, it is commonly utilized in research and development (R&D) and machine learning applications [22]. By applying selection, crossover, and mutation to a population of possible solutions and repeating the process over generations to identify the best results, GA mimics biological evolution. GA's main objective is to produce high-quality solutions for complicated problems by giving each iteration's most "fit" individuals priority. The probability of survival and reproduction

of a solution is determined by its fitness, which guarantees that the quality of succeeding generations increases with time [24]. For a given computing issue, the GA process ends when an optimal solution is found [25]. In deep learning applications, GA has been widely applied for hyperparameter tuning and dataset augmentation, improving the generalization ability of neural networks [26]. By introducing genetic-based augmentation techniques, GA helps enhance dataset variability, reducing overfitting and improving classification performance. This approach has been particularly beneficial in image classification tasks, where feature diversity is crucial for achieving high accuracy.

3. Training

This section describes the training procedure and parameter settings used in this research.

3.1. Datasets

The training process in this study followed a structured approach to ensure the effective classification of Romblon marble patterns (Figure 1). A comprehensive dataset of 6000 high-resolution images was collected from actual quarry sites in Romblon, reflecting the diverse textures, colors, and vein patterns characteristic of different marble classes [27]. Each class contained 1500 images, all captured using a DSLR camera under controlled lighting conditions to minimize inconsistencies. Expert annotators ensured the accuracy of class labels, providing a reliable foundation for model training [28]. To optimize model learning, the dataset was split into three subsets: 60% or 3600 marble images for training; 20% or 1200 marble images for validation, and 20% or 1200 marble images for testing [29]. Additionally, to enhance dataset variability and improve model generalization, a Genetic Algorithm (GA)-based image augmentation was applied to the training set. These are the rotation, flipping, scaling, and cropping that were utilized to introduce variations, ensuring the CNN models were exposed to diverse representations of marble patterns (Figure 2). This augmentation process played a crucial role in preventing overfitting and improving the robustness of the trained models. To guarantee a fair comparison of CNN model performance, the training process kept the hyperparameter settings constant throughout all experiments. Because categorical cross-entropy loss works well for multi-class classification issues, it was used as the loss function. Adaptive learning rate modifications were made using the Adam optimizer, which used adaptive gradient estimates and momentum to achieve the best convergence. Early stopping was implemented as an error correction mechanism, preventing unnecessary training iterations once validation loss ceased to improve. These training strategies, combined with GA-enhanced data augmentation, provided a structured and optimized approach to achieving high-accuracy marble classification models.

GA Fitness Function Formula

$$F(I) = \alpha \cdot S(I) + \beta \cdot E(I) + \gamma \cdot D(I) \quad (1)$$

GA Crossover Formula

$$I_{new} = W_1 I_{parent1} + W_2 I_{parent2} \quad (2)$$

Mutation Formula

$$I^I = I + M \times G \quad (3)$$

https://doi.org/10.3390/engproc2026143037

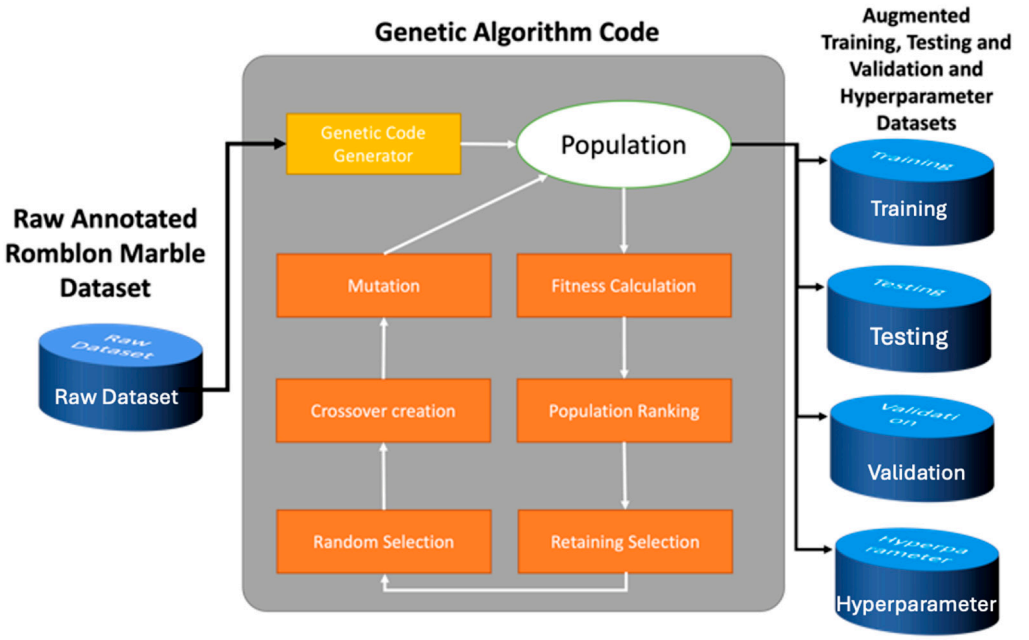


Figure 1. Genetic Algorithm code for image augmentation.

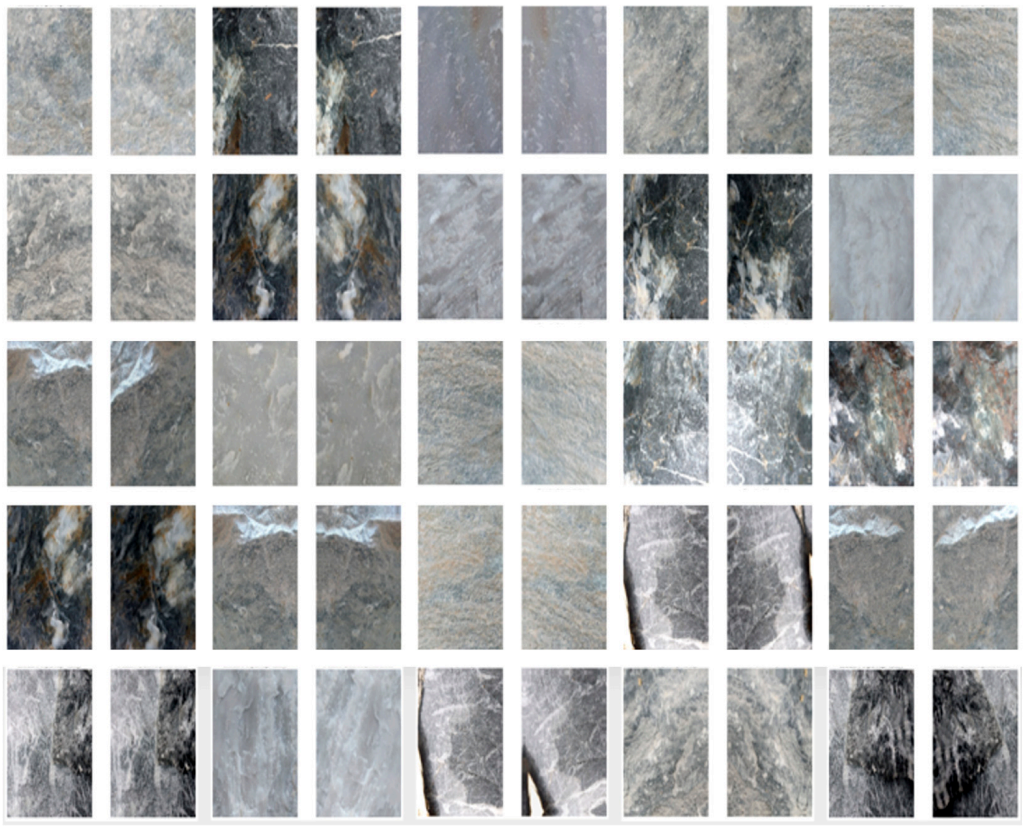


Figure 2. Genetic Algorithm Generated Augmented Images.

3.2. Procedure

The collected dataset is preprocessed and systematically incrementally evaluated across multiple CNN architectures. To ensure consistency, all models are trained using the same hyperparameter settings, allowing for an objective performance comparison. Since categorical cross-entropy loss works well for multi-class classification issues, it is utilized to quantify loss [30]. Accuracy serves as the primary performance metric, while the Adam

optimizer is applied due to its ability to compute adaptive learning rates for each parameter using momentum and adaptive gradient estimation [31].

For image augmentation, a Genetic Algorithm (GA) was applied to enhance dataset variability and improve model generalization. The GA was configured with parameter settings shown in Table 1, where a maximum of 20 generations and a population size of 10 were chosen to strike a compromise between computing cost and evolutionary search efficiency. A one-point crossover operator was used to combine features from two parent images, while a mutation rate of 0.01 ensured controlled randomness to maintain image diversity without compromising structural integrity. Tournament selection was implemented to prioritize fitter individuals for reproduction, with validation accuracy defined as the fitness function to directly align augmentation with model performance. An early stopping criterion with a five-generation patience was implemented to prevent needless calculations, ending the GA process when advancements reached a plateau.

Table 1. Genetic Algorithm Parameter Settings for GA-Based Image Augmentation.

Parameter	Value	Description
Population Size	10	Number of Individuals per Generation
Generations	20 (max)	Maximum Iteration
Crossover	One-point	Combines features on two parent images
Mutation Rate	0.01	Probability of introducing random changes
Selection	Tournament	Selects fitter individuals for reproduction
Fitness Function	Validation Accuracy	Guides GA toward model performance improvement
Early Stopping	Patience = 5	Stops if no improvement in 5 consecutive generations

The GA-augmented dataset was then used to train CNN models, which were implemented in Keras with ImageNet pre-trained weights to leverage transfer learning. By combining GA-driven augmentation with early stopping, the proposed procedure produced a diverse and high-quality set of training images, ultimately enhancing model robustness and leading to improved classification accuracy across different marble types [32].

3.3. Classification

Two experiments employing distinct dataset configurations were carried out to assess the performance of different CNN architectures. The first experiment involved training CNN models on the raw dataset, where no data augmentation was applied. This setup aimed to assess the models' capability to learn and classify marble textures with limited training samples. In order to increase the dataset's diversity and the models' capacity for generalization, the second experiment used GA-based data augmentation.

Standard assessment criteria, such as accuracy, recall, precision, F1-score, loss score, and ROC-AUC score, were used in both studies to assess model performance. The architectures tested included AlexNet, InceptionV3, VGG16, ResNet50, and MobileNet. Results demonstrated that models trained on the GA-augmented dataset exhibited significantly higher classification accuracy and lower loss, confirming that evolutionary augmentation techniques contribute to improved feature extraction and pattern recognition in marble classification tasks.

3.4. Evaluation

Evaluation Metrics

The performance of the different models can be explained by using evaluation measures. The research utilizes the Accuracy Score, F1 Score, and ROC-AUC Score as metrics of evaluation.

- Accuracy Score: The percentage of accurately predicted cases out of all the samples examined is known as classification accuracy. It provides information about how well

the classification algorithm recognizes the appropriate categories and is a key statistic for assessing model performance. By dividing the total number of accurate forecasts by the total number of predictions made, accuracy is calculated. The precision is then expressed as a percentage by multiplying the resultant fraction by 100 [33]. The formula for the accuracy score is as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

- Cross-Entropy Loss: By evaluating how closely a classification model's predicted probabilities match the actual class labels, cross-entropy loss gauges the model's efficacy. A perfect model with no misclassification mistakes is represented by a loss value of 0, which falls between 0 and 1 [30]. In order to improve forecast accuracy, the main goal of model training is to minimize this loss and get it as near to zero as possible. Although logistic loss and cross-entropy loss are frequently used interchangeably, they are not always the same. In particular, cross-entropy loss measures the discrepancy between the projected distribution produced by the classification model and the actual probability distribution seen in the dataset. In multi-class classification issues, this metric is frequently employed to make sure the model differentiates between several categories as well as it can [34,35].
- F1 Score: Evaluating classification models can be challenging, particularly when one model exhibits high precision but low recall, while another demonstrates strong recall but poor precision. Optimizing one metric often negatively affects the other, leading to significant discrepancies in model performance and increased confusion when making comparisons [36]. As a balanced metric that takes precision and recall into account at the same time, the F1-Score is utilized to overcome this problem. It offers a harmonic mean of the two, guaranteeing a more thorough assessment of the model's performance. When there is class imbalance in the dataset, the F1-Score is particularly useful, as it prevents misleading assessments that could arise from accuracy alone [37]. Since most real-world classification problems involve uneven class distributions, the F1-Score serves as a more reliable performance metric compared to accuracy.

$$F_1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} = \frac{2TP}{2TP + FP + FN} \quad (5)$$

- ROC-AUC Score: Model evaluation is not limited to numerical metrics alone; visual performance metrics such as the Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) also play a crucial role in assessing classification models [38]. The ROC curve graphically represents a model's performance by plotting the true positive rate (sensitivity) against the false positive rate, allowing for a comprehensive analysis of its classification capabilities across different threshold values. The AUC-ROC score quantifies a model's ability to generalize by summarizing the ROC curve into a single value. A higher AUC score indicates that the model is better at distinguishing between positive and negative classes. When comparing multiple models, the one with the highest AUC is generally considered superior, as it demonstrates greater accuracy in identifying correct classifications. Consequently, AUC serves as a key indicator of a model's overall reliability and effectiveness in real-world classification tasks [39].

3.5. Model Evaluation

The models in this study were assessed using key performance metrics, including accuracy score, cross-entropy loss, F1-score, and ROC-AUC score, to gauge their effectiveness

in classifying Romblon marble patterns. The categorical cross-entropy loss was chosen to measure how well the predicted class probabilities matched the actual class labels, while the SoftMax activation function was applied to convert raw outputs into probability values [40]. To optimize learning efficiency, the Adam optimizer was used, as it dynamically adjusts learning rates for each parameter, ensuring stable and faster convergence.

A comparison was made between Experiment 1, where CNN models were trained on raw datasets, and Experiment 2, which used GA-augmented datasets. The results indicated that InceptionV3, VGG19, and VGG16 performed exceptionally well in terms of accuracy, proving their capability to extract meaningful features for marble classification. On the other hand, ResNet50 and GA-based models recorded higher loss values, suggesting they had more difficulty in generalizing the dataset effectively.

As training progressed through different epochs, variations in loss trends (Figure 3) and accuracy (Figure 4) became noticeable. While some models, such as InceptionV3 and VGG16, consistently improved their performance with additional training, others experienced fluctuations, likely due to overfitting. To counteract this, early stopping was applied to ResNet50, preventing excessive training when validation loss stopped decreasing. This helped strike a balance between model learning and generalization while also reducing unnecessary computations.

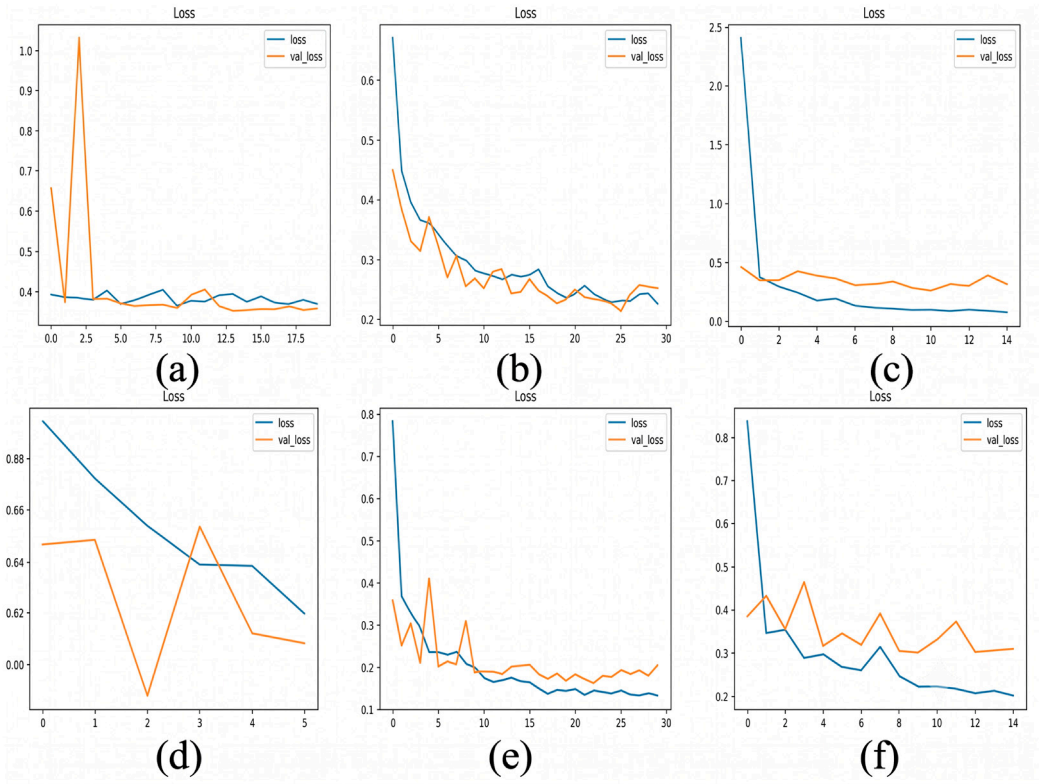


Figure 3. Loss Score graph of (**a**) AlexNet, (**b**) InceptionV3, (**c**) MobileNet, (**d**) ResNet50, (**e**) VGG16, and (**f**) VGG19 models using Raw Datasets.

The findings suggest that GA-based image augmentation played a significant role in enhancing model performance, with VGG16 achieving the highest accuracy at 94.68%, followed closely by InceptionV3 and MobileNetV2, both surpassing 90% accuracy. These results indicate that Genetic Algorithm-driven augmentation techniques can be highly beneficial for improving CNN-based marble classification, making them a practical solution for real-world applications in the marble industry.

https://doi.org/10.3390/engproc2026143037

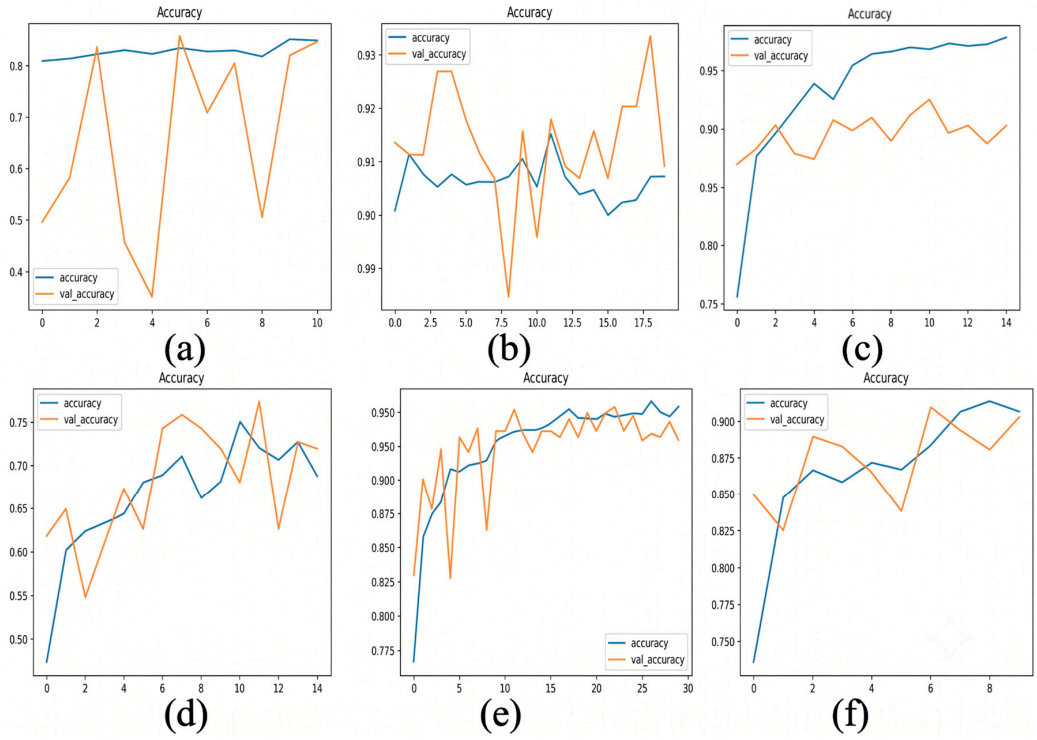


Figure 4. Accuracy graph of (**a**) AlexNet, (**b**) InceptionV3, (**c**) MobileNet, (**d**) ResNet50, (**e**) VGG16, and (**f**) VGG19 models using GA Augmented Dataset.

4. Discussion

The evaluation of CNN models trained on both the unprocessed dataset and the GA-augmented dataset highlights the substantial role of augmentation in improving classification accuracy and model generalization. As shown in Table 2, models trained solely on the raw dataset achieved relatively low accuracy, with InceptionV3 (35.33%) and AlexNet (34.67%) performing slightly better than other architectures. VGG16 and VGG19 achieved comparable scores of 33.78% and 33.56%, respectively, while ResNet50 showed the lowest performance at 31.11%. The consistently high entropy loss values (above 1.1) and low F1-scores (ranging from 0.29 to 0.35) indicate that the raw dataset provided insufficient variability, leading to poor convergence and limited generalization across all architectures.

Table 2. Comparison of Various Deep Learning Models Using Raw Dataset.

Model	Accuracy	Entropy Loss	F1 Score	ROU-AUC
AlexNet	34.67	1.1036	0.29	0.5443
InceptionV3	35.33	1.1369	0.34	0.4932
MobileNet	34.67	1.2010	0.32	0.5050
ResNet50	31.11	1.2012	0.35	0.5388
VGG16	33.78	1.2046	0.31	0.5021
VGG19	33.56	1.2031	0.30	0.5406

In contrast, Table 3 demonstrates that GA-based image augmentation resulted in substantial improvements in model performance. VGG16 achieved the highest accuracy at 94.68%, followed closely by MobileNet (92.68%) and InceptionV3 (92.46%). Significant gains were also observed in AlexNet, which improved from 34.67% to 85.14%, and ResNet50, which rose from 31.11% to 89.58%. These results suggest that GA-driven augmentation enriched feature diversity, enabling the CNNs to capture subtle marble texture variations more effectively. The consistent reduction in entropy loss values (dropping from >1.1

to below 0.34) further supports the role of augmentation in mitigating overfitting and promoting more stable learning.

Table 3. Comparison of Various Deep Learning Models Using GA Augmented Marble Dataset.

Model	Accuracy	Entropy Loss	F1 Score	ROU-AUC
AlexNet	85.14	0.3362	0.30	0.4914
InceptionV3	92.46	0.2370	0.33	0.4864
MobileNet	92.68	0.2717	0.29	0.5093
ResNet50	89.58	0.2504	0.27	0.4752
VGG16	94.68	0.1611	0.35	0.5076
VGG19	89.80	0.2545	0.31	0.4930

The Receiver Operating Characteristic (ROC) analysis provides an additional perspective. With the raw dataset (Table 2), the highest AUC values were obtained by ResNet50 (0.5388) and VGG19 (0.5406), while AlexNet (0.5443) also performed relatively better than expected compared to its accuracy. However, AUC values across all models remained close to 0.5, indicating limited separability between classes. After augmentation (Table 3), AUC values showed only minor changes: VGG16 reached 0.5076, MobileNet 0.5093, and VGG19 0.4930. Although these fluctuations are not proportional to the accuracy improvements, they suggest that GA augmentation improved correct classifications but did not substantially enhance model discriminability across all decision thresholds.

It is also worth noting that F1-scores remained modest (0.27–0.35) despite significant accuracy gains. This discrepancy may be attributed to class imbalance in the dataset, where certain marble types were more dominant than others. While GA augmentation increased dataset size and diversity, it may not have fully addressed the imbalance, resulting in relatively low F1-scores despite high overall accuracy.

When combined, these results demonstrate how well evolutionary augmentation methods work to maximize CNN-based marble classification. While MobileNet showed an ideal mix between accuracy and computational efficiency, making it a viable option for deployment in real-world scenarios, VGG16 turned out to be the most accurate model. At the same time, the mixed AUC and F1-score results highlight the need for further refinement. Future work may consider class-balanced augmentation strategies or hybrid optimization techniques to strengthen both accuracy and generalization across diverse CNN architectures.

5. Conclusions

This study examined the impact of Genetic Algorithm (GA)-based image augmentation on the performance of CNN for identifying Romblon marble patterns. Two experiments were conducted: the first evaluated CNN performance on a raw marble dataset, while the second assessed performance using a GA-augmented dataset. Results clearly showed that applying GA for image augmentation significantly improved classification accuracy across all tested CNN architectures. Models trained on the GA-augmented dataset consistently outperformed those trained on the raw dataset. Among the tested architectures, VGG16 achieved the highest accuracy of 94.68% after 30 epochs, while InceptionV3 and MobileNet also exceeded 90% accuracy. The analysis of loss scores and convergence patterns further confirmed that augmentation enhanced classification performance and reduced overfitting, enabling the models to generalize better.

A computational complexity analysis also revealed important trade-offs between accuracy and efficiency. While VGG16 achieved the highest accuracy, MobileNet's great accuracy balance made it the most sensible option for practical applications (92.68%) and

computational efficiency. These findings highlight the importance of selecting models that not only deliver high accuracy but are also suitable for deployment in real-time industrial settings. Although ROC-AUC scores did not improve proportionally with accuracy, reflecting some limitations in class separability, the overall results affirm the value of GA-based augmentation in boosting CNN learning capacity.

Future research may examine the effects of hyperparameter adjustment and class-balanced augmentation on CNN performance, as well as other optimization techniques like Particle Swarm Optimization (PSO) for feature selection. This study advances the technology of the marble business in Romblon and offers insightful information about automated marble classification.

Author Contributions: Conceptualization, M.R.G.F. and S.V.; methodology, M.R.G.F.; software, M.R.G.F.; validation, M.R.G.F. and S.V.; formal analysis, M.R.G.F.; investigation, M.R.G.F.; resources, M.R.G.F.; data curation, M.R.G.F.; writing—original draft preparation, M.R.G.F.; writing—review and editing, S.V.; visualization, M.R.G.F.; supervision, S.V.; project administration, M.R.G.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interests.

Abbreviations

The following abbreviations are used in this manuscript:

CNN	Convolutional Neural Network
GA	Genetic Algorithm
ROC	Receiver Operating Characteristic
AUC	Area Under Curve

References

1. Fos, P.J. Mapping Romblon Marble's Future. 2025. Available online: https://pia.gov.ph/news/mapping-romblon-marbles-future/ (accessed on 5 June 2025).
2. *Minerals Yearbook 2018, Volume I, Metals and Minerals*; U.S. Department of the Interior, U.S. Geological Survey: Reston, VA, USA, 2018.
3. Parrocha, A. Romblon Gears Up for Biggest MIMAROPA FESTIVAL—Philippine News Agency. 2017. Available online: https://www.pna.gov.ph/articles/1007758 (accessed on 5 June 2025).
4. An Inquiry into the Quarry: Lived Experiences and Narratives in Romblon's Mining Communities. 2021. Available online: https://chr.gov.ph/an-inquiry-into-the-quarry-lived-experiences-and-narratives-in-rombons-mining-communities (accessed on 5 June 2025).
5. Baykasolu, A.; Özbel, B.K. Modeling and solving a real-world cutting stock problem in the marble industry via mathematical programming and stochastic diffusion search approaches. *Comput. Oper. Res.* **2021**, *128*, 105173. [CrossRef]
6. Ouzounis, A.; Sidiropoulos, G.; Papakostas, G.; Sarafis, I.; Stamkos, A.; Solakis, G. Interpretable deep learning for marble tiles sorting. In *Proceedings of 2nd International Conference on Deep Learning Theory and Applications*; SciTePress: Setúbal, Portugal; pp. 101–108. [CrossRef]
7. Naranjo-Torres, J.; Mora, M.; Hernández-García, R.; Barrientos, R.J.; Fredes, C.; Valenzuela, A. A review of convolutional neural network applied to fruit image processing. *Appl. Sci.* **2020**, *10*, 3443. [CrossRef]
8. Sarfraz, M. (Ed.) *Digital Imaging*; IntechOpen: London, UK, 2020.
9. Avci, D.; Sert, E. An effective turkey marble classification system: Convolutional neural network with genetic algorithm—Wavelet kernel—Extreme learning machine. *Trait. Signal* **2021**, *38*, 1229–1235. [CrossRef]

10. Moghaddam, H.K.; Rajaei, A.; Moghaddam, H.K. Marble slabs classification system based on image processing (Ark marble mine in Birjand). *Civ. Eng. J.* **2018**, *4*, 107. [CrossRef]
11. Canayaz, M.; Uludağ, F. Marble classification using deep neural networks. *Eur. J. Tech.* **2020**, *10*, 52–63. [CrossRef]
12. Ather, M.; Khan, B.; Wang, Z.; Song, G. Automatic recognition and classification of granite tiles using convolutional neural networks (CNN). In *Proceedings of 3rd International Conference on Advances in Artificial Intelligence*; ACM: New York, NY, USA, 2020; pp. 193–197. [CrossRef]
13. Pençe, İ.; Şişeci Çeşmeli, M. Deep learning in marble slabs classification. *Techno-Science* **2019**, *2*, 21–26.
14. Karaali, A.; Eminağaoğlu, M. A convolutional neural network model for marble quality classification. *SN Appl. Sci.* **2020**, *2*, 1733. [CrossRef]
15. Mascarenhas, S.; Agarwal, M. A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for image classification. In Proceedings of the 2021 International Conference on Disruptive Technologies for Multi-Disciplinary Research and Applications (CENTCON), Bengaluru, India, 19–21 November 2021; pp. 96–99. [CrossRef]
16. Öktem, M.; Akosman, A.; Moral, T.; Kılıç, V. Derin öğrenme tabanlı mermer yüzeylerinin otomatik sınıflandırılması. *Eur. J. Sci. Technol.* **2021**, *26*, 73–77. [CrossRef]
17. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778. [CrossRef]
18. Zahisham, Z.; Lee, C.P.; Lim, K.M. Food recognition with ResNet-50. In Proceedings of the 2020 IEEE 2nd International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), Kota Kinabalu, Malaysia, 26–27 September 2020; pp. 1–5. [CrossRef]
19. Sarwinda, D.; Paradisa, R.H.; Bustamam, A.; Anggia, P. Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer. *Procedia Comput. Sci.* **2021**, *179*, 423–431. [CrossRef]
20. Torun, Y.; Akbas, M.R.; Celik, M.A.; Kaynar, O. Development of a machine vision system for marble classification. In Proceedings of the 2019 27th Signal Processing and Communications Applications Conference (SIU), Sivas, Turkey, 24–26 April 2019; pp. 1–4. [CrossRef]
21. Sidiropoulos, G.K.; Ouzounis, A.G.; Papakostas, G.A.; Sarafis, I.T.; Stamkos, A.; Solakis, G. Texture analysis for machine learning-based marble tile sorting. In Proceedings of the 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC), virtual, 27–30 January 2021; pp. 0045–0051. [CrossRef]
22. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826. [CrossRef]
23. Elmas, B. Classification varieties of marble and granite by convolutional neural networks with transfer learning method. *J. Fac. Eng. Archit. Gazi Univ.* **2022**, *37*, 985–1001. [CrossRef]
24. Goyal, G.; Srivastava, P.K.; Bisht, D.C.S. Genetic algorithm: A metaheuristic approach of optimization. In *Computational Intelligence*; Bisht, D.C., Ram, M., Eds.; De Gruyter: Berlin, Germany, 2020; pp. 27–44. [CrossRef]
25. Helwe, C.; Clavel, C.; Suchanek, F. Reasoning with transformer-based models: Deep learning, but shallow reasoning. In *2021 International Conference on Automated Knowledge Base Construction (AKBC)*; Institut Mines-Télécom: Palaiseau, France, 2021. Available online: https://hal-imt.archives-ouvertes.fr/hal-03344668 (accessed on 1 June 2025).
26. Mirjalili, S.; Dong, J.S.; Sadiq, A.S.; Faris, H. *Genetic Algorithm: Theory, Literature Review, and Application in Image Reconstruction*; Springer: Cham, Switzerland, 2020; pp. 69–85. [CrossRef]
27. Ho, Y.; Wookey, S. The real-world-weight cross-entropy loss function: Modeling the costs of mislabeling. *IEEE Access* **2020**, *8*, 4806–4813. [CrossRef]
28. Paullada, A.; Raji, I.D.; Bender, E.M.; Denton, E.; Hanna, A. Data and its (Dis)contents: A survey of dataset development and use in machine learning research. *Patterns* **2021**, *2*, 100336. [CrossRef] [PubMed]
29. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2017**, arXiv:1412.6980.
30. Song, H.; Kim, M.; Park, D.; Lee, J.-G. How does early stopping help generalization against label noise? *arXiv* **2020**, arXiv:1911.08059.
31. Li, J.; Gao, M.; D'Agostino, R. Evaluating classification accuracy for modern learning approaches. *Stat. Med.* **2019**, *38*, 2477–2503. [CrossRef] [PubMed]
32. Zhang, Z.; Sabuncu, M.R. Generalized cross-entropy loss for training deep neural networks with noisy labels. In *Proceedings of the 32nd International Conference on Neural Information Processing System*; ACM: New York, NY, USA, 2018; pp. 8792–8802.
33. Uzen, H.; Turkoglu, M.; Hanbay, D. Surface defect detection using deep U-Net network architectures. In Proceedings of the 2019 27th Signal Processing and Communications Applications Conference (SIU), Sivas, Turkey, 24–26 April 2019; pp. 1–4. [CrossRef]
34. Wardhani, N.W.S.; Rochayani, M.Y.; Iriany, A.; Sulistyono, A.D.; Lestantyo, P. Cross-validation metrics for evaluating classification performance on imbalanced data. In Proceedings of the 2019 International Conference on Computer, Control, Informatics and its Applications (IC3INA), Tangerang, Indonesia, 23–24 October 2019; pp. 14–18. [CrossRef]

35. Di Napoli, M.; Carotenuto, F.; Cevasco, A.; Confuorto, P.; Di Martire, D.; Firpo, M.; Pepe, G.; Raso, E. Machine learning ensemble modeling as a tool to improve landslide susceptibility mapping reliability. *Landslides* **2020**, *17*, 1897–1914. [CrossRef]
36. Mandal, I.; Pal, S. Modeling human health vulnerability using different machine learning algorithms in stone quarrying and crushing areas of Dwarka river basin, eastern India. *Adv. Space Res.* **2020**, *66*, 1351–1371. [CrossRef]
37. Jaskowiak, P.A.; Costa, I.G.; Campello, R.J.G.B. The area under the ROC curve as a measure of clustering quality. *Data Min. Knowl. Discov.* **2022**, *36*, 1219–1245. [CrossRef]
38. Zaman, S.M.; Hasan, M.M.; Sakline, R.I.; Das, D.; Alam, M.A. A comparative analysis of optimizers in recurrent neural networks for text classification. In Proceedings of the 2021 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), Brisbane, Australia, 8–10 December 2021; pp. 1–6.
39. Zhou, Y.; Wang, X.; Zhang, M.; Zhu, J.; Zheng, R.; Wu, Q. MPCE: A maximum probability-based cross-entropy loss function for neural network classification. *IEEE Access* **2019**, *7*, 146331–146341. [CrossRef]
40. Baldi, P. Autoencoders, unsupervised learning, and deep architectures. In *Proceedings of the 2011 International Conference on Unsupervised and Transfer Learning Workshop*; ACM: New York, NY, USA, 2011; pp. 37–50. Available online: https://arxiv.org/abs/1206.5538 (accessed on 5 June 2025).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.