

SARS-CoV-2 Detection in Genosensor Images via Deep Learning and Data Augmentation

Detecção de SARS-CoV-2 em imagens de genossensores por meio de aprendizado profundo e aumento de dados

Ana Beatriz S. Zerati¹, Luan B. Guerra¹, Osvaldo N. Oliveira Jr,² Lucas C. Ribas^{1*}

Abstract: This work addresses the challenge of classifying genosensor images obtained by scanning electron microscopy (SEM) for SARS-CoV-2 diagnosis. Due to the specialized nature of the data, image datasets are inherently small, which limits the effectiveness of deep learning models. To overcome this limitation, we investigated the impact of seven data augmentation strategies - Flipping, Shift, Random Erasing, AugMix, AutoAugment, RandAugment and TrivialAugment - on the performance of four Convolutional Neural Networks (CNNs): ResNet50, ResNet101, DenseNet121, and ConvNeXt-Tiny. Our analyses demonstrated that data augmentation is crucial for the success of the task, achieving a remarkable accuracy of 97.85%, surpassing both traditional feature extraction methods and previously tested deep learning approaches. Furthermore, our findings indicate that the choice of augmentation technique is critical, with aggressive methods proving detrimental for highly specialized microscopy data.

Keywords: data augmentation — genosensors — deep learning

Resumo: Este trabalho aborda o desafio da classificação de imagens de genossensores, obtidas por microscopia eletrônica de varredura (MEV), para o diagnóstico de SARS-CoV-2. Devido à natureza especializada dos dados, os conjuntos de imagens são inerentemente pequenos, o que limita a eficácia de modelos de deep learning. Para superar essa limitação, investigamos o impacto de sete estratégias de aumento de dados — Flipping, Shift, Random Erasing, AugMix, AutoAugment, RandAugment e TrivialAugment — no desempenho de quatro redes neurais convolucionais: ResNet50, ResNet101, DenseNet121 e ConvNeXt-Tiny. Nossas análises demonstraram que o aumento de dados é crucial para o sucesso da tarefa, alcançando uma notável acurácia de 97.85%, superando tanto os métodos tradicionais de extração de características quanto as abordagens de deep learning previamente testadas. Além disso, nossas descobertas indicam que a escolha da técnica de aumento de dados é crítica, já que métodos agressivos se mostraram prejudiciais para dados de microscopia altamente especializados.

Palavras-Chave: aumento de dados — genossensores — aprendizado profundo

¹ *Institute of Biosciences, Humanities and Exact Sciences, Universidade Estadual Paulista (Unesp), Brazil*

² *Instituto de Física de São Carlos, Universidade de São Paulo (USP), Brazil*

*Corresponding author: lucas.ribas@unesp.br

DOI: <http://dx.doi.org/10.22456/2175-2745.150894> • Received: 13/10/2025 • Accepted: 27/10/2025

CC BY-NC-ND 4.0 - This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

1. Introduction

The COVID-19 pandemic highlighted the critical need for rapid, accurate, and cost-effective diagnostic tools to contain the spread of infectious diseases. While gold-standard methods like real-time reverse transcription–polymerase chain reaction (RT-PCR) are highly accurate, they are often expensive, require specialized laboratory equipment, and can have long turnaround times. Molecular diagnosis using real-time RT-PCR takes at least 3 hours, including preparation of viral RNA, which can affect diagnostic accuracy [1]. This has

spurred the development of alternative diagnostic platforms, with genosensors emerging as a promising solution for point-of-care (POC) testing.

Genosensors are a class of biosensors that utilize nucleic acids to detect a wide range of biological analytes, including viral RNA [2]. They are highly valued for their exceptional sensitivity and specificity, crucial for identifying pathogens like SARS-CoV-2. Furthermore, genosensors offer key advantages over conventional methods, being relatively low-cost, rapid, and simple to use. The genosensor proposed by Soares *et al.* [3], for example, is capable of yielding a diagnosis in a

few minutes at a low cost.

The work of Soares *et al.* [3] leverages an image-based detection approach for genosensors, a strategy initially developed for prostate cancer diagnosis via the analysis of scanning electron microscopy (SEM) images [4]. This method is effective because the binding of target analytes induces specific and measurable changes in the biosensor's surface morphology. These changes in texture and morphology provide a rich visual signature that can be directly classified to determine a diagnosis.

While this image-based approach has proven successful, its application to emerging pathogens like SARS-CoV-2 presents a unique challenge. A previous study using this exact methodology on a SARS-CoV-2 genosensor dataset found that traditional handcrafted texture descriptors with an LDA classifier surprisingly outperformed initial deep learning (DL) models [3]. This outcome highlights a critical limitation: the scarcity of large, diverse datasets, which prevents data-hungry DL models from reaching their full potential [5, 6, 7]. This issue often leads to poor generalization and overfitting, underscoring the need for a solution to this data scarcity problem [8, 9].

To address this limitation, our work explores the full potential of deep learning by conducting an in-depth analysis supported by data augmentation. We hypothesize that synthetically expanding the genosensor image dataset will allow deep convolutional neural network (CNN) models to learn more robust and generalizable features. We also believe that this approach will enable models trained on augmented datasets to surpass the performance of both handcrafted descriptors and standard deep models in the detection task. This work contributes by providing an effective and scalable detection solution, while also demonstrating the critical role of data augmentation in enabling advanced machine learning on limited and specialized datasets.

2. Methods

This section details the experimental design, dataset characteristics, data augmentation protocol, model architectures, and evaluation metrics used in this study.

2.1 Dataset

The dataset used in this study consists of 200 scanning electron microscopy (SEM) images of genosensor units. Each image, with an original resolution of 1024×714 pixels and 8-bit grayscale depth, was obtained from a genosensor unit exposed to a specific analyte, including various concentrations of the target analyte, SARS-CoV-2, alongside negative controls (unexposed probes) and two types of interferents (HPV16 and PCA3). The interferent samples serve as control measurements to test the genosensor's specificity against non-target analytes. The distribution of samples across the different analytes and their corresponding class labels is detailed in Table 1.

Table 1. Composition of the dataset used. The table details each analyte, its corresponding diagnostic label, and the number of image samples. The dataset includes both negative control samples (Probe, HPV16, PCA3) and seven classes representing varying concentrations of SARS-CoV-2.

Analyte	Label	N of samples
Probe	Negative	30
HPV16	Negative	10
PCA3	Negative	10
SARS-CoV-2 10^{-6}molL^{-1}	6	30
SARS-CoV-2 10^{-8}molL^{-1}	8	20
SARS-CoV-2 $10^{-10} \text{molL}^{-1}$	10	20
SARS-CoV-2 $10^{-12} \text{molL}^{-1}$	12	20
SARS-CoV-2 $10^{-14} \text{molL}^{-1}$	14	20
SARS-CoV-2 $10^{-16} \text{molL}^{-1}$	16	20
SARS-CoV-2 $10^{-18} \text{molL}^{-1}$	18	20

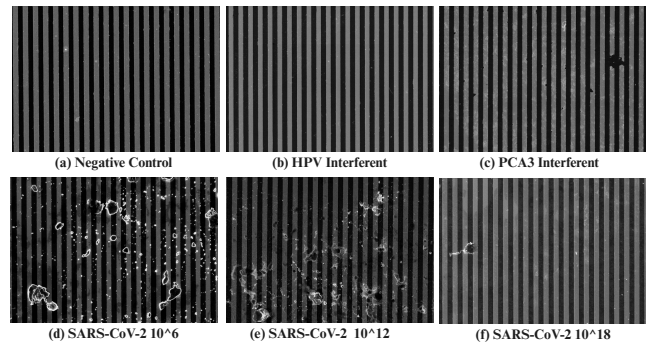


Figure 1. Example SEM images from the genosensor dataset, illustrating the morphological distinction between a negative control and a positive SARS-CoV-2 sample.

Figure 1 presents example images from the dataset, demonstrating the morphological differences between control and analyte-exposed samples.

2.2 Data Augmentation

To address the limited size of the dataset and enable deep learning models to learn robust and generalizable features, a set of data augmentation techniques was applied. These transformations were randomly performed on the training images to synthetically increase the dataset's variability. The following classic geometric transformations were implemented, motivated by the study conducted by Welsman *et al.* [10]:

- **Horizontal and Vertical Flips:** Images were randomly flipped horizontally and vertically. This is particularly useful for objects that have symmetric features.
- **Zoom and Shifting:** Random zooms (up to 1.2x) and small translations (up to 10% of the image size) were applied to simulate slight variations in image capture and scale.

We also tested modern methods that combine multiple transformations to improve model robustness and performance:

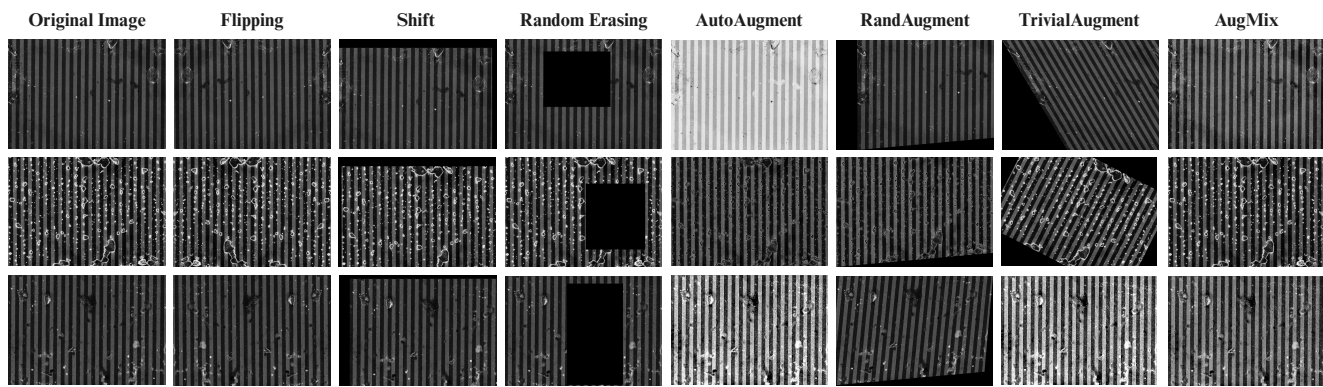


Figure 2. Examples of augmented images generated from three distinct original genosensor images. Each row illustrates the transformations applied by all evaluated data augmentation approaches, including Flipping, Shift, Random Erasing, AugMix, AutoAugment, RandAugment, and TrivialAugment.

- Random Erasing [11]: This technique randomly selects a rectangular area in the image and erases its pixels with a fixed value or random noise. The goal is to encourage the model to rely on other, un-occluded parts of the object for classification, making it more robust to occlusions and distractions in the data.
- AutoAugment [12]: This strategy uses a search algorithm (e.g., reinforcement learning) to automatically discover the best combination of data augmentation policies for a specific dataset. Instead of a manually selected set of transformations, AutoAugment learns which transformations and with what magnitudes should be applied to maximize validation accuracy. While highly effective, the search process is computationally expensive.
- RandAugment [13]: As a more efficient alternative to AutoAugment, RandAugment simplifies the process by eliminating the costly search procedure. It selects a random number of transformations from a predefined list and applies them all with the same, shared magnitude. This approach is much faster to implement and often achieves performance comparable to AutoAugment.
- TrivialAugment [14]: This is a further simplification of AutoAugment and RandAugment. It completely removes the magnitude parameter, so it simply selects one random transformation from a list and applies it with a randomly selected magnitude. Its "trivial" nature makes it extremely easy to use while still providing strong baseline results.
- AugMix [15]: Creates a "mixed" image by blending multiple augmented versions, improving robustness to common corruptions such as noise, blur, or compression artifacts. Although effective, it introduces additional computational overhead during preprocessing.

Figure 2 presents representative examples generated by each augmentation technique, illustrating the visual alterations introduced by these methods. In addition to their qualitative differences, the computational requirements of these tech-

niques also vary. AutoAugment incurs the highest computational cost due to the policy search process, which involves repeated training across large candidate sets. RandAugment and TrivialAugment reduce this cost by adopting fixed or randomized augmentation schemes that eliminate the need for search. AugMix introduces moderate overhead during preprocessing because it generates and blends multiple image variants. These methods therefore present distinct trade-offs between computational complexity and augmentation diversity, which are relevant when designing efficient training pipelines.

2.3 Proposed Learning Methodology

Our pipeline was designed to systematically evaluate the impact of different data augmentation strategies on the performance of deep learning models for SARS-CoV-2 genosensor image classification. The experimental procedure is demonstrated in Figure 3 and consists of the following steps: dataset partitioning, data augmentation application, model training and evaluation.

First, the original dataset of 200 images was randomly partitioned into three subsets: a training set, a validation set, and a test set, using 70/15/15 ratio. This partitioning is critical to prevent data leakage and ensure an unbiased evaluation of the models' performance on unseen data. Then, the data augmentation techniques described in Section 2.2 were applied exclusively to the training set. The validation and test sets remained untouched by these transformations to ensure a fair evaluation of each model's ability to generalize to real, unaugmented data.

We conducted a comparative study using four well-established deep learning architectures in image classification: ResNet-50, ResNet-101, DenseNet-121, and ConvNeXt-tiny. To address the data limitation, we employed models pre-trained on the ImageNet dataset. For each model and each data augmentation strategy, the training process was repeated three times to ensure the robustness of our results. The average performance across the three runs was considered for analysis. Each trained model was then evaluated on the held-out test set using a comprehensive set of metrics: accuracy, preci-

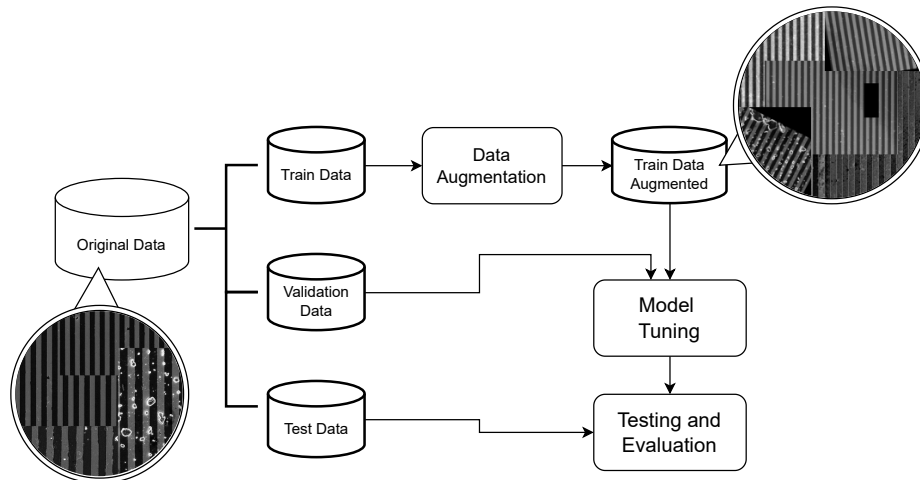


Figure 3. Diagram of the proposed methodology. The dataset is first split into training, validation, and testing sets using a 70/15/15 ratio. Data augmentation techniques are applied to the training set, followed by model training and hyperparameter tuning on the validation set. The final model performance is then evaluated on the test set.

sion, recall, and F1-score. These results were then compared across the different augmentation strategies and against the handcrafted descriptor methods from the study conducted by by Soares *et al.* [3] with the goal of validating our central hypothesis.

3. Results and Discussion

This section presents the results of our experiments, first analyzing the performance of different data augmentation (DA) strategies on the deep learning models and then comparing the best-performing approach to the prior state-of-the-art for this dataset.

3.1 Training and Augmentation Settings

The implementation of deep CNNs used the following libraries: Python 3.11.9, PyTorch 2.4.0, and Torchvision 0.19.0. The data augmentation strategies were applied with their default settings as provided by the Torchvision v0.19.0 library. The models were trained on both the original and augmented datasets using the Adam optimizer [16], with a learning rate of 10^{-4} and hyperparameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$, following the original implementation.

3.2 Impact of Data Augmentation

The results of the comparative analysis of deep learning models with various data augmentation strategies are summarized in Table 2. Our findings demonstrate that applying data augmentation significantly improves the performance of the deep learning models, particularly when compared to the baseline with no data augmentation.

As shown in Table 2, the ResNet-50 model, when trained with the Shift augmentation strategy, achieved the highest overall performance, with a remarkable accuracy of 97.85%. This result is a substantial improvement over the baseline

accuracy of 83.88% for the same model without data augmentation. Other classic techniques like Flipping also led to a notable increase in accuracy for ResNet-50, reaching 93.55%.

Interestingly, the performance of the advanced augmentation strategies was more varied. While certain methods proved highly effective, others failed to deliver a substantial improvement. Random Erasing demonstrated a strong performance, particularly benefiting the more complex architectures. The DenseNet-121 and ResNet-101 architectures, which have a greater capacity to learn intricate features, achieved their best results with this method (91.40% and 89.25% accuracy, respectively). This suggests that forcing these models to learn from partially-occluded images is an effective strategy for preventing overfitting and improving their ability to generalize. This finding can be attributed to Random Erasing's function as a regularizer, which combats the overfitting tendency of more complex models on a limited dataset and allows them to better utilize their full learning capacity.

Conversely, the other aggressive augmentation methods yielded the poorest results. TrivialAugment, AugMix, and AutoAugment often resulted in lower performance compared to the simpler techniques, and sometimes even the baseline without augmentation. The high standard deviation observed with these methods further indicates a lack of training stability, reinforcing the hypothesis that these aggressive augmentations can be detrimental. We hypothesize that more aggressive techniques, which perform drastic transformations or combine multiple effects, may inadvertently destroy the subtle, micro-level features that are critical for identifying the morphological changes on the genosensor's surface. Given the specialized nature of SEM images, these minute details are likely more important than large-scale geometric or photometric variations.

In this context, the performance of RandAugment is par-

Table 2. Comparative performance results of deep learning models for different data augmentation (DA) strategies. Values represent the mean and standard deviation over three runs. The best overall result is highlighted in bold, and the best results per model are underlined.

DA Strategy	Metric	Resnet 50	Resnet 101	Densenet 121	ConvNeXt Tiny
No Data Aug	Accuracy	83.875 ± 4.561	77.42 ± 4.567	80.647 ± 3.225	61.29 ± 9.68
	Precision	85.156 ± 4.787	77.327 ± 0.932	79.959 ± 11.332	59.196 ± 6.648
	Recall	83.073 ± 3.686	76.667 ± 0.00	78.889 ± 7.281	56.562 ± 14.493
	F1-score	81.8738 ± 4.561	76.066 ± 0.85	77.349 ± 8.667	53.427 ± 12.047
Flipping	Accuracy	93.55 ± 6.45	88.71 ± 2.277	79.57 ± 3.724	69.89 ± 3.724
	Precision	90.347 ± 8.36	87.674 ± 1.227	72.444 ± 5.501	59.302 ± 5.642
	Recall	88.368 ± 10.104	85.469 ± 2.872	73.611 ± 4.811	65.139 ± 4.167
	F1-score	87.669 ± 10.734	85.372 ± 2.412	71.279 ± 3.587	60.58 ± 5.221
Shift	Accuracy	97.847 ± 1.865	83.87 ± 9.122	89.247 ± 4.928	68.82 ± 3.723
	Precision	97.917 ± 1.804	89.584 ± 3.755	80.37 ± 10.876	59.662 ± 11.447
	Recall	97.222 ± 2.405	88.889 ± 2.406	81.945 ± 4.811	61.458 ± 1.804
	F1-score	97.142 ± 2.475	88.452 ± 2.61	79.399 ± 6.579	57.545 ± 5.214
Random Erasing	Accuracy	92.473 ± 4.924	<u>89.247 ± 10.367</u>	<u>91.397 ± 1.865</u>	67.643 ± 3.225
	Precision	90.625 ± 8.268	90.995 ± 9.008	87.732 ± 7.127	58.77 ± 9.889
	Recall	90.277 ± 6.365	87.222 ± 12.811	89.757 ± 2.105	59.444 ± 3.368
	F1-score	89.106 ± 7.203	87.413 ± 12.582	87.432 ± 3.725	57.369 ± 5.022
AugMix	Accuracy	80.647 ± 5.589	80.65 ± 0.00	82.8 ± 3.724	69.355 ± 2.284
	Precision	80.585 ± 6.375	76.296 ± 5.966	77.982 ± 0.544	62.016 ± 8.998
	Recall	77.604 ± 5.13	75.556 ± 0.962	77.779 ± 4.811	62.552 ± 5.966
	F1-score	76.129 ± 5.707	73.239 ± 2.591	76.451 ± 3.211	58.969 ± 6.297
AutoAugment	Accuracy	80.647 ± 3.225	78.497 ± 1.865	86.023 ± 4.924	67.743 ± 12.905
	Precision	71.289 ± 3.902	71.782 ± 1.32	78.495 ± 2.752	57.575 ± 13.011
	Recall	75.869 ± 2.962	75.07 ± 2.604	82.501 ± 5.464	60.625 ± 13.378
	F1-score	72.717 ± 2.815	70.885 ± 2.066	73.347 ± 4.677	55.89 ± 14.091
RandAugment	Accuracy	84.977 ± 9.254	80.647 ± 3.225	80.645 ± 4.561	72.147 ± 3.539
	Precision	81.898 ± 17.228	77.557 ± 6.702	72.292 ± 3.27	52.451 ± 11.936
	Recall	81.424 ± 10.525	75.001 ± 4.166	76.112 ± 4.276	61.146 ± 9.563
	F1-score	79.834 ± 13.511	73.101 ± 2.966	72.723 ± 3.392	55.512 ± 10.849
Trivial Augment	Accuracy	78.493 ± 4.928	68.71 ± 14.699	80.647 ± 3.225	61.29 ± 8.536
	Precision	79.662 ± 9.642	61.657 ± 19.246	75.046 ± 4.561	52.075 ± 9.561
	Recall	73.09 ± 5.321	64.757 ± 12.93	75.555 ± 3.369	52.535 ± 10.133
	F1-score	72.882 ± 7.012	61.033 ± 16.181	72.895 ± 2.695	50.097 ± 9.773

ticularly noteworthy. Its randomized nature means that it does not rely on a fixed policy, and its effectiveness may fluctuate accordingly. This variability suggests that in some runs, RandAugment may select a combination of transformations that is less destructive to the dataset’s critical micro-features, leading to a better outcome compared to its more rigid counterparts.

In conclusion, the results of this comparative analysis highlight the critical role of data augmentation in achieving high performance on this specialized dataset. We found that the effectiveness of data augmentation is highly dependent on both the chosen technique and the model’s architecture. Simple, targeted transformations, such as Flipping and Shift, proved to be the most effective, significantly outperforming the baseline. Conversely, the more aggressive and automated augmentation methods were largely ineffective, and sometimes even detrimental, likely due to their tendency to distort the subtle, micro-level features essential for classification. The findings also underscore the importance of regularization, as demonstrated by the success of Random Erasing in mitigating the overfitting tendency of the more complex deep learning architectures.

3.3 Comparison with the others

The results presented in Table 3 validate our central hypothesis that by addressing data scarcity with effective data augmentation, deep learning models can surpass the performance of traditional methods in the task of diagnosing SARS-CoV-2 with genosensor images.

Table 3. Accuracy comparison with the state-of-the-art for SARS-CoV-2 diagnosis. The table presents the performance of handcrafted feature extractors, previously tested deep learning models, and our approach, which sets a new accuracy record.

Methods	Accuracy (%)
Handcrafted extractors	
AHP	90.12 ± 1.92
CLBP	95.78 ± 0.97
GLDM	90.56 ± 1.17
LCFNN	91.36 ± 1.35
Deep Learning	
DenseNet201	94.48 ± 1.05
InceptionResNetV2	86.87 ± 1.60
MobileNet	89.12 ± 1.58
Our model	
ResNet-50 with Shift DA	97.85 ± 1.86

In the original study [3], traditional methods, particularly CLBP, which achieved an accuracy of 95.78%, outperformed the deep learning models they evaluated. Our approach, however, successfully challenges this outcome. Our best-performing model, the ResNet-50 with the Shift augmentation

strategy, achieved an accuracy of 97.85%, which not only surpasses the best deep learning performance from the original study but also outperforms all traditional techniques. This underscores the potential of fine-tuned deep learning models, when combined with carefully selected data augmentation, in the task of diagnosing SARS-CoV-2 with genosensor images.

4. Conclusion

In this work, we investigated the effectiveness of different data augmentation strategies for enhancing the performance of deep learning models in diagnosing SARS-CoV-2 from genosensor images acquired through scanning electron microscopy (SEM). The primary challenge addressed in this work was the inherent limitation of dataset size in this type of material application.

Our results demonstrate that data augmentation is a fundamental technique for overcoming sample scarcity, yielding significant improvements in model performance. The comparative analysis showed that the Shift strategy, when combined with the ResNet-50 architecture, achieved a remarkable accuracy of 97.85%.

Additionally, our findings indicate that the effectiveness of data augmentation is highly dependent on the chosen technique and the model’s architecture. While simpler, focused augmentation strategies were the most successful, more aggressive techniques such as TrivialAugment and AutoAugment were largely ineffective. This reinforces our hypothesis that they may inadvertently destroy the micro-textural features essential for classification in the context of microscopy images. We also observed that the regularization provided by methods like Random Erasing was crucial to prevent overfitting for more complex models.

In summary, this work validates the application of deep learning for the detection of SARS-CoV-2 from genosensor images and provides a methodological framework for selecting optimal data augmentation strategies in data-limited contexts. Future work will aim to generalize these findings to other genosensor-based diagnostic applications and to integrate explainability techniques, such as Grad-CAM, to visualize the critical micro-features that the models exploit. Additionally, exploring Vision Transformer architectures and strategies that better preserve local image characteristics could further enhance diagnostic performance across biomedical sensing applications.

Acknowledgements

The authors acknowledge support from the São Paulo Research Foundation (FAPESP), Brazil. Process Numbers, #2024/07241-8, #2024/16261-2 and #2023/04583-2. This study was also financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES).

Author Contributions

Luan B. Guerra was responsible for conceptualization, methodology, software, investigation, validation, writing (original draft), and writing (review and editing); Ana Beatriz S. Zerati contributed to methodology, investigation, validation, writing (original draft), and writing (review and editing); Osvaldo N. Oliveira Jr. contributed to conceptualization, resources, writing (review and editing), supervision, and funding acquisition; Lucas C. Ribas contributed to conceptualization, methodology, validation, resources, writing (original draft), writing (review and editing), supervision, project administration, and funding acquisition.

References

- [1] SEO, G. et al. Rapid detection of covid-19 causative virus (sars-cov-2) in human nasopharyngeal swab specimens using field-effect transistor-based biosensor. *ACS Nano*, American Chemical Society (ACS), v. 14, n. 4, p. 5135–5142, abr. 2020. ISSN 1936-086X. Disponível em: <http://dx.doi.org/10.1021/acsnano.0c02823>.
- [2] BABAEI, A. et al. Genosensors as an alternative diagnostic sensing approaches for specific detection of virus species: A review of common techniques and outcomes. *TrAC Trends in Analytical Chemistry*, Elsevier BV, v. 155, p. 116686, out. 2022. ISSN 0165-9936. Disponível em: <http://dx.doi.org/10.1016/j.trac.2022.116686>.
- [3] SOARES, J. C. et al. Detection of a sars-cov-2 sequence with genosensors using data analysis based on information visualization and machine learning techniques. *Materials Chemistry Frontiers*, Royal Society of Chemistry (RSC), v. 5, n. 15, p. 5658–5670, 2021. ISSN 2052-1537. Disponível em: <http://dx.doi.org/10.1039/D1QM00665G>.
- [4] RODRIGUES, V. C. et al. Electrochemical and optical detection and machine learning applied to images of genosensors for diagnosis of prostate cancer with the biomarker pca3. *Talanta*, Elsevier BV, v. 222, p. 121444, jan. 2021. ISSN 0039-9140. Disponível em: <http://dx.doi.org/10.1016/j.talanta.2020.121444>.
- [5] SHORTEN, C.; KHOSHGOFTAAR, T. M. A survey on image data augmentation for deep learning. *Journal of Big Data*, Springer Science and Business Media LLC, v. 6, n. 1, jul. 2019. ISSN 2196-1115. Disponível em: <http://dx.doi.org/10.1186/s40537-019-0197-0>.
- [6] GOCERI, E. Medical image data augmentation: techniques, comparisons and interpretations. *Artificial Intelligence Review*, Springer Science and Business Media LLC, v. 56, n. 11, p. 12561–12605, mar. 2023. ISSN 1573-7462. Disponível em: <http://dx.doi.org/10.1007/s10462-023-10453-z>.
- [7] CUNHA, I. et al. Machine learning in microscopy – insights, opportunities and challenges. *Journal of Cell Science*, The Company of Biologists, v. 137, n. 20, out. 2024. ISSN 1477-9137. Disponível em: <http://dx.doi.org/10.1242/jcs.262095>.
- [8] RIBAS, L. C.; CASACA, W.; FARES, R. T. Conditional generative adversarial networks and deep learning data augmentation: A multi-perspective data-driven survey across multiple application fields and classification architectures. *AI, MDPI*, v. 6, n. 2, p. 32, 2025.
- [9] SCABINI, L. et al. A comparative survey of vision transformers for feature extraction in texture analysis. *Journal of Imaging*, MDPI, v. 11, n. 9, p. 304, 2025.
- [10] WELSMAN, J. A. et al. Enhancing electron microscopy image classification using data augmentation. In: *SC24-W: Workshops of the International Conference for High Performance Computing, Networking, Storage and Analysis*. [S.l.: s.n.], 2024. p. 64–71.
- [11] ZHONG, Z. et al. Random erasing data augmentation. *Proceedings of the AAAI Conference on Artificial Intelligence*, Association for the Advancement of Artificial Intelligence (AAAI), v. 34, n. 07, p. 13001–13008, abr. 2020. ISSN 2159-5399. Disponível em: <http://dx.doi.org/10.1609/aaai.v34i07.7000>.
- [12] CUBUK, E. D. et al. Autoaugment: Learning augmentation strategies from data. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.]: IEEE, 2019. p. 113–123.
- [13] CUBUK, E. D. et al. Randaugment: Practical automated data augmentation with a reduced search space. In: LAROCHELLE, H. et al. (Ed.). *Advances in Neural Information Processing Systems*. [S.l.]: Curran Associates, Inc., 2020. v. 33, p. 18613–18624.
- [14] MÜLLER, S. G.; HUTTER, F. Trivialaugment: Tuning-free yet state-of-the-art data augmentation. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. Montreal, QC, Canada: IEEE, 2021. p. 774–782.
- [15] HENDRYCKS, D. et al. *AugMix: A Simple Data Processing Method to Improve Robustness and Uncertainty*. arXiv, 2019. Disponível em: <https://arxiv.org/abs/1912.02781>.
- [16] KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.