

Ensemble of learning transfer models with GAN for data augmentation: Application to skin cancer detection

Fernando Humberto de Almeida Moraes Neto^{1†}, Adriano Kamimura Suzuki¹, Francisco Louzada Neto¹, Ricardo Rocha², Hellen Oliveira da Paz².

¹Universidade de São Paulo; Instituto de Ciências Exatas; Departamento de Estatística; São Carlos-SP, Brasil.

²Universidade Federal da Bahia; Instituto de Matemática e Estatística; Departamento de Estatística; Salvador-BA, Brasil.

Abstract: Deep Learning models typically require large datasets to achieve high performance, which can be challenging when applied to medical data due to the difficulty in acquiring sufficient observations, particularly when the data is imbalanced between categories like pathological and normal cases. To address this, approaches like transfer learning, ensemble methods, and data augmentation through Generative Adversarial Networks (GANs) can be utilized to improve classification tasks on small datasets. In situations with imbalanced data, resampling techniques are also crucial. This research combines transfer learning with resampling methods to boost the prediction accuracy of minority class samples in small, imbalanced datasets. Additionally, techniques like data retracing and GAN-based augmentation are applied. The dataset used includes small, imbalanced images of skin cancer, aimed at classifying them as malignant or benign. The findings reveal that models using resampling techniques achieve better results, while those without resampling underperform. This underscores the benefit of resampling in enhancing prediction, particularly for the minority class. Furthermore, using GANs for data augmentation improves model performance over those that do not incorporate this technique.

Keywords: Convolutional Neural Network; Transfer Learning; Ensemble; Data augmentation.

Introduction

Convolutional Neural Networks (CNN) are models used to analyze images. But the performance of these models can be affected by the number of observations, for example when there are few observations available for training these models. These models require a large amount of data to perform well.

Several studies have been carried out with the aim of solving this problem. For example, Nanni et al. (2022) *et al* present an ensemble with several activation functions for image classification tasks with small data sets. Additionally, Saravanan et al. (2022) uses data augmentation to augment the observations from the database used of brain tumors. Another widely used approach is the use of transfer learning models, for example, Ju et al. (2022) propose a transfer learning model to be used in a small database to identify defects in images.

Another approach that can be used as data augmentation is the generation of synthetic data through Generative Adversarial Network (GAN), according to Chen et al. (2022) synthetic medical images coming from a GAN can be used to supply the small number of images that make up the most diverse medical image banks. Anjos et al. (2020) uses GAN to generate synthetic data and after generation uses the data to diagnose lung cancer.

Recently models involving combinations of transfer learning models have been proposed, Zheng et al. (2022) proposed a model using an Ensemble of Transfer Learning Models (ETCNN) to classify mammography images, the authors also used data augmentation in the training set. Vallabhajosyula et al. (2022) created an automatic detection model for detecting plant diseases, using transfer models and using data augmentation in the training set.

[†]Corresponding author: moraesfernando.mat@gmail.com

Manuscript received: 26/12/2024; Revised: 09/06/2025; Accepted: 02/07/2025

According to Hashemi et al. (2018) one of the biggest challenges in CNN training arises when the data is unbalanced, which is common in many medical imaging applications, such as lesion segmentation. Learning from an imbalanced data set presents a tricky problem in which traditional learning algorithms perform poorly. Traditional classifiers usually aim to optimize the overall accuracy without considering the relative distribution of each class Chakraborty and Chakraborty (2020).

To solve the problem in unbalanced databases, resampling techniques can be used to balance the training database. For example, Roy et al. (2022) uses oversampling and undersampling to balance the training set on a chest x-ray set to identify COVID-19.

This study aims to apply a new approach to improve the performance of CNN's in small unbalanced databases using transfer learning methods, ensemble, data augmentation with GAN and resampling techniques (undersampling and oversampling).

This article is organized as follows: related work, the methodology and dataset, the performance assessment and results, conclusions and information on future work are presented.

Related work

In the literature there are several works that use CNN's to detect cancer or other diseases: Alsaif et al. (2022) for detecting brain tumors, Samee et al. (2022) and Talatian Azad et al. (2022) for detecting breast cancer and Alrashedy et al. (2022) for detection of COVID-19. These models stand out in the area of image analysis for balanced databases and have provided scientific advances in several areas of knowledge.

Several works have been proposed to solve the imbalance in small unbalanced databases. The unequal (unbalanced) proportion of classes in image databases means that a CNN may have difficulty in matching data from the class with the smallest number of observations. It is very common to find medical data work that is unbalanced and some strategies are used to mitigate this imbalance.

Many authors use resampling techniques such as oversampling and undersampling. Manjunath et al. (2022) uses oversampling and undersampling to detect COVID-19 through X-rays. Mukhin et al. (2022) propose a data balancing method for biomedical data using oversampling. Ferreira and Suzuki (2024) explore the use of oversampling and undersampling combined with Weighted Focal Loss in Extreme Gradient Boosting (XGBoost) to reduce expected misclassification costs in credit scoring.

Another widely used approach is data augmentation, as CNN's need large data sets. Without the existence of many images in datasets, different deep learning models will not be able to learn and produce accurate models (Khalifa et al., 2022). There are different types of ways to augment data, for example, GANs have been widely used in computer vision for data augmentation (Mirza and Osindero, 2014). According to Huang et al. (2020) these networks do not require the data to follow any specific assumptions when modeling complex data, even with some implicit distributions, being naturally suitable for the data generation task.

Several authors research the use of GANs for data augmentation, Dai et al. (2022) and Lee and Park (2021) use data augmentation based on GANs to generate artificial data samples for smaller classes. Ali-Gombe and Elyan (2019) propose a new data augmentation approach using a GAN to deal with the class imbalance problem.

There are still works that use the combination of models and GANs, for example, Ding et al. (2023) propose a hybrid structure (RGAN-EL) combining generative adversarial networks and ensemble learning method to improve the performance of classifying unbalanced data.

Huang et al. (2020) propose a new ensemble structure based on GANs and an effective data cleaning way to solve the class imbalance problem for climate classification and they use the combination of transfer learning models (ETCINN) of learning and data augmentation based on GAN, Chatterjee et al. (2022) make this use for the automatic classification of plastic bottles, Baldine et al. (2024) for the classification of diseases in apple tree leaves, Bharathi Raja and Selvi Rajendran (2023) for detecting diseases caused by pathogens, insects and parasitic plants and Al-Rasheed et al. (2022) use it for automatic detection of breast cancer skin.

The main goal of this study is to show that using ETCNN with resampling and using a GAN for data augmentation is highly effective for predicting minority class observations in imbalanced databases.

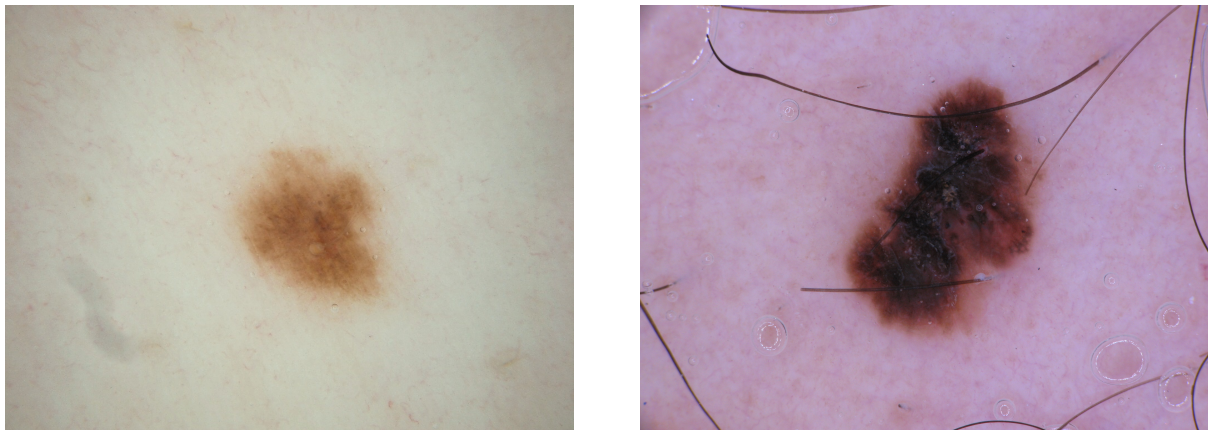
Methodology

Below are details about the ISIC 2016 dataset, transfer learning models, preprocessing, data augmentation with GAN, resampling, ensembles used, evaluation criteria and metrics, and the method.

ISIC Challenge Datasets

The database used in this work is provided by the International Skin Imaging Collaboration (ISIC) Challenge Datasets 2016 (Gutman et al., 2016), this is a small unbalanced database on malignant or benign dermatoscopic lesions, an example of the images present in this database is given in Figure 1. The distribution of images is presented in Table 1.

Figure 1: Sample mages of ISIC Dataset 2016: benign image (left) and malignant image (right).



Source: (Gutman et al., 2016)

Table 1: Number of elements belonging to each class in the ISIC dataset

Class	Number of elements (benign)	Number of elements (malignant)	Total
train	727	173	900
test	304	75	379

Source: From the authors (2025).

The images had different sizes and presented noise, so a pre-processing step was necessary, after pre-processing the final image had dimensions 256×256 .

To solve the classification problem when there are few observations of a given data, data augmentation techniques can be used. In order to obtain a good generalization of the model, a GAN was used in this work to increase the number of samples in the minority class instead of simply using common data augmentation.

GANs consist of two models. According to Chen et al. (2022), one model is the generator (G), which generates new data, while the other is the discriminator (D), which evaluates whether the generated data is real or synthetic.

Resampling was performed using undersampling and oversampling methods in the training set. As shown in Table 1, the subset size for undersampling consisted of 173 images from both the malignant and benign classes, while for oversampling, 727 images were used for each class.

Proposed transfer learning models

As highlighted by Goodfellow et al. (2016), Convolutional Neural Networks (CNNs) represent a class of Deep Learning architectures that incorporate the convolution operation in at least one of their layers. In the scope of this study, we chose to use previously trained CNNs, an approach known as transfer learning. The transfer learning models used in this research include: VGG16 (von Wangenheim, 2018), Densenets (Huang et al., 2017), Inception Resnetv2 (Szegedy et al., 2017), Resnets (He et al., 2016), Xception (Chollet, 2017) and MobileNet (Howard et al., 2017). These models are available in the Keras package, and additional details can be found at (Sarkar et al., 2018). For ResNet and DenseNet, some variations were tested. All experiments were performed on Google Colab GPUs (GOOGLE, 2017). The main code is available at : https://github.com/FernandooMoraes/Article-Ensemble-of-transfer-learning-models_GAN.

For training the networks, the K-fold with 10 folds was considered, the training algorithm used was Adagrad, we used dropout to avoid overfitting and early stopping as a stopping criterion, with 100 epochs, finally the loss function used was Cross-Entropy loss for binary classification.

The ensemble is performed as follows: if at least i models classify an instance as the minority class, the final prediction is assigned to this class. The parameter i varies from 1 to 13.

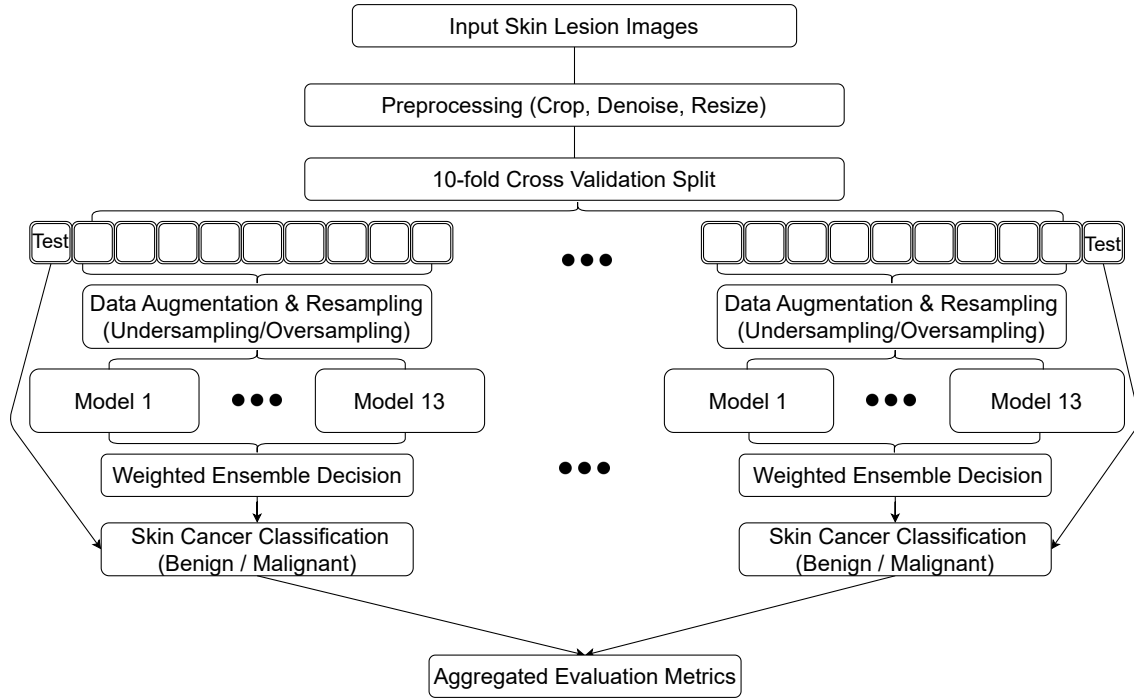
It should be noted that all models were trained with images from the training set, and, at the end, we proceeded to create the ensembles. It is important to highlight that the application of data augmentation, undersampling and oversampling is performed only on the training set of each fold.

The flowchart of the methodology used in this study is presented in Figure 2. The methodology consists of the steps: Pre-processing, K-fold, resampling, data augmentation with GAN, modeling of the 13 transfer learning CNNs, ensembles and model evaluation.

To evaluate the fit of the model in the classification problem, some performance measures are used based on the confusion matrix. The evaluation measures used were: precision (P), specificity (S) and F Beta Score. The F-beta score is a widely used performance measure for binary classification with unbalanced data (Lee et al., 2021). According to the same author, if $\beta < 1$ the metric emphasizes precision, whereas when $\beta > 1$ it prioritizes recall.

In this work we will adopt an F₂-score metric as an alternative to mitigate this balancing effect in the malignant class.

Figure 2: Representation of the modeling used step by step.



Source: from the authors (2025).

Results and Discussion

In this section the results on ISIC are presented. For modeling, 13 transfer learning models were used and after this 13 combinations were used to predict dermatological images of skin lesions.

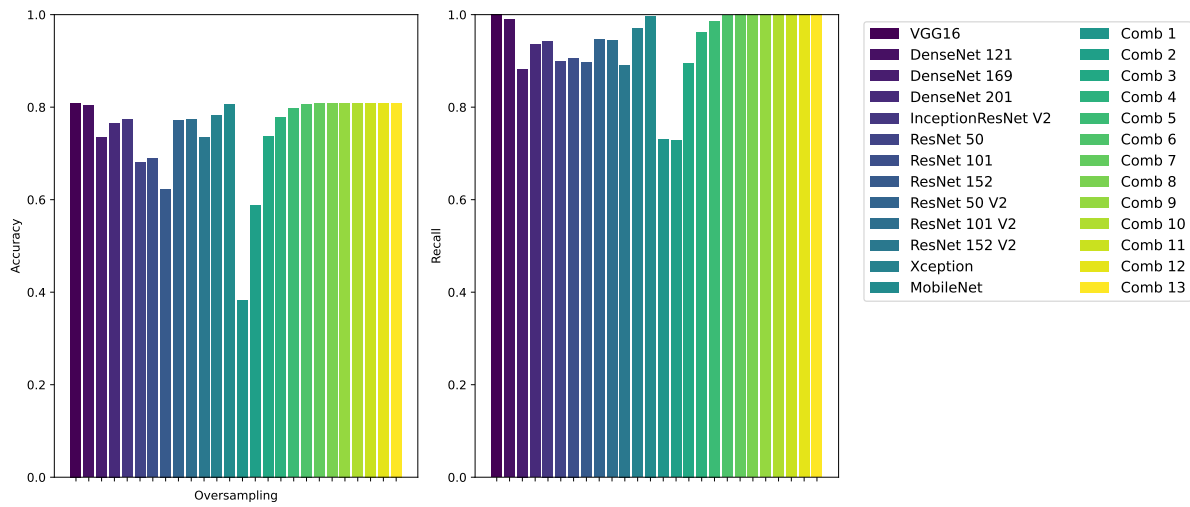
Figure 3 shows the accuracy and recall metrics for the individual models and the proposed modifications. Since only 20% of the observations in the original dataset belong to the malignant class, identifying these cases is challenging. This is reflected in a recall value close to 1 and an accuracy of approximately 80%.

This pattern of trade-off is well documented in the literature. According to Chakraborty and Chakraborty (2020) and Hashemi et al. (2018), CNNs trained on imbalanced datasets tend to prioritize the majority class, leading to inflated accuracy metrics that fail to reflect the model's performance on the minority class.

In turn, Figure 4 displays the Precision and Specificity metrics. Note that all individual models have specifications for identifying malignant observations. However, the proposed methods exhibit superior performance, particularly in the cases of pigments 1 and 2, which achieved values above 20%.

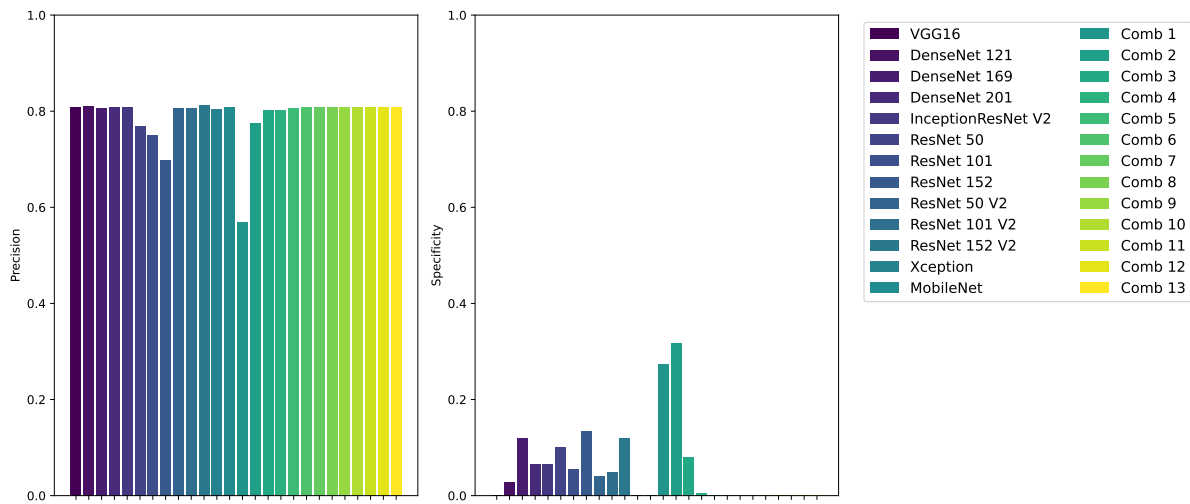
Finally, Figure 5 presents the results of the F_2 -scores for resampling methods (undersampling and oversampling) without data augmentation. Note that the F_2 -score of the best combination in oversampling reaches approximately 60%, outperforming the individual models. In the case of subsampling, it is also verified that the F_2 -score of the proposed changes exceeds the performance of the individual models. These results reinforce the benefits of ensemble approaches. Comb 1 and Comb 3 clearly outperformed individual models in terms of F_2 -score, demonstrating better sensitivity to the minority class. This corroborates the findings of Zheng et al. (2022) and Ding et al. (2023), which emphasize the advantages of ensemble learning in medical image classification under imbalance.

Figure 3: Result of Accuracy and Recall of models and combinations used with no data augmentation and resampling.



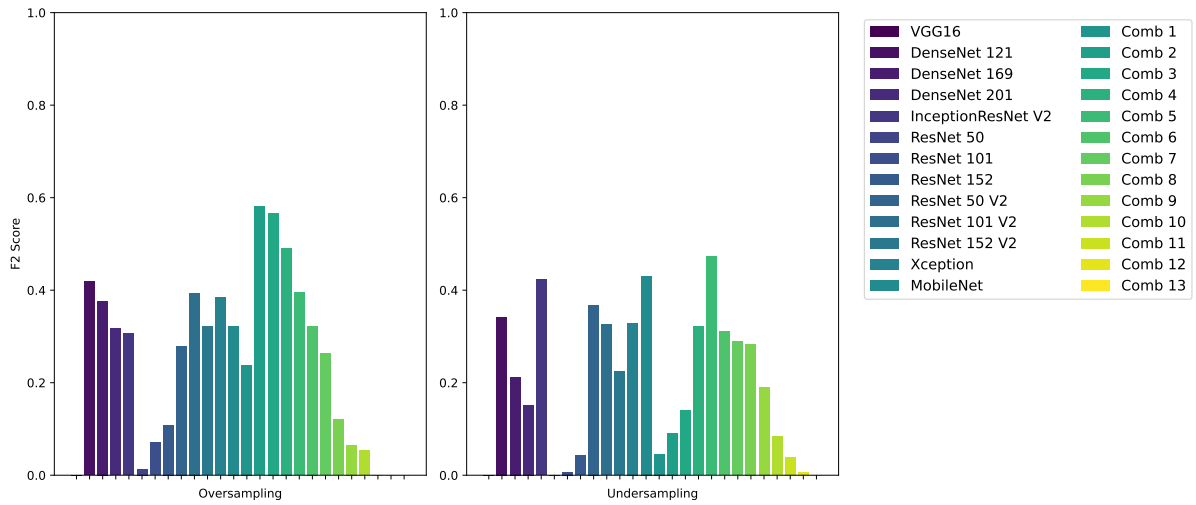
Source: from the authors (2025).

Figure 4: Result of Precision and Specificity of models and combinations used with no data augmentation and resampling.



Source: from the authors (2025).

Figure 5: Result of F_2 -score of models and combinations used with resampling and no data augmentation.



Source: from the authors (2025).

Table 2 presents the results referring to F_2 -score in the scenario I (no resampling) and in undersampling scenario II (best among scenarios with resampling). In scenario I, the model with the best F_2 -score is Inception Resnet V2, considering the combinations the best F_2 -score is combination 1. For scenario II, the model with the best F_2 -score Resnet 50, for the combinations the one with the best F_2 -score is combination 3.

Table 2: Result of the best F_2 -score for the models and combinations with data augmentation for scenario I and II.

Models	Scenario I	Scenario II
Inception_resnet_v2	0.34	0.31
Resnet50	0.21	0.39
Comb 1	0.46	0.23
Comb 3	0.30	0.60

Source: From the authors (2025).

From this result, it is concluded that the best F_2 -score are obtained by using Ensemble Transfer learning with resampling and data augmentation using GAN. As reported by Huang et al. (2020) and Chatterjee et al. (2022), GANs can help mitigate scarcity by generating diverse and plausible synthetic samples. In the two scenarios shown, ETCCN had better F_2 -scores than the individual models. Furthermore, it is important to note that although the best Undersampling models have approximately 60% F_2 -score, this metric is low.

Table 3 compares the proposed method with previous work on the ISIC 2016 dataset on the same test set, Zhou et al. (2020) achieved a marginally higher F_2 -score, and Codella et al. (2017), who achieved high specificity and precision. it's crucial to acknowledge the inherent limitations of comparing across diverse datasets and preprocessing methodologies.

Table 3: Comparison with other works using the test dataset provided by ISIC 2016.

	F ₂ -score	Specificity	Precision
Zhou et al. (2020)	0.49	0.99	-
	0.69	0.95	-
Majtner et al. (2016)		0.87	
Lopez et al. (2017)	-	0.79	0.80
Codella et al. (2017)		0.55	0.93
	-	0.94	0.21
Gutman et al. (2016)	-	0.92	-
	-	0.94	-
Oversampling	0.53	0.88	0.20

Source: From the authors (2025).

Our approach, integrating transfer learning, GAN-assisted data augmentation, and a carefully calibrated resampling strategy, appears to mitigate the impact of the ISIC 2016 dataset's class imbalance, resulting in a performance profile that is demonstrably robust. However, further investigation is warranted to establish more consistent and reliable metrics. While preliminary results suggest an advantage, a more comprehensive analysis incorporating a wider range of evaluation criteria and potentially exploring the relative contributions of each component within the ETCNN architecture is necessary to definitively conclude the extent of our model's superiority. It's essential to move beyond qualitative assessments and towards rigorous quantitative validation to fully characterize the model's performance characteristics.

Even if some metrics are lower than those reported in prior work, the proposed methodology offers valuable insights for future research, especially by highlighting the practical challenges encountered in real world scenarios with limited data availability.

A thorough analysis of the results indicates that the absence of resampling and data augmentation makes it difficult to accurately classify malignant cases. The combined approaches demonstrate superior metrics across all scenarios when compared to individual models. It is noteworthy that the combination that achieved the highest F₂-score was observed in scenario II, where there was an increase of around 53% in the data through the application of the GAN technique.

Results showed that combining transfer learning models with resampling and data augmentation techniques using GANs performed better in predicting the minority class in imbalanced databases. The primary contribution of this study lies in the introduction of an improved ETCNN, employing resampling and data augmentation with GAN to enhance the detection of malignant cancer in the analyzed dataset. This methodology has proven particularly effective in improving the detection of malignant lesions.

In summary, our findings support the view that tackling class imbalance in small medical datasets requires a multipronged approach. The ETCNN, combined with data augmentation via GANs, presents a promising methodology. The proposed framework offers a solid foundation for further research and practical deployment in real world, resource constrained medical contexts.

Final considerations

In this study, we introduce an approach that combines transfer learning models with resampling and data augmentation techniques. The goal is to build a CNN model capable of accurately identifying malignant cancer in small, imbalanced datasets. To demonstrate the effectiveness of these combinations, we conducted an experiment using a dataset of dermoscopic lesions, where we observed that the combined approaches stood out in the studied context.

We generally conclude that the proposed methodology has the ability to accurately predict observations belonging to the minority class in imbalanced datasets, making it a valuable tool for medical image analysis.

It is worth noting that the problem faced in this study is the difficulty of obtaining a significant number of observations in medical data sets, especially with an unequal distribution between the different classes (pathological and normal). To mitigate this problem, the research uses transfer learning techniques, ensembles, data augmentation through GANs, and resampling techniques.

This approach is especially useful for medical applications where obtaining a large number of balanced observations can be challenging, such as detecting skin cancer from dermoscopic images.

In future work, we plan to compare the performance of the current model, trained on imbalanced data, with a forthcoming model trained on balanced data, incorporating more real images in the malignant class. Additionally, we aim to conduct further studies on the proposed model, exploring alternative generalization functions, applying various resampling methods, and expanding the range of transfer learning models utilized.

This work explored the use of transfer learning models widely recognized for their performance in image classification tasks. The selection of these models, despite the existence of more recent and complex architectures, was motivated by their proven robustness and effectiveness in similar situations. The study exclusively used the ISIC 2016 dataset, which serves as a solid basis for skin cancer classification, especially in small and unbalanced databases. However, it is recognized that this approach may limit the assessment of the generalizability of the proposed method. Future research should investigate newer models and diversify datasets to validate the methodology in different scenarios.

In summary, our findings support the view that tackling class imbalance in small medical datasets requires a multipronged approach. The ETCNN, combined with data augmentation via GANs, presents a promising methodology. The proposed framework offers a solid foundation for further research and practical deployment in real world, resource constrained medical contexts.

Acknowledgements

Thanks to the National Council for Scientific and Technological Development (CNPQ) and Coordination for the Improvement of Higher Education Personnel (CAPES) for the financial support.

References

- Al-Rasheed, A., Ksibi, A., Ayadi, M., Alzahrani, A. I., Zakariah, M., and Ali Hakami, N. (2022). An ensemble of transfer learning models for the prediction of skin cancers with conditional generative adversarial networks. *Diagnostics*, 12(12):3145.
- Ali-Gombe, A. and Elyan, E. (2019). Mfc-gan: Class-imbalanced dataset classification using multiple fake class generative adversarial network. *Neurocomputing*, 361:212–221.
- Alrashedy, H. H. N., Almansour, A. F., Ibrahim, D. M., and Hammoudeh, M. A. A. (2022). Braingan: Brain mri image generation and classification framework using gan architectures and cnn models. *Sensors*, 22(11):4297.
- Alsaif, H., Guesmi, R., Alshammari, B. M., Hamrouni, T., Guesmi, T., Alzamil, A., and Belguesmi, L. (2022). A novel data augmentation-based brain tumor detection using convolutional neural network. *Applied Sciences*, 12(8):3773.

- Anjos, B. H. L. d. et al. (2020). Predcgan: uma abordagem para geração de nódulos pulmonares sintéticos gan usando pré-treinamento. Master's thesis, Universidade Federal de Alagoas, Maceió, AL.
- Baldine, R. B., Fonseca, K. V. d. S., and Ferreira, E. B. (2024). Deep learning classification of apple leaf diseases: comparison of neural networks. *Sigmae: Revista Eletrônica da Estatística*, 13(5).
- Bharathi Raja, N. and Selvi Rajendran, P. (2023). An efficient banana plant leaf disease classification using optimal ensemble deep transfer network. *Journal of Experimental & Theoretical Artificial Intelligence*, pages 1–24.
- Chakraborty, T. and Chakraborty, A. K. (2020). Superensemble classifier for improving predictions in imbalanced datasets. *Communications in Statistics: Case Studies, Data Analysis and Applications*, 6(2):123–141.
- Chatterjee, S., Hazra, D., Byun, Y.-C., and Kim, Y.-W. (2022). Enhancement of image classification using transfer learning and gan-based synthetic data augmentation. *Mathematics*, 10(9):1541.
- Chen, Y., Yang, X.-H., Wei, Z., Heidari, A. A., Zheng, N., Li, Z., Chen, H., Hu, H., Zhou, Q., and Guan, Q. (2022). Generative adversarial networks in medical image augmentation: a review. *Computers in Biology and Medicine*, page 105382.
- Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1251–1258.
- Codella, N. C., Nguyen, Q.-B., Pankanti, S., Gutman, D. A., Helba, B., Halpern, A. C., and Smith, J. R. (2017). Deep learning ensembles for melanoma recognition in dermoscopy images. *IBM Journal of Research and Development*, 61(4/5):5–1.
- Dai, W., Li, D., Tang, D., Wang, H., and Peng, Y. (2022). Deep learning approach for defective spot welds classification using small and class-imbalanced datasets. *Neurocomputing*, 477:46–60.
- Ding, H., Sun, Y., Wang, Z., Huang, N., Shen, Z., and Cui, X. (2023). Rgan-el: A gan and ensemble learning-based hybrid approach for imbalanced data classification. *Information Processing & Management*, 60(2):103235.
- Ferreira, G. A. and Suzuki, A. K. (2024). Adaptations of extreme gradient boosting for imbalanced datasets with application in credit scoring. *Sigmae: Revista Eletrônica da Estatística*, 13(4).
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press. <http://www.deeplearningbook.org>.
- GOOGLE (2017). Google colab. <https://colab.research.google.com/>. Acesso em: 6 jun. 2025.
- Gutman, D., Codella, N. C., Celebi, E., Helba, B., Marchetti, M., Mishra, N., and Halpern, A. (2016). Skin lesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (isbi) 2016, hosted by the international skin imaging collaboration (isic). *arXiv preprint arXiv:1605.01397*.
- Hashemi, S. R., Salehi, S. S. M., Erdogmus, D., Prabhu, S. P., Warfield, S. K., and Gholipour, A. (2018). Asymmetric loss functions and deep densely-connected networks for highly-imbalanced medical image segmentation: Application to multiple sclerosis lesion detection. *IEEE Access*, 7:1721–1735.

- He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.
- Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708.
- Huang, Y., Jin, Y., Li, Y., and Lin, Z. (2020). Towards imbalanced image classification: a generative adversarial network ensemble learning method. *IEEE Access*, 8:88399–88409.
- Ju, J., Zheng, H., Xu, X., Guo, Z., Zheng, Z., and Lin, M. (2022). Classification of jujube defects in small data sets based on transfer learning. *Neural Computing and Applications*, pages 1–14.
- Khalifa, N. E., Loey, M., and Mirjalili, S. (2022). A comprehensive survey of recent trends in deep learning for digital images augmentation. *Artificial Intelligence Review*, pages 1–27.
- Lee, J. and Park, K. (2021). Gan-based imbalanced data intrusion detection system. *Personal and Ubiquitous Computing*, 25:121–128.
- Lee, N., Yang, H., and Yoo, H. (2021). A surrogate loss function for optimization of f beta score in binary classification with imbalanced data. *arXiv preprint arXiv:2104.01459*.
- Lopez, A. R., Giro-i Nieto, X., Burdick, J., and Marques, O. (2017). Skin lesion classification from dermoscopic images using deep learning techniques. In *2017 13th IASTED international conference on biomedical engineering (BioMed)*, pages 49–54. IEEE.
- Majtner, T., Yildirim-Yayilgan, S., and Hardeberg, J. Y. (2016). Combining deep learning and hand-crafted features for skin lesion classification. pages 1–6.
- Manjunath, S. M., Gurjar, M., O’Kane, N., McCarren, A., and Gualano, L. (2022). Detection of covid 19 from an imbalanced chest x-ray image data set.
- Mirza, M. and Osindero, S. (2014). Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*.
- Mukhin, A. V., Kilbas, I. A., Paringer, R. A., Ilyasova, N. Y., and Kupriyanov, A. V. (2022). A method for balancing a multi-labeled biomedical dataset. *Integrated Computer-Aided Engineering*, (Preprint):1–17.
- Nanni, L., Brahnam, S., Paci, M., and Ghidoni, S. (2022). Comparison of different convolutional neural network activation functions and methods for building ensembles for small to midsize medical data sets. *Sensors*, 22(16):6129.
- Roy, S., Tyagi, M., Bansal, V., and Jain, V. (2022). Svd-clahe boosting and balanced loss function for covid-19 detection from an imbalanced chest x-ray dataset. *Computers in Biology and Medicine*, 150:106092.
- Samee, N. A., Atteia, G., Meshoul, S., Al-antari, M. A., and Kadah, Y. M. (2022). Deep learning cascaded feature selection framework for breast cancer classification: Hybrid cnn with univariate-based approach. *Mathematics*, 10(19):3631.

- Saravanan, T., Karthiha, K., Kavinkumar, R., Gokul, S., and Mishra, J. P. (2022). A novel machine learning scheme for face mask detection using pretrained convolutional neural network. *Materials Today: Proceedings*, 58:150–156.
- Sarkar, D., Bali, R., and Ghosh, T. (2018). *Hands-On Transfer Learning with Python: Implement advanced deep learning and neural network models using TensorFlow and Keras*. Packt Publishing Ltd.
- Szegedy, C., Ioffe, S., Vanhoucke, V., and Alemi, A. A. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning. In *Thirty-first AAAI conference on artificial intelligence*.
- Talatian Azad, S., Ahmadi, G., and Rezaeipana, A. (2022). An intelligent ensemble classification method based on multi-layer perceptron neural network and evolutionary algorithms for breast cancer diagnosis. *Journal of Experimental & Theoretical Artificial Intelligence*, 34(6):949–969.
- Vallabhajosyula, S., Sistla, V., and Kolli, V. K. K. (2022). Transfer learning-based deep ensemble neural network for plant leaf disease detection. *Journal of Plant Diseases and Protection*, 129(3):545–558.
- von Wangenheim, A. (2018). Deep learning::reconhecimento de imagens. http://www.lapix.ufsc.br/ensino/visao/visao-computacionaldeep-learning/deep-learningreconhecimento-de-imagens#Modelos_de_VGG. acessado: 20-06-2019.
- Zheng, Y., Li, C., Zhou, X., Chen, H., Xu, H., Li, Y., Zhang, H., Li, X., Sun, H., Huang, X., et al. (2022). Application of transfer learning and ensemble learning in image-level classification for breast histopathology. *Intelligent Medicine*.
- Zhou, Q., Ren, C., and Qi, S. (2020). An imbalanced r-stdp learning rule in spiking neural networks for medical image classification. *IEEE Access*, 8:224162–224177.