Convolutional Neural Networks Enriched by Handcrafted Attributes (Enriched-CNN): An Innovative Approach to Pattern Recognition in Histological Images

Luiz Fernando Segato dos Santos[1][a], Leandro Alves Neves[1][b], Alessandro Santana Martins[2][c], Guilherme Freire Roberto[3][d], Thaína Aparecida Azevedo Tosta[4][e] and Marcelo Zanchetta do Nascimento[5][f]

[1] *Department of Computer Science and Statistics (DCCE), São Paulo State University (UNESP), São José do Rio Preto, Brazil*
[2] *Federal Institute of Triângulo Mineiro (IFTM), Ituiutaba, Brazil*
[3] *Institute of Mathematics and Computer Science (ICMC), University of São Paulo (USP), São Carlos, Brazil*
[4] *Federal University of São Paulo (UNIFESP), São José dos Campos, Brazil*
[5] *Faculty of Computer Science (FACOM), Federal University of Uberlândia (UFU), Uberlândia, Brazil*

Keywords: Enriched Backpropagation, Fractal Techniques, Loss Functions, CNN Architectures, H&E Images.

Abstract: This paper presents a novel method called Enriched-CNN, designed to enrich CNN models using handcrafted features extracted from multiscale and multidimensional fractal techniques. These features are incorporated directly into the loss function during model training through specific strategies. The method was applied to three important histological datasets for studying and classifying H&E-stained samples. Several CNN architectures, such as ResNet, InceptionNet, EfficientNet, and others, were tested to understand the enrichment behavior in different scenarios. The best results achieved accuracy rates ranging from 93.75% to 100% for enrichment situations involving only 3 to 5 features. This paper also provides significant insights into the conditions that most contributed to the process and allowed competitiveness compared to the specialized literature, such as the possibility of composing models with minimal or no structural changes. This unique aspect enables the method to be applied to other types of neural architectures.

1 INTRODUCTION

Enriching convolutional models have been explored and applied in histopathological contexts to improve diagnostic support systems and pattern recognition (Roberto et al., 2021; Longo et al., 2023; Tenguam et al., 2024). This approach has yielded various benefits, such as improving model performance or resolving training issues like overfitting (Jahan et al., 2022). Some studies combine different types of attributes aiming to enrich their models. The most well-known attributes in this context are handcrafted and deep-learned.

Among handcrafted attributes, approaches based on multiscale and/or multidimensional fractal techniques stand out, especially in the histopathology field (Roberto et al., 2021; Ivanovici and Richard, 2011). Deep-learned attributes, on the other hand, include those obtained through convolutional neural networks (CNN) (Nanni et al., 2020). For instance, network models are applied to image samples, and the values obtained during training enable the formation of feature vectors. It is also crucial to highlight that the training of these neural networks and the optimal representation of data are facilitated by an algorithm known as backpropagation, which strengthens the most relevant weights throughout training. During this stage, a cost function, also known as loss function, generates a scalar value indicating how well a

[a] https://orcid.org/0009-0008-0851-4379
[b] https://orcid.org/0000-0001-8580-7054
[c] https://orcid.org/0000-0003-4635-5037
[d] https://orcid.org/0000-0001-5883-2983
[e] https://orcid.org/0000-0002-9291-8892
[f] https://orcid.org/0000-0003-3537-0178

Santos, L. F. S., Neves, L. A., Martins, A. S., Roberto, G. F., Tosta, T. A. A. and Zanchetta do Nascimento, M.
Convolutional Neural Networks Enriched by Handcrafted Attributes (Enriched-CNN): An Innovative Approach to Pattern Recognition in Histological Images.
DOI: 10.5220/0013277300003929

In *Proceedings of the 27th International Conference on Enterprise Information Systems (ICEIS 2025) - Volume 1*, pages 467-478
ISBN: 978-989-758-749-8; ISSN: 2184-4992

sample is classified. This value can be minimized or maximized by an optimizer, depending on the training objective.

Some combinations have integrated handcrafted and deep-learned attributes (Nanni et al., 2020). These combinations have contributed to various areas (Zheng et al., 2023), but some challenges still persist, such as providing appropriate dimensions and scale for the involved attributes (Cheng et al., 2023) or preventing redundancy in combining handcrafted and deep-learned features (Zheng et al., 2023). To address these divergences in the combination process, some modifications in the architectures are required, resulting in more complex models with specific adjustments for each architecture type (Zheng et al., 2023).

On the other hand, enrichment can also be explored through backpropagation. This process offers several strategies that incorporate rewards or penalties directly into the loss function, depending on the training objective. The possible types of incorporation into the loss function primarily occur through sum-based rules (Hosseini et al., 2023) or weighted sum-based rules (Wu et al., 2023). The information incorporated in this process is usually backpropagated to the network to update the model weights (Diao et al., 2023). Despite contributions on the topic in the medical imaging field (Diao et al., 2023), there is no research that has explored model enrichment through fractal descriptors directly in the loss function to investigate model performance and training behavior.

Therefore, researching how models could be enriched using handcrafted features such as multiscale and multidimensional fractal attributes, as well as through the loss function, and exploring the potential forms of enrichment, including multiple attribute combinations, is yet to be explored in the literature.

1.1 Research Directions and Contributions

Research has indicated that combining distinct attributes, such as deep-learned and handcrafted features, is a crucial path to improving pattern recognition systems, regardless of the application context (Roberto et al., 2021; Nanni et al., 2020). This is particularly evident when different attribute sources are considered (Sukegawa et al., 2022). However, despite attention modules and additional data fusion mechanisms improving results (Montalbo, 2022), the models developed using these strategies still have limitations (Zheng et al., 2023).

Alternatively, model enrichment research, primarily through backpropagation, has overcome some of these limitations, such as information redundancy from feature fusion processes and the incompatibility of feature dimensions and scales (Xu et al., 2022), using less complex models and achieving promising results. Moreover, backpropagation studies have enabled improvements in distinction rates and reduced training costs (Xu et al., 2022), leading to advancements in addressing vanishing gradient issues (Hu et al., 2021) and overfitting, especially in the presence of class imbalance (Zhang et al., 2024). Different types of loss functions have also been considered for various scenarios, contributing distinctively to model enrichment (Xu et al., 2022; Zhang et al., 2024).

When handcrafted features are incorporated into these functions through specific rules, such as weighted sum-based rules (Xu et al., 2022), the results have been encouraging. However, in the histological context, incorporating these important features directly into the loss function has not been investigated. Furthermore, these studies have not directly incorporated handcrafted attributes into the loss function but rather utilized domain metrics like retaining edge information (Edge Loss) and reducing image distortion (MSE Loss) during learning (Xu et al., 2022). These strategies represent more generic information in medical images, particularly histological images where pathologists explore patterns like cell clustering.

This raises a fundamental question: how would model learning be affected if it is enriched with attributes that are more aligned with the nature of histological images? This question serves as the primary motivation for this study. Additionally, using domain-specific knowledge descriptors, such as multiscale and multidimensional fractals, has helped overcome numerous challenges for this type of scenario (Tenguam et al., 2024; Longo et al., 2023). Therefore, studying these descriptors and their incorporation rules could enable investigations into new combinations between distinct feature groups. These combinations and their rules represent significant frontiers in the machine learning field and enhance diagnostic support systems.

In light of the above, the main contributions of this study are:

- A new method (Enriched-CNN) capable of enriching CNN models through loss functions using multiscale and multidimensional fractal attributes;
- Insights into the primary conditions and enrichment rules based on various neural architectures such as ResNet, InceptionNet, DenseNet, VGGNet and EfficientNet;

- Application of the method in relevant histological dataset representing breast cancer, colorectal cancer and liver tissue, providing information on the best enriched models and the necessary conditions for their study and classification.

2 METHODOLOGY

The proposed approach was divided into two stages to explore different enrichment strategies through handcrafted attributes. Stage 1 aims to extract local and global fractal attributes from each input image using multidimensional techniques like Fractal Dimension (D), Lacunarity (Λ) and Percolation ($PERC$). Stage 2 investigates the influence of primary enrichment rules through fractal attributes on the indicated architectures.

2.1 Stage 1 - Multiscale and Multidimensional Fractal Attributes

Various fractal techniques are available in the literature for image investigations. This study focuses on techniques from a multidimensional and multiscale perspective, such as probabilistic fractal dimension (Ivanovici and Richard, 2011), lacunarity (Ivanovici and Richard, 2009), and fractal percolation (Roberto et al., 2017), as these techniques provide complementary quantifications for colored images. The details are presented in the following subsections.

2.1.1 Probabilistic Approach-Based Fractal Dimension

Fractal dimension (D) was calculated based on the approach described by (Ivanovici and Richard, 2011). Given a colored RGB input image, each image pixel is represented by a 5D vector (x,y,r,g,b), where spatial coordinates (x,y) have color components (r,g,b). Then, a hypercube of side L is initially positioned in the upper-left corner of the image. On each iteration, this hypercube is dislocated from left to right and from top to bottom, covering all pixels in a process known as gliding-box (Ivanovici and Richard, 2011). The hypercube size is increased when the analysis reaches the lower-right corner of the image. For each displacement, an analysis is performed comparing the pixels contained within the hypercube. To do this, the central pixel of the hypercube, $F_c = f(x_c,y_c,r_c,g_c,b_c)$, is fixed, and a comparison is made with the rest of the pixels, including the central pixel itself, using a distance measure Δ. The pixels analyzed in this process are defined as $F_i = f(x_i,y_i,r_i,g_i,b_i)$. In this approach, the analysis is done through the Minkowski distance (Δ_{mink}), calculated as Equation (1):

$$\Delta_{mink} = max(|F_i(k_i) - F_c(k_c)|), \quad k \in r,g,b. \tag{1}$$

In this process, each pixel F_i with a distance Δ less than or equal to the scale size L is labelled as 1, indicating it belongs to the hypercube. Otherwise, it is labelled as 0. By counting these pixels, it is possible to construct the probability matrix $P(m,L)$ (Ivanovici and Richard, 2011), which characterizes the probability of m points being contained within the hypercube of side L.

With the construction of the $P(m,L)$ matrix, we can obtain the partial fractal dimension $N(L)$, which is associated with each hypercube size as defined by the Equation (2):

$$N(L) = \sum_{m=1}^{L^2} \frac{P(m,L)}{m}. \tag{2}$$

To obtain D, after calculating the $N(L)$ value for each L, the angular coefficient of the linear regression defined by $log\ L \times log\ N(L)$ enables us to obtain the probabilistic fractal dimension of the image.

2.1.2 Lacunarity

The multidimensional and multiscale method for calculating the LAC (Λ) of the images under investigation was based on the approach by Ivanovici (Ivanovici and Richard, 2009), using the same probability matrix for the fractal dimension as described in subsection 2.1.1. The metric was based on the first and second-order moments, as defined by Equations (3) and (4). The LAC (Λ) as a function of L, $\Lambda(L)$, was obtained from the distribution measure indicated in Equation (5).

$$\lambda(L) = \sum_{m=1}^{L^2} mP(m,L). \tag{3}$$

$$\lambda^2(L) = \sum_{m=1}^{L^2} m^2 P(m,L). \tag{4}$$

$$\Lambda(L) = \frac{\lambda^2(L) - (\lambda(L))^2}{(\lambda(L))^2}. \tag{5}$$

2.1.3 Multidimensional and Multiscale Percolation

Multidimensional and multiscale percolation ($PERC$) was calculated following the strategy outlined in (Roberto et al., 2017). To do this, percolation theory was applied to analyze pixel paths between one end of the image and the other. The method we used considers a multiscale approach using the gliding-box technique. Initially, hypercubes were defined with $L = 3$. This parameter is increased by two units after crossing the whole image from the top left to the bottom right. The relationship between the number of hypercubes T that have crossed an image with height H and width W, as a function of L, is given by:

$$T(L) = (H - L + 1) \times (W - L + 1), \quad L \leq min(H, W). \tag{6}$$

For each hypercube of size L, we applied a multidimensional approach similar to the one described in (Ivanovici and Richard, 2011) in subsection 2.1.1. Therefore, when the distance Δ has a value less than or equal to L, the pixel P is labelled as -1, indicating it represents a pore. Otherwise, it is labelled as 0, considered the background of the image.

Based on these comparisons, several clusters were formed, as described in (Roberto et al., 2017). From this process, we extracted three functions: the average number of clusters C; the ratio of percolating boxes Q; and the average coverage of the largest cluster M. To calculate the average number of clusters per box $C(L)$, we utilized the number of clusters in a single box (c_i), as a function of scale L, divided by the total number of boxes, as shown in the equation:

$$C(L) = \frac{\sum_{i=1}^{T(L)} c_i}{T(L)}. \tag{7}$$

The ratio of percolating boxes Q was obtained by counting the number of percolating boxes based on scale L. A box q_i is considered percolating if the ratio between the number of pixels labelled as pores (Ω_i) and the total number of pixels within the box (L^2) exceeds a percolation threshold p, defined as 0.59275 (Roberto et al., 2017). The ratio of percolating boxes as a function of L, ($Q(L)$), was obtained by dividing the total number of percolating boxes q_i by the total number of boxes T in a scale L. The expression is given in Equation (8):

$$Q(L) = \frac{\sum_{i=1}^{T(L)} q_i}{T(L)}. \tag{8}$$

Finally, the average coverage ratio of the largest cluster (M) was calculated by identifying the coverage ratio of the largest cluster in each box evaluated at scale L, as shown in Equation (9), where γ_i represents the largest cluster in a box i.

$$M(L) = \frac{\sum_{i=1}^{T(L)} \frac{\gamma_i}{L^2}}{T(L)}. \tag{9}$$

2.1.4 Local and Global Attributes

The fractal descriptors based on the probabilistic fractal dimension, lacunarity and percolation approaches were calculated with scale variations L, using the gliding-box method. In these cases, $L_{max} = 41$ was considered (Roberto et al., 2021), allowing for quantification of 20 different scales. The quantifications used in this study resulted in a set of 100 local attributes ($\mathscr{L}$) for each input image. This enabled us to define characteristic curves as a function of each attribute and the scale L. Therefore, for lacunarity, the curves were formed based on the local values as a function of the sliding hypercube dimension. In the percolation approach, the curves obtained were C, Q and M, representing percolating regions. The LAC and PERC curves were represented by scalar values to form attribute vectors. Based on these curves, the following metrics were extracted to generate global attributes: area under the curve (A), skewness (S), area ratio (Γ) and maximum point (MP) (Roberto et al., 2017).

The total number of attributes was dependent on each category investigated in this study. Table 1 presents the distribution of handcrafted attributes, comprising a structure of 116 attributes ($\mathscr{T}$), with 100 local ($\mathscr{L}$) and 16 global ($\mathscr{G}$), calculated as a function of distance Δ. Global and local attribute sets were analyzed to understand enrichment using different strategies, as described in the next sections.

2.2 Stage 2 - Enrichment Strategies

To apply the enrichment strategies, some models were obtained using transfer learning, which reduces the training time of the CNN model and enables analysis involving datasets with a smaller number of samples. Therefore, in this proposal, pre-trained networks in the ImageNet dataset were used (Roberto et al., 2021; Almaraz-Damian et al., 2020). Examples of architectures that can be investigated are VGGNet, Inception, ResNet, EfficientNet, DenseNet and others. Some of these architectures have already shown relevant results in medical image classification problems in various contexts (Rajinikanth et al., 2020) and also in the classification of histological images (Tenguam et al., 2024; Longo et al., 2023; de Oliveira et al., 2023). These models were treated as the basis for enrichment and therefore named as baseline, and were defined according to recommendations available in the literature (Tenguam et al., 2024; Longo et al., 2023; de Oliveira et al., 2023).

The enrichment process consists of incorporating fractal attributes into the loss function through

Table 1: Identification of the attribute name and the total number of handcrafted attributes.

Attribute	Number of Attributes	Sets
PERC - C, Q, M	60	Local ($\mathscr{L}$)
LAC	20	
Local D	20	
PERC metrics - C, Q, M	12	Global ($\mathscr{G}$)
LAC metrics	4	
Total Number of Attributes	**116**	All ($\mathscr{T} = \mathscr{L} \cup \mathscr{G}$)

different strategies of selection and normalization. Starting from a defined architecture for the baseline model, some layers are released for enrichment. The weights of these layers are updated during training, which, through backpropagation, incorporates the corresponding handcrafted features for each batch of trained samples (batch size = 32). At the end, the baseline model becomes an enriched model. Incorporation of features is contingent on the number of features selected. For instance, if the ReliefF algorithm identifies and chooses only the 3 most pertinent features from the overall feature set, training then proceeds by processing each given batch of samples exclusively with these 3 selected features while also applying a defined incorporation strategy, such as calculating the mean of their values. This approach allows for the progressive enrichment of the model as it encounters each batch of data across its multiple training iterations.

2.2.1 Preparation of Handcrafted Features for Incorporation

In this stage, normalization, selection and attribute incorporation into the loss function ($\mathcal{L}$) processes were established for enrichment. The normalization process was applied to the handcrafted fractal attribute set $\mathscr{T}$. Initially, we tested two normalization types: min-max (defined by Equations (10) and (11)) and z-score (defined by Equation (12)). Normalization is important to ensure that attributes are correctly incorporated, as their value range differs from the loss function where incorporation occurs. Additionally, this strategy strengthens the proposed enrichment approach because it avoids structural modifications or extensions in the architectures, which can lead to dimensional incompatibilities (Cheng et al., 2023). It also minimizes phenomena such as information redundancy (Zheng et al., 2023) and other limitations that often require new modifications or adjustments tailored to each architecture.

$$\tilde{\mathscr{T}_c} = \frac{\mathscr{T}_c - min(\mathscr{T})}{max(\mathscr{T}) - min(\mathscr{T})}\varepsilon, \tag{10}$$

where $\mathscr{T}_c \in \mathscr{T}$ and $\tilde{\mathscr{T}_c}$ represents each element $\mathscr{T}_c$ normalized by min-max according to the scale of values ε provided by the loss function.

$$\varepsilon = (max(\mathcal{L}) - min(\mathcal{L})) + min(\mathcal{L}). \tag{11}$$

$$\tilde{\mathscr{T}_z} = \frac{\mathscr{T}_z - \mathscr{T}_u}{\mathscr{T}_{sd}}, \tag{12}$$

where $\mathscr{T}_z \in \mathscr{T}$ and $\tilde{\mathscr{T}_z}$ represents each element $\mathscr{T}_z$ normalized by z-score, $\mathscr{T}_u$ is the mean of the $\mathscr{T}$ and $\mathscr{T}_{sd}$ is its standard deviation.

The normalized attributes $\tilde{\mathscr{T}}$ were subjected to a selection process using the ReliefF algorithm (Kononenko et al., 1997), resulting in a vector $\tilde{\mathscr{T}}^R$. Several selection tests were performed with the best values presented in the results section. This strategy allowed us to obtain the most relevant descriptors, enhance model interpretability and indicate which rules apply to enrichment through different attributes.

To incorporate the normalized and selected attributes $\tilde{\mathscr{T}}^R$ into the model, some adjustments were made since the loss function only accepts scalars. Some strategies were implemented, such as averaging the values ($\tilde{\mathscr{T}}_u^R$) and applying norms to feature vectors ($\|\tilde{\mathscr{T}}^R\|_p$) with $p = \{1, 2\}$, as tested according to Equation (13):

$$\mathcal{L} = Error(y_i, \hat{y}_i) + \alpha\|\tilde{\mathscr{T}}^R\|_p, \tag{13}$$

where α is a relevance coefficient that accounts for normalized fractal attributes. In this study, we considered $\alpha = 1$.

2.3 Comparisons and Tests

The enrichment process was tested on the ResNet, InceptionNet, DenseNet, EfficientNet and VGGNet architectures across different layers. First, we analyzed which layers in these architectures contained trainable parameters. Then, enrichment was applied in a combined manner across the layers. Each combination resulted in a new enriched CNN model (Enriched-CNN). The goal of this test was to identify the most suitable layer combinations for each architecture, considering commonly used performance evaluation and validation methods in the literature, such as accuracy (Acc), cross-validation, and others (Martinez et al., 2003). The results of the enriched models were compared to those of baseline models to validate the proposed approach.

Additionally, the models enriched using multiscale and multidimensional fractal attributes were compared to the classification results of these attributes using conventional machine learning algorithms representing different categories, such as SVM, Random Forest, KNN and Naive Bayes (Ponti Jr, 2011). Comparisons between the proposed models were conducted using the histological datasets described in the next subsection.

2.4 Application Context - Histological Image Datasets

Histological image datasets, especially those stained with Hematoxylin & Eosin (H&E), are essential for training CNN models to create classification systems. However, these datasets have some limitations, such as limited availability and diversity of samples, making it challenging to train these models for pattern recognition. This requires solutions like enrichment through handcrafted features (Diao et al., 2023). In the context, the effects and conditions imposed to validate the proposed enrichment approach were tested on several datasets, including Colorectal (CR) (Sirinukunwattana et al., 2017), Breast (UCSB) (Gelasca et al., 2008) and Liver tissue (LG) (Zahn et al., 2007). Examples from each dataset group (CR, UCSB and LG) are shown in Figures 1 to 3, respectively. Further information on the datasets is provided in Table 2.

3 RESULTS AND ANALYSIS

The proposed enrichment was implemented in the indicated architectures (ResNet, InceptionNet, DenseNet, VGGNet and EfficientNet) and applied to the following H&E datasets: CR, LG and UCSB. Following subsection 2.2, the pre-trained CNN models underwent normalization and feature selection steps before incorporating handcrafted features. The ReliefF algorithm was used for the selection process with the following feature count (parameter σ): 1, 3, 5, 10 and 20. These parameters were defined based on observations from relevant studies exploring this algorithm in the histological context (Longo et al., 2023). All tests were run three times, and the average of the results was considered when comparing the architectures and datasets. The results were defined through samples for training (70%) and testing (30%), using the holdout cross-validation method. In this study, we considered a fixed learning rate at $1e^{-3}$.

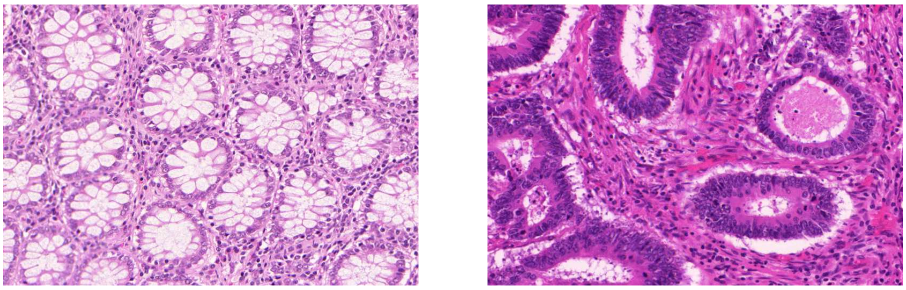
Firstly, we tested the min-max normalization. All attribute incorporation strategies into the loss function were applied ($\tilde{\mathcal{T}}_u^R$, $p = 1$ and $p = 2$). After applying the enrichment method with the defined strategies for each σ value, considering the CR dataset, we observed that except for the ResNet50 architecture, the enriched model achieved higher Acc rates, with the highest value at 100%, obtained using $\sigma = 10$ and $p = 2$, provided by the EfficientNetB2 architecture, as well as $\sigma = 20$ and $p = 2$ using the VGG19 architecture, outperforming the baseline models.

When testing the z-score normalization strategy under the same conditions as the min-max, as shown in Table 3, the results indicated that the highest Acc rate (100%) was achieved with EfficientNetB2 using $\sigma = 10$ and VGG19 using $\sigma = 5$. The most effective incorporation was obtained using the vector norm with $p = 2$. In turn, the highest average Acc value (96.73%) was achieved with $\tilde{\mathcal{T}}_u^R$ incorporation using $\sigma = 10$. Notably, this normalization strategy outperformed the baseline models in various combinations, namely $\sigma = 10$ and $p = 2$; $\sigma = 3$ and $p = 1$; and, finally, $\sigma = 1$ and $\sigma = 10$ using $\tilde{\mathcal{T}}_u^R$. This indicates that the z-score normalization strategy was more efficient in enriching models based on the ResNet50 architecture. In contrast, this only occurred in one situation ($\sigma = 1$ and $p = 1$) using the min-max strategy. Therefore, it is clear that normalization strategies are important in the context of enriching CNN models using handcrafted attributes.

For the LG and UCSB datasets, the enrichment process exhibited different behaviors compared to the CR dataset. The same normalization, selection and feature incorporation steps were performed for the LG and UCSB datasets with notable results highlighted in Tables 4 and 5, respectively.

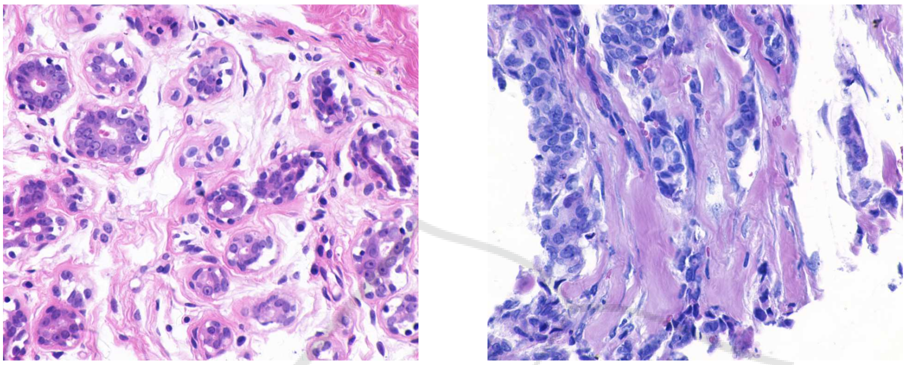
In the LG dataset, some of the most significant results were achieved with the vector norm strategy ($p = 2$) and min-max normalization. This combination achieved the highest average Acc rate among the architectures (88.78% with $\sigma = 10$). In terms of the highest Acc value, different combinations of normalization, feature selection and incorporation strategies achieved the 100% value using the EfficientNetB2 architecture, similar to the CR dataset. Furthermore, the results obtained with this dataset indicated that enrichment was more efficient than with the CR dataset, as the proposed models outperformed baseline models in most combinations.

In the UCSB dataset, in contrast to CR and LG, the



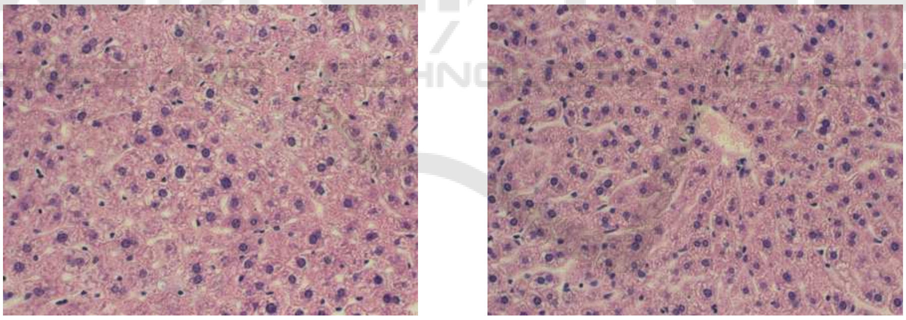
(a) (b)

Figure 1: Examples of H&E colorectal images: (a) benign and (b) malignant.



(a) (b)

Figure 2: Examples of H&E breast images: (a) benign and (b) malignant.



(a) (b)

Figure 3: Examples of H&E liver images: (a) male and (b) female.

winning combination in terms of the highest average Acc rate was vector norm incorporation ($p = 1$) with z-score normalization. This combination achieved the highest average Acc rate among the architectures (78.33% with $\sigma = 3$). The highest Acc value (93.75%) was achieved using $p = 1$, z-score and the VGG19 architecture, as well as $p = 1$, min-max and EfficientNetB2, and also $\tilde{\mathcal{T}}_u^R$, min-max and VGG19.

The results suggest that, except for the ResNet50 architecture, enriched models outperformed baseline models in all strategies for the CR and LG datasets, especially for the UCSB dataset. This is a significant contribution, as histological datasets often have limited samples, particularly the UCSB dataset with only 58 samples, which generally hinders the training of traditional CNN models.

Another noteworthy point is that the enrichment behavior differed depending on the image dataset used. Among the incorporation strategies, vector norm with $p = 2$ for the CR dataset and $p = 1$ for the UCSB dataset stood out. In the LG dataset, no particular strategy stood out. Regarding normalization strategies, their importance was evident in the enrichment process through backpropagation. Additionally,

Table 2: Information on the studied datasets.

Dataset	Description	Classes	Total Images	Dimensions (pixels)
CR	Colorectal tissue	2	165 (91+74)	567 × 430 to 775 × 522
UCSB	Breast tissue	2	58 (32+26)	896 × 768
LG	Liver tissue	2	265 (150+115)	417 × 312

Table 3: Acc rates (%) of enriched models compared to baseline models for various CNN architectures, considering the CR dataset, z-score normalization and feature incorporation using the vector norm with $p = 2$.

CNN Architectures	Enriched-CNN (z-score)					baseline model
	$\sigma = 1$	$\sigma = 3$	$\sigma = 5$	$\sigma = 10$	$\sigma = 20$	
ResNet50	55.10	65.99	66.67	81.63	55.10	**99.32**
InceptionV3	97.96	**98.64**	97.28	97.96	97.28	93.88
DenseNet121	97.96	97.28	97.28	**98.64**	97.28	96.60
EfficientNetB2	97.28	98.64	97.28	**100**	97.96	90.48
VGG19	97.96	98.64	**100**	97.28	99.32	91.16
Average ± SD	89.25 ± 0.19	91.84 ± 0.14	91.70 ± 0.14	**95.10 ± 0.08**	89.39 ± 0.19	94.29 ± 0.04

the min-max strategy yielded better results for the LG dataset, while the z-score strategy was prominent for other datasets. Regarding selection strategies, while no commonalities emerged across datasets, strategies involving few attributes ($\sigma = 1$ or $\sigma = 3$) consistently produced considerable performance in several situations. This behavior could be due to the enrichment process, which performed better in more challenging circumstances, as observed for the UCSB dataset. When $\sigma = 1$, it is implied that a single attribute is capable of promoting the necessary enrichment for the model. This highlights the importance of the selection strategies discussed here, further underscoring the innovation of the proposed enrichment methodology as a method for enhancing classification systems and assisting researchers in this field.

It is worth noting that training the enriched CNN models involved enrichment across various layers to identify the combinations yielding the best results, as discussed in section 2.3. Tests were performed on different layers of the indicated architectures. The best results are presented in Tables 3 to 5 for the three datasets studied. The layers achieving these best performances for each architecture are listed in Table 6.

To better understand the enrichment process in machine learning contexts, comparisons were made among models obtained through classification algorithms, as described in subsection 2.3. The algorithms used were SVM, Random Forest, Naive Bayes and KNN. These algorithms were combined in an ensemble decision process (Longo et al., 2023). Additionally, another combination known as ensemble descriptors was implemented. In this approach, fractal attributes were concatenated with deep-learned attributes extracted from the retrained baseline model based on the last fully connected layer. This layer selection was based on investigations and relevant results achieved in the context of histological images (Tenguam et al., 2024; Longo et al., 2023; de Oliveira et al., 2023). This strategy also acted as a form of CNN model enrichment (Tenguam et al., 2024; de Oliveira et al., 2023; Roberto et al., 2021). The fractal descriptors used here were the same as those used in the previously presented enriched models. This comparison is presented in Table 7 for $\sigma \geq 5$ values, with the best results highlighted in bold.

The results indicate that the proposed enrichment method outperformed, in most situations, traditional training with common classifiers in machine learning processes for all datasets tested. Notably, there were some instances where the enrichment showed less significant performance, particularly for the ResNet50 architecture. This could be attributed to the residual connection mechanism implemented in this architecture.

When $\sigma = 5$ and $\sigma = 10$, the proposed enrichment method showed noticeable improvements, particularly for the CR dataset. While no overall advantages emerged for other datasets, the proposed method outperformed ensemble models in most situations when EfficientNetB2 and VGG19 architectures were considered. This comparison served only to evaluate the feasibility of the proposal, as the implemented enrichment relied on selected fractal attributes, while ensemble models included both deep-learned and fractal attributes. This lack of equality in the comparison hindered a definitive conclusion. However, the proposed enrichment approach can be further explored using other types of attributes crucial in the context of H&E images, potentially leading to new findings and complementing the results achieved here.

Table 4: Acc rates (%) of enriched models compared to baseline models for various CNN architectures, considering the LG dataset, min-max normalization and feature incorporation using the vector norm with $p = 2$.

CNN Architectures	Enriched-CNN (min-max)					baseline model
	$\sigma = 1$	$\sigma = 3$	$\sigma = 5$	$\sigma = 10$	$\sigma = 20$	
ResNet50	67.09	**83.54**	77.64	78.06	80.59	81.44
InceptionV3	56.96	67.51	56.96	76.37	67.93	**83.12**
DenseNet121	81.86	91.98	89.03	**94.09**	91.56	80.17
EfficientNetB2	98.73	**99.58**	98.73	**99.58**	99.16	78.48
VGG19	95.78	92.40	**97.47**	95.78	96.20	89.87
Average ± SD	80.08 ± 0.18	87.00 ± 0.12	83.97 ± 0.17	**88.78** ± **0.11**	87.09 ± 0.13	82.62 ± 0.04

Table 5: Acc rates (%) of enriched models compared to baseline models for various CNN architectures, considering the UCSB dataset, z-score normalization and feature incorporation using the vector norm with $p = 1$.

CNN Architectures	Enriched-CNN (z-score)					baseline model
	$\sigma = 1$	$\sigma = 3$	$\sigma = 5$	$\sigma = 10$	$\sigma = 20$	
ResNet50	68.75	79.00	64.58	68.75	**79.17**	58.33
InceptionV3	60.42	62.50	**68.75**	60.42	66.67	64.58
DenseNet121	52.08	70.00	68.75	**75.00**	52.08	66.67
EfficientNetB2	83.33	89.58	89.58	**91.67**	87.50	75.00
VGG19	85.42	**93.75**	90.00	89.58	81.25	56.25
Average ± SD	70.00 ± 0.14	**78.33** ± **0.14**	76.67 ± 0.13	77.08 ± 0.13	73.33 ± 0.14	64.16 ± 0.07

Table 6: Layers that exhibited the best performance during the enrichment process based on the studied architectures.

CNN Architectures	Enriched layers indicated by name
ResNet50	*conv2_block1_3_bn*, *conv2_block2_3_bn*, *conv2_block3_3_bn*, *conv3_block1_0_bn*, *conv3_block1_3_bn*, *conv3_block2_3_bn*, *conv3_block3_3_bn*, *conv3_block4_3_bn*, *conv4_block1_0_bn*, *conv4_block2_3_bn*, *conv4_block3_3_bn*, *conv4_block6_3_bn*, *conv5_block1_3_bn*, *conv5_block2_3_bn*
InceptionV3	*batch_normalization_5*, *batch_normalization_7*, *batch_normalization_18*, *batch_normalization_21*, *batch_normalization_24*, *batch_normalization_26*, *batch_normalization_33*, *batch_normalization_39*, *batch_normalization_43*, *batch_normalization_48*, *batch_normalization_58*, *batch_normalization_71*, *batch_normalization_78*, *batch_normalization_82*, *batch_normalization_83*, *batch_normalization_87*, *batch_normalization_88*, *batch_normalization_93*
DenseNet121	*bn*
EfficientNetB2	*block1a_project_bn*, *block1b_project_bn*, *block2a_project_bn*, *block2b_project_bn*, *block2c_project_bn*, *block3c_project_bn*, *block4a_project_bn*, *block4c_project_bn*, *block4d_project_bn*, *block5a_project_bn*, *block5c_project_bn*, *block6a_project_bn*, *block6c_project_bn*, *block6e_project_bn*, *block7a_project_bn*
VGG19	*block5_conv4*

3.1 Comparative Overview

To emphasize the importance of the proposed method and its feasibility, the results presented here were compared with those of other established and relevant studies in this research area. The comparisons are presented in Table 8 for each H&E histological dataset.

Most studies presented in the table focus on ensemble descriptor strategies, where the best results are generally achieved through combinations involving deep-learned attributes. When only handcrafted attributes or combinations focused on these attributes are considered, the overall results are not impressive for most methods. The method proposed in this study introduced a new approach for CNN architectures to leverage handcrafted attributes, maximizing their potential during training. Moreover, this integration was subtle and did not significantly alter the architectures. This highlights the advantage of the proposed approach.

Table 7: Acc rates (%) of enriched models compared to those obtained using fractal attributes applied to machine learning algorithms and to those of models enriched through ensemble descriptor strategies using various CNN architectures, considering the CR, LG and UCSB datasets.

CNN Architectures	$\sigma = 5$		$\sigma = 10$		$\sigma = 20$	
	Enriched-CNN	Ensemble	Enriched-CNN	Ensemble	Enriched-CNN	Ensemble
CR						
ResNet50	66.67	**94.55**	81.63	**95.15**	55.10	**97.58**
InceptionV3	**97.28**	95.15	**97.96**	92.73	97.28	**97.58**
DenseNet121	**97.28**	95.76	**98.64**	97.58	97.28	**99.39**
EfficientNetB2	**97.28**	95.15	**100**	99.39	97.96	**99.39**
VGG19	**100**	95.15	**97.28**	96.36	**99.32**	97.58
Fractal Attributes	80.61		84.24		86.06	
LG						
ResNet50	77.64	**95.85**	78.06	**98.49**	80.59	**98.11**
InceptionV3	56.96	**88.68**	76.37	**88.30**	67.93	**95.85**
DenseNet121	89.03	**97.74**	94.09	**99.25**	91.56	**99.62**
EfficientNetB2	**98.73**	93.96	**99.58**	94.34	**99.16**	96.23
VGG19	**97.47**	92.83	**95.78**	94.34	**96.20**	94.34
Fractal Attributes	80.38		90.19		93.96	
UCSB						
ResNet50	64.58	**86.21**	68.75	**91.38**	79.17	**91.38**
InceptionV3	68.75	**81.03**	60.42	**87.93**	66.67	**87.93**
DenseNet121	68.75	**86.21**	75.00	**86.21**	52.08	**91.38**
EfficientNetB2	**89.58**	79.31	**91.67**	84.48	87.50	**89.66**
VGG19	**90.00**	74.14	**89.58**	74.14	81.25	**86.21**
Fractal Attributes	72.41		68.97		74.14	

4 CONCLUSIONS

This study developed a novel method for enriching CNN. In this method, deep-learned features were enriched through backpropagation using relevant fractal techniques commonly applied in H&E image contexts. The results achieved in the studied histological datasets indicate the feasibility of the proposed method, including indications of the combinations that contributed the most. It was important to investigate different selection, normalization and attribute incorporation strategies in the performance analysis of various CNN architectures. The method also highlighted how architectures can be enriched without major structural changes. This opens the door for applying the method to other types of architectures besides CNN. Comparison with related studies suggests that the method achieved notable performance with just a few fractal attributes, while other studies often utilized deep-learned attributes or combinations with more features. Moreover, the method consistently outperformed traditional training with the indicated architectures in most situations.

This opens up new possibilities for future research, such as applying enrichment to architectures beyond CNN, as well as exploring other attributes relevant to the investigated context. Investigating forms of incorporation beyond the loss function could also lead to new interpretations and potentially improve the classification systems developed.

ACKNOWLEDGEMENTS

This research was funded in part by the: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001; National Council for Scientific and Technological Development CNPq (grants #132940/2019-1, #313643/2021-0 and #311404/2021-9); the State of Minas Gerais Research Foundation - FAPEMIG (grant #APQ-00578-18); São Paulo Research Foundation - FAPESP (grant #2022/03020-1).

Table 8: Acc rates (%) of Enriched-CNN models compared to other techniques, considering the CR, LG and UCSB dataset.

Method	Approach	Type of Attributes	Number of attributes	Acc
CR				
Enriched-CNN	VGG19 enriched by attributes $\mathscr{L}$ (PERC and LAC) and $\mathscr{G}$ (S)	Handcrafted	**5**	100
(Longo et al., 2023)	DenseNet121 and EfficientNetB2	Ensemble of deep-learned	10	100
(Roberto et al., 2021)	ResNet50, D, LAC and PERC	Ensemble of deep-learned and handcrafted	300	99.39
(Dabass et al., 2019)	31-layered CNN	Deep-learned	-	96.97
LG				
(Di Ruberto et al., 2016)	Statistical Analysis and Texture Descriptors	Handcrafted	20	100
(Longo et al., 2023)	DenseNet121 and ResNet50	Ensemble of deep-learned	25	100
(Roberto et al., 2021)	ResNet50, D, LAC and PERC	Ensemble of deep-learned and handcrafted	300	99.62
Enriched-CNN	EfficientNetB2 enriched by attributes $\mathscr{L}$ (LAC) and $\mathscr{G}$ (Γ)	Handcrafted	**3**	99.58
(Andrearczyk and Whelan, 2017)	Texture CNN	Deep-learned	-	99.10
UCSB				
(Yu et al., 2019)	CNN, LBP, SURF, GLCM and others	Ensemble of deep-learned and handcrafted	319	96.67
(Longo et al., 2023)	DenseNet121 and EfficientNetB2	Ensemble of deep-learned	25	94.83
Enriched-CNN	VGG19 enriched by attributes $\mathscr{L}$ (PERC) and $\mathscr{G}$ (Γ)	Handcrafted	**3**	93.75
(Kausar et al., 2019)	Color normalization, Haar wavelet decomposition and a 16-layered CNN	Deep-learned	-	91.00
(Roberto et al., 2021)	ResNet50, D, LAC and PERC	Ensemble of deep-learned and handcrafted	300	89.66

REFERENCES

Almaraz-Damian, J.-A., Ponomaryov, V., Sadovnychiy, S., and Castillejos-Fernandez, H. (2020). Melanoma and nevus skin lesion classification using handcraft and deep learning feature fusion via mutual information measures. *Entropy*, 22(4):484.

Andrearczyk, V. and Whelan, P. F. (2017). Deep learning for biomedical texture image analysis. In *Proceedings of the Irish Machine Vision & Image Processing Conference*. Irish Pattern Recognition & Classification Society (IPRCS).

Cheng, C., Xu, T., and Wu, X.-J. (2023). Mufusion: A general unsupervised image fusion network based on memory unit. *Information Fusion*, 92:80–92.

Dabass, M., Vig, R., and Vashisth, S. (2019). Five-grade cancer classification of colon histology images via deep learning. In *Communication and Computing Systems: Proceedings of the 2nd International Conference on Communication and Computing Systems (ICCCS 2018), December 1-2, 2018, Gurgaon, India*, page 18. CRC Press.

de Oliveira, C. I., do Nascimento, M. Z., Roberto, G. F., Tosta, T. A., Martins, A. S., and Neves, L. A. (2023). Hybrid models for classifying histological images: An association of deep features by transfer learning with ensemble classifier. *Multimedia Tools and Applications*, pages 1–24.

Di Ruberto, C., Putzu, L., Arabnia, H., and Quoc-Nam, T. (2016). A feature learning framework for histology images classification. *Emerging trends in applications and infrastructures for computational biology, bioinformatics, and systems biology: systems and applications*, pages 37–48.

Diao, S., Luo, W., Hou, J., Lambo, R., Al-Kuhali, H. A., Zhao, H., Tian, Y., Xie, Y., Zaki, N., and Qin, W. (2023). Deep multi-magnification similarity learning for histopathological image classification. *IEEE Journal of Biomedical and Health Informatics*, 27(3):1535–1545.

Gelasca, E. D., Byun, J., Obara, B., and Manjunath, B. (2008). Evaluation and benchmark for biological image segmentation. In *2008 15th IEEE International Conference on Image Processing*, pages 1816–1819. IEEE.

Hosseini, V. R., Mehrizi, A. A., Gungor, A., and Afrouzi, H. H. (2023). Application of a physics-informed neural network to solve the steady-state bratu equation arising from solid biofuel combustion theory. *Fuel*, 332:125908.

Hu, Z., Zhang, J., and Ge, Y. (2021). Handling vanishing gradient problem using artificial derivative. *IEEE Access*, 9:22371–22377.

Ivanovici, M. and Richard, N. (2009). The lacunarity of colour fractal images. In *Image Processing (ICIP), 2009 16th IEEE International Conference on*, pages 453–456. IEEE.

Ivanovici, M. and Richard, N. (2011). Fractal dimension of color fractal images. *IEEE Transactions on Image Processing*, 20(1):227–235.

Jahan, I., Ahmed, M. F., Ali, M. O., and Jang, Y. M. (2022). Self-gated rectified linear unit for performance improvement of deep neural networks. *ICT Express*.

Kausar, T., Wang, M., Idrees, M., and Lu, Y. (2019). Hwd-cnn: Multi-class recognition in breast histopathology with haar wavelet decomposed image based convolution neural network. *Biocybernetics and Biomedical Engineering*, 39(4):967–982.

Kononenko, I., Šimec, E., and Robnik-Šikonja, M. (1997). Overcoming the myopia of inductive learning algorithms with relieff. *Applied Intelligence*, 7(1):39–55.

Longo, L. H. d. C., Roberto, G. F., Tosta, T. A., de Faria, P. R., Loyola, A. M., Cardoso, S. V., Silva, A. B., do Nascimento, M. Z., and Neves, L. A. (2023). Classification of multiple h&e images via an ensemble computational scheme. *Entropy*, 26(1):34.

Martinez, E. Z., Louzada-Neto, F., and Pereira, B. d. B. (2003). A curva roc para testes diagnósticos. *Cad. saúde colet.,(Rio J.)*, 11(1):7–31.

Montalbo, F. J. P. (2022). Diagnosing gastrointestinal diseases from endoscopy images through a multi-fused cnn with auxiliary layers, alpha dropouts, and a fusion residual block. *Biomedical Signal Processing and Control*, 76:103683.

Nanni, L., Ghidoni, S., Brahnam, S., Liu, S., and Zhang, L. (2020). Ensemble of handcrafted and deep learned features for cervical cell classification. In Nanni, L., Brahnam, S., Brattin, R., Ghidoni, S., and Jain, L., editors, *Deep Learners and Deep Learner Descriptors for Medical Applications. Intelligent Systems Reference Library*, volume 186, pages 117–135. Springer.

Ponti Jr, M. P. (2011). Combining classifiers: from the creation of ensembles to the decision fusion. In *2011 24th SIBGRAPI Conference on Graphics, Patterns, and Images Tutorials*, pages 1–10. IEEE.

Rajinikanth, V., Joseph Raj, A., Thanaraj, K., and Naik, G. (2020). A customized vgg19 network with concatenation of deep and handcrafted features for brain tumor detection. *Appl. Sci.*, 10(10):3429.

Roberto, G. F., Lumini, A., Neves, L. A., and do Nascimento, M. Z. (2021). Fractal neural network: A new ensemble of fractal geometry and convolutional neural networks for the classification of histology images. *Expert Systems with Applications*, 166:114103.

Roberto, G. F., Neves, L. A., Nascimento, M. Z., Tosta, T. A., Longo, L. C., Martins, A. S., and Faria, P. R. (2017). Features based on the percolation theory for quantification of non-hodgkin lymphomas. *Computers in biology and medicine*, 91:135–147.

Sirinukunwattana, K., Pluim, J. P., Chen, H., Qi, X., Heng, P.-A., Guo, Y. B., Wang, L. Y., Matuszewski, B. J., Bruni, E., Sanchez, U., et al. (2017). Gland segmentation in colon histology images: The glas challenge contest. *Medical image analysis*, 35:489–502.

Sukegawa, S., Fujimura, A., Taguchi, A., Yamamoto, N., Kitamura, A., Goto, R., Nakano, K., Takabatake, K., Kawai, H., Nagatsuka, H., et al. (2022). Identification of osteoporosis using ensemble deep learning model with panoramic radiographs and clinical covariates. *Scientific reports*, 12(1):1–10.

Tenguam, J. J., Longo, L. H. d. C., Roberto, G. F., Tosta, T. A., de Faria, P. R., Loyola, A. M., Cardoso, S. V., Silva, A. B., do Nascimento, M. Z., and Neves, L. A. (2024). Ensemble learning-based solutions: An approach for evaluating multiple features in the context of h&e histological images. *Applied Sciences*, 14(3):1084.

Wu, M., Zhu, C., Yang, J., Cheng, S., Yang, X., Gu, S., Xu, S., Wu, Y., Shen, W., Huang, S., et al. (2023). Exploring prognostic indicators in the pathological images of ovarian cancer based on a deep survival network. *Frontiers in Genetics*, 13:1069673.

Xu, L., Zhang, H., Song, L., and Lei, Y. (2022). Bi-mgan: Bidirectional t1-to-t2 mri images prediction using multi-generative multi-adversarial nets. *Biomedical Signal Processing and Control*, 78:103994.

Yu, C., Chen, H., Li, Y., Peng, Y., Li, J., and Yang, F. (2019). Breast cancer classification in pathological images based on hybrid features. *Multimed Tools Appl*, 78:21325–21345.

Zahn, J. M., Poosala, S., Owen, A. B., Ingram, D. K., Lustig, A., Carter, A., Weeraratna, A. T., Taub, D. D., Gorospe, M., Mazan-Mamczarz, K., et al. (2007). Agemap: a gene expression database for aging in mice. *PLoS genetics*, 3(11):e201.

Zhang, F., Lin, S., Xiao, X., Wang, Y., and Zhao, Y. (2024). Global attention network with multiscale feature fusion for infrared small target detection. *Optics & Laser Technology*, 168:110012.

Zheng, H., Hu, Z., Yang, L., Xu, A., Zheng, M., Zhang, C., and Li, K. (2023). Multi-feature collaborative fusion network with deep supervision for sar ship classification. *IEEE Transactions on Geoscience and Remote Sensing*.