

Illumination-targeting Data Augmentation for Monochrome Images

Andrzej Śluzek, Piotr Stachura
0000-0003-4148-2600, 0000-0002-7133-2729
Warsaw University of Life Sciences - SGGW,
Institute of Information Technology,
Nowoursynowska 159, bld.34,
02-776 Warsaw, Poland
Email: {andrzej_sluzek, piotr_stachura}@sggw.edu.pl

Abstract—This paper addresses the challenge of data augmentation for monochrome images, which are still highly relevant in machine vision. Current augmentation methods designed for natural scenes focus on geometric transformation and often fall short in simulating real-world illumination variations. We propose a novel approach that leverages a two-step monochrome image processing technique consisting of image pseudo-colorization followed by a specific decolorization scheme. Our method generates diverse and unpredictable intensity variations, effectively emulating illumination changes and, if only a specific category of color maps are used, preserving the naturalness of the resulting images. To exclude the results which closely replicate images already in the dataset, we rank these maps, identifying a subset that significantly alters illumination characteristics of the originals. A popular SSIM measure is used for that purpose. The proposed technique enhances the illumination diversity of datasets, offering a valuable tool for improving the robustness of AI/ML models in monochrome imagery.

NOTE: All figures are best viewed in high resolution.

I. MOTIVATION AND INTRODUCTION

DATA augmentation is a crucial AI/ML technique that artificially expands training datasets by generating modified samples, [1]. This process enhances robustness of machine vision models to unseen data, reduces overfitting, and improves performance, particularly in scenarios with data scarcity or imbalance, [2]. By leveraging existing data, the need for expensive and time-consuming data collection efforts is minimized.

The data augmentation problem for a specific category of visual data, i.e., monochrome images, is the subject of this paper. These images continue to be a popular and informative source, comparable in relevance to color images.

Commonly, the augmentation of monochrome visual data (e.g., [3], [4]) relies on geometric transformations, encompassing rotation, scaling, cropping, shearing, homography, and random occlusions. Illumination and intensity transformations are less frequently employed, largely due to their limited options, which primarily include adjustments to contrast and brightness, alongside standard image processing techniques like blurring or sharpening. Significantly, current augmentation methods often fall short in their ability to simulate the unpredictable intensity variations caused by real-world illumination changes, such as those arising from diverse weather or

visibility conditions.

Consequently, this paper focuses on augmenting visual data by addressing this specific limitation. Directly visualizing the effects of real-world illumination or visibility changes is generally infeasible due to the extremely complex computational demands of 3D scene reconstruction and illumination modeling based on identified or artificially embedded light sources. Instead, we propose utilizing an image processing technique capable of producing varying and unpredictable outcomes that can effectively emulate diverse illumination changes, even though the underlying physical characteristics of these changes remain unspecified and undetermined.

The proposed approach leverages a recently developed image decolorization algorithm [5], which has been subsequently adapted to enhance the quality of grayscale images in [6], [7]. The algorithm is a specific grayscaling scheme which effectively enhances the visual prominence of original color images by incrementally determining pixel intensities based on the colors and intensities of their already processed neighbors (details in [5]). Example results are shown in Fig. 1.



Fig. 1. Example images from COLOR250 dataset [8] decolorized using the method described in [5].

Subsequently, any monochrome image can be transformed into a color variant using pseudo-colorization, a common technique for visualizing grayscale data from diverse, primarily non-visual domains, e.g., [9], [10], and then decolorized using the aforementioned approach. We have observed that this two-step process can significantly enhance image perceptibility compared to the original grayscale image, given the application of a suitable *color map* in the initial step. Optimal color maps for this purpose are identified in [6], [7], both for general monochrome images (targeting primarily those from non-visual domains) [7]), and for images depicting natural

scenes, [6]. However, we also observed that for images of natural scenes, many color maps produce monochrome results that, regardless of improved perceptibility, can resemble the same scenes under different illumination conditions, as shown in Fig. 2.



Fig. 2. The original image (left) and its variants, generated using diversified color maps in the first step of our method.

This finding forms the core principle of this paper, as such images are incorporated into the augmented dataset.

First, Section II categorizes color maps employed in pseudo-colorization and identifies the sub-category suitable for the proposed data augmentation scheme. Within this context, the key features of the adopted decolorization scheme are also briefly highlighted. In Section III, those color maps are ranked so that only maps which statistically generate the most diverse (and different from the dataset originals) images are retained. In this way, we exclude the augmented images that basically replicate visual information already provided in the training dataset.

Finally, Section IV presents examples of this augmentation and discusses further aspects of the method, including its limitations.

II. PRINCIPLES OF THE METHOD

A. Remarks on Image Decolorization (more in [5])

The decolorization scheme presented in [5] assumes, but does not directly apply, linear *rgb-to-gray* mappings defined by:

$$I = k_R R + k_G G + k_B B, \quad (1)$$

where the coefficients $[k_R, k_G, k_B]$ are optimized either globally (e.g., [11], [12]) or locally (e.g., [13], [14]) to enhance the visual expressiveness of the resulting monochrome images.

Instead, we observe that a pixel with $[R, G, B]$ values can only be assigned an intensity I within the constrained range:

$$\min(R, G, B) \leq I \leq \max(R, G, B), \quad (2)$$

regardless of the specific values of $[k_R, k_G, k_B]$ in Eq.1.

Eq. 2 allows us to identify pixels with the narrowest range of feasible grayscale intensities. For these pixels, their grayscale values are either randomly selected from this restricted range or assigned deterministically when $\min(R, G, B) = \max(R, G, B)$.

These pixels serve as the starting points (the *initial list*) for the decolorization process, which employs a randomized variant of a *flood-fill* algorithm. Subsequently, pixels with already decolorized neighbors are assigned grayscale intensities (constrained by Eq. 2) in proportion to their color differences from those neighbors. This iterative process continues until all pixels are processed, completing the decolorization step.

Furthermore, as discussed below, Eq. 2 indirectly determines the color maps that can be effectively used in the first step (image pseudo-colorization) of the proposed image augmentation.

B. Color Maps

Various color maps are designed for diverse purposes (e.g., [15], [16]); however, we specifically focus on *sequential color maps*, where the luminance Y of assigned colors changes monotonically with intensity. That is, if $I_1 < I_2$ then $Y(RGB(I_1)) < Y(RGB(I_2))$. This constraint ensures that the pseudo-colored image preserves the perceptual order of luminance, which is important for maintaining the natural appearance of the augmented images.

To further preserve perceptual naturalness, we impose another constraint on the color maps, derived from Eq. 2. Specifically, we only use those sequential color maps that satisfy the condition:

$$\max[R(0), G(0), B(0)] < 0.5 < \min[R(255), G(255), B(255)]. \quad (3)$$

This condition implies that after pseudo-coloring and subsequent decolorization, the darkest pixels in the original image always remain darker than the brightest pixels. Combined with the fact that the average intensity in Eq. 2 generally increases for sequential color maps, the original hierarchy of intensities is statistically maintained. We will jointly refer to sequential maps satisfying Eq. 3 as SEQ3 maps.

At least 42 SEQ3 maps exist, as identified from the publicly available map catalog [17] and other sources.

Figure 3 provides examples of those categories of color maps.

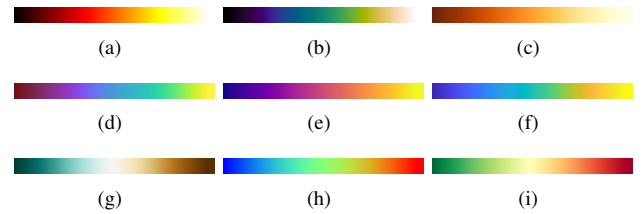


Fig. 3. Examples of: (a,b,c) sequential color maps satisfying Eq. 3 (SEQ3 maps); (d,e,f) non-SEQ3 sequential color maps; and (g,h,i) non-sequential color maps.

To further illustrate the significance of using only SEQ3 color maps, examples are provided in Fig. 4. The results obtained with SEQ3 maps effectively emulate alternative illumination conditions, whereas the image generated using a non-sequential map appears too unrealistic (despite its high visual clarity) to be considered an augmented variant of the original.

III. RESTRICTING THE NUMBER OF COLOR MAPS

Although our proposed method allows any SEQ3 map to generate realistically-looking alternatives of monochrome images, some maps yield outputs nearly indistinguishable from the originals. Such near-identical images are unsuitable for data augmentation as they essentially replicate existing

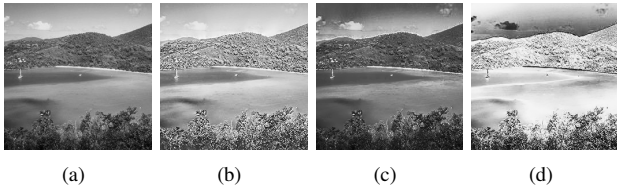


Fig. 4. Original image (a) compared with two variants using SEQ3 maps (b,c) and a variant using a non-sequential color map (d), which shows an unacceptable loss of naturalness.

dataset items in terms of overall visual impression. Instead, we should prioritize maps that introduce substantial variations in illumination characteristics. Illustrative examples of both scenarios are provided in Fig. 5.

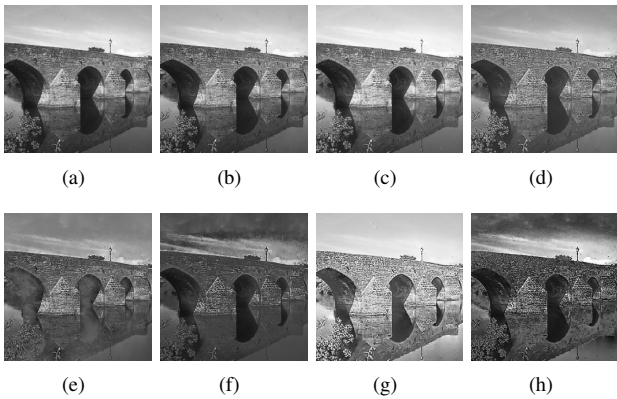


Fig. 5. (a) Original image; (b-d) variants closely resembling the original (*bone*, *greys*, *ice* SEQ3 maps); and (e-h) variants with significantly different illumination characteristics (*haline*, *inferno*, *OrRd*, *rainforest* SEQ3 maps).

To assess the statistical suitability of various SEQ3 maps for illumination-targeting data augmentation in monochrome imagery, we utilized a large number (N) of grayscale images randomly extracted from the widely used SUN dataset [18]. This dataset was selected to facilitate future research comparing the classification performance of diverse DL architectures trained on either color images or their grayscale versions.

Each test image was processed by our method using all (SEQ3(1) through SEQ3(42)) maps. Subsequently, we evaluated the illumination disparity between each generated image and its original. The most dissimilar results would be potential candidates for inclusion in the augmented dataset.

While visual inspection could assess this dissimilarity, its burden on human evaluators quickly becomes impractical even with a small number of test images. Therefore, after considering several common image similarity metrics, we opted for the structural similarity index measure (SSIM) for automatic evaluation.

The SSIM is defined by a weighted combination of three measures that broadly represent statistical similarities between the intensity, contrast, and structure of two images, X and Y :

$$SSIM(X, Y) = i(X, Y)^\alpha \times c(X, Y)^\beta \times s(X, Y)^\gamma \quad (4)$$

where $i(X, Y) = \frac{2\mu_X\mu_Y+c_1}{\mu_X^2+\mu_Y^2+c_1}$, $c(X, Y) = \frac{2\sigma_X\sigma_Y+c_2}{\sigma_X^2+\sigma_Y^2+c_2}$ and $s(X, Y) = \frac{\sigma_{XY}+c_3}{\sigma_X\sigma_Y+c_3}$, see [19].

The structural similarity index measure (SSIM) ranges from -1 to 1 , with a maximum value of 1 indicating identical images.

While lower SSIM values suggest various distortions between an original image and its processed version, in our specific scenario, the image structure is perfectly preserved. Therefore, intensity variations are the primary factor contributing to the SSIM-measured dissimilarity, making it a suitable metric for our focus on illumination changes.

Consequently, for each test images $X(n)$ (where $n = 1, \dots, N$), we calculate 42 SSIM values, denoted as their 42 $SSIM(X(n), X_i(n))$ (where $i=1, \dots, 42$). These values represent the similarity between the original image and its variants $X_i(n)$ generated using each of the 42 SEQ3 maps.

Subsequently, we estimate the norms of the SSIM value vectors across all test images.

$$\text{norm} \{[SSIM(X(1), X_i(1)), \dots, SSIM(X(N), X_i(N))]\} \quad (5)$$

The SEQ3 maps are ranked based on the increasing values of the metric defined in Eq. 5. Consequently, maps exhibiting the lowest vector norms are considered strong candidates for data augmentation, as they statistically produce images with significant dissimilarity from the originals. Conversely, maps with the highest vector norms tend to generate near-identical images, rendering them unsuitable for this purpose.

We explored various vector norms, primarily L^k norms with different values of k , both with and without outlier removal. Remarkably, the resulting rankings of SEQ3 maps remained nearly identical across these different norms, especially within the top (recommended maps) and bottom (least recommended maps) tiers. Specifically, the **top 5** and **bottom 5** map sets consistently contained the same members, with only minor variations in their internal order, as detailed in Table I, irrespective of the chosen norm.

TABLE I
RANKING OF SEQ3 MAPS, INDICATING THE TOP AND BOTTOM TIERS.

rank	1	2	3	4	5	...	38	39	40	41	42
map	<i>rainforest</i>	<i>hotM</i>	<i>hot</i>	<i>Y(OrBr)(OrRd)</i>	<i>OrRd(Y(OrBr))</i>		<i>ocean(ice)</i>	<i>ice(ocean)</i>	<i>copper</i>	<i>greys</i>	<i>bone</i>

Additionally, we have found that the maximum subset of color maps which are always on top of this ranking (although sometimes in a slightly varying order) irrespective of the norm applied in Eq. 5, contains 14 SEQ3 maps.

We further assessed, using SSIM, the cross-dissimilarity of the outputs generated by these color maps. Generally, the results exhibited substantial differences, with the notable exception of *hotM* and *hot*. These two maps, being nearly identical, understandably produced very similar images, suggesting that one should be excluded.

Consequently, for illumination-targeting data augmentation in monochrome natural scene imagery, we recommend employing the following color maps in the first step of our proposed monochrome-to-monochrome (via pseudo-coloring) image alternations:

(1) *rainforest*, (2) *hot*, (3) *YlOrBr*, (4) *OrRd*, (5) *YlOrRd*, (6) *oranges*, (7) *nuclear*, (8) *RdPu*, (9) *reds*, (10) *PuRd*, (11) *speed*, (12) *warm* and (13) *inferno*.

Detailed information on these maps can be found in [17] and other resources.

Specifically, Figures 3(a) 3(b) and 3(c) illustrate the *hot*, *rainforest* and *YlOrBr*, respectively.

Our limited-scale visual validation strongly supports the suitability of these automatically identified color maps.

IV. SUMMARY

A. Examples

Figures 6 to 9 illustrate the augmentation results achieved with the recommended color maps. These figures display original monochrome images of natural scenes and the corresponding outputs produced by a random selection of seven out of the thirteen recommended maps. While some image

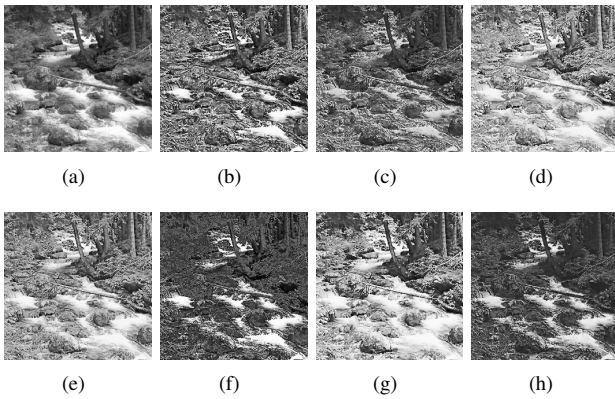


Fig. 6. (a) Original image; augmented variants obtained by using: (b) *hot*, (c) *inferno*, (d) *OrRd*, (e) *PuRd*, (f) *rainforest*, (g) *RdPu* and (h) *speed* SEQ3 maps.

fragments may show resemblance, their illumination characteristics diverge significantly in other parts.

B. Concluding Remarks

Overall, the SEQ3 color maps generate randomized and often unpredictable intensity changes while preserving the natural appearance of the original images. This means our proposed augmentation method effectively produces scene variations that could occur under various unspecified illumination conditions, an effect not achievable with traditional image processing techniques.

Nevertheless, it's important to highlight that our approach addresses only a limited range of potential illumination variations. Specifically, within the proposed model, we cannot simulate the effects of relocating light sources or changing their numbers. Despite this limitation, we believe our

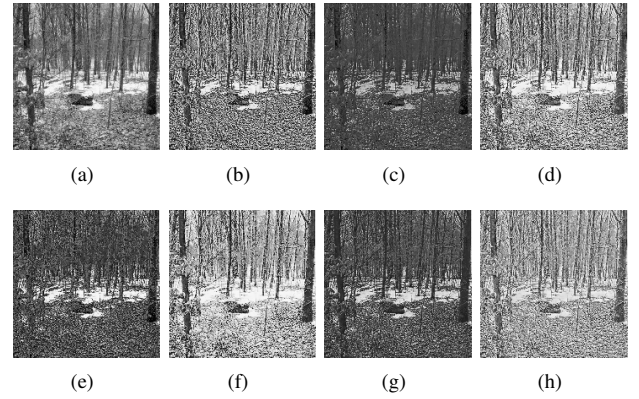


Fig. 7. (a) Original image; augmented variants obtained by using: (b) *hot*, (c) *nuclear*, (d) *oranges*, (e) *rainforest*, (f) *RdPu*, (g) *speed* and (h) *YlOrRd* SEQ3 maps.

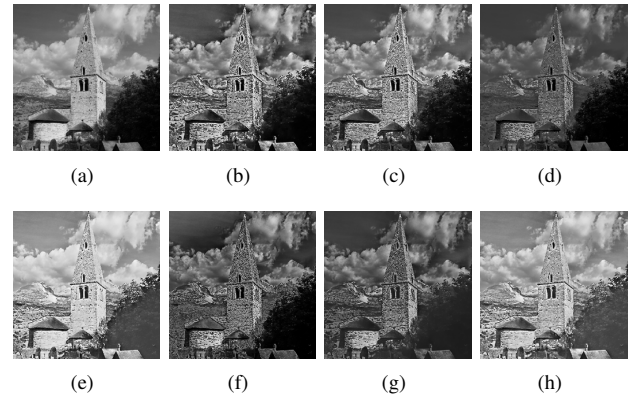


Fig. 8. (a) Original image; augmented variants obtained by using: (b) *hot*, (c) *inferno*, (d) *nuclear*, (e) *OrRd*, (f) *rainforest*, (g) *speed* and (h) *YlOrRd* SEQ3 maps.

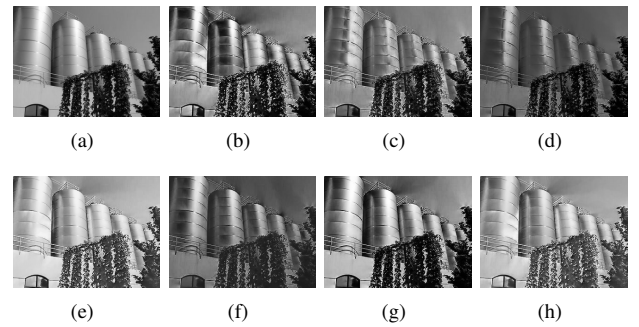


Fig. 9. (a) Original image; augmented variants obtained by using: (b) *hot*, (c) *inferno*, (d) *nuclear*, (e) *OrRd*, (f) *speed*, (g) *warm* and (h) *YlOrRd* SEQ3 maps.

method offers a valuable tool for augmenting training data in monochrome imagery, introducing items with unseen and diversified illumination characteristics. Notably, when combined with geometry-based augmentation, this method can yield sufficiently large and diverse training datasets, even when the availability of images from a particular domain is limited.

Finally, it should be noted that restricting the augmenta-

tion to only SEQ3 color maps is not a strict requirement. Generally, non-SEQ3 sequential color maps or non-sequential maps produce unacceptably unnatural outputs, as illustrated in Fig. 3(d). However, if extremely unusual and unpredictable visibility conditions are anticipated in a specific scenario, one might consider incorporating non-SEQ3 or even non-sequential maps into the data augmentation process. Fig. 10 provides an illustrative example of such a case.

Additionally, we have experimentally verified that our approach offers a superior level of intensity diversification compared to standard image processing techniques. We evaluated over ten common, *non-task-specific* image enhancement algorithms. Of these, only two - *image sharpening* and the *LoG* filter - produced statistically comparable SSIM dissimilarity results. However, even these methods were noticeably inferior to ours. It is important to note that image sharpening and the *LoG* filter primarily enhance edges. This means the visual differences observed with these techniques largely stem from structural modifications, not from the genuine intensity diversification our proposed method achieves.

In the future, we plan to extend our proposed data augmentation scheme to natural color images. Preliminary experimental results for this extension are shown in Figure 11.

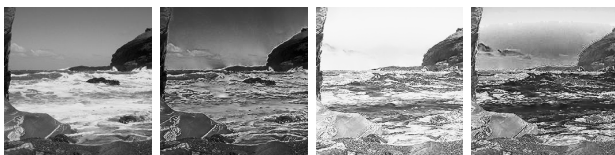


Fig. 10. Original image (left) and three augmented variants using non-SEQ3 color maps.



Fig. 11. (a) Original color image; augmented results obtained via SEQ3 color maps: (b) *inferno*, (c) *RdPu*, and (d) *rainforest*.

REFERENCES

- [1] A. Mumuni and F. Mumuni, "Data augmentation: A comprehensive survey of modern approaches," *Array*, vol. 16, p. 100258, 2022. doi: <https://doi.org/10.1016/j.array.2022.100258>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2590005622000911>
- [2] Y. Zhang, A. Yilmaz, M. Popa, and C. Brewster, "Analysis of the impact of data augmentation on the performance of deep learning models in multispectral food authenticity identification," in *Proc. 18th Conference on Computer Science and Intelligence Systems*, ser. Annals of Computer Science and Information Systems, vol. 35. IEEE, 2023. doi: 10.15439/2023F3643 p. 823–832.
- [3] A. Mikolajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," in *2018 International Interdisciplinary PhD Workshop (IIPHDW)*, 2018. doi: 10.1109/IIPHDW.2018.8388338
- [4] L. Nanni, M. Paci, S. Brahnam, and A. Lumini, "Comparison of different image data augmentation approaches," *Journal of Imaging*, vol. 7, no. 12, p. 222–239, 2014. doi: 10.3390/jimaging7120254
- [5] A. Śluzek, "Incremental image decolorization with randomizing factors," in *32nd European Signal Processing Conference (EUSIPCO)*, 2024. doi: 10.23919/EUSIPCO63174.2024.10715444 pp. 591–595.
- [6] —, "An unorthodox technique for enhancing monochrome images of natural scenes," in *Proc. 6th Polish Conference on Artificial Intelligence, Katowice, Poland*. Springer, 2025, in print.
- [7] —, "A new technique for enhanced monochrome visualization of non-visual data," in *Computational Science – ICCS 2025 Workshops*, M. Paszynski, A. S. Barnard, and Y. J. Zhang, Eds. Springer Nature Switzerland, 2025, pp. 51–65.
- [8] C. Lu, L. Xu, and J. Jia, "Contrast preserving decolorization with perception-based quality metrics," *International Journal of Computer Vision*, vol. 110, p. 222–239, 2014. doi: 10.1007/s11263-014-0732-6
- [9] N. S. Aghdam, M. M. Amin, M. E. Tavakol, and E. Ng, "Designing and comparing different color map algorithms for pseudo-coloring breast thermograms," *Journal of Medical Imaging and Health Informatics*, vol. 3, no. 4, pp. 487–493, 2013. doi: 10.1166/jmhi.2013.1191
- [10] A. Rahimian, M. Etchadtavakol, M. Moslehi, and E. Ng, "Comparing different algorithms for the pseudo-coloring of myocardial perfusion single-photon emission computed tomography images," *Journal of Imaging*, vol. 8, no. 12, p. 331, 12 2022. doi: 10.3390/jimaging8120331
- [11] Q. Liu, J. Xiong, L. Zhu, M. Zhang, and Y. Wang, "Extended RGB2Gray conversion model for efficient contrast preserving decolorization," *Multimedia Tools and Applications*, vol. 76, p. 14055–14074, 2017. doi: 10.1007/s11042-016-3748-9
- [12] T. Wu, C. Eising, M. Glavin, and E. Jones, "An efficient and effective image decolorization algorithm based on cumulative distribution function," *Journal of Imaging*, vol. 10, no. 3, 2024. doi: 10.3390/jimaging10030051
- [13] T. Wu and A. Toet, "Color-to-grayscale conversion through weighted multiresolution channel fusion," *Journal of Electronic Imaging*, vol. 23, no. 4, p. 043004, 2014. doi: 10.1117/1.JEI.23.4.043004
- [14] J. Yu, F. Li, and X. Lv, "Contrast preserving decolorization based on the weighted normalized l1 norm," *Multimedia Tools and Applications*, vol. 80, p. 31753–31782, 2021. doi: 10.1007/s11042-021-11172-9
- [15] M. Khan, Y. Gotoh, and N. Nida, "Medical image colorization for better visualization and segmentation," in *Medical Image Understanding and Analysis*, M. Valdés Hernández and V. González-Castro, Eds. Cham: Springer International Publishing, 2017. doi: 10.1007/978-3-319-60964-5_50 pp. 571–580.
- [16] K. Moreland, "Why we use bad color maps and what you can do about it," in *Proc. IS&T Int'l. Symp. on Electronic Imaging: Human Vision and Electronic Imaging*, vol. 28, 2016. doi: 10.2352/ISSN.2470-1173.2016.16.HVEI-133
- [17] "Colormap catalog," accessed on Jan. 5, 2025. [Online]. Available: <https://cmmap-docs.readthedocs.io/en/latest/catalog/>
- [18] J. Xiao, J. Hays, K. Ehinger, A. Oliva, and A. Torralba, "SUN database: Large-scale scene recognition from abbey to zoo," in *2010 IEEE Conference CVPR*, 2010. doi: 10.1109/CVPR.2010.5539970 pp. 3485–3492.
- [19] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004. doi: 10.1109/TIP.2003.819861