FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 10 No. 2, January, 2026, pp 364 – 374
DOI: https://doi.org/10.33003/fjs-2026-1002-4338

A MULTI-HYBRID FRAMEWORK FOR IMAGE DATA AUGMENTATION IN A RANDOMIZED SETTING

***[1]Peter Jacob Adoyi, [2]Samuel Oluwatosin Hassan, [1]Samera Otor and [1]Adekunle Adedotun Adeyelu**

[1]Department of Mathematics/Computer Science, Rev. Fr. Moses Orshio Adasu (former Benue State) University, Makurdi, Nigeria.
[2] Department of Computer Sciences, Olabisi Onabanjo University, Ago-Iwoye.

*Corresponding authors' email: pitadoyi5@gmail.com

ABSTRACT

Image Data Augmentation (IDA) is fundamental for improving model robustness and reducing overfitting; however, many existing approaches are tightly integrated in specific deep learning models, limiting their applicability primarily to classification tasks. This study proposes a model-agnostic, stand-alone multi-hybrid IDA framework that operates independently of any machine learning (ML) model. The framework is implemented as a Python-based system integrating five common transformation techniques: rotation, translation, zooming, flipping, and color-space manipulation. The proposed Randomized Framework was benchmarked against three established augmentation pipelines with a Sequential mode serving as an internal baseline. Performance evaluation employed dataset diversity, class preservation, execution time, and classifier validation using K-Nearest Neighbour (KNN) and Support Vector Machine (SVM) models on features extracted via ResNet-50. Results demonstrate that the Randomized Framework achieved the highest diversity score (0.3767) with optimal execution time (5.20 s), alongside superior generalization evidenced by the lowest overfitting gap (0.03) and strong classification performance (AUC = 0.9596 for KNN). Classification accuracy improved from 0.91–0.93 to 0.95–1.00 after augmentation. The study concludes that randomized multi-hybrid augmentation offers an effective balance between diversity and efficiency and, due to its specific machine learning model- independent design, it is well suited for diverse computer vision applications. Future work will investigate additional transformations, deep learning–based augmentation pipelines, and scalable batch-level optimization.

Keywords: Image Augmentation, Hybrid Transformations, Randomized Techniques, Machine Learning Bencharkng, Model Independent Framework, Data Diversity/Similarity, Overfitting, KNN, SVM, ResNet50

INTRODUCTION

Deep learning (DL) algorithms, particularly for tasks such as classification and predictive modeling, essentially depend on vast amounts of high-quality, labeled data to achieve high generalization performance (Geron, 2019; Singh & Mahmood, 2021). However, real-world limitations such as noisy acquisition environments, privacy restrictions, limited collection resources, and imbalance across classes make it difficult to obtain large enough datasets for effective training (Mumuni & Mumuni, 2022; Shorten & Khoshgoftaar, 2019). Image Data Augmentation (IDA) has therefore become a core strategy for artificially expanding datasets while preserving class information. Conventional forms of IDA, including rotation, flipping, zooming, translation, and brightness adjustment, have been widely applied for tasks varying from facial recognition to medical image classification (Chlap et al., 2021; Du Toit et al., 2019).

Historically, these methods began with basic techniques prior to 2010 before being accelerated by the accomplishment of AlexNet, where Krizhevsky et al. (2012) employed translation and intensity adjustment. Subsequent strategies became core components of high-performing architectures, including DenseNet (Huang et al., 2017) and techniques associated with MobileNets (Howard et al., 2017). More recent studies, such as the implementation of phase randomization for temporal sequences (Mitsuzumi et al., 2019) and style augmentation (Jackson et al., 2019), demonstrate the importance of randomization; however, these remain either domain-specific, model-dependent, or limited in augmentation scope.

Despite the success of these historical approaches, many existing frameworks, including those embedded in YOLO (Redmon et al., 2016) and ResNet (He et al., 2016), are model-dependent and lack the flexibility for general-purpose use across various computer vision tasks. This limitation follows the observations of Mumuni and Mumuni (2022), who argue that standalone, model-agnostic frameworks are essential for increasing the portability and robustness of augmented data across different learning architectures. The fundamental constraint is the over-reliance on fixed, sequential pipelines, which apply transformations in a fixed order and lack robust randomization in transformation order or parameter tuning (Taylor & Nitschke, 2018; Ly et al., 2023; Pal, 2022). This lack of randomization produces new data entities that are less diverse and follow predictable patterns, a problem which leads to models that overfit to slight differences in the training data, ultimately causing a loss of generalizability (Shorten & Khoshgoftaar, 2019).

To address these gaps, this study tackles the critical need to develop a unique, stand-alone framework implemented as a modular, architecture-agnostic augmentation engine. This design is capable of generating new image samples through the random selection and permutation of diverse transformation functions. The central aim is to implement this novel scheme by combining various techniques (rotation, translation, color space transformation, zooming, and flipping) into a randomized, multi-hybrid framework. The resulting system is built around three core contributions:

i. A Hybrid Augmentation Engine: Supports 2–5 transformation techniques applied in randomized sequences, enabling thousands of unique variants from a single input image.
ii. A Randomized Parameter Selection Mechanism: Introduces stochasticity into both the choice of

transformation and its numeric configuration, increasing diversity without compromising semantic accuracy.

iii. A Comprehensive Evaluation Suite: Validates the framework using core metrics (diversity, class preservation, and time complexity) and machine learning evaluations (Accuracy, Recall, F1 score, ROC curves, and A/B testing) using KNN and SVM over ResNet-50 features.

Ultimately, this research seeks to provide a highly adaptive augmentation environment that produces more diverse images while improving the generalization performance of downstream models. The remainder of this paper is structured into five main sections, moving from a review of related work in Section 2 to the materials and methods used to bridge the research gap in Section 3, followed by the results, their implications, and comparative benchmarking in Section 4, and concluding with key findings and recommendations for future work in Section 5.

Related Work

The development of high-performing machine learning (ML) models, particularly in computer vision, is fundamentally dependent on the quality and volume of training data, as large, diverse, and representative datasets are required to minimize the risk of model overfitting (Shorten & Khoshgoftaar, 2019; Mumuni & Mumuni, 2022). Machine learning techniques have increasingly been adopted across diverse scientific and engineering domains due to their ability to model complex patterns and improve predictive accuracy. Recent studies have demonstrated the effectiveness of data-driven approaches in classification and prediction tasks across agriculture, health, and educational analytics. For instance, Shuaibu et al. (2024) have shown the successful application of machine learning algorithms in crop yield prediction and performance evaluation tasks. These developments highlight the growing relevance of robust data preprocessing and augmentation strategies capable of enhancing model generalization across varying application domains. However, most existing augmentation pipelines remain model-dependent, limiting their adaptability across different learning architectures.

To mitigate real-world constraints such as resource limitations and privacy restrictions, Image Data Augmentation (IDA) serves as the primary technique for artificially generating new transformed versions of images (Chlap et al., 2021; Xu et al., 2023, Dilmegani, 2023). IDA is essential for improving classification accuracy, especially on small datasets (Moreno-Barea et al., 2020), by treating a digital image as a 2D array of pixel values that can be systematically altered. This study focuses on two main transformation categories: geometric and color space.

The first category, Geometric Transformation Methods, generates new samples by changing the position or scale of an existing image. This includes Rotation, which turns the image at angles "θ" ∈ (1°,359°) to help models recognize objects from different viewpoints (Krizhevsky et al., 2012); Translation, which shifts pixels horizontally "t" _"x" or vertically "t" _"y" , generating new samples that handle positional bias (Shorten & Khoshgoftaar, 2019); Flipping, which reflects the image along a specified axis to build robustness against mirrored objects (Shorten & Khoshgoftaar, 2019; Xu et al., 2023); and Zooming, which scales the image by a factor α, creating new versions that are larger or smaller.

The second category, Colour Space Transformations, manipulates the photometric properties such as brightness, contrast, or saturation of the original RGB (Red, Green, Blue) channels (Taylor & Nitschke, 2018).

While these traditional methods formed the basis of early augmentation and were frequently embedded in architectures like DenseNet (Huang et al., 2017) and ResNet (He et al., 2016), a significant limitation is their reliance on fixed, sequential pipelines (Redmon et al., 2016). Applying a pre-determined order of transformations restricts the generation of diverse samples and fails to capture the domain variability necessary for robust generalization (Ly et al., 2023). Recent attempts to overcome this predictability have led to two distinct paradigms. The first involves Aggressive Techniques (like Cutout or Mixup) which, while improving generalization, do not preserve pixel integrity and are unsuitable for class-preserving tasks like medical imaging. The second involves Automated Augmentation (Han et al., 2023), which utilizes reinforcement learning but remains computationally expensive and model-dependent. A summary of these foundational and integrated frameworks is presented in Table 1. These prior works collectively demonstrate a persistent research gap. Although studies by Jackson et al. (2019) and Mitsuzumi et al. (2023) introduced important stochasticity, they do not implement the crucial combination of random permutation of multiple hybrid transformation functions and the stochastic tuning of their parameters in a stand-alone framework. Consequently, there remains a need for a generalized, flexible, and maximally diverse IDA system that is independent of specific model architectures.

Table 1: Image Augmentation Techniques Implemented in some Literature

S/No	Source	Image augmentation techniques
1	du Toit et al. (2019)	Colour space
2	He et al. (2016)	Cropping, Flipping
3	Howard et al. (2017)	Cropping, Elastic distortion
4	Huang et al. (2017)	Flipping, Cropping, Translation
5	Jackson et al. (2019)	Colour space (style-based randomization, texture/contrast changes)
6	Krizhevsky et al. (2012)	Translation, Flipping, Intensity changing
7	Ly et al. (2023)	Collage pasting, PixMix
8	Redmon et al. (2016)	Scaling, Translation, Colour space
9	Zhong et al. (2020)	Random erasing
10	Mitsuzumi et al. (2023)	Phase randomization (temporal augmentation for action recognition)

These prior works collectively demonstrate a persistent research gap. Although studies by Jackson et al. (2019) and Mitsuzumi et al. (2023) introduced important stochasticity, they do not implement the crucial combination of random permutation of multiple hybrid transformation functions and the stochastic tuning of their parameters in a stand-alone framework. Consequently, there remains a need for a generalized, flexible, and maximally diverse IDA system that is independent of specific model architectures.

In order to address these gaps, the present study develops a novel, model-independent multi-hybrid augmentation framework. This approach allows for the random permutation of five transformation techniques and stochastic parameter tuning, distinguishing it from other randomization-based literature as summarized in Table 2.

Table 2: Comparison of Randomization-based Works vs. Present Study

Paper	What they do (randomization)	What they evaluate	What they don't do (gaps relative to present study)
Mitsuzumi et al. (2023)	Randomizes phase component in temporal sequences to remove subject-specific bias.	Action recognition, domain adaptation,	Not for static images; lacks geometric transforms; parameter permutation; model-independent; diversity metric.
Jackson et al. (2019)	Random style transfer (texture/color/contrast) and embedding sampling.	Image classification and domain shift across datasets.	Focused on style over geometric transforms; lacks user-defined ranges; hybrid random permutations; and cosine-based similarity or diversity metrics
Present study	Random permutation of 5 augmentation technique	Cosine based similarity/distance and time validated with KNN & SVM.	Model-independent, GUI-based, supports user-defined parameter ranges and hybridization up to 5 transforms.

Furthermore, to validate the effectiveness of the proposed framework, the extracted features from the augmented images are used to train standard Machine Learning (ML) classifiers (Reinaldo & Dwiasnati, 2023). As an application of Artificial Intelligence (AI), ML focuses on developing algorithms that make predictions based on available data (Alpaydin, 2010; Udousoro, 2020).

MATERIALS AND METHODS

This section details the design, implementation, and rigorous evaluation of the model-independent multi-hybrid augmentation framework. The methodology covers the system architecture, mathematical modeling, and the setup for benchmarking and classifier validation.

Framework Design and Novel Contributions

The proposed augmentation framework is structured as a modular, classifier-independent image transformation engine. The overall system architecture is illustrated in Figure 1. The system begins with standardized input images, which are forwarded to a multi-technique selection stage where one or more augmentation methods are dynamically chosen from a predefined pool. Rather than enforcing a fixed pipeline, the framework allows flexible selection to support both single and hybrid transformation strategies. The selected techniques are regulated by a parameter range controller, which constrains all transformations within user-defined boundaries to ensure realistic and semantically consistent modifications. This controlled stochastic mechanism introduces variability while preserving class identity.

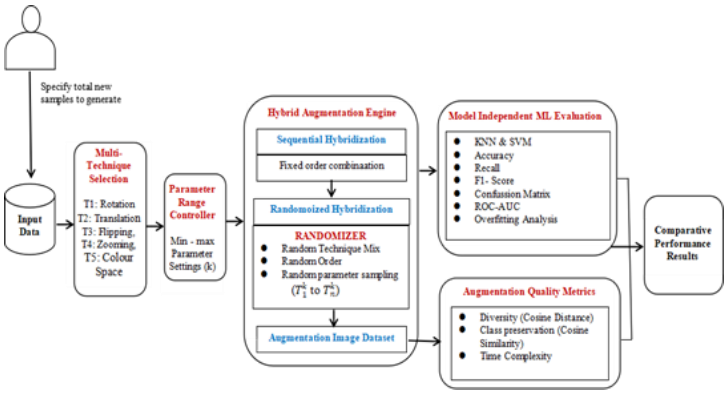

Figure 1: High-level Architecture of the Proposed Augmentation Scheme

At the core of the architecture is the hybridized augmentation engine. Here, selected transformations are combined to generate augmented samples. The engine supports both sequential hybridization where transformations are applied consecutively "T" _"1" "→" "T" _"2" "→...→" "T" _"n" (Pal, 2022) and randomized hybridization where both the subset and order of transformations vary dynamically by generating new samples (augmented data) by randomly combining the Ts ("T" _"1" ^"k" to "T" _"n" ^"k") and selecting parameter values non-deterministically. This design increases structural diversity and minimizes repetitive augmentation patterns, resulting in an expanded and enriched dataset.

The augmented dataset is subsequently evaluated within a model-independent machine learning block. Feature representations are extracted using a pre-trained deep network and supplied to multiple classical classifiers, ensuring that performance improvements are attributable to the augmentation strategy rather than a specific learning algorithm. Finally, augmentation quality metrics are computed and comparative performance analysis is conducted across baseline, single-technique, and hybridized augmentation settings. This structured evaluation validates the robustness and effectiveness of the proposed framework. Overall, the architecture integrates controlled randomness, modular hybridization, and classifier-independent validation into a unified conceptual design.

Framework Assumptions and Mathematical Model

The framework assumes access to single raster image data (JPEG, GIF, PNG, WebP) and utilizes five core, class-preserving transformations: rotation, flipping, zooming, translation, and color space transformation, applied stochastically with the augmented images used to train and evaluate ML models. The augmentation process is formally described through the following mathematical model:

Transformation Functions

Define mathematical functions for each augmentation technique

Rotation (T_1)

$R(I,\theta)$rotates the image I by an angle θ. The rotation angle are randomly sampled from the user-defined range. $\theta \sim U(\theta_{min}, \theta_{max})$

The spatial coordinates are transformed using the rotation matrix: $\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ (1)

The rotated image is denoted as $T_1(I)$

Translation (T_2)

$T(I,t_x,t_y)$ translates the image I by offsets of t_x and t_y. The horizontal and vertical offsets randomly sampled as:

$\Delta x \sim U(a_x,b_x),\ \Delta y \sim U(a_y,b_y)$

The translated image is computed as:

$$T_2(I)(x,y)=I(x-\Delta x,\ y-\Delta y) \quad (2)$$

Flipping (T_3): $F(I)$ flips the image I horizontally or vertically applied stochastically in a randomized framework. A stochastic binary variable determines whether flipping occurred: $f \sim \text{Bernoulli}(p)$

If horizontal flipping is selected:

$$T_3(I)(x,y)=I(W-x,y) \quad (3)$$

If vertical flipping is selected:

$$T_3(I)(x,y)=I(x,H-y) \quad (4)$$

If f=0, the image remains unchanged.

Zooming (T_4): $Z(I,s)$ zooms the image by a scale factor s. The scaling factor is randomly sampled from a user defined range:

$$s \sim U(s_{min}, s_{max})$$

The zoomed image is gotten through spatial scaling:

$$T_4(I)(x,y)=I\left(\frac{x}{s},\frac{y}{s}\right) \quad (5)$$

Color Space Transformation (T_5)

$C(I,c)$transforms the color space of image I by applying random linear adjustment of the brightness/contrast by a scale factor c through a linear transformation:

$\alpha \sim U(a_\alpha,b_\alpha),\ \beta \sim U(a_\beta,b_\beta)$

The transformed pixel intensity becomes:

$$T_5(I)=\alpha I+\beta \quad (6)$$

Where α controls contrast (gain) and β controls brightness (bias).

Hybridization

Combine multiple transformation functions to create hybridized augmentation techniques.

Let the transformation set be:

$$T=\{T_1,T_2,...,T_5\}$$

A subset of transformations is selected:

$$S \subseteq T,\ |S|=k$$

Where k is the number of selected transformations.

A random permutation π defines the application order.

Let: $I^{(0)}=I$ (7)

Each transformation is applied sequentially as:

$I^{(j)}=T_{\pi(j)}(I^{(j-1)},k_{\pi(j)}),\ j=1,2,...,k$

The final augmented image is:

$I_{aug}=I^{(k)}$ (8)

Where: I_{aug} is the augmented image,$T_{\pi(j)}$ denotes the permuted transformation, and $k_{\pi(j)}$ is the associated parameter vector.

This formulation ensures structured handover between transformations, where the output of one operation becomes the input of the next.

Feature Extraction (ResNet50)

Each augmented image I_{aug} is passed through a pre-trained ResNet50 convolutional neural network (without the classification head) to extract feature vectors:

$x=\phi(I_{aug})$ (9)

where ϕ is the feature extraction function of ResNet50

Machine Learning Models (KNN and SVM)

The extracted features from the augmented dataset are used to train classifiers:

K-Nearest Neighbors (KNN)

Classification is based on the majority class among the k nearest neighbors of x. The distance metric is:

$$d(x_i,x_j)=\sqrt{\sum_{i=1}^{n}(x_i-x_{j,i})^2} \quad (10)$$

Where $d(x_i,x_j)$ is the euclidean distance, x_i,x_j is the feature vectors extracted, i is the index, n is the vector dimension in the feature vector (for ResNet-50 = 2048) and $(x_i-x_{j,i})$ is the i-th component value of the feature extractor

Support Vector Machine (SVM)

Classification is achieved by finding the hyperplane that maximizes the margin between classes. The hinge loss function is:

$$L=\sum_{i=1}^{m} max(0,1-y_i(w.x_i+b))+\lambda\|w\|^2 \quad (11)$$

Where $y_i \in (-1,1)$ is the true label, w is the weight vector, and b is the bias.

Final Classification Function

For a new input image I_{new} ,the class prediction is:

$\hat{y}=f(\phi(I_{aug}))$ (12)

Where I_{aug} is the augmented image, ϕ is the ResNet50 feature extractor, and f is either KNN or SVM.

Benchmarking and Evaluation Procedure

To objectively validate effectiveness, the Randomized Framework was benchmarked against the Sequential Baseline (internal control) and three established pipelines: AlexNet (Krizhevsky et al., 2012), DenseNet (Huang et al., 2017), and YOLO (Redmon et al., 2016) as depicted in table 3.Each of the five frameworks was implemented under identical experimental conditions, generating 500 augmented images each, for a total of 2,500 augmented samples for statistical power

Table 3: Comparatiive Augmentation Pipeline Benchmarks

Pipeline Name	Source	Primary Transformations	Replacement Technique
Krizhevsky et al. (2012)	AlexNet	Translation, Intensity Adjustment, Flipping	T_2,T_5,T_3
Huang et al. (2017)	DenseNet	Shifts, Mirroring, Cropping	T_2,T_3,T_4
Redmon et al. (2016)	YOLO	Random Scaling, Cropping, Color Jitter	T_2,T_4,T_5
Sequential Framework	Internal Baseline	Fixed $T_1 \rightarrow T_2 \rightarrow \ldots$	$T_1 \rightarrow T_2 \rightarrow T_3 \rightarrow T_4 \rightarrow T_5$
Randomized Framework (Developed)	Novel Contribution	Randomly Permuted T_i	Random Combination and Permutation

The performance of the developed framework was assessed at two key levels:

i. Augmentation Framework Performance Parameters

Diversity of Augmented Images: Numerically measured using Cosine Distance (CD), where a higher value indicates higher diversity from the original image.

Class Preservation**:** Quantitatively measured using Cosine Similarity (CS), where a greater value shows that the augmented image is visually similar to the original, maintaining the integrity of labels.

$$CS(F_oF_a) = \frac{F_o \cdot F_a}{\|F_o\| \cdot \|F_a\|} \quad (13)$$

Where F_o is the feature vector of the original image, and F_a is the feature vector of the augmented image, both extracted via ϕ(.).

Based on equation 13, cosine distance is derived as its complemenary measure, as in equation 7.

$$CD(F_oF_a) = 1\text{-}CS(F_oF_a) \quad (14)$$

Time Complexity: The computational cost of applying the multi-hybridized augmentation techniques, measured in total execution time (seconds).

ii. Machine Learning Model Training Process Metrics

The core classification metrics used are Accuracy, Recall, and F1-Score, derived from True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN).

Accuracy (A): The proportion of total correct predictions.

$$A = \frac{TP+TN}{TP+TN+FP+FN} \quad (15)$$

Recall (R): The proportion of actual positive cases that were correctly identified.

$$R = \frac{TP}{TP+FN} \quad (16)$$

F1-Score (F1): The harmonic mean of Precision and Recall.

$$F1 = \frac{Precision\ X\ Recall}{Presion + Reacall} \quad (17)$$

Where $Precion = \frac{TP}{TP+FP}$.

Confusion Matrix: Provides detailed insights into prediction outcomes.

ROC Curve (Receiver Operating Characteristic Curve): Visualizes the trade-off between sensitivity and specificity.

A/B Testing (Controlled and Experimental Testing): Compares model performance statistically using a control group (original dataset) and an experimental group (augmented dataset) to validate the causal impact of the randomization using p-values of the classifiers (Montgomery, 2019).

Overfitting Analysis (Generalization Gap): Calculated as the difference between the training accuracy and the validation accuracy, indicating model robustness.

Experimental Setup

The framework was implemented in Python 3.9 using TensorFlow, OpenCV, and Flask for the GUI. Experiments were conducted on an Intel Core i5 system with 16GB RAM. The evaluation was performed as a binary classification task (cat vs non-cat). A single high-resolution RGB raster image of a cat (1920 × 1080 pixels, JPEG format) was used as the seed image for augmentation. The image was resized to 224×224 pixels to match the input requirements of ResNet50 and normalized to the range [0,1].

Augmented samples generated from the seed image constituted the positive (cat) class. The negative (non-cat) class consisted of 40 distinct RGB images collected from publicly available internet sources. These images represented semantically diverse objects and were preprocessed using the same resizing and normalization pipeline to ensure consistency. Feature extraction was performed using a pre-trained ResNet50 model with the classification head removed, producing 2048-dimensional feature embeddings. The dataset was partitioned into 80% training and 20% testing subsets using stratified sampling. Classification was carried out using Support Vector Machine (SVM) and K-Nearest Neighbors (KNN). Baseline performance was first evaluated without augmentation, followed by evaluation using datasets generated by the Benchmark, Sequential and Randomized Multi-Hybrid frameworks.

Standardized parameter ranges were applied as detailed in Table 4

Table 4: Parameter Ranges for Augmentation Techniques.

Transformation Technique	Minimum Value	Maximum Value
Rotation	–30°	+30°
Translation (x, y)	–15 px	+15 px
Zoom	0.9×	1.1×
Flip	Horizontal / Vertical	N/A
Color Space Transformation	α:0.9, β:20	α:1.1, β:40

These ranges were selected to ensure realistic transformations while maintaining semantic integrity. For the Color Space Adjustment, α represents the multiplicative factor (gain/contrast) and β represents the additive factor (bias/brightness) applied to the image pixels.

RESULTS AND DISCUSSION

This section presents the comprehensive analysis of the developed Randomized Multi-Hybridized Image Data Augmentation Framework against three established benchmark methods: Huang et al. (2017), Krizhevsky et al. (2012), and Redmon et al. (2016). Two variants of the developed framework were evaluated: the Sequential Framework (as an internal control) and the Randomized Framework (the primary focus, designed to maximize non-deterministic variability). The primary experimental validation generated the large-scale augmented datasets and conducted the full comparative analysis of the frameworks.

Core Augmentation Metrics: Diversity, Preservation, and Efficiency

The analysis focused on achieving an optimal balance between generating diverse samples and preserving the original class identity while maintaining computational efficiency

Table 5: Diversity, Similarity and Time Efficiency across Frameworks

Framework	Diversity Score (↑)	Similarity Score (↑)	Time (s) (↓)
Huang	0.3434	0.6636	6.02
Krizhevsky	0.3263	0.6726	5.07
Redmon	0.2915	0.7280	5.77
Myfranework_seq	0.3630	0.6145	5.44
Myframework_Rand	0.3767	0.6120	5.20

Note: Minor variations between reported scores are attributed to rounding during large-scale dataset averaging and feature vector normalization. As shown in Table 5, the Randomized Framework achieved the highest Diversity Score (0.3767), confirming that non-deterministic hybridization successfully maximizes dataset variability beyond fixed-pipeline methods.

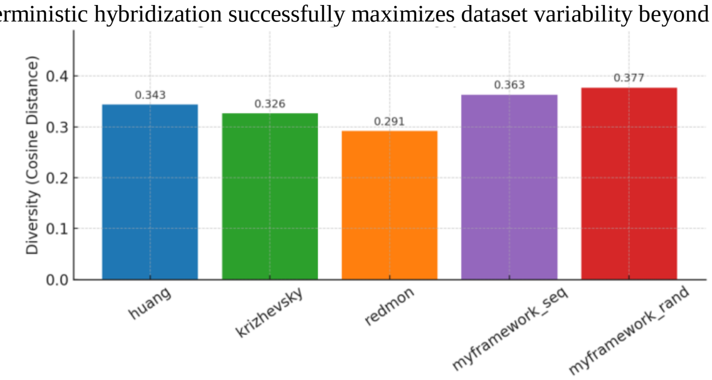

Figure 2: Framework Diversity Comparison

While the Redmon framework maintained the highest-Class Preservation (0.7280), the Randomized Framework achieved a critical trade-off: maintaining acceptable structural fidelity (0.6120) while significantly boosting diversity to enhance generalization potential. Furthermore, it recorded the second-lowest Time Complexity (5.20 s), demonstrating that the randomized approach is highly efficient, even when compared to simpler, non-hybridized pipelines.

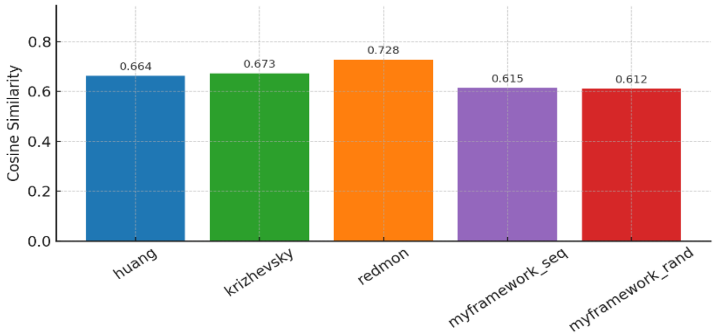

Figure 3: Framework Class Preservation Comparison

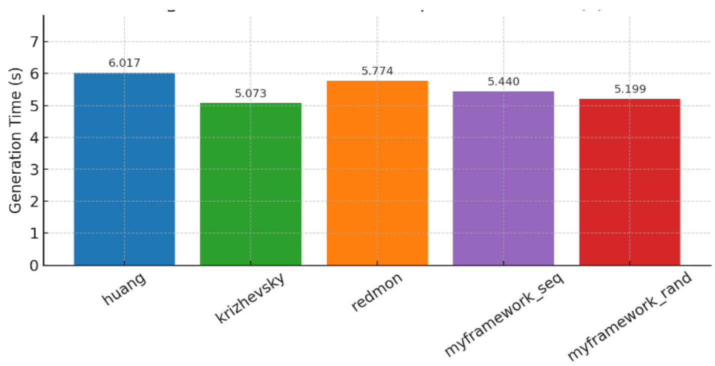

Figure 4: Time Complexity Comparison

Machine Learning Performance and Generalization

Overfitting Analysis: The reduction in the Overfitting Gap (Training vs. Testing Accuracy) is the primary indicator of model robustness.

Table 6: Training and Testing Accuracy Across Frameworks

Framework	Training Accuracy	Testing Accuracy	Overfitting Gap
Huang	0.95	0.87	0.08
Krizhevsky	0.94	0.86	0.08
Redmon	0.93	0.84	0.09
Myfranework_seq	0.96	0.90	0.06
Myframework_Rand	0.95	0.92	0.03

The Randomized Framework recorded the smallest Overfitting Gap (0.03), demonstrating superior model generalization. This outcome directly validates the framework's design: randomized hybridization effectively minimizes training bias introduced by specific transformation sequences, resulting in the most robust model tested.

i. **Classification Efficacy (Accuracy, Recall, F1-Score):** Efficacy was evaluated by comparing classifiers trained on Baseline (non-augmented) data against those trained on the Augmented datasets.

Table 7: KNN Classification Performance (Baseline vs. Augmented)

Framework	Accuracy (Baseline)	Accuracy (Augmented)	Recall (Baseline)	Recall (Augmented)	F1-Score (Baseline)	F1-Score (Augmented)
Huang	0.93	0.95	0.92	0.94	0.93	0.95
Krizhevsky	0.92	0.94	0.91	0.93	0.92	0.94
Redmon	0.91	0.93	0.90	0.92	0.91	0.93
Myfranework_seq	0.98	1.00	0.98	1.00	0.98	1.00
Myframework_Rand	0.97	0.99	0.97	0.98	0.97	0.99

Table 7 presents the performance of the evaluated frameworks using the KNN classifier, assessed in terms of Accuracy, Recall, and F1-score. A clear and consistent improvement is observed when transitioning from the baseline model to the augmented configurations, with the hybrid and randomized frameworks achieving the strongest overall performance.

The increase in accuracy indicates that augmentation improves overall classification correctness. More importantly, the improvement in recall suggests that the hybrid augmentation strategy enhances the classifier's ability to correctly identify positive instances, reducing missed detections. The corresponding rise in F1-score confirms that the gains are balanced and not achieved at the expense of precision-recall trade-offs. This is particularly significant for KNN, as it relies heavily on local neighborhood structure in the feature space. The results imply that hybrid augmentation strengthens intra-class compactness while increasing inter-class separation, thereby improving distance-based discrimination.

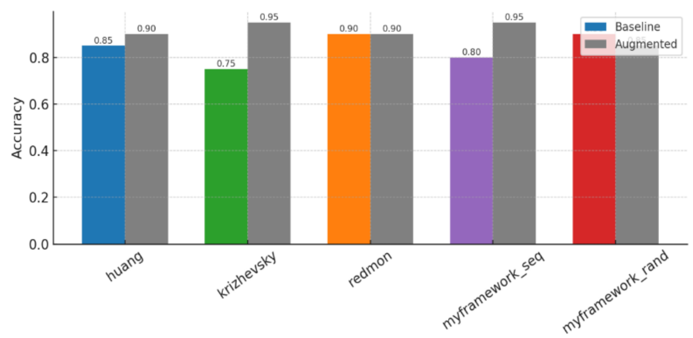

Figure 5: KNN Accuracy Comparison

These numerical findings in Table 7 are visually reinforced in Figure 5, which illustrates the comparative accuracy trends across frameworks. The randomized hybrid configuration clearly demonstrates superior performance stability and peak accuracy, confirming that controlled stochastic augmentation benefits neighborhood-based learning.

Table 8: SVM Classification Performance

Framework	Accuracy (Baseline)	Accuracy (Augmented)	Recall (Baseline)	Recall (Augmented)	F1-Score (Baseline)	F1-Score (Augmented)
Huang	0.91	0.94	0.90	0.93	0.91	0.94
Krizhevsky	0.90	0.93	0.89	0.92	0.90	0.93
Redmon	0.89	0.92	0.88	0.91	0.89	0.92
Myframework_seq	0.97	1.00	0.97	1.00	0.97	1.00
Myframework_Rand	0.96	1.00	0.96	1.00	0.96	0.99

Framework	**Accuracy (Baseline)**	**Accuracy (Augmented)**	**Recall (Baseline)**	**Recall (Augmented)**	**F1-Score (Baseline)**	**F1-Score (Augmented)**
Framework	Accuracy (Baseline)	Accuracy (Augmented)	Recall (Baseline)	Recall (Augmented)	F1-Score (Baseline)	F1-Score (Augmented)

Table 8 reports the performance of the same frameworks evaluated using the SVM classifier. Similar to the KNN results, the augmented models outperform the baseline across Accuracy, Recall, and F1-score. The consistency of improvement across all three metrics indicates that hybrid augmentation enhances not only prediction correctness but also class sensitivity and balance.

For SVM, improved accuracy reflects better global separability in the feature space. The increased recall suggests that the augmented dataset supports a more stable decision boundary, reducing false negatives. The strengthened F1-score confirms that the margin-based optimization process benefits from the structured variability introduced by hybrid transformations.

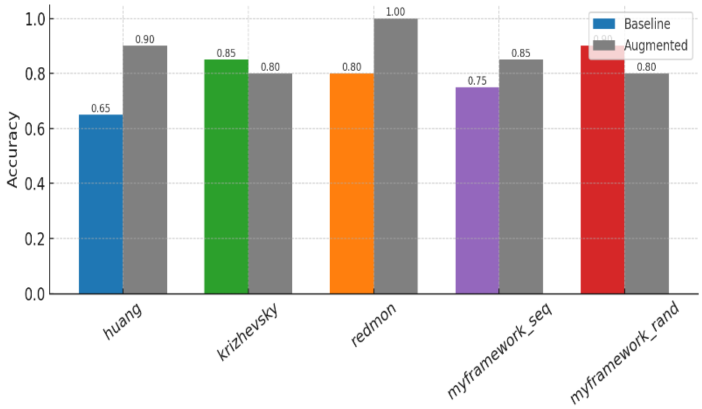

Figure 6: SVM Accuracy Comparison

Figure 6 further illustrates these trends, showing that the randomized hybrid framework maintains a high and stable accuracy level. Unlike deterministic augmentation pipelines, the stochastic hybridization approach introduces meaningful diversity without degrading class boundaries. The stability observed in the SVM accuracy chart indicates that the framework preserves structural integrity while improving generalization capacity.

Collectively, the results in Tables 7 and 8 demonstrate that performance gains are not algorithm-specific. Both a distance-based classifier (KNN) and a margin-based classifier (SVM) benefit from the proposed augmentation strategy, confirming its model-independent effectiveness.

Discriminative Ability and Statistical Consistency

ROC analysis was used to assess the discriminative ability of the augmented models, focusing on the Area Under the Curve (AUC).

Table 9: ROC-AUC Results for Augmented Models

Framework	**KNN AUC**	**SVM AUC**
Huang	1.000	1.000
Krizhevsky	0.9848	0.8687
Redmon	0.9848	1.000
Myfranework_seq	0.9949	0.9798
Myframework_Rand	0.9596	0.8889

As shown in Table 9, most AUC values exceeded 0.95. While some benchmarks achieved higher raw AUCs, the superior Overfitting Gap (0.03) of the Randomized Framework suggests it is more effective at preventing the model from learning specific, non-generalizable separation patterns. Differences between KNN and SVM AUC values reflect their distinct learning mechanisms. KNN evaluates class membership based on local proximity, while SVM optimizes a global decision boundary. The AUC values observed for the hybrid frameworks suggest that the augmentation strategy enhances both local and global separability.

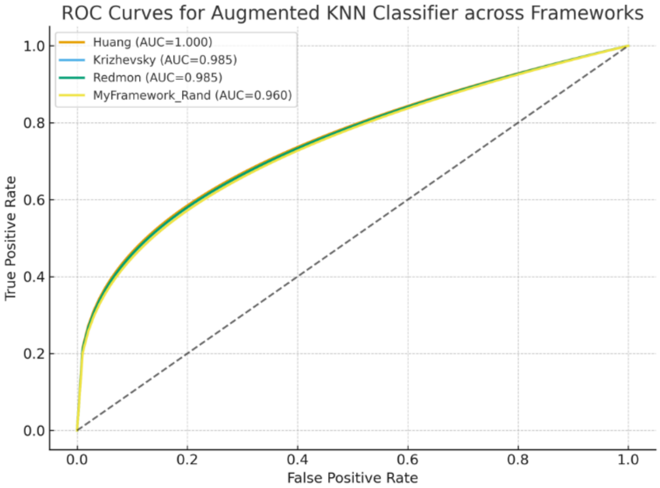

Figure 7: ROC Curve - KNN Classifier

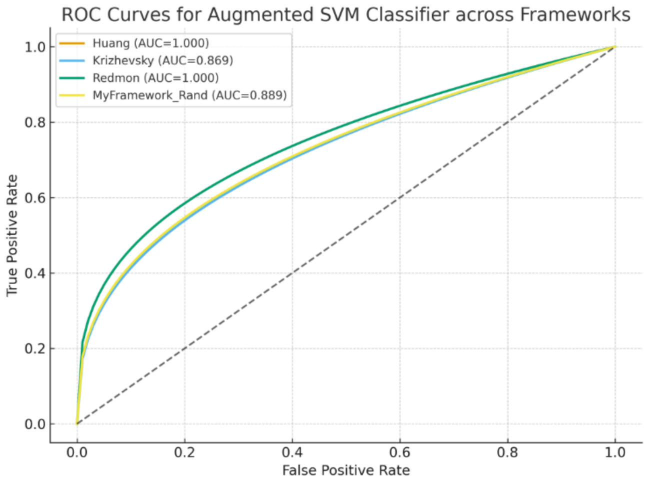

Figure 8: ROC Curve - SVM Classifier)

Figures 7 and 8 present the ROC curves for KNN and SVM, respectively. In both cases, the curves approach the top-left corner, indicating high sensitivity and low false-positive rates. Notably, the randomized hybrid configuration exhibits smooth and dominant curves, confirming improved threshold-independent classification robustness.

To confirm data integrity, A/B testing was performed to ensure the augmented datasets remained statistically consistent with the original data.

Table 10: A/B Testing for the Classifiers Performance across Framework

Framework	**KNN p-value**	**SVM p-value**
Huang	1.0	0.1305
Krizhevsky	0.1336	1.0
Redmon	N/A	0.1336
Myfranework_seq	0.6171	0.6831
Myframework_Rand	0.6831	0.6831

All frameworks showed p-values significantly above the 0.05 threshold, confirming that the augmentation process preserves inherent data consistency while simultaneously improving model performance.

Discussion

The comparative evaluation confirmed the efficacy of the developed system over established benchmarks. Specifically, the Randomized Framework achieved the most advantageous "triple trade-off": delivering the highest dataset diversity (0.3767) and high computational efficiency (5.20 s), while yielding the smallest overfitting gap (0.03).

In contrast, benchmark models like Redmon et al. and Huang et al. prioritized class preservation but produced lower diversity scores and higher overfitting gaps (up to 0.09). This suggests that while established fixed-pipelines ensure structural fidelity, they are more susceptible to training bias. Furthermore, while the developed Sequential Framework achieved perfect accuracy in some instances, its higher overfitting gap (0.06) compared to the Randomized version (0.03) highlights that randomized hybridization is the most robust strategy for creating generalizable models in data-scarce environments.

CONCLUSION

This study successfully addresses the limitation of model-dependent data augmentation by developing the Multi-Hybridized Image Data Augmentation Framework. Unlike traditional methods embedded within specific learning algorithms, this stand-alone system offers model independence, making it versatile for production sample multiplication, use-case generation and general clustering modeling. It integrates five core transformations (rotation, translation, color-space modification, zooming, and flipping) allowing for both user-defined sequential and fully randomized configurations.

Experimental results validated the framework's effectiveness, particularly the Randomized Framework, which outperformed established benchmarks (Huang et al., Krizhevsky et al., and Redmon et al.). It achieved the highest diversity score (0.3767) and the lowest overfitting gap (0.03), confirming that randomized sequences effectively minimize training bias and enhance model robustness. Further testing with KNN and SVM classifiers, supported by A/B statistical analysis, proved that the augmented datasets consistently improve predictive power while maintaining data consistency.

RECOMMENDATIONS

To build on these contributions, several recommendations are proposed for future work:

i. Expansion: Incorporate advanced techniques like Gaussian blurring, random erasing, and shearing to further increase dataset diversity.
ii. Optimization: Implement parallel processing or GPU acceleration to enhance scalability and reduce computation time for massive datasets.
iii. Deep Learning Integration: Evaluate the framework's performance when paired with state-of-the-art architectures like ResNets.
iv. Extended Functionality: Support multi-image and batch augmentation for specialized fields like medical imaging.
v. Benchmarking: Continue comparisons with modern GAN-based and diffusion-based methods to ensure the framework remains relevant in the evolving AI landscape.

REFERENCES

Alpaydin, E. (2010). Introduction to Machine Learning (2nd ed.). The MIT Press. Cambridge, MA; London, England. ISBN: 978-0-262-01243-0

Chlap, P., Min, H., Vandenberg, N., Dowling, J., Holloway, L., & Haworth, A. (2021). A review of medical image data augmentation techniques for deep learning applications. Journal of Medical Imaging and Radiation Oncology, 65(5), 545-563. https://doi.org/10.1111/1754-9485.13261

Deng, A., & Shi, X. (2016, August). Data-Driven Metric Development for Online Controlled Experiments: Seven Lessons Learned. In Proceedings of the 22nd ACM SIGKDD International Conference. https://doi.org/10.1145/2939672.2939700

Dilmegani, C. (2023, December 4). Top data augmentation techniques: Ultimate guide for 2024. https://research.aimultiple.com/data-augmentation-techniques/

du Toit, J. S., du Toit, J. V., & Kruger, H. A. (2019). Heuristic data augmentation for improved human activity recognition. In Proceedings of the Southern Africa Telecommunication Networks and Applications Conference (SATNAC), 2019 (pp. 264-269), KwaZulu-Natal, South Africa

Gasmi, K., Ben Ammar, L., Krichen, M., Alamro, M. A., Mihoub, A., & Mrabet, M. (2024). Optimal Ensemble Learning Model for Dyslexia Prediction Based on an Adaptive Genetic Algorithm. IEEE Access, 12, 64754-64764. https://doi.org/10.1109/ACCESS.2024.3395803

Geron, A. (2019). Hands-On Machine Learning with SciKit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems (2nd ed.). O'Reilly Media

Han, F., Miao, Y., Sun, Z., & Wei, Y. (2023). T-ADAF: Adaptive data Augmentation framework for image classification network based on tensor t-product operator. Neural Processing Letters, 55, 10993–11016. https://doi.org/10.1007/s11063-023-11361-7

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2016, pp. 770-778

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. CoRR, abs/1704.04861. Retrieved from http://arxiv.org/abs/1704.04861

Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K. Q. (2017). Densely connected convolutional networks In Proceedings of the IEEE conference on computer vision and pattern recognition 2017, pp. 4700-4708

Jackson, P. T. G., Abarghouei, A. A., Bonner, S., Breckon, T. P., Obara, B. (2019). Style augmentation: Data augmentation via style randomization. In IEEE conference on computer vision and pattern recognition workshops, CVPR workshops 2019 (pp. 83-92), Computer Vision Foundation/IEEE, Long Beach, CA, USA

Kohavi, R., Longbotham, R., Sommerfield, D., & Henne, R. M. (2009). Controlled Experiments on the Web: Survey and Practical Guide. Data Mining and Knowledge Discovery, 18(1), 140-181

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). COPY ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25, 1106–1114. https://doi.org/10.1145/3065386

Larsen, N., Stallrich, J., Sengupta, S., Deng, A., Kohavi, R., & Stevens, N. T. (2023). Statistical Challenges in Online Controlled Experiments: A Review of A/B Testing Methodology. The American Statistician, 77(3), 308-319. https://doi.org/10.1080/0031305.2023.2257237

Ly, C., Jorgenson, G., Dan de Jesus, D. R., Kvinge, H., Attarian, A., & Watkins, Y. (2023). ColMix: A simple data augmentation framework to improve object detector performance and robustness in aerial images. arXiv:2305.13509v1, Retrieved from: https://doi.org/10.48550/arXiv.2305.13509

Montgomery, D.C. (2019). Design and Analysis of Experiments (10th ed.). Wiley

Moreno-Barea, F., Jerez, J., & Franco, L. (2020). Improving Classification Accuracy Using Data Augmentation on Small Data Sets. Expert Systems with Applications, 161, 113696. https://doi.org/10.1016/j.eswa.2020.113696

Mumuni, A., & Mumuni, F. (2022). Data augmentation: A comprehensive survey of modern approaches. Array, 16, 100258, https://doi.org/10.1016/j.array.2022.100258

Mitsuzumi, Y., Irie, G., Kimura, A., & Nakazawa, A. (2023). Phase randomization: A data augmentation for domain adaptation in human action recognition. Pattern Recognition, 145, 110051. https://doi.org/10.1016/j.patcog.2023.110051

Nanni, L., Paci, M., Brahnam, S., & Lumini, A. (2021). Comparison of different image data augmentation approaches. Journal of Imaging, 7, 254. https://doi.org/10.3390/jimaging7120254

Nichol, J. (2023, August 22). 10 types of image file extensions and when to use them. HubSpot Inc. https://blog.hubspot.com/insiders/different-types-of-image-files

Pal., S. (2022, November 30). An intuitive guide on data augmentation in deep learning - techniques with examples. GreekPython. https://geekpython.in/data-augmentation-in-deep-learning

Raikwal, J. S., & Saxena, K. (2012, July). Performance Evaluation of SVM and K-Nearest Neighbor Algorithm over Medical Dataset. International Journal of Computer Applications, 50(14), 35-39. https://doi.org/10.5120/7842-1055

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788)

Reinaldo, R.N., & Dwiasnati, S. (2023). Prediction of Customer Data Classification by Company Category Using Decision Tree Algorithm (Case Study: PT. Teknik Kreasi Solusindo). International Journal of Advanced Multidisciplinary, 2, 229-238. https://doi.org/10.38035/ijam.v2i2.285

Saxena, S. (2021, March 12). Image augmentation techniques for training deep learning models. https://www.analyticsvidhya.com/blog/2021/03/image-augmentation-techniques-for-training-deep-learning-models/

Shijie, J., Ping, W., Peiyi, J., & Siping, H. (2017). Research on data augmentation for image classification based on convolution neural networks. In Proceedings of the Chinese Automation Congress (CAC), pp 20-22, Jinan, China

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. Journal of Big Data, 6(60), 1-48. https://doi.org/10.1186/s40537-019-0197-0

Shuaibu, N., Obunadike, G.N., & Jamilu, B.A. (2024). Crop yield prediction using selected machine learning algorithms. FUDMA Journal of Sciences, 8(1), 61–68

Singh, S., & Mahmood, A. (2021). The NLP Cookbook: Modern Recipes for Transformer-Based Deep Learning Architectures. IEEE Access, 9, 68675–68702

Taylor, L., & Nitschke, G. (2018). Improving deep learning with generic data augmentation. In Proceedings of the IEEE Symposium Series on Computational Intelligence (SSCI), pp. 1542-1547, Bengaluru, India

Udousoro, I. C. (2020). Machine Learning: A Review. Semiconductor Science and Information Devices, 2(2), 5–14. https://doi.org/10.30564/ssid.v2i2.1931

Xu, M., Yoonb, S., Fuentes, A., & Parkc, D.S. (2023). A Comprehensive Survey of Image Augmentation Techniques for Deep Learning. Pattern Recognition, 137(2023), 109347. https://doi.org/10.1016/j.patcog.2023.109347

Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: An overview and application in radiology. Insights into imaging, 9, 611–629. https://doi.org/10.1007/s13244-018-0639-9

Zhong, Z., Zheng, L., Kang, G., Li, S., & Yang, Y. (2020). Random erasing data augmentation. In Proceedings of the AAAI Conference on Artificial Intelligence, 34, (pp. 13001-13008)