International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(14s) 2026 pp. 718-728
https://doi.org/10.70917/ijcisim-2026-4278
Article

Adaptive Hybrid Seam Carving Through Stable Alpha-Regime Prediction: A Cross-Dataset Evaluation

Anusree K. Sanil[1], Venifa Mini G.[2]

[1]Noorul Islam Centre for Higher Education, Kanyakumari, Tamil Nadu, India.
Email: **anusreesanil21@gmail.com**
[2]Department of Computer Science and Engineering, Noorul Islam Centre for Higher Education, Kanyakumari, Tamil Nadu, India.
Email: **vmminijohnny@gmail.com**

Abstract: Hybrid seam carving combines gradient-based structural energy and saliency-based semantic energy using a weighting parameter, alpha. Using a fixed alpha for every image may not adequately represent variations in image content. This study investigates whether four lightweight image features—edge density, mean gradient magnitude, entropy, and saliency coverage—can predict broad alpha-weighting regimes. For each image, nine candidate alpha values were evaluated, and provisional labels were assigned according to similarity with a conventional seam-carving reference. These labels therefore represent reference-derived weighting preferences rather than perceptually optimal alpha values. Exact-value regression produced negative cross-validated R^2 values, while a preliminary three-class formulation failed to identify the Transition category reliably. The task was consequently reformulated as binary classification of stable Semantic ($\alpha \leq 0.4$) and Structural ($\alpha \geq 0.8$) regimes. Label-stability analysis showed that 88.5% of RetargetMe images and 91.4% of NRID images had near-maximum-SSIM alpha values confined to one reference-derived regime. On 61 stable binary RetargetMe samples, class-balanced Logistic Regression achieved a balanced accuracy of 0.728±0.157, a macro F1-score of 0.658±0.140, and a ROC-AUC of 0.728±0.194 under repeated stratified cross-validation. For the selected configuration, a permutation test produced $p=0.002$. On 23 independent stable binary NRID samples, the model obtained a balanced accuracy of 0.623, a macro F1-score of 0.518, and an MCC of 0.226. The findings demonstrate statistically significant reference-regime predictability within RetargetMe but only modest cross-dataset transfer. Therefore, the proposed framework should be interpreted as a feasibility study of lightweight reference-derived alpha-regime prediction rather than a complete solution for perceptually optimal alpha selection.
Keywords: Image Retargeting; Hybrid Seam Carving; Alpha-Regime Prediction; Label Stability; Logistic Regression; RetargetMe; NRID; Cross-Dataset Evaluation.

1. INTRODUCTION

Content-aware image retargeting changes image dimensions while attempting to preserve visually and semantically important content. Recent surveys have reviewed developments and challenges in AI-assisted seam-carving methods [1]. Unlike uniform

scaling, which may distort the entire image, seam carving iteratively removes or inserts connected paths of low-energy pixels known as seams [2]. The quality of the retargeted output therefore depends strongly on the energy function used to identify removable or insertable seams.

Traditional seam carving primarily uses gradient-based energy to preserve edges and structural information. Although effective for textured and architectural regions, gradient energy does not explicitly represent semantic importance and may remove visually significant objects such as faces, text, or foreground subjects. Saliency-based energy addresses this limitation by assigning greater importance to visually prominent regions. Hybrid seam-carving formulations combine gradient-based structural energy and saliency-based semantic energy as follows:

$$Ehybrid = \alpha Egradient + (1 - \alpha)Esaliency$$

α controls the relative contribution of the two energy components. Lower alpha values emphasize semantic saliency, whereas higher alpha values emphasize structural gradients.

Existing hybrid approaches commonly apply a single manually selected alpha value to every image [3]. This assumption overlooks substantial differences in image composition. An image containing a dominant foreground subject may require greater saliency weighting, whereas an architectural or highly textured image may benefit from stronger gradient weighting. Selecting alpha manually for each image is subjective, time-consuming, and unsuitable for automated or large-scale retargeting systems. Automatic image-dependent alpha selection is therefore necessary for making hybrid seam carving genuinely adaptive.

Predicting alpha directly presents two important challenges. First, different alpha values can produce very similar SSIM scores for the same image, making the identity of a single "best" alpha potentially uncertain. Second, the available benchmark datasets are modest and exhibit an imbalanced concentration of images in high-alpha regimes. Treating every maximum-SSIM alpha as an equally reliable provisional reference-derived label may consequently introduce unstable targets and misleading performance estimates.

To address these challenges, this study introduces a stable alpha-regime prediction framework. Nine candidate alpha values between 0.1 and 0.9 are evaluated for each image using seam-carving-specific reference outputs. All alpha values within 1% of the maximum SSIM are treated as near-optimal candidates. A label is considered regime-stable only when these candidates remain within the same Semantic, Transition, or Structural regime. Final modelling focuses on stable Semantic ($\alpha \leq 0.4$) and Structural ($\alpha \geq 0.8$) cases, while ambiguous transition and regime-unstable cases are not forced into potentially unreliable classes.

Four lightweight image descriptors—edge density, mean gradient magnitude, entropy, and saliency coverage—are used to predict the stable alpha regime. Logistic Regression, Support Vector Machine, and Random Forest classifiers are evaluated using repeated stratified cross-validation on RetargetMe. Because the resulting classes are imbalanced, balanced accuracy is used as the primary evaluation metric, supported by macro F1-score, ROC-AUC, Matthews correlation coefficient, and permutation testing. The selected model is then evaluated without retraining on the independent NRID dataset to assess cross-dataset transfer.

This paper investigates image-dependent alpha selection through a sequence of prediction formulations. Continuous regression and three-class classification are first evaluated as diagnostic experiments. Because neither formulation provides reliable performance, the final approach focuses on stable binary Semantic and Structural regimes. The regime labels are provisionally derived from similarity to a conventional seam-carving reference and are screened to exclude transition cases and images whose near-maximum-SSIM alpha values cross regime boundaries.

The contributions of this study are fourfold. First, it examines the limitations of exact-value regression and three-class classification for predicting hybrid-energy weights. Second, it introduces a label-stability procedure that identifies images whose near-maximum-SSIM alpha values remain within a single reference-derived weighting regime. Third, it evaluates a lightweight binary framework for distinguishing stable Semantic and Structural regimes using four image features. Fourth, it provides a statistically tested cross-dataset evaluation using repeated stratified cross-validation, permutation testing, and independent validation on NRID. The study does not claim perceptually optimal alpha prediction; instead, it determines the extent to which lightweight features can predict broad reference-derived weighting regimes.

The remainder of this paper is organized as follows. Section II reviews related work on seam carving, saliency-guided retargeting, and adaptive energy formulations. Section III explains feature extraction, provisional label generation, label-stability screening, and alpha-regime prediction. Section IV presents the datasets, models, implementation, and evaluation protocol. Section V discusses the experimental results and cross-dataset findings. Section VI outlines the limitations, and Section VII concludes the study and identifies directions for future research.

2. RELATED WORK

Seam carving was introduced by Avidan and Shamir [2] as a content-aware image-resizing method that removes or inserts connected paths of low-energy pixels. The original method primarily uses gradient-based energy to preserve edges and structural information. Although it reduces the uniform distortion associated with conventional scaling, gradient energy alone may not adequately represent semantic importance. Consequently, visually important low-gradient regions, such as faces, text, or foreground objects, may be distorted or removed.

Subsequent studies introduced improved energy functions, bidirectional seam operations, foreground constraints, saliency information, and combinations of seam carving with other retargeting operators. A comprehensive review by Asheghi et al. [4] identified content preservation, computational complexity, dataset variability, and perceptual-quality assessment as continuing challenges in image retargeting. Shi et al. [5] investigated bidirectional seam-carving operations, while Chen et al. [6] combined balanced energy information with foreground constraints to improve preservation of important regions. These approaches demonstrate that seam-carving performance depends strongly on the quality and relative influence of the energy components.

Learning-based image retargeting has increasingly been used to model high-level semantic information. Song et al. [7] proposed CarvingNet, a learning-based framework designed to perform content-aware image retargeting. Cho et al. [8] introduced weakly and self-supervised learning for content-aware deep image retargeting, reducing the dependence on manually prepared paired training data. Although learning-based methods can provide stronger semantic representations than handcrafted energy maps, they generally require greater training data, computational resources, and implementation complexity.

The authors' earlier work [3] combined Random Forest-based saliency estimation with gradient energy in a lightweight hybrid seam-carving framework. However, a fixed manually selected alpha value was applied across images. The present study extends that direction by investigating whether lightweight image features can predict broad, image-dependent weighting regimes. Unlike methods that directly learn complete retargeted outputs, the proposed approach focuses on selecting the relative contribution of existing structural and semantic energy terms. It also explicitly evaluates label stability, class imbalance, statistical significance, and transfer from RetargetMe to the independently held-out NRID dataset. The reference-derived labels are treated as provisional weighting preferences and not as perceptually validated optimal alpha values.

3. METHODOLOGY

A. Framework Overview

The proposed framework predicts an appropriate alpha-weighting regime for an input image before hybrid seam carving is performed. It consists of five main stages: image-feature extraction, candidate-alpha evaluation, provisional reference-derived alpha identification, label-stability screening, and stable binary alpha-regime prediction.

The hybrid seam-carving energy function combines gradient-based structural energy and saliency-based semantic energy as follows:

$$Ehybrid = \alpha Egradient + (1 - \alpha)Esaliency,$$

where Egradient represents gradient-based structural energy, Esaliency represents saliency-based semantic energy, and α controls the relative contribution of the two energy components. Lower alpha values assign greater importance to semantic saliency, whereas higher alpha values emphasize structural gradients.

For each benchmark image, hybrid seam carving is evaluated using nine candidate alpha values. Each retargeted output is compared with the corresponding seam-carving reference image using the Structural Similarity Index (SSIM). The alpha value producing the maximum SSIM is initially recorded as the provisional reference-derived alpha. Label-stability analysis is subsequently applied to determine whether this label represents a reliable Semantic, Transition, or Structural weighting regime.

B. Image-Feature Extraction

Four lightweight image features are extracted directly from each original image before resizing: edge density, mean gradient magnitude, image entropy, and saliency coverage. These features represent complementary structural, statistical, and semantic characteristics of the image.

1. Edge Density

Edge density is defined as the proportion of image pixels identified as edges using Canny edge detection:

Edge Density = Nedge / Ntotal,

where Nedge is the number of detected edge pixels and Ntotal is the total number of image pixels. A higher edge density indicates that the image contains a greater amount of structural boundary information.

2. Mean Gradient Magnitude

Horizontal and vertical image gradients are calculated using Sobel operators. The gradient magnitude at pixel position (x,y) is calculated as:

$$G(x,y) = \sqrt{[Gx(x,y)^2 + Gy(x,y)^2]},$$

where Gx and Gy represent the horizontal and vertical Sobel gradients. Mean gradient magnitude is then obtained by averaging the gradient magnitudes across all image pixels:

$$Mean\ Gradient = (1/N)\ \Sigma\ G(x,y)$$

Images containing strong textures, architectural elements, or prominent boundaries are expected to produce higher mean-gradient values.

3. *Image Entropy*

Image entropy measures the complexity and distribution of grayscale intensity information. It is calculated from the normalized grayscale histogram as:

$$Entropy = -\Sigma\ pi\ log_2(pi)$$

where pi represents the probability of occurrence of intensity level i. Higher entropy generally indicates greater intensity variation and visual complexity.

4. Saliency Coverage

A spectral-residual saliency detector is used to generate a normalized saliency map. Saliency coverage is defined as the proportion of pixels whose saliency value exceeds a threshold of 0.5:

Saliency Coverage = Nsalient / Ntotal,

where Nsalient is the number of pixels with normalized saliency values greater than 0.5. This feature estimates how widely visually prominent information is distributed throughout the image.

These four features were selected because they can be calculated efficiently for a previously unseen image without performing seam carving. Their low dimensionality also limits overfitting risk when working with modest benchmark datasets

C. Reference-Based Provisional Label Generation

For each image, nine candidate alpha values were evaluated:

$$A = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$$

For every candidate alpha, the normalized gradient-energy and saliency-energy maps were combined using the hybrid-energy equation. Dynamic programming was applied to identify the minimum-energy vertical seam, and seams were removed iteratively until the target dimensions of the corresponding benchmark reference were reached. Images requiring height reduction were transposed before vertical seam removal and returned to their original orientation after processing.

Each candidate output was compared with the corresponding conventional seam-carving reference using the Structural Similarity Index. The provisional reference-derived alpha was defined as

$$\alpha ref = arg\ max\ \{SSIM(I\alpha, ISC) \mid \alpha \in A\}$$

Here, Iα denotes the hybrid seam-carving output generated using candidate alpha α, while ISC denotes the corresponding conventional seam-carving reference.

The conventional seam-carving output was used as a method-consistent comparison target rather than combining outputs generated by different retargeting methods. Consequently, the provisional reference-derived alpha identifies the hybrid weighting that most closely approximates the reference seam-carving output. It should not be interpreted as a perceptually optimal alpha value.

D. Diagnostic Prediction Formulations

Continuous alpha regression was initially investigated using Support Vector Regression, Random Forest Regression, and a Multi-Layer Perceptron. The four image features were used as predictors, and the provisional reference-derived alpha was used as the continuous target variable. All three regression models produced negative cross-validated R^2 values, indicating that the exact reference-derived alpha value could not be predicted reliably using the available features.

The problem was subsequently evaluated using three alpha categories:

- Low: $\alpha \le 0.4$
- Medium: $0.4 < \alpha \le 0.7$
- High: $\alpha > 0.7$

Random Forest and Support Vector Machine classifiers were compared using stratified 5-fold cross-validation. However, the Medium category was not correctly identified by the final Random Forest model during RetargetMe cross-validation. The three-class performance was also strongly influenced by the majority High category. These results indicated that intermediate alpha values formed an ambiguous transition region rather than a clearly separable category.

The regression and three-class experiments were therefore retained as diagnostic analyses. They demonstrate the limitations of exact-value and intermediate-category prediction and motivate the final stable binary formulation.

E. Alpha-Label Stability Analysis

Selecting only the alpha value with the highest reference-based SSIM may produce an unstable provisional label when several candidate alpha values obtain nearly equivalent scores. A small numerical difference could assign an image to one regime even though an alpha from another regime produces almost the same similarity. Label-stability analysis was therefore introduced to identify provisional regime labels that remained consistent among near-maximum-SSIM candidates.

All candidate alpha values whose SSIM was at least 99% of the maximum SSIM for an image were included in the near-maximum set:

$$\mathrm{Anear} = \{\alpha \in A \mid \mathrm{SSIM}(\alpha) \ge 0.99 \times \max\beta{\in}A\ \mathrm{SSIM}(\beta)\}.$$

Each alpha in this set was assigned to one of three weighting regimes:

- Semantic regime: $\alpha \le 0.4$
- Transition regime: $0.5 \le \alpha \le 0.7$
- Structural regime: $\alpha \ge 0.8$

An image was considered regime-stable when all alpha values in its near-maximum set belonged to the same weighting regime. If the set contained alpha values from two or more regimes, the image was designated regime-unstable. This procedure evaluates the consistency of the reference-derived label; it does not establish perceptual optimality.

The analysis identified 69 regime-stable images among 78 RetargetMe samples, corresponding to 88.5% of the processed dataset. For NRID, 32 of 35 images were regime-stable, corresponding to 91.4%.

Stable Transition images were excluded from the final binary task because the preliminary three-class experiment showed that the Transition category was not separable using the current feature representation. Regime-unstable images were also excluded because they could not be assigned consistently to one reference-derived weighting regime. After applying both exclusion criteria, 61 RetargetMe images and 23 NRID images remained for final binary modelling and evaluation.

F. Stable Binary Alpha-Regime Prediction

The final prediction task distinguishes two stable alpha-weighting regimes:

Semantic: stable samples with $\alpha \leq 0.4$

Structural: stable samples with $\alpha \geq 0.8$

The stable RetargetMe binary dataset contains 61 images, consisting of 11 Semantic and 50 Structural samples. The independent stable NRID binary dataset contains 23 images, consisting of 6 Semantic and 17 Structural samples.

Logistic Regression, Support Vector Machine, and Random Forest classifiers are compared using the four lightweight image features. Because the Structural category substantially outnumbers the Semantic category, class-balanced weighting is used during model training.

Three feature configurations are also evaluated:

1. Original feature set: edge density, mean gradient magnitude, entropy, and saliency coverage.
2. Enhanced feature set: gradient variation, Laplacian variance, saliency mean, saliency variation, saliency concentration, largest salient-region proportion, and colourfulness.
3. Combined feature set: all original and enhanced features.

The expanded feature configurations do not improve repeated cross-validated balanced accuracy. The final model therefore uses the original four features with class-balanced Logistic Regression.

Feature standardization and classification are implemented within a single modelling pipeline. Standardization parameters are learned only from the training portion of each cross-validation split, preventing information from validation folds from influencing preprocessing.

For a previously unseen image, edge density, mean gradient magnitude, entropy, and saliency coverage are first extracted. The four features are standardized using parameters learned exclusively from the RetargetMe training data and are then passed to the trained Logistic Regression classifier. The classifier returns either a Semantic or Structural reference-derived weighting regime.

The proposed method predicts a broad reference-derived weighting regime rather than an exact continuous alpha value. A numerical alpha must subsequently be selected from within the predicted regime before hybrid seam carving is applied. The present experiment evaluates the reliability of regime classification but does not establish an optimal within-regime alpha for a previously unseen image. This distinction prevents the classification results from being interpreted as exact or perceptually optimal alpha prediction.

4.Experimental Setup

A. Datasets

1. *RetargetMe Dataset*

RetargetMe is used as the primary dataset for feature analysis, model comparison, cross-validation, and final model training. The benchmark contains 80 natural images representing a variety of scenes, including people, animals, architecture, natural landscapes, indoor environments, and images with multiple salient objects.

Each RetargetMe image is accompanied by outputs generated using different retargeting methods. In this study, only the seam-carving-specific reference output is used for candidate-alpha evaluation. This provides a method-consistent target and prevents the provisional alpha label from being influenced by outputs generated through fundamentally different retargeting operations such as cropping, warping, or multi-operator resizing. The resulting labels are therefore reference-derived weighting preferences and are not treated as perceptually optimal alpha values.

Seventy-eight of the 80 RetargetMe images were successfully processed using valid original and conventional seam-carving reference pairs. Six images requiring height reduction were processed using the transpose-based procedure described in Section III. Two images were excluded because their available seam-carving references required enlargement or did not provide a valid reduction configuration supported by the implemented seam-removal pipeline.

Label-stability analysis identified 69 regime-stable images among the 78 successfully processed samples. After excluding stable Transition images and regime-unstable cases, 61 images remained for final binary modelling:

- Structural regime: 50 images
- Semantic regime: 11 images

The resulting class distribution is substantially imbalanced. Consequently, ordinary accuracy is not used as the principal model-selection metric.

2. *NRID Dataset*

The NTHU Retargeting Image Dataset (NRID) is used exclusively for independent cross-dataset evaluation. It contains 35 source images and their corresponding seam-carving reference outputs. NRID includes both portrait and landscape orientations and presents image characteristics that differ from those in RetargetMe.

All 35 NRID images were processed using the same candidate-alpha evaluation and feature-extraction procedures applied to RetargetMe. Label-stability analysis identified 32 regime-stable images. After removing stable Transition cases and regime-unstable samples, 23 images remained for final binary evaluation:

- Structural regime: 17 images
- Semantic regime: 6 images

NRID samples were not merged with RetargetMe and were not used for feature selection, model selection, hyperparameter selection, cross-validation, or classifier training. The selected RetargetMe model was evaluated once on NRID without retraining.

B. Feature Configurations

Three feature configurations were evaluated to determine whether additional handcrafted descriptors improved alpha-regime prediction.

1. Original Feature Configuration

The original configuration contained the four features proposed for lightweight prediction:

- Edge density
- Mean gradient magnitude
- Image entropy
- Saliency coverage

2. Enhanced Feature Configuration

The enhanced configuration contained seven additional descriptors:

- Standard deviation of gradient magnitude
- Laplacian variance
- Mean saliency
- Standard deviation of saliency
- Saliency concentration
- Largest salient-region proportion
- Image colourfulness

3. Combined Feature Configuration

The three feature configurations were compared on RetargetMe to determine whether the additional handcrafted descriptors improved alpha-regime discrimination. The independent NRID dataset was not examined during feature selection. Feature-set selection was based primarily on balanced accuracy and macro F1-score because the reference-derived regime labels were imbalanced. The original four-feature configuration produced the strongest balanced performance and was therefore retained for the final stable binary experiment.

C. Models Compared

Three supervised classifiers were evaluated for stable binary alpha-regime prediction:

1. Logistic Regression
2. Support Vector Machine
3. Random Forest

Class-balanced weighting was used to reduce the influence of the majority Structural category. For Logistic Regression and SVM, feature standardization was included within the modelling pipeline. Random Forest was evaluated using 500 trees, a maximum depth of 5, a minimum leaf size of 2, class-balanced weighting, and a random state of 42.

The final Logistic Regression model used class-balanced weighting, a maximum of 3000 iterations, and a random state of 42. It was selected because the original four-feature Logistic Regression configuration produced the highest balanced accuracy and macro F1-score during the RetargetMe feature-and-model comparison.

D. Cross-Validation Protocol

Repeated stratified 5-fold cross-validation was used for model development on RetargetMe. Stratification preserved the approximate proportion of Semantic and Structural samples within each fold.

The complete cross-validation procedure was repeated ten times using different fold assignments generated from a fixed random state. This produced 50 evaluation folds for each model and feature configuration.

All preprocessing operations were contained within the machine-learning pipeline. Feature standardization was therefore fitted using only the training portion of each fold and subsequently applied to its validation portion. This prevented information leakage between training and validation data.

The model and feature configuration producing the highest mean balanced accuracy were selected. Macro F1-score and ROC-AUC were used as supporting measures when interpreting model performance.

E. Evaluation Metrics

Because the final binary dataset was imbalanced, balanced accuracy was used as the primary evaluation metric:

$$\text{Balanced Accuracy} = (\text{RecallSemantic} + \text{RecallStructural}) \,/\, 2$$

Balanced accuracy gives equal importance to both categories regardless of their sample counts.

The following additional metrics were reported:

- Accuracy: proportion of all correctly classified images.
- Precision: proportion of predicted samples belonging to a class that were correctly identified.
- Recall: proportion of actual samples belonging to a class that were correctly identified.
- Macro F1-score: unweighted mean of the Semantic and Structural F1-scores.
- ROC-AUC: ability of the classifier to rank the two regimes across classification thresholds.
- Matthews correlation coefficient: correlation between predicted and actual binary labels while accounting for class imbalance.
- Confusion matrix: distribution of correct and incorrect predictions for each class.

Ordinary accuracy was reported for completeness but was not treated as the primary indicator of performance because a majority-only prediction could achieve high accuracy by ignoring the minority Semantic class.

F. Permutation Significance Test

A label-permutation test was conducted to determine whether the observed RetargetMe balanced accuracy was greater than expected under random label assignment.

The Logistic Regression pipeline was evaluated using stratified 5-fold cross-validation. The observed balanced accuracy was then compared with results obtained from 500 random permutations of the class labels. The permutation p-value was calculated as the proportion of permuted scores equal to or greater than the observed score, with the standard correction for the finite number of permutations.

The resulting p-value was 0.002, indicating that the performance of the selected Logistic Regression configuration was statistically significant at the 0.05 level under random label assignment. This test was conditional on the selected model and feature configuration and did not repeat the complete feature-and-model selection process within each permutation.

G. Independent Cross-Dataset Evaluation

After model selection, the final Logistic Regression pipeline was fitted using all 61 stable binary RetargetMe samples. The fitted model was then applied directly to the 23 stable binary NRID samples.

No NRID samples were used for feature selection, model selection, model fitting, threshold adjustment, or estimation of preprocessing parameters. The NRID evaluation therefore measures transfer from RetargetMe to an independently processed dataset.

Performance was reported using accuracy, balanced accuracy, macro F1-score, Matthews correlation coefficient, class-specific precision and recall, and a confusion matrix. The model was also compared with the majority Structural baseline on NRID.

H. Implementation Environment

The complete framework was implemented in Python using Google Colab. OpenCV was used for image loading, colour conversion, Canny edge detection, Sobel-gradient computation, spectral-residual saliency detection, seam identification, and seam removal. Scikit-image was used to calculate SSIM and PSNR.

Pandas and NumPy were used for dataset processing and numerical operations. Scikit-learn was used for feature standardization, regression, classification, repeated stratified cross-validation, permutation testing, model evaluation, and confusion-matrix generation. Trained models were stored using joblib.

A random state of 42 was used for reproducible fold generation and model initialization wherever supported.

5.Results and Discussion

Initial experiments showed that predicting the exact alpha value was unsuitable for the available feature representation. Random Forest Regression, Support Vector Regression, and the Multi-Layer Perceptron produced negative cross-validated R^2 values of −0.202, −0.612, and −0.654, respectively. A preliminary three-class classifier also failed to identify the Medium category and achieved an accuracy of 0.615, below the majority-class baseline of 0.705. These findings motivated the stable binary-regime formulation. Label-stability analysis showed that 69 of 78 RetargetMe images (88.5%) and 32 of 35 NRID images (91.4%) had near-maximum-SSIM alpha values confined to a single reference-derived regime. After excluding Transition and regime-unstable cases, 61 RetargetMe samples and 23 NRID samples remained for final binary modelling. Comparison of Logistic Regression, SVM, and Random Forest using original, enhanced, and combined feature sets showed that the original four-feature Logistic Regression model achieved the highest balanced performance.

On the 61 stable RetargetMe samples, class-balanced Logistic Regression achieved an accuracy of 0.746 ± 0.116, balanced accuracy of 0.728 ± 0.157, macro F1-score of 0.658 ± 0.140, and ROC-AUC of 0.728 ± 0.194 under repeated stratified cross-validation. Although a majority Structural classifier could obtain higher ordinary accuracy because of the 50:11 class imbalance, its balanced accuracy would remain 0.50 because it would fail to identify every Semantic image. Balanced accuracy was therefore treated as the principal performance measure. A permutation test produced an observed balanced accuracy of 0.753 compared with a mean permuted score of 0.494, resulting in $p = 0.002$ Conditional on the selected model and feature configuration, this result provides statistically significant evidence that the four lightweight image features contain information for distinguishing stable Semantic and Structural reference-derived regimes within RetargetMe.

Table I. Consolidated Experimental Results

Evaluation measure	RetargetMe	NRID
Images initially analysed	78	35
Regime-stable images	69 (88.5%)	32 (91.4%)
Stable binary samples	61	23
Semantic/Structural distribution	11/50	6/17
Accuracy	0.746 ± 0.116	0.522
Balanced accuracy	0.728 ± 0.157	0.623
Macro F1-score	0.658 ± 0.140	0.518
ROC-AUC	0.728 ± 0.194	—
MCC	—	0.226
Permutation p-value	0.002	—

The final Logistic Regression model was trained on all 61 stable RetargetMe samples and applied without retraining to the 23 stable NRID samples. It achieved a balanced accuracy of 0.623, macro F1-score of 0.518, and MCC of 0.226, exceeding the majority-class balanced accuracy of 0.50 and macro F1-score of approximately 0.425. The model correctly identified five of six Semantic images but only seven of seventeen Structural images. It therefore demonstrated strong sensitivity to the minority Semantic regime but tended to overpredict this category on NRID.

Table II. Confusion Matrix on Stable NRID Samples

True/Predicted	Semantic	Structural	Total
Semantic	5	1	6
Structural	10	7	17
Total	15	8	23

Overall, the results demonstrate that stable alpha-regime prediction is more appropriate than exact regression or three-class classification for the present dataset. Label-stability screening reduced the influence of uncertain near-equivalent alpha values, while balanced Logistic Regression produced statistically significant RetargetMe performance. The independent NRID results indicate modest cross-dataset transfer, although the lower Structural recall and modest MCC show that generalization is not yet robust. The proposed method should therefore be interpreted as a statistically supported approach to predicting reference-derived alpha regimes rather than perceptually optimal alpha values. Further improvement requires larger balanced datasets, richer semantic representations, and perceptual validation.

6. Limitations

The study has several limitations. The final training set contains only 61 stable binary RetargetMe samples, including 11 Semantic cases, while the independent NRID evaluation contains only 23 samples. This modest sample size and class imbalance limit the precision and generalizability of the performance estimates. The provisional labels were derived by measuring similarity to conventional seam-carving references; therefore, they represent reference-consistent weighting preferences rather than perceptually optimal alpha values. Transition-regime and regime-unstable images were excluded from final binary modelling, so the classifier does not cover the complete alpha continuum. Label-stability screening requires candidate-alpha evaluation during dataset construction and is therefore a provisional-label quality procedure rather than a confidence detector for unseen images. Furthermore, the classifier predicts only a broad regime and does not determine the exact alpha that should be selected within that regime. The permutation result was conditional on the selected model and feature configuration and did not account for the complete model-selection process. Finally, the NRID results showed limited Structural recall and only modest cross-dataset transfer, while quantitative evaluation was restricted to image reduction and enlargement was demonstrated only qualitatively.

7. Conclusion and Future Work

This study introduced a framework for predicting stable, reference-derived alpha-weighting regimes in hybrid seam carving. The contributions are threefold. First, it introduces a label-stability procedure that examines near-maximum-SSIM candidates and prevents uncertain provisional labels from being treated as reliable. Second, it reformulates alpha selection as the prediction of stable Semantic and Structural regimes using four lightweight image features. Third, it provides cross-validated, permutation-based, and independent cross-dataset evaluation using RetargetMe and NRID.

Label-stability analysis identified consistent reference-derived regimes for 88.5% of RetargetMe images and 91.4% of NRID images. On stable RetargetMe samples, class-balanced Logistic Regression achieved a balanced accuracy of 0.728 ± 0.157, macro F1-score of 0.658 ± 0.140, and ROC-AUC of 0.728 ± 0.194. Conditional on the selected model and feature configuration, the permutation result of $p = 0.002$ provided statistically significant evidence of predictive information within RetargetMe. Independent NRID evaluation produced a balanced accuracy of 0.623, macro F1-score of 0.518, and MCC of 0.226, indicating measurable but limited cross-dataset transfer.

The findings indicate that lightweight image descriptors contain information for predicting broad reference-derived weighting regimes. However, these regimes represent similarity to conventional seam-carving references and should not be interpreted as perceptually optimal alpha values. The framework also does not determine the exact alpha to use within a predicted regime. Future work will incorporate perceptual or expert-derived labels, expand and balance the datasets, investigate learned semantic representations, model Transition cases probabilistically, develop confidence estimation for unseen images, determine appropriate within-regime alpha values, and extend quantitative evaluation to image enlargement.

References

1. A. K. Sanil and V. Mini G., "A Survey of Developments and Challenges in Image Retargeting Using AI-Powered Seam Carving," Journal of Advancement in Software Engineering and Testing, vol. 9, no. 1, pp. 1–13, 2025.
2. S. Avidan and A. Shamir, "Seam Carving for Content-Aware Image Resizing," ACM Transactions on Graphics, vol. 26, no. 3, Art. no. 10, 2007, doi: 10.1145/1276377.1276390.
3. A. K. Sanil and V. Mini G., "Improving Seam Carving with Feature-Based Saliency Estimation: A Hybrid Model for Image Retargeting," in Proceedings of the International Conference on Future of Connected Systems—IoT, AI and 5G (FoCS 2025), pp. 99–103, 2025, ISBN: 978-93-6665-684-7.
4. B. Asheghi, P. Salehpour, A. M. Khiavi, and M. Hashemzadeh, "A Comprehensive Review on Content-Aware Image Retargeting: From Classical to State-of-the-Art Methods," Signal Processing, vol. 195, Art. no. 108496, 2022, doi: 10.1016/j.sigpro.2022.108496.
5. M. Shi, G. Peng, L. Yang, and D. Xu, "Optimal Bi-Directional Seam Carving for Content-Aware Image Resizing," in Entertainment for Education: Digital Techniques and Systems, Lecture Notes in Computer Science, vol. 6249. Berlin, Heidelberg: Springer, 2010, pp. 456–467.
6. J. Chen, L. Miao, and X. Liu, "Image Re-Targeting with Balanced Energy Map and Foreground Constraint," World Wide Web, vol. 14, no. 3, pp. 281–292, 2011, doi: 10.1007/s11280-010-0105-1.
7. E. Song, M. Lee, and S. Lee, "CarvingNet: Content-Guided Seam Carving Using Deep Convolution Neural Network," IEEE Access, vol. 7, pp. 284–292, 2019, doi: 10.1109/ACCESS.2018.2885347.
8. D. Cho, J. Park, T.-H. Oh, Y.-W. Tai, and I. S. Kweon, "Weakly- and Self-Supervised Learning for Content-Aware Deep Image Retargeting," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017, pp. 4558–4567.