ACS Sustainable Resource Management

pubs.acs.org/acssrm

Article

Open Access

This article is licensed under CC-BY 4.0

Physically Informed Multi-Magnification Learning for Scalable, Non-Destructive Assessment of Polypropylene Degradation in Recycled Materials

Published as part of ACS Sustainable Resource Management special issue "Artificial Intelligence (AI) Tools: A New Frontier for Sustainability".

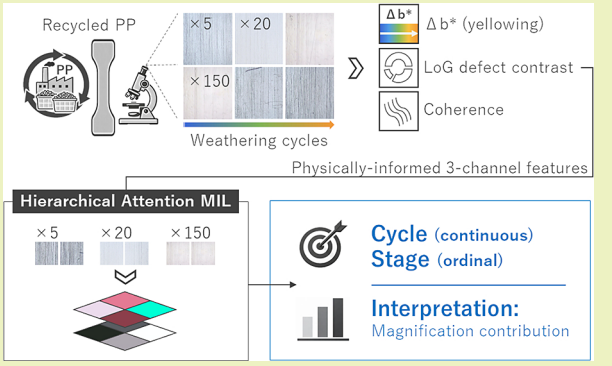
Kazuki Hammura, Kiyotaka Hitomi, Kenji Nagata, and Masanobu Naito*

Cite This: ACS Sustain. Resour. Manag. 2026, 3, 3048−3057

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Reliable quality control of recycled polypropylene (PP) is essential for circular plastics management, but conventional degradation tests are often destructive or unsuitable for large-scale screening. This study introduces a physically informed, multi-magnification learning framework for non-destructive and scalable assessment of PP degradation from optical micrographs. A three-channel feature representation captures baseline-referenced yellowing, defect-associated contrast, and microstructural anisotropy. To address multi-scale variability, magnification-aware image patches are integrated via a hierarchical attention-based multiple instance learning model. The model jointly predicts continuous weathering cycles and ordinal degradation stages, enabling quantitative evaluation of heterogeneous material conditions. On the validation set, the approach achieved a balanced accuracy of 0.747, a stage-decoded mean absolute error of 5.6 cycles, and a quadratic weighted kappa of 0.908, with an accuracy within ±1 stage of 0.97. Notably, the model identified 10-cycle-related image signals despite the absence of a clear increase in attenuated total reflection infrared carbonyl index under the same conditions. These results demonstrate the feasibility of AI-enabled optical micrograph analysis as a rapid and non-destructive screening approach for weathering-related surface-state assessment under the controlled post-industrial recycled (PIR) PP conditions examined in this study. The framework can therefore serve as a basis for high-throughput image-based screening of weathering-related surface states, although extension to broader recycled PP streams and practical quality attributes requires further validation.



KEYWORDS: *polypropylene, optical micrographs, multiple instance learning, hierarchical attention, weathering, non-destructive evaluation, circular economy, recycled polymers*

■ INTRODUCTION

The transition toward a circular economy requires reliable strategies for the reuse and recycling of polymeric materials. Polypropylene (PP), one of the most widely used plastics in packaging, automotive components, and consumer products, is increasingly incorporated into recycled material streams.[1] However, during service and recycling processes, PP is exposed to environmental stressors such as ultraviolet radiation, heat, and mechanical loading, leading to degradation and heterogeneous material states.[2−4] This variability poses a critical challenge for ensuring the quality, safety, and performance of recycled PP in practical applications. A key bottleneck in sustainable polymer utilization is the lack of rapid, scalable, and non-destructive methods for evaluating degradation states.

Conventional techniques, including mechanical testing and chemical analysis, are often destructive, labor-intensive, and unsuitable for high-throughput screening. Consequently, they are difficult to integrate into real-world recycling workflows, where large volumes of materials must be assessed efficiently. In addition, polymer degradation progresses continuously and heterogeneously, making discrete classification based on limited

Received: May 7, 2026
Revised: August 15, 2026
Accepted: August 17, 2026
Published: September 2, 2026



ACS Publications

© 2026 The Authors. Published by American Chemical Society

Downloaded from pubs.acs.org/asrmcd/article-pdf/3/9/3048/6819859 6/rm-2026-003077.pdf by NATL INST FOR MATLS SCIENCE (NIMS), Masanobu Naito on 13 October 2026

measurements insufficient for reliable quality control. Because photo-oxidative degradation of PP initiates at the surface and propagates inward, surface morphology contains valuable information about environmental exposure history. Optical micrographs provide a cost-effective and widely accessible means of capturing such surface features, including discoloration and defect formation in the skin layer.[5,6] Notably, early-stage degradation may not be clearly reflected in conventional chemical indices, highlighting the potential of image-based approaches for detecting subtle surface changes. However, the interpretation of optical micrographs is inherently challenging due to spatial heterogeneity and strong dependence on magnification, while degradation itself is a multi-scale phenomenon.

Recent advances in artificial intelligence (AI) and machine learning offer new opportunities to extract meaningful information from complex image data. In particular, multiple instance learning (MIL) enables the aggregation of patch-level information into group-level predictions under weak supervision and has demonstrated strong performance in domains such as computational pathology.[7−9] Nevertheless, the application of AI to optical micrographs of polymer materials remains limited, especially in the context of multi-magnification data and heterogeneous degradation states. Furthermore, approaches that explicitly incorporate physically meaningful representations while addressing multi-scale variability are still underdeveloped.

This study introduces a physically informed, multi-magnification learning framework for the non-destructive assessment of PP weathering from optical micrographs. The method introduces a three-channel feature representation to capture color change, defect-associated contrast, and microstructural anisotropy. To address magnification-dependent variability, image patches are extracted using a magnification-aware strategy and integrated via a hierarchical attention-based MIL model. In addition, degradation is modeled as an ordinal variable to enable consistent evaluation across stages. By combining physically grounded feature design with scalable AI-based analysis, the proposed approach provides a proof-of-concept foundation for high-throughput, non-destructive screening of weathering-related surface states in controlled PIR-derived PP systems.

■ MATERIALS AND METHODS

Materials, Specimen Preparation, and Accelerated Weathering

To introduce controlled material variability, a post-industrial recycled PP (rPP) and two types of virgin PP (vPP) were used in this study. The rPP was supplied by Nakasada Resin Co., Ltd. and was produced from post-industrial PP scrap generated during the manufacture of medical-use products. The supplied rPP was relatively clean and visually homogeneous. The vPP grades were Prime Polypro E100-GPL (Prime Polymer Co., Ltd.) and Novatec EA9 (Japan Polypropylene Corporation). The tested compositions consisted of neat rPP and rPP/vPP blends. To introduce controlled composition variability, the rPP was separately compounded with each vPP grade at rPP:vPP mass ratios of 75:25, 50:50, and 25:75, while neat rPP corresponded to rPP:vPP = 100:0. Dumbbell-shaped specimens (Japanese Industrial Standards (JIS) K7139 A1) were fabricated using a twin-screw compounder, followed by extrusion and injection molding. Fifteen specimens were prepared for each compounded composition and twelve for neat rPP.

Accelerated weathering was conducted using a xenon-arc weathering tester under JIS K7350-2 cycle No. 10 conditions (1 cycle: 2 h light and water spray). Accelerated weathering was performed for 10, 30, 50, 70, and 100 cycles, and optical micrographs were acquired before and after weathering. The before-weathering images were assigned to the 0-cycle state. The 10-cycle condition was applied only to the compounded specimens.

The rPP used in this study was a post-industrial recycled PP supplied as a recycled feedstock, rather than a model recycled material prepared by artificially degrading virgin PP under controlled laboratory conditions. Therefore, its prior thermal, processing, additive, and oxidation histories were not fully controlled in this study. The initial state for modeling was operationally defined as the molded, before-weathering state of each material system, which was assigned to 0 cycles. Accordingly, the weathering-cycle labels used for model training should be interpreted as the imposed xenon-arc weathering history relative to the molded 0-cycle state of each material system, rather than as an absolute measure of the total degradation history of the recycled feedstock.

Optical Microscopy, Data Organization, and Prediction Task

Optical micrographs were acquired using a digital microscope, VK-X3000, Keyence Corporation, at magnifications of 5×, 10×, 20×, 50×, and 150×. In total, 1020 images were collected across all conditions. Representative optical micrographs used as model inputs are shown in Figure 1. These images were organized into 68 state-specific groups, hereafter referred to as MIL bags. Each bag corresponded to a unique composition-weathering-cycle-condition-state combination and contained 15 images, consisting of 3 replicate specimens acquired at 5 magnifications.

Representative optical micrographs of dumbbell-shaped PP specimens before (0 cycles; left) and after xenon-arc weathering (right). Weathering induces progressive yellowing and texture changes. Voids are observed on recycled PP surfaces, and vertical striations are consistently present due to the molding process. The scale bars are 500, 50, 50, and 5 μm for 5×, 20×, 50×, and 150× images respectively.

The task was formulated as a group-level prediction problem, where each bag was assigned a degradation label corresponding to the weathering cycle. Thus, the terms "weathering cycle" and "degradation stage" in the prediction task denote the imposed accelerated-weathering stage relative to the molded 0-cycle reference state and do not represent an absolute cumulative degradation history. The model was trained to predict both a continuous cycle value and an ordinal degradation stage. This formulation enables quantitative evaluation of heterogeneous material states within the present controlled weathering dataset and provides a basis for future high-throughput image-based screening.

Three-Channel Physically Informed Feature Image

Each RGB micrograph was transformed into a three-channel feature image designed to encode physically meaningful characteristics of PP degradation: baseline-referenced color change, defect-associated LoG contrast, and microstructural anisotropy. Representative examples of the RGB input and the resulting three feature channels are shown in Figure 2. This representation was designed to enhance interpretability while enabling robust learning from heterogeneous micrographs.

RGB micrographs were transformed into three channels representing Δb^*, baseline-referenced color change; multi-scale LoG defect contrast; and structure-tensor coherence, representing microstructural anisotropy. Each row shows one material, weathering, and magnification condition.

Channel 0: Baseline-Referenced Color Change. To quantify yellowing associated with photo-oxidative degradation, each RGB image was converted to the CIELAB color space and the blue–yellow component was extracted as $b = \mathrm{CIELAB}[:, :, 2] - 128$. The b channel was robustly normalized using the 1st and 99th percentiles: $b_{01}(x) = \mathrm{clip}\left(\frac{b(x)-p_1}{p_{99}-p_1}, 0, 1\right)$, where p_1 and p_{99} denote the 1st and 99th percentiles of b, respectively.

Downloaded from pubs.acs.org/asrmcd/article-pdf/3/9/3048/6819859/m-2026-003077.pdf by NATL INST FOR MATLS SCIENCE (NIMS), Masanobu Naito on 13 October 2026

https://doi.org/10.1021/acssusresmgt.6c00307
ACS Sustain. Resour. Manag. 2026, 3, 3048−3057

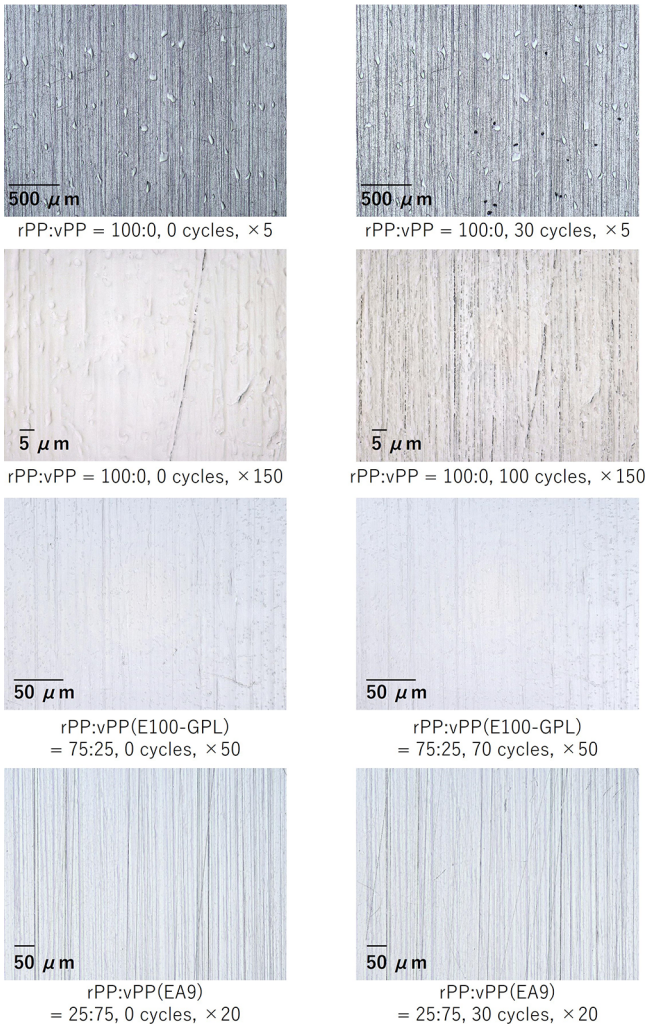


Figure 1. Representative optical micrographs.

To account for baseline differences among material compositions and magnifications, a baseline value b_0 was subtracted. The baseline b_0 was defined as the median b_{01} value of a corresponding minimum-cycle image, 0 cycles in this study, with the same magnification and rPP/vPP ratio. If no corresponding reference image was available, b_0 was set to the median b_{01} value of the current image.

The relative yellow-blue shift was then calculated as $\Delta b^*(x) = b_{01}(x) - b_0$. The resulting Δb^* values were symmetrically clipped to $[-\delta, +\delta]$, with $\delta = 1.0$, linearly rescaled to $[0, 1]$, and quantized to 8-bit unsigned integers.

This baseline-referenced design is most directly applicable when a composition- and magnification-matched 0-cycle reference image is available, as in the present post-industrial recycled (PIR) PP material set. The Δb^* channel was not intended to fully remove arbitrary initial coloration caused by pigments, fillers, surface dirt, or contamination in highly heterogeneous post-consumer recycled streams.

Channel 1: Multi-Scale LoG Defect Contrast. To capture defect-associated contrast patterns, including ring-like structures on the specimen surface, a multi-scale Laplacian-of-Gaussian (LoG) response was computed. For each pixel x, the maximum absolute response across three scales was retained: $R(x) = \max_i |[(\nabla^2 G_{\sigma_i}) \times I](x)|$, where $G_{\sigma i}$ is a Gaussian kernel with standard deviation σ_i, ∇^2 denotes the Laplacian operator, and * denotes convolution. The base scale was defined as $\sigma_0 = \max\left[0.6, 3.0 \times \frac{M}{M_{\text{ref}}}\right]$, where M is the image magnification and $M_{\text{ref}} = 150$. The three LoG scales were then set to $\sigma_i \in \{0.8\sigma_0, 1.0\sigma_0, 1.3\sigma_0\}$. The resulting response map was normalized to the range $[0, 255]$ and used as the second channel of the feature image.

Channel 2: Structure-Tensor Coherence. To quantify microstructural anisotropy, structure-tensor coherence was computed from image gradients. Image gradients g_x and g_y were first obtained using Sobel filters. The structure tensor components were calculated as $A_{xx} = ⊠g_x^2⊠$, $A_{xy} = ⊠g_x g_y⊠$, $A_{yy} = ⊠g_y^2⊠$, where ⊠·⊠ denotes Gaussian smoothing. The smoothing scale was adjusted according to magnification as $\sigma = \text{clip}\left(0.8 \times \frac{M_{\text{ref}}}{M}, 0.6, 6.0\right)$. Let $\lambda_1 \geq \lambda_2$ be the eigenvalues of the local structure tensor.

The coherence was computed as $C(x) = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2 + \varepsilon}$, where ε is a small constant for numerical stability. The coherence values were clipped to $[0, 1]$ and quantized to 8-bit unsigned integers. The three channels, Δb^*, multi-scale LoG defect contrast, and structure-tensor coherence, were stacked to form a physically informed feature image. This representation was designed to reduce reliance on purely data-driven feature extraction while preserving degradation-related cues that are interpretable in terms of color change, defect formation, and anisotropic surface texture.

Multi-Magnification Bag Construction (MIL Input)

The problem was formulated within a MIL framework, where each composition-weathering-cycle group constituted one bag. Here, a group denotes the unit used for model splitting and prediction, rather than an individual physical dumbbell specimen. Micrographs acquired at multiple magnifications for each group were first converted into three-channel feature images as described in the section "Three-channel physically informed feature image." Multiple square patches were then extracted from these feature images and aggregated into a single MIL bag.

Patch extraction used a magnification-dependent nominal crop-size rule, $C_M^{\text{nom}} = \text{round}\left(C_{\text{ref}} \frac{M_{\text{ref}}}{M}\right)$, with $M_{\text{ref}} = 150$ and $C_{\text{ref}} = 512$ pixels. The nominal size was clipped to 224–2048 pixels, adjusted to an odd integer, and then further limited by the available non-border image region. Because the maximum image size was 1020 pixels and an 8% border was excluded on each side, the effective maximum crop size was 855 pixels. The resulting effective crop sizes were 855 pixels for 5×–50× images and 511 pixels for 150× images. All crops were resized to 224 × 224 pixels.

From each selected micrograph, eight square patches were sampled. During training, patch locations were sampled randomly from valid image regions. During validation, patch locations were sampled deterministically using a stable seed to ensure reproducibility. The resulting three-channel patches were converted to $[0, 1]$ floating-point tensors and normalized using the ImageNet mean and standard deviation. During training, instance dropout was applied with $p_{\text{drop}} = 0.05$, and each training bag was capped at 24 instances.

This bag construction strategy allowed heterogeneous information from multiple magnifications to be integrated while maintaining computational efficiency.

ConvNeXt Backbone and Hierarchical Attention MIL

The proposed model consists of a patch encoder, a hierarchical attention-based MIL aggregator, a regression head, and a supporting rank-consistent ordinal regression (CORAL) head.[10,11] Given a bag of patches and their corresponding magnification indices, each patch was first mapped to a fixed-dimensional embedding using a ConvNeXt-Tiny[12] backbone pretrained on ImageNet. The classification head of the backbone was removed, and global average pooling was applied to the final feature maps to obtain a patch embedding with dimensionality $d = 768$.

Downloaded from pubs.acs.org/asrmcd/article-pdf/3/9/3048/6819859/m-2026-003077.pdf by NATL INST FOR MATLS SCIENCE (NIMS), Masanobu Naito on 13 October 2026

https://doi.org/10.1021/acssusresmgt.6c00307
ACS Sustain. Resour. Manag. 2026, 3, 3048−3057

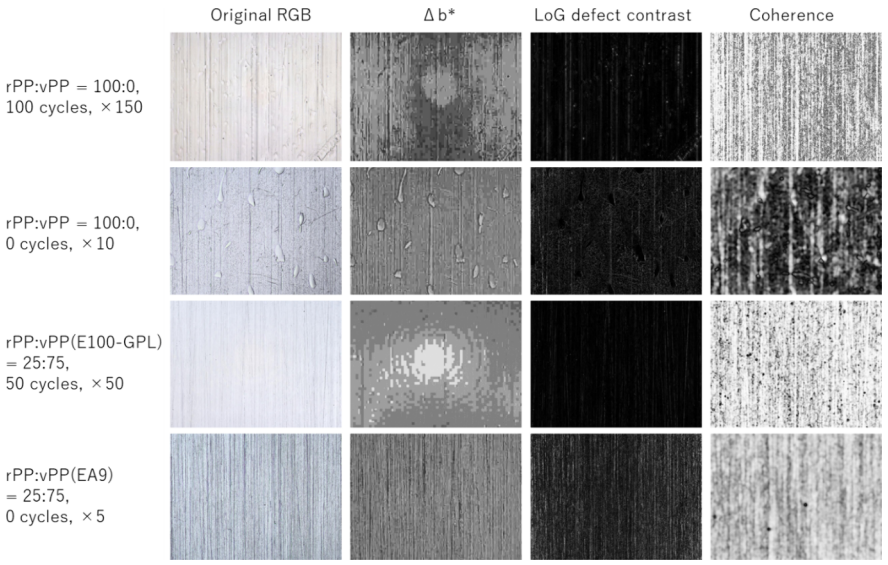


Figure 2. Representative three-channel feature images.

To integrate multi-magnification information, a hierarchical attention pooling mechanism was employed. In the first stage, patch embeddings were grouped by magnification and aggregated within each magnification using a multilayer perceptron (MLP) -based attention module consisting of Linear-tanh-Linear layers with softmax temperature τ. This produced magnification-level representations. A learnable magnification embedding was then concatenated to each magnification-level representation to explicitly encode magnification identity.

In the second stage, the set of magnification-level representations was aggregated using another attention pooling module to obtain a single group-level embedding. This hierarchical design allows the model to learn both which patches are informative within each magnification and which magnifications contribute most strongly to the final prediction. The resulting group-level embedding was used for continuous weathering-cycle regression and ordinal degradation-stage prediction. The overall architecture is summarized in Figure 3. A detailed implementation diagram, including within- and across-magnification attention pooling and magnification embeddings, is provided in Figure S1 in the Supporting Information.

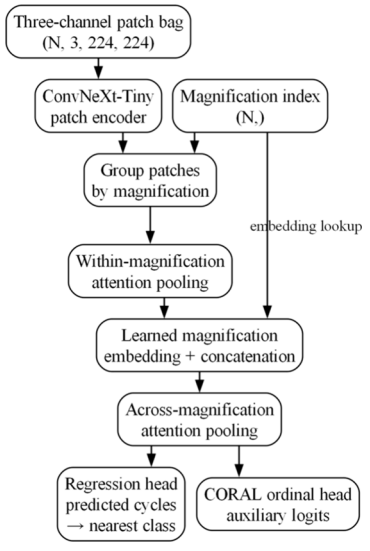


Figure 3. Overview of the proposed hierarchical attention-based MIL architecture.

Multi-magnification patches are encoded by a ConvNeXt-Tiny backbone, grouped by magnification, and aggregated in two stages using MLP-based attention (Linear-tanh-Linear). The resulting group-level representation is used for both continuous cycle regression and ordinal stage prediction via a CORAL head.

Training and Evaluation

The network was trained to jointly predict a continuous weathering-cycle value and an ordinal degradation stage. For each group-level bag with class index y, corresponding to cycle value c_y, a weighted sum of regression and ordinal losses was minimized.

The regression loss L_{reg} was the Huber, or smooth L1, loss between the predicted cycle value and c_y, with a weight of 0.8. The ordinal loss L_{CORAL} followed the CORAL formulation. The model outputs $K - 1$ threshold logits for K ordered stages, and binary ordinal targets were defined as $t_k = \mathbf{1}[y > k]$. Binary cross-entropy with logits was used without label smoothing, with a weight of 0.8. In the present run, the CORAL bias was initialized from cumulative class priors, and class-frequency-based positive weights were used in the binary cross-entropy term.

To encourage ordinal consistency, a monotonicity penalty was added. For ordinal probabilities $q_k = \text{sigmoid}(\text{logit}_k)$, the monotonicity penalty was defined as $L_{\text{mono}} = \text{mean}(\text{ReLU}(q_{k+1} - q_k))$, with a weight of $\lambda_{\text{mono}} = 0.1$. Training was performed using AdamW with a learning rate of 2×10^{-4} and a weight decay of 5×10^{-4}. A reduced learning rate, multiplied by 0.1, was used for the pretrained backbone, and no weight decay was applied to bias, normalization, and embedding parameters. The ConvNeXt backbone was frozen for the first three epochs and then unfrozen.

Within the attention blocks, the softmax temperature τ was linearly annealed from 1.2 to 0.8 over the first eight epochs. Gaussian score noise was annealed from 0.15 to 0.03 over the first eight epochs, and the top-k attention fraction was annealed from 1.0 to 0.9 over the first 10 epochs.

To mitigate class imbalance, training groups were sampled at the group level using a balanced group sampler. Minority classes were oversampled to the maximum class count multiplied by an epoch multiplier of 6, while classes were interleaved to avoid long consecutive runs from the same class.

Downloaded from pubs.acs.org/asrmcd/article-pdf/3/9/3048/6819859/m-2026-003077.pdf by NATL INST FOR MATLS SCIENCE (NIMS), Masanobu Naito on 13 October 2026

https://doi.org/10.1021/acssusresmgt.6c00307
ACS Sustain. Resour. Manag. 2026, 3, 3048−3057

Model performance was evaluated at the group level, where each group corresponded to one MIL bag. The dataset was split by group so that no group appeared in both the training and validation sets. A stratified group split was used to ensure that all six degradation stages, 0, 10, 30, 50, 70, and 100 cycles, were represented in both sets.

The training set contained 17, 3, 4, 3, 3, and 4 groups for the 0, 10, 30, 50, 70, and 100-cycle classes, respectively. The validation set contained 17, 3, 3, 4, 4, and 3 groups for the same classes. All reported metrics were calculated on the validation set using the best checkpoint selected at epoch 8, where the validation quadratic weighted kappa (QWK) and balanced accuracy were highest in this run.

Stage predictions were obtained by mapping the regression output to the nearest stage value among the six weathering stages. Accordingly, the mean absolute error (MAE) reported in this study denotes a stage-decoded MAE expressed in cycle units. Because the model predicts ordered degradation stages, balanced accuracy, stage-decoded MAE, QWK, and within ±1-stage accuracy were reported.

ATR-IR Spectroscopy and Carbonyl-Index Calculation

Attenuated total reflection infrared (ATR-IR) spectroscopy was performed to compare the image-based degradation response with a conventional chemical descriptor of PP oxidation. Spectra were acquired from the weathered specimen surface using an IRXross Fourier transform infrared spectrometer (Shimadzu Corporation) equipped with a GladiATR diamond ATR accessory (PIKE Technologies, Inc.) operated at an incident angle of 45°. Spectra were recorded over 4000–450 cm^{-1} at a resolution of 4 cm^{-1} by co-adding 32 scans after background acquisition under ambient conditions. The spectra were baseline-corrected before index calculation. The carbonyl index was defined as Carbonyl index = $\frac{A_{1715}}{A_{1456}}$, where A_{1715} and A_{1456} denote the absorbance values at 1715 and 1456 cm^{-1}, respectively. The band at 1715 cm^{-1} was used as the carbonyl-related signal, whereas the band at 1456 cm^{-1} was used as an internal reference for PP. The carbonyl index was calculated for each measured spectrum, and summary statistics were obtained across replicates for each weathering cycle.

■ RESULTS AND DISCUSSION

Optical Signatures of PP Weathering

Representative optical micrographs at different degradation stages are shown in Figure 1. The surfaces of all dumbbell-shaped specimens consistently showed vertical striations attributable to the molding process. With increasing weathering cycles, progressive yellowing, enhanced ring-like contrast associated with surface defects, and changes in microstructural anisotropy became visually apparent across magnifications.

Model Performance and Magnification-Level Attention

The quantitative performance of the proposed method on the validation set is shown below. The model achieved a balanced accuracy of 0.747 and a stage-decoded MAE of 5.6 cycles. The QWK reached 0.908. This indicates strong ordinal agreement between predicted and true degradation stages while accounting for their ordinal nature. The within ±1-stage accuracy was 0.97, showing that nearly all predictions fell within one adjacent stage of the ground truth. These results suggest that the model captures ordinal stage progression of degradation. In addition to predictive performance, attention diagnostics revealed stage-dependent magnification usage. For the 0- and 10-cycle groups, attention was concentrated on higher-magnification micrographs, particularly 150× and 50×, suggesting that discrimination at the earliest boundary relied mainly on fine surface features (Supplementary Figure S2).

Checkpoint Selection and Validation Behavior

Model selection was performed using validation performance across the heterogeneous material conditions represented in the validation set. Specifically, the best-performing model was selected using the highest QWK on the validation set, which is appropriate for evaluating ordinal predictions of degradation stages. Training dynamics are summarized in Figure S3. The validation QWK reached its maximum at epoch 8, at which point the model also achieved the lowest stage-decoded MAE of 5.6 cycles and the highest balanced accuracy (0.747). Accordingly, the checkpoint at epoch 8 was selected as the final model for evaluation. After epoch 8, validation metrics exhibited fluctuations without further improvement, indicating the onset of overfitting. Early stopping was therefore applied, with training terminated at epoch 13 based on the lack of improvement in validation QWK. Although the evaluation metrics were concordant at the selected checkpoint, the performance estimates are specific to this single validation split. Future work will include repeated experiments with different data splits and random seeds to assess statistical stability. This selection strategy provided a consistent basis for choosing the final checkpoint under the present experimental conditions.

Confusion Matrix Analysis

The confusion matrix of the best-performing model on the six-stage validation set is shown in Figure 4. Out of 34 validation groups, 24 were classified correctly. Nine of the 10 misclassifications were confined to adjacent stages, consistent with the ordinal nature of the task and the high within ±1-stage accuracy. The dominant ambiguity occurred at the earliest boundary: all three 10-cycle validation groups were classified as 10 cycles rather than 0 cycles, whereas 6 unweathered groups were assigned to the 10-cycle class. One non-adjacent error was observed for a 70-cycle group predicted as 10 cycles. These results suggest that the model detected a 10-cycle-related signal at the group level, although the 0–10 cycle boundary remained the main source of uncertainty.

Row-wise normalized confusion matrix for the six-stage validation set consisting of 0, 10, 30, 50, 70, and 100 weathering cycles. Raw counts are shown in each cell, and color indicates the fraction of each true class. Most misclassifications were confined to adjacent stages, whereas the principal ambiguity occurred at the 0–10 cycle boundary.

Early-Stage Comparison with ATR-IR Spectroscopy and Carbonyl Index

To relate the model behavior at the 0–10 cycle boundary to a conventional chemical descriptor and to place the early-stage response in the context of the full weathering trajectory, ATR-IR spectroscopy was examined across the weathering series. Representative ATR-IR spectra are shown in Figure 5a. The spectra showed no clear development of a carbonyl-related absorption feature around 1715 cm^{-1} at the early weathering stages under the present measurement conditions, whereas a more pronounced change was observed at the most advanced weathering stage. The carbonyl index was defined as $\frac{A_{1715}}{A_{1456}}$. As shown in Figure 5b, the carbonyl index remained nearly unchanged from 0 to 70 cycles, showing only small fluctuations without a clear monotonic increase across the 6 compositions. In contrast, a pronounced increase was observed at 100 cycles, where the mean carbonyl index rose markedly relative to the

Downloaded from pubs.acs.org/asrmcd/article-pdf/3/9/3048/6819596/rm-2026-003077.pdf by NATL INST FOR MATLS SCIENCE (NIMS), Masanobu Naito on 13 October 2026

https://doi.org/10.1021/acssusresmgt.6c00307
ACS Sustain. Resour. Manag. 2026, 3, 3048−3057

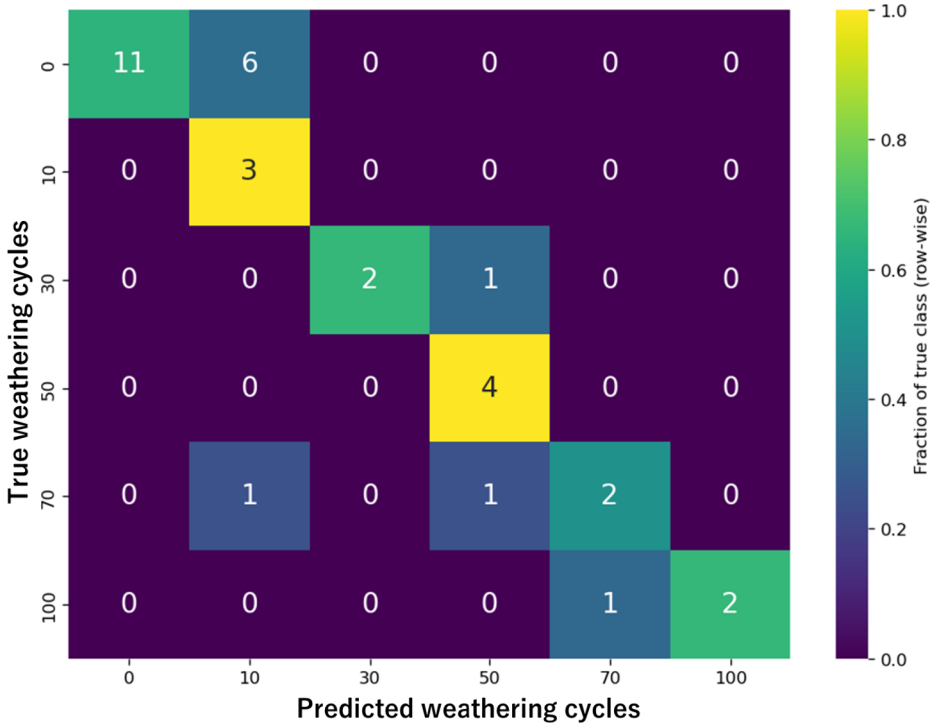


Figure 4. Validation confusion matrix for micrograph-based MIL prediction of PP weathering stage.

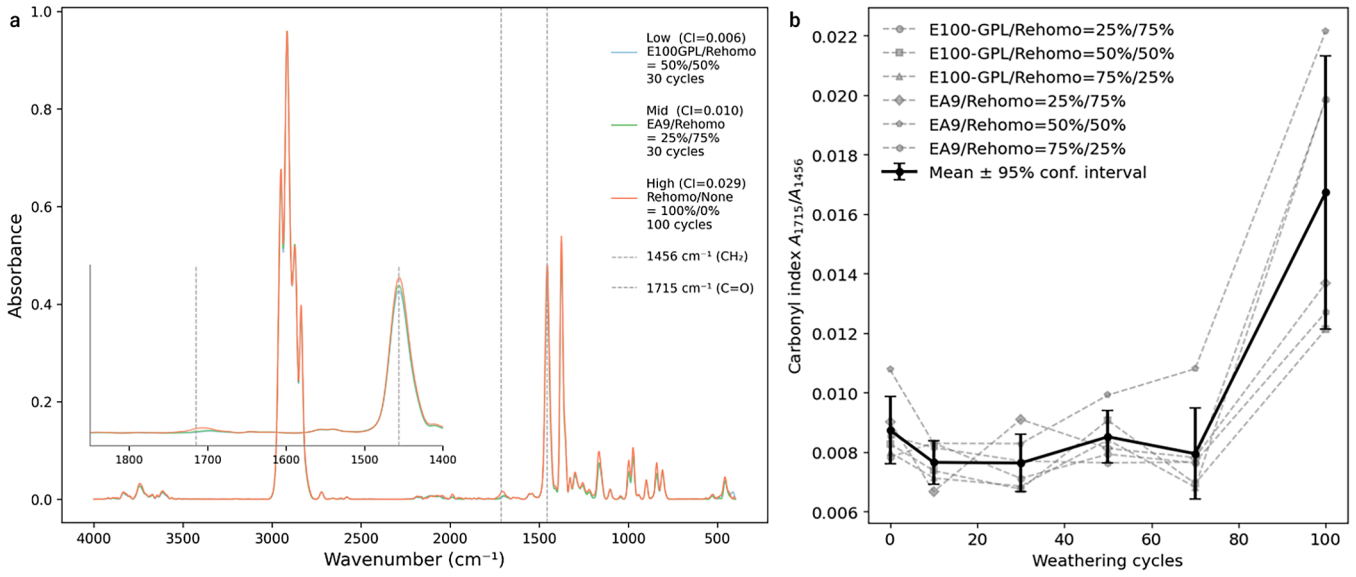


Figure 5. (a, b) ATR-IR spectral analysis and carbonyl-index evolution of weathered PP specimens.

earlier stages. Thus, under the present measurement conditions, the 10-cycle stage was not accompanied by a clear or systematic increase in carbonyl index, and substantial carbonyl growth became evident only at the most advanced weathering stage.

(a) Representative baseline-corrected ATR-IR spectra selected to span low, intermediate, and high carbonyl-index values. The inset expands the 1850–1400 cm^{-1} region. Dashed vertical lines indicate the PP CH_2 reference band at 1456 cm^{-1} and the carbonyl-related C=O band at ca. 1715 cm^{-1} used for carbonyl-index calculation. The material composition, weathering cycle, and carbonyl index for each spectrum are indicated in the panel legend. (b) Carbonyl index, defined as $\frac{A_{1715}}{A_{1456}}$, plotted as a function of weathering cycles for each rPP/vPP composition. Gray dashed lines indicate individual compositions, and the black line with error bars shows the mean ± 95% confidence interval. The carbonyl index remained nearly unchanged from 0 to 70 cycles and increased markedly at 100 cycles.

In contrast, the micrograph-based MIL model assigned all three held-out 10-cycle validation groups to the 10-cycle class

https://doi.org/10.1021/acssusresmgt.6c00307
ACS Sustain. Resour. Manag. 2026, 3, 3048−3057

Downloaded from pubs.acs.org/asrmcd/article-pdf/3/9/3048/6819856/rm-2026-003077.pdf by NATL INST FOR MATLS SCIENCE (NIMS), Masanobu Naito on 13 October 2026

rather than the unweathered class, although the six 0-cycle groups assigned to the 10-cycle class indicate that this earliest boundary remained uncertain. These observations suggest that optical micrographs may be sensitive to early surface-level changes, whereas ATR-IR more clearly reflects later-stage oxidative evolution; thus, the two methods provide complementary information on weathering progression.

Composition-Dependent Feature Variation and Prediction Behavior

The rPP/vPP blends were introduced to include controlled material variability in the present PIR-derived PP dataset. To clarify how this variability affected both the input representation and the model output, a composition-dependent post hoc analysis was conducted using the same feature images, final checkpoint, and validation split as those used for the main validation results. For the feature-level analysis, image-level means of the three physically informed channels, namely Δb^*, LoG defect contrast, and structure-tensor coherence, were aggregated at the MIL-bag level and then summarized by rPP content, vPP grade, weathering cycle, and magnification. In addition to the baseline-subtracted feature changes shown in Figure 6a, the raw composition-dependent feature-channel means, feature trajectories stratified by rPP content and vPP grade, and feature-value distributions are provided in Figures S4–S8. The corresponding numerical summaries are listed in Tables S5 and S7. This normalization was used only for analysis and visualization, and no model retraining was performed. Because these subgroup analyses were based on a limited number of validation bags, they are interpreted as descriptive diagnostics of composition-dependent behavior rather than as independently powered estimates of generalization performance for each composition.

The three feature channels exhibited both composition-dependent baseline variations and weathering-dependent changes. At the 0-cycle state, the mean LoG defect-contrast

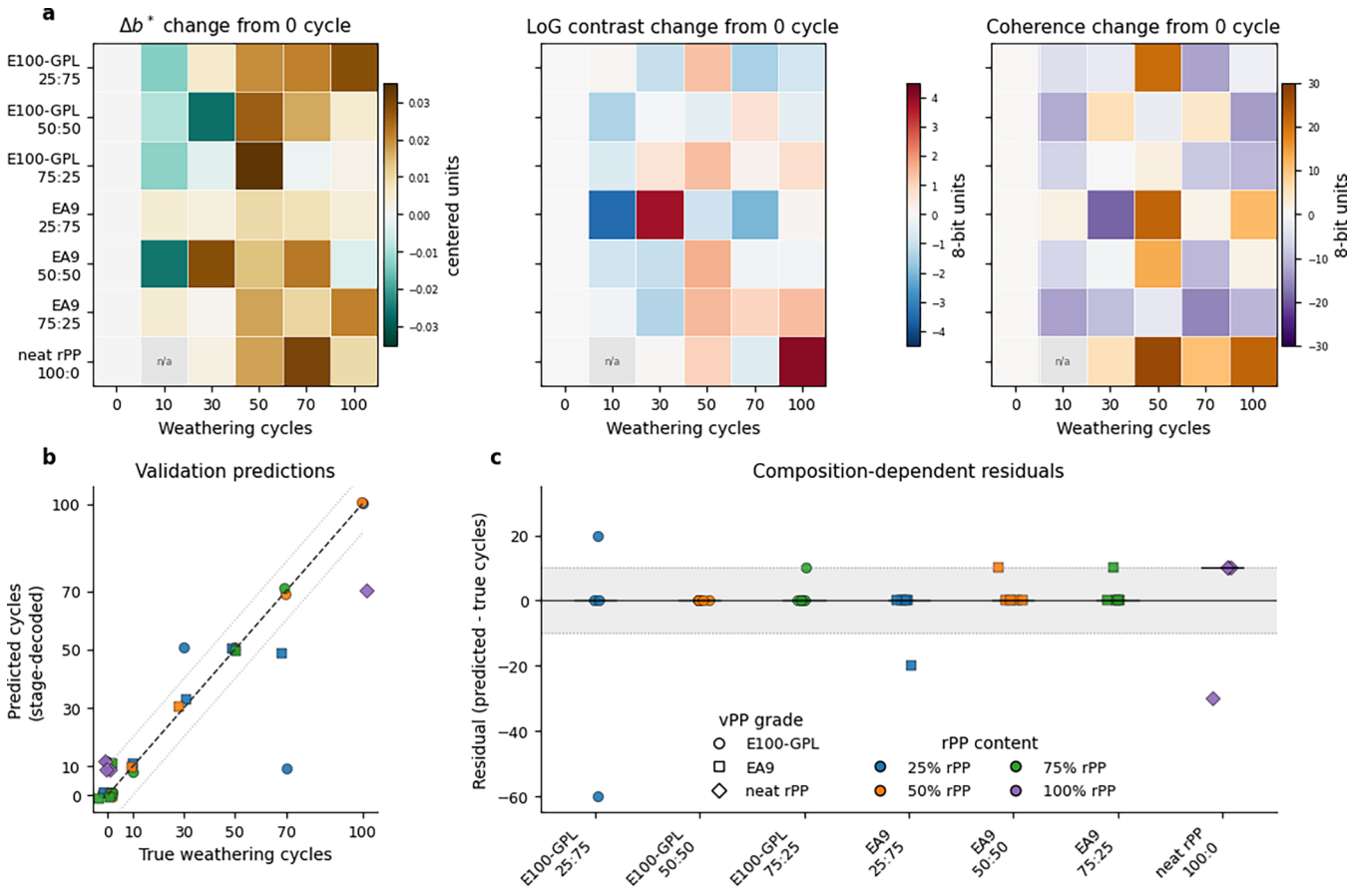


Figure 6. Composition-dependent feature variation and validation prediction behavior. (a) Composition-dependent changes in the three physically informed feature channels from the corresponding 0-cycle state. Image-level means of Δb^*, LoG defect contrast, and structure-tensor coherence were first aggregated at the MIL-bag level and then averaged for each composition–weathering-cycle condition. The value at 0 cycles for the same composition was subtracted to visualize weathering-related changes relative to the composition-specific baseline. Δb^* is shown in centered units, whereas LoG defect contrast and coherence are shown in 8-bit intensity units. The 10-cycle condition for neat rPP is marked as n/a because this condition was not included for neat rPP. (b) Group-level validation predictions obtained using the same final checkpoint and validation split as in Figure 4. Continuous regression outputs were mapped to the nearest weathering-cycle stage for stage-decoded evaluation. Marker color indicates rPP content, and marker shape indicates vPP grade. The dashed diagonal line indicates exact prediction, and the dotted lines are shown as visual references. (c) Composition-dependent prediction residuals for the same validation groups. Residuals were calculated as stage-decoded predicted cycle minus true weathering cycle. The horizontal solid line indicates zero residual, and the shaded region indicates ±10 cycles as a visual guide. Marker color indicates rPP content, and marker shape indicates vPP grade. The results show high ordinal agreement across most rPP/vPP blend compositions, while composition-dependent uncertainty remains for some subgroups.

https://doi.org/10.1021/acssusresmgt.6c00307
ACS Sustain. Resour. Manag. 2026, 3, 3048−3057

Downloaded from pubs.acs.org/asrmcd/article-pdf/3/9/3048/6819596/rm-2026-003077.pdf by NATL INST FOR MATLS SCIENCE (NIMS), Masanobu Naito on 13 October 2026

value increased with rPP content from 14.95 for the rPP 25% blends to 20.00 for neat rPP, whereas the mean structure-tensor coherence decreased from 170.1 to 140.0. This trend indicates that defect-associated contrast and anisotropic texture were influenced by the initial rPP/vPP composition, even before additional accelerated weathering. After subtraction of the composition-specific 0-cycle baseline, Δb^* generally showed positive shifts at later weathering stages for several compositions, while LoG defect contrast and coherence exhibited composition-dependent and non-monotonic responses. Differences between the two vPP grades were also detectable, although the rPP-content effect was more apparent in the group-level summaries. These observations indicate that the physically informed representation should be interpreted as capturing weathering-related surface changes superimposed on composition-dependent baseline morphology, rather than as composition-invariant degradation descriptors.

Composition-dependent prediction behavior was then evaluated by exporting group-level validation predictions from the final checkpoint. The resulting predictions and corresponding residuals are shown in Figure 6b,c, respectively. For the rPP/vPP blend subset, the exact accuracy was 0.800, the stage-decoded MAE was 4.33 cycles, the QWK was 0.917, and the within ±1-stage accuracy was 0.967. The rPP 50% and rPP 75% subsets showed low stage-decoded MAE values of 0.91 and 2.00 cycles, respectively. In contrast, the rPP 25% subset showed a larger MAE of 11.11 cycles and contained one non-adjacent error. The two vPP grades showed similar exact accuracies in this split, 0.800 for both E100-GPL and EA9, although the E100-GPL subset contained one non-adjacent error and showed a larger MAE than the EA9 subset. The neat rPP subset showed no exact predictions in this validation split, but all four predictions remained within one adjacent ordinal stage.

Because the number of validation bags per full composition subgroup was limited, full 6 × 6 confusion matrices for each subgroup were sparse. Therefore, composition-dependent prediction behavior was summarized using group-level predictions, residuals, exact accuracy, stage-decoded MAE, QWK, and within ±1-stage accuracy. The subgroup metrics stratified by rPP content, vPP grade, and full composition are provided in Tables S2–S4 and visualized in Figure S9, and the complete group-level validation predictions are listed in Table S6. Overall, the additional analysis indicates that the model retained high ordinal agreement across most rPP/vPP blend compositions, while composition-dependent uncertainty remained, particularly for the rPP 25% subset and neat rPP. Thus, the present results support the feasibility of the proposed approach within the controlled rPP/vPP blend set incorporating PIR feedstock examined here, but they should not be interpreted as demonstrating composition-independent robustness for arbitrary recycled PP streams.

This composition-dependent behavior provides an important basis for defining the scope of the present study, as discussed below.

SCOPE AND LIMITATIONS OF THE PRESENT STUDY

Because the imposed weathering was superimposed on the pre-existing history of the recycled feedstock, the learned image features may include information from both the initial recycled-material state and the subsequent accelerated weathering response. The present model therefore predicts weathering-related surface states within the controlled material set examined here and should not be interpreted as a universal predictor of absolute degradation history.

The Δb^*-based channel was designed to quantify baseline-referenced yellowing when a corresponding 0-cycle reference image with the same composition and magnification is available. This design is suitable for the relatively clean PIR-derived PP materials examined in this study. However, strongly colored, black, pigmented, filled, contaminated, or post-consumer recycled PP streams may exhibit large initial color and surface-appearance variations that are unrelated to weathering-induced yellowing. For such materials, Δb^* alone may not be sufficient to distinguish degradation-induced color change from product-history- or contamination-derived appearance differences. Extension to heterogeneous PCR streams will require broader training datasets, color calibration or color-invariant feature designs, contamination masking, and validation across colored and filled materials.

The present model should also not be interpreted as a direct predictor of processability or mechanical performance. Practical recycled-PP quality attributes such as melt flow rate or melt flow index, molecular weight, tensile strength, elongation at break, impact strength, embrittlement, and oxidation stability were not included as training targets or validation endpoints in this study. Therefore, the predicted cycle or ordinal stage represents an image-based, weathering-related surface-state label under the controlled accelerated-weathering protocol, rather than a direct measure of deterioration in bulk material performance. Establishing quantitative links between the predicted image-based stage and practical quality attributes remains an important subject for future work.

CONCLUSION

A physically informed, multi-magnification learning framework was developed for the non-destructive assessment of PP weathering from optical micrographs. By integrating a three-channel feature representation with magnification-aware patch extraction and a hierarchical attention-based MIL model, the proposed approach enabled group-level prediction of imposed weathering stages while providing interpretable information across magnifications and image regions. The results demonstrate that the model can capture early-stage image signals associated with PP weathering that are not clearly reflected in ATR-IR carbonyl index measurements under the present conditions, highlighting the complementary role of image-based analysis.

The additional composition-dependent analysis showed that the three physically informed feature channels contained both weathering-dependent changes and composition-dependent baseline variations. The model retained high ordinal agreement across most rPP/vPP blend compositions, while composition-dependent uncertainty remained for some subgroups. These findings support the feasibility of the proposed approach as a non-destructive image-based screening framework for weathering-related surface changes in the controlled PIR-derived PP systems examined here.

From a sustainability perspective, such a framework may contribute to high-throughput screening and data-driven quality assessment by providing a rapid, non-destructive indicator of weathering-related surface states. However, the present model should not be interpreted as a direct predictor of processability

Downloaded from pubs.acs.org/asrmcd/article-pdf/3/9/3048/6819596/rm-2026-003077.pdf by NATL INST FOR MATLS SCIENCE (NIMS), Masanobu Naito on 13 October 2026

https://doi.org/10.1021/acssusresmgt.6c00307
ACS Sustain. Resour. Manag. 2026, 3, 3048−3057

or mechanical performance, because practical recycled-PP quality attributes such as melt flow rate, molecular weight, tensile properties, impact strength, embrittlement, and oxidation stability were not included as training targets or validation endpoints. Future work will focus on validating the method across broader PIR and PCR sources, colors, additives, contaminants, processing histories, optical imaging conditions, and degradation conditions, as well as establishing quantitative links between image-based weathering-stage predictions and practical recycled-PP quality attributes.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssusresmgt.6c00307.

Model and training configuration; detailed hierarchical attention-based MIL architecture; weathering-cycle-wise inter-magnification attention diagnostics on the validation set; training dynamics and checkpoint selection; composition-dependent feature analyses; subgroup-level validation metrics; group-level validation predictions; and composition-cycle feature summaries (DOCX)

■ AUTHOR INFORMATION

Corresponding Author

Masanobu Naito – *Graduate School of Science and Technology, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan; Research Center for Macromolecules and Biomaterials, National Institute for Materials Science (NIMS), 1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan;* orcid.org/0000-0001-7198-819X; Email: NAITO.Masanobu@nims.go.jp

Authors

Kazuki Hammura – *Graduate School of Science and Technology, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan; Asahi-Kasei Corporation, 1-1-2 Yurakucho, Chiyoda-ku, Tokyo 100-0006, Japan;* orcid.org/0009-0000-9671-5059

Kiyotaka Hitomi – *Research Center for Macromolecules and Biomaterials, National Institute for Materials Science (NIMS), 1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan*

Kenji Nagata – *Center for Basic Research on Materials, National Institute for Materials Science (NIMS), 1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan*

Complete contact information is available at: https://pubs.acs.org/doi/10.1021/acssusresmgt.6c00307

Author Contributions

K.H. performed the data analysis, implemented the computational workflow, prepared the figures, and drafted the manuscript. K.H. prepared the specimens and conducted the experimental work, including accelerated weathering, optical microscopy, and ATR-IR measurements. K.N. provided guidance on research strategy and data interpretation. M.N. conceived and supervised the study, acquired funding, and reviewed and edited the manuscript. All authors approved the final manuscript.

Funding Sources

This work was supported by the JSPS KAKENHI Grant-in-Aid for Scientific Research (B) Grant Number 23K26724. M.N. acknowledges the support of JST CREST (JPMJCR19J3), the MEXT Program: Data Creation and Utilization-Type Material Research and Development Project Grant Number JPMXP1122714694, and the Cross-ministerial Strategic Innovation Promotion Program (SIP) JPJ012290.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors used an AI language model to assist with the language editing and organization of the manuscript. All scientific content, data analysis, interpretation, and conclusions were reviewed and approved by the authors.

■ REFERENCES

(1) Hossain, M. T.; Shahid, M. A.; Mahmud, N.; Habib, A.; Rana, M. M.; Khan, S. A.; Hossain, M. D. Research and application of polypropylene: A review. *Discover Nano* **2024**, *19*, 2.

(2) Rabello, M. S.; White, J. R. Crystallization and melting behaviour of photodegraded polypropylene—I Chemi-crystallization. *Polymer* **1997**, *38* (26), 6379–6387.

(3) Craig, I. H.; White, J. R.; Kin, P. C. Crystallization and chemi-crystallization of recycled photo-degraded polypropylene. *Polymer* **2005**, *46* (2), 505–512.

(4) Aurrekoetxea, J.; Sarrionandia, M. A.; Urrutibeascoa, I.; Maspoch, M. L. Effects of recycling on the microstructure and the mechanical properties of isotactic polypropylene. *J. Mater. Sci.* **2001**, *36* (11), 2607–2613.

(5) Fujiyama, M.; Wakino, T.; Kawasaki, Y. Structure of skin layer in injection-molded polypropylene. *J. Appl. Polym. Sci.* **1988**, *35* (1), 29–49.

(6) Fechine, G. J. M.; Demarquette, N. R. Cracking formation on the surface of extruded photodegraded polypropylene plates. *Polym. Eng. Sci.* **2008**, *48* (2), 365–372.

(7) Hashimoto, N.; Fukushima, D.; Koga, R.; Takagi, Y.; Ko, K.; Kohno, K.; Nakaguro, M.; Nakamura, S.; Hontani, H.; Takeuchi, I.Multi-scale domain-adversarial multiple-instance CNN for cancer subtype classification with unannotated histopathological images. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA, June 14–19; IEEE: Piscataway, NJ, **2020**; pp 3851–3860.

(8) Dietterich, T. G.; Lathrop, R. H.; Lozano-Pérez, T. Solving the multiple instance problem with axis-parallel rectangles. *Artif Intell* **1997**, *89* (1–2), 31–71.

(9) Ilse, M.; Tomczak, J. M.; Welling, M.Attention-based deep multiple instance learning. *Proceedings of the 35th International Conference on Machine Learning*, Stockholm, Sweden, July 10–15, 2018; Dy, J.; Krause, A., Eds.; Proceedings of Machine Learning Research; PMLR, **2018**, vol *80*, pp 2127–2136.

(10) Cao, W.; Mirjalili, V.; Raschka, S. Rank consistent ordinal regression for neural networks with application to age estimation. *Pattern Recogn. Lett.* **2020**, *140*, 325–331.

(11) Zhang, H.; Zhao, M.; Liu, M.; Luo, J.; Guan, Y.; Zhang, J.; Xia, Y.; Zhang, D.; Zhou, X.; Fan, L.; Liu, S.; Zhou, S. K.Hierarchical multiple instance learning for COPD grading with relatively specific similarity. In: *Medical Image Computing and Computer Assisted Intervention—MICCAI 2024, Marrakesh, Morocco, October 6–10, 2024*; Linguraru, M. G.; Dou, Q.; Feragen, A.; Giannarou, S.; Glocker, B.; Lekadir, K.; Schnabel, J. A.; eds.; Lecture Notes in Computer Science; Springer Nature Switzerland:Cham, Switzerland, vol *15001*; **2024**, 536–545.

Downloaded from pubs.acs.org/asrmcd/article-pdf/3/9/3048/6819859/rm-2026-003077.pdf by NATL INST FOR MATLS SCIENCE (NIMS), Masanobu Naito on 13 October 2026

(12) Liu, Z.; Mao, H.; Wu, C.-Y.; Feichtenhofer, C.; Darrell, T.; Xie, S.A ConvNet for the 2020s. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, New Orleans, LA, June 19–24; IEEE: Piscataway, NJ, **2022**; pp 11966–11976.

Downloaded from pubs.acs.org/asrmcd/article-pdf/3/9/3048/6819859 6/rm-2026-003077.pdf by NATL INST FOR MATLS SCIENCE (NIMS), Masanobu Naito on 13 October 2026

https://doi.org/10.1021/acssusresmgt.6c00307
ACS Sustain. Resour. Manag. 2026, 3, 3048−3057