Computer Vision for Predictive Rating Analysis of Coating Durability

Industry Mentors: Dr. Katie Daisey, Research Scientist, Arkema;
Doug Mall, Applications Specialist, Arkema

CHEG/CISC/ECEG/MSEG 848
Spring 2024

Dr. Arthi Jayaraman, Professor; Tristan Meyers TA

May 13, 2024

Justin Labombard
Ph.D.
Department of Electrical and Computer Engineering
The University of Delaware

Ana Maria Mosquera Rodriguez
Ph.D.
Department of Material Science and Engineering
The University of Delaware

Tasia Walker
Ph.D.
Department of Chemistry
Delaware State University

Abstract

Arkema, a global leader in specialty chemicals and advanced materials, offers a diverse range of key products and services, including specialty adhesives, polyamide resins, electroactive polymers, UV, LED, and EB curable resins, among others. Partnering with Arkema in this project, we aimed to develop a robust computer vision and predictive modeling framework for assessing and evaluating the longevity of coatings using image data. Utilizing Python and PyTorch with an NVIDIA GPU 3070Ti, we designed models to identify weathering effects on coating panels and predict their performance over time. The coated panels undergo evaluation in accordance with the American Society for Testing and Materials (ASTM) standards. The ASTM Visual Standard employs a numerical grading system based on the appearance of the coating, ranging from 10 (no discernible failure mode) to 1 (evident failure mode). The coated panels' performance was evaluated for the developed model based on their resistance to chalking, flaking, and cracking, as well as the accumulation of mildew and algae. The identification of failure modes significantly contributed to the training and testing of the developed model. Our initial results indicate the model's effective quantification and prediction capabilities, offering substantial improvements over traditional visual inspection methods. This advancement aids in precise and replicable performance assessments in the coatings industry.

1. Introduction

Coatings have been employed for centuries across various sectors of society primarily to protect and enhance the appearance of materials. The application of coatings has expanded the growth of social and industrial advancements.[1] As global demands on material performance intensify, driven by environmental considerations, technological advancements, and consumer expectations, the need for high-performance coatings has never been more critical. In this context, the ability to accurately assess and predict the durability and performance of coatings under different environmental conditions becomes essential for the development of new materials and for maintaining the competitiveness of coating manufacturers.[2,3]

Traditionally, the assessment of coating durability at Arkema, as in much of the industry, has relied heavily on visual inspections and manual grading systems based on ASTM standards. These standards facilitate the systematic grading of various failure modes such as Cracking & Checking, Discoloration, Flaking, Erosion, Chalking, and Mildew/Algae formation. While useful, these manual methods are inherently subjective and often lead to inconsistencies in data and limitations in scalability and repeatability.[3,4] Figure 1 is a visual representation of two selected failure modes of the coated panels from 2018 to 2022.

Figure 1. Images of panels displaying cracking & checking (left) and discoloration (right) [Images provided by Arkema]

Emerging technologies in computer vision and machine learning offer transformative potential for the coatings industry. These technologies can automate and standardize the analysis of coating performance, reducing human error, and increasing the efficiency of the assessment processes. Notably, advancements in machine learning have facilitated the development of models capable of identifying complex patterns and predicting material behaviors from large datasets, which are infeasible for human evaluators.[5] Recent studies have similarly emphasized the integration of machine learning in coating assessments. Researchers have explored the application of machine learning to predict the erosion of thermal barrier coatings at elevated temperatures, employing various algorithms to analyze the performance of coatings used in gas turbine engines.[6] Additional research presented an AI-based image recognition tool developed to assist inspectors in coating condition assessments, showcasing the practical applications of machine learning in maritime asset maintenance.[7] A similar approach to our model was developed with the use of an automated CNN-based coating breakdown and corrosion (CBC) assessment system for marine and offshore structures. The system utilizes a customized unmanned aerial system (UAS) to collect visual and thermal data from inaccessible areas. Deep learning-based object detection and instance-aware semantic segmentation are used to identify and quantify different types of CBC, including surface CBC and edge CBC. The system can generate an inspection report according to international standards, and the results are more accurate and consistent than manual inspection by surveyors.[8]

For our project, we leverage the above cutting-edge computer vision and predictive modeling technologies to enhance the precision and reliability of coating assessments. By utilizing a machine learning architecture, specifically the convolutional neural network (CNN), this project not only automates the detection of weathering effects on coatings but also predicts the long-term performance of these materials from historical data.

Project Goal

Our work aims to address the aforementioned challenges with coatings assessment by developing a computational model that objectively and quantitatively assesses the performance of coated panels. One model must evaluate the coatings across four metrics: cracking, flaking, chalking, and mildew and algae. Another model must forecast the performance of the coatings in the future using current data trends. Our project is split into three major objectives:

- **Objective 1*:*** *Image Splitting*: Split the images into samples for a machine-learning model.
- **Objective 2:** *Rating Predictor*: Develop a model that predicts ratings given an image sample.
- **Objective 3:** *Future Rating Predictor*: Modify the developed rating predictor to forecast ratings with a given series of historical and current image samples.

2. Computational Approach

2.1 Technical Framework and Computational Resources

We utilized the Python programming language and the PyTorch library to implement machine learning models. Computations were performed using an NVIDIA GPU 3070Ti, which provided the necessary power for processing large datasets and complex algorithms efficiently.

2.2 Data Preprocessing

The preprocessing of image data is a crucial step in ensuring the accuracy of our models. The process included:

- Splitting the images of panels into individual sections and matching these sections to the labels provided in a separate Excel file.
- Random Cropping: Initial cropping of images to 100×100 pixels to reduce dimensionality and focus on relevant features.
- Resizing: Images were resized to 256×256 pixels using bilinear interpolation to standardize the input size while preserving image quality.
- Central Cropping: A further crop centered to 224×224 pixels to maintain uniformity across all inputs.
- Rescaling: To normalize the data, each pixel value was rescaled to a range of [0, 1].
- Normalization: The rescaled images were then normalized using a mean of [0.485, 0.456, 0.406] and a standard deviation of [0.229, 0.224, 0.225], aligning them with the pre-trained VGG19 model requirements.[9]

This is what the rack images provided by Arkema look like.

We apply **histogram equalization** to keep consistent colors between images.

By using our app, we can segment into individual sections.

Figure 2. Preprocessing steps of the images for segmentation.

2.3 Dataset Adequacy

The current dataset size of 312 samples, which are split into 250 training images, 47 testing images, and 15 validation images, is insufficient for developing highly accurate predictive models. By cropping the images into five smaller images, the overall dataset size, as well as the size of each split, is increased by five-times (5x) the original dataset size. This was carried out in a uniform manner and at random. For the uniform 5x crop dataset, each image was cropped into five parts (top left, top right, bottom left, bottom right, center), resulting in 1,250 training images, 235 test images, and 75 validation images. For the random crop dataset, we evaluated how the model performed when looking at inconsistent crop locations instead of five consistent crop locations. This change helped mitigate the lack of samples, but this limitation still affected how well the models could generalize without overfitting.

2.4 Model Development

The models used are (CNN)-VGG19: This deep learning model is renowned for its effectiveness in image classification and feature extraction tasks.[8] First and Second-Order Statistics: These models provide a statistical texture analysis by evaluating pixel distribution and relationships, offering insights into the surface properties of coatings.[9] Style Transfer for Texture Analysis: Leveraging artistic style transfer techniques, this model assesses the texture of coatings by transforming their visual appearance, which can reveal underlying degradation patterns.[11]

Method 1: Convolutional Neural Network (CNN) with Original Images

We employ VGG19, a deep CNN pre-trained on the ImageNet dataset.[12] ImageNet is a large visual database designed for visual object recognition software research, containing over 14 million images categorized into over 20,000 classes. The pre-training on such a comprehensive dataset allows VGG19 to have a highly developed ability to recognize a wide range of features and textures, making it exceptionally suitable for analyzing complex visual data such as coated panels.[9,12] We chose the VGG19 network due to its simplicity and ability to find texture information, as seen in Style Transfer.

- **Transfer Learning:** Given the limited dataset, transfer learning is employed, where the core of the network remains unchanged, and only the final layers are trained to adapt to our specific task of identifying different failure modes. Our VGG19 model consists of four heads, each set to predict one of the four failure modes.

Method 2: First and Second-Order Statistics

This method employs first and second-order statistical metrics derived from the image data to perform detailed texture analysis of the coatings. The primary goal is to capture essential textural information that can indicate various states of coating degradation.

- **First-Order Statistics**

First-order statistics are used to characterize the distribution of gray levels within an image based on the histogram. Metrics such as mean, variance, energy, and entropy provide a quantitative description of the image texture, reflecting the basic intensity distribution characteristics.

- **Second-Order Statistics**

Since first-order statistics do not account for the spatial distribution of pixels, second-order statistics are crucial for more detailed texture characterization. We utilize the Gray-Level Co-occurrence Matrix (GLCM) and the Neighboring Gray-Level Dependence Matrix (NGLDM) to extract deeper textural patterns.

GLCM: Measures the frequency of pixel pairs (with specific gray levels) at a given distance and orientation, capturing the spatial relationships between pixels.

NGLDM: Focuses on the dependence of gray levels neighboring each pixel, offering insights into the textural uniformity or variability.

The incorporation of statistical methods for characterizing surface textures is supported by previous research, which has highlighted the effectiveness of these metrics in capturing the nuanced textures of coating surfaces.[10] This directly informs the approach taken for our project. We used a Multi-Layer Perceptron (MLP) with these statistics to acquire predictions for the failure modes. The number of layers and parameters for the MLP are treated as hyperparameters and optimized using the validation set. This MLP-based network is significantly smaller in size compared to the VGG19 model, with approximately 70,000 times fewer parameters, depending on the specific hyperparameters chosen. This smaller model size can be an advantageous factor, particularly in scenarios with limited computational resources.

Method 3: Style Transfer for Texture Analysis

Style transfer is a technique that modifies the "style" of an image while preserving its "content." By leveraging the capabilities of CNNs, particularly the VGG19 model, we can analyze the textural features from coating images, which are essential for accurately assessing their condition. The method effectively uses embeddings within a pre-trained network to capture and transfer styles between images, which can be applied to texture analysis to highlight areas of potential degradation or wear in coating images.[11]

3. Results

Our project utilized three machine learning models and image processing techniques to assess coating degradation with three datasets. Figure 3 presents the count of each rating for each type of coating degradation. The plots show that the failure modes are skewed towards higher ratings, meaning most of the samples received had little to no degradation. The performance of each model across different preprocessing techniques is summarized in Table 1, which provides detailed accuracy metrics for each selected type of coating degradation: Cracking, Flaking, Chalking, and Mildew/Algae.

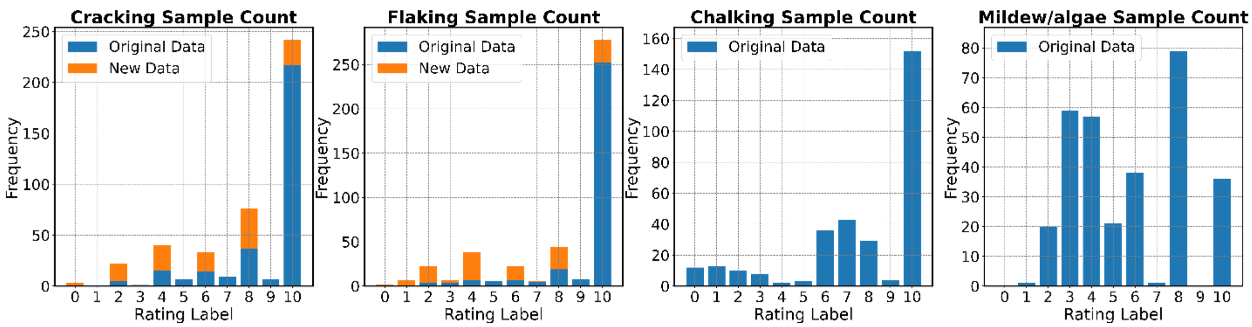

Figure 3. The graphical representation of the number of samples for each of the four failure modes (Cracking, Flaking, Chalking, and Mildew/Algae).

	Cracking	Flaking	Chalking	Mildew/Algae
VGG19 5x Crop	45.75 ± 1.61	53.14 ± 1.67	47.05 ± 3.46	29.27 ± 3.97
VGG19 Random Crop	56.11 ± 1.08	**69.89 ± 1.33**	61.89 ± 5.45	**40.11 ± 4.98**
First & Second-Order Statistics	52.83 ± 0.00	64.15 ± 0.00	48.00 ± 3.32	36.50 ± 6.34
Style Transfer 5x Crop	44.56 ± 1.22	53.32 ± 2.46	54.11 ± 0.10	31.62 ± 1.41
Style Transfer Random Crop	**57.78 ± 2.38**	68.15 ± 1.67	**68.11 ± 1.60**	39.96 ± 3.44

Table 1. The models mean and standard deviation results for accuracy performance for each of the four failure modes.

Model Performance and Computational Considerations

VGG19 Model:
The comparison between the 5x crop and random crop methods highlights significant differences in model performance across all degradation types. Notably, the random crop method consistently yields higher accuracy. The most pronounced improvement is seen in the detection of cracking and flaking, where the random crop method shows an increase of approximately 18-20 percentage points in accuracy compared to the 5x crop method. Similarly, for chalking and mildew/algae, the random crop method improves accuracy by roughly 13 and 20 percentage points, respectively.

First and Second-Order Statistics Model:
Overall, the performance of the first and second-order statistics model highlights the potential for these statistical measures in texture analysis for coating degradation. The higher accuracy in detecting cracking and flaking indicates that variations in pixel intensity distributions (first order) and spatial relationships (second order) are strong indicators of these conditions. Lower accuracy in detecting chalking and mildew/algae might be due to the more nuanced textural changes associated with these types of degradation, which are potentially less distinct in the original unprocessed images. The small size of this network paired with the performance reveals how useful the texture information is to this problem.

Style Transfer Model:
The style transfer model shows promising results, particularly when applied to images processed with the random crop method. This method improves detection capabilities by enhancing textural distinctions that are key indicators of degradation. The improvement in detection rates across all metrics when using random cropping suggests that this preprocessing technique helps expose more diverse textural features to the model.

4. Conclusion

Our evaluation of different models for coating degradation—VGG19, first and second-order statistics, and style transfer—highlights a trade-off between model complexity, accuracy, and computational efficiency. The best-performing model for this task is the style transfer network, which is the largest network among the models considered. This model leverages the concepts of style transfer to extract texture information that is useful for making accurate predictions. The second-largest model is the VGG19 network, which was trained using transfer learning techniques. This approach allowed the model to effectively handle the small size of the dataset, and its performance fell in the middle range compared to the other models. The worst-performing model was the statistical model, which was the smallest of the three models presented. However, despite its smaller size, this model still managed to achieve results that were comparable to the other, larger models, utilizing a much smaller number of parameters.

The current dataset size is insufficient for effectively training more complex models, such as VGG19 and the style transfer network, or providing adequate results for future rating prediction. This limitation could potentially lead to generalization issues when applying these models to real-world applications. In contrast, the first and second-order statistics model, which utilizes a significantly smaller MLP with only 4,655 parameters, offers a more practical and cost-effective alternative. Despite its smaller size, this model can deliver respectable accuracy, making it a viable option when computational resources are limited. Furthermore, the operational costs for running high-parameter models, such as VGG19 and style transfer, are substantially greater than those for simpler models, making the first—and second-order statistics models more suitable for continuous use in cost-sensitive environments.

Open access/open-source code links:
Our GitHub page holds the code used for segmenting the images using two mouse clicks as well as the code for the models trained in our work. We provide this open-source code to help future work on this project and other related projects. The GitHub page will come with the final version of this document.

Acknowledgments: We would like to express our utmost appreciation and gratitude to Dr. Arthi Jayaraman for her guidance and support, our industry Mentors, Katie Daisey and Doug Mall, for their insight into industry needs, and the NRT-Midas Program for the learning opportunity.

References

(1) Makhlouf, A. S. H. (2011). Current and advanced coating technologies for industrial applications. In Elsevier eBooks (pp. 3–23). https://doi.org/10.1533/9780857094902.1.3

(2) *Evaluation of Weathering Effects: Visual Inspections and Instrumental Measurements*.

(3) Adeboye, S., Adebowale, A. D., Siyanbola, T. O., & Ajanaku, K. O. (2023). Coatings and the environment: a review of problems, progress and prospects. IOP Conference Series. Earth and Environmental Science, 1197(1), 012012. https://doi.org/10.1088/1755-1315/1197/1/012012

(4) Cogulet, A., Blanchet, P., & Landry, V. (2019). Evaluation of the impacts of four weathering methods on two acrylic paints: Showcasing distinctions and particularities. Coatings, 9(2), 121. https://doi.org/10.3390/coatings9020121

(5) Rannetbauer, W., Hambrock, C., Hubmer, S., & Ramlau, R. (2024). Enhancing predictive quality in HVOF Coating Technology: A Comparative analysis of Machine Learning techniques. Procedia Computer Science, 232, 1377–1387. https://doi.org/10.1016/j.procs.2024.01.136

(6) Liu, Y., Ravichandran, R., Chen, K., & Patnaik, P. C. (2021). Application of Machine Learning to Solid Particle Erosion of APS-TBC and EB-PVD TBC at Elevated Temperatures. Coatings, 11(7), 845. https://doi.org/10.3390/coatings11070845

(7) Wen, F., Gu, H., Wang, B., Vitali, T., & Rasberry, H. (2020). The Development of a Machine Learning-Based Image Recognition Tool for Coating Inspections. https://doi.org/10.4043/30504-ms

(8) Liu, H., Tan, E., Cai, Z., Yin, X., & Zhen, Y. (2018). CNN-based automatic coating inspection system. Advances in Science, Technology and Engineering Systems Journal, 3(6), 469–478. https://doi.org/10.25046/aj030655

(9) Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. CoRR, abs/1409.1556.

(10) Akhil, V. M., Raghav, G. R., Arunachalam, N., & Desamsetti, S. (2020). Image Data-Based surface texture characterization and prediction using machine learning approaches for additive manufacturing. Journal of Computing and Information Science in Engineering, 20(2). https://doi.org/10.1115/1.4045719

(11) Gatys, L. A., Ecker, A. S., & Bethge, M. (2016). Image Style Transfer Using Convolutional Neural Networks. The IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016-December, 2414–2423. https://doi.org/10.1109/cvpr.2016.265

(12) Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M. S., Berg, A. C., & Li, F. (2015). ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision, 115(3), 211–252. https://doi.org/10.1007/s11263-015-0816-y

(13) Toet, Alexander & Wu, Tirui. (2014). Efficient contrast enhancement through log-power histogram modification. Journal of Electronic Imaging. 23. 063017. 10.1117/1.JEI.23.6.063017

Appendix

Appendix A: Histogram Equalization

Histogram Equalization is a method used to improve the contrast of a grayscale image. The method redistributes the pixel intensities of the image to span a wider range. Figure 4 shows what histogram equalization does to a grayscale image. The original image (A) is dark and contains pixel values centered around black (intensity=0) and the equalized image (B) is lighter with a wider range of pixel intensities.

Figure 4. Example of histogram equalization[13]. The right image (A) is the original image and the image on the left (B) is the equalized image.

Since the lighting conditions of our images are variable, we used equalization as an attempt to normalize the brightness between images. Our exact approach to applying histogram equalization on 3-channel images is explained below:

1. Convert RGB image to HSL. HSL is required to have a separate channel for light (L).
2. Apply histogram equalization to the L-channel. This step is used to increase the contrast of the image.
3. Convert the image back from HSL into RGB. The image is now back to the original format with an increase in contrast.

Appendix B: Full Table of Results

The full table of results shown in Table 2 is an extension of Table 1. This table shows all of the possible dataset/model combinations and provides some insight into both the performance of the models as well as the data itself. The style transfer model performs the best on chalking and mildew/algae in every case, showing that these two failure modes are closely tied to texture. The performance difference between the style transfer model and the VGG19 transfer learning model is minimal for cracking and flaking which shows that the VGG19 model already has the embeddings necessary for a correct prediction without the additional network. The statistics model performs the worst consistently, but due to its small size, it still offers a competitive result for a small fraction of the price.

	Standard Dataset (440 samples)			5x Crop (2,200 samples)			Random Crop (2,200 samples)		
Failure Modes	VGG19	Stats	Style	VGG19	Stats	Style	VGG19	Stats	Style
Cracking	60.19 (2.14)	52.83 (0.00)	**61.70** **(2.08)**	**45.75** **(1.61)**	42.26 (0.02)	44.56 (1.22)	56.11 (1.08)	52.83 (0.34)	**57.78** **(2.38)**
Flaking	**74.90** **(1.70)**	64.15 (0.00)	74.53 (3.40)	53.14 (1.67)	51.32 (0.00)	**53.32** **(2.46)**	**69.89** **(1.33)**	64.15 (0.00)	68.15 (1.67)
*Chalking	69.00 (2.13)	48.00 (3.32)	**69.67** **(1.79)**	47.05 (3.46)	41.50 (1.22)	**54.11** **(0.10)**	61.89 (5.45)	49.20 (0.44)	**68.11** **(1.60)**
*Mildew/Algae	39.58 (3.61)	36.50 (6.34)	**59.33** **(2.00)**	29.27 (3.97)	22.33 (1.88)	**31.62** **(1.41)**	**40.11** **(4.98)**	32.67 (2.71)	39.96 (3.44)

Table 2. Full table of accuracy values over all possible dataset/model combinations. This table is the full version of Table 1 with average accuracy values on top and the standard deviation of the accuracy values on the bottom in parentheses. * New labels were not provided for these failure modes; we use only the original data for these failure modes.