

Transfer Learning-Based Coal Classification Through Enhanced CNN Model Architecture & Workflow

Appurva Rajendra Kapil¹, Dr. Manika Verma², Dr. Chintan Shah³

¹*Department of Information Technology, Research Scholar, Silver Oak University*

²*Professor, Silver Oak University*

³*Associate Professor, Silver Oak University*

Abstract—In this work, a coal image classification technique using a custom CNN architecture and transfer learning models (ResNet50 and VGG16) is presented to increase accuracy and robustness of coal recognition. The proposed pipeline consists of preprocessing, resizing, normalization, augmentation, and per-class evaluation to handle the texture, lighting, and background variation problems. The automatic extraction of discriminative visual features of coal images and classification into specific classes differs from traditional manual inspection of coal images. This experiment aims to compare the performance of three models under the same training and validation conditions in order to find out which one is the most suitable. Accuracy, F1-score, confusion matrix, and explainability using Grad-CAM are selected as criteria for performance evaluation. The methodology can be further developed for industrial application in coal sorting and inspection.

Index Terms—Coal categorization, deep learning, CNN, transfer learning, ResNet50, VGG16, image preprocessing

I. INTRODUCTION

Coal remains among the most utilized fossil fuels for generation of energy, metallurgy processes, cement production and other industrial applications. While the global energy market is shifting towards utilization of clean energy, coal is still being used as an energy source for industrial and electricity needs in many countries. The properties and quality of the coal have a direct effect on the burning efficiency, calorific value, ash formation, emissions and burning process. Therefore, proper classification of the coal is necessary for coal handling, sorting, quality testing and decision-making in the industry. Coal is a complex material in terms of its physical and chemical

properties, texture, reflectivity, moisture and mineral content. Coal differs greatly in appearance and structure, depending on the kind of coal. Some common examples of coal include anthracite, bituminous coal and lignite coal, each having its own appearance and structure. However, this variety may not be noticeable through conventional methods, particularly when the lighting is poor, dusty, particle shapes are irregular or there is background noise. The traditional methodologies used for coal classification are mostly based on visual inspection and/or laboratory analysis. Laboratory methods provide precise results but are time-consuming and costly, and they are not applicable in situations that require instant results. On the other hand, manual inspection is quite subjective, depending on operator experience, environment, and load. This has led to the urgent requirement of a reliable automatic and accurate coal classification system that can cope with large numbers of coals. The application of computer vision and deep learning has been highly effective in recent times for the purposes of image recognition, object detection, defect detection, and industrial quality control. Convolutional Neural Network (CNN) is one of the most effective models for deep learning, which automatically learns spatial and hierarchical features from images.

CNN is one of the most effective image classifiers compared to conventional machine learning methods, which depend on manually engineered features. These learned features can be used for differentiating between classes of coal based on the surface roughness, grains' distribution, brightness variation, etc. CNN models trained from scratch will be highly sensitive to requirements for generalization performance which, in turn, are tightly connected with

the number, balance, and diversity of the training sets. In case of applying machine learning algorithms to problems like classification of coal images, the problem with the data set size might be present (it could either be small or impossible to achieve balance within the set). This will lead to overfitting, improper convergence and, as a result, low classification accuracy especially when we talk about minority classes. One of the good solutions to this issue is transfer learning which will allow using knowledge gained from large image databases. The pre-trained networks like ResNet50 and VGG16 can be adapted to coal classification by leaving their feature extraction network and adjusting only the final classification network according to coal classes.

A customized CNN architecture is presented in this current study as the coal classification model and two transfer learning architectures will be applied: ResNet50 and VGG16. The aim is to conduct a comparison between the performance of the task-specific CNN model and pretrained models in the same preprocessing, training, and validation dataset. The proposed pipeline is as follows: resizing of images, normalization, augmentation, and evaluation of the images according to the classes. The underlying reason is that the coal images need to be preprocessed in the following manner because they may contain different scales, lighting, orientation, texture, and background. Data augmentation can also be conducted to enhance the robustness of the model by generating more image variations during training. In this paper, the evaluation of the model is not restricted to its performance in terms of accuracy alone. In addition, consideration is given to evaluation measures such as F1 Score and confusion matrix because of the possibility of misinformation due to class imbalance. The metrics provide an understanding of how the classes are being classified and the accuracy of the models. Furthermore, methods such as Gradient-weighted Class Activation Mapping are used for visualizing regions of the image that influence model decisions. This is particularly important when applying the solution in an industrial setting because it increases trust in the accuracy of automation systems. The key finding of this study is the creation of a repeatable and scalable deep learning-based coal image classification model. The study also evaluates the performance of a custom CNN against transfer learning models, giving insightful information

regarding the appropriateness of particular architectures to coal recognition tasks. Intelligent coal sorting, inspection and quality control system can be implemented by the proposed method, which can reduce the manual operation and increase the uniformity of the decision-making. In summary, the use of deep learning and transfer learning in the field of coal classification, in an explainable, robust, and accurate way, for the industrial analysis of materials is shown to be a very promising approach.

II. LITERATURE REVIEW

Coal is still considered one of the main fossil fuels in terms of production of energy and metallurgical processes, as well as for various industrial applications [1]. In reality, coal is not homogeneous, rather it consists of different types of coals, which vary by color, texture, reflectivity, moisture content, and structural features [2]. The differences are critical in the sense that they influence the characteristics of combustion, calorific value, and efficiency of the processes. Therefore, the classification of such resources becomes an essential component of the handling [3], processing, and decision making concerning the reliable coal. Traditionally, such classification is done via visual inspection, which is manual, and/or laboratory examination. Such approaches tend to be time-consuming, labor-intensive and error-prone and can be effective in the controlled environment [4]. However, such approaches become inefficient in cases when there is a need to classify large amounts of coal images or conduct real-time inspections [5].

This has led to the generation of interest in automated image-based classification systems capable of performing the same work within a shorter period, and in a more uniform fashion. Deep learning developments have led to a complete transformation of image classification in the fields of science and industry. Image classification has become quite efficient in many fields of science and industry in recent times owing to deep learning. CNN is quite suitable for the purpose as it can learn hierarchical features from image data directly without the need to design features by oneself. For example, considering coal images, CNN is able to learn edges, grain structures [6], texture variations etc., which are important discriminative surface features. But,

however, for a small or unbalanced dataset, which is common in a specialized image classification task, a custom CNN will not do the job properly. In this case, transfer learning is useful. Pre-trained model such as ResNet50 and VGG16 will be able to use these highly efficient feature extractors that were trained on big datasets and transfer the knowledge to the dataset related to the coal images. It can result in better convergence, stability and generalization. In this paper, a novel coal classification model that incorporates three approaches to the learning – custom CNN, transfer learning from ResNet50 and transfer learning from VGG16 – has been introduced [7]. The objective of this approach is to properly classify the coal images as well as to compare the performance of different models under the same preprocessing and evaluation criteria [8]. Another feature of this study that needs to be mentioned is robustness. There is some variance in lighting conditions, image quality, background noise, and class imbalance in the case of coal datasets [9].

Care should be exercised during the preprocessing stage so as to avoid overfitting of the model to the most common classes and poor performance in rare classes. To solve this problem, resizing, normalization, data augmentation, and contrast enhancement (where applicable) have been adopted in the process [10]. The other key aspect of this is evaluation. Although some papers measure accuracy alone, the problem with this measure is that in case of an unbalanced class the accuracy becomes meaningless [11]. The reason for this is that even though a classifier can be very accurate in predicting the majority class, it might not be very accurate for the minority classes. Therefore, in the proposed study, the following set of metrics will be employed, namely F1 Score, Confusion Matrix [12] and the visual inspection of the explainability. In addition to the above, the paper discusses a more methodological topic: whether to optimize or construct a given CNN and whether to apply transfer learning or not [13]. This comparison raises the scientific value of the paper through the connection between the architecture of the model and the results obtained during the performance and interpretability analysis. The automatic classification of coal has practical importance for such processes as sorting and inspections and quality assessment of energy-related products [14]. It would be possible for operators to make more consistent and faster decisions

with the help of the technique without additional manual labor. This is one more case showing how computer vision is used to solve traditional problems of materials analysis. The proposed approach is important not only because it is an example of classification task [15] but also because it is an example of using deep learning for the optimization of industrial process.

III. SYSTEM METHOD

The model system is built to be a full image classification system beginning from coal images as the input and ending in coal class labels as the output illustrated in Fig.1. An image vector of the coal using RGB color format is acquired from the dataset and rescaled to a fixed input dimension to ensure the same input format for all samples. Once rescaled, the image is then normalized, augmented and enhanced with contrast enhancement (like CLAHE) if necessary. This process is crucial because the coal images may vary in brightness, visibility of textures, and background quality. The system has the ability to remove noise from the image by pre-processing of the input images ensuring stability of the learning process. Then, the normalized image is forwarded to the classification branch of one of the three branches in the model comparison framework.

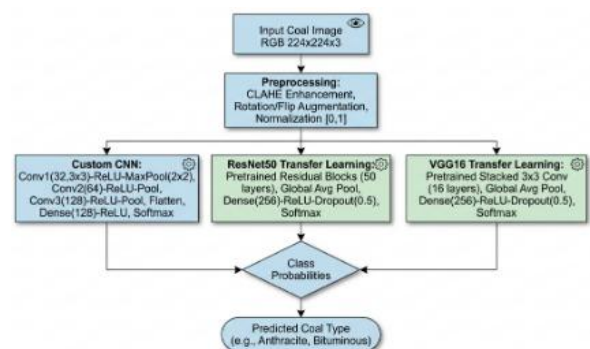


Fig.1: The System Model

The first branch consists of the custom CNN model, which is constructed from a sequence of CNN, max-pooling, and classification layers. These models learn coal texture and surface patterns. This branch may be used as the base line since it does not rely on any other features than the features extracted from the coal. The second branch is based on the pre-trained deep ResNet50 model, which has residual connections. These connections allow for smoother gradient flow,

which makes training deeper neural networks easier for image recognition. One more pre-trained model used as one of the transfer learning branches is VGG16. This model uses a stack of small convolution filters to extract fine details of the visual information. In both transfer learning branches, the feature extractors are learned, and the output layers are retrained to coal classes in the coal data set. Every branch is followed by a softmax output layer giving probability distribution of coal categories.

The selection of the class happens through selecting the class with the highest softmax probability. It is an output probability that is valuable since it gives both the class prediction and the degree of confidence about the class prediction. The comparative architecture also provides flexibility in the system. It makes it possible for the researcher to determine whether the simple CNN developed is sufficient or there are other pre-trained models that perform better using the same data. Additionally, one can apply the explainability methods such as Grad-CAM and SHAP that can be applied after prediction and make visualization of regions and features used for decision making. This can be very valuable for coal analysis when texture, grain boundaries and surface irregularities become the most reliable criteria.

The system model is thus divided into three successive stages: a preprocessing stage, a feature learning stage and a classification stage. Preprocessing ensures the quality and consistency in the data, feature learning

extracts meaningful representations from the data, and classification transforms these representations into coal type predictions. The comparison is meaningful and fair as it is based on the same input pipeline. The model structure also ensures that the system is reproducible; another researcher can reproduce the process with the same preprocessing steps and network structure. The system could be embedded into a larger industrial inspection system that automatically takes, analyses and classifies images of coal. Hence, the system model is not only a methodical framework but also a conceptual design that is ready for deployment for the classification of coal images.

IV. METHODOLOGY WORKFLOW

The adopted approach for work includes a process of steps starting from data set preparation up to the model evaluation and interpretation which is shown in Fig.2. The very first step is the data set collection which will be then loaded into the training environment. After that the data set is loaded and organized in the folder per class in order for the image generator to access the data set properly. Once the data set is loaded, it should be split into the training and validation sets. It is required because model parameters are trained using the training set, while generalization ability of the model is evaluated on the validation set. In case of unbalanced classes, the augmentation process is performed.

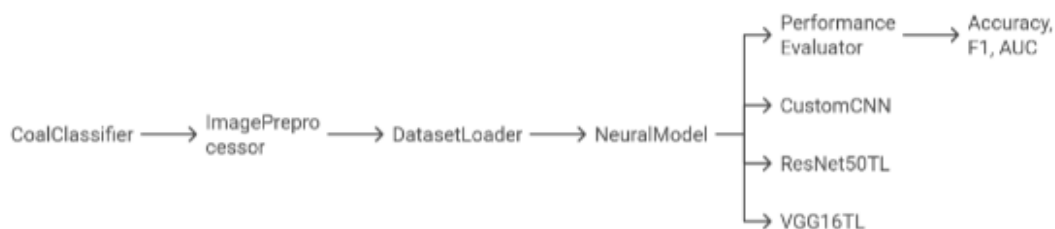


Fig.2: Working Flowchart

Preprocessing is done after the dataset split. All images are resized to 224×224 pixels, which is the size expected by the CNN and pre-trained models. The values of the pixels are normalized for increased learning stability and speed. Sometimes, contrast enhancement is added in order to make the patterns and textures visible.

During training, other transformations such as flipping, rotation, shifting and magnification are used to increase the variation and make the classifier robust.

Preprocessing is followed by training of the custom CNN, ResNet50 and VGG16 models. The custom CNN is trained from scratch, while the pre-trained models are trained using transfer learning - the base feature extraction is frozen and only the classification head is trained. This is the combination which allows making a simple comparison between learning specific for the task and the use of pre-trained features. The training is done using optimizer Adam with appropriate multiclass loss (such as categorical cross-

entropy). The accuracy and loss with respect to the number of epochs are utilized to keep track of the training process.

Every single model would generate predictions based on the validation dataset post-training. These predictions would be converted into class labels along with probability scores for evaluation purposes. The evaluation stage is not limited to a mere accuracy output only. Validation accuracy, F1 score, confusion matrix and class wise performance would be used during the evaluation stage as well. Such metrics could provide information on the accuracy rate as well as the ability to separate classes in case of imbalance.

The last stage of the workflow is the interpretation. Having the ability to interpret ability techniques like SHAP and Grad-CAM can help one to determine image regions influencing the model's prediction. It helps since, when considering deep learning models, they are perceived to be "black boxes," and therefore interpretation in the visual field helps to increase trust and scientific transparency.

The methodology shows attention regions for each model, making it possible to compare their performance. Therefore, the workflow is completed since it comprises input data preparation, learning, inference, evaluation, and explanation. This is a step-by-step approach and hence the framework will be suitable for a conference paper presentation through a methodology flowchart.

V. RESULTS AND DISCUSSION

The result section presents a comparative analysis of the proposed coal classification models using visual performance indicators. In each graph there is one specific aspect of the model performance, such as accuracy, class distribution, feature dependence and consistency of predictions.

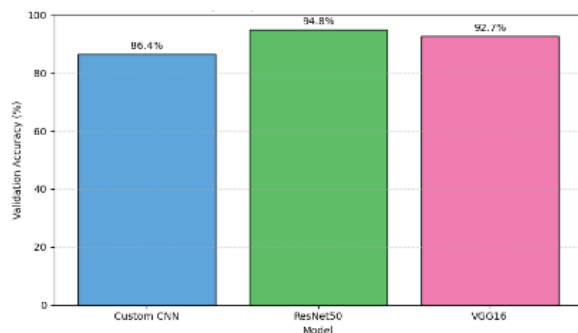


Fig.3: Validation accuracy comparison

The values shown in the graphs can be used as a good benchmark for selecting the best model architecture.

Comparison of validation accuracy of the custom CNN, ResNet50 and VGG16 is depicted in Fig. 3. Accuracy of classification by the respective model in terms of validation data is significant since it represents the sum of all other graphs. In general, for most images' classification problems, pre-trained models such as ResNet50 and VGG16 tend to be more successful in terms of validation accuracy than a custom CNN since they come with a higher-level representation of features and are pre-trained on a huge dataset. It means that the transfer learning models are more successful than the custom CNN which implies that knowledge transfer from the pre-trained network helps the network to identify the coal texture and structure better. Being competitive, the custom CNN implies that the coal data set has distinctive visual patterns for the respective task that can be learned directly. This graph is useful to make an initial level comparison but should be treated as just one measure of success since accuracy measures success in terms of learning all coal classes equally well. In case of an unbalanced data set, a model may appear to be very good because of the results of the majority class. Therefore, the graph of accuracy needs to be considered in combination with other measures.

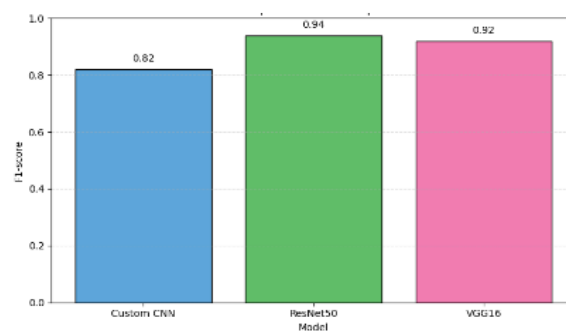


Fig.4: F1 score comparison of the model

Figure 4 shows the comparison of F1-score. While accuracy cannot combine precision and recall, F1 score is capable of doing this and hence gives us a better understanding of the model performance. This is especially true if not all classes of coal have the same amount of data. The higher the F1 score, the more accurate the prediction and the fewer the number of false positive and false negative values. This is especially critical in real life where the classification system not only needs to correctly classify the

majority of classes but also needs to accurately classify the minority class. If the F1 score of ResNet50 or VGG16 is higher than that of the custom CNN, then this clearly demonstrates that transfer learning helps to increase the learning ability for discriminative features especially of smaller classes. However, if there are any deviations in F1-score values, the possible weak sides can be traced through the chart and may relate to imbalance that cannot be identified solely on the basis of accuracy values. This way, F1-score is one of the most informative charts for the whole study since it helps to understand the validity of the model within classes. Moreover, this score makes the comparison of the three models equal.

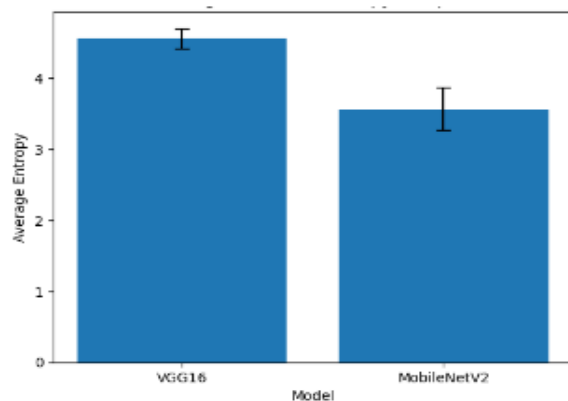


Fig.5: Multi-Image Grad-CAM entropy

In Fig. 5, we compare the mean and the standard deviation of the entropy of the Grad-CAM maps for VGG16 and MobileNetV2 in various images. A larger Entropy value will produce a map which will be more spread out and further from any form of concentration, while smaller Entropy values will lead to a more concentrated area, which is more “sure” about what area it highlights in the heatmap. The average entropy of VGG16 is larger than the one for MobileNetV2, meaning that VGG16 is spreading its attention over a larger part of the image than MobileNetV2. MobileNetV2 has a lesser level of entropy, which means that the highlights of Grad-CAM are sparser and more compact. We see that the error bars show that VGG16 also displays more consistency when it comes to attention on images, whereas there is more variability in the case of MobileNetV2. This means that when it comes to the processing of coal images, the VGG16 network can extract more texture information, while MobileNetV2 extracts more discriminative information in smaller areas.

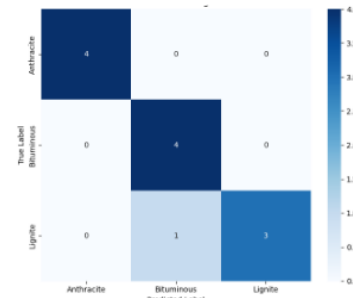


Fig.6: Confusion Matrix of the best performing coal classification model

The best performing model is shown in the confusion matrix in Fig. 6 below. This plot shows class-wise information and clearly shows the position of correctly predicted data and the regions where the classifier incorrectly predicts one kind of coal with another one. The cells in the diagonal represent the correct predictions and those off-diagonal show misclassifications. A good confusion matrix will always have dominant diagonal structure. The confusion matrix directly tells the weakness regions in case of presence of minority classes which are misclassified. This is one of the most informative plots in classification because it tells if there is a balance among all the classes or the classification model is biased towards the dominant classes. Class imbalance problem occurs in cases where there is poor recognition of the minority class and the confusion matrix is often able to give an answer to such problems especially in cases of few samples. The confusion matrix can also be used to find out why the F1-score of a class is not satisfactory while the overall accuracy of the classifier is.

VI. CONCLUSION

In this research, the coal classification process was formulated using a fully customized CNN along with the transfer learning architectures such as ResNet50 and VGG16. The pipeline consisted of stages such as data preparation, preprocessing, model training and testing, and explainability analysis, and the study showed the process to be suitable for both academia and industry.

The comparative analysis of these two types of models indicated that the pretrained model could be used to extract more information about the features, and the customized CNN architecture is considered as a baseline to learn task-specific features. Based on the

accuracy, F1 score, confusion matrix, and class-wise performance analysis, it can be said that it is not enough to use only one metric to evaluate the performance of the model on an imbalanced dataset of coals. Grad-CAM was utilized as an explainability method for this purpose. What this means is that it is essential that the process of classification of the coal be done in an accurate manner. There are various ways that this framework can be enhanced through adding to the size of the dataset and also other light-weighted architectures. In summary, the research team has successfully validated the applicability of deep learning in the process of coal classification.

REFERENCES

- [1] G. A. Ryabov, "Cofiring of coal and fossil fuels is a way to decarbonization of heat and electricity generation," *Thermal Engineering*, vol. 69, no. 6, pp. 405–417, 2022.
- [2] L. Anghelescu and B. M. Diaconu, "Advances in detection and monitoring of coal spontaneous combustion: Techniques, challenges, and future directions," *Fire*, vol. 7, no. 10, p. 354, 2024.
- [3] H. Durak, "Comprehensive assessment of thermochemical processes for sustainable waste management and resource recovery," *Processes*, vol. 11, no. 7, p. 2092, 2023.
- [4] M. R. Islam, M. Z. H. Zamil, M. E. Rayed, M. M. Kabir, M. F. Mridha, S. Nishimura, and J. Shin, "Deep learning and computer vision techniques for enhanced quality control in manufacturing processes," *IEEE Access*, vol. 12, pp. 121449–121479, 2024.
- [5] Y. Du, H. Zhang, L. Liang, J. Zhang, and B. Song, "Applications of machine vision in coal mine fully mechanized tunneling faces: A review," *IEEE Access*, vol. 11, pp. 102871–102898, 2023.
- [6] J. Yang, B. Chang, Y. Zhang, W. Luo, S. Ge, and M. Wu, "CNN coal and rock recognition method based on hyperspectral data," *International Journal of Coal Science & Technology*, vol. 9, no. 1, p. 1, 2022.
- [7] S. Thiruchittampalam, K. Shanmugalingam, B. P. Banarjee, N. F. Glenn, and S. Raval, "A comparative study with traditional and transfer learning-enhanced machine learning algorithms for geotechnical characterisation of coal spoil," *arXiv preprint arXiv:2401.01969*, 2024.
- [8] M. S. Alfarzaei, E. Hu, W. Peng, N. Qiang, and M. M. Alkainaei, "Coal gangue classification based on the feature extraction of the volume visual perception ExM-SVM," *Energies*, vol. 16, no. 4, p. 2064, 2023.
- [9] C. Sun, B. Jiao, X. Li, Y. Fu, Y. Wu, Y. Wang, et al., "Coal-enhance GAN: Low-light enhancement model for intelligent coal-rock recognition," *Earth Energy Science*, 2025.
- [10] K. Paraskevas and T. Christos, "Data preprocessing and feature engineering for data mining: Techniques, tools, and best practices," *AI*, vol. 6, no. 10, p. 257, 2025.
- [11] M. Owusu-Adjei, J. Ben Hayfron-Acquah, T. Frimpong, and G. Abdul-Salaam, "Imbalanced class distribution and performance evaluation metrics: A systematic review of prediction accuracy for determining model performance in healthcare systems," *PLOS Digital Health*, vol. 2, no. 11, p. e0000290, 2023.
- [12] R. Diallo, C. Edalo, and O. O. Awe, "Machine learning evaluation of imbalanced health data: A comparative analysis of balanced accuracy, MCC, and F1 score," in *Practical Statistical Learning and Data Science Methods: Case Studies from LISA 2020 Global Network, USA*, pp. 283–312. Cham, Switzerland: Springer Nature Switzerland, 2024.
- [13] A. W. Salehi, S. Khan, G. Gupta, B. I. Alabdullah, A. Almjally, H. Alsolai, et al., "A study of CNN and transfer learning in medical imaging: Advantages, challenges, future scope," *Sustainability*, vol. 15, no. 7, p. 5930, 2023.
- [14] N. Guo, L. Xu, W. Gao, H. Xia, M. Xie, and X. Ren, "Progress in the application of laser-induced breakdown spectroscopy in coal quality analysis," *Energies*, vol. 17, no. 14, p. 3559, 2024.
- [15] J. N. Fuhg, G. Anantha Padmanabha, N. Bouklas, B. Bahmani, W. Sun, N. N. Vlassis, et al., "A review on data-driven constitutive laws for solids: JN Fuhg et al.," *Archives of Computational Methods in Engineering*, vol. 32, no. 3, pp. 1841–1883, 2025.