Determining HEDP Foams' Quality with Multi-View Deep Learning Classification

Nadav Schneider^{1, 2}, Matan Rusanovsky^{3, 4}, Raz Gvishi⁵ and Gal Oren^{3, 6 \Box}}

¹Technology Division, Soreq Nuclear Research Center - Yavne, Israel ²Israel Atomic Energy Commission

³Scientific Computing Center, Nuclear Research Center – Negev, Israel

⁴Department of Physics, Nuclear Research Center – Negev, Israel

⁵Applied Physics Division, Soreq Nuclear Research Center - Yavne, Israel

⁶Department of Computer Science, Technion – Israel Institute of Technology, Israel

nadavsc@soreq.gov.il, matanr@nrcn.org.il, rgvishi@soreq.gov.il galoren@cs.technion.ac.il

Abstract—High energy density physics (HEDP) experiments commonly involve a dynamic wave-front propagating inside a lowdensity foam. This effect affects its density and hence, its transparency. A common problem in foam production is the creation of defective foams. Accurate information on their dimension and homogeneity is required to classify the foams' quality. Therefore, those parameters are being characterized using a 3D-measuring laser confocal microscope. For each foam, five images are taken: two 2D images representing the top and bottom surface foam planes and three images of side cross-sections from 3D scannings. An expert has to do the complicated, harsh, and exhausting work of manually classifying the foam's quality through the image set and only then determine whether the foam can be used in experiments or not. Currently, quality has two binary levels of normal vs. defective. At the same time, experts are commonly required to classify a sub-class of normal-defective, i.e., foams that are defective but might be sufficient for the needed experiment. This sub-class is problematic due to inconclusive judgment that is primarily intuitive. In this work, we present a novel state-of-the-art multi-view deep learning classification model that mimics the physicist's perspective by automatically determining the foams' quality classification and thus aids the expert. Our model achieved 86% accuracy on upper and lower surface foam planes and 82% on the entire set, suggesting interesting heuristics to the problem. A significant added value in this work is the ability to regress the foam quality instead of binary deduction and even explain the decision visually. The source code used in this work, as well as other relevant

sources, are available at: https://github.com/Scientific-Computing-Lab-NRCN/Multi-View-Foams.git.

Index Terms—HEDP, Low-Density Foams, Aerogel, Multi-View Classification, Deep Learning, LIME.

I. INTRODUCTION

A. HEDP Foams and Aerogels

High energy density physics (HEDP) experiments [1]–[5] commonly involve a dynamic wave-front propagating inside a low-density foam. This effect affects its density and hence, its transparency. The analysis of the experimental measurements required accurate information on the dimension and homogeny of the foam. Therefore, a dimension and homogeny characterization of the foam using a 3D-measuring laser confocal microscope is needed. For each foam, five images were taken:

two 2D images representing the upper surface foam plane (henceforth, 'top' and 'bottom' images) and three other images of side cross-sections from 3D-scanning images (henceforth, 'profiles' images) (Fig. 1, 2).

The foam used in the HEDP experiments is commonly aerogel. Aerogels are a large family of materials, generally defined as extremely low-density solids (more than 90% porosity, less than 200mg/cm³ density) [6], [7]. Aerogels can be composed of metals or dielectric materials and can be produced in pure, hybrid, and doped forms. For these reasons, aerogels can exhibit a diverse range of properties (chemical and physical) that could also be tailored for a specific application. Aerogels are signified by their unique properties, especially low density and high specific surface area (commonly above $\sim 600 \text{m}^2/\text{g}$), resulting in an excellent insulating material for heat, electric and acoustic. Therefore, their applications range from thermal insulators [8]-[11], acoustic insulation [12], [13], catalyst supports, electrode materials and fuel cell [14]-[16], random lasers matrices [17], space micrometeorites collectors [18], bio-medical [19], drug-delivery [20]-[22], cosmetic, lightweight magnetic actuator [23], Cherenkov radiators [24], [25] and HEDP measurements [2], [4].



Fig. 1: Illustration of the acquired microscope's images (left) from a given foam (right).



Fig. 2: Five images of an original example from the data set. From left to right: top and bottom views and three profiles.

B. Determinning HEDP Foams' Quality: Current Status

Aerogels are commonly prepared from Si-alkoxide precursors through sol-gel chemistry [26]–[29]. The gel formed is made of a few nanometers thick of fragile walls in a random structure surrounded by meso-sized pores. To dry these delicate materials, supercritical drying is necessary. Transferring the pore fluid to the supercritical phase makes it possible to vent it out with no capillary forces [30]. Thus, supercritical drying is essential to achieve dry material without a collapse of the fine porous structure. A significant limitation of dry aerogels is mechanical fragility and tending to suffer from cracks [31]-[33]. Therefore, it is required to characterize each sample for its quality as a foam. The expert searches for characteristics such as scratches, dirt, and dark stains in the top and bottom views images, implying a deep hole inside the foam. As a complementary, the profile images help confirm or refute the initial assumptions. This complicated, harsh, and exhausting work of manually classifying the foams' quality through the image set is mandatory to decide whether said foams can be used in experiments (such as HEDP). Currently, quality has binary levels of normal vs. defective. Experts are also commonly required to classify a sub-class of normal-defective, i.e., foams that are defective in some way but might be sufficient for a given experiment. This sub-class is problematic due to inconclusive judgment that is primarily intuitive. Thus, the need to devise a new, precise, explainable, consistent, and objective classification of said foams' quality is essential.

C. Suggested New Approach: Multi-View Classification

In order to mimic the physicist's intuitive perspective and automatically determine the foams' quality even in borderline cases, there is a need to devise either an intelligent or even a learned model, which will be able to self-conclude the desired classification [34], [35]. Moreover, as the image set, in this case, is diverse – both in 3D perspective (top and bottom views) and in acquisition technology (optics vs. x-ray), standard classification methodologies that do not consider perspective and domain heterogeneity will fail. Thus, the recently novel multi-view approach [36]–[38] was chosen to solve this problem correctly.

Multi-view classification models are designed for cases when the final decision depends on different features in different images [39]. Multi-view approach suggests looking over an object from several points of view (both spatially and visually), which is technically achieved by extracting unique features of the object from all the viewpoints, concatenating them into one vector, and making the decision on it. The idea has two implementation approaches: (1) Classical feature extractions and machine learning algorithms [40], [41]. (2) Deep multiview adjusted convolutional neural networks [39], [42]. In this work, we examined both of the methods as our data set is comparably small for initial learning.

1) Features Extraction and Machine Learning: Before the age of deep learning, there was a need to "manually" extract features from given images [43]. This technique aims to numerically distinguish and represent unique features of the image [43]. Several feature extractor algorithms were devised over the years for different tasks. Some of the famous ones are SIFT (Scale Invariant Feature Transform) and SURF (Speeded-Up Robust Features) [44]. The main focus of these feature extractors is extracting unique features such as scale, illumination, rotation, distortion, etc. Most feature extractors collect key points of interest in the image and produce a descriptor for each. The descriptors are usually numerical vectors with constant size.

In the multi-view approach, the descriptors are collected from the different images and grouped into centroids [40], [41] using an unsupervised model such as K-Means [45]. These centroids are like a "vocabulary of visual words" [46]. The next stage is finding visual words for each image and a vector creation that counts the number of features belonging to each centroid of visual words. The result is a (1, K) properties vector representing an image, and using this process on all the images in the data set yields a (N, K) matrix while N represents the number of images and K represents the vocabulary size. Afterward, classification is done on the matrix using a machine learning algorithm such as logistic regression, supported vector machine (SVM) [41], etc. Finally, solution adaptation into multi-view is done by creating a feature vector for all images in one example, using a feature extractor, and counting the visual words for all images (Fig. 3 A).

Nevertheless, in this work, no learning or understanding could be achieved using this method, suggesting either not enough data was collected for the data set or no sufficient learning was made. As such, there was a need for a comprehensive deep-learning approach.

2) Deep Multi-View Adjusted Convolutional Neural Networks: Following the success of convolutional neural networks in computer vision problems [47], there was an attempt to adjust these networks to multi-view classification [48]. One approach is to take the feature map of every view and stack them. For example, if every feature map is in size $K \cdot K \cdot C$, a union of the five maps is executed along the channels' dimensions, and one matrix of $K \cdot K \cdot 5C$ of feature maps is created. This approach assumes there is an order of orientation between the different views. The other approach is pooling the feature maps (view pooling), which is the chosen approach for the given problem (although orientation exists, the images do not naturally precept a single object from different angles). This approach unites all the views' features without assuming an order (Fig. 3 B).

In computer vision, when there are merely several examples similar to the given problem, as in our case, transfer learning,



Fig. 3: (A) The multi-view concept, (B) the chosen multi-view deep learning implementation approach, and (C) represents the modified Resnet in each CNN block that was intentionally reduced to prevent overfit due to a lack of data.

which uses prior knowledge accumulated from another model trained on another data set, is usually used [49]. ResNet is the selected pre-trained model for the given problem [50], trained on the ImageNet data set [51], which contains millions of high-resolution labeled images and 22 thousand different classes. This model is designed for object detection tasks using feature extraction, and with an addition of neuron layers at the end of the network, one can finetune for a given task.

II. DATA SET CREATION

While many data sets are created with strict rules and repeatable methods for an adequate examination, our data set was created solely for an expert's classification purpose. In the absence of reproducible settings for the 3D laser microscope and an arranged process of collecting the data, a considerable amount of pre-processing is required to fit the data set into a machine learning model. Furthermore, since only experts can label these examples, the labeled data is rare, and there is no real option to reach extensive labeled data. Consequently, our goal in this work was to develop an end-to-end model with the constraint of a severe shortage in data. Iteratively, we added batches of new labeled images when they were ready while designing the model as a few-shot model. This iterative process results in a data set with 95 labeled examples, each consisting of 5 images, i.e., 475 images in total.

We notice that the normal-defective labels are given when an example does not have an obvious decision for its quality. Without significant model learning, which can be achieved using lots of normal and defective examples, one can not expect to classify samples even an expert could not. Therefore, we refer to the normal-defective labels as defective labels under the strict assumption of "if there is a doubt, there is no doubt." This assumption is necessary for the given problem because we can not afford to ignore examples due to their tiny amount. However, the severe assumption above is not always correct, leading to a trade-off between the measured metrics; accuracy and AUC (Section IV). Thus, we check our model on different learning configurations (Fig. 6).

III. PRE-LEARNING PROCESS

An initial running of the multi-view deep convolutional neural networks on the data set has been done, and a quickly evolved overfit been observed. There are two common reasons for overfitting: a few training examples and an over-complicated model [52]. Therefore, solutions such as data augmentation, reduction, and pre-processing (Table I, Fig. 4) are done to overcome the few physics-guided examples issue while preserving the original and relevant properties for decision.



TABLE I: Data set images before and after pre-processing.

A. Data Augmentation

Orientation conservation is easily destroyed when creating new examples with augmentation. Nonetheless, since training deep learning models require lots of examples, an augmentation approach has been attempted. Classical augmentations such as brightness, contrast adjustments, and rotations contribute to the robustness and generalization of the model and, therefore, decrease overfit. Several rotations permutations have been chosen: -10° to 10° with a 5° jump for the profiles and 0° to

 270° with a 90° jump for the top and bottom views images. This wide range is possible due to the symmetry of the images (circles). In addition, gaussian noise, brightness, and contrast adjustments have been performed (Fig. 4).



Fig. 4: Five images of an original example from the data set after pre-process and data augmentation stages. From left to right: top and bottom views and three profiles.

B. Data Reduction and Pre-Processing

Pre-process helps focus the model on relevant features for classification. An automatic physics-guided pre-processing tool has been developed to create a proper data reduction that preserves the original and relevant properties of the foam with the following actions:

- 1) *RGB to grayscale conversion*: classification of the foams is independent of their color; therefore, the conversion prevents the model from learning irrelevant features.
- 2) Intensity quantization: a division of the pixels' values (which range from 0 to 255) to 10 bins. In many images, the black background is not an absolute black but a collection of values near zero. The model might search for a reason around these pixels' variance, and bins division prevents it.
- Circle extraction: only the central circle is relevant for classification. Extracting this circle¹ and masking the ring and other outliers helps the model focus on relevant features (Fig. 5).
- 4) *Circles bounding*: minimizing the black background solves the problem of variance in the circles' position and conserves homogeneity between the images.
- 5) Profiles centering and padding: padding black pixels for the background of profiles if necessary and centering the profiles. These operations conserve homogeneity and constant size between the profiles. Further manual cropping has been done to keep only the relevant scope.

IV. THE LEARNING PROCESS AND ITS RESULTS

In cases of a shortage of data, transfer learning is usually a decent solution. However, as mentioned above, an overcomplicated model prunes to overfit. Thus, we applied the following methods to the model to overcome the problem:

1) Simplify the model by cutting some of its last layers (Fig. 3 C). The 18 last convolutional layers have been

¹In order to focus the model on relevant features, an automatic algorithm for circle extraction in the top and bottom views images is required. Canny edge detection [53] using erosion and dilation is not enough to effectively mask the circles. Hence, a designated algorithm has been developed. The algorithm uses permutations of circles in different radii and locations around the center to find the circle with minimal bright pixels and maximal dark pixels (Fig. 5). The algorithm assumes that the background will always be brighter than the circle. Pixel's definition of bright or dark is changeable.



Fig. 5: Automatic circle extractor algorithm. Choosing the circle with minimal bright pixels and maximal dark pixels.

cut, and the network's last layer has been reduced to 128 neurons.

 Changing input dimension for 1x244x244 instead of 3x224x224 – since ResNet has been trained on a colored data set – differently from our input examples which are in grayscale – the dimensional change decreases lots of unnecessary weights.

Overall, there are 1,341,890 learnable parameters instead of 21,799,674 – about 6% of the original network.

Searching for the optimal performance, we survey 6 model configurations (Fig. 6). Specifically, three one-view models with solely top view, bottom view, and a combination of both; three multi-view models with both top and bottom views; all three profiles; and a full group of five images. Each of the 6 models was trained on the above-stated data set with and without normal-defective examples (marked in Table II and Table III as ND – normal-defected).



Fig. 6: Six learning configurations based on the different types of models and image sets.

We hereby elaborate on the relevant performance metrics (IV-A), the effect of data additions on the model performances (IV-B), the yielded experimental results (IV-C) and the possibility for explainability (IV-D).

A. Performance Metrics

Two metrics have been used for measuring the models' performance. The first is the model's accuracy in the epoch with the minimum loss (Table II). This measurement indicates the model's performance on the particular test set. Another metric that has been used is the area under curve (AUC) [54] (Table



TABLE IV: Learning curves (accuracy and loss) of all models. +ND and -ND stand for normal-defective and represent the models' training and testing with or without normal-defective examples. MV stands for a multi-view model and OV stands for a one-view model. Best models' cells are shadowed.

III). AUC measures the model's generalization capabilities and, therefore, its performance not only on a certain test set but on many. An additional noteworthy parameter is the loss graph, indicating if convergence and real learning have occurred.

B. Additional Data Effect

To understand the impact of the data set size on the resulted model performances, a comparison between two train sets has been performed: The train set mentioned above was compared to itself with subtraction of 20 examples, while the test set remains similar for both sets. This test has been performed to determine unequivocally that there is a need for continued expansion of the data set to achieve better results and that the model truthfully learns from the data and does not memorize it. The addition of 20 training examples has shown an unambiguous improvement in the models' performances – an improvement of 8% for every model's accuracy on average and a 16.5% improvement for every model's AUC on average.

C. Experimental Results

The data set split was a 70:30 train-to-test ratio while maintaining a balanced ratio between classes to prevent the model from learning from skewed data. Overall, there are 66 training examples with 23, 21, and 22 normal, normal-defective, and defective labeling correspondingly and 29 test examples with 10, 10, and 9 normal, normal-defective, and defective labeling correspondingly. The models have run on 32GB Tesla V100 GPU, using Pytorch [55]. Accuracy and AUC measures are shown in Table II and Table III, respectively. Loss and accuracy trends are presented in Table IV. Conclusively, we show that the multi-view model with top-bottom views has the best accuracy – 86%, and the multi-view model with the entire group has the best AUC – 84%.

D. Model Explainability

Physics-oriented models generally need to ensure that the outcome predictions are based on actual, relevant features and not on imaginary correlations. Thus, LIME [56] algorithm was implemented on the model predictions to study the connections between inputs and outputs. Specifically on the one-view top model as it is more intuitive to comprehend. In LIME's output, contributing areas for normal or defective decisions are marked as green or red, respectively. White stain is colored under the red area while the green area is mostly clear – both truthfully explaining the original rationale (Fig. 7).

V. CONCLUSIONS AND FUTURE WORK

While there is high accuracy for the multi-view model with top-bottom views (86%), its AUC is relatively low (71%), meaning the model managed to perform well on the test set, but its generalization capability is incomplete. Nevertheless, the loss graph shows a significant learning process without any overfit. On the other hand, the multi-view model with the entire group (full group) has the best AUC (84%), and its accuracy is relatively high as well (82%). However, there is an obvious overfit in the early epochs of the training. Most of our multi-view models show better results with the normal-defective examples in accuracy and AUC. A possible explanation may lie in the nature of normal-defective examples. Those examples are similar to normal, however, still have tiny defects. The action of labeling those examples as defective may contribute to learning these nuances and improve the classification.

For future work, we can name the apparent need for data set expansion while maintaining class balance. Also, further investigation of different pre-trained models for better transfer learning can be helpful. For example, the Inception pre-trained model [57], which has been trained on vast x-rays data set, may be more fitted for extracting features from the profiles. Finally, we suggest examining the possibility of concatenating the views' feature maps which preserves and assumes order, similarly to the given problem, instead of max-pooling them.



Original LIME's Output Fig. 7: Model's explainability using LIME. Contributing areas for normal or defective decision are green and red, respectively.

ACKNOWLEDGMENT

This work was supported by the Pazy foundation and the Lynn and William Frankel Center for Computer Science. Computational support was provided by the NegevHPC project [58]. The authors would like to thank Galit Bar⁵, Guy Lazovski⁵ and Muriel Tzadka¹ for foam samples preparation.

REFERENCES

- K. Mussack *et al.*, "Simulating hohlraum dynamics and radiation flow for pleiades experiments on nif," in *APS Division of Plasma Physics Meeting Abstracts*, vol. 53, 2011, pp. GP9–124.
- [2] M. Baer, "A numerical study of shock wave reflections on low density foam," *Shock Waves*, vol. 2, no. 2, pp. 121–124, 1992.
- [3] H. M. Johns *et al.*, "A temperature profile diagnostic for radiation waves on omega-60," *High Energy Density Physics*, vol. 39, p. 100939, 2021.
 [4] Y. Wu *et al.*, "Dynamics of laser produced plasma from foam targets
- [4] Y. Wu *et al.*, "Dynamics of laser produced plasma from foam targets for future nanolithography devices and x-ray sources," *Scientific Reports*, vol. 11, no. 1, pp. 1–14, 2021.
- [5] R. Harel *et al.*, "Complete deep computer-vision methodology for investigating hydrodynamic instabilities," in *International Conference* on High Performance Computing. Springer, 2020, pp. 61–80.
- [6] O. Figovsky et al., Green nanotechnology. Jenny Stanford Publishing, 2017.
- [7] D. W. Schaefer et al., "Structure of random porous materials: silica aerogel," *Physical review letters*, vol. 56, no. 20, p. 2199, 1986.
- [8] G. Wei et al., "Thermal conductivities study on silica aerogel and its composite insulation materials," *International Journal of Heat and Mass Transfer*, vol. 54, no. 11-12, pp. 2355–2366, 2011.
- [9] M. Reim et al., "S., m. arduini-schuster, hp ebert, j. fricke, j," Silica aerogel granulate material for thermal insulation and daylighting, Solar Energy, vol. 79, pp. 131–139, 2005.

- [10] B. Yoldas *et al.*, "Chemical engineering of aerogel morphology formed under nonsupercritical conditions for thermal insulation," *Chemistry of materials*, vol. 12, no. 8, pp. 2475–2484, 2000.
- [11] E. Zhang et al., "Insulating and robust ceramic nanorod aerogels with high-temperature resistance over 1400 c," ACS Applied Materials & Interfaces, vol. 13, no. 17, pp. 20548–20558, 2021.
- [12] S. Takeshita *et al.*, "Structural and acoustic properties of transparent chitosan aerogel," *Materials Letters*, vol. 254, pp. 258–261, 2019.
- [13] K. Wu et al., "Lightweight and flexible phenolic aerogels with threedimensional foam reinforcement for acoustic and thermal insulation," *Industrial & Engineering Chemistry Research*, vol. 60, no. 3, pp. 1241– 1249, 2021.
- [14] Y. Gu et al., "Electronic structure tuning in ni3fen/r-go aerogel toward bifunctional electrocatalyst for overall water splitting," ACS Nano, vol. 12, no. 1, pp. 245–253, Jan. 2018.
- [15] M. A. Worsley *et al.*, "Synthesis of graphene aerogel with high electrical conductivity," *Journal of the American Chemical Society*, vol. 132, no. 40, pp. 14067–14069, 2010.
- [16] Peles-Strahl et al., "Bipyridine modified conjugated carbon aerogels as a platform for the electrocatalysis of oxygen reduction reaction," Advanced Functional Materials, vol. 31, no. 26, p. 2100163, 2021.
- [17] N. U. Wetter *et al.*, "Dynamic random lasing in silica aerogel doped with rhodamine 6g," *RSC advances*, vol. 8, no. 52, pp. 29678–29685, 2018.
- [18] M. J. Burchell *et al.*, "Characteristics of cometary dust tracks in stardust aerogel and laboratory calibrations," *Meteoritics & Planetary Science*, vol. 43, no. 1-2, pp. 23–40, 2008.
- [19] N. Ganonyan *et al.*, "Entrapment of enzymes in silica aerogels," *Materials Today*, vol. 33, pp. 24–35, 2020.
- [20] G. Nir et al., "Gradual hydrophobization of silica aerogel for controlled drug release," RSC advances, vol. 11, no. 14, pp. 7824–7838, 2021.
- [21] García-González *et al.*, "Polysaccharide-based aerogel microspheres for oral drug delivery," *Carbohydrate polymers*, vol. 117, pp. 797–806, 2015.
- [22] J. Zhao et al., "Polyethylenimine-grafted cellulose nanofibril aerogels as versatile vehicles for drug delivery," ACS applied materials & interfaces, vol. 7, no. 4, pp. 2607–2615, 2015.
- [23] N. Ganonyan et al., "Ultralight monolithic magnetite aerogel," Applied Materials Today, vol. 22, p. 100955, 2021.
- [24] D. Asner et al., "Experimental study of aerogel cherenkov detectors for particle identification," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, vol. 374, no. 3, pp. 286–292, 1996.
- [25] T. Sumiyoshi et al., "Silica aerogel cherenkov counter for the kek bfactory experiment," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, vol. 433, no. 1-2, pp. 385–391, 1999.
- [26] C. J. Brinker et al., Sol-gel science: the physics and chemistry of sol-gel processing. Academic press, 2013.
- [27] R. Gvishi, "Monolith sol-gel materials," chapter in [The Sol-Gel Handbook], Eds: Levy D. and Zayat M., in publication by Wiley-VCH, 2014.
- [28] I. Smirnova *et al.*, "Aerogel production: Current status, research directions, and future opportunities," *The Journal of Supercritical Fluids*, vol. 134, pp. 228–233, 2018.
- [29] G. Lazovski et al., "A simple method for preparation of silica aerogels doped with monodispersed nanoparticles in homogeneous concentration," *The Journal of Supercritical Fluids*, vol. 159, p. 104496, 2020.
- [30] L. Guy et al., "A procedure to synthesize silica aerogels in a wide range of densities by a single-step base catalyzed recipe," *Journal of Porous Materials*, vol. 28, no. 4, pp. 1227–1236, 2021.
- [31] A. Sedova *et al.*, "Reinforcing silica aerogels with tungsten disulfide nanotubes," *The Journal of Supercritical Fluids*, vol. 106, pp. 9–15, 2015.
- [32] R. Haj-Ali *et al.*, "Mechanical characterization of aerogel materials with digital image correlation," *Microporous and Mesoporous Materials*, vol. 226, pp. 44–52, 2016.
- [33] A. Sedova *et al.*, "Silica aerogels as hosting matrices for ws2 nanotubes and their optical characterization," *Journal of materials science*, vol. 55, no. 18, pp. 7612–7623, 2020.
- [34] M. Rusanovsky *et al.*, "An end-to-end computer vision methodology for quantitative metallography," *Scientific Reports*, vol. 12, no. 1, pp. 1–27, 2022.

- [35] R. Matan et al., "Mlography: An automated quantitative metallography model for impurities anomaly detection using novel data mining and deep learning approach," arXiv preprint arXiv:2003.04226, 2020.
- [36] C. Xu et al., "A survey on multi-view learning," arXiv preprint arXiv:1304.5634, 2013.
- [37] S. Sun, "A survey of multi-view machine learning," *Neural computing and applications*, vol. 23, no. 7, pp. 2031–2038, 2013.
- [38] J. Zhao et al., "Multi-view learning overview: Recent progress and new challenges," *Information Fusion*, vol. 38, pp. 43–54, 2017.
- [39] M. Seeland *et al.*, "Multi-view classification with convolutional neural networks," *Plos one*, vol. 16, no. 1, 2021.
- [40] F. Nie et al., "Multi-view clustering and semi-supervised classification with adaptive neighbours," in *Thirty-first AAAI conference on artificial intelligence*, 2017.
- [41] L. Houthuys et al., "Multi-view least squares support vector machines classification," *Neurocomputing*, vol. 282, pp. 78–88, 2018.
- [42] M. Kan et al., "Multi-view deep network for cross-view classification," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 4847–4855.
- [43] G. Kumar et al., "A detailed review of feature extraction in image processing systems," in 2014 Fourth international conference on advanced computing & communication technologies. IEEE, 2014, pp. 5–12.
- [44] L. Juan et al., "A comparison of sift, pca-sift and surf," International Journal of Image Processing (IJIP), vol. 3, no. 4, pp. 143–152, 2009.
- [45] J. A. Hartigan *et al.*, "Algorithm as 136: A k-means clustering algorithm," *Journal of the royal statistical society. series c (applied statistics)*, vol. 28, no. 1, pp. 100–108, 1979.
- [46] S. Tafasca, "Multi-view image classification: From logistic regression to multi-view convolutional neural networks (mvcnn)," https://toward sdatascience.com/multi-view-image-classification-427c69720f30, December 24, 2019, [Online].
- [47] S. Khan *et al.*, "A guide to convolutional neural networks for computer vision," *Synthesis lectures on computer vision*, vol. 8, no. 1, pp. 1–207, 2018.
- [48] H. Su et al., "Multi-view convolutional neural networks for 3d shape recognition," in Proceedings of the IEEE international conference on computer vision, 2015, pp. 945–953.
- [49] K. Weiss et al., "A survey of transfer learning," Journal of Big data, vol. 3, no. 1, pp. 1–40, 2016.
- [50] K. He et al., "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.
- [51] O. Russakovsky et al., "Imagenet large scale visual recognition challenge," International journal of computer vision, vol. 115, no. 3, pp. 211–252, 2015.
- [52] S. Kadam et al., "Review and analysis of zero, one and few shot learning approaches," in *International Conference on Intelligent Systems Design* and Applications. Springer, 2018, pp. 100–112.
- [53] J. Canny, "A computational approach to edge detection," *IEEE Transactions on pattern analysis and machine intelligence*, no. 6, pp. 679–698, 1986.
- [54] C. Marzban, "The roc curve and the area under it as performance measures," *Weather and Forecasting*, vol. 19, no. 6, pp. 1106–1114, 2004.
- [55] A. Paszke et al., "Pytorch: An imperative style, high-performance deep learning library," Advances in neural information processing systems, vol. 32, 2019.
- [56] M. T. Ribeiro et al., ""why should i trust you?": Explaining the predictions of any classifier," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016, pp. 1135–1144.
- [57] C. Szegedy et al., "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Thirty-first AAAI conference on* artificial intelligence, 2017.
- [58] R. I. Park, "NegevHPC Project," https://www.negevhpc.com, 2019, [Online].