Parameters Correlation of Deep Learning Model for Visual Inspection in the Automotive Industry

Charbel El Hachem
FEMTO-ST Institute, CNRS
Univ. Bourgogne Franche-Comté (UBFC)
Belfort, France
charbel.el_hachem@univ-fcomte.fr

Anass El Houd
FEMTO-ST Institute, CNRS
Univ. Bourgogne Franche-Comté (UBFC)
Belfort, France
anass.elhoud@femto-st.fr

Loïc Painvin
R&D center
Faurecia Clean Mobility
Bavans, France
loic.painvin@faurecia.com

Gilles Perrot
FEMTO-ST Institute, CNRS
Univ. Bourgogne Franche-Comté (UBFC)
Belfort, France
gilles.perrot@univ-fcomte.fr

Raphaël Couturier
FEMTO-ST Institute, CNRS
Univ. Bourgogne Franche-Comté (UBFC)
Belfort, France
raphael.couturier@univ-fcomte.fr

***Abstract*—Welding seams' visual inspection is manually operated by humans in different fields, which makes the result of inspection subjective and expensive. At present, the integration of deep learning methods for welds classification is a research focus in engineering applications. This work intends to apprehend and emphasize the contribution of deep learning model explainability and the improvement of welding seams classification accuracy: two of the main metrics affecting production lines and cost in the automotive industry.**
This paper introduces a novel hybrid method that relies on combining the model prediction scores and visual explanation heatmap. The results reaches higher accuracy than a traditional deep learning model, by at least 18,2%. New perspectives in explaining and interpreting deep learning models are presented in this contribution.

***Index Terms*—Visual Inspection, Industry 4.0, Deep Learning, Model Explainability Heatmap, Hybrid Classification**

I. Introduction

Welding is a manufacturing process consisting in joining two or more elements in a permanent way while ensuring continuity between these elements [1]. The assembly is done either by heating, by pressure, or by the two combined.
Welding defects affect manufacturing plants, in which Faurecia is a major actor on a global scale. Currently, there is a lot of research work being done on the automation of quality control [2], [3] and visual inspection of welding seams [4]. Automatic systems has been developed based on artificial intelligence (AI) to check the quality of welding seams in plants. These systems are evaluated solely by their accuracy, the percentage of true positive images and true negative images. However, they have some shortcomings and errors in the classification of welding seam quality even when their accuracy is high [4]. Model interpretation is becoming a primary evaluation metric as well as its performance [5]. A compromise between explainability and accuracy is more and more necessary in industrial applications.
The proposed approach in this study consists of:

1) Displaying heatmap of a deep learning model trained with tensorflow.
2) Correlating heatmap results and traditional deep learning classes' scores.
3) Proposing a hybrid approach: integrating machine learning classifiers in order to improve overall accuracy.

This paper is organized as follows: Section II introduces the problem statement and motivation to set up this study. Section III presents an overview of the existing work done in the area of welding seam inspection and model explainability. Section IV displays the novel hybrid approach. In Section V, the implementation details are mentioned. Section VI examines the experimental results. The last section gives the concluding remarks.

II. Problem Statement

A. Deep learning accuracy calculation

Since weld defects are not always present in production, images from both OK and Not OK (NOK) classes will not have the same occurrence, so this is an unbalanced dataset problem [6].
Welding seam's classification was initially solved by applying data augmentation techniques [7]: a deep learning model based on MobileNet architecture has been trained on a set of images corresponding to four welding seams. Figure 1 shows how each welding seam's dataset is classified.

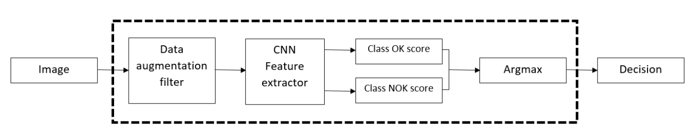

Fig. 1. Decision-making approach for MobileNet architecture

Supported by Faurecia

The proposed solution has reached the client's requirements: 97% of NOK predicted as NOK on weld1 and weld4. The challenge remains present on weld2 and weld3 with an average of 57,5% and 80,5% of accuracy reached respectively. Hence the need to look for other methods to overcome this problem. Table I displays the best accuracy results, obtained when applying data augmentation techniques on MobileNet architecture.

TABLE I
BEST RESULTS REACHED WITH DATA AUGMENTATION TECHNIQUES ON WELD1, WELD2, WELD3 AND WELD4

Weld Number	Data augmentation filter	Accuracy
Weld1	Box blur	98,5%
Weld2	Flip	57,5%
Weld3	Rotation	80,5%
Weld4	No data augmentation	96,5%

B. Model Reliability

Model reliability is considered as an important criterion when choosing the best model [8]–[10]: even if the accuracy target is achieved, the model might use a biased part of the image to classify it which causes a drop of accuracy when faced with new data.
Heatmap visualization is a graphical representation of numerical data. It uses matrices where each cell represents the intensity of the studied event. In deep learning, heatmap is used to represent the weights distribution relative to the model's decision-making [11]. Warm colors represent high weights in the deep learning model while cool colors represent low-value weights.

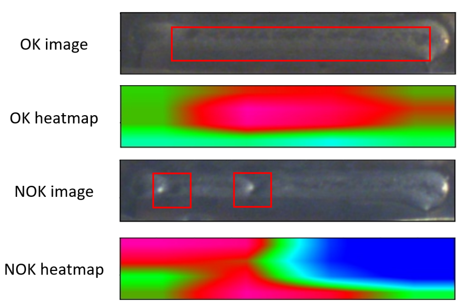

Fig. 2. Heatmap visualization of weld2 (OK and NOK images)

Figure 2 shows an example of OK/NOK images and their relative heatmap visualization. The red square in 'OK image' is the zone of interest on which the model should be based on when classifying an OK image. The 'OK heatmap' shows the section on which the model relied on for classification (warm colors sections). 'NOK image' shows that there are two defects in the NOK image (inside the red squares). While the 'NOK heatmap' shows that the warm colors are not correctly distributed: the warm colors are not super-imposed with the zone of interest displayed in 'NOK image'. This means that the model used a biased zone of interest in its decision-making. Thus, this NOK image has been a misclassified weld by the model.
Although the model has good overall accuracy on the dataset, it is necessary to study the reliability of the model: in addition to its accuracy, a possible correlation between the activation heatmap results and the model classes' scores can be identified.

III. RELATED WORKS

A. Visual Explanation Heatmap

The explainability of the AI model reinforces the credibility of the obtained results and allows to evaluate the reliability of the model and its behavior if a partial change occurred in data [12]–[14]:
Selvaraju et al. [15] have first proposed the Grad-CAM method, a visualization technique able to tell which parts of a given image led the trained convolutional neural network (CNN) to its final classification decision. This method makes it easy to debug the decision process of a CNN, especially in the case of misclassification. The result of this method is an activation heatmap indicating the parts of the image that have contributed the most to the final decision of the network.
In the same context, Zhang et al. [12] have applied model explainability to interpret the deep learning models trained to classify multiple sclerosis types in the brain using the Grad-CAM method. The experimental results showed that Grad-CAM gives the best heatmap localizing ability, and CNNs with a global average pooling layer and pre-trained weights had the best classification performance.
The Grad-Cam method has proven its ability in explaining depp learning model's decision-making. This method will be added to MobileNet architecture for accuracy's further improvement.

B. Hybrid Methods

Ahlawat et al. [16] have implemented a hybrid method using CNN and Support-vector machine (SVM) classifier for handwritten digit recognition: CNN works as an automatic feature extractor and SVM works as a binary classifier. Their results showed that the hybrid approach achieved an accuracy of 99.28% on the MNIST dataset.
Soumaya et al. [17] have tested a hybrid classification model using a genetic algorithm and SVM to detect Parkinson's disease. Their method attempts to give an accuracy of 80% and 72.50% using two kernels of SVM. The hybrid method seems to ensure the optimization of the classification system by minimizing the dimension of the features vector and maximizing the accuracy.
In the same context, Ahammad et al. [18] have suggested a new CNN-deep segmentation-based boosting classifier for spinal cord injury prediction. This method gives 10% improvement on the classification rate.
Another hybrid method is developed by Liu et al. [19] for CO_2 welding defects detection by using CNN for primary features extraction and long short-term memory (LSTM) for feature fusion. The algorithm reaches 94% of prediction accuracy.

Hybrid methods can reach higher accuracy when the right parameters are correlated together. In this contribution, traditional classes' scores will be combined with the red color ratio displayed in heatmaps.

IV. Methods

A. Model Explainability

The main approach of this work is based on the Grad-CAM method, introduced by Selvaraju et al. [15]. This method assigns importance to each position in the last convolutional layer by producing a coarse localization map of important regions in the image. It computes the linear combination of activations, weighted by the corresponding output weights of the predicted class. The resulting class activation mapping is then resampled to the size of the input image. Grad-cam allows to validate that the deep learning model is looking at the correct patterns.

B. Heatmap Analysis

The calculation of the Red Color Ratio (RCR) is done using K-nearest neighbor (KNN). This algorithm identifies the pixels' closest color in the heatmap with K = 3 representing the RGB color system. The Red Color Ratio is calculated as follows:

$$RCR = \frac{\text{Number of Red Cluster pixels}}{\text{Total number of pixels}} \times 100$$

Figure 3 shows the distribution of classes' scores related to weld3 images. Each blue dot is a defective weld (NOK image), and the orange triangle is a non-defective weld in reality (OK image). Defining a threshold in order to separate both clusters is not possible with classes' scores.

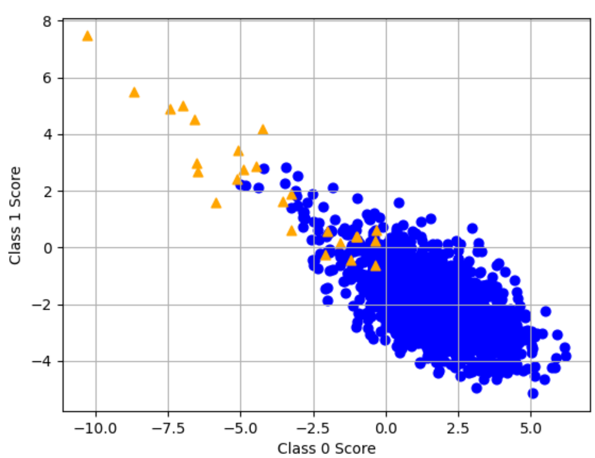

Fig. 3. Distribution of weld3 images per classes' scores

A better visualization of both clusters is represented in Figure 4: correlating the Red Color Ratio (RCR) parameter with classes' scores. This correlation offers better visibility on the cluster's distribution, compared to Figure 3.

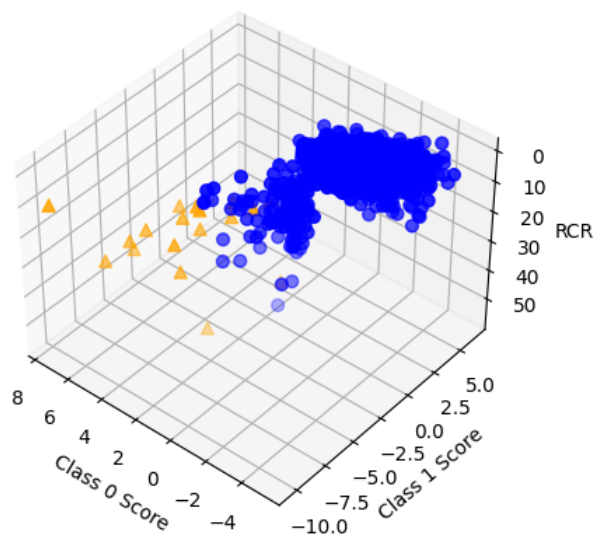

Fig. 4. Correlating classes' scores with Red Color Ratio for weld3 images

This correlation will be used for the rest of this study. Having numerical input data, a machine learning classifier should be used to assign a class label to these data inputs.

C. Machine Learning Classifiers

Machine learning classifiers have been tested for the improvement of the decision-making:

1) XGboost Classifier: a gradient boosting algorithm that offers a panel of hyperparameters. It is possible to have total control over the implementation of Gradient Boosting [20]. The chosen booster is "gbtree" which uses a tree-based boosting. The step size shrinkage is set to 0,5 and the used sampling method is the uniform selection.
2) The decision tree uses the Gini function to measure the quality of the split. Gini index measures the probability of an observation [21]. Used to identify the degree of a particular variable being wrongly classified when it is randomly chosen. The max depth is set to the default value so that the tree nodes are expanded until all leaves are pure or contain less than 2 elements.
3) SVM classifiers: machine learning algorithms that solve classification and regression problems; known for their strong theoretical guarantees and their great flexibility. SVM projects data into a higher-dimensional space and makes them separable. It becomes a universal approximator [22]: with enough data, the algorithm can always find the best possible boundary to separate two classes. In this study, two SVM kernels are used:
   - Linear Kernel: This is the case of a linear classifier, without space change. The data are assumed to be linearly separable.
   - 5-Polynomial Kernel: This is the case where a polynomial transformation is applied in order to

change the space. In this study, polynomial degree is equal to five.

Adding a machine learning classifier changed the way each welding seam has been classified (previously presented in Figure 1). This paper's proposed classification is presented in Figure 5.

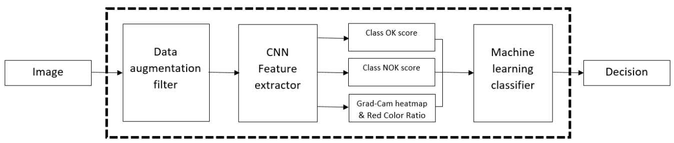

Fig. 5. Proposed decision-making approach: adding Grad-Cam heatmap and Red Color Ratio

V. Implementation Details

A. Data Collection

The dataset used in the paper [4] has been collected in order to improve the accuracy of detecting defective parts. The client's requirements is 97% of weld defects detection. Weld2 and weld3 haven't reached the target. Many data augmentation filters have been applied on these two welds. One dataset has been chosen for each of these welds, based on the potential improvement of their accuracy, as below:

1) Flip filter has been selected for weld2: previously reached 15% on OK images and 100% on NOK images (an average of 57,5%).
2) Rotation filter has been selected for weld3: previously reached 93% on OK images and 68% on NOK images (an average of 80,5%).

B. Experimental Environement

The experimental environment is powered by Intel i5 CPU, 2.30 GHz with 64-bit, Windows 10 system and 8 GB memory. The software programming environment is Python. It uses both Keras and Tensorflow as backend. Rmsprop is selected as the optimizer of the MobileNet model. The chosen learning rate decay type is exponential starting with a value of 0.01 and ending with a value equal to 0.0001. The model has been trained with 9000 epochs.

VI. Experimental Results

As in Table II, four of the machine learning classifiers have been applied following the proposed approach.
Weld2 reaches an overall accuracy of:

1) 99,3% with XGBoost
2) 98,1% with Decision Tree
3) 98,8% with SVM linear and SVM Poly5

Compared to the traditional approach, weld2's accuracy is improved by +41,8%.
Weld3 reaches an overall accuracy of:

1) 98,7% with XGBoost and SVM Poly5
2) 97,2% with Decision Tree
3) 98,1% with SVM linear

Compared to the traditional approach, weld3's accuracy is improved by +18,2%.

TABLE II
Proposed decision making results when applied on weld2 and weld3

Welding Seams	Accuracy per Classifier			
	XGboost	***Decision Tree***	***SVM Linear***	***SVM Poly5***
Weld 2	**99,3%**	98,1%	98,8%	98,8%
Weld 3	**98,7%**	97,2%	98,1%	**98,7%**

These results prove that adding a statistical machine learning classifier after the feature extractor and class activation heatmap does increase the overall accuracy of the model for both welds. The target accuracy is reached, with a tiny advantage to XGboost classifier.

VII. Conclusion & Future directions

In this paper, a hybrid approach of CNN-Machine Learning Classifier is proposed for welds defects classification. This approach adds a new reliability score calculated using the Grad-CAM heatmap.
The hybrid approach proposed in this paper reaches high accuracy on weld defects classification. The highest accuracy improvement was by +41,8% for weld2 using MobileNet-XGboost classifier and by +18,2% for the weld 3 using MobileNet-SVM Poly5 Kernel or MobileNet-XGboost classifier.
This work presents new model-driven optimization methods to improve the accuracy of vision systems. In future work, this approach can be applied on other deep learning architectures to validate its efficiency.

Acknowledgment

This work was done as a part of a CIFRE (N 2018/1029) project with Faurecia, funded by the Ministry of Higher Education and Research of France, managed by the Association Nationale de la Recherche et de la Technologie (ANRT) and was partially supported by the EIPHI Graduate School (contract "ANR-17-EURE0002").
The first author would like to thank Faurecia plant's staff for their availability.

References

[1] Baicun Wang, S Jack Hu, Lei Sun, and Theodor Freiheit. Intelligent welding system technologies: State-of-the-art review and perspectives. *Journal of Manufacturing Systems*, 56:373–391, 2020.

[2] Ernie Mazuin Mohd Yusof, Mohd Ismail Yusof, Rafidah Ali, Izzat Hilmi Harjimi, and Qasidah Kamarul Bahrin. Welding station monitoring system using internet of thing (iot). *Indonesian Journal of Electrical Engineering and Computer Science*, 18(3):1319–1330, 2020.

[3] Raoul Schönhof, Artem Werner, Jannes Elstner, Boldizsar Zopcsak, Ramez Awad, and Marco Huber. Feature visualization within an automated design assessment leveraging explainable artificial intelligence methods. *Procedia CIRP*, 100:331–336, 2021.

[4] Charbel El Hachem, Gilles Perrot, Loïc Painvin, and Raphaël Couturier. Automation of quality control in the automotive industry using deep learning algorithms. In *2021 International Conference on Computer, Control and Robotics (ICCCR)*, pages 123–127. IEEE, 2021.

[5] Wojciech Samek, Thomas Wiegand, and Klaus-Robert Müller. Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. *arXiv preprint arXiv:1708.08296*, 2017.

[6] Petr Mrozek, John Panneerselvam, and Ovidiu Bagdasar. Efficient resampling for fraud detection during anonymised credit card transactions with unbalanced datasets. In *2020 IEEE/ACM 13th International Conference on Utility and Cloud Computing (UCC)*, pages 426–433. IEEE, 2020.

[7] Charbel El Hachem, Gilles Perrot, Loïc Painvin, Jean-Baptiste Ernst-Desmulier, and Raphaël Couturier. Welding seam classification in the automotive industry using deep learning algorithms. In *2021 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*, pages 235–240. IEEE, 2021.

[8] Alex John London. Artificial intelligence and black-box medical decisions: accuracy versus explainability. *Hastings Center Report*, 49(1):15–21, 2019.

[9] Basemah Alshemali and Jugal Kalita. Improving the reliability of deep neural networks in nlp: A review. *Knowledge-Based Systems*, 191:105210, 2020.

[10] Yang Liu, Zelin Zhang, Xiang Liu, Lei Wang, and Xuhui Xia. Performance evaluation of a deep learning based wet coal image classification. *Minerals Engineering*, 171:107126, 2021.

[11] Omdena. https://omdena.com/blog/heatmap-machine-learning/.

[12] Yunyan Zhang, Daphne Hong, Daniel McClement, Olayinka Oladosu, Glen Pridham, and Garth Slaney. Grad-cam helps interpret the deep learning models trained to classify multiple sclerosis types using clinical brain magnetic resonance imaging. *Journal of Neuroscience Methods*, 353:109098, 2021.

[13] Yusuke Osawa, Keiichi Watanuki, Kazunori Kaede, and Keiichi Muramatsu. Visualization of features in multivariate gait data: Use of a deep learning for the visualization of body parts and their timing during gait training. In *International Conference on Applied Human Factors and Ergonomics*, pages 1007–1013. Springer, 2020.

[14] Jennie Crosby, Sophia Chen, Feng Li, Heber MacMahon, and Maryellen Giger. Network output visualization to uncover limitations of deep learning detection of pneumothorax. In *Medical Imaging 2020: Image Perception, Observer Performance, and Technology Assessment*, volume 11316, page 113160O. International Society for Optics and Photonics, 2020.

[15] Ramprasaath R Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision*, pages 618–626, 2017.

[16] Savita Ahlawat and Amit Choudhary. Hybrid cnn-svm classifier for handwritten digit recognition. *Procedia Computer Science*, 167:2554–2560, 2020.

[17] Zayrit Soumaya, Belhoussine Drissi Taoufiq, Nsiri Benayad, Korkmaz Yunus, and Ammoumou Abdelkrim. The detection of parkinson disease using the genetic algorithm and svm classifier. *Applied Acoustics*, 171:107528, 2021.

[18] Sk Hasane Ahammad, V Rajesh, Md Zia Ur Rahman, and Aimé Lay-Ekuakille. A hybrid cnn-based segmentation and boosting classifier for real time sensor spinal cord injury data. *IEEE Sensors Journal*, 20(17):10092–10101, 2020.

[19] Tianyuan Liu, Jinsong Bao, Junliang Wang, and Yiming Zhang. A hybrid cnn–lstm algorithm for online defect recognition of co2 welding. *Sensors*, 18(12):4369, 2018.

[20] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. *CoRR*, abs/1603.02754, 2016.

[21] quantinsti. https://blog.quantinsti.com/gini-index/.

[22] Junping Wang, Quanshi Chen, and Yong Chen. Rbf kernel based support vector machine with universal approximation and its application. In *International symposium on neural networks*, pages 512–517. Springer, 2004.