



Article

An Objective Metallographic Analysis Approach Based on Advanced Image Processing Techniques

Xabier Sarrionandia ^{1,2}, Javier Nieves ^{1,*} , Beñat Bravo ¹, Iker Pastor-López ²  and Pablo G. Bringas ² 

¹ AZTERLAN, Basque Research and Technology Alliance (BRTA), Aliendale Auzunea 6, 48200 Durango, Spain

² Department of Mechanics, Design and Industrial Management, Faculty of Engineering, University of Deusto, Unibertsitate Etorbidea 24, 48007 Bilbao, Spain

* Correspondence: jnieves@azterlan.es

Abstract: Metallographic analyses of nodular iron casting methods are based on visual comparisons according to measuring standards. Specifically, the microstructure is analyzed in a subjective manner by comparing the extracted image from the microscope to pre-defined image templates. The achieved classifications can be confused, due to the fact that the features extracted by a human being could be interpreted differently depending on many variables, such as the conditions of the observer. In particular, this kind of problem represents an uncertainty when classifying metallic properties, which can influence the integrity of castings that play critical roles in safety devices or structures. Although there are existing solutions working with extracted images and applying some computer vision techniques to manage the measurements of the microstructure, those results are not too accurate. In fact, they are not able to characterize all specific features of the image and, they cannot be adapted to several characterization methods depending on the specific regulation or customer. Hence, in order to solve this problem, we propose a framework to improve and automatize the evaluations by combining classical machine vision techniques for feature extraction and deep learning technologies, to objectively make classifications. To adapt to the real analysis environments, all included inputs in our models were gathered directly from the historical repository of metallurgy from the Azterlan Research Centre (labeled using expert knowledge from engineers). The proposed approach concludes that these techniques (a classification under a pipeline of deep neural networks and the quality classification using an ANN classifier) are viable to carry out the extraction and classification of metallographic features with great accuracy and time, and it is possible to deploy software with the models to work on real-time situations. Moreover, this method provides a direct way to classify the metallurgical quality of the molten metal, allowing us to determine the possible behaviors of the final produced parts.

Keywords: artificial vision; machine learning; deep learning; metallography; classification



Citation: Sarrionandia, X.; Nieves, J.; Bravo, B.; Pastor-López, I.; Briñas, P.G. An Objective Metallographic Analysis Approach Based on Advanced Image Processing Techniques. *J. Manuf. Mater. Process.* **2023**, *7*, 17. <https://doi.org/10.3390/jmmp7010017>

Academic Editor: Steven Y. Liang

Received: 18 November 2022

Revised: 21 December 2022

Accepted: 27 December 2022

Published: 4 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Metallurgy (as a process of production and transformation of materials) has allowed society to evolve. More specifically, the foundry remains one of the central axes of the world economy [1]. A huge number of parts are made in foundries all over the world to combine and create a more complex system. Some of those parts are security components used in several industries, e.g., brake calipers that help the braking systems of motorized vehicles, the propellers that allow ships to move, the mechanisms that are in charge of moving the flaps of the wings in an airplane, or the trigger and the firing system in a firearm.

The foundry, despite being a fundamental axis of society, is still at a lower level of development in terms of digitalization and the application of advanced intelligent systems compared to other industries of similar importance. In addition, current trends encourage the production of ever smaller and more precise components. Thus, any tiny aspect or characteristic of the process can influence the results of the final manufactured parts.

The foundry is the place where the process of transforming metals into castings takes place. In other words, the operators introduce the already molten metals into a mold. Then, once the cooling process is finished, the final casting is obtained. Specifically, Pattnaik et al. [2] detailed the process on which this research work is focused on (a process in which molten metals are poured into molds, where there are other related tasks, such as the preparation of those molds or cores, and the finishing tasks of the manufactured parts). Since the two most influential elements are (primarily) the metal and the mold, improvements to any of these will positively affect the final result.

This research paper focuses on analyzing one of the two key topics, specifically, metal. As it was registered in the Foundry Technology book [3], we can highlight the following types of iron castings.

- *White cast iron* shows that all of the carbon is combined in the form of cementite. This type of manufacturing offers great hardness and fragility. Furthermore, it provides great resistance to wear and abrasion [4].
- In *gray cast iron* [5], in contrast to what happens in white cast iron, the carbon is in the form of sheets. This type of cast iron has good resistance to compression, and its mechanical properties are suitable for a large number of activities.
- Lastly, in *nodular cast iron* (also known as ductile iron or spheroidal graphite) [6], carbon crystallizes in the form of spherical nodules, inhibiting the creation of cracks by releasing the matrices of stress points. This type of cast iron has much more hardness and resistance to fatigue. For this reason, it is especially useful for elements that require high resistance, such as safety components for the automotive industry.

The differences between the aforementioned types of iron, as well as the extraction of information to determine the characteristics of the metal, can be identified through the analysis of the microstructure. More accurately, this work was conducted under some regulations, such as UNE EN ISO 945-1 [7], UNE EN ISO 945-2 [8], and UNE EN ISO 945-4 [9] where the manner of proceeding (in which the characteristics must be analyzed and the methodology must be conducted) is explained. This activity, taking into account these regulations, is carried out by extracting metallography, in other words, images with microscopes, and evaluating them. Although there are several regulations, all work was conducted through subjective methods, i.e., after the generation of the sample and the extraction of the metallography was accomplished; a comparison was carried out based on the knowledge of the person who performed it. As this work depends on a human being, the classification and the analysis could be affected by many human factors [10], such as (i) fatigue, (ii) the predisposition of obtaining a result that is expected, or (iii) the error in the accuracy of the evaluation. Therefore, it is needed to find a new and different approach to conduct it, avoiding these difficulties.

As a first approach to carrying out this type of analysis and removing the objective way of doing it, there are several tools to automate (or at least facilitate) it. Specifically, there are different software tools, such as ImageJ [11], AmGuss [12], or software that comes with microscopes [13], attempting to avoid this type of subjective analysis. Despite the progress made, these tools have several problems associated with the way they perform these analyses. Some of those problems that we discovered in our daily work are the following:

1. The classification is based solely on mathematical formulae, leaving out the ones that do not meet the criteria. The software extracts some characteristics from images using computer vision and they apply formulas defined in the regulations they include.
2. Subjectivity is still present even if measurements were objectively made.
3. The software implements regulations that, perhaps, are not needed for the customer or by a certification procedure. For instance, AmGuss only includes DIN EN ISO 945 and ASTM-A247. Sometimes customers ask for custom calculations that are not possible using this software.

4. The analysis software mainly focuses on analyzing the image to extract features and characteristics, but they do not accurately determine the final quality of the iron that was manufactured. Hence, new functionality needs to be added to them.
5. In the end, these solutions are not believed to be deployed for real-time analyses in manufacturing plants.

Along with these software applications, there are some investigations and approaches that also use advanced artificial intelligence and machine vision techniques to solve very similar problems. Currently, artificial intelligence based on deep learning (DL) is opening up new horizons in image analysis research and promises to revolutionize the microscopy field. For instance, there are recent developments in DL applied to microscopy with the aim of classifying and detecting anatomical elements in the wood [14], helping in cancer detection [15], or illustrating the techniques with images from microscopes [16]. DL is widely used for classifying and detecting defects in extracted images [17] and for metrology measurements, such as in this specific work focusing on semiconductors [18].

On the one hand, these approaches are the 'axes' of believing that they could be used for metallography. However, to our knowledge, there is no research work using these techniques for metallography. In this case, related to the microstructure of metal materials, the work presented by Rodrigues et al. [19] focused on the classification of the different types of cast iron (i.e., gray, white, or ductile). The authors extracted a series of characteristics that they used as inputs into different statistical classifiers, such as K-nearest neighbors [20], K-means [21], Bayesian networks [22], support vector machines [23], and artificial neural networks [24]. The research achieved outstanding results, obtaining high accuracy, and proving that these techniques could be helpful for the analyses.

Against this background and attempting to fight the five aforementioned problems in the current software, we propose a specific approach that, using well-known algorithms and solutions, allows us to advance the current state of software solutions. Moreover, thanks to AI methods instead of specific formulas extracted from regulations, our approach can easily change the regulation or create a new custom one. As most of the world's safety parts are made from ductile iron, we focused our research on the analysis and extraction of characteristics of this type of foundry casting.

To accomplish this objective, we used a "Divide et Impera" methodology [25], splitting the original problem into a sequence of more affordable tasks, namely, (1) characterizing the sizes of the nodules, (2) determining the shapes of the nodules, and (3) performing a characterization of the metallurgical quality.

First, regarding nodule characterization aspects, a series of analyses based on usual machine vision techniques were carried out to calculate the size of each graphite nodule in relation to the whole part. Second, for the classification of the shapes of the nodules, three deep learning methods were tested and contrasted. Two of them are the widely known VGG16 [26] and VGG19 [27] networks, which, starting from a pre-trained set with a large image data collection, can be customized by learning the input layer and the output layer over the specific problem to solve. The last one is a custom DL model developed specifically for this research work (a classification pipeline using a conditional combination of binary convolutional neural networks), which was able to obtain better results than previous well-known methods. Finally, the determination of the quality of the metal was carried out through the calculation of similarity between the characteristics of the metallography, the classification of the shapes of the nodules achieved with our classification pipeline, and the representative centroids of the different classes to be categorized. Those centroids were created using the extracted knowledge from several metallurgical engineers from the Azterlan Research Centre.

The remainder of this paper is organized as follows. In Section 2, we describe the work and the steps carried out to achieve the creation. Moreover, this section discusses the data generation and methodology. Section 3 describes the results that were achieved with the proposed approach. Here, we will introduce the metrics needed for all developed models as well as the final solution created for the foundries. Finally, Sections 4 and 5 discuss the

solution, future work, and improvements that can be made to our proposal. In addition, we conclude the article and close the research introduced here.

2. Materials and Methods

2.1. Methodology and Evaluation Methods

Traditional metallurgical analyses in ductile cast iron pieces, as we explained in Section 1, are based on visual comparisons, and this interpretation can be turned into a supervised learning problem to overcome this situation. Furthermore, some recent approaches, such as in [28,29], in addition to previously cited ones, made this activity more precise, applying artificial vision techniques to obtain data from images. Thus, we propose combining machine vision tasks with DL classifications to resolve already-explained problems.

The main goals and challenges of the microstructure inspection are in the determination of cast iron properties and verifying metallurgic quality. This is a complex problem that takes into account, first, numerical data, similar to the work Amit Sata et al. conducted [30]. In this research work, the authors resolved defect detection when managing numerical data through the employment of Bayes inference. Moreover, image shapes, such as in the work by Praveen Damacharla et al. [31], involved automating steel surface defect detection using deep learning.

Therefore, we propose combining machine vision tasks with DL classifications to resolve the aforementioned problems. However, due to its complexity, the best way to achieve all of these purposes is through a simplification of the general problem using the well-known “Divide and Conquer” methodology. This methodology reduces the initial problem into small and more affordable challenges that, by solving and combining them, will lead to a global solution to the initial problem. This methodology is widely used to deal with legal issues [32], mathematical calculations [33], and computer science problems (specifically in parallel processing) [34]. Hence, and considering this idea, the steps defined in the research and development of our solution are detailed below:

1. *Identification of the problem and the challenges to overcome.* The purpose of this first step is to extract the background and context of the problem to be solved. In other words, we work to become aware of what we are attempting to solve.
2. *Acquisition of knowledge.* The second step involves the acquisition of knowledge at a high level, providing the general vision necessary to start the investigation. Later, when we are working on a much more specific topic about the challenge to be faced, we will study that topic in a more specific way.
3. *Division in challenges.* Based on the idea of the divisions already mentioned, in this step, we define the challenges faced. In our work, the following challenges were identified: (i) data extraction, (ii) handling nodule sizes, (iii) handling nodule shapes, and (iv) metal quality classification. For each of them, the following sub-phases were carried out:
 - (a) Acquisition of specific knowledge. Once the topic was defined, in this stage, we increased the knowledge to solve this problem. Many times, this acquisition was directly related to the exploration and learning of the production process that we were optimizing.
 - (b) Definition of the experiment and the techniques used. At this point, the specific research and experiments were defined for each of the challenges faced.
 - (c) Evaluation. At this stage, we carried out the defined experiment and obtained the results of the approximation that was defined.
 - (d) Analysis. Once the previous stage was complete, an analysis process was carried out on the data collected during the specific experimentation designed.
 - (e) Interpretation of the results. When a solution for each identified challenge was created, we combined all of them, coming up with a final interpretation based on all of those results.

In some of the identified challenges, the use of ML classifiers is needed. In this way, these classifiers are tested against previously labeled data, in other words, by applying a supervised learning evaluation. The outcome is characterized by the accuracy values, where the performances of the models can be evaluated. Similar performing models vary in how they fail; thus they need to be further studied. That is why confusion matrices are calculated. This type of visualization illustrates how the classifier works and, in case of a bad classification, researchers can know if the deviation is important or not (as well as the direction) [35]. Apart from that, values, such as the mean average error (MAE), the root mean squared error (RMSE) [36], and the standard deviation between each model's folds are also taken into account. These error rates provide some insight into the issues and the nature of the errors committed by classifiers.

Given the lengths that it would take to provide detailed explanations of the methodology's steps for the identified challenges, the authors chose to provide a more concise explanation—one that is clear and extensive enough to understand the problem and facilitate the reproducibility of the solution we present. Therefore, the aspects associated with each of these challenges will be detailed below.

2.2. Data Extraction

Our research work is based on data-driven artificial intelligence models. Thus, data acquisition is one of the most important parts of the selected methodology. In our case, data acquisition was conducted by gathering historical records of metallography. Precisely, all images were provided by Azterlan. This research center stores (in a huge database) all tests and analyses that were conducted in the last years. The generated and stored images are made using a microscope in the conditions indicated by the UNE EN ISO 945 standards [7–9] for metallographic analyses. The photographs belonged to various metallic pieces from different foundries after being analyzed and categorized by metallurgical experts. The most important problem, for this research, is that these images must be categorized. Moreover, as we explained before, some applications are not able to do this accurately. For this purpose, those images were classified under the criteria of experts from Azterlan. More accurately, Azterlan, performed the analysis by extracting the diameter of the nodules using ImageJ [11] software as an auxiliary tool. Due to the results obtained by this type of software, it was not as perfect as we needed it to be for this research; manual work was conducted to solve misclassifications. In summary, after extracting some automatic information using a third-party software, all results were reviewed and corrected by domain experts.

For this dataset generation, several evaluations were conducted regarding the following topics.

- *Nodule shapes.* Taking into account the given results by third-party software, one new script was generated to segment the images and extract each nodule. Once it was done, the results were automatically classified into different folders. On the one hand, some of them were labeled and inserted in different folders related to the nodule shape, and on the other hand, in a special folder with nodules without classification. The aforementioned review task focused on validating the extracted information, as well as classifying the non-processed nodules.
- *Metal Quality.* The graphite distribution together with the nodule density was used to determine the quality of the solidified metal. The higher graphite density with a smaller graphite size distribution indicates that the metal has enough expansion capacity to avoid the austenitic contraction that occurs at the end of solidification. In this way, higher densities and graphite distributions with smaller graphites present less danger in terms of internal sanity defect apparitions. On the contrary, if there is a sample with a low graphite density with a distribution of coarser graphites, this indicates that at the end of solidification the precipitated graphites are insufficient at preventing the metallic contraction of the austenite, leading to micro-shrinkage. The analysis was conducted by extracting the diameter of the nodules using ImageJ [11],

to divide them into different groups, ranging from 5 to 60 μm . The results were plotted in a histogram-type graph to definitively be interpreted by experts; they were evaluated and a label was assigned to the metallography. In the end, the resultant dataset consisted of labeled images referencing five different groups: (i) optimal, (ii) good, (iii) regular, (iv) bad, and (v) pessimal; from a higher density and smaller size to the contrary. The labeling criteria were carried out by using a comparison method. Accurately, the labeled dataset was processed by extracting individual nodules and measuring them. Secondly, the average graphite diameters were calculated and divided into the groups explained in Section 2.5. Third, the average values between the metallography belonging to the same group were calculated to establish a centroid of each possible classification label. Finally, each metallography was compared to all centroids using cosine similarity [37] as the measurement. The biggest similarity was the label to be applied to our metallography.

2.3. Handling Nodule Sizes

As the first detected challenge, the nodule size values were used to determine how the distribution of graphites impacted the ductility of the metal. Regarding the previous explanation, the distribution and the sizes of graphites play important roles in explaining the behavior of the metal. Considerably sized nodules indicate that the graphite is condensed in a few spots, pointing out the deficiencies in the casting process.

For our research work, there was a need to build the extraction process from scratch in order to avoid third-party software usage. To extract all possible information from each nodule, we used the OpenCV library [38]. The employed method is similar to the algorithms presented in [39,40]. In our specific case, our machine vision development extracted the following parameters: the number of nodules mm^2 of each nodule and the summation, the average of the graphite area, the percentage of particles in the area, and a percentage of the nodularity. Later, other information will be added but will be related to the classifications of the nodule shapes (i.e., the percentage of each nodule in the different groups previously defined).

We built a custom solution because it helped to speed up the overall process by computing only necessary values. In order to complete this, the available metallography was binarized as the first step. Thus, the contour detection task used to obtain the nodules worked better, due to the fact that there was no need to process any other additional misleading information.

After that, we extracted the size features from the graphites (as depicted in Figure 1). On the one hand, values as diameters (l_m) were calculated by drawing the minimum enclosing circles for the nodule and measuring them directly as the standards dictate. We discarded particles with diameters smaller than 5 μm because they are inclusions (or microporosity) that do not influence the solidification morphology.

On the other hand, the graphite area was obtained by subtracting black pixels (A) to the total area (A_m). This last value was not used in the classification, but it contributed to the final evaluation of the metallography.

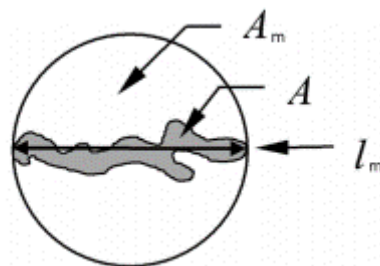


Figure 1. Diameter measuring method where the nodule was circumscribed within a circle with diameter l_m and area A_m ; the nodule area is A .

The metallographic images have standard sizes, in accordance with the ISO methodology, from which we obtained the conversion factor relation from the pixels to the actual lengths. Applying this relation to calculate the diameters of the graphites, we categorize them into 12 different groups that ranged from 5 μm each, starting from a minimum diameter of 5 μm and collecting all of the nodules bigger than 60 μm in the last group (i.e., the groups collected graphites with diameters in the ranges of 5–10 μm , 10–15 μm , ... 55–60 μm , and >60 μm , respectively). This division will be employed lately to classify the quality of the metal.

2.4. Handling Nodule Shapes

In addition to nodule sizes, their shapes also make significant contributions to the metallurgical quality. Moreover, shape classification is needed to determine the quality of the metal. At first glance, the graphite image contour classification problem seems suitable for deep learning classifiers.

Visual categorization using these approaches was quite successful throughout the years and is effective at classifying everything from animals [41] to landscapes [42]. Concerning this problem, there were similar approaches to nodule shape classifications. Regarding the usage of shape classification in the metallurgic sector, Damacharla et al. [31] applied segmentation to real-time images using computer vision and then successfully trained a deep learning network for steel surface defect detection. This research has a similar procedure to our classification problem.

In the same way, similar shape classification problems were tackled. Esteva et al. [43] classified skin cancer by using DL to differentiate it from the more common skin lesions. The nets developed were the best dermatologists while classifying the samples, proving that this type of solution can be used as an auxiliary tool for medics or it could also be included in mobile devices to help people. Moen et al. [44] also described methods on how to face the problem of cellular analysis using deep learning nets and computer vision, emphasizing the usefulness of these methods and discussing why they are suitable for this type of analysis. Edge detection, shape classification, and image extraction are some of the techniques that are recommended, and most of them can be used in nodule shape classification problems as they bear resemblances.

As for the graphite dataset problem, nodules are extracted from the metallography, obtaining a binarized rectangle containing the graphite. The validation set also follows the same logic as the former one, so these data are ordered by shape type and quality size. Usually, the categorization is made using mathematical formulas, such as by calculating circularity and the aspect ratio to determine which group they belong to. Circularity involves the compactness of an object that encloses the most area for a given perimeter. Specifically, this is calculated as per Equation (1), where A is the total area of the object and p^2 is the square value of the perimeter, obtaining a f_{circ} value that indicates how much of a circle it is, with 1 being a perfect circle and 0 a straight line.

$$f_{circ} = \frac{4\pi A}{p^2} \quad (1)$$

$$A_R = \frac{d_{max}}{d_{min}} \quad (2)$$

Regarding the aspect ratio, it defines the relationship between the width and height. Equation (2) illustrates (A_R); it was calculated based on the largest (d_{max}) and smallest (d_{min}) orthogonal diameter of the minimum rectangle that encloses the nodule, with the smallest diameter being value 1. Equations (1) and (2) combined are used to classify the nodule type depending on the values obtained (see Table 1 for more information), circularity (CIRC) ranging from 0 to 1, and aspect ratio (AR) from 1 (perfect square) to infinity.

Table 1. Nodule-type classification rules indicating the belonging group.

Team Sheet		
Type III	CIRC. [0,0.7]	AR. [2, ∞)
	CIRC. [0,0.5)	AR. (1.5, 2)
	CIRC. [0,0.35)	AR. [1, 1.5]
Type V	CIRC. [0.35,0.77)	AR. [1, 1.5]
	CIRC. [0.5,1]	AR. (1.5, 2)
	CIRC. [0.7,1]	AR. [2, ∞)
Type VI	CIRC. [0.77,1]	AR. (1, 1.5]

With this categorization, the goal is to obtain an estimation of the nodule dataset and let the deep learning net deduce the relevant characteristics in each class.

The labels used to represent the nodules were based on the ISO standard classification [7–9] ranging from I to VI. Despite this, the first and second groups were mixed with the third one, due to the fact that the properties obtained through Type III nodules and inferior ones were not required for ductile cast iron. In fact, these are considered vermicular graphites and they are counterproductive for this type of iron [45]. Moreover, the similarities between the fourth and fifth groups are high, which is why they are gathered into the same class. To summarize, nodules were divided into the following groups:

- *Type III (Figure 2a)*: this group encompassed the first three types of the ISO classification, the ones considered elongated or amorphous, far from the perfect circular-shaped graphites.
- *Type V (Figure 2b)*: considered as both Types IV and V, and far more similar to circles than the first group, but still presented the deformities.
- *Type VI (Figure 2c)*: the last types of graphites were very circular and proportional; they were considered the ideal shapes.

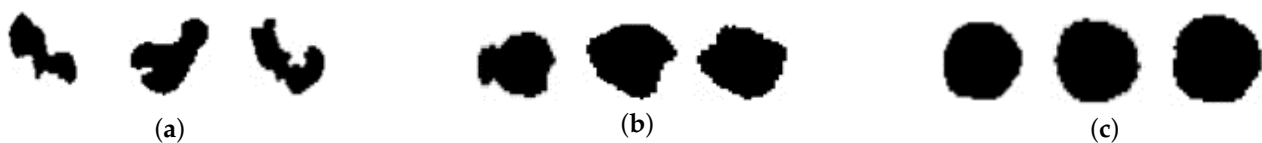


Figure 2. Examples of (a) Type III nodules presenting deformities and elongated shapes, (b) Type V nodules presenting slight deviations from circular shapes, and (c) Type VI nodules resembling circles with little to no deformities.

The working datasets are not balanced, as the Type III group had fewer nodules than the other (35,000 against approximately 100,000 for the other two classes). To fix this issue, we randomly picked the same number of nodules from every class, so we could balance the dataset and still obtain a representative number from each one. The images needed to be processed by a deep learning network, and as Shorten et al. detailed in their survey [46], data augmentation techniques strengthen deep learning models due to the increase of available data without needing new samples. As it is considered good practice to reinforce the classifiers, we selected techniques, such as horizontal and vertical flips and rotations, obtaining around 140,000 images for each label as seen in Table 2.

Table 2. Nodule dataset of the multiclass classifier.

	Type III	Type V	Type VI
Original	35,000	60,000	107,000
Augmented	140,000	140,000	140,000

This kind of experiment is more resource-demanding when images are involved in large learning networks. In our case, to make the shape classifications, three different

models were tested: two pretrained well-known models, such as VGG16 and VGG19 [47], and the third one with a custom architecture based on a net used to classify letters and numbers (as seen in Figure 3). The VGG classifiers were trained on the ImageNet dataset which contained over 14 million images over 1000 classes, which let the model learn about, which were the important features of entities, such as corners, shapes, and so on. The greatest part about these models is that the input and output layers can be configured freely, so it would be possible to adapt it to other types of classification problems without forgetting what it learned.

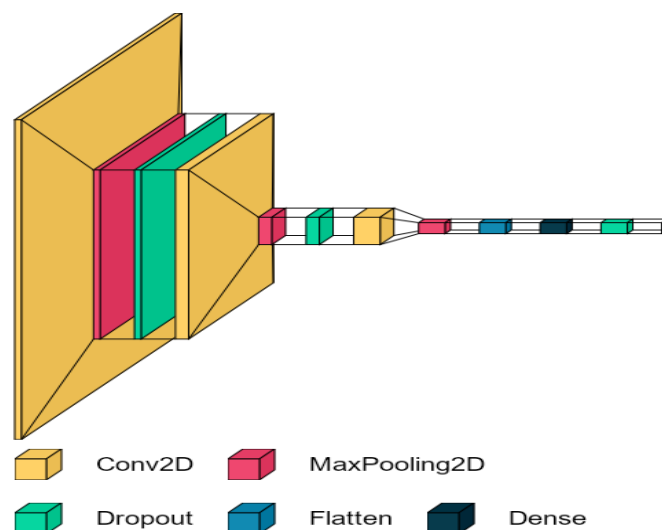


Figure 3. Custom architecture of the deep learning classifier.

The aforementioned three models will be tested against the nodule shape classification. Once the best is determined, the optimizers are checked to see if there are notable differences between them.

To fine-tune the classification process, a pipeline of binary classifiers is proposed instead of a single multiclass one. The application of this strategy is based on the idea of “one-vs-all” defended by Rifkin et al. [48]. The architecture involves implementing a specialization pattern, distinguishing each of the models from two classes instead of all of them. The classifiers will learn different features as they need to differentiate in binary ways.

The pipeline was trained in a different way than the multiclass classifier. For this method, the data needed to be changed and re-adapted to a new perspective. The models made binary choices, but the nodules not belonging to the first classes evaluated needed to be grouped together. The first classifier proposed differentiated between the Type III and V classes as seen in Table 3a. Training samples from Type VI would also be classified as Type V, due to their resemblance. Nevertheless, the second classifier did not include Type III, as they were already labeled on the first step. These two types had more original samples than the first one, which is why this dataset, as seen in Table 3b, consisted of all original samples with data augmentation applied. The two classes had more similarities between them than the previous ones; hence, it was critical to use most of the possible samples.

Table 3. Nodule dataset for the (a) first and (b) second classifiers of the pipeline.

	a			b	
	Type III	Type V		Type V	Type VI
Original	35,000	60,000	Original	60,000	107,000
Base	35,000	35,000	Base	60,000	60,000
Augmented	140,000	140,000	Augmented	300,000	300,000

The nodule shape has an extra validation method as it behaves differently from the other techniques. The Grad-CAM method [49] indicates the points of interest in the deep learning net, which are the areas of attention that lead to classifying an image into its corresponding label. Based on neuron activation on each pixel, the model is capable of turning these into understandable values. To represent these results, nodule images were depicted with an overlapped heatmap, which manifested where the layers focused on when classifying (drawn with bright colors). The scale used for the pictures was “inferno”, as shown in Figure 4.



Figure 4. Inferno color scale.

2.5. Metal Quality Classification

Throughout the years, statistical methods were used to solve a variety of similar problems. Classical methods, such as K-nearest neighbor [20] and the Bayes networks [22] still serve as fast techniques that achieve great results. Fragassa et al. [50] predicted the hardness properties of cast iron based on statistical and machine learning methods, obtaining great accuracy in their forecast.

These algorithms are also proven to be efficient in other aspects, such as in the research by Wisnu et al. [51], who compared both methods in order to analyze customer satisfaction concerning digital payments. In this case, both seemed adequate, but an algorithm as simple as KNN achieved the highest accuracy.

The metallurgical area often presents distinct problems where machine learning methods are useful, although not every one of them works as intended. Santos et al. [52] experimented with several methods, including support vector machines (SVMs) [23] and decision trees [53], among other classical statistical methods for the prediction of dross in ductile iron casting productions. They concluded that some of the random forest (RF) variations performed quite well, although instance-based classifiers (as the first ones mentioned) were the worst for this kind of activity. In contrast with the last article mentioned, Gola et al. [54] used SVM and focused on classifying the microstructures of low-carbon steels. The model achieved high accuracy by reducing the parameters evaluated, simplifying the model, and working more efficiently. Moreover, artificial neural networks (ANNs) [24] seem to learn specific problems efficiently. Their adaptability is their strength, and given enough data, their performances are outstanding. As shown by Rodrigues et al. [19], who classified cast iron (composed of carbon, iron, and silicon) into three variants, with the neural network performing quite well. As with classical methods, ANNs also adapt to other types of situations. In the case of Patricio and Orellana [55], they classified the risks in savings and credit cooperatives with the same accuracy as financial entities.

Regarding metal quality, the distribution of particle sizes affects the metallurgic properties as they are indicators of temperature evolution during the casting process. Thus, once the groups defined in Section 2.3 are populated, the data are prepared to be processed by different classifiers; moreover, 1000 images were used for cross-validation, with a value of $K = 10$, while the other 110 were saved for validation. This set was balanced with the aim of proving the robustness of the algorithm and checking the best scoring model performance on new data. The resulting dataset consisted of 1110 types of metallography with the sizes labeled; 1000 samples for cross-validation training were divided 10-fold, and 110 were divided from the original dataset for the validation part.

After this, a variety of classifiers were tested. Since these algorithms accept different parameters, some variations were considered as candidates. Afterward, the best scoring methods were compared using the measurement values mentioned in Section 2.1.

The methods compared were as follows:

- **K-nearest neighbor (KNN):** This is a supervised classification algorithm that uses the Euclidean distance to measure how close the learned values are in relation to the ones

that need to be classified. The method was tested with numbers between 1 and 8 to see how the model reacted to different values in order to evaluate accuracy.

- **Bayesian network classifier:** This is a graphical probabilistic model for the multivariate analysis based on the Bayes theorem. Various search algorithms were used to determine which was the best, due to behavioral variations: (i) K2, (ii) hill climber, (iii) LAGD hill climber, (iv) tree augmented naïve, and (v) simulated annealing.
- **Artificial neural networks (ANNs):** These are a collection of interconnected units called neurons that are distributed between layers. They compute the outputs based on the input values and non-linear activation functions. The networks used consist of multilayer perceptrons, input, a single hidden layer, and an output layer, which use backpropagation for training. The input had 12 different values, representing the classes mentioned, 8 neurons on the hidden layer, and 5 outputs representing the classification. Two of them were used and the differences between them involved the tolerances and the usage of conjugate gradient descent.
- **Support vector machine (SVM):** Data represented in an n-dimensional space are divided into two regions by a hyperplane to attempt to maximize the margin between distinct classes. In order to diversify the surfaces created, multiple kernels such as (i) poly kernel, (ii) normalized poly kernel, (iii) radial basis function kernel, and (iv) Pearson VII were used.
- **Decision trees:** These are logical construct diagrams based on successive rules or conditions that represent solutions for problems. Different tree sizes were tested; numbers ranged from 50 to 500 in batches of 50. Apart from that, another distinct method based on the C4.5 algorithm was used.

3. Results

Once the research work was conducted, the authors used the measurements and validations explained in Section 2. As a summary, all results, regarding the three challenges defined in our methodology, were satisfactory, giving an idea of how this approach can help to solve the initially defined problems. The following subsections attempt to explain the achieved results and show how astonishing the use of machine vision and DL is in this kind of problem.

3.1. Nodule Size Results

Regarding the work of extracting features from the images, a direct comparison was made between our development and other third-party tools. In this way, the same metallography was analyzed through our script based on the OpenCV library and a third-party tool, to later compare the values provided by both systems. In all cases, the same results were achieved for all images, as long as the configurations of all of them were the same. The only thing we can say in favor of our script is that avoiding interactions with painting devices until the end of the job produces a small acceleration of the calculation. The improvement is a few microseconds. Perhaps it is not relevant for single metallography, but usually, the analysis is conducted in a batch of images. In that case, the same data would be obtained in the order of about 700 microseconds (average value reached during the tests) less by metallography.

3.2. Nodule Shape Results

As explained in Section 2.4, in order to classify the shapes of the nodules, DL models were chosen. Among them, we had VGG16 and VGG19, where the best was the second one. The success rate achieved was 93%. As we showed with the confusion matrix in Table 4a, a high error was obtained when making the Type V classifications. Thus, a different form of classification was carried out; the binary classifier pipeline is explained. The success rate rose to 95% and the stability of these classifications increased, as shown by the confusion matrix grouped in Table 4b.

Table 4. Confusion matrix results for the single classifier in the (a) unique and (b) pipeline models.

		a		
		Predicted		
		Type III	Type V	Type VI
Actual	Type III	94.64	5.36	0
	Type V	0.58	72.83	26.59
	Type VI	0	3.3	95.88
		b		
		Predicted		
		Type III	Type V	Type VI
Actual	Type III	89.29	10.71	0
	Type V	3.47	75.72	20.81
	Type VI	0	1.56	98.44

For this type of classification system, the use of Grad-CAM as validation is common. Grad-CAM allows one to visualize which features are being used by the DL-based classification system. In order to not extend ourselves into showing too many images, we will focus on illustrating and commenting on the results of the best classifier—the classification pipeline.

Thus, Figures 5–7 show that through coloring (via the Inferno color scale), the key points in which the classification models are used. Accurately, in the case of Figure 5 and 6, they are perfectly adjusted to the surfaces of the nodules. The same can be observed in Type III and Type V nodules. It is true that in the last of them they appear a little more displaced, but one can still determine the shape and structure of the nodule. In summary, the model detects elongation and deformities as indicators in a very accurate way.

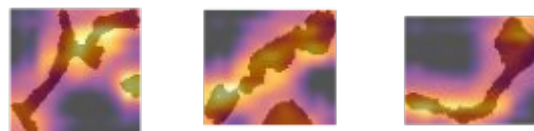


Figure 5. Type III nodule-significant shapes colorized by Grad-CAM.

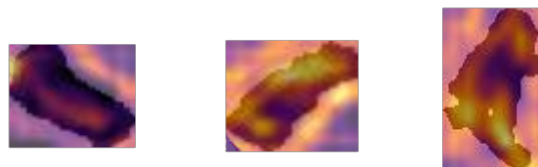


Figure 6. Type V nodule shapes and contours colorized by Grad-CAM.

Nevertheless, the last ones, Figure 7, have peculiar results, where the vivid part is observed to center on the middle and four main corners. Essentially, we thought as this kind of nodule is really round, the model checks a circle-like shape with no holes or deformities in the center. Again, this behavior is totally related to the expected shape of the graphite, giving a good accuracy of the model.

In summary, DL models seem to work by checking the main parts of the nodules depending on their shapes. These key points are representative of the classification and, because of this, they perform well, providing highly accurate values.

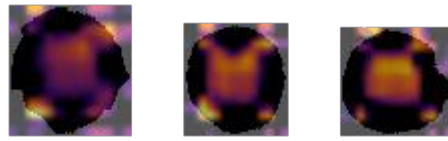


Figure 7. Type VI nodule corners and middle parts colorized by Grad-CAM.

3.3. Metallurgical Quality Results

As said before, to make the metallurgical quality classification, we employed statistical classifiers. To perform this task, from the testing dataset, some specific information in the form of (i) a group and size division, and (ii) counts were extracted (for more information, see Section 2.5). Then, the characterization of each metallography was sent to different ML classifiers. These results are presented in Table 5. This table shows the different classifiers and their configurations (in other words, the learning algorithms, kernels, or values of n used for creating them). In order to make the table more ‘readable’, the best performance is colored in gray and the best classifier for each algorithm group is highlighted in bold text.

As Table 5 shows, the majority of classifiers performed well. All of them achieved accuracies of 85%. Some simple classifiers, such as KNN and Bayesian classifiers, obtained good results. Specifically, the best KNN method (with k equal to 7) scored an accuracy of 91%. However, as seen in Table 6c, the error nature from the confusion matrix is high regarding the classes from both ends, while the others remain close to the actual class. The best (tree-augmented) Bayesian classifier obtained an accuracy of 88%. Moreover, it obtained the highest notorious error rate while classifying contrary classes, as the Table 6d confusion matrix shows.

The best kernel for support vector machines was Pearson VII. It provided an accuracy of 95% over the dataset, although the classification failures spread through the least quality labels, as seen in Table 6e. In this research, several decision trees were tested, but all of them performed similarly, obtaining accuracies of around 93%. The error dispersion is shown in Table 6f.

Table 5. Results and characteristics of the nodule size classification. Results in bold are the best from each algorithm type, and the one colored in gray has the highest accuracy (%) and lowest error overall. All results (i.e., accuracy, MAE, and RMSE) include the standard deviation for each fold in the k-fold cross validation.

Classifier	Accuracy (%)	MAE	RMSE
Bayesian Network (K2)	85.07 ± 3.29	0.06 ± 0.01	0.22 ± 0.02
Bayesian Network (Hill Climber)	85.07 ± 3.29	0.06 ± 0.01	0.22 ± 0.02
Bayesian Network (LAGD Hill Climber)	87.87 ± 3.04	0.06 ± 0.01	0.19 ± 0.02
Bayesian Network (Tree Augmented Naïve)	88.13 ± 3.26	0.06 ± 0.01	0.19 ± 0.02
Bayesian Network (Simulated Annealing)	88.5 ± 3.23	0.06 ± 0.01	0.19 ± 0.03
Naïve Bayes	85.99 ± 3.07	0.06 ± 0.01	0.21 ± 0.02
SVM (Poly Kernel)	92.78 ± 2.33	0.24 ± 0	0.32 ± 0
SVM (Normalized Poly Kernel)	91.28 ± 2.47	0.24 ± 0	0.32 ± 0
SVM (RBF Kernel)	56.28 ± 2.51	0.27 ± 0	0.36 ± 0
SVM (Puk)	95.08 ± 1.95	0.03 ± 0	0.32 ± 0
ANN (MLP+Conjugate gradient descent, 500 ep.)	97.84 ± 1.38	0.02 ± 0	0.09 ± 0.02
ANN (MLP, 500 ep.)	96.76 ± 1.61	0.05 ± 0.01	0.1 ± 0.02
KNN (K = 1)	87.85 ± 2.89	0.05 ± 0.01	0.22 ± 0.03
KNN (K = 2)	86.48 ± 3.18	0.05 ± 0.01	0.19 ± 0.02
KNN (K = 3)	90.62 ± 2.68	0.05 ± 0.01	0.17 ± 0.02
KNN (K = 4)	89.95 ± 2.76	0.05 ± 0.01	0.17 ± 0.02
KNN (K = 5)	90.98 ± 2.54	0.06 ± 0.01	0.17 ± 0.02
KNN (K = 6)	90.49 ± 2.68	0.06 ± 0.01	0.17 ± 0.02
KNN (K = 7)	91 ± 2.66	0.06 ± 0.01	0.17 ± 0.02

Table 5. Cont.

Classifier	Accuracy (%)	MAE	RMSE
KNN (K = 8)	90.51 ± 2.84	0.06 ± 0.01	0.17 ± 0.02
Tree (J48)	87.36 ± 3.26	0.05 ± 0.01	0.22 ± 0.03
Random forest (n = 50)	92.95 ± 2.36	0.06 ± 0.01	0.15±0.01
Random forest (n =100)	93.01±2.32	0.06 ± 0.01	0.15 ± 0.01
Random forest (n = 150)	93.07 ± 2.33	0.06 ± 0.01	0.15 ± 0.01
Random forest (n = 200)	93.12 ± 2.24	0.06 ± 0.01	0.15 ± 0.01
Random forest (n = 250)	93.12 ± 2.12	0.06 ± 0.01	0.15 ± 0.01
Random forest (n = 300)	93.06 ± 2.21	0.06 ± 0.01	0.15 ± 0.01
Random forest (n = 350)	93.15 ± 2.17	0.06 ± 0.01	0.15 ± 0.01
Random forest (n = 400)	93.16 ± 2.15	0.06 ± 0.01	0.15±0.01
Random forest (n = 450)	93.14 ± 2.25	0.06 ± 0.01	0.15 ± 0.01
Random forest (n = 500)	93.16 ± 2.31	0.06 ± 0.01	0.15 ± 0.01

The best performances were provided by artificial neural networks. The neural network that did not use the conjugate gradient descent obtained a precision of around 97% with a good distribution of errors as they were close to the actual class (see Table 6a). In contrast, when the conjugate gradient was used, the accuracy increased to 98%, despite the fact that the distribution of false positives was wider, yet smaller than in the previous model (illustrated by Table 6b).

Table 6. Confusion matrix results for the (a) first neural network, (b) second neural network, (c) KNN, (d) Bayesian network, (e) support vector machines, and (f) random forest models.

a						b							
		Predicted							Predicted				
		Optimal	Good	Type VI	Bad	Pessimal			Optimal	Good	Type VI	Bad	Pessimal
Actual	Optimal	99.36	0.63	0	0	0	Actual	Optimal	100	0	0	0	0
	Good	3	96	0	0	1		Good	1	98	0	1	0
	Regular	0	0.96	98.54	0	0.48		Regular	0	0.48	98.54	0	0.96
	Bad	0	0	1.1	97.24	1.57		Bad	0	0	1.18	97.63	1.57
	Pessimal	0	0	0	2.7	97.29		Pessimal	0.67	0	0	2.03	97.29
c						d							
		Predicted							Predicted				
		Optimal	Good	Type VI	Bad	Pessimal			Optimal	Good	Type VI	Bad	Pessimal
Actual	Optimal	92.72	1.21	0	0	6.06	Actual	Optimal	90.38	4.49	0	0	5.13
	Good	5.43	90.22	4.35	0	0		Good	6.25	78.13	15.62	0	0
	Regular	0	3.74	91.58	4.44	0.23		Regular	0.24	4.6	90.8	4.36	0
	Bad	0	0	5.43	88.37	6.2		Bad	0	0	7.17	89.64	3.19
	Pessimal	0.78	0	0.78	3.1	95.35		Pessimal	7.05	0	1.92	5.13	85.9
e						f							
		Predicted							Predicted				
		Optimal	Good	Type VI	Bad	Pessimal			Optimal	Good	Type VI	Bad	Pessimal
Actual	Optimal	94.97	1.89	0	0	3.14	Actual	Optimal	91.19	3.14	0	0	5.66
	Good	1.98	88.12	9.9	0	0		Good	2.97	89.1	7.92	0	0
	Regular	0	0	97.81	1.7	0.48		Regular	0	0.72	95.62	3.4	0.24
	Bad	0	0	1.99	94.42	3.58		Bad	0	0	5.58	92.03	2.39
	Pessimal	0.66	0	0	5.33	94		Pessimal	1.33	0	0.66	4.66	93.33

4. Discussion

The research work that used machine vision techniques was performed as expected. In addition, given the accuracy in the extraction of information, among the obtained accelerations, they are good options to solve this problem. However, better working methods could still be found, such as the use of CUDA programming for obtaining much greater acceleration.

The rest of the experiments are more complex; for the discussion of each of them, we created a specific section for each one.

4.1. Nodule Shape Results Discussion

Regarding the nodule shape classification problem, it is clear that both VGG16 and VGG19 fail to understand the problem, most likely because they are not fit to deal with these kinds of shapes. The main model used for the classification problem was very effective from the start, obtaining a 93% accuracy, but there was a minor issue with the images and classes. Type V and VI nodules are relatively similar; here is where the model increases its error deviations. Although at this moment its accuracy is lower than the pipeline, the single classifier model is faster to train than the pipeline model. Hence, further work could be conducted to find a better option to create only one DL classifier.

Moreover, considering that accuracy improves while the complexity of the pipeline increases, it is possible that results would improve using a more complicated structure. Apart from that, different architectures for each classifier could be useful to test, since different problems require adequate solutions.

Finally, we must remember that we are using a data-driven approach; thus, the dataset has a crucial role in the classifiers, and one of the best practices to develop a strong model is data augmentation. This problem has one advantage—nodules can be rotated at any angle without losing any characteristics related to the shape. The nodule can be rotated and flipped, vastly increasing the dataset size. Still, this operation is demanding. As this consists of research, it is possible to improve these results and close the gap between the two conflicting types of graphites by dedicating enough time and resources.

4.2. Metallurgical Quality Results Discussion

Before the experimentation in this area, we attempted to process the metallography directly using a deep learning network on the labeled images. However, we discarded this approach due to its low accuracy. Further research with different prediction models could prove if the idea could be done. Hence, this will be considered feasible in future work to reduce the number of calculations for this purpose. This new way of doing this will provide the classification in one step only.

As observed in the results of the nodule size classification, most of the models achieved exceptional accuracy. After all the tests, even the classic methods performed well, displaying great results. Nevertheless, ANNs proved to be more accurate at the cost of longer training times than the previous ones. Despite having success in the experimentation, many classifiers have not been tested. To obtain an overview of the best classifiers, it will be necessary to test and compare more than the ones previously discussed as the actual accuracy could improve.

Moreover, the experimentation could benefit by using an ensemble of classifiers [56] to obtain the best results. For example, the majority of voting [57] could be included among them as a method of agreement and to prevent unilateral decisions.

5. Conclusions

Usually, metallographic analyses are based on visual comparisons made by human beings. Thus, different types of characteristics are determined based on what a person believes or thinks. This way of working, in addition to being totally subjective, is subscribed to multiple possibilities of failure or deviations introduced by the variability of the human

being. That is why, through this research work, a way to be more precise was sought, allowing us to calculate everything in a much more objective way.

In the same way, when these results were made with more advanced microscopes, we obtained better results due to the measurement device-applied image processing algorithms. However, the results provided are not as accurate as needed. Hence, in this research work, we developed a new technique for making this kind of analysis, which is capable of classifying nodules from metallography in an objective way. The obtained models under a deep learning architecture improve accuracy in the classification tasks. In fact, they are able to give classification results for all nodules and, taking into account expert validations, with better accuracy than the classical methods by extracting numerical values for the image generated by the microscope. Although in this research work we only applied one classification method, the employment of this kind of model allowed us to create a huge classification repository, depending on the customer or the regulation, by easily selecting and loading a specific model to make this work.

As the experiments were limited due to time and hardware, further research needs to be done, applying different architectures, methods, and data augmentation techniques in order to strengthen the classification model. Exceptionally, there could be errors or defects while cutting metal pieces that produce small malformations on the nodules. In order to detect them, another classifier could be created to manually or automatically divide them into their original shapes.

The whole microstructure analysis was not complete without the pearlite and ferrite details. This was carried out by chemically attacking the metallic piece and measuring the proportions from the metallography. The resultant image was not like the ones used for this experiment; thus, a completely new approach needs to be implemented. Moreover, the lamellar cast iron faces a similar issue. Contrary to nodules, the graphite part was distributed in line-like shapes across the iron part, which is why the same approach would be valid to solve the quality determination problem for this type of iron. Instead of classifying separated circles, these lines are interconnected in most cases, increasing the categorization difficulty. The main idea behind the development of these two iron analysis types is to build microstructure examination software to improve and ease labor.

Regarding other types of material, aluminum analysis faces an identical situation to that in this research. Nodule extraction, measurements, and classifications provide a strong start to adapt this analysis to the different domains. Further research must be carried out on how it could be solved with a similar approach.

To summarize, using the methods proposed in this paper, both nodule size and shape classification were successfully achieved; therefore, we completed all of the objectives set out for this research and proved the usefulness of deep learning networks and artificial neural networks. Our main contributions to this research work are as follows. We:

1. Presented an approach to manage the characterization of metallographic images.
2. Provided a deep learning-based method to classify nodule characteristics.
3. Introduced a multiple-level classifier (classification pipeline) to improve the type of classification.
4. Completed the feature extraction by determining the quality of metal made thanks to the employment of an ANN classification.
5. Proposed a new way of classification, which is easy to manage and make different classifications based on different conditions.
6. Defined how it could be created via a small software prototype that is able to handle all metallographic analyses in offline and online manners.

Author Contributions: Conceptualization, X.S.; B.B. and J.N.; methodology, J.N. and I.P.-L.; software, X.S.; validation, J.N., B.B., I.P.-L. and P.G.B.; investigation, X.S.; resources, B.B.; data generation, B.B. and X.S.; writing—original draft preparation, X.S. and J.N.; writing—review and editing, J.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The dataset has been generated by extracting information from different private projects of the Azterlan research center. However, the way to generate it has been explained, allowing readers to reproduce it.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. *Metal Casting Market Size, Share & Trends Analysis Report By Material (Aluminum, Iron, Steel), By Application (Automotive & Transportation, Building & Construction, Industrial), By Region, And Segment Forecasts, 2020–2025*; Market Analysis Report GVR-3-68038-298-3; Grand View Research: San Francisco, CA, USA, 2019.
2. Pattnaik, S.; Karunakar, D.B.; Jha, P.K. Developments in investment casting process—a review. *J. Mater. Process. Technol.* **2012**, *212*, 2332–2348. [[CrossRef](#)]
3. Beeley, P. *Foundry Technology*; Elsevier: Amsterdam, The Netherlands, 2001.
4. Ngqase, M.; Pan, X. An overview on types of white cast irons and high chromium white cast irons. *J. Phys. Conf. Ser.* **2020**, *1495*, 012023.
5. Collini, L.; Nicoletto, G.; Konečná, R. Microstructure and mechanical properties of pearlitic gray cast iron. *Mater. Sci. Eng. A* **2008**, *488*, 529–539. [[CrossRef](#)]
6. Özdemir, N.; Aksoy, M.; Orhan, N. Effect of graphite shape in vacuum-free diffusion bonding of nodular cast iron with gray cast iron. *J. Mater. Process. Technol.* **2003**, *141*, 228–233. [[CrossRef](#)]
7. *ISO 945-1:2008*; Part 1: Clasificación del Grafito por análisis Visual. Standard, International Organization for Standardization: Geneva, CH, USA, 2012.
8. *ISO 945-2:2011*; Part 2: Graphite Classification by Image Analysis. Standard, International Organization for Standardization, Geneva, CH, USA, 2011.
9. *ISO 945-4:2019*; Part 4: Test Method for Evaluating Nodularity in Spheroidal Graphite Cast Irons. Standard, International Organization for Standardization, Geneva, CH, USA, 2019.
10. Reason, J. *Human Error*; Cambridge University Press: Cambridge, UK, 1990.
11. Schneider, C.A.; Rasband, W.S.; Eliceiri, K.W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **2012**, *9*, 671–675. [[CrossRef](#)]
12. Gesellschaft zur Förderung Angewandter Informatik. AmGuss. Available online: <https://www.gfai.de/entwicklungen/bildverarbeitung/amguss> (accessed on 15 October 2021).
13. Leica. Leica Application Suite X (LAS X). Available online: <https://www.leica-microsystems.com/products/microscope-software/p/leica-las-x-ls/> (accessed on 26 December 2022).
14. Espinosa, L.F.; Herrera, R.J.; Polanco-Tapia, C. Segmentation of anatomical elements in wood microscopic images using artificial vision techniques. *Maderas Cienc. Technol.* **2015**, *17*, 735–748.
15. Gómez, F.J.P. Application of Artificial Vision Algorithms to Images of Microscopy and Spectroscopy for the Improvement of Cancer Diagnosis. Ph.D. Thesis, Universitat Politècnica de València, Valencia, Spain, 2018.
16. von Chamier, L.; Laine, R.F.; Henriques, R. Artificial intelligence for microscopy: What you should know. *Biochem. Soc. Trans.* **2019**, *47*, 1029–1040. [[CrossRef](#)]
17. Batool, U.; Shapiai, M.I.; Tahir, M.; Ismail, Z.H.; Zakaria, N.J.; Elfakharany, A. A systematic review of deep learning for silicon wafer defect recognition. *IEEE Access* **2021**, *9*, 116572–116593. [[CrossRef](#)]
18. Tin, T.C.; Tan, S.C.; Lee, C.K. Virtual Metrology in Semiconductor Fabrication Foundry Using Deep Learning Neural Networks. *IEEE Access* **2022**, *10*, 81960–81973. [[CrossRef](#)]
19. Rodrigues, D.d.A.; Santos, G.P.d.; Fernandes, M.C.; Santos, J.C.d.; Freitas, F.N.C.; Rebouças Filho, P.P. Classificação automática do tipo de ferro fundido utilizando reconhecimento de padrões em imagens de microscopia. *Matéria (Rio Jan.)* **2017**, *22*. [[CrossRef](#)]
20. Peterson, L.E. K-nearest neighbor. *Scholarpedia* **2009**, *4*, 1883. [[CrossRef](#)]
21. Hamerly, G.; Elkan, C. Learning the k in k-means. *Adv. Neural Inf. Process. Syst.* **2003**, *16*, 281–288.
22. Pearl, J. *Reverend Bayes on Inference Engines: A Distributed Hierarchical Approach*; Cognitive Systems Laboratory, School of Engineering and Applied Science, University of California: Los Angeles, CA, USA, 1982.
23. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer Science & Business Media: Berkeley, CA, USA, 2013.
24. Krogh, A. What are artificial neural networks? *Nat. Biotechnol.* **2008**, *26*, 195–197. [[CrossRef](#)] [[PubMed](#)]
25. Knuth, D. *Art of Computer Programming, The: Volume 3: Sorting and Searching*; Addison-Wesley: Reading, MA, USA, 1968.
26. Qassim, H.; Verma, A.; Feinzimer, D. Compressed residual-VGG16 CNN model for big data places image recognition. In Proceedings of the 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 8–10 January 2018; pp. 169–175. [[CrossRef](#)]
27. Narendra, T.; Sankaran, A.; Vijaykeerthy, D.; Mani, S. Explaining deep learning models using causal inference. *arXiv* **2018**, arXiv:1811.04376.
28. Constante, P.; Gordon, A.; Chang, O.; Pruna, E.; Acuna, F.; Escobar, I. Artificial Vision Techniques to Optimize Strawberry’s Industrial Classification. *IEEE Lat. Am. Trans.* **2016**, *14*, 2576–2581. [[CrossRef](#)]

29. Draghici, S. A neural network based artificial vision system for licence plate recognition. *Int. J. Neural Syst.* **1997**, *8*, 113–126. [[CrossRef](#)]
30. Sata, A.; Ravi, B. Bayesian inference-based investment-casting defect analysis system for industrial application. *Int. J. Adv. Manuf. Technol.* **2017**, *90*, 3301–3315. [[CrossRef](#)]
31. Damacharla, P.; Ringenberg, A.R.M.V.J.; Javaid, A.Y. TLU-Net: A Deep Learning Approach for Automatic Steel Surface Defect Detection. *arXiv* **2021**, arXiv:2101.06915.
32. Posner, E.A.; Spier, K.E.; Vermeule, A. Divide and conquer. *J. Leg. Anal.* **2010**, *2*, 417–471. [[CrossRef](#)]
33. Mackey, L.; Jordan, M.; Talwalkar, A. Divide-and-conquer matrix factorization. *Adv. Neural Inf. Process. Syst.* **2011**, *24*.
34. Horowitz, E.; Zorat, A. Divide-and-conquer for parallel processing. *IEEE Trans. Comput.* **1983**, *32*, 582–585. [[CrossRef](#)]
35. Düntsch, I.; Gediga, G. Confusion matrices and rough set data analysis. *J. Phys. Conf. Ser.* **2019**, *2019*, 012055.
36. Chai, T.; Draxler, R.R. Root mean square error (RMSE) or mean absolute error (MAE). *Geosci. Model Dev. Discuss.* **2014**, *7*, 1525–1534.
37. Wang, P.; Wang, J.; Wei, G.; Wei, C. Similarity measures of q-rung orthopair fuzzy sets based on cosine function and their applications. *Mathematics* **2019**, *7*, 340. [[CrossRef](#)]
38. Bradski, G.; Kaehler, A. *Learning OpenCV: Computer Vision with the OpenCV Library*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2008.
39. Pohudina, O.; Kritskiy, D.; Bykov, A.; Szalay, T. Method for identifying and counting objects. In *Integrated Computer Technologies in Mechanical Engineering*; Springer: Cham, Switzerland, 2020; pp. 161–172.
40. Ni, J.; Khan, Z.; Wang, S.; Wang, K.; Haider, S.K. Automatic detection and counting of circular shaped overlapped objects using circular hough transform and contour detection. In Proceedings of the 2016 12th World Congress on Intelligent Control and Automation (WCICA), Guilin, China, 12–15 June 2016; pp. 2902–2906.
41. Zhuang, P.; Xing, L.; Liu, Y.; Guo, S.; Qiao, Y. Marine Animal Detection and Recognition with Advanced Deep Learning Models. In Proceedings of the CLEF (Working Notes), Dublin, Ireland, 11–14 September, 2017.
42. Buscombe, D.; Ritchie, A.C. Landscape classification with deep neural networks. *Geosciences* **2018**, *8*, 244. [[CrossRef](#)]
43. Esteva, A.; Kuprel, B.; Novoa, R.A.; Ko, J.; Swetter, S.M.; Blau, H.M.; Thrun, S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **2017**, *542*, 115–118. [[CrossRef](#)]
44. Moen, E.; Bannon, D.; Kudo, T.; Graf, W.; Covert, M.; Valen, D.V. Deep learning for cellular image analysis. *Nat. Methods* **2019**, *16*, 1233–1246. [[CrossRef](#)]
45. Guo, Q.; Zhong, Y.; Dong, T.; Gao, P.; Guo, Y.; Li, J. Effects of vermicular graphite rate on the oxidation resistance and mechanical properties of vermicular graphite iron. *J. Alloys Compd.* **2018**, *765*, 213–220. [[CrossRef](#)]
46. Shorten, C.; Khoshgoftaar, T.M. A survey on image data augmentation for deep learning. *J. Big Data* **2019**, *6*, 1–48. [[CrossRef](#)]
47. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
48. Rifkin, R.; Klautau, A. In defense of one-vs-all classification. *J. Mach. Learn. Res.* **2004**, *5*, 101–141.
49. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *Int. J. Comput. Vis.* **2020**, *128*, 336–359. [[CrossRef](#)]
50. Fragassa, C.; Babic, M.; dos Santos, E.D. Machine Learning Approaches to Predict the Hardness of Cast Iron. *Tribol. Ind.* **2020**, *42*, 1–9. [[CrossRef](#)]
51. Wisnu, H.; Afif, M.; Ruldevyani, Y. Sentiment analysis on customer satisfaction of digital payment in Indonesia: A comparative study using KNN and Naïve Bayes. *J. Phys. Conf. Ser.* **2020**, *1444*, 012034.
52. Santos, I.; Nieves, J.; Bringas, P.G.; Zabala, A.; Sertucha, J. Supervised learning classification for dross prediction in ductile iron casting production. In Proceedings of the 2013 IEEE 8th Conference on Industrial Electronics and Applications (ICIEA), Melbourne, VIC, Australia, 19–21 June 2013; pp. 1749–1754.
53. Quinlan, J.R. Induction of decision trees. *Mach. Learn.* **1986**, *1*, 81–106. [[CrossRef](#)]
54. Gola, J.; Webel, J.; Britz, D.; Guitar, A.; Staudt, T.; Winter, M.; Mücklich, F. Objective microstructure classification by support vector machine (SVM) using a combination of morphological parameters and textural features for low carbon steels. *Comput. Mater. Sci.* **2019**, *160*, 186–196. [[CrossRef](#)]
55. Patricio Peña, M.; Orellana, J.M. Red neuronal para clasificación de riesgo en cooperativas de ahorro y crédito. *Congr. Cienc. Tecnol.* **2018**, *13*, 121–124.
56. Freund, Y.; Schapire, R.E. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.* **1997**, *55*, 119–139. [[CrossRef](#)]
57. Penrose, L.S. The elementary statistics of majority voting. *J. R. Stat. Soc.* **1946**, *109*, 53–57. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.