Anomaly Detection for High Precision Foundries

Javier Nieves, Igor Santos, Xabier Ugarte-Pedrero and Pablo G. Bringas
S^{3}Lab, DeustoTech - Computing
University of Deusto
Bilbao, Spain
Email: {jnieves, isantos, xabier.ugarte, pablo.garcia.bringas}@deusto.es

***Abstract*—Microshrinkages are known as probably the most difficult defects to avoid in high-precision foundry. This failure renders the casting invalid, with the subsequent cost increment. Modelling the foundry process as an expert knowledge cloud allows machine learning algorithms to foresee the value of a certain variable, in this case, the probability that a microshrinkage appears within a casting. However, this approach needs to label every instance for generating the model that can classify the castings. In this paper, we present a new approach for detecting faulty castings inspired on anomaly detection methods. This approach represents correct castings as feature vectors of information extracted from the foundry process. Thereby, a casting is classified as correct or not correct by measuring its deviation to the representation of normality (i.e., correct castings). We show that this method achieves good accuracy rates to reduce the cost and testing time in foundry production.**

I. Introduction

The casting production or the foundry process is considered as one of the main factors influencing and improving the development of the world economy. Since the ancient time, such as it is showed in Biblical verses, Egyptian drawings or illustrations on Greek vases, it has already existed an advanced casting handicraft. When considering the development of foundry engineering in a historical aspect we tend to connect it with the development of the human civilisation and the attribution of a high position among the oldest world professions. Consequently, the casting production has always been considered like an ancient-magic-surrounded activity.

The actual capacity of the world's casting production, which is higher than 60 million metric tones per year, is strongly diversified [1]. The last decade brought significant changes in the world map of the greatest casting producers. Currently, the biggest producer is China, closely followed by Europe. These and other countries supply key pieces to many other industries, such as automotive, naval, weapon and aeronautic. Therefore, the foundry process is subject to very strict safety controls to assure the quality of the manufactured castings because, as one may think, the tiniest defect may become fatal.

The techniques for the assurance of failure-free foundry processes are exhaustive production control and diverse simulation techniques [2]. Many of the techniques used can only be applied when the casting is done. Thus, when a faulty casting is detected, it must be remelted, which can be translated into a cost increment. There are also other tests that have to destroy the castings, called destructive inspections, e.g., the tests for checking the mechanical properties. Likewise, these tests must be done when the casting is finished and the destruction of the casting also means a cost increment.

Unfortunately, these methods are still incapable of preventing what is known to be the most difficult flaw in ductile iron castings, namely the microshrinkage. More specifically, this imperfection, also called secondary contraction, consists of tiny porosities that appear inside the casting when it is cooling down. The difficulty of its detection is due to the fact that almost all the parameters of the foundry process influence the apparition of microshrinkages.

Indeed, the problem of the microshrinkage apparition is very difficult to solve due to the following reasons [3]: (i) A huge amount of data is required to be managed that it is not prioritised or categorised in any way, (ii) it is very hard to find cause-effect relationships between the variables of the system and (iii) the human knowledge used in this task usually tends to be subjective, incomplete and not subjected to any test [4].

Currently, machine learning is being used increasingly in the field of metallurgy in order to solve the aforementioned problems. One of the most widely used methods is the application of neural networks in several aspects such us classifying foundry pieces [5], optimising casting parameters [6], detection of causes of casting defects [7] and in other problems [8]. Similarly, other experiments involving K-nearest neighbour algorithm include fault detection of semiconductor manufacturing processes [9]. Bayesian networks are also used as previous methods in Bayesian Neural networks methodology (e.g., predicting the ferrite number in stainless steel [10]).

In our previous work, we tested several machine-learning classifiers [11], [12], e.g., Bayesian networks, support vector machines, decision trees, artificial neural networks among others, to identify which is the best classifier to predict microshrinkages and, also, to reduce the noise in the data gathering process produced by the foundry workers [13].

Despite the good accuracy level achieved in the previous works, this kind of approaches requires the whole dataset to be labelled, i.e., every instance of the dataset must be classified in correct or faulty casting before the training of the machine-learning models. The reason underlying this task is that the labelled dataset is the knowledge base employed by the machine-learning classifiers to generate the complete classification model. However, the labelling work is a hard, expensive and time-consuming work.

Given this background, we present a method to classify castings and to foresee microshrinkages that is highly inspired in anomaly detection methods. This type of approach is capable of determining whether a casting contains the secondary contraction defect or not by comparing features of the casting extracted from the foundry process with a dataset composed only by correct castings. Thus, when a casting under prediction presents a considerable deviation to what is considered as usual (the previously labelled correct castings), it is considered that the casting has some flaw, and specifically in this case, there is a high probability that the casting has some microshrinkages. This method deals with the aforementioned problem, achieving a reduction in the number of castings required to be labelled.

Anomaly detection techniques have been applied in the industrial damage detection domain. This domain can be classified into two others: (i) dealing with defects in mechanical components and (ii) dealing with defects in physical structures. For instance, this type of anomaly-based approach was previously used for fault detection in mechanical units [14] and, also, for structural defect detection [15].

Summarising, our main contributions are: (i) we select a set of variables extracted from the foundry process to determine whether a casting has a microshrinkage or not and we provide a relevance measure for each variable based on information gain, (ii) we propose an anomaly-detection-based architecture for microshrinkage prediction, by means of weighted comparison against a dataset composed of only correct castings and (iii) we evaluate the method using three different deviation measures and show that this method can achieve high accuracy rates while reducing the number of labelled castings required.

The remainder of the paper is organised as follows. Section II details the casting production process and presents tone of the most difficult defect to avoid, the microshrinkage. Section III introduces in deeper detail the concept of the anomaly detection method explaining the distance measures tested in this research. Section IV describes the experiments performed and examines the obtained results. Finally, section V discusses the main implications of our results while describes the main limitations of our approach and outlines the avenues of future work.

II. Foundry Process and Microshrinkages

The foundry process can be considered as one of the axes of our society. However, a task that seems simple becomes complex due to the hard conditions in which it is carried out. Besides the casting process, the foundry workers produce castings that are close to the final product shape, i.e., 'near-net shape' components. To this end, the production has to pass through several stages in which the castings are transformed to obtain the final casting.

Although all of the foundry process are not equal, the work flow performed in foundries is very similar to the work-flow shown in Fig. 1. The most important stages are the following [16]:

- **Pattern making:** In this step, moulds (exteriors) or cores (interiors) are produced in wood, metal or resin for being used to create the sand moulds in which the castings will be made.
- **Sand mould and core making:** The sand mould is the most widely extended method for ferrous castings. Sand is mixed with clay and water or other chemical binders. Next, the specialised machines create the two halves of the mould and join them together to provide a container in which the metals are poured into.
- **Metal melting:** In this process (see 1 in Fig. 1), raw materials are melt and mixed. Molten metal is prepared in a furnace and depending on the choice of the furnace, the quality, the quantity and the throughput of the melt change.
- **Casting and separation:** Once the mixture is made, the molten material is poured into the sand mould. It can be done using various types of ladles or, in high volume foundries, automated pouring furnaces. Later, the metal begins to cool. This step (see 2 in Fig. 1) is really important because the majority of the defects can appear during this phase. Finally, when the casting has been cooled enough to maintain the shape, the casting is separated from the sand. The removed sand is recovered for further uses.
- **Removal of runners and risers:** Some parts of the casting that had been used to help in the previous processes are then removed. They can be detached by knocking off, sawing or cutting.
- **Finishing:** To finish the whole process some actions are usually performed, e.g., cleaning the residual sand, heat treatment and rectification of defects by welding.

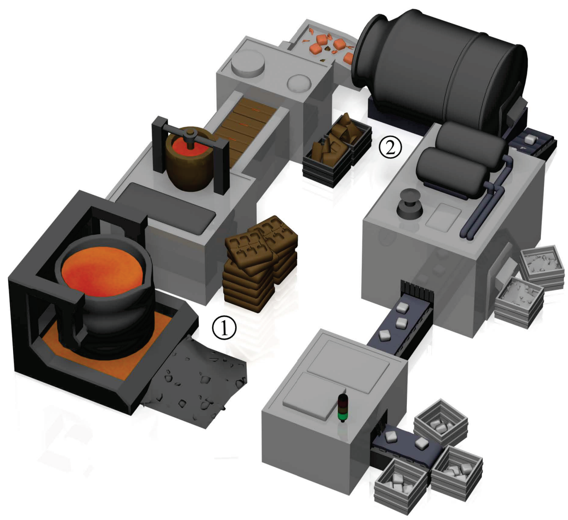

Fig. 1. Foundry process work-flow showing the different phases castings have to pass through. More accurately, in 1 it is showed the metal melting step, and in 2 it is showed the casting preparation and separation step.

The complexity of detecting faulty castings using an *ex-ante* method arises principally from the high number of variables that participate in the production process and, therefore, may influence on the final design of a casting.

In consequence, the foundry process is simplified to solve the aforementioned problem. In our case, the main variables to control in order to predict the faulty castings can be classified into metal-related and mould-related categories. Metal-related variables are divided into the following categories:

- **Composition:** Type of treatment, inoculation and charges [17].
- **Nucleation potential and melt quality:** Obtained by means of a thermal analysis program [18], [19], [20].
- **Pouring:** Duration of the pouring process and temperature.

Mould-related variables can be split into the following categories:

- **Sand:** Type of additives used, sand-specific features and carrying out of previous test or not.
- **Moulding:** Machine used and moulding parameters.

The dimension and geometry of the casting also play a very important role in this practice and, thus, we included several variables to control these two features. We also took into account other parameters regarding the configuration of each machine working in the manufacturing process [21]. In this way, we represent the castings with 24 variables [11].

A casting defect is an irregularity in the casting. Defects are defined as conditions that make a casting to be corrected or rejected. There are several defects that affect metal castings such as, shrinkages, gas porosities or pouring metal defects [16].

In this paper, we deal with microshrinkages. This kind of defects usually appears during the cooling phase of the metal but it cannot be noticed until the production is accomplished. This flaw consists of a form of filamentary shrinkage in which the cavities are very small but large in number and can be distributed over a significant area of the casting, i.e., a minuscule internal porosities or cavities. It is produced because the metals are less dense as a liquid than as a solid and the density of the metal increases and it solidifies while the volume decreases in parallel. During this process, diminutive, microscopically undetectable interdendritic voids may appear leading to a reduction of the castings hardness and, in the cases faced here (where the casting is a part of a very sensitive piece), rendering the piece useless [22].

The way to examine castings is the usage of non-destructive inspections. The most common techniques are X-ray and ultrasound emissions. Unfortunately, both require suitable devices, specialised staff and quite a long time to analyse all the parts. Moreover, every test have to be done once the casting is done. Therefore, post-production inspection is not an economical alternative to the pre-production detection of microshrinkages.

Although we have already obtained overall significant results through a machine-learning-based approach predicting those imperfections [11], [23], [24], [25], [26], [27], [12], [13], these approaches require a manual labour to label every casting within the dataset. This process can be specially time-consuming for several machine-learning models and hinders a subsequent cost increment. Note that in the year 2009, China, which is the biggest producer of castings in the world, produced 35.3 million tons of castings [1] and Europe, the second producer, make 12 million tons of castings [1]. Although not all the castings were labelled, the cost of the foundry workers developing labelling tasks would be too high. Therefore, if only a little piece of the production is labelled, the cost of the prediction preprocessing steps will be reduced.

Therefore, we present here an anomaly-based approach that only requires labelling the correct castings and that measures the deviations of the inspected pieces with these previous labelled castings. Such an approach will indeed reduce the efforts of labelling castings, working with less information available in beforehand.

To this end, as we mentioned before, we manage 24 variables extracted from the foundry process. We apply relevance weights to each characteristic based on Information Gain (IG) [28], which provides a measure for each characteristic that shows its importance to consider whether a casting is valid or not. These weights were calculated from a real dataset acquired from a foundry specialised in safety and precision components for the automotive industry. The dataset is composed of 690 correct castings and 261 faulty castings, and serve to obtain a better distance rating among samples.

III. Anomaly Detection

Through the features described in the previous section, our method represents valid castings as points in the feature space. When a casting is being inspected our method starts by computing the values of the point in the feature space. This point is then compared with the previously calculated points of the valid foundry castings.

To this end, distance measures are required. We have used the following distance measures:

- **Manhattan Distance.** This distance between two points v and u is the sum of the lengths of the projections of the line segment between the points onto the coordinate axes:

$$d(x,i) = \sum_{i=0}^{n} |x_i - y_i| \quad (1)$$

  where x is the first point; y is the second point; and x_i and y_i are the ith component of first and second point, respectively.
- **Euclidean Distance.** This distance is the length of the line segment connecting two points. It is calculated as:

$$d(x,y) = \sum_{i=0}^{n} \sqrt{v_i^2 - u_i^2} \quad (2)$$

  where x is the first point; y is the second point; and x_i and y_i are the ith component of first and second point, respectively.
- **Cosine Similarity.** It is a measure of similarity between two vectors by finding the cosine of the angle between them [29]. Since we are measuring distance and not

similarity we have used $1 - Cosine\ Similarity$ as a distance measure:

$$d(x, y) = 1 - \cos(\theta) = 1 - \frac{\vec{v} \cdot \vec{u}}{||\vec{v}|| \cdot ||\vec{u}||} \quad (3)$$

where $\vec{v}$ is the vector from the origin of the feature space to the first point x, $\vec{u}$ is the vector from the origin of the feature space to the second point y, $\vec{v} \cdot \vec{u}$ is the inner product of $\vec{v}$ and $\vec{u}$. $||\vec{v}|| \cdot ||\vec{u}||$ is the cross product of $\vec{v}$ and $\vec{u}$. This distance ranges from 0 to 1, where 1 means that the two evidences are completely different and 0 means that the evidences are the same (i.e., the vectors are orthogonal between them).

By means of these measures, we are able to compute the deviation of a casting respect to a set of not faulty castings. Since we have to compute this measure with the points representing valid castings, a combination metric is required in order to obtain a final value of distance which considers every measure performed. To this end, our system employs 3 simplistic rules: (i) select the mean value, (ii) select the lowest distance value and (iii) select the highest value of the computed distances. In this way, when our method inspects a casting a final distance value is acquired, which will depend on both the chosen distance measure and combination rule.

IV. Empirical Validation

To evaluate our anomaly-based faulty casting detector, we collected a dataset from a foundry specialised in safety and precision components for the automotive industry, principally in disk-brake support with a production over 45,000 tons a year.

The experiments were focused exclusively in the microshrinkage prediction. Note that, as aforementioned, microshrinkages have internal presence, hence, the evaluation must be done according to non-destructive X-ray, first, and ultrasound testing techniques thenceforth to ensure that even the smallest microshrinkages are found [4].

The acceptance/rejection criterion of the studied models resembles the one applied by the final requirements of the customer (i.e., in the examined cases, the automotive industry). According to the very restrictive quality standards imposed by these clients, pieces flawed with an invalid microshrinkage must be rejected.

In the validation, we worked with two different references, i.e., type of pieces and, to evaluate the proposed method, with the results of the non-destructive X-ray and ultrasound inspections from 951 production stocks performed in beforehand. The dataset comprises 690 correct castings and 261 faulty castings.

Specifically, we followed the next configuration for the empirical validation:

1) **Cross validation:** Despite the small dataset, we had to use as much of the available information to obtain a proper evaluation of the data. To this end, we performed a 5-fold cross-validation [30] over the correct castings to divide them into 5 different divisions of 552 castings for representing normality and 138 for testing. In this way, each fold is composed of 552 not faulty castings that will be used as representation of normality and 399 testing castings, from which 138 are valid castings and 261 are faulty castings.
2) **Calculating distances and combination rules:** We extracted the characteristics described in Section II and employed the 3 different measures and the 3 different combination rules described in Section III to obtain a final measure of deviation for each testing evidence. More accurately, we applied the following distances: (i) the Manhattan Distance, (ii) the Euclidean Distance and (iii) the Cosine Similarity. For the combination rules we have tested the followings: (i) the mean value, (ii) the lowest distance and (iii) the highest value.
3) **Defining thresholds:** For each measure and combination rule, we established 10 different thresholds to determine whether a casting is valid or not.
4) **Testing the method:** We evaluated accuracy by measuring False Negative Ratio (FNR) and False Positive Ratio (FPR).
   In particular, FNR is defined as:
   $$FNR(\beta) = \frac{FN}{FN + TP} \quad (4)$$
   where TP is the number of faulty castings correctly classified (true positives) and FN is the number of faulty castings misclassified as valid castings (false negatives).
   On the other hand, FPR is defined as
   $$FPR(\alpha) = \frac{FP}{FP + TN} \quad (5)$$
   where FP is the number of valid castings incorrectly detected as faulty castings while TN is the number of valid castings correctly classified.

Table I shows the obtained results. Euclidean and Manhattan distances, despite of consuming less processing time, have achieved better results than cosine-similarity-based distance for the tested thresholds. Moreover, our anomaly-based faulty casting detector, for each distance measure, accomplished its best results selecting the mean value for computing the final distance of a casting respect to the not faulty castings. In particular, our detector is able to detect more than 85% of faulty castings (using Manhattan distance), maintaining the rate of misclassified castings lower than 15%. However, all the distances obtained similar results. Euclidean distance achieves more than 84% of accuracy and cosine-similarity-based obtains more than an 83% of accuracy.

Comparing with our previous works focused on microshrinkage [11], [12], the anomaly-based method achieves similar results as many of the previously evaluated classifiers. In fact, this method improves the behaviour of K-nearest neighbour (from lower than 81% of accuracy to higher than 85%). In the same way, the presented approach improves others classifiers such us Bayesian networks learned using K2 and Hill climber

TABLE I
RESULTS FOR DIFFERENT COMBINATION RULES AND DISTANCE MEASURES. THE RESULTS IN BOLD ARE THE BEST FOR EACH COMBINATION RULE AND DISTANCE MEASURE. OUR METHOD IS ABLE TO DETECT MORE THAN 85 % OF THE FAULTY CASTINGS WHILE MAINTAINING FPR LOWER THAN 15 %.

	$1 - Cosine\ Similarity$			$EuclideanDistance$			$ManhattanDistance$		
Combination	**Threshold**	**FNR**	**FPR**	**Threshold**	**FNR**	**FPR**	**Threshold**	**FNR**	**FPR**
Mean	0.10957	0.00000	0.92754	0.12932	0.00000	0.98406	0.23561	0.00000	0.99588
	0.14659	0.06207	0.51014	0.14613	0.03831	0.68406	0.27519	0.03448	0.74928
	0.18361	0.11111	0.25217	0.16293	0.08659	0.36377	0.31478	0.08352	0.38841
	0.22063	0.15709	0.17971	**0.17974**	**0.15172**	**0.27681**	**0.35436**	**0.14789**	**0.29420**
	0.25765	0.16092	0.15797	0.19654	0.15709	0.27246	0.39395	0.26820	0.18116
	0.29468	**0.16782**	**0.14203**	0.21335	0.33487	0.14058	0.43353	0.46207	0.08406
	0.33170	0.27739	0.10435	0.23015	0.61609	0.04638	0.47312	0.65977	0.03913
	0.36872	0.36935	0.05652	0.24696	0.73103	0.01304	0.51270	0.77778	0.01449
	0.40574	0.57701	0.02754	0.26376	0.78697	0.01304	0.55229	0.83525	0.00725
	0.44276	0.72414	0.00000	0.28057	0.89195	0.00000	0.59187	0.93180	0.00000
Maximum	0.43994	0.00000	0.98116	0.22137	0.00000	1.00000	0.50955	0.00000	1.00000
	0.50212	0.04215	0.71449	0.24592	0.03448	0.91594	0.56456	0.01992	0.95942
	0.56431	0.07663	0.58406	0.27048	0.05747	0.78406	0.61956	0.04444	0.81449
	0.62649	0.11418	0.38116	0.29503	0.11877	0.62319	0.67457	0.12337	0.56957
	0.68867	0.15479	0.24638	0.31959	0.13793	0.41449	**0.72957**	**0.26897**	**0.32319**
	0.75086	0.16092	0.17826	**0.34414**	**0.18008**	**0.26522**	0.78458	0.49579	0.14783
	0.81304	**0.16092**	**0.16232**	0.36870	0.34483	0.20000	0.83958	0.68966	0.07826
	0.87522	0.17241	0.15652	0.39325	0.66284	0.08696	0.89549	0.81456	0.01884
	0.93741	0.36935	0.11739	0.41781	0.79464	0.03333	0.94959	0.92107	0.00870
	0.99959	0.99080	0.00000	0.44236	0.98161	0.00000	1.00460	0.98161	0.00000
Minimum	0.00018	0.00000	0.83913	0.00423	0.00000	0.92754	0.00740	0.00000	1.00000
	0.01207	**0.39923**	**0.07826**	**0.01604**	**0.22146**	**0.23913**	0.11820	0.71034	1.00000
	0.02396	0.59387	0.03188	0.02785	0.36398	0.14493	0.22900	0.99004	1.00000
	0.03585	0.70728	0.01594	0.03967	0.49732	0.08986	0.33980	1.00000	1.00000
	0.04774	0.71648	0.00870	0.05148	0.63908	0.06957	0.45060	1.00000	1.00000
	0.05962	0.72567	0.00435	0.06329	0.71571	0.02609	0.56140	1.00000	0.96377
	0.07151	0.72567	0.00435	0.07510	0.72261	0.01739	0.67220	1.00000	0.58261
	0.08340	0.73333	0.00290	0.08692	0.72567	0.00580	0.78300	1.00000	0.15507
	0.09529	0.73563	0.00145	0.09873	0.73410	0.00290	0.89380	1.00000	0.02174
	0.10718	0.73563	0.00000	0.11054	0.73563	0.00000	**1.00460**	**1.00000**	**0.0000**

algorithms and Naïve Bayes. Otherwise, the Bayesian networks learnt with Tree Augmented Naïve (TAN) and Artificial Neural Networks (ANN), using Multilayer Perceptron and Backpropagation algorithm achieved a close behaviour to this anomaly-based approach. Anyhow, this method cannot improve other classifiers such as Support Vector Machines, or Decision Trees: higher than 90% of accuracy.

Besides, FNR and FPR established the cost of misclassification. It is important to set the cost of false negatives ($1-FNR$) and false positives, in other words, establish whether is better to classify a faulty casting as correct casting or to classify a correct casting as a casting with a microshrinkage. In particular, since our method is developed to predict castings as correct or faulty, one may think that is more important to minimise the false negative ratio (castings which have a microshrinkage but we do not detect in beforehand) than to fail indicating that our well done casting may have a microshrinkage. In fact, this requirement comes from foundries because they adopt a conservative point of view to assure their cost ratios.

Even though the system has not raised a 100% accuracy level, it has achieved significant results (more than 85%) that may positively impact the production of high precision foundries. Our method reduces in a significant way the cost and the duration of the actual testing methods and, also, detects microshrinkages in beforehand.

To include the anomaly-based detection method in foundries, the behaviour of the system can be performed in the following way: (i) when the system detects that a microshrinkage ought to appear, the operator may change the factors to produce this reference or the reference (and, thus, skip the cost of having to manufacture it again) or (ii) being a part of a Model Predictive Control system [31], specifically, inside the prediction step. Using the second approach the system can foresee defects and determine the required changes to solve them in an automatic or semi-automatic way.

V. DISCUSSION AND CONCLUSIONS

Foreseeing the apparition of microshrinkages in ductile iron castings is one of the most hard challenges in foundry-related research. Our work in [3], [11] pioneered the application of Artificial Intelligence to the prediction of microshrinkages.

This time, our main contribution is the anomaly-detection-based approach employed for miscroshrinkage detection. In contrast to our previous approaches, this method only needs to previously label the correct castings and measures the deviation of the castings under inspection respect to normality (castings without the microshrinkage defect). Although anomaly detection systems tend to produce high error rates (specially, false positives), our experimental results show low values. This proofs the validity of our initial hypothesis.

Anyway, it presents several limitations that should be studied in a further work. Firstly, we cannot identify different levels of warnings as we did in our previous works. In this case, we only can classify the castings as correct or faulty. Nevertheless, we could compute it using another anomaly detection techniques such as clustering-based or nearest-neighbour-based anomaly detection.

Secondly, this kind of techniques based on the measurement of distances cannot achieve good results if the training data is disperse. In other words, if the normality cannot be represented as a compact group of instances, the threshold that allows to divide the instances into correct and faulty cannot be set correctly. Nevertheless, this fact can be solved due to the nature of the production process: all the castings are always produced in similar way. Hence, generated vectors of castings are close to each other, representing the normality easily and, therefore, allowing to measure the distances between correct and faulty castings.

Finally, it is important to consider efficiency and processing time. Our system compares each casting against a relative big dataset (244 vectors for each fold). Despite Euclidean and Manhattan distances are easy to compute, cosine distance and more complex distance measures such as Mahalanobis distance may take too much time to process every casting under analysis. For this reason, in further work we will emphasise on improving the system efficiency by reducing the whole dataset to a limited amount of samples which is sufficiently representative.

REFERENCES

[1] "44^{th} census of world casting production," Modern Casting, Tech. Rep., 2010, monthly report, edited annually with the data concerning the number of casting houses and the world casting production in the year preceding the issue.

[2] J. Sertucha, A. Loizaga, and R. Suárez, "Improvement opportunities for simulation tools," in *Proceedings of the 16th European Conference and Exhibition on Digital Simulation for Virtual Engineering*, 2006, invited talk.

[3] Y. Penya, P. García Bringas, and A. , Zabala, "Advanced fault prediction in high-precision foundry production," in *Proceedings of the* 6^{th} *IEEE International Conference on Industrial Informatics*, 2008, pp. 1673–1677.

[4] A. Zabala, R. Suárez, and J. Izaga, "Iron castings, advanced prediction tools, foundry process control and knowledge management," in *Proceedings of the* 68^th *WFC - World Foundry Congress*, 2008, pp. 355–360.

[5] A. Lazaro, I. Serrano, J. Oria, and C. de Miguel, "Ultrasonic sensing classification of foundry pieces applying neuralnetworks," in 5^th *International Workshop on Advanced Motion Control*, 1998, pp. 653–658.

[6] P. Zhang, Z. Xu, and F. Du, "Optimizing casting parameters of ingot based on neural network and genetic algorithm," in *ICNC '08: Proceedings of the 2008 Fourth International Conference on Natural Computation*. Washington, DC, USA: IEEE Computer Society, 2008, pp. 545–548.

[7] M. Perzyk and A. Kochanski, "Detection of causes of casting defects assisted by artificial neural networks," in *Proceedings of the Institution of Mechanical Engineers Part B Journal of Engineering Manufacture*, vol. 217, no. 9, 2003, pp. 1279–1284.

[8] H. Bhadeshia, "Neural networks in materials science," *ISIJ international*, vol. 39, no. 10, pp. 966–979, 1999.

[9] H. Peter and J. Wang, "Fault detection using the k-nearest neighbor rule for semiconductor manufacturing processes," *IEEE Transactions on Semiconductor Manufacturing*, vol. 20, no. 4.

[10] M. Vasudevan, M. Muruganath, and A. K. Bhaduri, "Application of bayesian neural network for modelling and prediction of ferrite number in austenitic stainless steel welds," *ser. Mathematical Modelling of Weld Phenomena - VI. London: Institute of Materials*, pp. 1079–1100, 2002.

[11] I. Santos, J. Nieves, Y. K. Penya, and P. G. Bringas, "Optimising machine-learning-based fault prediction in foundry production." in *Proceedings of the* 2^{nd} *International Symposium on Distributed Computing and Artificial Intelligence (DCAI),. S. Omatu et al. (Eds.): IWANN 2009, Part II, LNCS 5518, Springer-Verlag Berlin Heidelberg*, 2009, pp. 553–560.

[12] I. Santos, J. Nieves, and P. Bringas, "Enhancing fault prediction on automatic foundry processes," in *World Automation Congress (WAC), 2010*. IEEE, 2010, pp. 1–6.

[13] I. Santos, J. Nieves, , Y. K. Penya, and P. G. Bringas, "Towards noise and error reduction on foundry data gathering processes." in *Proceedings of the International Symposium on Industrial Electronics (ISIE)*, 2010, 1765—1770.

[14] R. Fujimaki, T. Yairi, and K. Machida, "An approach to spacecraft anomaly detection problem using kernel feature space," in *KDD '05: Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*. New York, NY, USA: ACM, 2005, pp. 401–410.

[15] G. Hollier and J. Austin, "Novelty detection for strain-gauge degradation using maximally correlated components," in *Proceedings of the European Symposium on Artificial Neural Networks, ESANN2002, Bruges*, 2002, pp. 257–262.

[16] S. Kalpakjian and S. Schmid, *Manufacturing engineering and technology*, 2005.

[17] C. J. F. and R. Ríos, "A fracture mechanics study of nodular iron," *Revista de Metalurgía*, vol. 35, no. 5, pp. 279–291, 1999.

[18] R. Gonzaga-Cinco and J. Fernández-Carrasquilla, "Mecanical properties dependency on chemical composition of spheroidal graphite cast iron," *Revista de Metalurgia*, vol. 42, pp. 91–102, March–April 2006.

[19] M. Hecht and F. Condet, "Shape of graphite and usual tensile properties of sg cast iron: Part 1," *Fonderie, Fondeur d'aujourd'hui*, vol. 212, pp. 14–28, 2002.

[20] P. Larrañaga, J. Sertucha, and R. Suárez, "Análisis del proceso de solidificación en fundiciones grafíticas esferoidales." *Revista de Metalurgia*, vol. 42, no. 4, pp. 244–255, 2006.

[21] J. Sertucha, R. Suárez, J. Legazpi, and P. Gacetabeitia, "Influence of moulding conditions and mould characteristics on the contraction defects appearance in ductile iron castings." *Revista de Metalurgia*, vol. 43, no. 2, pp. 188–195, 2007.

[22] "Inoculation alloy against microshrinkage cracking for treating cast iron castings," patent US 2005/0180876 A 1.

[23] J. Nieves, I. Santos, Y. K. Penya, S. Rojas, M. Salazar, and P. G. Bringas, "Mechanical properties prediction in high-precision foundry production." in *Proceedings of the* 7^{th} *IEEE International Conference on Industrial Informatics (INDIN 09)*, 2009, pp. 31–36.

[24] I. Santos, J. Nieves, Y. K. Penya, and P. G. Bringas, "Machine-learning-based mechanical properties prediction in foundry production," in *In Proceedings of ICROS-SICE International Joint Conference (ICCAS-SICE)*, 2009, pp. 4536–4541.

[25] I.Santos, J. Nieves, P. Bringas, and Y. Penya, "Machine-learning-based defect prediction in highprecision foundry production," in *Structural Steel and Castings: Shapes and Standards, Properties and Applications*, L. M. Becker, Ed. Nova Publishers, 2010, pp. 259–276.

[26] J. Nieves, I. Santos, and P. Bringas, "Overcoming data gathering errors for the prediction of mechanical properties on high precision foundries," in *World Automation Congress (WAC), 2010*. IEEE, 2010, pp. 1–6.

[27] J. Nieves, I. Santos, Y. K. Penya, F. Brezo, and P. G. Bringas, "Enhanced foundry production control," in *Proceedings of the 21st International Conference on Database and Expert Systems Applications (DEXA). Lecture Notes in Computer Science 6262, Springer-Verlag Berlin Heidelberg*, 2010, pp. 213–220.

[28] J. Kent, "Information gain and a general measure of correlation," *Biometrika*, vol. 70, no. 1, pp. 163–173, 1983.

[29] S. Tata and J. Patel, "Estimating the Selectivity of tf-idf based Cosine Similarity Predicates," *SIGMOD Record*, vol. 36, no. 2, pp. 75–80, 2007.

[30] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proceedings of the International Joint Conference on Artificial Intelligence*, vol. 14, 1995, pp. 1137–1145.

[31] M. Morari and J. H. Lee, "Model predictive control: Past, present and future," *Computers and Chemical Engineering*, vol. 23, pp. 667–682, 1997.