https://doi.org/10.2352/ISSN.2470-1173.2018.09.IRIACV-277


# Part Quality Assessment using Convolution Neural Networks in High Pressure Die Casting

***Kelly Cashion, Nilesh Powar, Robert De Neff, Robert Kress; University of Dayton Research Institute, Dayton, Ohio; Honda North America, Marysville, Ohio; Honda Transmission Manufacturing, Russells Point, Ohio***

## Abstract

*High pressure die casting (HPDC) has been developed since the late nineteenth century for a breadth of manufacturing applications. The process forms molten metal into molds at high temperatures given a complex array of parameters and variables that are challenging to observe. We used a set of thermal cameras to capture imagery of the die used as a mold during its cooling process between part productions. This data was used to train a convolution neural network to assess the quality of the part just produced based on the thermal characteristics of the surface of the die. The system achieved 90% accuracy when distinguishing between parts that met quality standards and parts that did not.*

## Introduction

High pressure die casting (HPDC) is used to create aluminum, zinc, magnesium, and lead alloy parts by pouring or injecting molten metal into a die (mold). The process has been widely used in manufacturing since its initial development for typecasting in the 19th century, especially to produce high volume components [3]. Although years of practice has developed a robust base of die casting expertise, much of the process remains arcane due to the difficulty of observing the complex array of parameters in real time. While the process has been refined over the last century, there are still a numerous unknowns within any given casting cycle. In the worldwide push toward smart factories [1, 2, 4, 7, 10], HPDC is a unique challenge due to the intensity of the environment that needs to be observed. Any sensors integrated into an HPDC machine must not only be able to withstand the high temperatures but also observe accurately despite the harsh environment created by the process.

To offer additional insights into the process — with the objective of increasing efficiency and reducing the production of defective parts — we developed a real time thermal monitoring system that can be integrated directly into the die casting process. Data collected by this system was then used in conjunction with the results of downstream part quality testing to train a model that can assess if a part that has just been cast is likely to be of good quality. The remaining sections of this paper delve into the context of our task, the specifics of our method, and the results we have achieved so far.

## Context

### *High Pressure Die Casting*

The objective of HPDC is to produce good quality parts. In order to do so, process parameters have to be carefully managed to efficiently create high quality parts. These parameters include: injection speed, injection pressure, die temperature, melt pressure, flow rate, metal temperature, and alloy chemistry [10]. Managing these parameters is a challenging task that has historically been done using the expertise of engineers and traditional statistical approaches like taguchi methods [8]. In recent years, there has been a growing push towards smart factories that integrate advanced data analytics into manufacturing processes to improve quality and increase efficiency. These efforts include integration of data analytics into die casting as discussed in [1, 2, 4, 7, 10].

In order to understand where data analytics best fits in the die casting process, an overview of the process itself is necessary. A high level depiction is shown in Figure 1 [3].

In high pressure die casting, molten metal is added to the die which has a cavity that acts as a mold for a part. The time it takes for the metal to solidify, as well as the final material characteristics of the part being cast, are heavily dependent on the thermal characteristics not only of the metal but also of the die surface [10]. Failure to optimize these parameters can not only cause the process to be inefficient but can also lead to several types of defects in the final part [1]. Common defects include gas and air porosity, inclusions, shrinkage, and cracks [1]. Porosity defects in particular are well tested and quantified as part of the normal manufacturing process [1]. For our work, we will focus in particular on porosity defects because their close correlation with die surface temperature [1, 10] and also their prevalence in die casting production.

### *Thermal Imaging*

Thermal imaging is widely used for heat sensitive and low-light applications where infrared signals are relevant. Unlike a standard camera, thermal cameras are not impacted by light (barring significant associated heat), however there are several other key characteristics of the environment and the observed target that impact the infrared signals and must be accounted for in order to obtain accurate measurements [9]. These parameters include: (1) atmospheric temperature, (2) window temperature, (3) window transmission, (4) reflected temperature, (5) object distance, (6) relative humidity, and (7) emissivity. These parameters are diagrammed in Figure 2.

The window refers to an external lens. The transmission rating of the window and the emissivity of the observed object are both reflective of material properties of each respectively [9]. By accounting for these parameters accurately, a thermal camera can provide accurate measures of observed temperatures within its hardware specifications.

Thermal imaging has particular applications in die casting and has been utilized to observe the critical thermal characteristics of die surfaces and other high heat processes in manufacturing [4, 9]. We build on that work by integrating real time thermal camera observations and advanced image processing techniques

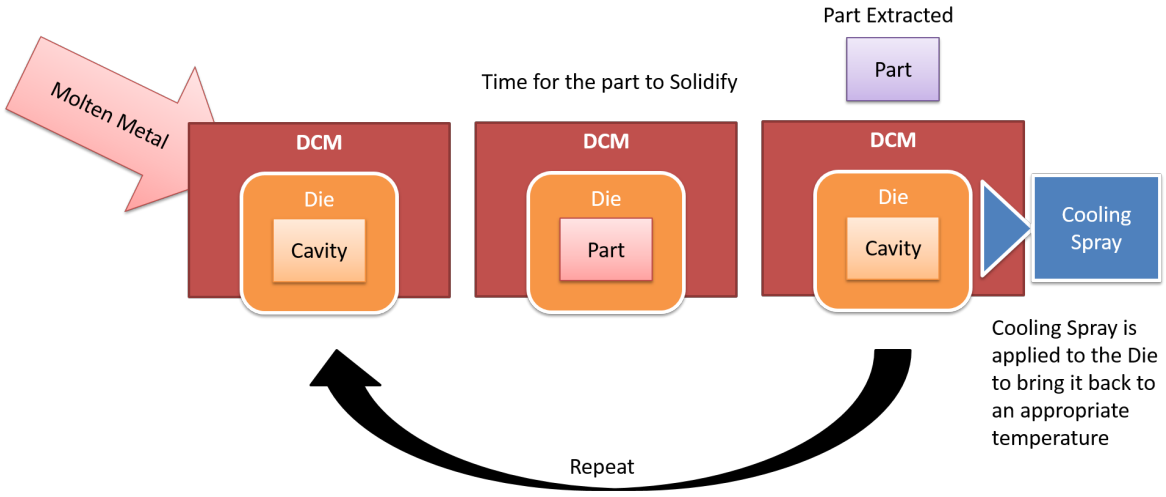


Figure 1: Die Casting Cycle

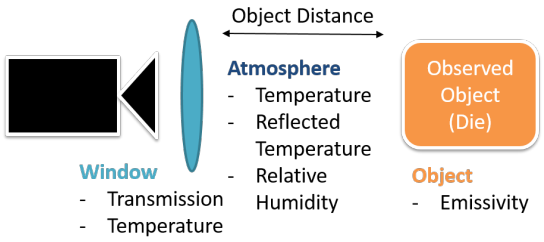


Figure 2: Thermal Parameters

to determine part quality at cast time.

## Method

Our method for HPDC analysis starts with an automated thermal imaging software controlling two radiometric thermal cameras. To collect accurate data, we tune the thermography parameters to account for environment variables and material properties in order to achieve accurate temperature readings. Finally, we take that data and train a convolution neural network (CNN) to infer the quality of the produced part based on the thermal images of the die. An overview of how our approach fits into the traditional manufacturing process is shown in Figure 3.

Traditionally, parts extracted from die casting go to machining, which adds significant value to the part. From machining they are then sent to quality control to test for a range of defects including porosity defects. Any parts with porosity above a given threshold are discarded losing all the value added during machining in addition to the original cost of producing the part in casting. Our proposed method uses a thermal camera system to capture images at the time the part is first cast and uses those images to determine if the part has a porosity defect. With our system, the machining phase can be bypassed for these parts saving significant time and materials.

### *Image Capture*

Our data capture system leverages two thermal cameras each positioned to view a different side of the die. For our experiment we have one camera positioned to look at the moving side of the die - the side that opens to remove the part - and one positioned to look at the fixed side. These two angles capture about 90% of the die surface and include a view of all critical components. During the die casting cycle after a part has been taken out of the die, but before new metal is poured in, a cooling agent is applied to the die as shown in Figure 1 [3]. Our system captures thermal images of the die shortly before and shortly after this cooling spray. This allows us to see not only the heat signature of the die but also observe if it is cooling evenly.

### *Thermography Parameter Tuning*

In order to calculate accurate temperature values based on the infrared light received by the cameras, we must account for the effects of various environment parameters and material properties. The thermography parameters we configure are: (1) atmospheric temperature, (2) window temperature, (3) window transmission, (4) reflected temperature, (5) object distance, (6) relative humidity, and (7) emissivity [9]. To tune the parameters, we performed an experiment where we collected temperature data off a die at 10 degree intervals between 100 °C and 300 °C. During this experiment we used a temperature probe to measure the temperature of the die at four key points around the main heating coils. Based on this acquired ground truth and the captured thermal images, we used a genetic algorithm to test different combinations of six of the parameters (object distance was held constant) and determine the best combination to optimize our accuracy within the desired temperature range.

### *Data Analysis*

With an automated data capture system integrated into several die cast machines and the themography parameters tuned, we were able to start learning from the data. We collected two and a half months of radiometric thermal images. Each part created during this time was also tested for several possible defects including porosity and leaks as part of the standard manufacturing process. Experts link the issues of high porosity and off scale leaks to abnormal thermal characteristics in the die casting process [1]. Given this knowledge and using the test results for each part as a ground truth, we identify two defect cases of interest - POR (high porosity) and OFF_SCALE_LEAKER (off scale leaks). We limit our data to one part model and remove other types of failure cases

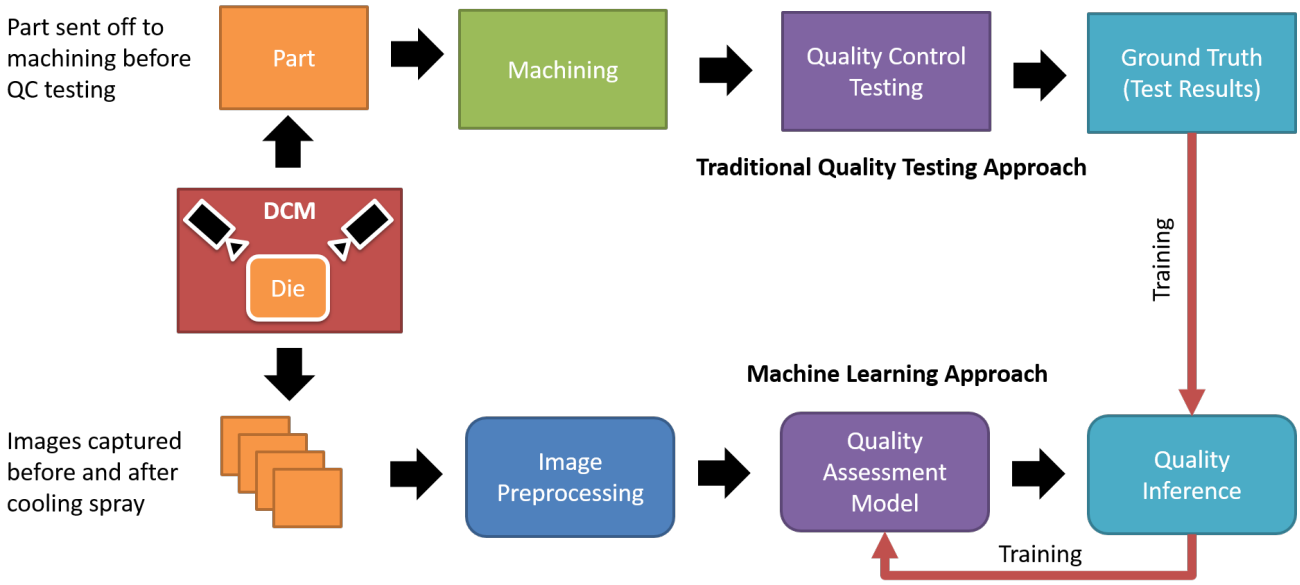


Figure 3: Die Casting Inspection and Data Analytics

that cannot be clearly linked to the die casting process.

After this initial filtering of the experimental data set, we have 671,648 images across 167,912 cycles including 166,187 cycles of GOOD parts, 706 cycles of POR (porosity) parts, and 1,019 cycles of OFF_SCALE_LEAKER (off scale leak) parts. Each cycle / part has four images - moving before spray, moving after spray, fixed before spray, and fixed after spray. These images are arranged in a 2x2 grid. Each individual image is an RGB image of relative temperature values utilizing an iron-bow color map. The individual images are 320 x 256 pixels. The combined image is 640 x 512 pixels.

We design a CNN classifier architecture initially targeting three output classes - GOOD, POR, OFF_SCALE_LEAKER. CNN is chosen due to the wide success of the algorithm in image processing applications [5]. Our CNN utilizes common layers: convolution with zero-padding and rectified linear (ReLu) activation, max pooling, and fully connected. An overview of our architecture is shown in Figure 4

Each convolution layer uses 32 3x3 filters with stride of 1, zero padding of 1 and rectified linear activation. The max pooling layer utilizes a 2x2 kernel with stride of 2. The fully connected layer has 128 nodes. Finally, the output layer has three output classes - POR, OFF_SCALE_LEAKER, and GOOD.

## Results

Our data analysis is split into three phases: our initial approach, a refined approach, and a generalized approach. In all three cases we used the same (with the exception of the output layer) CNN architecture and hyper-parameters for classification. The initial approach looked at discretely identifying all three classes - GOOD, POR (porosity), and OFF_SCALE_LEAKER. The refined approach changed our output layer to distinguish between GOOD and BAD parts. Finally, our generalized approach took a broader subset of the data to improve GOOD classification rates.

### *Initial Approach*

Initially, we looked at both our failure cases - high porosity (POR) and off scale leaker (OFF_SCALE_LEAKER) - as independent classes. It is important to remember that OFF_SCALE_LEAKER is a more significant version of a POR (porosity) defect. The three classes can be considered broad categories along the same scale, with GOOD reflecting any parts that test below a given porosity range, POR reflecting any parts that test above an initial threshold but below that of a full scale leaker, and OFF_SCALE_LEAKER reflects any parts that pass a significant threshold of porosity, dramatically failing the leak test.

After removing incomplete image sets, our data set contains 112,534 GOOD, 500 POR, and 701 OFF_SCALE_LEAKER. The discrepancy between the GOOD image count and the BAD image count is approximately 100 to 1. To accommodate, we significantly down-sampled our set of GOOD images. Up-sampling of the POR and OFF_SCALE_LEAKER was also considered, but resulted in severe over-training. For the initial approach, we separated out a test set of 100 GOOD, 100 POR, and 100 OFF_SCALE_LEAKER image sets. Our training set was 600 GOOD images - selected from a common subset of our 2.5 month period - along with the remaining 400 POR images and 500 of the OFF_SCALE_LEAKER images.

After 100 epochs, our initial algorithm performed with an accuracy of 70.3% on the test set. The accuracy/ recall is broken down for each subcategory - GOOD and BAD - and by nuanced class - GOOD, POR, and OFF_SCALE_LEAKER. The results are shown in Figures 5 and 6.

As can be seen in both figures, the GOOD classification rate is high - 93%. The results shown in Figure 6 further indicate that the OFF_SCALE_LEAKER is comparable at 87%. Our loss in accuracy largely comes from poor recall in the POR class. To break this down further, we look at the confusion matrix shown in Table 1.

From Table 1 we can clearly discern that the confusion in classifications lies predominantly in POR image sets being misclassified as OFF_SCALE_LEAKER image sets. We hypothe-

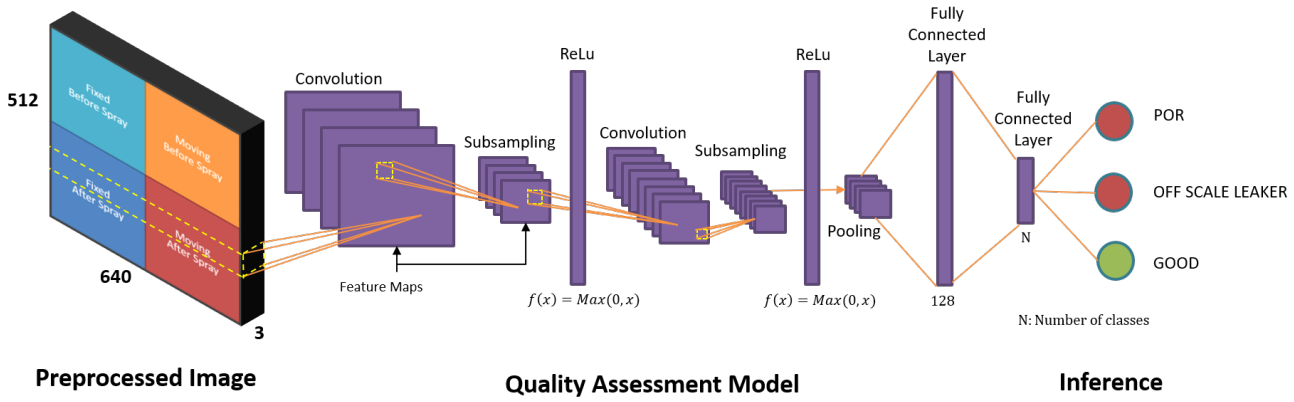


Figure 4: Algorithm Architecture

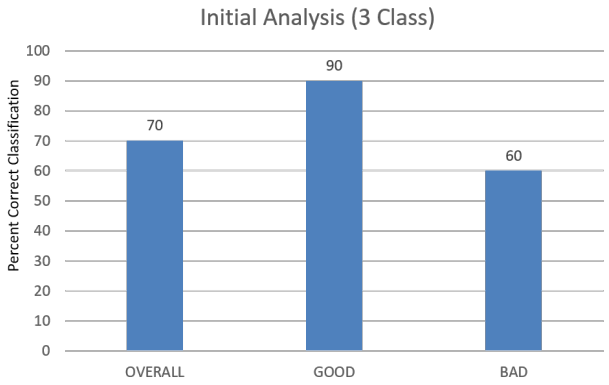


Figure 5: Initial Approach - Accuracy by Subcategory


Figure 6: Initial Approach - Accuracy by Class

size that the characteristics in the images that distinguish POR and the higher tier OFF_SCALE_LEAKER classifications are extremely similar due to the relationship between those two failure types - OFF_SCALE_LEAKER is simply a more extreme version of the POR defect. Due to our small image set, our algorithm does not have enough training examples to distinguish clearly between these two classes. However, the confusion matrix also shows us that not only were very few GOOD images misclassified as either type of BAD image, but none of the POR or OFF_SCALE_LEAKER images were misclassified as GOOD. We quantify this performance by calculating recall, precision and F-score for each class as described in [6]. The results are shown in Table 2.

The recall and precision for both POR and OFF_SCALE_LEAKER are low, resulting in poor F Scores of 0.43 and 0.66 for each class respectively. However, the recall and precision of GOOD are extremely high resulting in an excellent F-Score of 0.96. Despite the poor performance of the algorithm in identifying bad images, this iteration of the model can be utilized for determination of GOOD parts.

### *Refined Approach*

Based on our results in the initial approach, we adjust the algorithm to look at only two classes - GOOD and BAD. The primary objective of our research is the ability to distinguish between the GOOD and BAD, i.e., to determine the worth of a part before the added expense of further machining. For this refined approach we not only utilize the same architecture (with exception of the output layer class number), but also utilize the same training and test sets as for our initial approach. The results broken down by accuracy / recall at both the categorical level (GOOD and BAD) and the nuanced class level (GOOD, POR, and OFF_SCALE_LEAKER) are shown in Figures 7 and 8 respectively.

Unsurprisingly, given the clear distinction between GOOD images and the combined POR and OFF_SCALE_LEAKER set

Table 1: Initial (3 Class) Confusion Matrix

| | Predicted | | |
|---|---|---|---|
| Actual | POR | OFF | GOOD |
| POR | 31 (31.0%) | 69 (69.0%) | 0 (0.0%) |
| OFF | 13 (13.0%) | 87 (8.07%) | 0 (0.0%) |
| GOOD | 1 (1.0%) | 6 (6.0%) | 93 (93.0%) |

Table 2: Initial Approach Metrics

| | Recall | Precision | F-Score |
|---|---|---|---|
| POR | 0.31 | 0.69 | 0.43 |
| OFF | 0.87 | 0.54 | 0.66 |
| GOOD | 0.93 | 1.00 | 0.96 |
| AVERAGE | 0.70 | 0.74 | 0.69 |

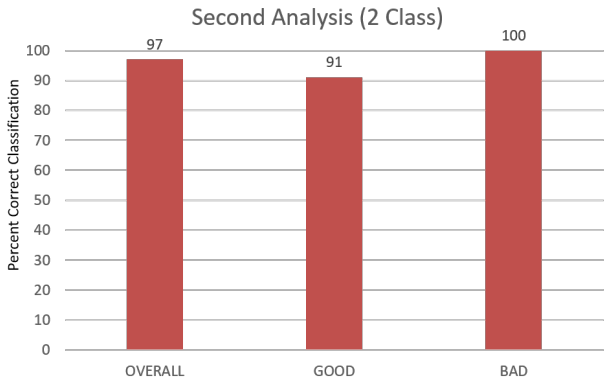

Figure 7: Refined Approach - Accuracy by Subcategory

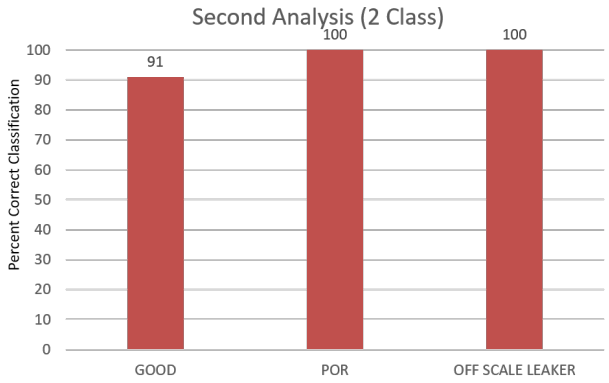

Figure 8: Refined Approach - Accuracy by Class

in the initial approach, our accuracy is greatly improved in this two class case. The overall accuracy is 95.7% with the GOOD accuracy dropping slightly to 87% and the classification of both POR and OFF_SCALE_LEAKER images as BAD at 100%. Our confusion matrix is shown in Table 3 and our precision and recall for both the GOOD and BAD classes is shown in Table 4.

Table 3: Refined Confusion Matrix

| | Predicted | |
|---|---|---|
| Actual | BAD | GOOD |
| BAD | 200 (100%) | 0 (0.0%) |
| GOOD | 13 (13.0%) | 87 (87.0%) |

Table 4: Refined Approach Metrics

| | Recall | Precision | F-Score |
|---|---|---|---|
| BAD | 1.00 | 0.94 | 0.97 |
| GOOD | 0.87 | 1.00 | 0.93 |
| AVERAGE | 0.94 | 0.97 | 0.95 |

With this refined experiment both our BAD images and GOOD images have high F-Scores of 0.97 and 0.93 respectively. It also remains the case that the GOOD images in this set have a perfect precision on the test set, just as in the initial results. However, for a final application, it would be better to tune the algorithm to maximize the precision of BAD image set classifications while maintaining a good recall on bad parts in order to ensure that no GOOD parts are mistakenly scrapped. In that case, the precision of GOOD parts might decrease, but still greatly reduce the number of POR and OFF_SCALE_LEAKER parts, saving significant machining costs.

### *Generalized Approach*

Toward the goal of finding greater distinction between GOOD and BAD parts we look at a more generalized subset of the GOOD images. We still down-sample the GOOD image set significantly - from 112,534 to 900 training images and 100 testing images. Additionally, we sample the GOOD, POR, and OFF_SCALE_LEAKER across the entire 2.5 months in order to split training and test sets (90:10). For the GOOD images, this results in a broader set of variations in candidates offering a greater generalization of our algorithm. The architecture remains the same and we continue looking at the problem as two rather than three class. The accuracy of each subcategory and subclass are shown in Figures 9 and 10 respectively.

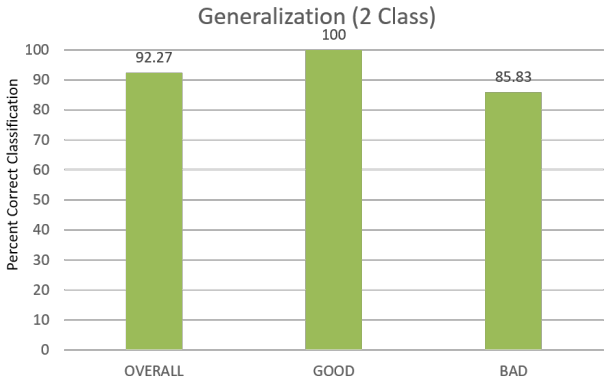

Figure 9: Generalized Approach - Accuracy by Subcategory

Figure 10: Generalized Approach - Accuracy by Class

The overall accuracy drops to 81.4%. This result is lower than the refined approach but still higher than our initial three class approach. Breaking it down into the confusion matrix and metrics we can look at the nuances of recall and precision for both the GOOD and BAD cases. The confusion matrix and metrics are shown in Tables 5 and 6 respectively.

Consider our sample objective from the refined analysis - optimize around BAD precision without sacrificing too much recall. In Table 6, we can see that the precision for BAD images and

Table 5: Generalized Confusion Matrix

| | Predicted | |
|---|---|---|
| Actual | BAD | GOOD |
| BAD | 83 (69.2%) | 37 (30.8%) |
| GOOD | 4 (4.0%) | 96 (96.0%) |

Table 6: Generalized Approach Metrics

| | Recall | Precision | F-Score |
|---|---|---|---|
| BAD | 0.69 | 0.95 | 0.80 |
| GOOD | 0.96 | 0.72 | 0.82 |
| AVERAGE | 0.83 | 0.84 | 0.81 |

the recall for GOOD (respectively) have both increased from the refined approach results. This increase in GOOD classification recall is very positive - indicating that we are far less likely to scrap GOOD parts based on our algorithm. The drop in recall for BAD and respectively precision for GOOD is indicative of our limited BAD samples compared to GOOD samples in our training set. While better at generalizing against variations in the GOOD samples, we do not have the same kind of data to draw on to further generalize the BAD samples. This shortcoming directly inspires some of our next steps.

## Conclusion

Our results to date indicate that there are distinct characteristics that the CNN we've trained can use to distinguish between a good part and a part with a porosity or leak based defect. However, our results also demonstrated the limitations of our current data-set - robust in good images, but limited for defect samples. Our continuing work begins with the acquisition of additional defect samples or the development of a tool to simulate defects for training. From there, the next steps will be further refinement of the approach and the integration of additional process parameters and sensor data. The more information we can acquire about each die casting cycle, the better we will be able to transition from simply classifying a part's quality to informing the die casting engineers why quality was bad for a given cycle. Extracting the 'why' will take this application from simple cost savings on machining to real process improvement.

## References


[1] Franco Bonollo and Nicola Gramegna, The MUSIC guide to key-parameters in High Pressure Die Casting, Assomet srl and Enginsoft SpA (2014).

[2] Nicola Gramegna and Franco Bonollo, Smart Control and Cognitive System Applied to the HPDC Foundry 4.0: A Robust and Competitive Methodology Developed Under EU-FP7 Music Project (2016).

[3] Alan Kaye and Arthur Street, Die Casting Metallurgy: Butterworths Monographs in Materials. Elsevier, 2016.

[4] L. X. Kong, F.H. She, W. M. Gao, S. Nahavandi, and P.D. Hodgson, Integrated optimization system for high pressure die casting processes. Journal of materials processing technology, 201 (1-3), pgs. 629-634 (2008)

[5] Weibo Liu, Zidong Want, Ziaohui Liu, Nianyin Zeng, Yurong Liu, and Fuad E. Alsaadi, A survey of deep neural network architectures and their applications, Neurocomputing, 234, pgs. 11-26 (2017)

[6] David Martin Powers, Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation (2011).

[7] Michael Rix, Bernd Kujat, Tobias Meisen, and Sabina Jeschke, An agile information processing framework for high pressure die casting applications in modern manufacturing systems, In Automation, Communication and Cybernetics in Science and Engineering, Springer, Cham, pgs. 957-969 (2016).

[8] G. P. Syrcos, Die casting process optimization using Taguchi methods, Journal of materials processing technology 135, 1, pgs. 68-74 (2003).

[9] Thomas Williams, Thermal imaging cameras: characteristics and performance. CRC Press. (2009)

[10] Prasad KDV Yarlagadda and Eric Cheng Wei Chiang, A neural network system for the prediction of process parameters in pressure die casting, Journal of Materials Processing Technology, 89, pgs. 583-590 (1999).


## Author Biography

*Kelly A. Cashion received her B.S. in Computer Engineering from the University of Dayton (UD) in 2007 and her M.S. in Electrical Engineering from UD in 2016. She has worked for the UD Research Institute (UDRI) for 5 years as a software engineering, machine learning, and deep learning scientist. Her work includes prototyping and testing a non-line of sight tool positioning system and development of a robust thermal imaging system for high pressure die casting.*

*Nilesh U. Powar received his Bachelor of Engineering in Electronics from University of Bombay, India in 1999, a M.S. in Computer Engineering from Wright State University in 2002, and a PhD degree in engineering from UD in 2013. He is currently a senior researcher at UDRI and has 15+ years of experience in the field of image processing, statistical pattern recognition, machine learning, and system integration.*

*Robert De Neff received his B.S. in Manufacturing Engineering from Western Michigan University in 1986. He has worked in the field of Automotive Aluminum Casting for the past 32 years. He is a member of the American Foundry Society and is currently serving on the North American Die Casting Association Board of Governors.*

*Robert M. Kress received his B.S. in Mechanical Engineering from The Ohio State University in 2012. He has worked at the Honda Transmission Manufacturing plant in the High Pressure Die Cast department for the past 5 years. The first 2 years were spent as a quality engineer, and the last 3 years have been with the equipment team as an equipment engineer.*