DBS Business Review 05 | 2023
ARTICLE | dbsbusinessreview.ie

Comparative Study of Image Processing Algorithms to Detect Defects in Cast Components

Nikith Muralidhar
Postgraduate
MSc Data Analytics, Dublin Business School
Dublin, Ireland

Abstract

The non-destructive evaluation (NDE) of cast components is essential in the industrial sector. These cast components are prone to abnormalities like blowholes. Including such faults in the components reduces the fatigue life, which increases the risk of catastrophic accidents. Currently, humans analyse cast components using a variety of manual techniques. In order to create a category that would do away with manual testing, we suggest an automatic method for identifying flaws in casts. The method uses Convolutional Neural Networks (CNN) and Support Vector Classifiers (SVC) to determine whether a cast component has a defect or not and looks for them. The strategy may be advantageous to human examiners because it lightens their burden and greatly minimises the possibility of using defective parts in manufactured goods, according to the hypothesis.
Keywords: Image classification, CNN, Deep learning, Machine Learning, Defect detection.

Introduction

Quality control is a crucial component of all industrial processes. Due to competition in the manufacturing industry, manufacturers must increase their product yield while adhering to tight quality control guidelines. To address the growing demand for high-quality products, intelligent visual inspection equipment is essential in manufacturing lines (Ferguson et al., 2017). A part or product is created via casting, which is the act of pouring molten metal into a mould. Inconsistencies in the product brought on by the casting process frequently detract from its quality. Early detection of defective products during the production process is possible with timely screening of these faults, which can help save time and money. It is ideal to automate quality control wherever it is feasible to assure accurate and economical inspection, especially if fault detection must be carried out repeatedly throughout the manufacturing process. The primary drivers promoting the deployment of automated inspection systems are higher inspection rates, better quality requirements, and the need for more quantitative product evaluation unrestricted by human fatigue (Ferguson et al., 2017). Non-Destructive Examination (NDE) is a procedure used on completed components. NDE is a technique for examining modifications to a material, component, structure, or system, without causing damage to the original part, to look for defects, discontinuities, or atypical changes. Depending on the various types of faults, the right NDE techniques for castings must be used. There may be surface, subsurface, or inner problems, but it's crucial to carefully inspect all three areas for discontinuities. The proper NDE techniques for various fault types are shown in Table 1.

Table 1: Types of defects and their inspection method

Defects	Inspection Method
Surface	Visual testing with Liquid Penetration
Sub-Surface	Magnetic Particle testing
Internal	Ultrasound or Radiography

Need for visual inspection

The manufacturing industry is under pressure to create and apply automated inspection techniques for the detection of casting faults. Depending on the part's design and manufacturing technique, such faults may be imperfections like fractures or tearing, inclusions from chemical reactions or foreign objects in the molten metal, or porosity brought on by gases or shrinkage. Each of these faults carries the risk of producing a fracture site and serving as a stress raiser. Casting fracture could result from pressure on this location. Castings such as steering knuckles, control arms, transmission mounts, impellers, and cross members are essential structural components (Chen, Miao, and Ming, 2011). These castings must be trustworthy and of the highest standard. Because they are open 24 hours a day, 7 days a week, today's casting producers require optical and X-ray inspection systems that can function at production line rates. Dynamic analysis is possible with imaging technology, which lowers false rejects. The technician can operate several machines simultaneously and only needs a little training. The following items can be visually inspected:

- Metals
- Non-Ferrous Metals
- Glass
- PVC
- Ceramics
- Food and cosmetic items

Most industries use at least one sort of visual examination. Comparing visual inspection to other types of inspection, its advantages include being inexpensive, requiring little in the way of equipment, and requiring no prior knowledge from the technician.

Research problem

Any produced component that is manually examined must go through a labour-intensive, subjective, time-consuming, and potentially biased process. For assisting inspectors in checking welds and casts based on a set of criteria, an automatic inspection system is advised. In comparison to human inspectors, this technology would produce evaluation results that are more trustworthy, consistent, efficient, and objective (Dong, Taylor, and Cootes, 2021).

Methodology

Crisp DM

Under the umbrella of data science, the topic of maximising the value of data has become substantially more volumetric and sophisticated. However, it has also significantly increased in exploratory-ness. In contrast to the usual data mining process, which starts with specified business goals that translate into a clear data mining task, which finally turns data to knowledge, data-driven and knowledge-driven stages interact. In other words, the methods for maximising the value of data have evolved with data itself. CRISP-DM is the name of a data mining process model that seems to be independent of industry. The iterative process from business knowledge to implementation consists of six stages.

- Business Understanding

The business comprehension phase's main goal is to comprehend a project's objectives and specifications. The remaining three tasks in this phase, except for job number three, are essential project management procedures that are applicable to the majority of projects.

- Understanding the Data

The next stage is data understanding. It focuses attention on locating, obtaining, and evaluating data sets that might help you meet project objectives in addition to building the foundation of business understanding.

- Preparing the Data

Data preparation is one of the most important and time-consuming data mining processes. According to estimates, data preparation takes up a significant portion of a project's time and work. This cost can be decreased by devoting adequate time and effort to the first stages of business and data understanding, but the data preparation process would still require a lot of work.

- Modelling the Data

Modelling is typically done across several iterations. Before fine-tuning a model or going back to the data preparation step for any necessary adjustments by their preferred model, data miners often run several models with the default settings. Rarely can an enterprise successfully address a data mining challenge with a single model and a single execution. Data mining is so fascinating because it allows you to examine a problem from a variety of perspectives.

- Evaluation

Make sure the models developed during the modelling phase satisfy the technical and practical specifications of the previously defined data mining success criteria.

But one should evaluate the results of the work using the project's original set of commercial success criteria. This is crucial to ensuring that the business can use the results produced.

- Deployment

Deployment is the process of using new insights to enhance the business. Utilising data mining findings to implement change within your organisation is another definition of deployment. These results will undoubtedly be useful for planning and making marketing-related decisions, even though they might not be officially included 12 into your information systems.

Understanding the Dataset

During the casting process, a liquid material is poured into a mould with a hollow chamber in the desired shape, and the material is then allowed to solidify. The purpose of this data collection is to identify casting problems. Casting faults are unwanted flaws that occur during the metal casting process. There are many different types of casting flaws, including metallurgical flaws, blow holes, pinholes, burrs, shrinkage flaws, flaws in the material used to make the mould, and errors in the metal-pouring procedure. Defects are not desirable in the casting industry. Every industry has a quality control division whose job it is to weed out these inferior goods. The main problem, however, is that this examination process is manual. This process takes a long time and is not totally precise due to human error. The entire order could be rejected if there was a mistake. The corporation suffers a large loss as a result.

The dataset includes items made by the casting industry. The impeller for the submersible pump can be seen in every image in the dataset. There are 7348 total photos in the collection. Each of these photographs has a resolution of (300*300) pixels and is in grayscale. All of the images have previously undergone enhancement. Additionally, there were 512x512 grayscale images. This data collection does not involve any augmentation. Included in this are 781 photographs that have been classified as defective and 519 that have been classified as non-defective. Figure 1 illustrates the two types of categories—defective and excellent products. Two files were created from the data that was used to train and test the classification model. Both the train and test folders contain the subfolders def front and ok front.

1. Train: def front contains 3758 and ok front contains 2875 images

2. Test: def front contains 453 and ok front contains 262 images

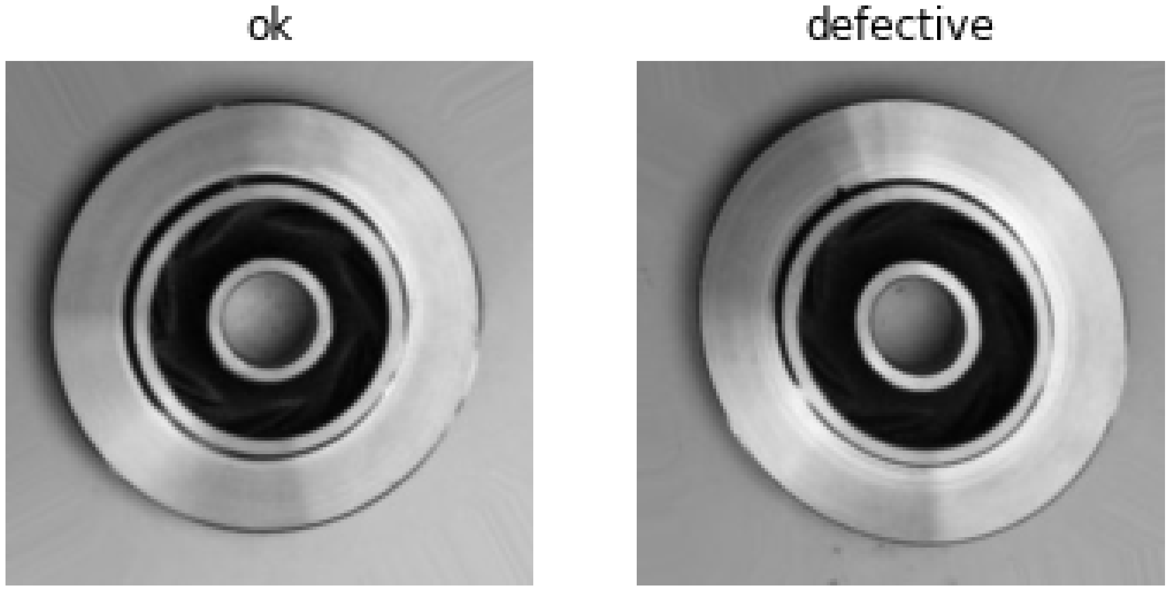


Figure 1: Sample images from the dataset. Image on the left shows a non-defective product. The image on the right shows a defective product.

Splitting the data

How to split the data for training and testing is a basic issue that machine learning and deep learning (ML/DL) practitioners face. Even if it appears straightforward at first, its complexity cannot be determined unless it is carefully examined. Inaccurate training and testing sets may have unforeseen consequences on the model's output. It could result in the data being either over- or under-fitted, which would lead to erroneous conclusions from the model. The data should be divided into three sets: a train set, a test set, and a holdout cross-validation or development (dev) set. These sets should have the following information and serve the following purposes:

1. Train Set: The data that would be input into the model would be included in the train set. Simply said, this information would inform our model. A regression model, for example, might discover gradients to lower the cost function by making use of the examples in this data. These gradients will then be applied to save costs and improve data prediction. Figure 2 shows the division between "ok" and "not ok" photos.

2. Validation Set: The trained model is validated using the development set. Since it serves as the foundation for our model evaluation, this setting is the most crucial. If there is a significant discrepancy between the error on the training set and the error on the development set, the model is overfitting and has a high variance.

3. Test Set: The data that were used to assess the trained and approved model are in the test set. It demonstrates the effectiveness of our entire model and the likelihood that it will forecast an irrational event. Several assessment criteria (such as

precision, recall, accuracy, etc.) may be utilised to gauge the effectiveness of our strategy.

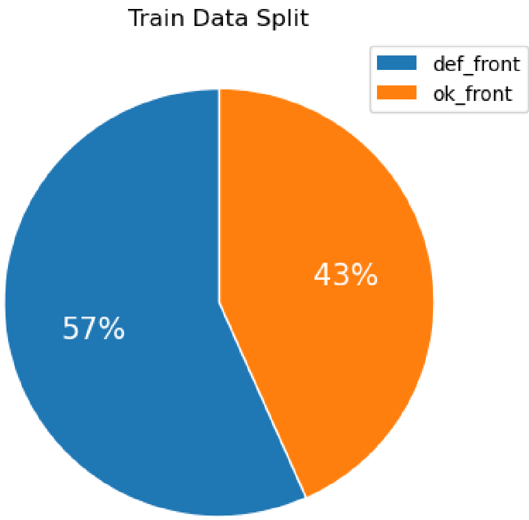


Figure 2: Split percentage of the dataset.

Image Augmentation

Image augmentation is a technique that modifies the original image in numerous ways to produce multiple altered versions of the same image. Depending on the augmentation techniques utilised, such as shifting, rotating, flipping, etc., each duplicate image differs from the others in a few significant ways. The target class is unaffected by these edits to the original image; they only provide a different perspective for capturing the object in real life. These methods of picture augmentation not only expand the size of one's dataset but also provide some variation, which enhances the model's ability to generalise to unseen data. Additionally, the model becomes more trustworthy when it is trained on recent, minimally altered photographs. However, it leaves it out of the original corpus of images; It merely transmits the modified images. The model would then repeatedly see the actual images, overfitting our model. Another advantage is that ImageDataGenerator requires less RAM. This is because when we don't use this class, all the photographs load at once. However, we load the images in chunks while using it, which consumes significantly less RAM. In general, convolutional neural networks (CNN) is the most popular algorithm for any image processing algorithms since it serves as the foundation for more complex algorithms. This method is exclusively utilised for CNN because it works well for deep learning algorithms. The split count across the Training, Validation, and Testing sets is displayed in Figure 3.

Several crucial variables for this strategy include the ones listed below:

- Directory: This is the location of the aren't folder, which includes a subdirectory containing all the pictures for the various classes.
- Target_size: The input image's size.
- Color_mode: If the photographs are black and white, set colour mode to grayscale; otherwise, set it to RGB for colourful images.
- Batch_size: The size of the data batches.
- Class_mode: Set to categorical for 2-D one-hot encoded labels and binary for 1-D binary labels.
- Seed: set to replicate the outcome is a seed.

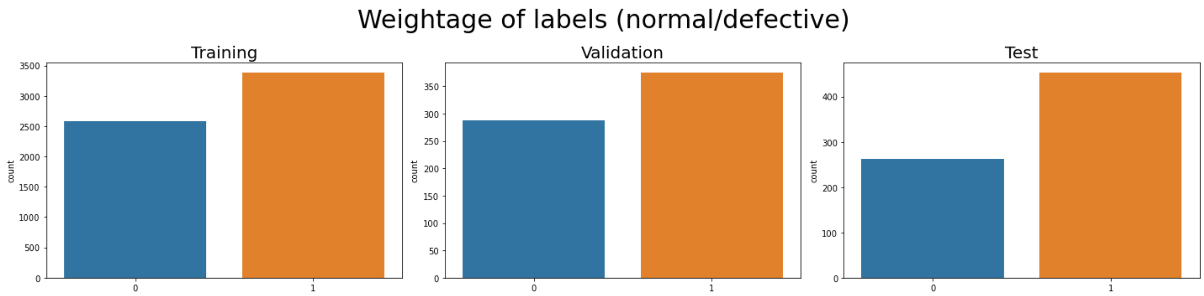


Figure 3: Split percentage of the dataset.

Model Building

Convolutional Neural Networks (CNN) and Support Vector Classifiers (SVC) were chosen for this problem statement because they serve as the foundation for sophisticated image processing techniques. CNN carries out feature extraction automatically, and SVM functions as a binary classifier.

Results

Analysing results of the CNN model

One of the main strategies used to fight this overfitting pestilence known as regularisation is early stopping. The model tries to follow the loss function on the training data by changing the parameters. As the validation set, a distinct set of data is maintained. The validation set's loss function is recorded as training proceeds. Instead of running through all the epochs, the model quits when the validation set has not improved (also known as early stopping). The graphs in Figure 4 demonstrate that the model is not overfitting the data, and this is supported by the fact that both the train loss and the validation loss simultaneously went to zero. Additionally, accuracy on both the train and val sides moves closer to 100%. The model was designed to run for 100 epochs; however it was stopped after 28 because the validation did not improve.

The callback function was implemented as it allowed for the following:

- Save the model at regular intervals
- Callback can be used on fit, predict and evaluation of the model

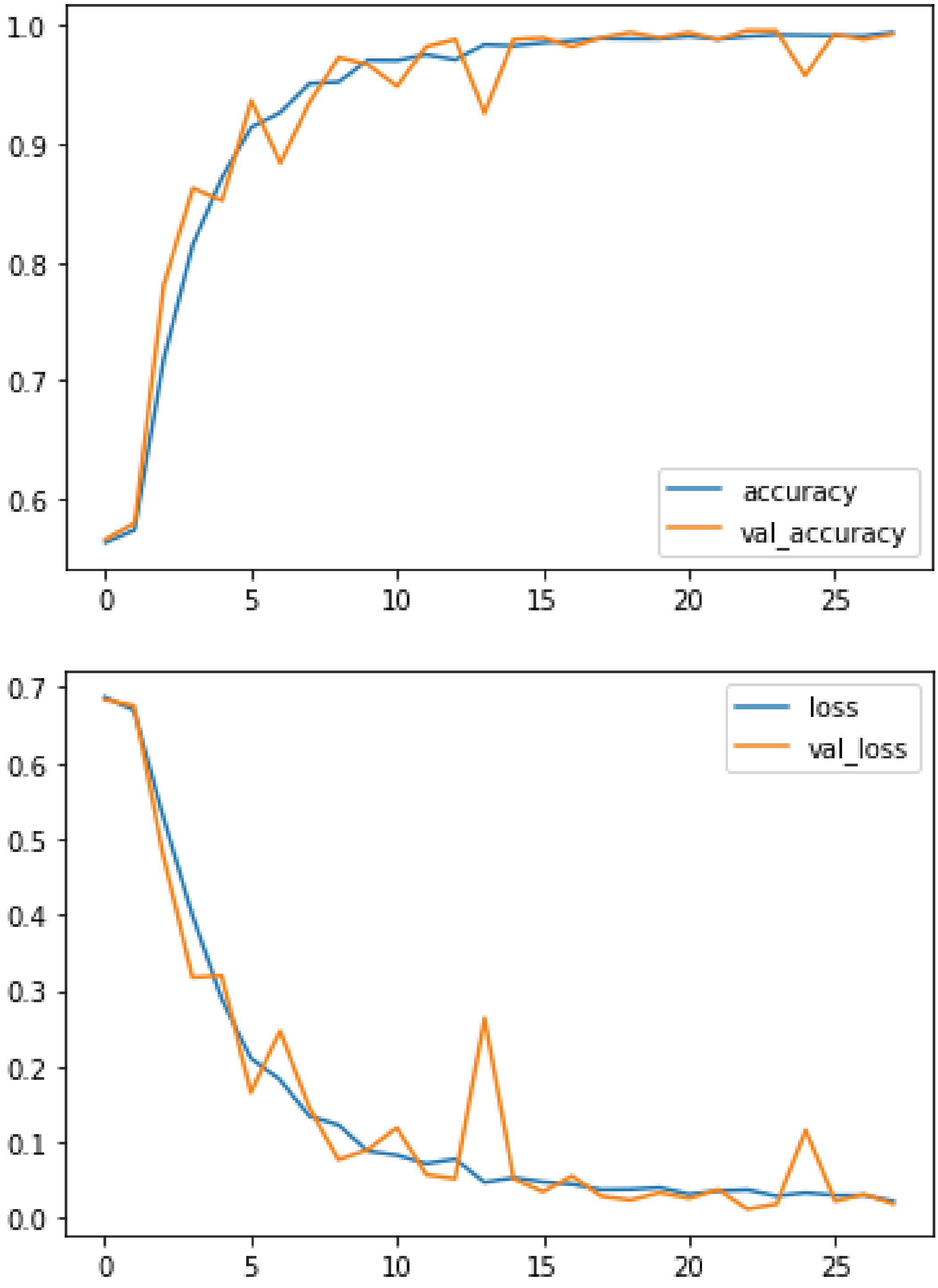


Figure 4: CNN Model performance

Analysing results of SVC model

In a graph, a classification model's receiver operating characteristic (ROC) illustrates how well it performs at each level of categorization. The number of False Positives and True Positives increases when the categorization criterion is lowered because more objects are categorised as positive. AUC measures the entire two-dimensional region that is covered by the entire ROC curve. AUC offers a comprehensive evaluation of performance across all available categorization criteria. The probability that the model values a randomly selected positive example higher than a randomly selected negative example is known as the AUC. The huge AUC in Figure 5 of the SVC ROC curve shows the model's strong performance.

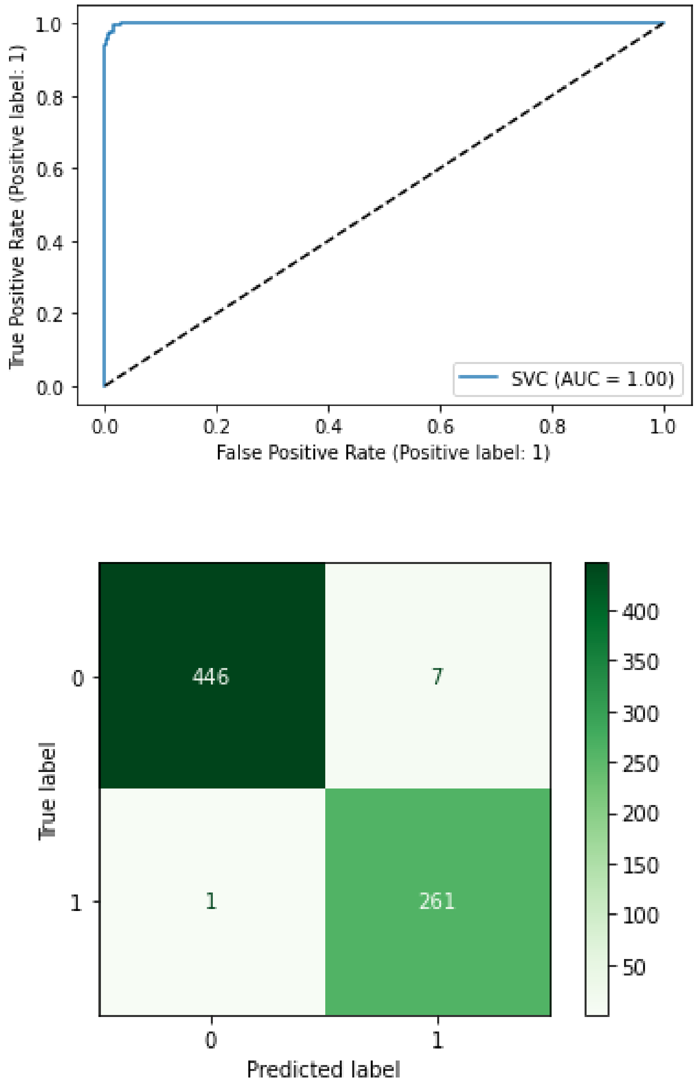


Figure 5: SVC Model performance

Model Comparison

The classification performance of the SVM and CNN models was compared. Figure 6's findings show that the CNN has the highest level of classification accuracy for the dataset. The SVM model ignores the dataset's rich spatial information and only uses the spectral signature of each image pixel. Using additional spatial information acquired from surrounding pixels, the methodologies for classifying defects' classification accuracy are improved. Since spatial information can improve classification accuracy, a CNN uses rich spatial characteristics at multiple scales to reflect the spatial structure of the data and classify each pixel in the image. A high precision rate indicates that the classifiers are performing as they ought to. A significant portion of all pertinent incidents have been recovered, as indicated by a high fractional value of recall. It's interesting to note that CNN's accuracy is higher at 99.86% compared to SVC's lesser accuracy of 98.88%. It is necessary to consider the precision, recall, and accuracy values when determining the discriminative capability of risk models. In this situation, a CNN is preferable to an SVC.

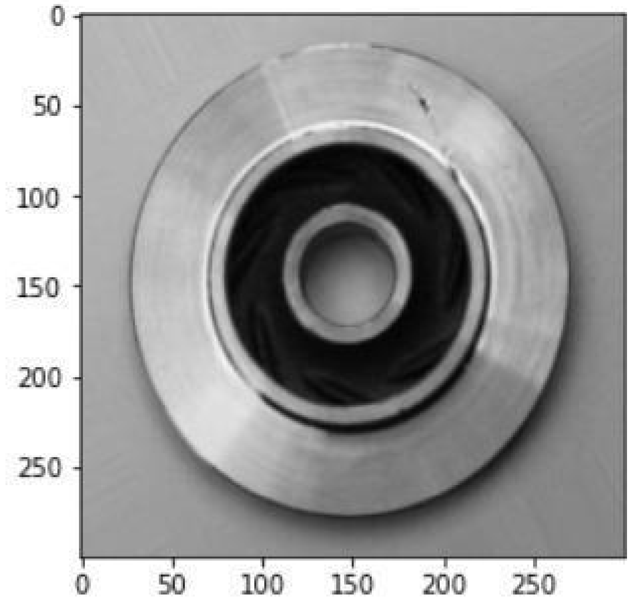
	precision		recall		f1-score		support	
	CNN	SVC	CNN	SVC	CNN	SVC	CNN	SVC
Metrics								
0	0.996198	0.997763	1.000000	0.984547	0.998095	0.991111	262.000000	453.000000
1	1.000000	0.973881	0.997792	0.996183	0.998895	0.984906	453.000000	262.000000
accuracy	0.998601	0.988811	0.998601	0.988811	0.998601	0.988811	0.998601	0.988811
macro avg	0.998099	0.985822	0.998896	0.990365	0.998495	0.988008	715.000000	715.000000
weighted avg	0.998607	0.989012	0.998601	0.988811	0.998602	0.988837	715.000000	715.000000

Figure 6: Model comparison

Deployment

Deployment, the action of incorporating a machine learning model into an already-existing production environment, enables you to use data to help you make informed business decisions. This is the final stage of the machine learning life cycle.

```
1/1 [==============================] - 0s 62ms/step
[[1.]]
WARNING: Component is Defective
```




```
1/1 [==============================] - 0s 94ms/step
[[0.]]
Component is not defective!
```

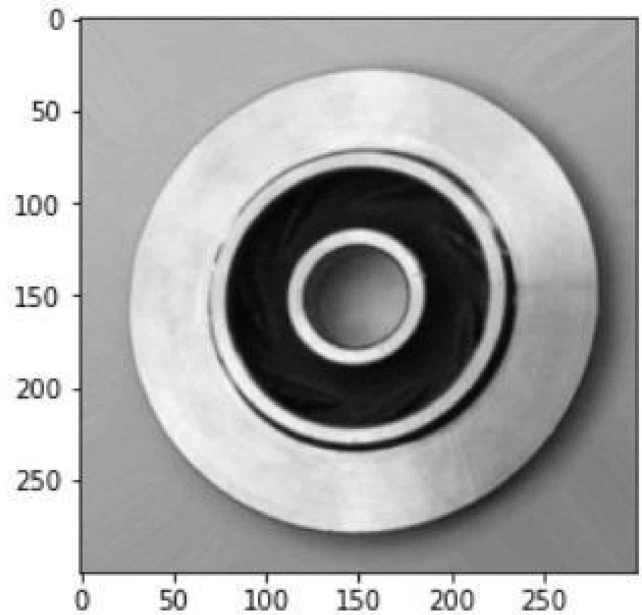


Figure 7: Custom predictions

The "cnn model.hdf5" file is used to store the best-performing model. You can call this whenever you want and use it to make predictions. The model has been saved now. The saved model can be called to generate personalised forecasts. A for loop and a specific file path for the user to upload are included in the code. The model will be able to provide a binary output (Defect or not) once the path has been provided. This is demonstrated in Figure 7.

Conclusion and Future work

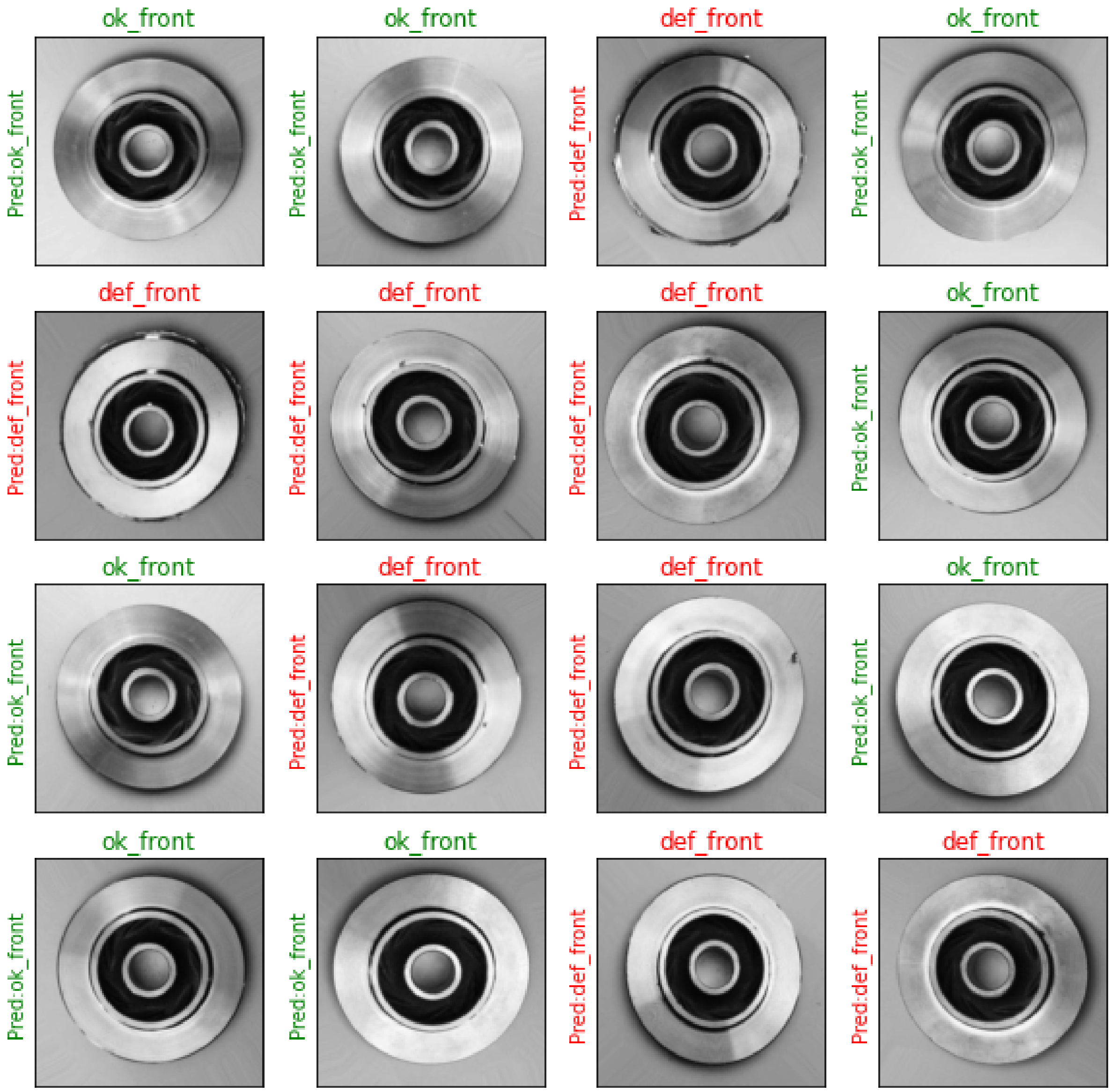
Conclusion

CNN are a serious contender for image recognition. CNN can evaluate sequence data, but they are best at sorting through large amounts of image data to find non-linear associations. For these classifications, SVC is a margin classifier that supports a number of kernels. When class labels are quite lengthy, SVC has problems predicting the classes. Although parallelization is a challenge for SVC as well, it is a feature of the CNN design. Accuracy ratings do not reflect reality when compared to visual perception of the overall classification of each image. However, data must be ready for kernel window processing because to the way CNN functions. Based on past research, the Kernel window for this inquiry was 7*7 pixels. It was important to use layers of the CNN in order to reduce the dimensions of the data that are decreasing noisy bands. The approach required to be improved in order to increase the dataset training samples. SVC classifiers have fared better in terms of accuracy in image classification despite the accuracy of 99.86% claimed by CNN. Figure 8 below shows predictions from both models CNN and SVC. Even when only a few photographs were used for training, our method still produced good results.

This implies that the suggested technique is appropriate for identifying anomalies in both other photos.

Key takeaways:

- Early stopping technique is utilised to avoid unnecessary under-fitting and over-fitting due to usage of too few or too many epochs in our model.
- Data Augmentation the Keras ImageDataGenerator class' primary advantage is that it is intended to give real-time data augmentation. Meaning that while your model is still being trained, it is instantly producing enhanced pictures.
- We built and trained our model without any pre-trained weights. We used the callback function built in keras to save our model.


ok_front
Pred:ok_front
ok_front
Pred:ok_front
def_front
Pred:def_front
ok_front
Pred:ok_front
def_front
Pred:def_front
def_front
Pred:def_front
def_front
Pred:def_front
ok_front
Pred:ok_front
ok_front
Pred:ok_front
def_front
Pred:def_front
def_front
Pred:def_front
ok_front
Pred:ok_front
ok_front
Pred:ok_front
ok_front
Pred:ok_front
def_front
Pred:def_front
def_front
Pred:def_front

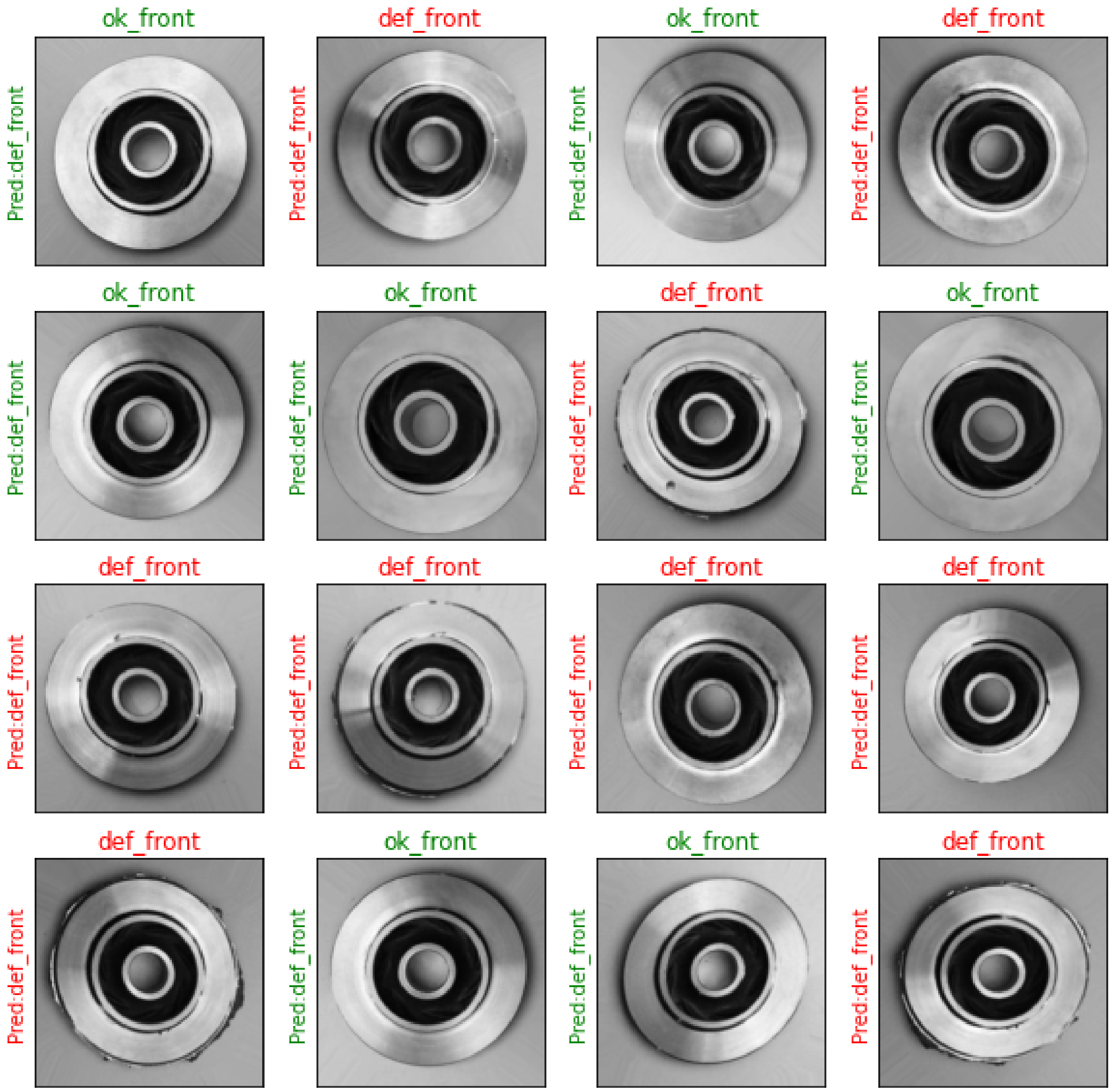


Figure 8: Predictions SVC (left) and CNN (right)

Future work

Only the defect classification utilising two methods is covered in this study. In the future, the use of Region Based CNNs (R-CNNs), which would use an image localization technique to find flaws in the cast components could lead to new insights. In essence, this method would outline a bounding box around the area of interest. A box would be drawn around the defective area in our scenario. Weld/cast pictures databases, on the other hand, are few. As demonstrated in our study, we may train our model on more recent photos to enhance its functionality. This has the potential to be employed in industries like aerospace and automotive where a high level of safety is required. The stored model can easily be set up on Streamlit or Heroku. This would essentially allow the user to upload their own images and the model should be able to predict the class.

References

Chen, W., Miao, Z. and Ming, D. (2011) 'Automated Inspection Using X-Ray Imaging'. *2011IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications*. Changsa, China, 16-18 November. pp. 1769–1772. Available at: http://doi.org/10.1109/TrustCom.2011.247

Dong, X., Taylor, C.J. and Cootes, T.F. (2021) 'A Random Forest-Based Automatic Inspection System for Aerospace Welds in X-Ray Images', *IEEE Transactions on Automation Science and Engineering*, 18(4), pp. 2128–2141. Available at: https://dx.doi.org/10.1109/TASE.2020.3039115

Ferguson, M., Ak, R., Lee, Y.T. and Law, K.H (2017) 'Automatic localization of casting defects with convolutional neural networks'. *2017 IEEE International Conference on Big Data (Big Data)*. Boston, Massachusetts, USA, 11-14 December., pp. 1726–1735. Available at: https://dx.doi.org/10.1109/BigData.2017.8258115

.