



e-ISSN: 2319-8753 | p-ISSN: 2320-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Issue 5, May 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.512

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com



AI-Based Visual Inspection System for Defect Detection in Bulk Materials

Rishikesh Vishnu ¹, Rajaganesh ²

U.G. Student, Department of Mechanical Engineering, St. Josephs Institute of Technology, Semmancheri, Chennai, Tamil nadu, India ¹

U.G. Student, Department of Mechanical Engineering, St. Josephs Institute of Technology, Semmancheri, Chennai, Tamil nadu, India ²

ABSTRACT: Detecting defects in bulk materials is one of the challenging problems for industries worldwide and is currently very manual and time consuming. Manually inspecting the bulk production of products for defect detection is a tiresome task and could be prone to human errors. This study is aimed at making AI-based visual inspection system for defect detection in bulk materials, particularly steel. When steel sheets are manufactured it comes in contact with many heavy machines for the purpose of drying, cutting, rolling, heating due to these some defects occur on the surface. Through this study, defects in steel were detected by applying transfer learning to 12600 steel images using a pretrained Xception, Unet and Mask-RCNN convolutional neural network models. After making suitable changes to the models, we compare the efficiencies obtained by each of the model. The metric that was used to evaluate the performance of the models was dice-coefficient. Among the models unet achieved a dice-coefficient of 0.492.

However, while applying these models on other material's dataset, we obtained a much lesser accuracy.

KEYWORDS: Defect detection, Computer Vision, Deep learning, CNN

I. INTRODUCTION

We propose to investigate the problem of predicting defects in bulk materials, particularly steel. Since we come from a mechanical engineering background, we're familiar with the importance of quality assurance for bulk materials. Steel is one of the most important building materials of modern times. And the production process of flat sheet steel is especially delicate. From heating and rolling to drying and cutting, several machines touch flat steel by the time it's ready to ship, shipping steel can be extremely expensive, and if the steel performs differently than expected, the purchaser takes on large risks. Steel that underperforms can damage existing parts in the project and/or equipment used to work it. Manufacturers must also meet federal regulations when supplying steel. Steel suppliers are therefore under high pressure to ensure their quality assurance is both robust and inexpensive. Hence, automating the inspection process would accelerate the steel sheet production.

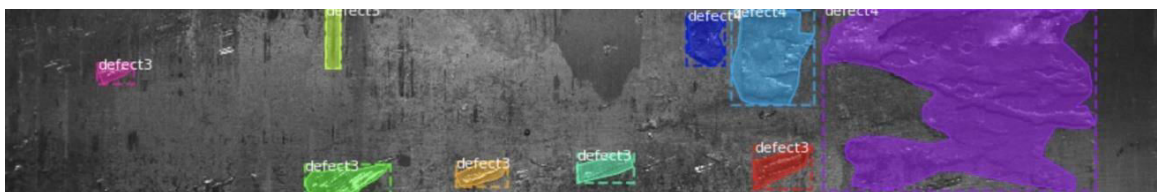
II. RELATEDWORK

For the proposed study, the input is an image of a steel sheet and the output is a same-size segmented image with dense defect area marks. Since the output is a pixel-wise label for the input image, it's essentially a segmentation task. Within image segmentation, there are two different sub-task, semantic segmentation and instance segmentation. Semantic segmentation could be considered as a per pixel classification problem. The widely used method for semantic segmentation using CNN is the Fully Convolutional Network (FCN), which converts fully connected layers into 1x1 convolutional layers and achieves end-to end per-pixel prediction. Compared with semantic segmentation, instance segmentation aims to predict not only class label, but also pixel-wise instance mask to localize varying numbers of instances presented in images.

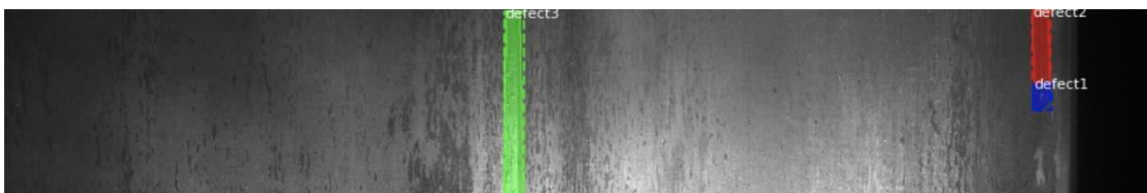
III.METHODOLOGY

Dataset and features

The dataset we are using is obtained from a Kaggle competition called Severstal steel defect detection. The dataset contains images of steel sheet with and without defects, defect class labels and defect area marks. The images provided are of size $1600 \times 256 \times 1$ and totals 12600 counts. Out of all the images, 5934 images are with defects and 6666 images are without defects. There are four labels of defects 1, 2, 3, and 4. In the images with defects, 899 images are of class 1 defect, 277 images are of class 2 defect, 5150 images are of class 3 defects, and 801 images are of class 4 defect. Additionally, 6239 images have only 1 class of defect, 425 images have two classes of defects, and only 2 images have 3 classes of defects.



(a) Sample image 1 with defect instances



(b) Sample image 2 with defect instances

Figure 1: Sample image with instance-level defects

Xception

Xception is a convolutional neural network that is 71 layers deep and acts as our baseline model. We removed the last FCN of the original Xception model and added an atrous convolution layer, which enlarges the feature map by linear interpolation.

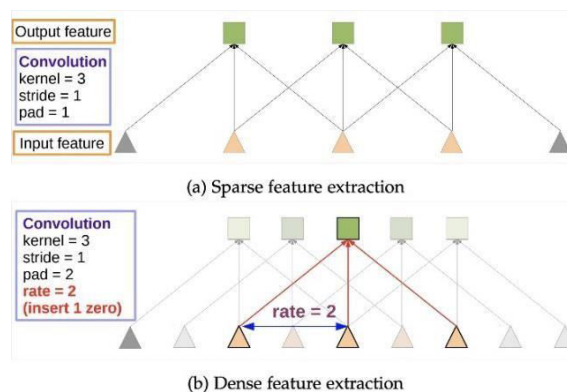


Figure 2: Illustration of atrous convolution in 1-D



Mask R-CNN

Mask R-CNN’s main idea revolves around instance segmentation, which identifies object outlines at the pixel level. The main idea of mask R-CNN is obtained from Fast R-CNN. It consists of two stages: the first stage scans the image and generates proposals (Region Proposal Network (RPN)) and the second stage classifies the proposals and generates bounding boxes and masks. The final output of Mask R-CNN is stack of masks of the size of input image with bounding box coordinate and defect classes.

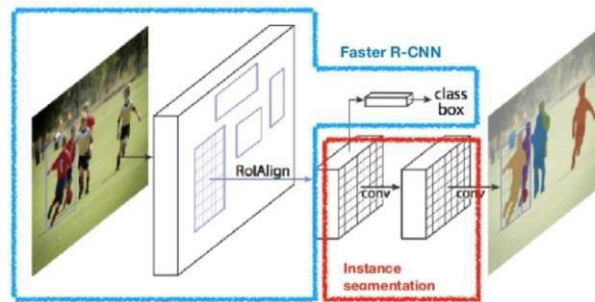


Figure 3: Mask R-CNN Structure

Unet

U-Net is a Fully Convolutional Network model. This model is more accurate in terms of pixel-based image segmentation. U-Net is used convert image into vector by using feature mapping and uses same mapping to convert it into image which maintains the structure integrity of an image. Its architecture consists of three parts: the contraction, the bottleneck and expansion section. The contraction section is used to learn more about the features of the image as it consists of 3x3 convolution layers followed by 2x2 max pool layer. In expansion layer important features of the image are learned while reconstructing the image. This layer consists of 3x3 CNN layer followed by 2x2 un-sampling layer. It uses loss weighing scheme where higher weights are at the border of the segmented objects which causes to classify each pixel as one class. The U-net architecture is shown in Fig.4

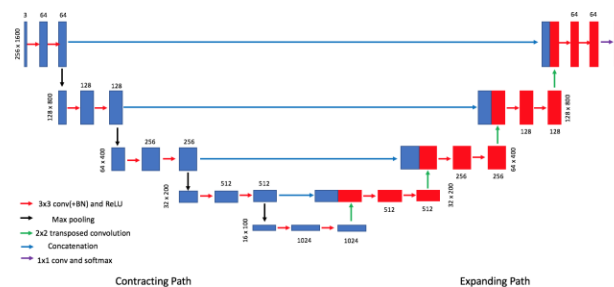


Figure 4: U-net Architecture



IV. EXPERIMENTAL RESULTS

Dice Coefficient

The Dice coefficient can be used to compare the pixel wise agreement between a predicted segmentation and its corresponding ground truth. The formula is given by:

$$Dice(X,Y)=\frac{2 \cdot |X \cap Y|}{|X| + |Y|}$$

Where X is the predicted set of pixels and Y is the ground truth. The Dice coefficient is defined to be 1 when both X and Y are empty

Xception

We train our Xception-based model for 6 epochs. Learning rate during the training process is set to 1e – 3. Adam optimizer is used. The binary cross-entropy loss, dice score and IOU score during the training are plotted in Fig

5. In the inference phase, images are first resized to 256 * 256 to feed as the input of our model. The output of the model are then resized back to 256 * 1600 as the final predictions. Dice score is used to evaluate our model. Our model achieves 0.369 on positive dice score and 0.978 on negative dice score. We also show a qualitative result of our model in Fig 5

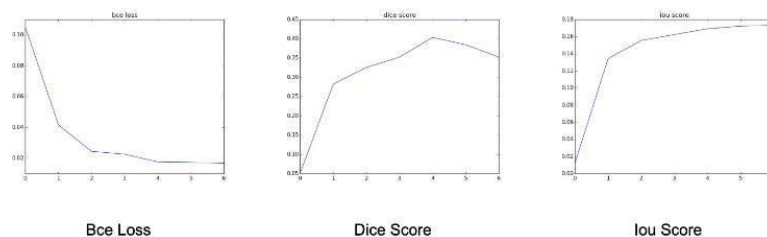


Figure 5: Binary cross-entropy loss, dice score and IoU score during the training process of Xception model

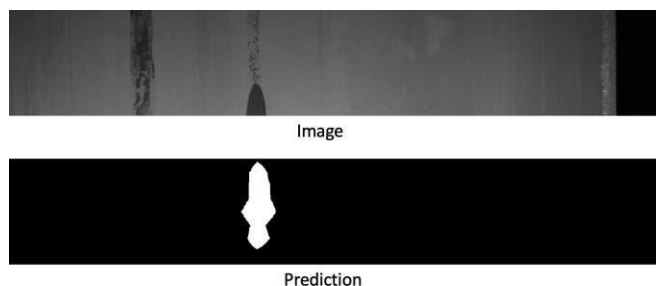


Figure 6: Qualitative result of Xception model.



Mask-RCNN

The model is not able to provide accurate shape and location of the defect as seen in the labelled image, which is shown in Fig. 7b and 7c. The training loss drops to around 1.1 and remains steady throughout the training process.

The average positive Dice score is 0.083 and the negative Dice score is 0.898. Throughout the training process, the rpn_box loss keeps at a high value and accounts for the majority of the loss. The result shows that the Mask RCNN network is not an ideal model for the defect detection task. The training loss and validation loss is shown in the Fig. 7a.

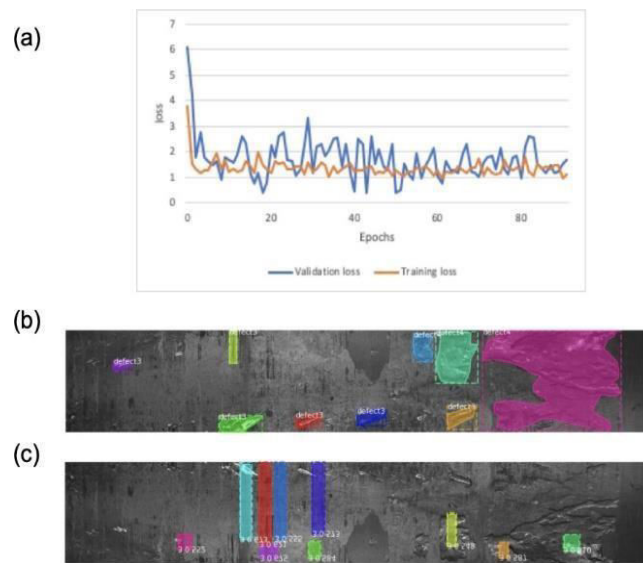


Figure 7: Results of the Mask R-CNN.

(a) Loss of the training and the validation sets. (b) A sample labelled image containing instances of defect. (c) The defect instances prediction result generated by Mask R-CNN

U-Net

We used Dice loss function during training, which is defined as

$$L_{dice} = 1 - \frac{1}{B} \sum_{i=1}^B \frac{2 \sum_{j=1}^n y_j^{(i)} \hat{y}_j^{(i)}}{\sum_{j=1}^n y_j^{(i)} + \sum_{j=1}^n \hat{y}_j^{(i)}}$$

where B is the batch size, $n = 4 \times 256 \times 1600$ is the total number of pixels of the 4 defect classes, $y_j^{(i)}$ is the true label of each pixel (1 for defect and 0 for normal), and $\hat{y}_j^{(i)}$ is the predicted probability. A batch size of 4 was used because of computational limitations. We used Adam as our optimizer like in Xception model. The loss values and the positive Dice score during training are shown in Fig. 8a and 8b, respectively. After 20 epochs, the positive and negative Dice scores are 0.492 and 0.977. The U-net model showed good performance for predicting defects of class 3 (Fig. 8c).

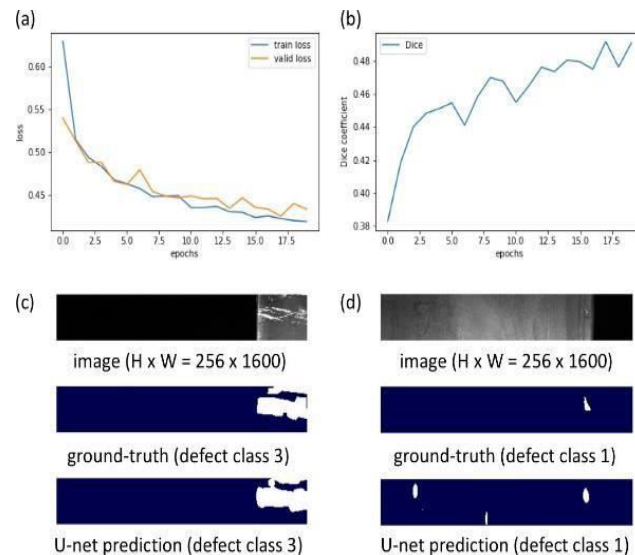


Figure 8: Results of the unet

(a) Loss of the training and validation sets (b) Average positive dice score of the validation set during training
 (c) A sample image containing defect of class 1

IV. FUTUREWORK

Due to the imbalance between different kinds of labels, a large dataset with equal number of examples for each label would have improved the performance of our models. Data augmentation would have also helped. Other than that, we could train an embedding specifically for steel images for feature extraction such that we could train a logistic regression quickly whenever a new defect a feature is of interest. Lastly, effort could be put to predict the severity of each defect kind so that the steel factory could also leverage the severity information to determine whether to dispose the steel sheet.

V. CONCLUSION

This project leverages the segmentation and instance segmentation computer vision techniques to solve the steel defect detection problem. The models outputs the pixel wise defected area with corresponding labels. Among all the models, the U-Net performs the best with a positive dice coefficient of 0.492.

REFERENCES

- [1] Severstal: Steel defectdetection.
- [2] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille. Deeplab: Semantic image segmentationwith deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, 40(4):834–848,2017.
- [3] F.Chollet.Xception:Deeplearningwithdepthwiseparableconvolutions.InProceedingsoftheIEEEconference on computer vision and pattern recognition, pages 1251–1258,2017.
- [4] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele. The cityscapes dataset for semantic urban scene understanding. In Proceedings of the IEEE conference on computervision and pattern recognition, pages 3213–3223,2016.
- [5] R. Girshick. Fast r-cnn. In Proceedings of the IEEE international conference on computer vision, pages 1440–1448, 2015.
- [6] K. He, G. Gkioxari, P. Dollar, and R. Girshick. Mask r-cnn. In Proceedings of the IEEE international conferenceon computer vision, pages 2961–2969,2017.



- [7] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan and S. Belongie. Feature pyramid networks for object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 2117–2125,2017.
- [8] H. Noh, S. Hong, and B. Han. Learning deconvolution network for semantic segmentation. In Proceedings of the IEEE international conference on computer vision, pages 1520– 1528,2015.
- [9] B. Romera-Paredes and P. H. S. Torr. Recurrent instance segmentation. In European conference on computer vision, pages 312–329. Springer,2016.
- [10] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, pages 234–241. Springer, 2015.
- [11] L. Shen, Z. Lin, and Q. Huang. Relay backpropagation for effective learning of deep convolutional neural networks. In European conference on computer vision, pages 467–482. Springer,2016.



INNO  **SPACE**
SJIF Scientific Journal Impact Factor

**Impact Factor:
7.512**

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details