

Semi-Siamese Network for Robust Change Detection Across Different Domains with Applications to 3D Printing*

Yushuo Niu^{1[0000-0001-8077-1005]}, Ethan Chadwick^{2[0000-0002-1927-3856]},
Anson W. K. Ma^{2[0000-0002-2865-5776]}, and Qian Yang^{1[0000-0001-5519-1092]}

¹ Computer Science & Engineering Department, University of Connecticut
371 Fairfield Way, Unit 4155, Storrs, CT 06269-4155
{yushuo.niu,qyang}@uconn.edu

² Chemical & Biomolecular Engineering Department, University of Connecticut
97 North Eagleville Road, Unit 3136, Storrs, CT 06269-3136
{ethan.chadwick,anson.ma}@uconn.edu

Abstract. Automatic defect detection for 3D printing processes, which shares many characteristics with change detection problems, is a vital step for quality control of 3D printed products. However, there are some critical challenges in the current state of practice. First, existing methods for computer vision-based process monitoring typically work well only under specific camera viewpoints and lighting situations, requiring expensive pre-processing, alignment, and camera setups. Second, many defect detection techniques are specific to pre-defined defect patterns and/or print schematics. In this work, we approach the defect detection problem using a novel Semi-Siamese deep learning model that directly compares a reference schematic of the desired print and a camera image of the achieved print. The model then solves an image segmentation problem, precisely identifying the locations of defects of different types with respect to the reference schematic. Our model is designed to enable comparison of heterogeneous images from different domains while being robust against perturbations in the imaging setup such as different camera angles and illumination. Crucially, we show that our simple architecture, which is easy to pre-train for enhanced performance on new datasets, outperforms more complex state-of-the-art approaches based on generative adversarial networks and transformers. Using our model, defect localization predictions can be made in less than half a second per layer using a standard MacBook Pro while achieving an F1-score of more than 0.9, demonstrating the efficacy of using our method for *in situ* defect detection in 3D printing.

* This work was supported by the National Science Foundation under grant number IIP-1822157 (Phase I IUCRC at University of Connecticut: Center for Science of Heterogeneous Additive Printing of 3D Materials (SHAP3D)). A. W. K. Ma acknowledges support from UConn via the United Technologies Corporation Professorship in Engineering Innovation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or the sponsors.

Keywords: change detection · defect localization · semi-siamese neural network · domain adaptation · 3D printing

1 Introduction

Defect detection methods that can provide feedback in real-time is of significant interest to the additive manufacturing community in order to save on materials cost, printing time, and most importantly, to ensure the quality of printed parts. A key advantage of 3D printing technology that can be leveraged to enable *in situ* defect detection is that 3D objects are printed layer by layer (Figure 1). Thus, each 2D layer of the object can be imaged and probed for internal defects; unlike traditional manufacturing processes, it is not necessary to wait to analyze the fully printed 3D object, and the interior of the object can be probed as the object is being constructed.

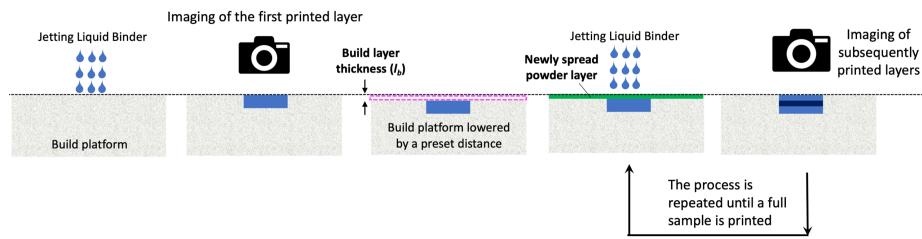


Fig. 1: Schematic diagram of the 3D printing process for binder jet 3D printing with layer-wise imaging during the print.

In this work, we propose a novel defect detection method for 3D printing that poses the problem as one of change detection between a desired reference schematic and a camera image of the printed layer (Figure 2). In the change detection problem in computer vision, two images such as satellite images of land or surveillance images of streets, are compared for differences. There are several challenges common to both the defect detection and change detection problems: the need to pre-process and pre-align images due to changes in camera angle and lighting, which result in significant and sometimes expensive limitations to the camera setup that must be used, and the data-hungry nature of this complex comparison problem. Additionally, the images we would like to compare for 3D printing are from heterogeneous sources: one is a reference print schematic and the other is a noisy camera image of the actual printed result. In this work, we utilize one-shot learning techniques [11] to develop a novel deep learning architecture that can provide fast and precise localization of defects robust to

camera angle and lighting perturbations. A key characteristic of our model is that its relatively lightweight and simple architecture can be easily pre-trained to adapt to new datasets; in fact we show that pre-training enables our simple model to outperform more complex methods based on state-of-the-art techniques such as transformers. The simplicity and flexibility of our model will enable it to be highly transferable to different industrial settings for 3D printing, without requiring careful camera setups and application-specific model customization that is both expensive and time-consuming. Our proposed approach of building on change detection methods from computer vision for tackling the challenges of defect detection is to the best of our knowledge a new direction in the 3D printing field.

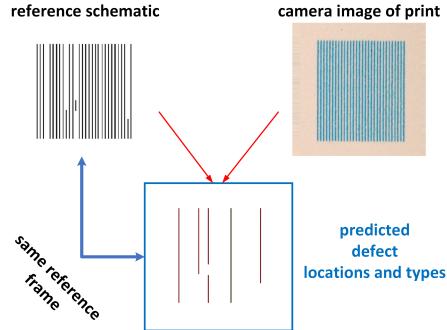


Fig. 2: Our robust defect localization model takes as input a reference print schematic and a camera image of the print, and predicts the precise location and type of defects with respect to the frame of reference of the print schematic. In the predicted image, white corresponds to no defects, red to over-extrusion and green to under-extrusion. This model can be used for *in situ* defect detection: it requires 0.419 seconds for prediction while printing a single layer on an inkjet-based 3D printer requires tens of seconds or less.

2 Related work

Change detection is a fundamental task in computer vision, with many important applications such as analysis of satellite imagery for agricultural monitoring [10], urban planning [17], and disaster assessment [21], among others. A large body of work has thus been built starting from at least the 1980s using methods such as change vector analysis [14]. To handle perturbations such as misalignment and varied lighting, techniques such as incorporating active camera relocation have been proposed [5]. Many state-of-the-art methods today are now based on deep learning, ranging from autoencoders to Siamese neural networks to recurrent neural networks, and various combinations thereof [20]. Recently, several

methods based on combining convolutional neural networks (CNN) with Siamese architectures have been proposed. One of the earlier such methods, ChangeNet, uses a combination of ResNet, fully connected, and deconvolution blocks in its Siamese branches [22]. It is designed to handle different lighting and seasonal conditions in the images, but like most existing methods assumes aligned or nearly aligned image pairs. Interestingly, the architecture is different from traditional Siamese architectures in that the deconvolution layers are not required to have the same weights. This is reminiscent of our proposed Semi-Siamese architecture which we will discuss in Section 3; however, we will propose the opposite - the deconvolution layers are the portion of our architecture that are required to share the same weights. Another interesting recent approach uses a Siamese pair of autoencoders [15], where the change map is generated based on the learned latent representations. However, this method also assumes coregistered images and can only learn approximate change locations in addition to a classification of whether changes have occurred. A recently proposed architecture that enables fast pixel-level change detection is FC-Siam-diff, a fully convolutional encoder-decoder network with skip connections [3]. In this model, there are two encoders that share the same architecture and weights, while there is only one decoder. However, this model again assumes coregistered images. Finally, the challenge of dealing with images that are not necessarily coregistered, with differences in lighting, camera viewpoint, and zoom, was addressed in Sakurada & Okatani [18] and with CosimNet [6]. The former uses features learned from CNNs trained for large-scale object recognition in combination with superpixel segmentation to address the problem of change detection in pairs of vehicular, omnidirectional images such as those from Google Street View. The latter, CosimNet, uses the DeeplabV2 model as a backbone and proposes various modifications to the loss function to provide robustness to perturbations [6]. Nevertheless, both of these methods still assume that the images being compared are from the same domain, e.g. they are both camera or satellite images, rather than from different domains such as a camera image versus a schematic. More recently, heterogeneous change detection has been addressed using generative adversarial networks [12] and transformers [2].

Despite its importance for additive manufacturing, defect detection has traditionally been a challenging task. First, there are many different types of defects that may be of interest, including defects caused by missing jets, inconsistent jets, angled jets, and cracks in powder bed material, just to name a few that are relevant to inkjet-based 3D printing; other technologies such as fused deposition modeling have their own set of defects. Many heuristic-based methods such as computing the entropy of depth maps have consequently been developed to address specific defect types [4]. In recent years, both classical machine learning methods such as support vector machines utilizing human-engineered features [8] and deep learning-based methods utilizing convolutional neural networks have begun to be developed to enable more powerful defect detection [9, 19]. However, many of these methods require large amounts of labeled experimental data, which is difficult to obtain. They also typically require fixed, high-resolution camera

setups, and cannot easily handle differences in camera angle and lighting. For example, one group of methods is based on denoising autoencoders [7], where the idea is that an autoencoder is trained to take as input a “noisy” (defective) image and output its non-defective counterpart. Then, differences between the input and output can be used to identify defects. An advantage of this approach is that it does not require a large amount of labeled experimental data; however, unlike change detection approaches which can handle general differences, this approach can only handle a pre-defined range of defects, since it must be trained to be able to remove them from the output.

3 Semi-Siamese defect detection model

Our proposed model consists of two major components: a novel Semi-Siamese architecture based on U-Net [16] branches, and a fully convolutional network (FCN) to reconstruct the final defect detection mask. The input to the model is a pair of 2D images corresponding to a particular layer during 3D printing: the reference schematic images of the desired print pattern, $I_{ref} \in R^{H \times W \times 3}$, and the camera images of the printed result, $I_{cam} \in R^{H \times W \times 3}$. The image pair (I_{ref}, I_{cam}) is first fed into a Semi-Siamese network to generate a pair of feature maps (F_{ref}, F_{cam}) of the same dimensions as the input. In contrast to standard Siamese networks and existing Semi-Siamese networks which use different decoders, a simple but key innovation of our architecture is that the feature extraction sections of each branch (encoder) do not share the same weights; only the reconstruction section (decoder) share the same weights. This is important for our defect detection problem because the camera image and reference schematic come from different domains. In order to compare them, we first use *different* feature extraction functions to transform them to the same latent feature space, after which we reconstruct them both back into a comparable reference frame using the *same* reconstruction function. Then, their difference is calculated to generate a change map. It is important to calculate the change map from the reconstructed images in a comparable reference frame rather than from the latent feature space in order to enable highly precise pixel-wise defect localization. The final FCN is used to fully transform the change map from this comparable reference frame back to the reference frame of the schematic image.

3.1 Transfer learning from U-Net models

As described above, the Semi-Siamese branches of our model are based on the U-Net architecture. This choice is made to leverage the ability of U-nets to produce high resolution outputs [16], enabling precise localization of defects upon comparison of the outputs (F_{ref}, F_{cam}) from each branch. In order to further improve the performance of our model, we first utilize transfer learning from a U-Net model with the same architecture as our Semi-Siamese branches. This U-Net model takes as input a perturbed camera image, and outputs a transformation of the image into the same reference frame as its corresponding reference print schematic. When trained on a fixed number of reference schematics, this U-Net can be used for detecting defects by directly comparing a camera image

transformed into the reference frame with its corresponding print schematic. However, it is important to note that this architecture cannot handle arbitrary print schematics. Suppose that we would like to detect defects in a print corresponding to a schematic (called “schematic-new”) that is similar to a print schematic that the model was previously trained on (called “schematic-old”), but that looks like a perturbed version of it. Then any camera images of a perfect print of “schematic-new” might be erroneously transformed by the model back into the reference frame of “schematic-old”. Now when compared with “schematic-new”, many defects will be detected, even though no defects occurred in the actual print. Thus, this U-Net architecture cannot be used on its own to handle defect detection for arbitrary desired print schematics. We will instead use this U-Net model pre-trained on a small set of reference schematics to initialize the weights of each branch of our Semi-Siamese model. This allows us to initialize the Semi-Siamese model in such a way that it can offset perturbations for some limited sets of camera images and reference schematics. We then continue training to fine-tune these weights to be able to handle arbitrary reference print schematics and perturbed camera images. As we show in the results, this ability to pre-train the U-Net to initialize our Semi-Siamese model is key to high performance on novel problems. We note that while we have utilized a U-Net backbone for our Semi-Siamese branches, we can replace it with any state-of-the-art encoder-decoder architecture of choice.

3.2 Semi-Siamese network architecture

Our deep learning model begins with two U-Net branches sharing an identical architecture. Each U-Net has five fully convolutional blocks to do downsampling (feature extraction) and four convolutional blocks to do upsampling (reconstruction). Each feature extraction block is composed of two 3×3 convolutional layers followed by a batch normalization layer and a rectified linear unit (ReLU) activation. For the first four feature extraction blocks, there is a 2D max pooling layer after each block, where each max pooling layer has pool size 2×2 and strides of 2. For each of the first four feature extraction blocks, the size of the feature maps is thus reduced by half, while the number of channels is doubled. In the last feature extraction block, there is no max pooling layer, so the size of the feature map remains the same and only the number of channels is doubled. For the reconstruction blocks, each block starts with 3×3 convolution layers followed by a batch normalization layer and ReLU activation. Analogous to the feature extraction block, the size is doubled each time but the number of channels is reduced by half. Before each reconstruction layer, there is a 2D transposed convolutional layer for upsampling (upsampling layer). Skip connections link the output from the max pooling layers to the corresponding upsampling layers. From these Semi-Siamese branches, a pair of feature maps are generated and their difference is calculated to get the change map. This change map is then fed into the remaining FCN, which generates the final change mask \hat{Y} giving the predicted probability for each pixel of whether it corresponds to a location with no defect, over-extrusion, or under-extrusion. A full schematic of the proposed architecture is shown in Figure 3.

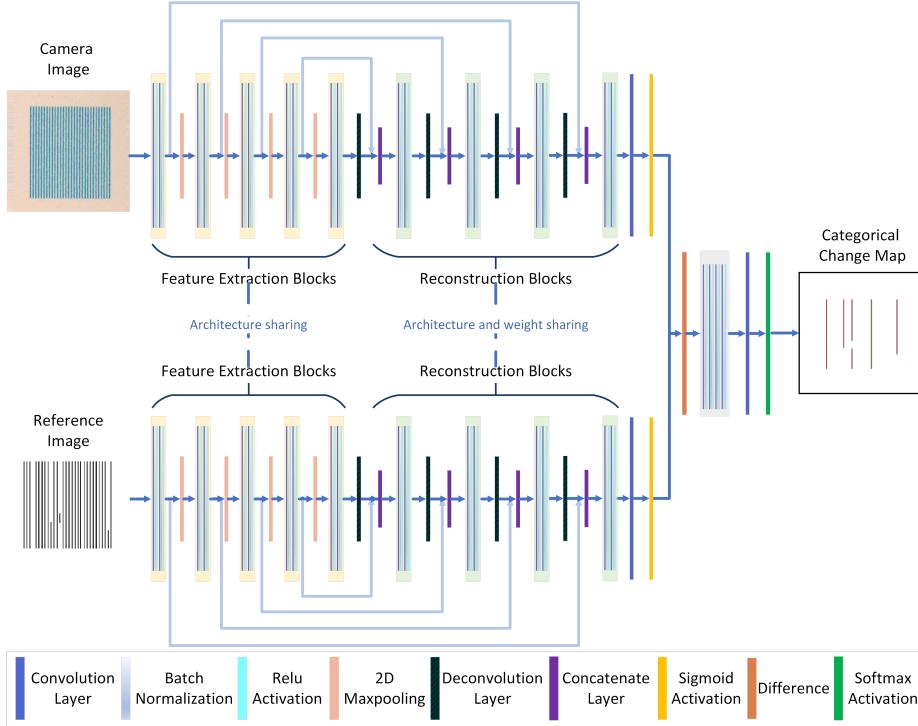


Fig. 3: Full architecture for our model. Different types of layers are labeled by color, as indicated in the legend. The Semi-Siamese branches each consist of an underlying U-Net architecture, with a feature extraction (encoder) section, a reconstruction (decoder) section, and skip connections between corresponding layers.

3.3 Training Objective

Our model uses focal loss [13] as the loss function in order to address the imbalance in change detection datasets between easy-to-classify background pixels and the smaller number of foreground pixels where changes may occur. We note that this imbalance is more pronounced in datasets from our 3D printing application than in benchmark datasets for change detection typically derived from satellite imagery. The focal loss leverages a modulating term on the cross entropy loss in order to focus learning on hard samples by reducing the contributions to the training loss of samples for which the model is confidently correct (easy samples). The equation for the focal loss is

$$\mathcal{FL}(\mathbf{y}, \mathbf{p}) = - \sum_{i=1}^N \alpha_i (1 - p_i)^\gamma \log(p_i) \quad (1)$$

where p_i is the predicted probability for the true class y_i for pixel i , and $(1 - p_i)^\gamma$ is the modulating term, with tunable focusing parameter $\gamma \geq 0$. The hyperparameters α_i are additional weighting parameters that re-balance the contribution of each pixel to the loss, typically based on its true class. Here we denote \mathbf{y} as the vector of all y_i and \mathbf{p} as the vector of all p_i .

4 Experiments

We compare our model against three change detection methods that represent different existing state-of-the-art approaches: ChangeNet [22], which utilizes a ResNet backbone and Semi-Siamese branches with shared encoders; BIT [2], which utilizes transformers; and DTCDN, which utilizes generative adversarial networks [12].

First, we compare our model with existing methods for heterogeneous change detection across different image domains using the benchmark Wuhan dataset [1], which consists of pairs of optical and synthetic aperture radar (SAR) images. Note that this dataset only involves binary classification and does not contain significant perturbations in image alignment and angle; there is no benchmark dataset for heterogeneous change detection in the multi-class case. To test the application of our model to 3D printing with three-class classification (no defect, under-extrusion, and over-extrusion), and also to demonstrate our model’s robustness to perturbations in camera angle and lighting, we then created our own experimental dataset consisting of pairs of (1) reference print schematics and (2) top-down camera images of inkjet-based 3D printing on powder bed material, where Dimatix Blue model fluid is used on Visijet core powder. To simplify dataset generation, our set of reference schematic images consist only of images with vertical lines of varying length spaced closely together. The corresponding camera images are taken at various angles, so that a simple template matching approach would not be able to easily achieve pixel-wise accuracy in defect localization. In future work, we will expand to more complex schematic images. However, the efficacy of our model in precisely localizing defects can be sufficiently demonstrated using this dataset.

4.1 Full dataset generation using data augmentation

Since generating an experimental dataset is time-consuming, and with a real 3D printer also incurs significant material costs, we start with only a limited dataset of 57 pairs of experimental images. We then use data augmentation to significantly increase our dataset size by adding perturbations in camera angle to existing camera images. Note that lighting perturbations come naturally from the camera images being taken with no special lighting setup. We do not make any changes to the reference schematics. The types of perturbations we use in the data augmentation includes zoom, rotation, shear, and position (width and height) shift. We emphasize that here we use data augmentation to create our initial full dataset, in contrast to the typical setting in computer vision where data augmentation should only be used in training. The perturbations to the

camera image given by data augmentation correspond to artificial new “experiments” of different camera setups. To prevent data leakage, we separate the training, validation, and test sets by reference schematic. The final dataset consists of 16400 training, 560 validation, and 560 test images, where 41 underlying schematic images are used in the training data, and 8 schematic images each were used for validation and test.

To create defective image pairs, we match the camera image from one reference image with a different reference schematic. Since all of the images correspond to perfect non-defective prints for their true corresponding schematic, we can precisely localize the “defects” in the defective image pairs by comparing the camera image’s true reference schematic with the given new schematic. To generate the training, validation, and test sets, we first randomly select either one underlying schematic for non-defective examples, or two different underlying schematics for defective examples. Then we randomly pick corresponding camera images from among the perturbed variations in our augmented dataset. We note that it is important to balance the dataset between defective and non-defective pairs; otherwise the trained model tends to predict the presence of some defects even for non-defective pairs.

5 Results

Table 1 shows the results of different change detection methods, as well as an ablation study on the Semi-Siamese and transfer learning components of our approach, on our generated vertical line dataset. Note that due to the imbalance between classes in this dataset, the under-extrusion class was the most difficult to correctly identify. We report the F1-score for each class, as well as the averaged macro F1-score. Our Semi-Siamese model with initialization was able to achieve significantly higher performance than the other methods on identifying under-extrusion. Even without initialization, using all of the same hyperparameters for handling the imbalanced dataset, the Semi-Siamese model outperformed other methods including BIT. In Figure 4, we provide both visual and quantitative comparisons of each method on several example pairs of images. Compared to

Table 1: Performance comparison on the vertical line dataset. Note that due to the imbalance between classes in this dataset, the under-extrusion class was the most difficult to correctly identify.

Method	Accuracy	macro F1-score	no-defect F1-score	over-extrusion F1-score	under-extrusion F1-score
ChangeNet	0.9842	0.5646	0.9920	0.6431	0.0588
DTCDN	0.9572	0.7613	0.9772	0.6970	0.6097
BIT	0.9957	0.9267	0.9978	0.9555	0.8268
Unet	0.9774	0.7155	0.9884	0.6859	0.4722
Semi-Siam (w/o init)	0.9962	0.9406	0.9981	0.9709	0.8529
Siamese (w/ init)	0.9658	0.6652	0.9823	0.4629	0.5503
Semi-Siam (w/ init)	0.9972	0.9517	0.9986	0.9503	0.9061

Table 2: Performance comparison on the Wuhan dataset [1]. DTCDN results are from Ref. [23] since it was difficult to reproduce the high performance without modifications, as also noted in [23].

Method	Precision	Recall	IOU	F1-score
ChangeNet	0.6555	0.6326	0.5232	0.6420
DTCDN*	0.6742	0.6536	0.5492	0.6629
BIT	0.6678	0.6859	0.5564	0.6759
Semi-Siam (w/ init)	0.6714	0.7247	0.5659	0.6905
Siamese (w/ init)	0.6571	0.6833	0.5476	0.6681
Semi-Siam (w/ init)	0.7306	0.7263	0.6113	0.7284

Table 3: All training times are based on 200 epochs (not including pre-training the GAN in the case of DTCDN, and pre-training the U-Net in the case of Semi-Siamese with initialization) on the vertical line dataset on a 4-GPU workstation.

Method	Training	Setup	Prediction
ChangeNet	178.14 h	1797.155 ms	423.406 ms
DTCDN	GAN:36.74 h U-Net++:56.28 h	GAN: 1577.17 ms U-Net++:150.958 ms	GAN:10410.228 ms U-Net++:490.128 ms
BIT	105.39 h	569.835 ms	342.935 ms
U-Net	19.34 h	136.563 ms	213.76 ms
Semi-Siam (w/o init)	48.31 h	206.847 ms	419.283 ms
Siamese (w/ init)	36.83 h	179.195 ms	408.974 ms
Semi-Siam (w/ init)	U-Net:9.34 h Semi-Siam: 48.31 h	206.847 ms	419.283 ms

existing methods, in most cases our model is able to capture the defect locations more precisely and with less noise.

Table 2 shows the results of different change detection methods (including an ablation study similar to above) on the benchmark Wuhan dataset, and Figure 5 provides visual and quantitative comparisons of each method on sample pairs of images. We note that visually, our Semi-Siamese model with initialization is able to better reproduce the smooth shape of the ground truth mask. In the second case where the macro F1-score of DTCDN and BIT outperform our method, the key difference is the detection of a change in a round region in the top left quadrant. While this is not labeled in the ground truth, on close inspection one can see that this is not necessarily inconsistent with the SAR image.

In addition to being more accurate, our method is comparatively simpler and more lightweight. From Table 3, we can see that our model takes significantly less time to train than ChangeNet, DTCDN, and BIT.

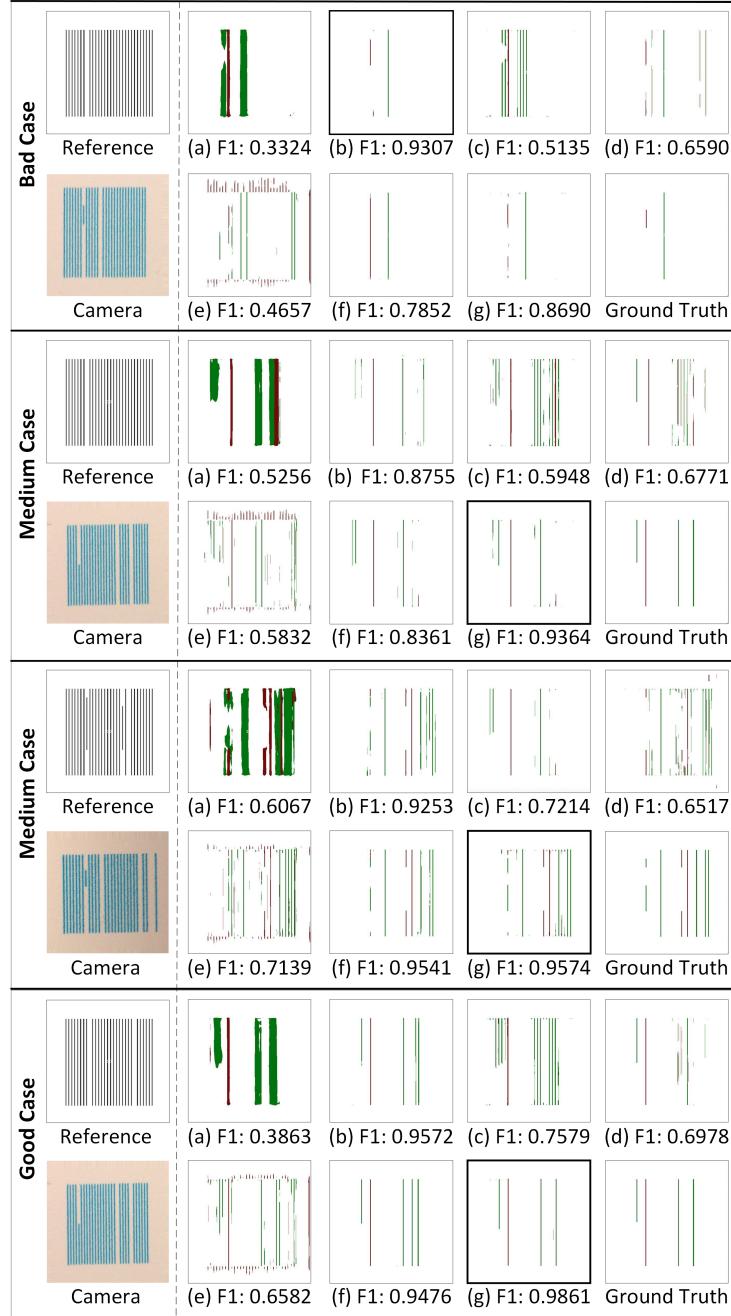


Fig. 4: Visual and quantitative (macro F1-score) comparison of different models on the vertical line dataset. The methods from (a)-(g) are (a) ChangeNet, (b) BIT, (c) DTCDN, (d) U-Net, (e) Siamese model with initialization, (f) Semi-Siamese model without initialization, (g) Semi-Siamese model with initialization. White, red and green correspond to no defects, over-extrusion, and under-extrusion, respectively.

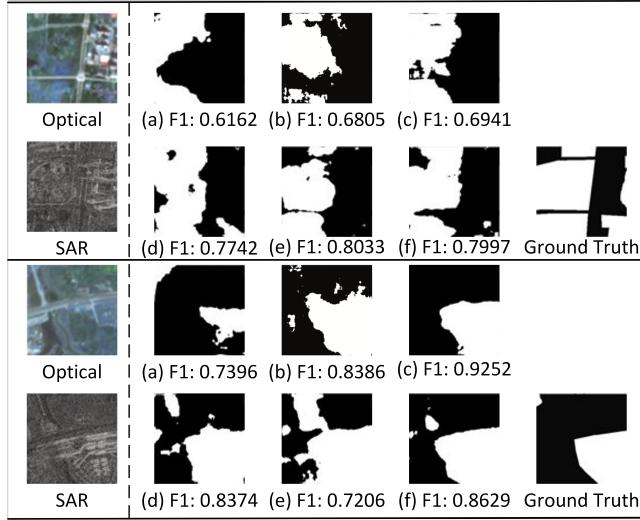


Fig. 5: Visual and quantitative comparison of different models on the Wuhan dataset [1]. The methods from (a)-(f) are (a) ChangeNet, (b) DTCNN, (c) BIT, (d) Siamese model with initialization, (e) Semi-Siamese model without initialization, and (f) Semi-Siamese model with initialization. DTCNN results are from Ref. [23] as it was difficult to reproduce the high performance without modifications, as also noted in [23].

6 Conclusions

We have developed a new deep learning-based method for change detection where (i) a camera image is being compared against a desired schematic rather than another camera image, and (ii) perturbations to the camera angle and lighting do not need to be pre-corrected, and coregistration is not necessary. This novel Semi-Siamese model can be applied to obtain precise *in situ* pixel-wise defect localization for each layer of a 3D print, enabling rapid detection of internal defects, ensuring the quality of 3D printed parts and saving time and material costs. While an acknowledged limitation of this method is that it does not directly handle defects in the z-direction in a *single* layer, due to the ability to observe each printed layer at various perturbed camera angles, large z-direction defects in the top layer will likely project onto the 2D camera image in such a way as to appear as in-plane defects. Robust handling of these types of z-directional defects will be explored in future work.

Defect detection for 3D printing is an important industrial challenge that to the best of our knowledge is being addressed with change detection techniques for the first time in this work. The key benefit of utilizing the change detection framework is that it is not necessary to pre-define the desired print schematic, nor to have a large set of annotated data for each defect type. Our model is capable of detecting defects in a few seconds with more than 90%

accuracy, and performs better than many different more complicated state-of-the-art approaches: ResNet-based Semi-Siamese models with shared encoders (ChangeNet), generative adversarial network (GAN)-based models (DTCDN), and transformer-based models (BIT), on not only our custom 3D printing dataset but also on the benchmark heterogeneous change detection Wuhan dataset. The simplicity of our model makes it possible to easily achieve good performance on new problems - it is only necessary to pre-train a U-Net (or other encoder-decoder backbone) and then transfer learn from that onto the Semi-Siamese architecture. The robustness of our algorithm to camera angle and lighting perturbations while enabling domain adaptation, as well as its lower training data requirements, will enable flexibility for utilizing this model in different industrial settings.

References

1. Caltagirone, F., Capuzi, A., Coletta, A., De Luca, G.F., ScorzaFava, E., Leonardi, R., Rivola, S., Fagioli, S., Angino, G., L'Abbate, M., Piemontese, M., Zampolini Faustini, E., Torre, A., De Libero, C., Esposito, P.G.: The cosmo-skymed dual use earth observation program: Development, qualification, and results of the commissioning of the overall constellation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **7**(7), 2754–2762 (2014). <https://doi.org/10.1109/JSTARS.2014.2317287>
2. Chen, H., Qi, Z., Shi, Z.: Remote sensing image change detection with transformers. *IEEE Transactions on Geoscience and Remote Sensing* **60**, 1–14 (2021). <https://doi.org/10.1109/TGRS.2021.3095166>
3. Daudt, R.C., Saux, B.L., Boulch, A.: Fully convolutional siamese networks for change detection. *2018 25th IEEE International Conference on Image Processing (ICIP)* pp. 4063–4067 (2018). <https://doi.org/10.1109/icip.2018.8451652>
4. Fastowicz, J., Grudziński, M., Teclaw, M., Okarma, K.: Objective 3d printed surface quality assessment based on entropy of depth maps. *Entropy* **21** (2019). <https://doi.org/10.3390/e21010097>
5. Feng, W., Tian, F., Zhang, Q., Zhang, N., Wan, L., Sun, J.: Fine-grained change detection of misaligned scenes with varied illuminations. *2015 IEEE International Conference on Computer Vision (ICCV)* pp. 1260–1268 (2015). <https://doi.org/10.1109/iccv.2015.149>
6. Guo, E., Fu, X., Zhu, J., Deng, M., Liu, Y., Zhu, Q., Li, H.: Learning to measure change: Fully convolutional siamese metric networks for scene change detection. *ArXiv* **abs/1810.09111** (2018)
7. Han, Y., Yu, H.: Fabric defect detection system using stacked convolutional denoising auto-encoders trained with synthetic defect data. *Applied Sciences* **10**, 2511 (2020). <https://doi.org/10.3390/app10072511>
8. Jacobsmühlen, J.Z., Kleszczynski, S., Witt, G., Merhof, D.: Detection of elevated regions in surface images from laser beam melting processes. *IECON 2015 - 41st Annual Conference of the IEEE Industrial Electronics Society* pp. 001270–001275 (2015). <https://doi.org/10.1109/iecon.2015.7392275>
9. Jin, Z., Zhang, Z., Gu, G.: Autonomous in-situ correction of fused deposition modeling printers using computer vision and deep learning. *Manufacturing Letters* **22**, 11–15 (2019). <https://doi.org/10.1016/j.mfglet.2019.09.005>

10. Khan, S.H., He, X., Porikli, F., Bennamoun, M.: Forest change detection in incomplete satellite images with deep neural networks. *IEEE Transactions on Geoscience and Remote Sensing* **55**(9), 5407–5423 (2017). <https://doi.org/10.1109/tgrs.2017.2707528>
11. Koch, G.R.: Siamese neural networks for one-shot image recognition. In: ICML Deep Learning Workshop (2015)
12. Li, X., Du, Z., Huang, Y., Tan, Z.: A deep translation (gan) based change detection network for optical and sar remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing* **179**, 14–34 (2021). <https://doi.org/10.1016/j.isprsjprs.2021.07.007>
13. Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal loss for dense object detection. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 2980–2988 (2017). <https://doi.org/10.1109/TPAMI.2018.2858826>
14. Malila, W.A.: Change vector analysis: An approach for detecting forest changes with landsat. In: Proceedings of the 6th Annual Symposium on Machine Processing of Remotely Sensed Data (1980)
15. Mesquita, D.B., Santos, R.F.D., Macharet, D., Campos, M., Nascimento, E.R.: Fully convolutional siamese autoencoder for change detection in uav aerial images. *IEEE Geoscience and Remote Sensing Letters* **17**, 1455–1459 (2020). <https://doi.org/10.1109/lgrs.2019.2945906>
16. Ronneberger, O., P.Fischer, Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI). LNCS, vol. 9351, pp. 234–241. Springer (2015). https://doi.org/10.1007/978-3-662-54345-0_3
17. Saha, S., Bovolo, F., Bruzzone, L.: Building change detection in vhr sar images via unsupervised deep transcoding. *IEEE Transactions on Geoscience and Remote Sensing* **59**(3), 1917–1929 (2021). <https://doi.org/10.1109/tgrs.2020.3000296>
18. Sakurada, K., Okatani, T.: Change detection from a street image pair using cnn features and superpixel segmentation. In: BMVC (2015). <https://doi.org/10.5244/c.29.61>
19. Scime, L., Siddle, D., Baird, S.T., Paquit, V.: Layer-wise anomaly detection and classification for powder bed additive manufacturing processes: A machine-agnostic algorithm for real-time pixel-wise semantic segmentation. *Additive manufacturing* **36**, 101453 (2020). <https://doi.org/10.1016/j.addma.2020.101453>
20. Shi, W., Zhang, M., Zhang, R., Chen, S., Zhan, Z.: Change detection based on artificial intelligence: State-of-the-art and challenges. *Remote Sensing* **12**(10) (2020). <https://doi.org/10.3390/rs12101688>
21. Sublime, J., Kalinicheva, E.: Automatic post-disaster damage mapping using deep-learning techniques for change detection: Case study of the tohoku tsunami. *Remote Sensing* **11**(9), 1123 (2019). <https://doi.org/10.3390/rs11091123>
22. Varghese, A., Gubbi, J., Ramaswamy, A., Balamuralidhar, P.: Changenet: A deep learning architecture for visual change detection. In: ECCV Workshops (2018). https://doi.org/10.1007/978-3-030-11012-3_10
23. Zhang, C., Feng, Y., Hu, L., Tapete, D., Pan, L., Liang, Z., Cigna, F., Yue, P.: A domain adaptation neural network for change detection with heterogeneous optical and sar remote sensing images. *International Journal of Applied Earth Observation and Geoinformation* **109**, 102769 (2022). <https://doi.org/10.1016/j.jag.2022.102769>