3D Defect Detection and Metrology of HBMs using Semi-Supervised Deep Learning

Ramanpreet Singh Pahwa[1, 2], Richard Chang[1], Wang Jie[1], Zhao Ziyuan [1,2], Cai Lile[1, 2], Xu Xun[1, 2], Foo Chuan Sheng[1, 2], Chong Ser Choong[3], Vempati Srinivasa Rao[3],
[1]Institute for Infocomm Research (I^{2}R), Agency for Science, Technology and Research (A*STAR), Singapore
[2]Artificial Intelligence, Analytics And Informatics (AI^3), Agency for Science, Technology and Research (A*STAR), Singapore
[3]Institute of Microelectronics (IME), Agency for Science, Technology and Research (A*STAR), Singapore
{ramanpreet_pahwa, richard_chang, zhao_ziyuan, wang_jie, caill, xu_xun, foo_chuan_sheng}@i2r.a-star.edu.sg, {chongsc, vempati}@ime.a-star.edu.sg

***Abstract*—3D Deep Learning has made tremendous progress recently and is being widely used in various fields, such as medical imaging, robotics, and autonomous vehicle driving, to identify and segment various structures. In this work, we leverage the recent developments in 3D semi-supervised learning to develop state-of-the-art models for 3D object detection and segmentation for various buried structures such as memory and logic die. We briefly describe our approach to fabricating, generating 3D scans, and annotating these samples. Thereafter, we explain our approach to locating these buried structures by demonstrating how semi-supervised learning is adopted to leverage vast amounts of available unlabeled data to improve both detection and segmentation performance. We also develop a metrology package that performs post-processing and outputs various important metrics for each package, such as void-to-solder ratio, pad misalignment, solder extrusion, and bond line thickness. Overall, we observe an improvement of up to 16% in object detection and 6% in 3D segmentation. Our final metrology results show a mean error of less than 1.24um for BLT and 0.753um for Pad misalignment when compared to the ground-truth labeled data.**

***Index Terms*—3D Semi-Supervised Learning, Multi-view Object detection, 3D Multi-class segmentation, 3D Metrology, HBM failure analysis**

I. Introduction

Buried defects in miniaturized package interconnects are a major cause of low yield in 2.5D-3D packages. Detecting these defects is challenging and time-consuming. The most common failure analysis methodology entails a destructive process in which a semiconductor package is cross-sectioned. After that, we can use optical inspection to identify embedded process defects such as solder extrusions, pad misalignment, embedded voids, and solder shorts. Usually, this failure analysis step is done manually with the assistance of off-the-shelf image processing techniques. However, this approach requires significant effort, man-hours, domain expertise, and expensive tools. Additionally, it only provides information in a single 2D plane and we need to repeat the process if we require additional information in the neighborhood regions. The new 3D X-ray machines are able to provide an acceptable resolution to examine and analyze buried features such as Through Silicon Vias (TSVs), micro-bumps, and other metallic structures. This can be an incredible non-destructive methodology for failure analysis in the future. Currently, the scanning takes anywhere between 2-8 hours per sample and some regions suffer from glaring artifacts [1]. Thus, currently, the tools may be impractical to be deployed in a real setting. However, like any technology, more advancements in the future will help reduce the acquisition and improve the quality of these scans.

Artificial Intelligence (AI) has significantly shaped technological progress in various fields such as visual surveillance [2], predictive maintenance [3], object detection, and image segmentation [4]–[7]. In particular, with the advancement of computing and additional research focus on efficient machine learning techniques, the scientific community has shown great progress in the 3D Deep Learning field [8]–[10]. Recently, deep learning has been adopted to perform this task of 2D-3D detection and segmentation in buried packages [11]–[13].

Usually, one requires a huge amount of labeled data to train an accurate model with deep learning-based approaches. This requires a lot of man-hours and can be extremely expensive. In a fast-moving world where a chip may go through multiple revisions in a year, companies may not have the time and logistics to spend months developing deep learning models for failure analysis that may become obsolete in less than a year.

We advance our prior work [14], [15] where we developed fully supervised learning (FSL) based 2D object detection and segmentation models followed by a 2D semi-supervised segmentation learning (SSL) based approach. We introduce a novel hierarchical consistency regularized mean teacher framework for performing 3D Object detection and segmentation on 3D X-ray scans consisting of dense High Bandwidth

This research is supported by Career Development Fund (Grant no. C210812046) and AME Programmatic Funds (Grant No. A20H6b0151) which are supported and administered by the Agency for Science, Technology and Research (A*STAR).

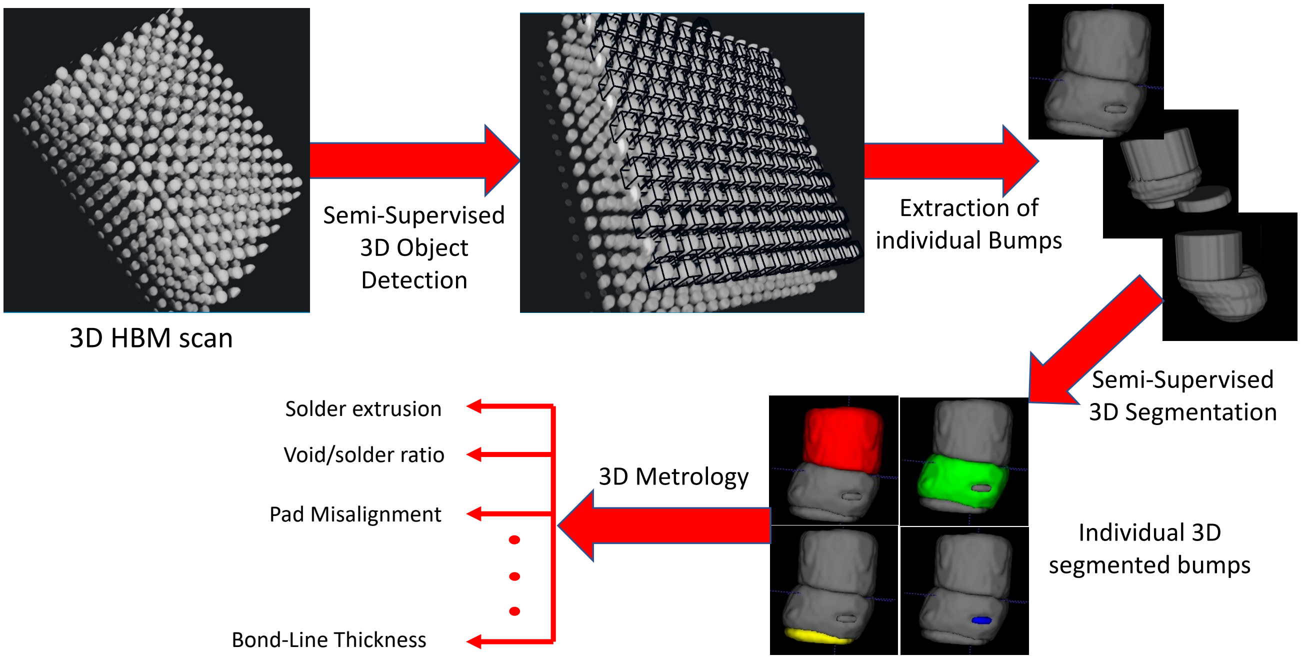


Fig. 1: Our methodology to perform accurate 3D metrology on 2.5D-3D bumps. First, we locate 3D scans of each bump individually using our multi-view Semi-Supervised object detection. Second, we employ our 3D Semi-Supervised Segmentation model to identify each component of the bump. Finally, we use our custom 3D metrology Python toolbox to measure and identify various defects such as Pad Misalignment, Void-to-Solder Ratio, Bond Line Thickness, etc.

Memory (HBM) packages with only a handful of 3D labeled data. This approach follows an efficient and sophisticated AI-based automated attribute measurement technique that provides critical information about the HBMs such as Bond Line Thickness (BLT), solder extrusion, void-to-solder ratio, and pad misalignment.

We use a three-step process for our Semi-Supervised based AI-based automated HBM metrology as shown in Fig. 1. First, we describe our novel approach to identifying HBMs using multiple views - a step-up from our previous single-view 2D approach. We extend this approach further by using Semi-Supervised Object detection to improve performance. After locating these bumps in 3D X-ray Machine (XRM) data, we crop these bumps into individually isolated Regions of interest (RoI). In the second step, these RoIs are processed by another novel semi-supervised based segmentation model that is capable of identifying various components such as Copper Pillar, Copper pad, Solder, and void. Finally, in the third step, we use our custom sophisticated computer vision package developed specifically for 3D HBMs that analyzes each bump individually to output a personalized report on solder extrusion, bond-line thickness, solder extrusion, void-to-solder ratio, etc.

Our contributions can be summarized below:

- Multi-view SSL object detection approach to identify each HBM bump accurately.
- 3D SSL semantic segmentation to identify each individual component of HBM.
- Custom computer-vision-based package to provide 3D Metrology details on each package for failure analysis.

We present the related works in Sec. II. We summarize our fabrication, 3D scanning, and annotation methodology in Sec. III-A. Thereafter, we provide a detailed explanation of our multi-view Semi-Supervised Object detection approach in Sec. III-B, 3D Semi-Supervised image segmentation methodology in Sec. III-C and our custom 3D HBM Metrology approach in Sec. III-D. Sec. IV investigates our object detection, and 3D segmentation approach and demonstrates our capabilities by displaying end results and providing critical metrology information to enable end-users to make strategic decisions. Finally, we conclude this work in Sec. V and discuss the gaps in the current approach and potential future directions.

II. Related Work

A. Object detection

Many object detectors have been introduced and improved versions are regularly published [16]–[19]. Although their performance improves over time, the amount of data required to train the models is still huge and public datasets are still used for training or evaluating their performance. Using these detectors on specific data would require training or fine-tuning the network. However, to the best of our knowledge, there is no public dataset on semiconductors that focus on HBM bump detection. One solution is to manually annotate our data which is time-consuming and costly since it requires domain knowledge. We focus on semi-supervised learning where only

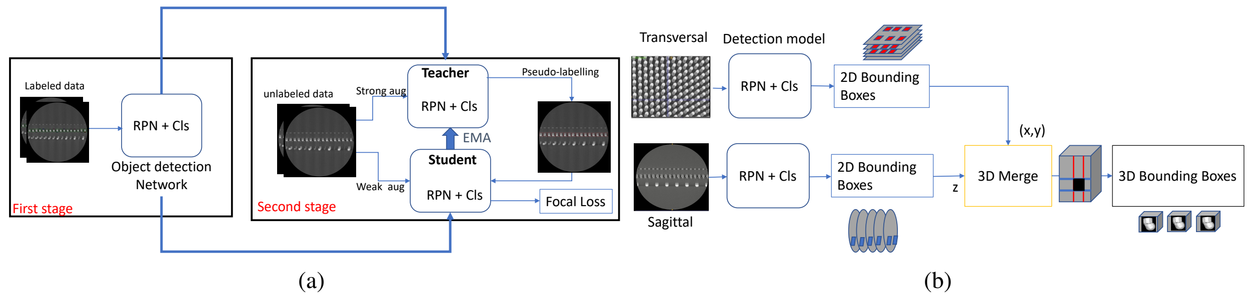


Fig. 2: (a) Semi-supervised training overview. It includes a first stage (burn-in) with labeled data and a second stage with unlabelled data and a student-teacher mutual learning framework. (b) 3D-slice-and-fuse approach with the detection model. Both detectors will run on transversal and sagittal views and output 3D bounding boxes corresponding to the bumps in the scans.

a small amount of labeled training data is required and the detection network is able to use unlabeled data to improve its performance.

Semi-supervised object detectors have been introduced in the literature [20], [21] and allow to reduce the requirement on labeled data which is very difficult to obtain in the semiconductor field. Different approaches have been proposed, from generating pseudo-labels from the model trained on partially labeled data to using data augmentations and contrastive learning to improve the training process and using specific loss functions. The most common approach involves a student-model teacher framework consisting of two stages. The first stage called the burn-in stage, where the detection model is pre-trained on labeled data. In the second stage, the weights are shared with both the student and teacher models. Different data augmentations techniques are applied to both student and teacher models and the consistency of the outputs is enforced to update the teacher's weights using Exponential Moving Average (EMA) [22].

3D object detection has been widely studied for different applications such as smart robot navigation [23], but they focus on LiDAR pointcloud data which is different from voxelized scans in semiconductors. Some approaches have also been introduced in medical imaging [24], [25] for liver and kidney detection in X-ray scans. Our previous slice-and-fuse method [14], [15] is applied to logic and memory bumps.

B. 3D Semantic Segmentation

3D semantic segmentation task is widely used for medical imaging [26] and semiconductor field [27]. Both scenarios require dense mask predictions to reveal and identify the internal structures present in 3D regions. Annotation difficulty, data complexity, and class imbalance [28] are some of the major challenges in 3D segmentation. Some recent improvements have introduced strong data augmentations for better generalization ability [29] and adopted various loss functions [30], [31].

The recent success of semi-supervised learning emerges under various tasks involving limited labeled data. Several self-ensembling methods, such as mean-teacher models [20], have been used to ensure that the predictions remain consistent with the inputs subjected to various perturbations between the student and teacher models. Based on the classic V-Net [32] backbone, several advancements have been made to enhance the teacher-student training, such as imposing geometric constraints for shape-awareness [33], introducing representation consistency under multi-task framework [34] and exploiting latent information of network [35]. In this work, we use V-Net as our backbone and explore the impacts of dataset complexity and ground-truth annotation.

III. Our Approach

A. Data Collection

This section briefly describes our approach to fabricating, scanning, and annotating the 2.5D Test Vehicles (TV). The fabricated test vehicles resemble contemporary High-Performance Computing packages, particularly logic, and memory bumps. Both types of wafers are temporarily bonded to a carrier and thinned to $50\,\mu m$. The memory wafers are diced and the chips are stacked individually on top of the logic wafer to form High Bandwidth Memory cubes. The readers can refer to [36], [37] for further details on memory and logic bumps fabrication.

In our design and fabrication step, we induce defects in our packages purposely by using sub-optimal packaging parameters to generate defective data consisting of solder extrusions, voids, pad misalignments, etc. that can be used to train our deep learning models. After the fabrication step, we use a 3D XRM scanner to generate our 3D scans. The samples are placed inside the tool and incrementally rotated from -5° through 185° to obtain raw 3D X-ray scans of 1 billion voxel resolution.

We primarily used ITK-SNAP [38] to label the HBMs. The Background is labeled as 0, Copper Pillars are labeled as 1, Solder as 2, the Void as 3, and Copper Pads as 4. Subject experts who have worked on failure analysis for multiple years helped train the labelers in locating the correct boundaries for each component. However, due to Beam Hardening artifacts [1], finding the correct boundaries of Pads, Pillars, and other components is extremely tough with subject experts widely varying in their opinions. As such, we are less focused on benchmarking the segmentation results against groundtruth and more interested in whether our approach is able to provide meaningful information to the users about 3D metrology.

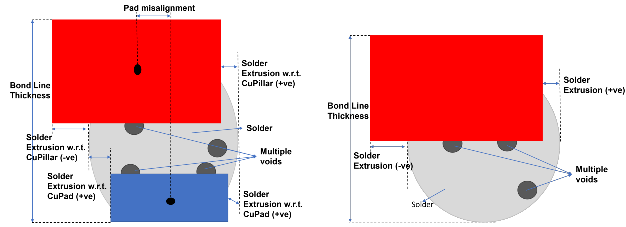
B. Object detection

The purpose of the object detection step in our approach is to detect and extract each individual bump from the overall scans for further processing. Given the size of each scan ($1000 \times 1000 \times 1000$ voxels), this pre-processing step is required for segmentation and defect detection in individual bumps due to GPU memory limitations. Furthermore, extracting individual bumps allows the segmentation and defect detection sections to process the bumps at their full resolution without any downsampling which may impact segmentation and metrology accuracy.

We introduce a multi-view semi-supervised detection framework that incorporates our 3D slice-and-fuse approach. We combine the robustness of our 3D slice-and-fuse but with less required labeled data for better efficiency and productivity. The unbiased teacher semi-supervised network [21] is used as the backbone. Figure 2 shows the architecture of the network and our 3D slice-and-fuse process. For each scan sc_k, we first extract n_t individual slices $I^{t,sc_k}_{i=0,\ldots,n}$ from the transversal view and n_s slices $I^{s,sc_k}_{i=0,\ldots,n}$ from the sagittal view. Two detection networks are trained and used to detect the bumps, one for each view, sagittal and transversal since each bump has a different appearance in these views.

For training purposes, we split the data into two sets, S^t_l and S^t_u representing our labeled and unlabelled data respectively. The student-teacher mutual learning process is applied here. First, a burn-in stage is performed with S^t_l as shown in Fig. 2(a). Second, the weights are duplicated to the student and teacher model. The teacher model generates pseudo-labels on the unlabeled data S^t_u with some weak data augmentations and the student uses the pseudo-labels after applying robust data augmentations for its own training process. The knowledge that the student has learned will then be transferred to the teacher model through the exponential moving average (EMA) on the network's weights. The same approach is adopted for the sagittal view. These outputs from the two views are then incorporated into the 3D slice-and-fuse approach.

The first 2D detection phase of the approach consists of processing each individual input slice I_i to the trained models and outputting the list of p bounding boxes $[x, y, w, h]^i_p$ corresponding to the detected bumps. This phase is applied to both sagittal and transversal views making it more robust to bumps' defects. The sagittal view provides the critical information on the dimension z and the transversal view provides the information on the dimension x, y. The bounding boxes are then assembled to create a 3D bounding box $[x, y, z, w, h, d]^k_p$ corresponding to the location and dimensions of the identified bump in the scan sc^k. The 3D individual bumps are extracted for further processing as shown in Fig. 2(b).



(a) Memory bump (b) Logic Bump

Fig. 3: Our 3D metrology features for HBM bumps. The features are computed in 3D using cross-sectional results. We only show a 2D slice illustration for ease of understanding the metrology approach. BLT for logic only includes the vertical height of the solder and Copper Pillar.

C. 3D semantic segmentation

Semantic segmentation is applied at the bump level to differentiate the volumetric structure and thus identify manufacturing defects in the 3D metrology step. The normal structure of memory and logic bump consists of four foreground components: copper pillar (Cu-Pillar), solder, copper pad (Cu-Pad), and void as regions of interest. By accurately depicting internal volumetric structures as shown in Fig. 1, we can facilitate further study on 3D metrology.

In this work, we face issues of limited labeled data and imbalance distribution of label classes. To investigate the effect of different percentages of labeled data, as well as the impact from the sparse nature of void appearance, we aim to achieve robust performance under data with less number of labeled samples. Inspired by the mean-teacher paradigm originated from the classification task, we adopt it to segmentation problems with 3D data. We adopt 3D V-Net as the backbone model. For both memory and logic bump inputs, we examine the performance against multiple metrics with various numbers of labeled samples and a consistent number of unlabeled samples.

D. 3D Metrology

We develop a custom python-based metrology module to measure specific features that are important for 2.5D-3D bump failure analysis. In particular, we measure four features considered most important according to our in-house experts namely, Bond-Line Thickness (BLT), Solder-to-void-ratio, pad misalignment, and solder extrusion. These features are described in Fig. 3.

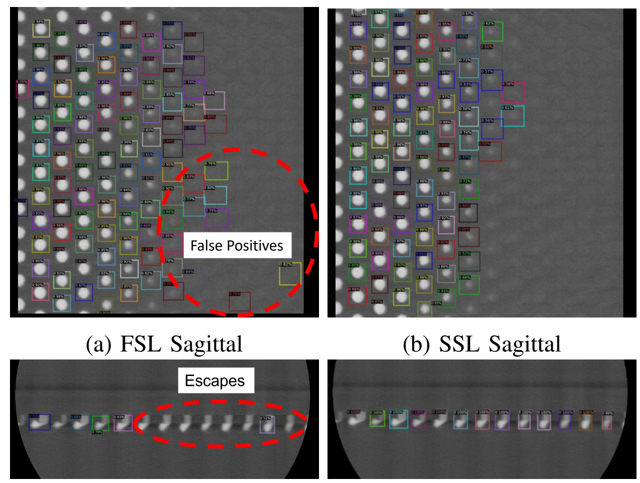
After we obtain the predicted output from our 3D segmentation model, we perform a series of post-processing steps. First, we ensure that the bumps are vertically aligned i.e. the bumps are upright in a common view. Second, we perform a series of morphological functions - dilation and erosion to ensure any pixels inside the solder, Copper Pillar, or Copper Pad that were labeled as background are correctly labeled as the relevant component. Third, we overlay the predicted voids on top of the newly cleaned predictions as void areas would have been reassigned as solder after the second step. Lastly,

Logic Die		1%		2%		5%		10%	
IoU @ 0.5:0.95		Precision	Recall	Precision	Recall	Precision	Recall	Precision	Recall
Sagittal	SSL	**0.795**	**0.836**	**0.801**	**0.841**	**0.81**	**0.845**	**0.814**	**0.846**
	FSL	0.776	0.806	0.781	0.814	0.782	0.817	0.809	0.848
Transversal	SSL	**0.788**	**0.824**	**0.782**	**0.814**	**0.782**	**0.814**	**0.782**	**0.814**
	FSL	0.714	0.753	0.659	0.703	0.679	0.725	0.701	0.739
Memory Die		1%		2%		5%		10%	
Sagittal	SSL	**0.769**	**0.801**	**0.764**	**0.81**	**0.788**	**0.827**	**0.798**	**0.832**
	FSL	0.607	0.662	0.631	0.68	0.634	0.685	0.648	0.685
Transversal	SSL	**0.764**	**0.812**	**0.781**	**0.81**	**0.798**	**0.834**	**0.824**	**0.853**
	FSL	0.723	0.784	0.76	0.79	0.784	0.824	0.803	0.846

TABLE I: We report FSL and SSL object detection Intersection over Union (IoU) for Logic and Memory die. Our SSL approach provides more accurate results than FSL approach on both sagittal and transversal views by up to 8% and 16% mAP on logic and memory bumps respectively.

we retain all the predicted voxels for each class that is within a specific threshold to the center of mass (CoM) of each class. This removes any leftover clusters or neighborhood component predictions that might be seen in the cropped individual 3D bumps.

After the pre-processing step, we identify the features needed to perform our metrology - mainly the CoM, top, left, right and bottommost regions for each individual bump. The features, shown in Fig. 3 are computed for each bump and provided to the domain experts to make critical decisions for failure analysis.



(a) FSL Sagittal (b) SSL Sagittal

(c) FSL Transversal (d) SSL Transversal

Fig. 4: Detection results with FSL(a,c) and SSL (b,d) models on logic bumps (5% split). Our approach is able to address issues such as false positives and escapes on the sagittal and transversal views.

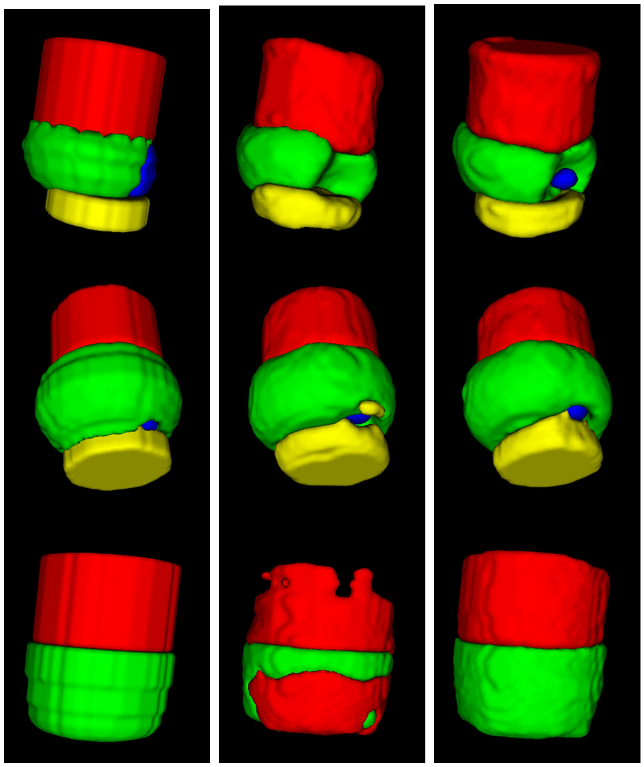
IV. Experiments

We collected and annotated a varied dataset consisting of different defects to validate our SSL-based approach. In total, we used $12,849$, $4,486$, and $4,593$ slices for training, validation, and testing for Object detection experiments. We also used 76, 36 training 3D scans and 13, 7 testing scans for memory and logic bumps respectively to evaluate our SSL-based 3D segmentation approach. Our workstation for this work includes an Intel i9-10900X CPU processor with an NVIDIA TITAN RTX 24GB GPU containing $4,608$ cuda cores.

A. Object detection

We split our dataset into different amounts of labeled data from 0.1% to 10% for our experiments. We use a FasterRCNN [18] backbone for object detection. The weak data augmentation method is a random horizontal flip and the strong augmentation methods include adding color jittering, grayscale, Gaussian blur, and cutout patches. Mean Average Precision (mAP) is used as our evaluation metric which consists of Intersection-over-Union (IoU) calculation estimating the quality of the predicted bounding boxes compared to the ground truth data [27] on different thresholds from 0.5 to 0.95. We also study the recall rates of the detector to evaluate the escapes or False Negatives. The training parameters are as follows. The learning rate is 0.01, the initial number of steps (burn-in stage) is $2,000$, and The total number of steps is $10,000$. The EMA rate is 0.996.

We first evaluate the efficacy of the 2D semi-supervised object detection on the individual slices on both sagittal and transversal views and compare it with the FSL method. Table I presents the results for the logic and memory bumps. We observe that our proposed SSL model outperforms the FSL model by up to 8% mAP for each data split for logic bumps and up to 16% for memory bumps. We also note that the performance improves when for higher splits when more labeled data is used. As the data distribution in logic bumps is more diverse and defective compared to memory bumps they are more difficult to predict with less labeled data. Given the specificity of the data and the low distribution, we notice that our SSL network performs well even with a very small amount of labeled data (1%). This demonstrates the reduced requirement for labeled data with our SSL framework. We also show the detection results on some slices for both sagittal and transversal views to highlight these differences in Fig. 4. We observe more false detections on the FSL model compared to ours. Finally, we notice a significant improvement in the number of bumps that are successfully extracted thanks to the combination of both sagittal and transversal views. With the



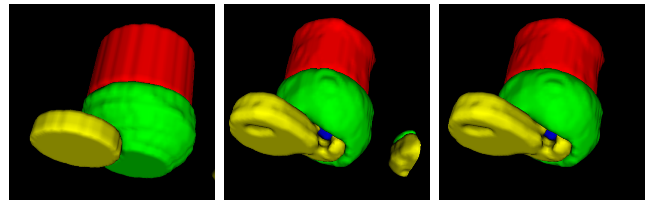
(a) Groundtruth (b) V-Net output (c) Ours

Fig. 5: We show some examples of groundtruth annotations, V-Net, and our output from our 3D segmentation models. Our approach provides more consistent and less erroneous results than our baseline.

previous method [27], some bumps could not be extracted due to defects (i.e. solder extrusion). Overall, the proposed method has extracted 511 bumps whereas the previous method only extracted 263 bumps. The implementation of the method includes 3 phases, the first phase detects the bumps on the individual slices, the second estimates the 3D bounding boxes in the scan and finally, the third phase extracts the bump into individual files. Given the structure of the data and our processing scripts, all phases can be parallelized which reduces the processing time. On average, our total processing time was about $5.14s$ per bump without any serious optimization.

B. 3D semantic segmentation

Subsets of 2.5%, 5%, 10%, 50%, and 100% labeled data are employed for training. In our framework, 3D V-Net structure is adopted as the backbone. We use multi-class Dice loss for the student model and Mean Squared Error (MSE) to update the consistency loss function for the teacher model [20]. The learning rate is set to 0.01 and step-down learning rate interval of every $5,000$ iterations. We adopt a learning rate warm-up of 300 iterations and train the backbone network from scratch. The EMA update rate is set at $\alpha = 0.999$ and the consistency weight is set to $\gamma = 0.01$. In the early phase of training, there is a linear ramping-up stage sustained for 40 epochs until it



(a) Groundtruth (b) Our raw output (c) Post-processed

Fig. 6: Our metrology package cleans up neighborhood cropped regions or any small false positive clusters to provide more accurate metrology results.

trains at full scale. All experiments last for $10,000$ iterations for the training phase which takes approximately 100 minutes.

We evaluate our model performance using four quantitative metrics: multi-class Dice coefficient, Jaccard coefficient, 95% Hausdorff Distance (95HD), and average surface distance (ASD). Table II show the Dice and IoU performance between FSL and SSL training under various percentages of selected labeled data. We observe that the model performance increases along with the addition of labeled data, and our SSL results consistently achieve higher metrics compared with baseline V-Net FSL runs. Visual results for some test samples are shown in Fig. 5.

In addition, activation functions typically serve as the main contributor to non-linearity during training. Our previous work has made attempts on comparing the effects from SiLU (swish function), SELU, and GELU [39]–[41]. However, we have not achieved promising performance in the 3D setting, which leaves room for further study.

C. 3D Metrology

We perform the metrology measurements with and without the post-processing step as described in Sec. III-D. We demonstrate the improvement in metrology measurements in Table III. We observe that aligning, cleaning, and performing neighborhood clustering drastically improves the results as most of the cropped predicted bumps include false positives for copper pillars, pads, and solder in addition to neighborhood components at the edges as shown in Fig. 6. Our final metrology results show a mean error of less than $1.24um$ for BLT and $0.753um$ for Pad misalignment when compared to the ground-truth labeled data.

V. Conclusion

In this work, we have presented a novel approach for 3D metrology using state-of-the-art 3D Semi-Supervised Deep Learning based object detection and semantic segmentation models. We explained in detail how we perform object detection on multiple views and combine the results to improve bump detection performance. This is followed by 3D Semi-Supervised semantic segmentation to identify various components in the individual structures such as Copper Pillar, Copper Pad, voids, and solder regions. We also described our post-processing approach to analyze the predicted samples

	2.5%		5%		10%		50%		100%	
	Dice	IoU	Dice	IoU	Dice	IoU	Dice	IoU	Dice	IoU
Memory Die										
V-Net (FSL)	79.89	64.86	85.14	**77.16**	81.49	**73.57**	83.26	75.63	87.89	80.19
Ours (SSL)	**80.63**	**72.38**	**85.50**	70.25	**86.69**	71.75	**86.10**	**82.03**	**88.67**	**82.81**
Logic Die										
V-Net (FSL)	81.82	75.07	**84.51**	**78.74**	84.51	78.85	85.26	79.85	**84.34**	**78.55**
Ours (SSL)	**84.22**	**78.54**	84.33	78.58	**84.80**	**79.66**	**85.65**	**80.54**	83.79	78.27

TABLE II: We report the FSL and SSL based 3D semantic segmentation results for memory and Logic die. SSL approach is generally able to identify the segments more accurately. A total of 76 memory die and 36 logic die were labeled for training our deep learning models.

Metrology Error	V-Net	Ours	Post-processed	Improvement
Memory Die				
BLT	5.654	1.292	**1.238**	0.054
SE	3.924	4.530	**1.602**	2.928
PM	2.286	0.859	**0.753**	0.106
V2SR	0.069	0.0626	**0.0594**	0.003
Logic Die				
BLT	4.55	6.65	**2.22**	4.43
SE	**0.54**	1.30	1.12	0.18
V2SR	**0.002**	0.004	0.004	0

TABLE III: We display the mean error for metrology features such as Bond-Line Thickness, solder extrusion, Pad misalignment (in um) between groundtruth, Vnet, raw, and post-processed predictions. We observe that our SSL-based approach segments the die more accurately and our metrology package further improves the results significantly.

and provided critical 3D metrology information such as Bond Line Thickness, Pad Misalignment, Solder Extrusion, etc. Our unique approach provides fast and accurate detection, segmentation, and metrology on high-resolution 3D voxelized XRM scans.

This approach when integrated with 3D acquisition systems has great potential in reducing the failure analysis time and thus improving yield in the long run. Additionally, our methodology can be customized for each unique user to potentially classify these packages as pass or fail depending on the user's self-defined criteria.

In the future, we intend to investigate the appropriate augmentations that work better for our Semi-Supervised segmentation approach and integrate continual learning to adapt to new design iterations of bumps for faster turn-around time in the failure analysis pipeline.

ACKNOWLEDGMENT

We would like to thank the Institute of Microelectronics (IME) for their invaluable support and domain expertise in designing, fabricating, and scanning the HBMs. This research is supported by Career Development Fund (Grant no. C210812046) and AME Programmatic Funds (Grant No. A20H6b0151) which are supported and administered by the Agency for Science, Technology and Research (A*STAR).

REFERENCES

[1] Rodney A Brooks and Giovanni Di Chiro, "Beam hardening in X-ray reconstructive tomography," *Physics in medicine & biology*, vol. 21, no. 3, pp. 390, 1976.

[2] Ramanpreet Singh Pahwa, Kennard Yanting Chan, Jiamin Bai, Vincensius Billy Saputra, Minh N. Do, and Shaohui Foong, "Dense 3D Reconstruction for Visual Tunnel Inspection using Unmanned Aerial vehicle," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Nov 2019, pp. 7025–7032.

[3] Ramanpreet Singh Pahwa, Jin Chao, Jestine Paul, Yiqun Li, Ma Tin Lay Nwe, S. Xie, Ashish James, Arulmurugan Ambikapathi, Zeng Zeng, and Vijay R. Chandrasekhar, "FaultNet: Faulty Rail-Valves Detection using Deep Learning and Computer Vision," in *IEEE Intelligent Transportation Systems Conference (ITSC)*, 2019, pp. 559–566.

[4] Ramanpreet Singh Pahwa, Tian Tsong Ng, and Minh N. Do, "Tracking objects using 3D object proposals," in *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, Dec 2017, pp. 1657–1660.

[5] Ziyuan Zhao, M. Xu, P. Qian, Ramanpreet Singh Pahwa, and Richard Chang, "DA-CIL: Towards domain adaptive class-incremental 3D object detection," in *British Machine Vision Conference (BMVC)*, 2022.

[6] Ramanpreet Singh Pahwa, Jiangbo Lu, Nianjuan Jiang, Tian Tsong Ng, and Minh N. Do, "Locating 3D Object Proposals: A Depth-Based Online Approach," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 3, pp. 626–639, 2018.

[7] Ramanpreet Singh Pahwa, Richard Chang, Wang Jie, Sankeerthana Satini, Chandrashekar Viswanathan, Du Yiming, Vernica Jain, Chen Tai Pang, and Wan Kong Wah, "A survey on object detection performance with different data distributions," in *International Conference on Social Robotics (ICSR)*. 2021, pp. 553–563, Springer International Publishing.

[8] Quang-Hieu Pham, Pierre Sevestre, Ramanpreet Singh Pahwa, Huijing Zhan, Chun Ho Pang, Yuda Chen, Armin Mustafa, Vijay Chandrasekhar, and Jie Lin, "A*3D Dataset: Towards Autonomous Driving in Challenging Environments," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 2267–2273.

[9] Tin Lay Nwe, Oo Zaw Min, Saisubramaniam Gopalakrishnan, Dongyun Lin, Shitala Prasad, Sheng Dong, Yiqun Li, and Ramanpreet Singh Pahwa, "Improving 3D Brain Tumor Segmentation With Predict-Refine Mechanism Using Saliency And Feature Maps," in *IEEE International Conference on Image Processing (ICIP)*, 2020, pp. 2671–2675.

[10] Ramanpreet Singh Pahwa, Wei Kiat Leong, Shaohui Foong, Karianto Leman, and Minh N. Do, "Feature-less stitching of cylindrical tunnel," in *IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS)*, June 2018, pp. 502–507.

[11] Ramanpreet Singh Pahwa, Saisubramaniam Gopalakrishnan, Huang Su, Ong Ee Ping, H. Dai, David Ho Soon Wee, Ren Qin, and Vempati S. Rao, "Automated Void Detection in TSVs from 2D X-Ray Scans using Supervised Learning with 3D X-Ray Scans," in *IEEE 71st Electronic Components and Technology Conference (ECTC)*, 2021, pp. 842–849.

[12] Ma Tin Lay Nwe, Oo Zaw Min, Richard Chang, Dongyun Lin, Shitala Prasad, Sheng Dong, Yiqun Li, and Ramanpreet Singh Pahwa, "On the use of component structural characteristics for voxel segmentation in semicon 3D images," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2022.

[13] Lile Cai, Ramanpreet Singh Pahwa, Xun Xu, Jie Wang, Richard Chang, Lining Zhang, and Chuan-Sheng Foo, "Exploring active learning for semiconductor defect segmentation," in *IEEE International Conference on Image Processing (ICIP)*, 2022, pp. 1796–1800.

[14] Ramanpreet Singh Pahwa, Richard Chang, Wang Jie, Xu Xun, Oo Zaw Min, Foo Chuan Sheng, Chong Ser Choong, and Vempati Srinivasa Rao, "Automated detection and segmentation of hbms in 3d x-ray images using semi-supervised deep learning," in *IEEE 72nd Electronic Components and Technology Conference (ECTC)*, 2022, pp. 1890–1897.

[15] Wang Jie, Richard Chang, Xu Xun, Cai Lile, Chuang Sheng Foo, and Ramanpreet Singh Pahwa, "Improved bump detection and defect identification for hbms using refined machine learning approach," in *IEEE 24th Electronics Packaging Technology Conference (EPTC)*, 2022, pp. 848–853.

[16] Chien-Yao Wang, Alexey Bochkovskiy, and Hong-Yuan Mark Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," *arXiv preprint arXiv:2207.02696*, 2022.

[17] A. Bochkovskiy, C-Y Wang, and H.Y. Liao, "Yolov4: Optimal speed and accuracy of object detection," 04 2020.

[18] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in *Advances in Neural Information Processing Systems*. 2015, vol. 28, pp. 91–99, Curran Associates, Inc.

[19] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," *Computing Research Repository (CoRR)*, vol. abs/1804.02767, 2018.

[20] Antti Tarvainen and Harri Valpola, "Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results," in *Advances in neural information processing systems*, 2017.

[21] Yen-Cheng Liu, Chih-Yao Ma, Zijian He, Chia-Wen Kuo, Kan Chen, Peizhao Zhang, Bichen Wu, Zsolt Kira, and Peter Vajda, "Unbiased teacher for semi-supervised object detection," *CoRR*, vol. abs/2102.09480, 2021.

[22] Kaiming He, Haoqi Fan, Yuxin Wu, Saining Xie, and Ross Girshick, "Momentum contrast for unsupervised visual representation learning," June 2020.

[23] Richard Chang, Ramanpreet Singh Pahwa, Jie Wang, Lawrence Chen, Sankeerthana Satini, Kong Wah Wan, and David Hsu, "Creating semi-supervised learning-based adaptable object detection models for autonomous service robot," in *12th Conference on Learning Factories (CLF)*, 2022.

[24] Imran A Tariq J Tu S Alzahrani A. Rahman H, Bukht TFN, "A deep learning approach for liver and tumor segmentation in ct images using resunet.," *Bioengineering (Basel)*, 2022.

[25] A. Lemay, "Kidney recognition in ct using yolov3," 2019.

[26] Wenjia "Bai, Ozan Oktay, Matthew Sinclair, Hideaki Suzuki, Martin Rajchl, Giacomo Tarroni, Ben Glocker, Andrew King, Paul M. Matthews, and Daniel" Rueckert, ""semi-supervised learning for network-based cardiac mr image segmentation"," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. 2017, pp. 253–260, Springer International Publishing.

[27] Ramanpreet Singh Pahwa, Ma Tin Lay Nwe, Richard Chang, Oo Zaw Min, Wang Jie, Saisubramaniam Gopalakrishnan, David Ho Soon Wee, Ren Qin, Vempati S. Rao, H. Dai, J.T. Neumann, Ramani Pichumani, and Tom Gregorich, "Automated Attribute Measurements of Buried Package Features in 3D X-ray Images using Deep Learning," in *IEEE 71st Electronic Components and Technology Conference (ECTC)*, 2021, pp. 2196–2204.

[28] Zeju Li, Konstantinos Kamnitsas, and Ben Glocker, "Analyzing overfitting under class imbalance in neural networks for image segmentation," *IEEE Transactions on Medical Imaging*, vol. 40, pp. 1065–1077, 2021.

[29] Geoff French, Samuli Laine, Timo Aila, Michal Mackiewicz, and Graham Finlayson, "Semi-supervised semantic segmentation needs strong, varied perturbations," 2019.

[30] Weiyang Liu, Yandong Wen, Zhiding Yu, and Meng Yang, "Large-margin softmax loss for convolutional neural networks," 2016.

[31] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár, "Focal loss for dense object detection," 2017.

[32] F. Milletari, N. Navab, and S. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *International Conference on 3D Vision (3DV)*, 2016, pp. 565–571.

[33] Shuailin Li, Chuyu Zhang, and Xuming He, "Shape-aware semi-supervised 3d semantic segmentation for medical images," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, pp. 552–561. Springer International Publishing, 2020.

[34] Xiangde Luo, Jieneng Chen, Tao Song, and Guotai Wang, "Semi-supervised medical image segmentation through dual-task consistency," 2020.

[35] Shumeng Li, Ziyuan Zhao, Kaixin Xu, Zeng Zeng, and Cuntai Guan, "Hierarchical consistency regularized mean teacher for semi-supervised 3d left atrium segmentation," in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine &amp Biology Society (EMBC)*. nov 2021, IEEE.

[36] Ramanpreet Singh Pahwa, Soon Wee Ho, Ren Qin, Richard Chang, Oo Zaw Min, Wang Jie, Vempati Srinivasa Rao, Tin Lay Nwe, Yanjing Yang, Jens Timo Neumann, Ramani Pichumani, and Thomas Gregorich, "Machine-Learning Based Methodologies for 3D X-Ray Measurement, Characterization and Optimization for Buried Structures in Advanced IC Packages," in *International Wafer Level Packaging Conference (IWLPC)*, 2020, pp. 01–07.

[37] Ramanpreet Singh Pahwa, Tin Lay Nwe, Richard Chang, Wang Jie, Oo Zaw Min, Soon Wee Ho, Ren Qin, Vempati Srinivasa Rao, Yanjing Yang, Jens Timo Neumann, Ramani Pichumani, and Tom Gregorich, "Deep Learning Analysis of 3D X-ray Images for Automated Object Detection and Attribute Measurement of Buried Package Features," in *IEEE 22nd Electronics Packaging Technology Conference (EPTC)*, 2020, pp. 221–227.

[38] P. A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *Neuroimage*, vol. 31, no. 3, pp. 1116–1128, 2006.

[39] Stefan Elfwing, Eiji Uchibe, and Kenji Doya, "Sigmoid-weighted linear units for neural network function approximation in reinforcement learning," 2017.

[40] Günter Klambauer, Thomas Unterthiner, Andreas Mayr, and Sepp Hochreiter, "Self-normalizing neural networks," 2017.

[41] Dan Hendrycks and Kevin Gimpel, "Gaussian error linear units (gelus)," 2016.