Deep learning for zero-defect inkjet-printing of electronics

1st Flaig Minnette
Functional Surfaces and Nano-structures
PROFACTOR GmbH
4407 Steyr, Austria
minnette.flaig@profactor.at

2nd Zambal Sebastian
Machine Vision
PROFACTOR GmbH
4407 Steyr, Austria
sebastian.zambal@profactor.at

***Abstract*—We present a vision system for automatic calculation of inkjet printed electronic structures. By adapting the printed structures to individual parts, the system is able to correct misalignments and deviations from previous process steps simply by adjusting the printed image. We propose a vision system that acquires high-resolution images after the pick & place step. A deep neural network is used to read this information and output a suitable print image for the subsequent step of inkjet printing. We evaluate the vision system in a set of experiments. The system reaches an intersection over union (IoU) of 82% on our data set. This shows the potential for future zero-defect additive manufacturing in electronics industry.**

***Index Terms*—inkjet printed electronics, deep learning**

I. Introduction

Inkjet printing is an interesting additive manufacturing technique that has a high potential in the field of electronics production. While it is not (yet) possible to print complete electronic devices, inkjet printing supports adaptive assembly and integration of individual electronic components.

In this work we propose a system that uses optical inspection to automatically infer image masks via inline inspection. These masks serve as an important basis to steer subsequent inkjet printing processes. In a larger context, the interplay between inline inspection to adaptive inkjet-printing has the potential to introduce real zero-defect manufacturing in electronics production. The background for the present work is manufacturing of RADAR and LiDAR electronic sensors as used for advanced driver-assistance systems and autonomous driving.

In the proposed approach, a printed circuit board (PCB) is manufactured with a cavity intended to enclose a bare die. The cavity is filled with glue and the bare die is picked from a wafer and placed into the cavity. Since the bare dies are slightly smaller than the cavity, this results in small gaps between the bare die and the border of the cavity. The long-term goal is to perform contacting of the chip via inkjet printed conductive paths. Therefore, it is important to fill the gap with isolating material first. The vision system presented in this work addresses this gap-filling task.

Work presented in this paper has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958472 (project TINKER) and from the project ZeRoboJet (Austrian Institute of Technology).

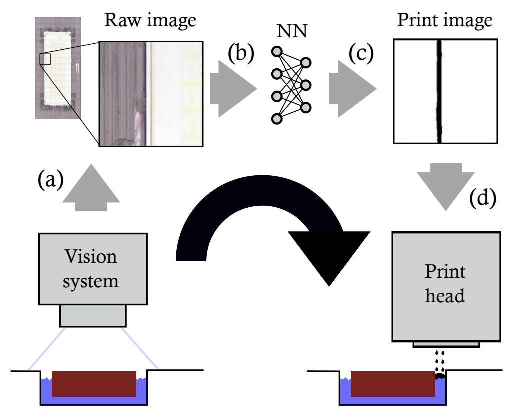

Fig. 1. Overview about the adaptive gap-filling process: Image acquisition (a), gap detection via neural network (b, c), and printing of calculated mask image (d).

Figure 1 illustrates the complete process of inspection and adaptive gap-filling. An image from top view is acquired via inline inspection. The image is fed to a neural network that outputs a segmentation of the gap around the bare die. The output image is then used as a mask that is inkjet-printed with isolating material to fill the gap. In this work, we focus on a method for mask calculation based on deep neural networks.

II. Related Work

Automated optical inspection is commonly used in industry for quality control of electronic devices. It is typically implemented with a system consisting of one or more machine vision cameras and several light sources. The systems mostly rely on reference comparison approaches. The recorded data is fed to a specifically trained algorithm that can detect flaws or defects on the surface of the test object by comparing it to its ideal form [1]. Extensive research has been done in applying such systems for detecting different errors on PCBs [2].

Vision systems have also been used to check bare dies using reference comparison algorithms [3]. Furthermore, a framework for the inspection of the attachment of the bare die to the PCB in regards to excess of insufficient glue was proposed [4]. In this work, pixel-based vectors for all regions of interest are created. Different machine learning algorithms for classification are investigated in order to classify regions as faulty if there was excess or insufficient glue. The best results were achieved using support vector machines [4]. In general, support vector machines have shown great potential for industrial applications in regards to classification [5], [6]. However, in many image recognition tasks it has been shown that deep learning approaches are more powerful, as they do not need as much data pre-processing and are able to generalise better [7]. Implementation of these approaches is mostly prevented by the enormous amount of data necessary to train these networks.

To bypass this problem [8] suggest using transfer learning and fine-tuning to adapt the deep learning network to the respective circumstances. They used a source network trained on the ImageNet 2012 dataset and performed transfer learning and one epoch of fine-tuning on an industrial optical inspection dataset provided by DAGM They receive an outstanding classification accuracy of 99.95%. The aim was to detect, whether the sample had defects or not and to classify the samples to one of six patterns resulting in a total of 12 classes. With 6,900 labeled images, the DAGM dataset still consists of a large amount of data which can be difficult to obtain in an industrial environment. However, what this study shows is that a network does not have to be trained from scratch on the target data in order to perform well. This can be used as an advantage when developing machine learning tools for production.

Inkjet printing is an additive manufacturing technique suitable for manufacturing in the micro- and nano-scale. Additive manufacturing has gained on popularity in recent years due to advantages like waste reduction and the large variety of applicable materials. This enables the application of additive manufacturing in many different sectors like the aerospace industry, the biomedical sector, or the electronics industry [9].

Inkjet technology is used to accurately generate free-flying fluid droplets and deposit them in precise locations on a substrate. While being the leading printer technology for graphical printing in the home and small office markets, inkjet technology has also shown increasing potential in the field of micro-scale manufacturing [10]. Several features make inkjet printing well suited for manufacturing, especially in the field of printed electronics. Inkjet printing is a digital process and can therefore be easily adjusted in real time. This allows each product in a sequence to be made differently, even to just adapt to misalignment of the product. Its digital nature also has a cost advantage compared to methods that require a physical mask or template. For inkjet printing the pattern to be printed is completely digital [11].

Another advantage is that inkjet printing is a non-contact method which enables accurate processing of fragile or non-planar substrates. Furthermore, inkjet printing enables the use of a wide range of materials, spanning from metals, ceramics, and polymers to biological materials like living cells. The main aspect is, that the substance must be fluid. Finally, inkjet printing is modular and scalable. Mutliple print-heads can be used simultaneously and their position varied. Two print-heads can be placed side-by-side resulting in a wider pattern or in a row in order to layer different materials. The latter is known from graphical inkjet printers for the home and small office sector, where often up to four colors are used [11].

There are several possible process routes that can be taken to employ inkjet printing for manufacturing. The most common technique is direct material printing. The droplets for the inkjet printing are generated by the flow of the liquid ink through a nozzle with a small opening and is formed by surface tension. The drop generation method ”drop-on-demand” where the ink is emitted through short jets and only forms drops when required, can achieve drop diameters lying between 10 and 100µm [11].

One of the main limitations of inkjet printing is that the feasible resolution depends on both the size of the final printed drop after its solidification, drying or curing, and the precise placement on the substrate. The latter is influenced by the movement accuracy of the print head or the substrate. The falling movement of the droplets also plays a role here, which is influenced by the aerodynamic and electrostatic properties of the droplet. These properties in turn depend on other factors. This limits the smallest feasible feature size to approximately 10µm when using direct material deposition [11].

In order to deploy machine vision methods for online-monitoring and adaptation of inkjet printing processes, the use of deep neural network is promising. The machine vision problem to be solved for gap detection essentially is segmantic segmentation. While neural networks for pure classification tasks rely on fully-connected layers [12], networks for segmentation rely on convolutional layers. Convolutional networks have proven to show better results for segmentation [13]. The key factor in the development of networks for semantic segmentation is to find ways to retain the spatial information of the input. Besides fully connected networks, there are approaches for semantic segmentation that are either region-based or weakly supervised semantic segmentation [14]. However, for this work, only fully convolutional networks are considered.

Chen et al. [15] propose using atrous convolutions to recover spatial information of the input throughout the network. Similarly like pooling layers, they are used to reduce the number of parameters that need to be learned. However, the degree of downsampling with atrous convolutional layers is much smaller than with pooling layers enabling the network to retain more spatial information.

III. Data Acquisition and Preparation

The images used in this work are extracted from the image shown on top in figure 2. The image shows a bare die embedded into a surrounding substrate made of polycarbonate. The image was created with a Keyence VHX-5000 digital microscope. It is a composition of a number of individual high-resolution images covering smaller sections of the sample. The full image has a size of 16746x7353 pixels. A single pixel within the image covers approximately 1.1µm of the sample.

The actual gap width is in the range of 10µm to 20µm. Figure 3 shows a sketch of how the bare die is inserted into the cavity of a substrate filled with glue substrate and indicates the gap between the bare die and the substrate.

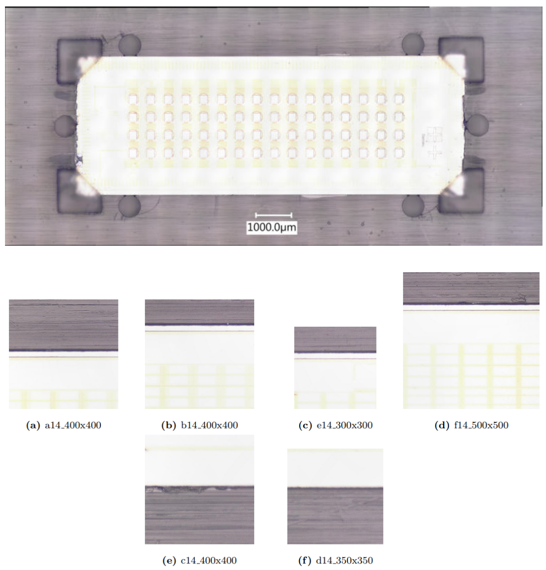

Fig. 2. Input images.

For inline inspection, processed images are on one hand required to sufficiently capture the gap. On the other hand, the image size should not be excessively large in order to enable reasonable image processing.

Pretrained neural networks often require a certain input size. Common input sizes are 224x224 [16] or 256x256 pixels [17]. The input data for the neural network training are images of a component consisting of a bare die and the substrate material of the PCB. In the planned production line, these images will be captured, and immediately fed to the tool without any further preprocessing besides those steps integrated into the tool.

As mentioned before, feeding the whole image to a deep neural network would end in an extensive demand on computational power and memory on the one hand and clash with size restraints of pretrained models on the other hand. Therefore, the main image is cropped into manageable sub-images. These sub-images are then used for model training and evaluation.

In total, six sequences of sub-images are extracted, each covering a horizontal section of the main image. Four sequences (a, b, e, and f) are taken from the top gap and two (c and d) from the bottom one. The image extraction is automated. For each sequence the position of the first sub-image within the main image is defined by selecting the coordinates of its upper left corner and adding the desired sub-image width. The corresponding horizontal section is then split into uniformly sized square images by moving the filter to the right by one sub-image width as illustrated in figure 4. The gap along the complete image width is captured by a set of squared sub-images.

To enhance the dataset, the position of the first sub-image and/or the image size of the crop are varied for each of the sequences. The basic crop size is 424 x 424 pixels (sequence a, b, c). The crops of sequence d are of size 371 x 371 pixels, the crops of sequence e are 318 x 318 pixels and the cops of sequence f have a size of 530 x 530 pixels. Additionally, the horizontal starting points are slightly different as can be seen by comparing the crops of sequence a and b. The odd values of the crop sizes were chosen such that the corresponding dimensions in the actual sample are even values. As preparation for further processing steps after the gap segmentation, the latter are recorded in the naming of the crops together with the name of the sequence. Sequence b is shifted by 224 pixels to the right compared to sequence a.

As neural network training is done in a supervised fashion, ground truth masks need to be provided that represent the gap regions. The ground truth is manually defined for the complete high-resolution image. In total, 128 images are acquired. They are divided in a training set consisting of 82 images, a validation set with 20 and a test set with 26 images.

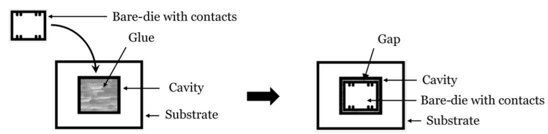

Fig. 3. Illustration of bare die insertion into cavity of PCB board.

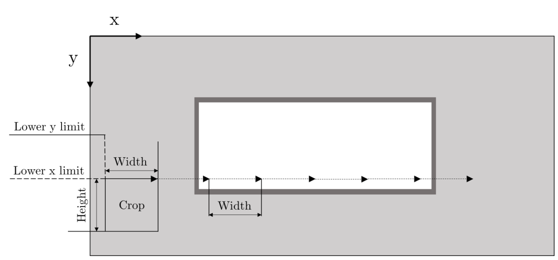

Fig. 4. The complete high-resolution image is split into multiple sub-images along gaps.

IV. Model Training

As outlined above, the task of gap-detection is equivalent to image segmentation. The input is an image acquired by the vision system. Expected output is a mask image that separates gap from background. Three fully convolutional networks were trained to accomplish the segmentation task:

1) **U-Net**: The first network architecture used is a U-Net developed and pretrained for abnormality segmentation in brain MRI by [17]. The U-Net was first introduced by Ronneberger et al. [18] and revolutionised the approach to semantic segmentation. The network is a fully convolutional network, initially developed for biomedical image segmentation. It consists of an encoder path that captures the context of the image and a symmetric

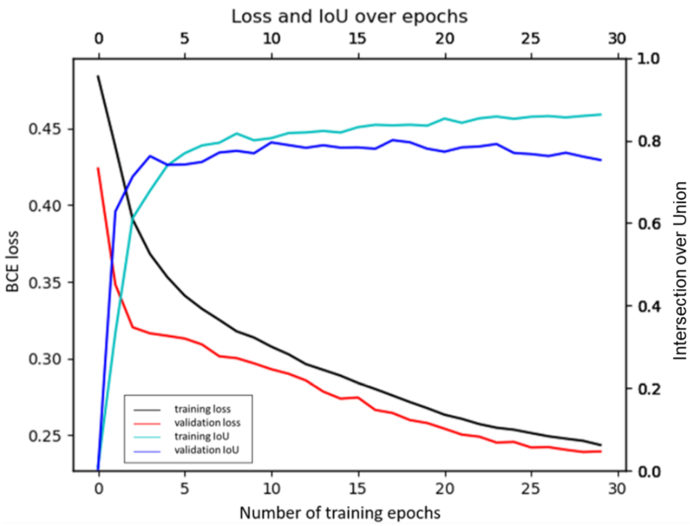

Fig. 5. Training progress for the U-Net:

decoder path that enables precise localisation. High-resolution features of the encoder blocks are fed to the horizontally corresponding decoder blocks where they are combined with the upsampled output.

2) **ResNet50 with fully connected decoder**: The second network architecture, is a ResNet50 [12] where the fully connected layers of the classifier are exchanged with convolutional layers as decoder as proposed by [13]. In addition to the convolutional decoder, the skip connections of the ResNet architecture help retain spatial information necessary for the segmentation task.
3) **ResNet50 with DeepLabV3**: The third network architecture is a ResNet50 with the DeepLabV3 head as decoder which was proposed by [15]. The DeepLabV3 utilizes atrous convolutions to retain the spatial information of the input.

Both ResNet based models are pretrained on the COCO dataset [19]. The models are trained using the adaptive moment estimation (Adam) algorithm [20]. As a loss function binary cross entropy was selected. The models are trained on 82 sub-images that are resized to 256x256. Data augmentation is applied by randomly rotating the training samples. The networks are trained over 30 epochs using two different learning rates, $1.0e^{-4}$ and $1.0e^{-5}$. After every training epoch, the models are tested on a validation data set consisting of 18 images. The model weights that performed best on the validation data are then used on the test data for a final model assessment.

V. Results

The success of the segmentation task is determined by the accuracy with which the gap is predicted based on the input image. For our evaluations, the success of the segmentation task is determined by the accuracy with which the gap is predicted based on the input image. To evaluate the prediction accuracy, every pixel value of the forecast image is compared to the pixel values in a ground truth mask that was prepared manually for the corresponding input image. Since the gap makes up a comparatively small part of the image, intersection over union (IoU) is chosen as evaluation measure.

Figure 5 shows the decrease of loss and increase in IoU over the progress of model training for the U-Net at a learning rate of $1.0e^{-5}$. While the training IoU keeps increasing, the validation IoU slowly starts to decrease in the last epochs, suggesting slight overfitting. The validation IoU reaches a maximum level of approximately 80% throughout training, but scores an even higher IoU of 82.45% on the test set. This corresponds well to values reported in similar applications. [21] perform segmentation of cracks in railway infrastructures and achieve an IoU of 81% and [22] obtain an IoU of 83% in a publicly available dataset on images with cracks in concrete. While the IoU of the U-Net trained with a higher learning rate is satisfying as well, the IoU level of the two other models is much lower with an IoU of 60.31% for the next best model, the ResNet50 with DeepLabV3 head.

Since the present task is more comparable to the original task the U-Net was developed for, it is logical that the U-Net performs better.

TABLE I
RESULTS OF MODEL TRAINING

Network architecture	Learning rate	IoU
U-Net	$1.0e^{-4}$	0.7826
U-Net	$1.0e^{-5}$	**0.8245**
ResNet50 + FCN	$1.0e^{-4}$	0.4386
ResNet50 + FCN	$1.0e^{-5}$	0.5187
ResNet50 + DeepLabV3	$1.0e^{-4}$	0.6031
ResNet50 + DeepLabV3	$1.0e^{-5}$	0.3096

Figure 6 illustrates a set of input images with the prediction results and the corresponding ground truth masks. The left four columns are the results from the U-Net, the right four columns the results of the ResNet50 with DeepLabV3 head on the same input images. The U-Net predicts the gaps with much higher confidence. In addition, the prediction seems to be sensible to irregularities on the gap border to the substrate that are not identified in the ground truth. The border to the bare die, however, is predicted quite accurately. From neural network output, a print image can be generated for use in the subsequent inkjet printing process.

VI. Conclusions and Future Work

We introduce a method for optical inspection and adaptive calculation of print images for printed electronics. The calculation of print images is accomplished via deployment of a deep neural network. For this, three deep neural networks are trained with transfer learning and fine-tuning on 82 images extracted from a sample consisting of a bare die placed into the cavity of a substrate. The images each contain sections of the gap located between the substrate and the bare die. A comparison of the results shows that the shallowest of the three architectures performs best on the analysed problem. The deep

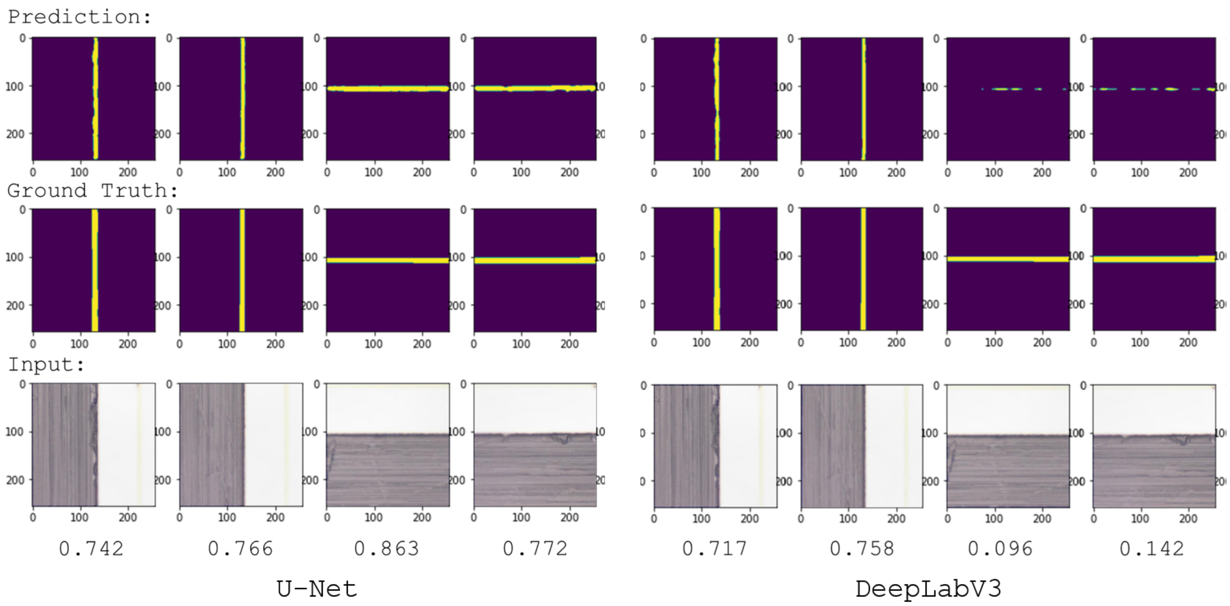

Fig. 6. Prediction results on four test images. The top row shows the prediction, the second row the ground truth and the bottom row the input image. Under each column the corresponding IoU is noted. The left four columns, show the results of the U-Net, the right four columns are the corresponding results of the ResNet50 model with DeepLabV3 head.

neural network is capable of detecting irregularities on the gap border. The use case investigated here, relates to filling of a gap after the placement of the bare die into the cavity on a PCB. Similar methods seem promising for other process steps like contacting of the bare die or layer-by-layer printing. While the practical implementation of this use case still needs to be examined, optical inspection together with inkjet printing is a strong combination to enable zero-defect manufacturing for printed electronics.

To improve the performance of the deep neural network on the analyzed problem, overfitting has to be prevented. For this, we plan to improve its generalization ability by introducing the deep neural network to additional data samples. As with many applications in industry, acquisition of large amounts of data is difficult. Hence, the use of artificial data (e. g. using Domain Randomization or Domain Adaptation) seems promising. Furthermore, there are other applications that are very similar to the one outlined in this paper. For example, contacting of bare dies via inkjet-printing would be an interesting use case that we plan to cover with our vision system.

References

[1] J. Jiang, J. Cheng, and D. Tao, "Color biological features-based solder paste defects detection and classification on printed circuit boards," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 2, no. 9, pp. 1536–1544, 2012.

[2] V. Chaudhary, I. R. Dave, and K. P. Upla, "Automatic visual inspection of printed circuit board for defect detection and classification," in *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*. IEEE, 2017, pp. 732–737.

[3] C. Weng and T. Saeger, "Combining vision inspection and bare die packaging for high volume manufacturing," in *International Conference on Compound Semiconductor Manufacturing Technology*, 2013.

[4] T. Vafeiadis, N. Dimitriou, D. Ioannidis, T. Wotherspoon, G. Tinker, and D. Tzovaras, "A framework for inspection of dies attachment on pcb utilizing machine learning techniques," *Journal of Management Analytics*, vol. 5, no. 2, pp. 81–94, 2018.

[5] L. M. R. Baccarini, V. V. R. e Silva, B. R. De Menezes, and W. M. Caminhas, "SVM practical industrial application for mechanical faults diagnostic," *Expert Systems with Applications*, vol. 38, no. 6, pp. 6980–6984, 2011.

[6] G. Duan, H. Wang, Z. Liu, and Y.-W. Chen, "A machine learning-based framework for automatic visual inspection of microdrill bits in pcb production," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 6, pp. 1679–1689, 2012.

[7] J. Masci, U. Meier, D. Ciresan, J. Schmidhuber, and G. Fricout, "Steel defect classification with max-pooling convolutional neural networks," in *The 2012 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2012, pp. 1–6.

[8] S. Kim, W. Kim, Y.-K. Noh, and F. C. Park, "Transfer learning for automated optical inspection," in *2017 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2017, pp. 2517–2524.

[9] T. D. Ngo, A. Kashani, G. Imbalzano, K. T. Nguyen, and D. Hui, "Additive manufacturing (3d printing): A review of materials, methods, applications and challenges," *Composites Part B: Engineering*, vol. 143, pp. 172 – 196, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1359836817342944

[10] P. J. Smith and D. H. Shin, *Inkjet-based micromanufacturing*. John Wiley & Sons, 2012.

[11] I. M. Hutchings and G. D. Martin, *Inkjet technology for digital fabrication*. John Wiley & Sons, 2012.

[12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[13] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[14] Y. Guo, Y. Liu, T. Georgiou, and M. S. Lew, "A review of semantic segmentation using deep neural networks," *International journal of multimedia information retrieval*, vol. 7, no. 2, pp. 87–93, 2018.

[15] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," *arXiv preprint arXiv:1706.05587*, 2017.

[16] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556*, 2014.

[17] M. Buda, A. Saha, and M. A. Mazurowski, “Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm,” *Computers in Biology and Medicine*, vol. 109, 2019.

[18] 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*. Springer, 2015, pp. 234–241.

[19] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, “Microsoft coco: Common objects in context,” in *European conference on computer vision*. Springer, 2014, pp. 740–755.

[20] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.

[21] J. S. Lee, S. H. Hwang, I. Y. Choi, and Y. Choi, “Estimation of crack width based on shape-sensitive kernels and semantic segmentation,” *Structural Control and Health Monitoring*, vol. 27, no. 4, 2020.

[22] Y. Pan, G. Zhang, and L. Zhang, “A spatial-channel hierarchical deep learning network for pixel-level automated crack detection,” *Automation in Construction*, vol. 119, 2020.