

Deteksi Kerusakan Permukaan Material Industri dan Permukaan Material Komposit Polimer Berdasarkan Gambar Mikrostruktur dengan Metode *Mask-Region based Convolutional Neural Network* (Mask R-CNN)

Fajar Astuti Hermawati¹, Dame Jeremia Hutapea², dan I Made Kastiawan³

^{1 2} Teknik Informatika, Universitas 17 Agustus 1945 Surabaya

³ Teknik Mesin, Universitas 17 Agustus 1945 Surabaya

Email Correspondence Author: fajarastuti@untag-sby.ac.id

Abstract — In the industrial world, there is a duty to ensure the quality of industrial materials is free from damage, especially on the product's surface. Often to carry out material surface quality control is done manually by humans who are experts in identifying complex material surface damage. To carry out the identification takes time and is not efficient. This research is aimed at detecting material damage on the surface of industrial products. With computer vision technology, it is possible to classify material damage on industrial products such as cracked, scratched, and crushed material surfaces. The dataset used is 399 data which contains 52 data with defects and 347 data without defects. Damaged areas were detected and segmented using the Mask-Region-based Convolutional Neural Network (Mask R-CNN) method. This method consists of two main steps: detecting the damaged area using the Faster R-CNN method and the segmentation stage of the area using a deep convolutional encoder-decoder Segnet network. From the segmentation results obtained, we can measure the percentage of damage area and the distribution of damage on the material's surface to determine its strength.

Keyword — Industrial Materials, Mask R-CNN, Polymer Composites, Segmentation, Segnet.

Abstrak — Dalam dunia industri, terdapat tugas untuk memastikan kualitas dari material industri bebas dari kerusakan terutama pada permukaan produk. Seringkali untuk melakukan kontrol kualitas permukaan material dilakukan secara manual oleh manusia yang ahli dibidangnya untuk mengidentifikasi kerusakan permukaan material yang kompleks. Untuk melakukan identifikasi tersebut memerlukan waktu dan tidak efisien. Penelitian ini ditujukan untuk melakukan deteksi kerusakan material pada permukaan produk industri. Dengan teknologi *computer vision* dapat dilakukan klasifikasi kerusakan material pada permukaan produk industri seperti permukaan material yang retak, tergores dan tertimpa benda. Dataset yang digunakan sebanyak 399 data yang berisikan data dengan *defect* sebanyak 52 data dan 347 data tanpa *defect*. Area yang mengalami kerusakan dideteksi dan disegmentasi dengan metode *mask-region based convolutional neural network* (mask r-cnn). Metode ini terdiri dari dua langkah utama, yaitu deteksi area kerusakan menggunakan metode *faster r-cnn* dan tahap segmentasi area tersebut menggunakan jaringan *segnet deep convolutional encoder – decoder*. Dari hasil segmentasi yang diperoleh dapat diukur prosentase luasan kerusakan serta sebaran kerusakan pada permukaan material untuk menentukan kekuatannya.

Kata kunci : Mask R-CNN, Material Industri, Komposit Polimer, Segmentasi, Segnet.

I. PENDAHULUAN

Perkembangan teknologi komputer pada *machine vision* telah banyak digunakan terutama untuk mendeteksi kerusakan material pada produk industri. Sejumlah besar algoritma untuk melakukan deteksi kerusakan permukaan telah muncul. Algoritma ini diklasifikasikan dengan metode manual, serta dengan metode dengan *machine learning*. Metode berbasis *machine learning* menggabungkan fitur statistik dengan terutama untuk mengekstrak fitur statistik dari permukaan kerusakan, dan kemudian menggunakan algoritma *machine learning* untuk mempelajari fitur agar mendapatkan deteksi kerusakan permukaan [1].

Kerusakan permukaan material pada dunia industri merupakan bagian penting dari produksi industri dan berdampak buruk pada kualitas dan fungsionalitas terhadap material industri. Kerusakan permukaan dengan pendeteksian manual memerlukan waktu yang lama dan hasil yang didapatkan tingkat keakuratan yang dihasilkan rendah [2]. Dengan menggunakan teknologi dari *Artificial Intelligence & Visual Computing* secara sistematis dapat melakukan penyelesaian dengan berbasis *machine vision*.

Permasalahan dari pendeteksian kerusakan pada material industri dapat diatasi dengan metode *Mask Region based Convolutional Neural Network* dan segmentasi dengan *segnet deep convolutional encoder – decoder* dan dari gambar – gambar material industri yang diambil dengan menggunakan mikroskop. Dengan adanya gambar – gambar material industri maka dapat dilakukan pengklasifikasian untuk mengetahui jenis kerusakan pada material industri serta mendapatkan tingkat akurasi (*confidence level*) pada kerusakan material tersebut.

II. DASAR TEORI

A. Convolutional Neural Network

Neural Network dapat diterapkan pada tugas-tugas *computer vision*, untuk mendapatkan kinerja generalisasi yang baik, ada baiknya untuk memasukkan pengetahuan sebelumnya ke dalam arsitektur jaringan. *Convolutional Neural Network* bertujuan untuk menggunakan informasi

spasial antara piksel gambar [3]. Convolutional Neural Network terdiri dari beberapa layer. Convolutional Layer merupakan bagian yang melakukan operasi konvolusi yaitu mengkombinasikan linier filter terhadap daerah lokal. Layer ini yang pertama kali menerima gambar yang diinputkan pada arsitektur. Bentuk layer ini adalah sebuah filter dengan panjang (pixel), tinggi (pixel), dan tebal sesuai dengan channel image data yang diinputkan. Ketiga filter ini akan bergeser keseluruhan bagian gambar. Pergeseran tersebut akan melakukan operasi “dot” antara input dan nilai dari filter tersebut sehingga akan menghasilkan output yang disebut sebagai activation map atau feature map [4]. Pooling Layer atau subsampling adalah pengurangan ukuran matriks. Terdapat dua macam pooling yang sering digunakan yaitu average pooling dan max pooling [5]. Fully Connected Layer digunakan sebagai beberapa lapisan terakhir. Dengan menghilangkan batasan dapat meringkas informasi yang disampaikan oleh lapisan tingkat bawah dengan lebih baik dalam melihat keputusan akhir.

B. Mask Region based Convolutional Neural Network (R-CNN)

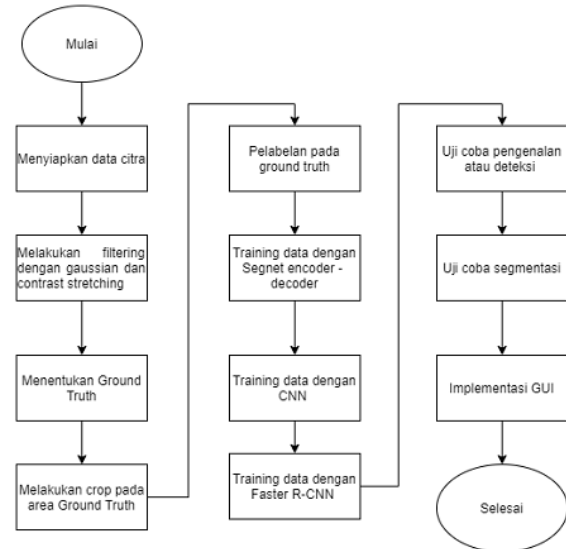
Mask R-CNN secara konseptual sederhana: Faster R-CNN memiliki dua output untuk setiap kandidat objek, label kelas dan offset kotak pembatas; untuk ini menambahkan cabang ketiga yang menampilkan objek mask. Mask R-CNN dengan demikian merupakan ide yang natural dan intuitif. Tetapi mask output tambahan berbeda dari output kelas dan kotak, yang membutuhkan ekstraksi tata letak spasial suatu objek yang jauh lebih halus. Selanjutnya, memperkenalkan elemen-elemen kunci dari Mask R-CNN, termasuk penyelarasan piksel-ke-piksel, yang merupakan bagian utama dari Fast / Faster R-CNN yang hilang [6].

Metode Faster R-CNN terdiri dari dua tahap yaitu mengekstrak dan melatih menggunakan Region Proposal Network (RPN) dan mengklasifikasikan fitur yang diperoleh. Dalam proses pelatihan terdapat empat tahap yaitu: a) Melatih jaringan R-CNN, b) Melatih jaringan faster RCNN dengan menggunakan RPN, c) Melatih ulang RPN menggunakan faster R-CNN, d) Melatih ulang faster R-CNN dengan RPN yang diperbarui [7].

III. METODOLOGI PENELITIAN

Penelitian ini mengimplementasikan metode Mask Region based Convolutional Neural Network (Mask R-CNN) dan menerapkan kategori kerusakan yaitu defect dan non defect. Tahapan penelitian deteksi kerusakan permukaan ini terdiri dari beberapa tahapan seperti pada Gambar 1. Tahap pertama adalah menyiapkan data citra yang akan diolah untuk melakukan segmentasi dan data training. Kemudian melakukan filtering dengan Gaussian dan contrast stretching untuk mengurangi noise dan meningkatkan kualitas citra agar lebih mudah saat dilakukan untuk pendeteksian maupun segmentasi. Sebelum dilakukan proses training, perlu ditentukan ground truth dengan menentukan lokasi pada data gambar yang terdapat defect dan void. Setelah itu daerah yang ditandai, dipotong

(*cropping*) sesuai ukuran daerah tersebut menjadi kecil untuk mempermudah dalam tahap training data. Tahap pelabelan pada *ground truth* dibutuhkan untuk melakukan training data dengan *segnet encoder – decoder* untuk mendapatkan hasil segmentasi dari data gambar yang telah diolah.



Gambar 1. Tahapan Penelitian Deteksi Kerusakan Pada Permukaan Material

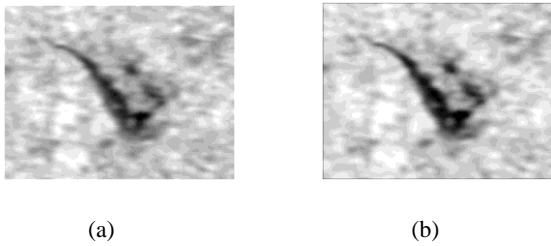
Data yang telah di crop pada area tertentu dari *Ground Truth* setelah itu akan digunakan kembali untuk ditraining menggunakan segnet encoder dan decoder untuk mendapatkan hasil segmentasi. Pada segnet dilakukan training data dengan menggunakan model pre-trained yaitu vgg16. Vgg16 adalah model convolutional neural network yang memiliki 16 kedalaman layer dengan parameter input size 224 x 224 x 3 dimana 3 adalah channel warna RGB.

A. Objek Penelitian

Objek penelitian ini adalah permukaan material industri yaitu material dari produk setengah jadi seperti komutator listrik. Pada permukaan komutator listrik diamati apakah terdapat defect atau tidak. Dataset berasal dari KolektorSDD (*Surface Defect Dataset*) yang berisikan 399 gambar dengan 52 gambar yang terdapat defect dan 347 gambar tanpa defect [8]. Material juga menggunakan komposit polimer dengan serat sisal sebagai penguatnya. Polimer dicetak menjadi cetakan persegi panjang, tebal 10 mm, lebar 20 mm dan panjang 200 mm dengan tekanan 20 kg / mm². Gambar mikrostruktur permukaan polimer ditangkap menggunakan kamera digital [9], [10].

B. Preprocessing

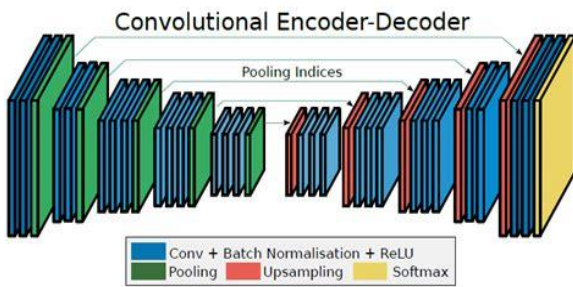
Tahap preprocessing bertujuan untuk menghilangkan noise pada gambar mikrostruktur. Kami menerapkan metode Gaussian filter 3x3 dengan standar derivasi 1.2 dan contrast stretching filter. Gambar 2(a) dan 2(b) menunjukkan hasil citra sebelum dan sesudah preprocessing.



Gambar 2. Gambar mikrostruktur: (a) sebelum dilakukan preprocessing dan (b) sesudah preprocessing

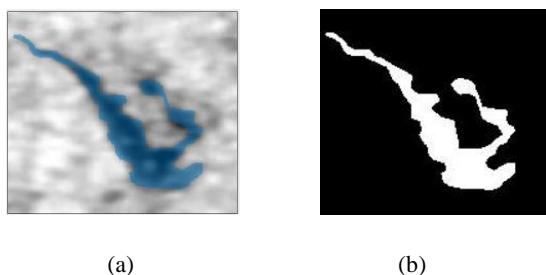
C. Segmentasi Segnet Deep Convolutional Encoder – Decoder

Segmentasi Segnet adalah arsitektur untuk semantik segmentasi yang efisien [11]. Segmentasi Segnet menggunakan piksel per piksel pada citra seperti menentukan kelas objek dan menentukan piksel kelas background. Segnet memiliki jaringan encoder dan decoder, setiap lapisan encoder dan decoder memiliki 13 lapisan seperti pada Gambar 3.



Gambar 3. Arsitektur segnet encoder – decoder [11]

Jaringan encoder dirancang untuk klasifikasi objek. Jaringan segmentasi segnet yang digunakan adalah jaringan dari Visual Geometry Group 16 (VGG16). Pada lapisan terakhir VGG16 pixelClassificationLayers digantikan dengan soft-max classifier untuk menghasilkan probabilitas kelas untuk setiap piksel. Tahap segmentasi ini digunakan untuk memberikan label pada daerah yang piksel yang terdeteksi kerusakan seperti Gambar 4(a) dan 4(b).



Gambar 4. (a) Gambar yang telah dilabeli dan (b) bentuk biner dari label

D. Convolutional Neural Network (CNN) Arsitektur

Convolutional Neural Network menggunakan jaringan yang telah di latih dahulu (pre-trained) yaitu alexnet.

AlexNet juga jaringan dari Convolutional Neural Network dengan kedalaman 8 layer yaitu 5 convolutional layer dan 3 fully connected layer dengan parameter input size 227 x 227 x 3. Alexnet dapat mengklasifikasikan gambar dalam 1000 kategori objek, seperti mouse, botol dan banyak objek lainnya. Keseluruhan arsitektur CNN alexnet dapat dilihat pada Tabel 1.

TABEL I
ARSITEKTUR JARINGAN PRETRAINED CNN ALEXNET

Layer	Size	Kernel Size
Input	Image	227 x 227 x 3
1	Convolution	55 x 55 x 96
	Max Pooling	27 x 27 x 96
2	Convolution	27 x 27 x 256
	Max Pooling	13 x 13 x 256
3	Convolution	13 x 13 x 384
4	Convolution	13 x 13 x 384
5	Convolution	13 x 13 x 256
	Max Pooling	6 x 6 x 256
6	Fully Connected	9216
7	Fully Connected	4096
8	Fully Connected	4096
Output	Fully Connected	1000

E. Faster R-CNN

Pada tahap ini bertujuan untuk melakukan deteksi dari object yang telah di training pada tahap CNN dengan model jaringan alexnet. Faster R-CNN menggunakan Fast R-CNN dan Region Proposal Network (RPN) sebagai arsitektur utamanya dan mengklasifikasikan berdasarkan fitur yang diperoleh. RPN menghasilkan beberapa bounding box dimana setiap box mempunyai lokasi terdapat object atau non object. Lokasi objek dari wilayah yang diketahui disebut Region of Interest (RoI). Setelah mendapatkan RoI selanjutnya di identifikasi pada RPN apakah hasil RoI tersebut sesuai dengan kelasnya atau tidak dengan classifier dari CNN.

Metode Faster R-CNN terdiri dari dua tahap yaitu mengekstrak dan melatih menggunakan Region Proposal Network (RPN) dan mengklasifikasikan fitur yang diperoleh. Dalam proses pelatihan terdapat empat tahap yaitu: a) Melatih jaringan R-CNN, b) Melatih jaringan faster R-CNN dengan menggunakan RPN, c) Melatih ulang RPN menggunakan faster R-CNN, d) Melatih ulang faster R-CNN dengan RPN yang diperbarui [7].

F. Perhitungan Performa

Pada percobaan hasil segmentasi diukur dengan menggunakan Precision, Sensitivity, Specificity, dan Dice Similarity. Dice Similarity (D) adalah ukuran kesamaan yang digunakan untuk mengukur tumpang tindih antara dua wilayah secara spasial. Sensitivity atau biasa disebut dengan True Positive Ratio (TPR) dan specificity disebut dengan True Negative Ratio (TNR). Hasil akurasi diperoleh dari

rata – rata nilai True Positive (TP) dan True Negative (TN). TP sebagai rasio perpotongan antara O_{SR}^M dan O_{GT} region yang di cakup oleh dan TN adalah rasio region yang tidak tercakup oleh O_{SR}^M dan O_{GT} sampai diluar wilayah yang dirumuskan sebagai berikut [12], [13] :

$$D = \frac{2 | O_{SR}^M \cap O_{GT} |}{| O_{GT} | + | O_{SR}^M |} \quad (1)$$

Dimana O_{GT} adalah ground truth yang didapat dari segmentasi manual dan O_{SR}^M adalah hasil segmentasi.

IV. HASIL DAN PEMBAHASAN

Sistem ini dibangun dengan menggunakan software Matlab dan berjalan pada laptop dengan Processor i3-7020U3 dengan dilengkapi GPU Nvidia GeForce MX110.

A. Hasil Segmentasi dengan Segnet

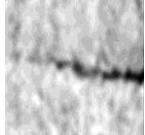
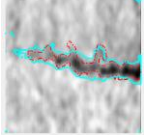
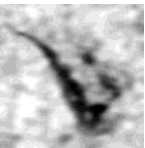
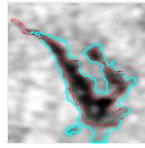
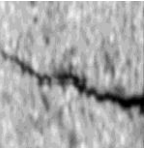
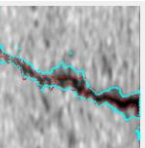

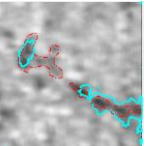
Pada tahap segmentasi kategori defect ini dilakukan training data dengan segnet deep convolutional encoder – decoder dan menggunakan model VGG16 untuk training. Tahap ini mengambil ukuran citra menjadi 224 x 224 pada 52 data yang terdapat defect. Dari hasil training data didapatkan akurasi sebesar 86,51 % dengan 50 epochs.

Tabel 2 menunjukkan beberapa contoh hasil segmentasi dengan menggunakan Segnet. Performa dari hasil segmentasi dihitung menggunakan *Precision*, *sensitivity (True Positive Ratio)*, *specificity (True Negative Ratio)* dan *Dice similarity*. Pada hasil segmentasi garis berwarna merah merupakan data ground truth dan garis berwarna biru muda adalah hasil segmentasi dengan menggunakan segnet.

Pada contoh pertama pada Tabel 2, hasil segmentasi mempunyai nilai *sensitivity* 0.9203 namun nilai *specificity*nya hanya 0.4678. Demikian juga untuk contoh gambar kedua, nilai *specificity*nya sangat kecil yaitu 0.2602 dengan nilai *sensitivity* yang besar yaitu 0.9488. Bahkan pada contoh gambar ketiga nilai *specificity*nya hanya sebesar 0.2602 dengan nilai *sensitivity* mencapai 0.9704. Hal ini dapat terjadi karena hasil segmentasi mempunyai area yang lebih kecil dan ada di dalam area *ground truth*. Sedangkan rata-rata nilai dari dice similarity mencapai 0.7 dan precision mencapai 0.6. Hal ini disebabkan bentuk obyek yang disegmentasi cenderung kasar dan mempunyai kurva tepi yang bergelombang atau tidak homogen. Untuk area defect yang terlalu kecil seperti pada contoh gambar keempat akan menghasilkan rata-rata performa yang hampir sama, dengan nilai *sensitivity* 0.5020, nilai *specificity* 0.5495.

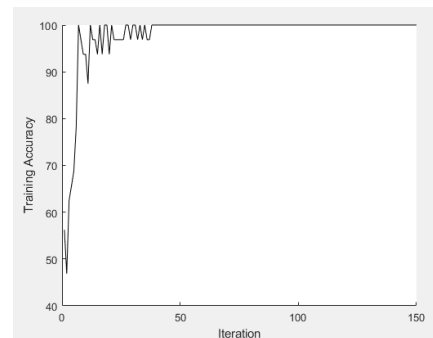
TABEL II

CONTOH HASIL SEGMENTASI DAN UKURAN PERFORMA

Gambar asli	Hasil segmentasi	Ukuran Performa
		P=0.6336 TPR=0.9203 TNR=0.4678 D=0.7505
		P=0,5619 TPR=0,9488 TNR=0,2602 D=0,7058
		P=0.5327 TPR=0.9704 TNR=0.1488 D=0.6878
		P=0.5270 TPR=0.5020 TNR=0.5495 D=0.5142

B. Hasil Deteksi Obyek dengan Faster RCNN

Pada tahap ini dibagi menjadi dua sub tahapan yaitu proses training dengan CNN yang digunakan untuk mendapatkan data sebelum masuk ke tahap Faster R-CNN dan mendapatkan hasil akurasi dari training sebesar 90.91 %. Grafik akurasi pada proses training dapat dilihat pada Gambar



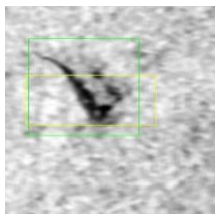
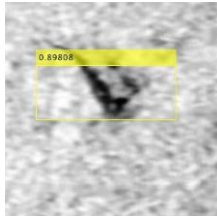
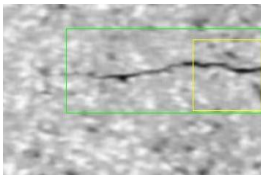

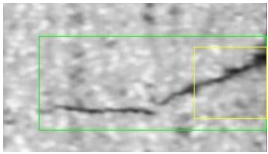
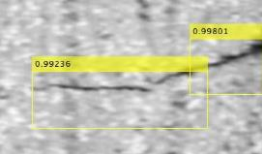
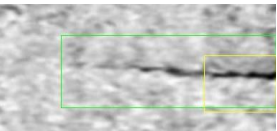

Gambar 5. Grafik akurasi pada proses training dengan CNN

Selanjutnya pada tahap Faster R-CNN terdiri dari 4 tahap: (1) Tahap pertama adalah melatih Region Proposal Network (RPN). Tahap ini akan menghasilkan beberapa bounding box untuk mendapatkan lokasi dari object atau non object (2) Tahap kedua adalah ekstraksi Region Proposal dari citra yang telah dilatih. Tahap ini menggunakan *Region of Interest (RoI) pool* dari setiap bounding box untuk melakukan klasifikasi (3) Tahap ketiga adalah melakukan training ulang dengan bobot baru dengan Fast R-CNN (4)

Tahap keempat adalah melakukan training ulang dengan Region Proposal Network yang telah diperbarui.

Tabel 3 menunjukkan hasil deteksi area defect menggunakan Faster R-CNN Pada gambar perbandingan garis hijau adalah letak dari ground truth dan garis kuning adalah hasil deteksi dengan Faster R-CNN. Confidence level merupakan tingkat kepercayaan tentang objek yang dideteksi.

TABEL III
CONTOH HASIL DETEKSI AREA DEFFECT DENGAN FASTER R-CNN

Hasil perbandingan deteksi dengan ground truth	Nilai confidence level pada pendeteksian
	
	
	
	

Hasil deteksi area pada gambar kedua sampai gambar keempat menghasilkan area yang mempunyai nilai confidence terbesar jauh lebih kecil dari area yang diinginkan. Hal ini bisa dilihat karena area yang dideteksi mempunyai bentuk memanjang dengan ketebalan area yang semakin kecil pada ujung kiri. Sehingga area yang dikenali cenderung pada bagian yang tebal di bagian kanan. Sedangkan pada contoh hasil pada gambar pertama, area yang dikenali lebih baik karena bentuknya tidak memanjang, meskipun pada bagian ujung area yang menipis tidak bisa dideteksi dengan baik.

V. KESIMPULAN

Dengan penelitian yang telah dilakukan, sistem ini dapat mendeteksi kerusakan yang ada pada gambar mikrostruktur dengan meningkatkan mutu kualitas citra untuk pendeteksian dan segmentasi sehingga meningkatkan efisiensi dalam mendeteksi kerusakan permukaan material industri pada gambar mikrostruktur. Sistem ini diperlukan pengembangan lebih lanjut dengan melakukan training dengan model yang lebih dalam seperti model pre-trained resnet50 dan resnet101 supaya mendapatkan hasil training yang lebih baik lagi.

DAFTAR ACUAN

- [1] F. Zhou, G. Liu, F. Xu, and H. Deng, "A generic automated surface defect detection based on a bilinear model," *Applied Sciences (Switzerland)*, vol. 9, no. 15, pp. 1–17, 2019.
- [2] X. Tao, D. Zhang, W. Ma, X. Liu, and De Xu, "Automatic metallic surface defect detection and recognition with convolutional neural networks," *Applied Sciences (Switzerland)*, vol. 8, no. 9, pp. 1–15, 2018.
- [3] J. Koushik, "Understanding Convolutional Neural Networks," in *29th Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, Spain*, 2016, no. 3, pp. 1–6.
- [4] R. D. Nurfitra and G. Ariyanto, "Implementasi Deep Learning Berbasis Tensorflow Untuk Pengenalan Sidik Jari," *Emitor: Jurnal Teknik Elektro*, vol. 18, no. 01, pp. 22–27, 2018.
- [5] M. B. Bejiga, A. Zeggada, A. Nouffidj, and F. Melgani, "A convolutional neural network approach for assisting avalanche search and rescue operations with UAV imagery," *Remote Sensing*, vol. 9, no. 2, pp. 693–696, 2017.
- [6] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 386–397, 2020.
- [7] F. A. Hermawati, H. Tjandrasa, and N. Suciati, "Combination of Aggregated Channel Features (ACF) Detector and Faster R-CNN to Improve Object Detection Performance in Fetal Ultrasound Images," *International Journal of Intelligent Engineering & System*, vol. 11, no. 6, pp. 65–74, 2018.
- [8] D. Tabernik, S. Šela, J. Skvarč, and D. Skočaj, "Segmentation-based deep-learning approach for surface-defect detection," *Journal of Intelligent Manufacturing*, vol. 31, no. 3, pp. 759–776, 2020.
- [9] F. A. Hermawati, I. M. Kastawan, and Muhyin, "Digital Microscopy Image Enhancement Technique for Microstructure Image Analysis of Bottom Ash Particle Polymer Composites," in *Parinov I., Chang SH., Long B. (eds) Advanced Materials*, Springer Proceedings in Materials, vol. 6. Springer, Cham., 2020, pp. 235–244.
- [10] F. A. Hermawati, I. M. Kastawan, and Muhyin, "Automatic Filler Dispersion Quantification in Microstructure Images of Bottom Ash Reinforced Polymer Composite," *Journal of Hunan University (Natural Sciences)*, vol. 48, no. 2, pp. 36–47, 2021.
- [11] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder

- Architecture for Image Segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.
- [12] F. A. Hermawati, H. Tjandrasa, and N. Suciati, “hybrid speckle noise reduction method for abdominal circumference segmentation of fetal ultrasound images,” *International Journal of Electrical and Computer Engineering*, vol. 8, no. 3, 2018.
- [13] F. A. Hermawati, H. Tjandrasa, and N. Suciati, “Phase-based thresholding schemes for segmentation of fetal thigh cross-sectional region in ultrasound images,” *Journal of King Saud University - Computer and Information Sciences*, no. xxxx, 2021.