Article

Research on Automated Defect Classification Based on Visual Sensing and Convolutional Neural Network-Support Vector Machine for GTA-Assisted Droplet Deposition Manufacturing Process

Chen Ma ¹, Haifei Dang ², Jun Du ¹, Pengfei He ¹, Minbo Jiang ¹ and Zhengying Wei ^{1,*}

¹ State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, China; chenm12@stu.xjtu.edu.cn (C.M.); jundu2010@xjtu.edu.cn (J.D.); pengfeihe@stu.xjtu.edu.cn (P.H.); jmb3975@stu.xjtu.edu.cn (M.J.)

² School of Foreign Languages, Lanzhou City University, Lanzhou 730070, China; dhf@lzcw.edu.cn

* Correspondence: zywei@mail.xjtu.edu.cn

Abstract: This paper proposes a novel metal additive manufacturing process, which is a composition of gas tungsten arc (GTA) and droplet deposition manufacturing (DDM). Due to complex physical metallurgical processes involved, such as droplet impact, spreading, surface pre-melting, etc., defects, including lack of fusion, overflow and discontinuity of deposited layers always occur. To assure the quality of GTA-assisted DDM-ed parts, online monitoring based on visual sensing has been implemented. The current study also focuses on automated defect classification to avoid low efficiency and bias of manual recognition by the way of convolutional neural network-support vector machine (CNN-SVM). The best accuracy of 98.9%, with an execution time of about 12 milliseconds to handle an image, proved our model can be enough to use in real-time feedback control of the process.

Keywords: gas tungsten arc (GTA); droplet deposition manufacturing (DDM); visual sensing; defect classification; convolutional neural network-support vector machine (CNN-SVM)

Citation: Ma, C.; Dang, H.; Du, J.; He, P.; Jiang, M.; Wei, Z. Research on Automated Defect Classification Based on Visual Sensing and Convolutional Neural Network-Support Vector Machine for GTA-Assisted Droplet Deposition Manufacturing Process. *Metals* **2021**, *11*, 639. <https://doi.org/10.3390/met11040639>

Academic Editors: Babak Shalchi Amirkhiz and Yongho Sohn

Received: 5 March 2021

Accepted: 10 April 2021

Published: 14 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Additive manufacturing (AM) is revolutionary compared to traditional processing methods in creating complex 3D-shaped components. Among the different additive manufacturing techniques, wire and arc additive manufacturing (WAAM), which combines an electric arc as heat source and wire as the feedstock material, is suitable to produce large metallic parts owing to the high deposition rates being significantly larger than powder as the feedstock [1]. Gas metal arc (GMA), gas tungsten arc (GTA) and plasma arc (PA) are the most used processes in WAAM. They all need external filler materials and a high energy-density arc heat source under an inert shielding gas [2]. Comparing with wires as feedstock, we developed a new metal additive manufacturing process in this paper, which uses fused droplets as feedstock combined with variable-polarity GTA to form aluminum alloy. The solid cylindrical aluminum alloy is inductively heated in a graphite crucible to a molten state. At the same time, a certain argon pressure is applied to make the fused aluminum alloy droplets flow out of the nozzle and fall into the molten pool by GTA. The deposited layer is formed after the liquid metal solidification. Figure 1 shows the schematic diagrams of the two different processes, where Figure 1a is the schematic diagram of WAAM, and Figure 1b is the schematic diagram of the process we presented.

The GTA process is accompanied by a highly non-linear heat source, and there are several input parameters to consider, such as the welding voltage and current, the process speed, the shielding gas flow and the type of materials [3,4]. To make the process relatively stable and restrain defects to form a good shape, non-destructive testing (NDT) plays a vital role in implementing online monitoring without changing or damaging the nature

and structure of the parts. Therefore, it is economical in different levels of development and maintenance [5].

Figure 1. The schematic diagrams of wire and arc additive manufacturing (WAAM) and gas tungsten arc (GTA)-assisted droplet deposition manufacturing (DDM). (a) WAAM; (b) GTA-assisted DDM.

In the past few years, several typical NDT techniques such as computerized tomography (CT), radiographic testing (RT), ultrasonic testing (UT), magnetic particle inspection (MPI) and eddy current testing (ET) have been applied to the field of metal AM [6]. Chabot et al. [7] applied a phased array ultrasonic testing (PAUT) to WAAM components. With the help of X-ray radiography, the PAUT method finished defect size detection from 0.6 to 1 mm for aluminum alloy parts. Bento et al. [8] developed an eddy current testing (ECT) system where the customized ECT probes were able to locate artificial defects: at depths up to 5 mm; with a thickness as small as 350 μm ; with the probe up to 5 mm away from the inspected sample surface. Wu et al. [9] used an infrared monochrome pyrometer (IMP) for accurately identifying simulated cracks on the surface of a laser metal deposition (LMD) sample. To detect lack-of-fusion defects, Montazeri et al. [10] captured the dynamic phenomena around the melt pool region by a spectrometer and an optical camera during directed energy deposition (DED). Chang et al. [11] proposed a method based on the position information of electron beam speckle to realize the three-dimensional reconstruction of the surface of the deposited parts in the process of electron beam freeform fabrication (EBF3).

As a very important method of NDT, a visual sensing system is widely used in online monitoring of metal AM. In addition, a lot of image processing algorithms suitable for different processes have been designed in order to improve the detection stability of systems. Zhuang et al. [12] proposed k-nearest neighbor (KNN) classification algorithms based on contour curve-KNN (CC-KNN) and locality preserving projection-KNN (LPP-KNN) effectively performed in vision and spectral analysis. Yu et al. [13] established the visual sensing system to capture every frame of the molten pool images matched for the actual weld location in the GMA AM process. A back propagation (BP) neural network was used to extract the shape and location features of the molten pool. Xia et al. [14] developed a visual sensing system working with a robot and a cold metal transfer (CMT) welder. The adaptive Wiener filter and the Canny algorithm were utilized to obtain information in welding pool images. Aminzadeh et al. [15] developed and trained a statistical Bayesian classifier to classify the quality of the build that signifies the defective or unacceptable build layers during laser powder bed fusion (LPBF).

Deep learning (DL) algorithms have recently grabbed the attention of scientists due to their strong ability to learn high-level features from raw data, and in most cases, they are much better than traditional algorithms in terms of accuracy and robustness. Convolutional

neural networks (CNN) are particularly more used in computer vision tasks which include image classification, object detection, segmentation and so on [16–19]. However, restricted by the computation performance and datasets, CNN fell out of use for several years until AlexNet was proposed by Krizhevsky for the ImageNet competition in 2012 [20,21]. Subsequently, VGGNet [22] and GoogleNet [23] were proposed considering the width and depth of the network respectively. ResNet [24] proved that the depth of the network can be increased very deeply. With the rise of state-of-the-art CNN architectures, researchers have introduced them to the field of metal AM. Cui et al. [25] used the Missouri S&T dataset (optical microscope images of LMD parts) to train and investigate Hyper-parameters including kernel size and the number of layers of their CNN model. Kwon et al. [26] applied CNN to melt-pool images with respect to six laser power labels in selective laser melting (SLM). The classification failure rate was under 0.01. Yin et al. [27] adopted CNN to analyze the welding process parameters and the weld dimensions from twin-wire CMT welding of 5083 aluminum alloy. Zhang et al. [28] presented the application of deep learning framework for automated surface quality inspection in recognition of under-melt, beautiful-weld and over-melt categories in LPBF. The classification accuracy of the finally developed model on the UB-Moog dataset is 0.82 by optimizing hyper parameters. Wang et al. [29], based on previous work [13], developed a prediction network (PredNet) to predict the change of molten pool shape 140ms in advance. Through regression network (SERes), the predicted results were regressed to the accurate weld reinforcement information of the deposited layer in advance. Tomaz et al. [30] realized multi-objective optimization during the GTAW process with the help of an artificial neural network (ANN) and a genetic algorithm (GA). The optimal welding parameters, including welding current = 222 A, welding speed = 25 cm/min, nozzle deflection distance = 8 mm, travel angle = 25°, were determined, and the determination coefficient (R²) and RMSE value of all response parameters were satisfactory, and the R² of all the data remained higher than 0.65.

The theorem called “No Free Lunch” states that no algorithm can perform well on all problems, so the objective of this work is to explore a good CNN-SVM-based model with the best possible optimizer function, good learning rate and a varied number of epochs to identify the common defects with best accuracy in GTA-assisted DDM. The results can be used for quasi-real-time (layer-wise) process control, further process decisions or corrective actions.

The remainder of this paper is organized as follows. In Section 2, we introduce the GTA-assisted DDM experiment platform and the CNN-SVM architecture in detail. In Section 3, the hyperparameters optimization is introduced in detail, including performance evaluation and the visualization of CNN features. The conclusion is summarized in Section 4.

2. Experiment Platform and Methods

2.1. Experiment Platform

GTA-assisted DDM experiment platform combined the GTAW platform (Fronius, Pettenbach, Austria), melting type high frequency induction heating equipment (SPG50K-15AB, ShuangPing Power Technology Co., Ltd., ShenZhen, China) and a visual sensing system (Mikrotron GmbH, Unterschleissheim, Germany), which are shown in Figure 2. The GTA welding platform included a welding power supply (Fronius Magicwave 3000 Job G/F, Fronius, Pettenbach, Austria) and a TIG robot welding torch (TBi RT20, TBi Industries GmbH, Fernwald-Steinbach, Germany). The melted aluminum 2024 was processed into a cylinder with a diameter of 60 mm, height of 80 mm and put in the graphite crucible. Aluminum 2024 grade was selected as work material because it is high-strength duralumin and extensively used in high-load parts such as skeletons and skins on aircraft. The chemical composition of that is shown in Table 1. The size of the substrate was 220 mm × 220 mm × 10 mm and a 65° inclination angle formed between the welding torch and the substrate. The visual sensing system consisted of a CMOS camera (EoSens CL: CAMMC1362, Mikrotron GmbH, Unterschleissheim, Germany) with an optical lens (AT-X 100 mm F2.8, Kenko Tokina Co., Ltd., Tokyo, Japan), an image acquisition card

(Xtium-CL MX4, Teledyne DALSA, Waterloo, ON, Canada) and other data storage devices. The high-speed camera was fixed on a tripod (Benro IF28+) and focused on the back edge of the deposition layer through the glass panel on the front of the glovebox. Once focused by adjusting the aperture to the maximum position, the aperture and the camera exposure time were adjusted at the same time to darken the field of view on the monitor. During the experiment, the droplets falling into the molten pool accurately matched each frame of image to a specific welding position.

(a)

(b)

Figure 2. GTA-assisted DDM experiment platform. (a) Schematic diagram; (b) physical diagram.

Table 1. The chemical composition of melted material (wt.%).

Elements	Cu	Mn	Mg	Zn	Al
Composition	4.75	0.5	1.56	0.25	Others

2.2. Experiment Methods

GTA-assisted DDM was mainly affected by parameters such as forming speed (T_s), forming current (I_p), forming flux (Qv), substrate temperature (T_b), etc. The overall dimensions of a good Al-2024 single-pass deposited layer were generally stable, and the surface of the deposited layer had fish-scale patterns, as shown in Figure 3c. It can be seen that the deposited layer spreads continuously and is metallurgically bonded well to the substrate. Figure 3a shows the cross section of the deposited layer. Table 2 lists the welding parameters as a standard baseline to achieve good processing conditions. Images of good Al-2024 single-pass deposited layers were captured and recorded, then several defects were introduced by altering process parameters one at a time. Finally, morphology images of defective deposited layers were also recorded.

Figure 3. Good Al-2024 single-pass deposited layer. (a) Cross-section of deposited layer; (b) Partial enlargement of deposited layer; (c) Deposited layer with good morphology.

Table 2. Standard baseline process parameters.

Current	Forming Flux	Substrate Temp	Travel Speed	Shield Gas Flux
260 A	140 mm ³ /s	280 °C	8 mm/s	15 L/min

Defects like lack of fusion between the deposited layer and the substrate mean that overflow and discontinuity of the deposited layer are prone to appear when process parameters change. Lack of fusion is mainly caused by insufficient heat input. When $I_p \leq 200$ A or $T_b \leq 220$ °C, it is difficult to spread the deposited layer and bond with the substrate poorly. Overflow is mainly due to excessive heat input and forming flux, when $I_p \geq 280$ A or $Qv \geq 200$ mm³/s, the heat accumulation is serious and the molten droplets cannot be completely absorbed by the molten pool. Discontinuity is mainly caused by the excessive forming speed, when $T_s \geq 30$ mm/s, the droplets will not fall into the molten pool continuously, causing fluctuations in the outer dimensions of the deposited layer. Figure 4. shows the macroscopic morphology and raw captured images with a size of 1280×1024 pixels for the above three common defects at different process parameters, which provided enough information about the defects.

Figure 4. The macroscopic morphology of common defects. (a) Lack of fusion ($I_p \leq 200$ A or $T_b \leq 220$ °C); (b) overflow ($I_p \geq 280$ A or $Q_v \geq 200$ mm³/s); (c) discontinuity ($T_s \geq 30$ mm/s).

2.3. Preprocessing

The size of raw captured images (1280×1024 pixels) had an approximate size of 1.2 MB for each one, which contained a substantial number of black pixels surrounding the deposited layer and the welding arc as seen in Figure 4. As a result of the hardware constraints, during the training stage of the algorithms, where the model receiving a higher resolution requires significantly more GPU memory. The subsampling operation is necessary before creating the deep learning dataset [31]. The reliability and speed of the algorithms are improved by region of interest (ROI) segmentation, histogram equalization and image filtering [32]. Figure 5 shows four types of data after ROI segmentation. The segmented images with 326×495 pixels were compressed to 155 kB to 160 kB. There were “good” (582), “lack of fusion” (641), “overflow” (589) and “discontinuity” (588). Among them, data at the beginning or end of the process that seriously affected the sample were given up to minimize the impact of sample imbalance on performance of algorithms.

2.4. Data Augmentation

Generally speaking, parameters of many deep CNN architectures are in the millions; a lot of data was required to make these parameters work correctly for training. It does not work to rely on new data completely in AM because of time and economic costs [33]. Data augmentation needs to be used to improve the model generalization ability due to the high diversity of welding conditions [34]. In this paper, scaling as well as translation, rotation, flipping, adding “Salt and Pepper” noise and changing the lighting condition was applied to create more data to make the algorithms have a better generalization effect. The result of the original image “good” processed by the above data augmentation methods was shown in Figure 6. Where scaling was achieved by three scale factors of 0.9, 0.75, 0.6, translation was done by moving 20% left, right, up and down respectively and rotation was finished by sequentially rotating 6° once a time from -30° to 30° . The number of each class after data augmentation was given in Table 3 and they were split into two main subsets: training and testing, with approximatively 75% and 25%, respectively.

Figure 5. Four types of data after region of interest (ROI) segmentation. “good” (582) vs. “lack of fusion” (641) vs. “overflow” (589) vs. “discontinuity” (588).

(a)

(b)

Figure 6. The result of the original image “good” processed by data augmentation. (a) the resized image of the original image “good” with a size of 224×224 pixels; (b) the result in the order of flipping, changing the lighting condition, adding “Salt and Pepper” noise, scaling, translation and rotation.

Table 3. The number of images in each category.

	Good	Lack of Fusion	Overflow	Discontinuity	Total
Train	6984	8136	7212	7056	29,388
Test	2328	2712	2404	2352	9796

2.5. CNN+SVM Architecture

CNN consists of multiple, repeating components which are stacked in layers: convolution, pooling, fully connected and classifier layers. Among them, local receptive field, shared weights and pooling are the three most important concepts [19,20]. The model in this paper draws on the ideas of the classic AlexNet model and linear SVM. The overall structure is composed of two parts: feature extraction and classification, which is shown in Figure 7.

Figure 7. The schematic of convolutional neural network-support vector machine (CNN-SVM) model.

The feature extraction consists of five convolutional layers named C1–C5 with a size of corresponding filters: 11×11 , 5×5 , 3×3 respectively and three maximum pooling layers named P1–P3. It can be described as follows:

$$x^l = f(b^l + \sum_j \sum_i w_{i,j}^l x_{i,j}^{l-1}) \quad (1)$$

$$f(x) = \text{ReLU}(x) = \begin{cases} x & x > 0 \\ 0.01x & x \leq 0 \end{cases} \quad (2)$$

$$a_j = \max_{N \times N} (a_i^{N \times N} u(n, n)) \quad (3)$$

where (J, I) denotes the size of the filters, J is the height of the filters, and I is the width of the filters. b^l denotes the bias of the convolutional layer. x^{l-1} denotes the output of the previous layer. w^l denotes the weight of the convolutional layer. $f(x)$ is the nonlinear activation function rectified linear units (ReLU) shown as Equation (2). Pooling operations are shown as Equation (3), which is a form of non-linear down-sampling and used to replace the output of a certain location with a summary statistic of nearby [34,35]. Finally, batch normalization was used for centering and normalization of the images and applied before the fully connected layers instead of local response normalization (LRN) for its unobvious regularization effects.

It comes to the classification stage. On the one hand, output of P3 was flattened to transfer to the fully connected layers with a 0.4 dropout to prevent overfitting of the data.

The was model constructed here with the better choice of optimizer function between stochastic gradient descent (SGD) and adaptive moment estimation (Adam). Varied learning rates ranging from 1×10^{-3} to 1×10^{-5} were selected and optimized by minimizing cross-entropy loss function. That can be described as:

$$H(p, q) = -\sum_x p(x) \log q(x) \quad (4)$$

where H is the cross entropy, which can determine how close the actual output is to the expected output, p is the expected output of sample x and $\log q(x)$ is the logarithm of the actual output of sample x .

On the other hand, the top layer is replaced by a linear SVM (L2-SVM), which minimizes the squared hinge loss, shown as Equation (5). Output of the pooling layer, M3, is given to the SVM algorithm and the classification performs.

$$\min_w \frac{1}{2} w^T w + C \sum_{n=1}^N \text{Max}(1 - w^T x_n t_n, 0)^2 \quad (5)$$

where w is the normal vector of classification hyperplane, C is a regularization parameter, x_n is the feature vector, t_n is the label it belongs to.

The overall training procedure was performed on a PC with an Intel Core i5-4570 CPU, RAM of 12 GB, a NVIDIA RTX 2080 Ti GPU, running Windows 10 professional with CUDA 9.0 libraries. The software environment was Python 3.6 using TensorFlow 1.13.1.

2.6. Evaluation Metrics

The commonly used evaluation metrics were described as Equations (6)–(9):

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \quad (6)$$

$$\text{Precision} = TP / (TP + FP) \quad (7)$$

$$\text{Recall} = TP / (FN + TP) \quad (8)$$

$$F\text{score} = (2 \times \text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (9)$$

Among them, accuracy is defined as the number of correct predictions, precision refers to how many of all positive samples are correctly identified, also called positive predictive value (PPV), recall is a measure of the degree of discrimination of all positive samples, which tells how many samples judged as positive are true. F score indicates the overall performance of the precision and recall, which is the harmonic mean of precision and recall. True Positive (TP) describes the number when the true value is positive and the model is considered to be positive; True Negative (TN) counts the number when the true value is negative and the model is considered to be negative; False Positive (FP) illustrates the number when the true value is negative but the model is considered to be positive; False Negative (FN) represents the number when the true value is positive but the model is considered to be negative.

3. Results and Discussion

3.1. Tuning of CNN Architecture

The classic AlexNet model has more than 60 million parameters and more than 650,000 neurons. If the parameter adjustment is not ideal, the performance will not be represented well. Batch size, learning rate, epochs, weight initialization, dropout, the choice of optimizer and data normalization are several indicators we are mainly concerned about. Figure 8 shows the effect of a varied batch size on the stability of model convergence, which presents loss and accuracy after training 500 epochs. A larger batch size is conducive to more stable model convergence but constrained by hardware. Hence, the maximum batch size used in our model is 128. It is particularly emphasized that networks can

converge unless the weight initialization of the fully connected is 1×10^{-4} . The weights of other layers can be initialized using a Gaussian distribution with a mean value of 0 and a standard deviation of 0.01 or 0.1. The neuron bias of all layers is initialized to a constant 0.1.

Figure 8. The effect of varied batch size on the stability of model convergence. (a,b) batch size = 16; (c,d) batch size = 32; (e,f) batch size = 64; (g,h) batch size = 128.

The change-of-learning rate is often combined with the optimization function used in model training. This study compared two different optimization functions of SGD and Adam [36,37], and then analyzes the impact of learning rate on model convergence. SGD randomly selects a point for calculating the fastest descent direction instead of traversing the entire training data set. This can greatly speed up the iteration speed while taking the local optimal solution into account. It can be described as:

$$J(\theta) = \frac{1}{2} \sum_{i=1}^m (y^i - h_{\theta}(x^i))^2 \quad (10)$$

$$\theta'_j = \theta_j + \alpha(y^i - h_{\theta}(x^i))x_j^i \quad (11)$$

where, $J(\theta)$ is the loss function that needs to be minimized, m is the sample batches, $h_{\theta}(x^i)$ is the function of parameter θ fitting to the sample, α is called learning rate.

The Adam algorithm calculates step size to be updated by calculating the first moment estimation and second moment estimation of the gradient, which are the mean of the gradient and the decentralized variance of the gradient, to design independent adaptive learning rates for different parameters. Figures 9 and 10 show the process of training models with different learning rates and optimization functions. It can be seen that the selection of the learning rate has a great influence on whether the model converges. When the learning rate is 1×10^{-4} and Adam is used, the model converges with the highest accuracy of convergence. If the learning rate is set too large, the larger training step makes the parameters oscillate back and forth on both sides of the optimal solution. If the learning rate is too small, the convergence speed will be greatly reduced and the final model will not reach the best accuracy.

Figure 9. Cont.

Figure 9. Using SGD to train the model with different learning rates. (a,b) learning rate = 1×10^{-3} ; (c,d) learning rate = 1×10^{-4} ; (e,f) learning rate = 1×10^{-5} .

Figure 10. Using Adam to train the model with different learning rates. (a,b) learning rate = 1×10^{-3} ; (c,d) learning rate = 1×10^{-4} ; (e,f) learning rate = 1×10^{-5} .

3.2. Performance Evaluation

The precision and recall of our model were evaluated on the test dataset according to evaluation metrics in Section 2.4 and the results were reported in Table 4. Overall, F score can basically reach 0.9, indicating good classification performance. At the same time, a comparative experiment of KNN, single use of SVM or CNN and our model was carried out in order to highlight our model in terms of recognition accuracy and efficiency. The results are listed in Table 5. It can be seen that our model has an accuracy of 98.9% with a time of about 12 milliseconds to recognize an image, which would be enough to use in a real-time feedback control of our process.

Table 4. Precision, recall, and F score of our model on the test dataset.

Class	Precision	Recall	F Score
Good	0.93	0.96	0.945
Lack of fusion	0.92	0.89	0.905
Overflow	0.95	0.94	0.945
Discontinuity	0.91	0.88	0.895

Table 5. Results of accuracy and time for different classification methods.

Method	Accuracy	Time (s)
KNN	0.926	0.61
SVM	0.864	0.13
CNN	0.96	0.008
Our model	0.989	0.012

3.3. The Visualization of CNN Features

The visualization of the feature map after convolution can help us understand what each layer of convolution has learned and show the process of abstracting features layer by layer. Figure 11 shows the feature maps of four samples representing “overflow”, “good”, “discontinuity” and “lack of fusion” after having been learned by different layers. Some feature maps focus on the background of images, others are more inclined to the outline of images. The 96 and 256 feature maps from the first two convolutional layers C1, C2 are basically about information of edges, stripes and grayscale where the shape of different classes can be seen clearly. Starting from C3, a larger number of 3×3 convolution kernels are used. Compared with the previous convolution layer, the 3×3 convolution kernel has a larger receptive field on the convolution output of this step. This expansion of the receptive field allows the convolutional layer to combine low-level features (lines, edges) into higher-level features (curves, textures) and this more abstract expression makes it less visually interpretable or difficult to differentiate.

Figure 11. The feature maps of four samples representing “overflow”, “good”, “discontinuity” and “lack of fusion” after having been learned by different layers.

4. Conclusions

Quality monitoring based on visual sensing was applied to a novel metal additive manufacturing process, GTA-assisted DDM, in this paper.

(1) A large number of process experiments were implemented with parameters including forming speed (T_s), forming current (I_p), forming flux (Q_v) and substrate temperature (T_b) deviating from a standard baseline. Different kinds of morphology images of deposited layers were obtained: “good”, “lack of fusion”, “overflow” and “discontinuity”.

(2) Translation, rotation, flipping, adding “Salt and Pepper” noise and other data augmentation methods were used to expand the original image dataset and reduce the cost of the process experiments.

(3) Training of a CNN-SVM model based on AlexNet and linear support vector machine was completed, where a batch size of 128 and learning rate of $1e-4$ on Adam function were determined as the optimal value, and the output of trained P3 as a feature was transferred to the SVM for classification. The results showed that F scores of our model could reach 0.9 mostly, and compared to KNN, or single use of SVM or CNN, it had a best test set accuracy of 98.9% as well as an execution time of about 12 milliseconds to handle an image, which provided a more sufficient control time in the GTA-assisted DDM.

Author Contributions: Conceptualization, methodology, writing—original draft, C.M.; proofreading, editing, H.D.; project administration, J.D. and P.H.; feasibility analysis, M.J.; supervision, Z.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (51775420), and the Key Research and Development program of China (2016YFB1100401HZ).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

GTA	Gas tungsten arc
T_s	Forming speed
I_p	Forming current
Q_v	Forming flux
T_b	Substrate temperature
C1–C5	Convolutional layer with size of corresponding filters: $n_x \times n_y$, stride: t and padding (Yes or Not)
P1–P3	Pooling layer
ReLU	rectified linear units specified as $f(x)$ for nonlinear activation of neurons
α	learning rate used during training
EPOCHs	Number of passes over the data set during training
DROPOUT	Dropout, specified as 0.4 for fully connected layers to prevent overfitting
KNN	k-nearest neighbor
SGD	Stochastic gradient descent
Adam	Adaptive moment estimation

References

- Williams, S.W.; Martina, F.; Addison, A.C.; Ding, J.; Pardal, G.; Colegrove, P. Wire plus Arc Additive Manufacturing. *Mater. Sci. Technol.* **2016**, *32*, 641–647. [\[CrossRef\]](#)
- Rodrigues, T.A.; Duarte, V.; Miranda, R.M.; Santos, T.G.; Oliveira, J.P. Current Status and Perspectives on Wire and Arc Additive Manufacturing (WAAM). *Materials* **2019**, *12*, 1121. [\[CrossRef\]](#)
- Rodriguez-Cobo, L.; Ruiz-Lombera, R.; Conde, O.M.; Lopez-Higuera, J.M.; Cobo, A.; Mirapeix, J. Feasibility study of Hierarchical Temporal Memories applied to welding diagnostics. *Sens. Actuator A Phys.* **2013**, *204*, 58–66. [\[CrossRef\]](#)
- Arora, H.; Kumar, V.; Prakash, C.; Pimenov, D.; Singh, M.; Vasudev, H.; Singh, V. Analysis of Sensitization in Austenitic Stainless Steel-Welded Joint. In *Advances in Mechanical Engineering*; J.B. Metzler: Jalandhar, India, 2021; pp. 13–23.
- Gao, X.D.; Li, G.H.; Chen, Z.Q.; Lan, C.Z.; Li, Y.F.; Gao, P.P. Modeling for detecting weld defects based on magneto-optical imaging. *Appl. Optics* **2018**, *57*, 6110–6119. [\[CrossRef\]](#)
- Florence, S.E.; Samsingh, R.V.; Babureddy, V. Artificial intelligence based defect classification for weld joints. In Proceedings of the IOP Conference Series: Materials Science and Engineering, Bandung, Indonesia, 9–13 July 2018; IOP Publishing: Kattankulathur, India, 2018; Volume 402, p. 012159.
- Chabot, A.; Laroche, N.; Carcreff, E.; Rauch, M.; Hascoet, J.Y. Towards defect monitoring for metallic additive manufacturing components using phased array ultrasonic testing. *J. Intell. Manuf.* **2020**, *31*, 1191–1201. [\[CrossRef\]](#)
- Bento, J.B.; Lopez, A.; Pires, I.; Quintino, L.; Santos, T.G. Non-destructive testing for wire plus arc additive manufacturing of aluminium parts. *Addit. Manuf.* **2019**, *29*, 100782.
- Wu, Y.; Cui, B.; Xiao, Y. Crack Detection during Laser Metal Deposition by Infrared Monochrome Pyrometer. *Materials* **2020**, *13*, 5643. [\[CrossRef\]](#)
- Montazeri, M.; Nassar, A.R.; Stutzman, C.B.; Rao, P. Heterogeneous sensor-based condition monitoring in directed energy deposition. *Addit. Manuf.* **2019**, *30*, 100916. [\[CrossRef\]](#)
- Chang, S.H.; Zhang, H.Y.; Xu, H.Y.; Sang, X.H.; Wang, L.; Du, D.; Chang, B.H. Online Measurement of Deposit Surface in Electron Beam Freeform Fabrication. *Sensors* **2019**, *19*, 4001. [\[CrossRef\]](#)
- Zhao, Z.; Guo, Y.T.; Bai, L.F.; Wang, K.H.; Han, J. Quality monitoring in wire-arc additive manufacturing based on cooperative awareness of spectrum and vision. *Optik* **2019**, *181*, 351–360. [\[CrossRef\]](#)
- Yu, R.W.; Zhao, Z.; Bai, L.F.; Han, J. Prediction of Weld Reinforcement Based on Vision Sensing in GMA Additive Manufacturing Process. *Metals* **2020**, *10*, 1041. [\[CrossRef\]](#)
- Xia, C.Y.; Pan, Z.X.; Zhang, S.Y.; Polden, J.; Wang, L.; Li, H.J.; Xu, Y.L.; Chen, S.B. Model predictive control of layer width in wire arc additive manufacturing. *J. Manuf. Process* **2020**, *58*, 179–186. [\[CrossRef\]](#)
- Aminzadeh, M.; Kurfess, T.R. Online quality inspection using Bayesian classification in powder-bed additive manufacturing from high-resolution visual camera images. *J. Intell. Manuf.* **2019**, *30*, 2505–2523. [\[CrossRef\]](#)
- Tao, W.; Leu, M.C.; Yin, Z. American Sign Language alphabet recognition using Convolutional Neural Networks with multiview augmentation and inference fusion. *Eng. Appl. Artif. Intell.* **2018**, *76*, 202–213. [\[CrossRef\]](#)
- Li, K.; Wu, Z.; Peng, K.-C.; Ernst, J.; Fu, Y. Tell Me Where to Look: Guided Attention Inference Network. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 27 February 2018; pp. 9215–9223.
- Wang, T.; Chen, Y.; Qiao, M.; Snoussi, H. A fast and robust convolutional neural network-based defect detection model in product quality control. *Int. J. Adv. Manuf. Technol.* **2018**, *94*, 3465–3471. [\[CrossRef\]](#)
- Azimi, S.M.; Britz, D.; Engstler, M.; Fritz, M.; Mücklich, F. Advanced steel microstructural classification by deep learning methods. *Sci. Rep.* **2018**, *8*, 2128. [\[CrossRef\]](#)
- Lecun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [\[CrossRef\]](#)

21. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, 1097–1105. [\[CrossRef\]](#)
22. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
23. Szegedy, C.; Liu, W.; Jia, Y.Q.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 23–28 June 2015; pp. 1–9.
24. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
25. Cui, W.Y.; Zhang, Y.L.; Zhang, X.C.; Li, L.; Liou, F. Metal Additive Manufacturing Parts Inspection Using Convolutional Neural Network. *Appl. Sci.* **2020**, *10*, 545. [\[CrossRef\]](#)
26. Kwon, O.; Kim, H.G.; Ham, M.J.; Kim, W.; Kim, G.H.; Cho, J.H.; Kim, N.I.; Kim, K. A deep neural network for classification of melt-pool images in metal additive manufacturing. *J. Intell. Manuf.* **2020**, *31*, 375–386. [\[CrossRef\]](#)
27. Yin, L.M.; Wang, J.Z.; Hu, H.Q.; Han, S.G.; Zhang, Y.P. Prediction of weld formation in 5083 aluminum alloy by twin-wire CMT welding based on deep learning. *Weld. World.* **2019**, *63*, 947–955. [\[CrossRef\]](#)
28. Zhang, B.B.; Jaiswal, P.; Rai, R.; Guerrier, P.; Baggs, G. Convolutional neural network-based inspection of metal additive manufacturing parts. *Rapid Prototyp. J.* **2019**, *25*, 530–540. [\[CrossRef\]](#)
29. Wang, Y.M.; Zhang, C.R.; Lu, J.; Bai, L.F.; Zhao, Z.; Han, J. Weld Reinforcement Analysis Based on Long-Term Prediction of Molten Pool Image in Additive Manufacturing. *IEEE Access* **2020**, *8*, 69908–69918. [\[CrossRef\]](#)
30. Tomaz, I.d.V.; Colaço, F.H.G.; Sarfraz, S.; Pimenov, D.Y.; Gupta, M.K.; Pintaude, G. Investigations on quality characteristics in gas tungsten arc welding process using artificial neural network integrated with genetic algorithm. *Int. J. Adv. Manuf. Technol.* **2021**, 1–15. [\[CrossRef\]](#)
31. Bacioiv, D.; Melton, G.; Papaelias, M.; Shaw, R. Automated defect classification of Aluminium 5083 TIG welding using HDR camera and neural networks. *J. Manuf. Process* **2019**, *45*, 603–613. [\[CrossRef\]](#)
32. Yahia, N.B.; Belhadj, T.; Brag, S.; Zghal, A. Automatic detection of welding defects using radiography with a neural approach. In Proceedings of the 11th International Conference on the Mechanical Behavior of Materials (ICM), Como, Italy, 5–9 June 2011; p. 10.
33. Gaikwad, A.; Imani, F.; Yang, H.; Reutzel, E.; Rao, P. In Situ Monitoring of Thin-Wall Build Quality in Laser Powder Bed Fusion Using Deep Learning. *Smart Sustain. Manuf. Syst.* **2019**, *3*, 98–121. [\[CrossRef\]](#)
34. Zhang, Z.F.; Wen, G.R.; Chen, S.B. Weld image deep learning-based on-line defects detection using convolutional neural networks for Al alloy in robotic arc welding. *J. Manuf. Process.* **2019**, *45*, 208–216. [\[CrossRef\]](#)
35. Zhu, H.X.; Ge, W.M.; Liu, Z.Z. Deep Learning-Based Classification of Weld Surface Defects. *Appl. Sci.* **2019**, *9*, 3312. [\[CrossRef\]](#)
36. Kingma, D.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.
37. Duchi, J.; Hazan, E.; Singer, Y. Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. *J. Mach. Learn. Res.* **2011**, *12*, 2121–2159.