

Article

Transfer Learning-Based Artificial Neural Network for Predicting Weld Line Occurrence through Process Simulations and Molding Trials

Giacomo Baruffa *, Andrea Pieressa, Marco Sorgato *  and Giovanni Lucchetta 

Department of Industrial Engineering, University of Padova, Via Venezia 1, 35131 Padova, Italy; andrea.pieressa@phd.unipd.it (A.P.); giovanni.lucchetta@unipd.it (G.L.)

* Correspondence: giacomo.baruffa@unipd.it (G.B.); marco.sorgato@unipd.it (M.S.); Tel.: +39-3477258385 (G.B.); +39-340-3228609 (M.S.)

Abstract: Optimizing process parameters to minimize defects remains an important challenge in injection molding (IM). Machine learning (ML) techniques offer promise in this regard, but their application often requires extensive datasets. Transfer learning (TL) emerges as a solution to this problem, leveraging knowledge from related tasks to enhance model training and performance. This study explores TL's viability in predicting weld line visibility in injection-molded components using artificial neural networks (ANNs). TL techniques are employed to transfer knowledge between datasets related to different components. Furthermore, both source datasets obtained from simulations and experimental tests are used during the study. In order to use process simulations to obtain data regarding the presence of surface defects, it was necessary to correlate an output variable of the simulations with the experimental observations. The results demonstrate TL's efficacy in reducing the data required for training predictive models, with simulations proving to be a cost-effective alternative to experimental data. TL from simulations achieves comparable predictive metric values to those of the non-pre-trained network, but with an 83% reduction in the required data for the target dataset. Overall, transfer learning shows promise in streamlining injection molding optimization and reducing manufacturing costs.

Keywords: injection molding; machine learning; transfer learning; weld lines; surface quality prediction; digital twin; artificial intelligence; artificial neural networks; CAE; part similarity



Citation: Baruffa, G.; Pieressa, A.; Sorgato, M.; Lucchetta, G. Transfer Learning-Based Artificial Neural Network for Predicting Weld Line Occurrence through Process Simulations and Molding Trials. *J. Manuf. Mater. Process.* **2024**, *8*, 98. <https://doi.org/10.3390/jmmp8030098>

Academic Editor: Yu-Ming Zhang

Received: 2 April 2024

Revised: 29 April 2024

Accepted: 8 May 2024

Published: 9 May 2024



Copyright: © 2024 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

In modern manufacturing, injection molding is a cornerstone process renowned for its versatility and efficiency in producing intricate components across diverse industries [1–3]. Its significance lies in mass production and in the ability to fabricate parts with high surface quality, intricate geometry, and minimal defects. However, achieving the desired level of quality and minimizing defects in injection-molded components remains a perpetual challenge for manufacturers.

Ensuring the production of components without defects holds paramount importance across industries for several compelling reasons [4,5]. Moreover, defect-free components minimize the need for post-production inspection and rework, streamlining the manufacturing process and reducing operational costs [6].

However, achieving flawless components, particularly those with stringent aesthetic standards, presents a formidable challenge when operators need to set process parameters during process setup or production startup. This is because injection molding is a complex and non-linear process that involves many process parameters, such as temperature, pressure, injection speed, cooling time, and mold design [7–12]. These parameters impact the quality of molded parts and exhibit non-linear relationships with quality attributes [13]. The traditional way of setting molding process parameters is by trial and error [14,15]. This

means that the operators are tasked to repeat the injection molding process during the machine setup phase. After each shot, molding staff must use their process knowledge to fine-tune the parameters and gradually improve the product quality [16]. However, the job market is becoming more competitive [17,18], making it harder to find experienced staff who can perform such tasks.

In recent years, machine learning has emerged as a powerful tool in the realm of manufacturing optimization, offering unprecedented capabilities in data analysis, pattern recognition, and predictive modeling [19]. By harnessing vast datasets, ML algorithms can discern intricate relationships and trends without prior knowledge of these relations. In the context of injection molding, ML techniques hold immense promise in guiding the identification of optimal process parameters to achieve desired aesthetic characteristics. ML models are especially well-suited for analyzing the complex interplay between input and output quality parameters in injection molding processes.

Zhang et al. [20] employed optimal Latin hypercube design (Opt LHD), an ensemble-based fuzzy neural network (EBFNN), and multi-objective particle swarm optimization (MOPSO) to achieve multi-objective optimization. The objects of the optimization were the warpage and the clamping force. The optimized parameters were then tested, and as a result, the reduction on the warpage was 60%, and the peak clamping force showed a 38% reduction. Tsai and Luo [21] utilized artificial neural networks combined with genetic algorithms (GAs) to develop an inverse model. This model aimed to determine the optimal process parameters required to minimize warpage in molded lenses. The process parameters optimized by the hybrid ANN and GA resulted in better lens form accuracy than the Taguchi experiment; the form accuracy improved by 13.36%. Ke and Huang [22] utilized an ANN to analyze pressure signals and predict product width. Through experimental validation, the authors demonstrated the effectiveness of the ANN in accurately forecasting quality attributes. The model achieved an accuracy rate exceeding 92% for product width.

Many other researchers have proposed approaches based on applying ANNs to injection molding. Some of the outputs of the IM process studied were mechanical properties [23–25], shrinkage and warpage [26–28], gate location [29], and other process outputs [30–32].

All of the previously cited articles have highlighted an essential advantage of ML: the possibility of moving from a reactive to a preventive approach in managing the quality of components made by injection molding [33]. However, the previous studies also had a significant limitation: a dedicated dataset was generated with each application of ML, and such datasets can be challenging to obtain. In addition, machine learning models require a large amount of data to achieve acceptable performance, which makes ML unattractive in an industrial setting [34,35].

One particularly promising avenue within the domain of machine learning for injection molding optimization is the application of transfer learning. Transfer learning leverages knowledge acquired from training on one task to improve training and performance on a related task [36]. It is intuitive to say that successful transfer learning depends on the level of similarity of the tasks and their domains [37]; differences may prevent successful transfer learning [38]. Successful TL (positive transfer) can be manifested in three ways: higher quality of the model than the non-pre-trained model, higher speed of convergence of the transferred model, and higher generalization ability of the transferred model. The presence of any of the previous three cases indicates positive transfer.

In recent years, transfer learning has become increasingly employed in injection molding. However, few articles have been found regarding the use of TL in press setup. Tercan et al. [39] explored the application of transfer learning in bridging the gap between real-world and simulated data for machine learning in injection molding. The authors investigated how pre-trained models can leverage knowledge from simulation data to enhance performance when applied to real-world scenarios. Through experimental validation, the authors demonstrated the potential to reduce the number of experimental data points required to train the injection molding digital twin via transfer learning applications.

Lockner et al. [40] explored the application of transfer learning between injection molding processes characterized by different polymer materials. The authors proposed transfer learning as a valuable tool for reusing existing data to train multiple models. In this case, the data were obtained by simulating the injection molding process of a single component while considering numerous materials. Through experimental validation, the study demonstrated the efficacy of transfer learning in improving the performance of artificial neural networks and reducing the amount of data required to train a new model. This publication uses artificial neural networks to model the relationships between six molding parameters and part weight. A publication by Lockner and Hopmann [41] applied transfer learning between the injection molding processes of different parts. In this study, the artificial neural network learned the relationship between six injection molding parameters and the part weight. The best results are obtained when transfer learning is applied between similar parts. Gim et al.'s [42] study focused on transferring knowledge from one production site to another, characterized by different molding equipment, to enhance efficiency and robustness. In this research, artificial neural networks are trained using only experimental data to predict the presence of surface defects and surface gloss from molding parameters. The authors demonstrate the effectiveness of transfer learning in improving the performance of machine learning models and reducing the dimension of the training dataset.

Previous studies have shown that the transfer learning approach offers several notable advantages over non-transferred ML models. Firstly, transfer learning reduces the need for extensive data collection and model training specific to the target application. Moreover, transfer learning enhances the robustness and generalization capabilities of ML models, mitigating the risk of overfitting and ensuring reliable performance across diverse production scenarios. Economically, this translates into tangible cost savings by minimizing experimentation iterations, material wastage, and downtime associated with fine-tuning process parameters. However, it is also clear that these preliminary studies on applying TL to injection molding have primarily focused on predicting part weight as the only quality characteristic. While the part weight is handy for assessing the structural performance of parts and allows for numerous considerations regarding process execution, it does not indicate the part's surface quality and aesthetic defects. Therefore, the potential application of transfer learning for predicting aesthetic and surface defects must be investigated further.

In this study, the aim is to apply transfer learning to streamline the training of an artificial neural network tasked with predicting the visibility of weld lines in an industrial component. The input features for this prediction will be the process parameters. Knowledge transfer will occur between injection molding processes related to different parts. Both source datasets obtained from experimental tests and numerical simulations of the injection molding process will be considered. The primary challenge in applying transfer learning from simulations to surface defect prediction lies in the fact that the visibility of defects is not a direct output of the simulations. To overcome this problem, a phenomenological study on the formation of the weld lines was deemed necessary to correlate the simulations' outputs to the weld line visibility.

2. Materials and Methods

2.1. Materials and Parts

The component used to obtain all the source datasets (part 1 or component 1) was a plate for electrical sockets. These parts were manufactured using black polycarbonate (Makrolon 2405, Covestro, Leverkusen, Germany). A thermal analysis using differential scanning calorimetry (Q200, TA Instruments, New Castle, DE, USA), following the ASTM E1356-98 standard [43], was used to characterize the no-flow temperature of the polymer under investigation.

The component was selected due to its stringent surface quality requirements, where even minor surface defects can threaten the acceptability of the plate. In this part, defects are especially conspicuous owing to its opaque nature. Weld line defects were identified as the primary cause of non-conformity for this product, thus prompting the focus on

resolving this issue. Figure 1 depicts a photo of the product with a weld line defect (right) and without a defect (left).

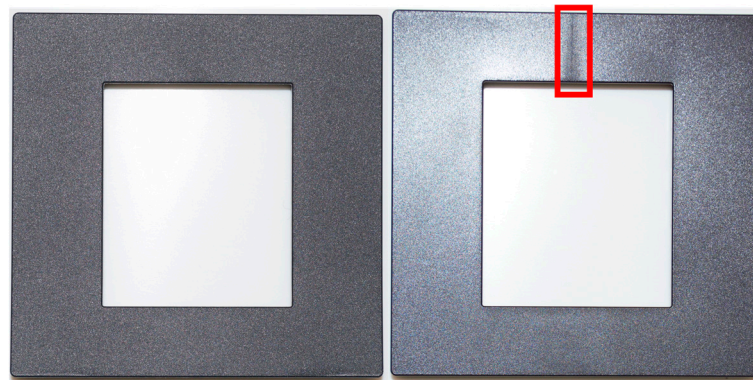


Figure 1. Part 1 with (right) and without (left) weld line defect.

Conversely, the part used to obtain the target dataset was a cover of an electrical socket (part or component 2). These parts were manufactured using the same black polycarbonate employed for part 1 (Makrolon 2405, Covestro, Leverkusen, Germany). Also, on the cover, aesthetic defects were apparent due to the black color of the part and its surface opacity. In this case, weld lines are the most challenging defects to eliminate. Figure 2 shows a photo of component 2 with a weld line defect (right) and without defect (left).

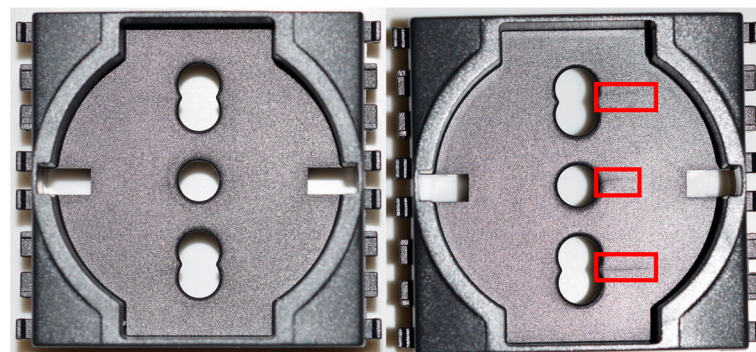


Figure 2. Part 2 with (right) and without (left) weld line defect.

It is crucial to highlight some differences between part 1 and part 2 that may pose challenges for knowledge transfer between the components. Firstly, the significant geometric variation between the parts presents initial obstacles. Regarding geometry, a critical issue arises from the difference in thickness between the parts, particularly in the area where weld lines form. The plate in the defective area has a thickness of 1.8 mm, whereas the cover has a thickness of 2 mm, representing a thickness increase of 11.11%. These thickness differences lead to variations in material flow, thereby influencing the formation of defects. Additionally, it is worth noting that the weld line formed in the plate due to the frontal encounter of two flow fronts (resulting in a cold weld line). In contrast, in the cover, the defect arose from the lateral encounter of two flow fronts (resulting in a hot weld line).

2.2. Simulations

For part 1, experimental and artificial (synthetic) datasets were created. The synthetic dataset was generated from numerical simulations. Throughout this study, all simulations were conducted using Moldex3D software (*Moldex3D 2022 R4*). Within the software's working environment, the CAD models of the component, the hot runner system, and the cooling circuit were imported (Figure 3). It was preferred to import the CAD of all components to improve the accuracy of the numerical analyses.

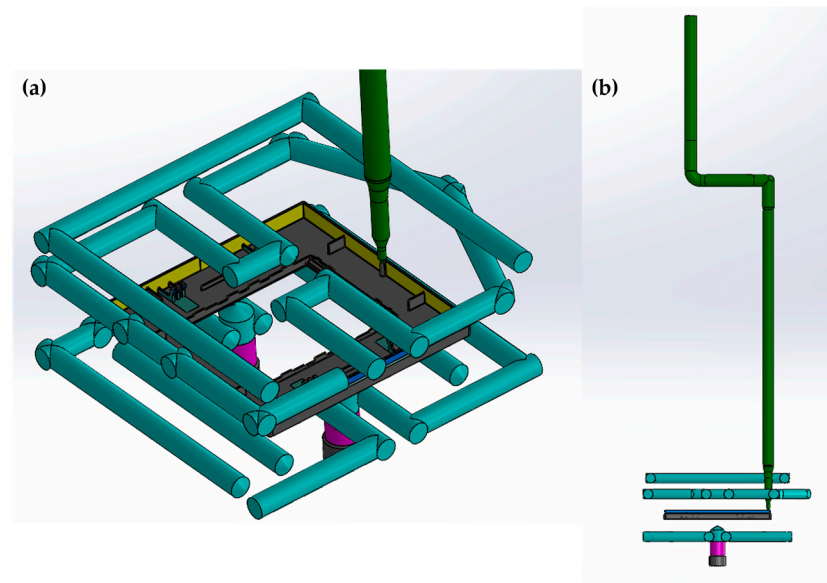


Figure 3. CAD model including part, hot runner system, and cooling channels: (a) Detailed view of the part and insert's cooling channels, (b) overview of part, cooling channels, and hot runner system.

Some modifications to the various models were necessary to obtain the simulation results in acceptable timeframes. First, all low-relief writings and logos were removed from the part, eliminating features that unnecessarily burden the mesh and extend calculation times without enhancing result accuracy. Given its particularly complex geometry, a second significant change was made to the cooling circuit. Mold cooling was achieved through a series of cooling lines operating in parallel, each placed on a different plate of the mold. However, within the Moldex3D environment, only the cooling line closest to the part was modeled, and thus, only the cooling line related to the mold insert was imported. This decision aimed to reduce computational time, considering that these cooling channels dissipate most of the heat brought to the mold by the molten material. In contrast, the remaining channels were situated at a much greater distance from the cavity.

The decision to utilize Moldex3D to simulate the injection molding process stemmed from its unique mesh structure capabilities. Specifically, a mesh structure with five boundary layers was chosen for the part. This boundary layer meshing (BLM) technique significantly enhanced simulation accuracy near the part's surface, where the defect under study formed. Additionally, a 3-boundary layer mesh (3 BLM) was employed for the cooling channels' mesh, while tetragonal elements discretized the feeding circuit and mold plate. The sizes of the mesh elements were determined through a balance between calculation accuracy and computational efficiency, resulting in element sizes of 0.8 mm for the part, 1.2 mm for the feeding circuit, and 2 mm for the cooling circuit.

The artificial dataset for part 1 comprises 90 points, each obtained from individual simulations. Process parameters were systematically varied within these simulations using a Latin hypercube sampling (LHS) design of experiments. For this study, five key process parameters were considered: mold temperature (T_{mold}), injection speed (V_{inj}), packing pressure (P_h), switch-over point (so), and nozzle temperature (T_{no}). These parameters were selected based on operators' frequent variation during the iterative machine setup process. In Figure 4, the mesh morphology in the region where the weld line formed can be observed.

The LHS experimental design was chosen because it distributes sampling points almost randomly within the domain of individual variables. This approach minimizes overlap between experimental conditions, improves sampling efficiency, and reduces correlations among multidimensional variables [23,44].

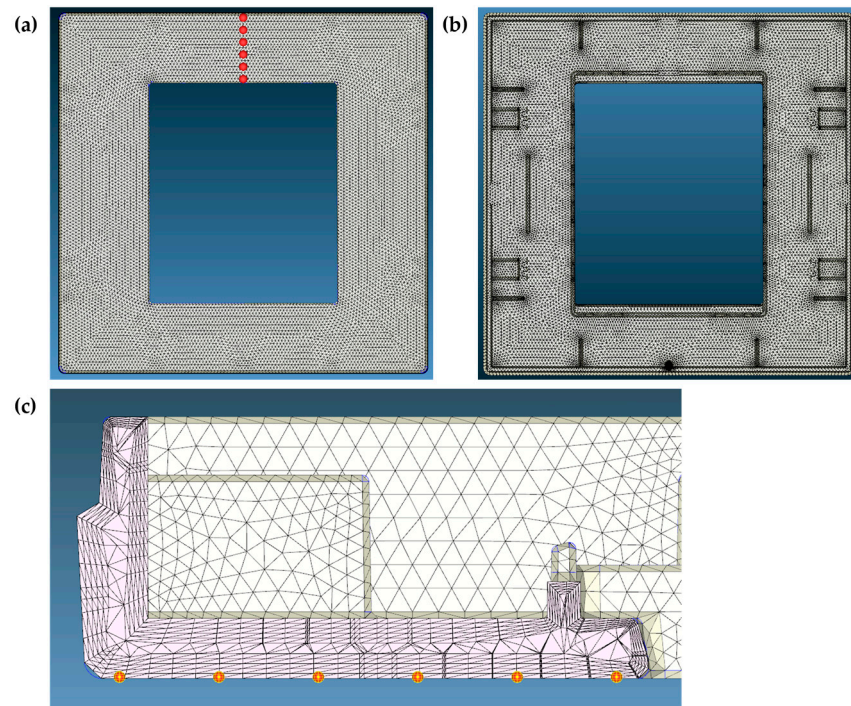


Figure 4. Representation of the Part 1 mesh: (a) global view of the mesh and probe points, (b) global view of the back of the component, (c) detailed view of the mesh in the defect formation area.

Table 1 displays the range of socket process parameters explored through simulations and the experimental campaign.

Table 1. Range of plate’s process parameters investigated through simulations and experimental campaign.

Part 1 Process Parameters		
	Minimum	Maximum
Mold temperature [°C]	90	130
Injection speed [mm/s]	5	100
Switch over point [mm]	5	14
Holding pressure [bar]	800	1300
Nozzle temperature [°C]	250	320

2.3. Weld Lines Visibility

As mentioned before, the visibility of surface defects is not a direct output of the simulations of the injection molding process. To perform the knowledge transfer from synthetic data (obtained from simulations), it is necessary to establish a correlation between one of the output variables of the simulations and the experimental observations. After some attempts, it was decided to use the frozen layer ratio (FLR) to evaluate the visibility of the defects. The FLR was sampled along the entire weld line using six probe points (Figure 4) at the end of filling time (when the weld line forms). Therefore, each simulation was assigned a FLR value equal to the average of the six sampled values. To correlate simulations with experimental observations, five tests were conducted on the injection molding machine, with process parameters varied randomly, followed by numerical simulations of these tests. Utilizing FLR values obtained from these simulations and the five molded parts, it was possible to establish a threshold FLR value that effectively distinguished points with defects from those without defects. Achieving 100% accuracy in this classification was not mandatory, as this dataset was solely intended for initializing the weights of the ANNs.

The threshold value of the FLR was then compared with the FLR value associated with each simulation. If the sampled FLR exceeded the threshold value, the weld line was

deemed visible, and the simulation was assigned to class 1. Conversely, if the FLR was equal to or lower than the threshold value, the defect was considered not visible, and the simulation was assigned to class 0.

2.4. Molding Trials

For both part 1 and part 2, experimental datasets were compiled by replicating the molding process. Throughout the experimental campaigns, the five process parameters were systematically varied using a Latin hypercube sampling (LHS) design of experiment. In total, 140 experimental points were gathered for part 2, and 200 for part 1. All experimental tests were performed on a hydraulic injection molding machine (SynErgy 800.230, Netstal, Näfels Glarus, Switzerland). Table 2 shows the range of variation in the input parameters for part 2; some differences from the values of part 1 can be noted (the process parameters for component 1 are reported in Table 1).

Table 2. Range of cover's process parameters investigated through molding trials.

Part 2 Process Parameters		
	Minimum	Maximum
Mold temperature [°C]	60	120
Injection speed [mm/s]	20	100
Switch over point [mm]	4	18
Holding pressure [bar]	700	1200
Nozzle temperature [°C]	260	330

For every experimental point, two samples were obtained to ensure consistency in the molding cycle's output. Between two consecutive sampling points, the molding process was repeated three times to validate the actual variation in the process parameters. Each collected sample underwent visual evaluation by a machine operator and a quality control officer. The validation process involved a visual check, to be carried out under various lighting conditions. It was decided to use this classification methodology in accordance with the parts manufacturer's acceptability standards. Components exhibiting weld lines were categorized as class 1, while defect-free components were assigned to class 0.

2.5. Transfer Learning Approach

As mentioned before, the cost and difficulty of obtaining training datasets limit the application of machine learning in injection molding and, more generally, in industrial production. Transfer learning proposes a possible solution for this problem. At the core of transfer learning is initializing neural network weights through pre-training on a source dataset. Subsequently, a second training phase is conducted on the target dataset to fine-tune the neural network's weights. The result of the pre-training brings the model closer to the optimum, thus making the ANN converge faster to the optimum during the transfer. Transfer learning demonstrates its utility by reducing the amount of data needed to develop a model for the target dataset, and enhancing the model's overall performance.

This paper investigates the feasibility of employing transfer learning to develop a machine learning model capable of predicting the visibility of weld lines based on five process parameters, which serve as input features for the artificial neural network.

The aim is to evaluate the potential for transferring knowledge between datasets derived from molding processes of different components, specifically from part 1 to part 2. The source dataset comprises data obtained from both simulations and experimental tests, while the target dataset consists solely of data obtained from experimental tests. Section 2.2, Section 2.3 and Section 2.4 provide detailed descriptions of the data collection processes. In Figure 5, a flowchart describing the transfer learning process can be observed.

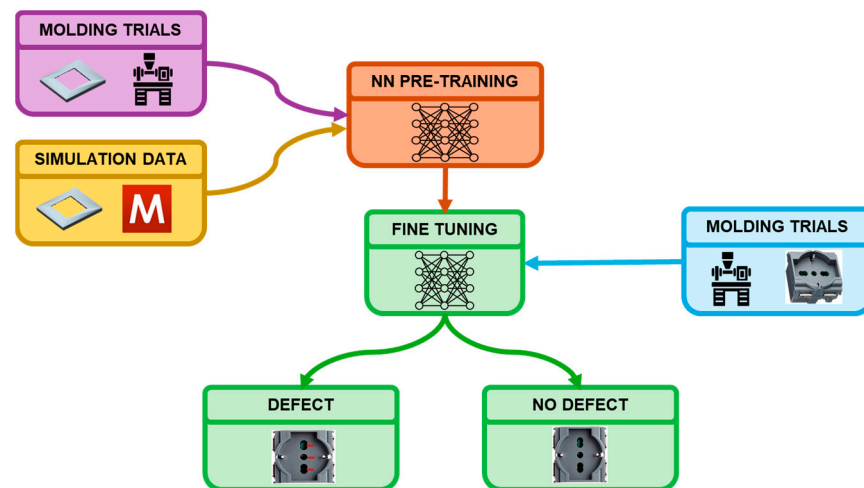


Figure 5. Flowchart describing the learning transfer: each step schematically shows how each dataset was obtained and the component to which it refers.

In this work, two different approaches to TL will be compared [39]. The first TL approach is called “soft start” (SS). It involves the initialization of the internal parameters of the network through training on the source dataset. Then, all the internal parameters of the network are updated through a second training (also called “fine-tuning”) on the target dataset. It is essential to specify that during the transfer learning, the structure of the neural network is not modified in any way.

The second mode of transfer is called “random initialization” (RI), and involves the pre-training of a neural network consisting of a single hidden layer (HL) on the source dataset. After the first training, the output layer is removed, and a series of HLs with randomly initialized weights are added to the network. The weights of the network’s transferred HL thus obtained (initialized with the first training) are kept non-trainable. Then, a second training is performed on the target dataset. It is essential to acknowledge that utilizing RI-type TL, which involves randomly initializing the weights of layers added after the pre-training, may initially yield poorer performance compared to the SS model. However, it is hoped that the new weights, combined with the knowledge transfer operated through the first HL, can lead the model to a different, possibly better, minimum of the loss function than previously found.

Finally, one last network was trained using only the data related to the target dataset. This network (called the control network or CN) served as a reference to evaluate the effects of transfer learning.

The training process was repeated by progressively increasing the number of data points related to the target dataset considered in the knowledge transfer. At each iteration, the performance of the network was evaluated using three metrics: accuracy, AUC, and recall. The use of the three metrics was necessary because all of the datasets were unbalanced in favor of class 1, and therefore, the use of accuracy alone could lead to misleading results. At each iteration, a five-fold cross-validation was conducted to assess the reproducibility of the results and subsequently compute the standard deviation of the various metrics. During the multiple iterations of the validation process (as the number of data used for fine-tuning varied), no changes were made in the structure or hyperparameters of the ANN. During the five-fold cross-validation operation, the process of dividing the data into train sets and test sets was repeated five times, varying the data in each set each time (at each iteration, data not used for training are used for validation). By comparing the results of the forecasts and the class assigned to each piece of data, it was possible to calculate the metrics.

2.6. Hyperparameter Tuning

To optimize the performance of the neural network, before applying the concepts of TL, it was necessary to determine the hyperparameters of the ANN. The determination of the hyperparameters was carried out in two successive steps. In the first step, the morphology of the network was determined, i.e., the number of hidden layers and the number of neurons per HL. This first objective was achieved by observing the curves representing the evolution of the loss function and the accuracy during the training (varying the number of training epochs). For both the loss function and the accuracy, the curves relative to the training and test datasets were generated; the position and the relative trend between these curves provided essential information about the training progress. For every morphology assessed, the curves were observed five times, with the distribution of data between the training set and test set varied on each occasion. It was decided to proceed this way to determine the network architecture that minimized the risk of overfitting or underfitting. The appearance of one of the previous two phenomena could have compromised the effectiveness of the knowledge transfer. During the second step of the hyperparameter optimization, the remaining hyperparameters were determined by performing a Gridsearch [45].

All implementations were performed in Python (version 3.10.9), utilizing the Python modules TensorFlow (version 2.12.0), Keras (version 2.12.0), and Scikit-Learn (version 1.2.1).

3. Results and Discussion

3.1. FLR Threshold

After completing the data collection, the threshold value of FLR was determined using the procedures described in Section 2.3. Table 3 displays the outcome of this process. The threshold was found to be 5%. The threshold value was used to split the simulations into classes “0” and “1”. The data thus obtained were used exclusively in the pre-training phase. The fine-tuning phase followed this first training, so the threshold value did not necessarily have to divide the collected data perfectly.

Table 3. Experimental tests used to find the threshold value of the FLR.

FLR Threshold Value						
T_mold	T_no	Vinj	so	Ph	FLR	Weld Line Presence
[°C]	[°C]	[mm/s]	[mm]	[bar]	[%]	
130	305	30	6.5	700	3.69	0
125	285	20	6.5	700	4.10	1
120	305	45	4.0	900	4.69	0
119	261	87	9.5	500	5.11	1
115	301	22	4.5	849	5.43	1

FLR was chosen for this classification as it is not solely influenced by mold temperature. Table 3 demonstrates how similar mold temperature values corresponded to sufficiently different FLR values. This is important because FLR allowed for the simultaneous consideration of the effect of multiple process parameters.

3.2. Hyperparameter Optimization

The optimization of the internal parameters was repeated for the transfer learning from simulation data and the transfer learning from experimental data. An artificial neural network with 4HL with four neurons each was chosen in both cases. The remaining hyperparameters are visible in Tables 4 and 5.

It is worth noting that in RI-type transfer learning, the network undergoes pre-training with a single hidden layer. Subsequently, the output layer is replaced with three new hidden layers and a new output layer, all initialized with random weights. Consequently, the network structure consistently comprised four hidden layers, each housing four neurons. Finally, before the knowledge transfer, an early-stopping callback was added to the ANN in all models to mitigate the risk of overfitting even further.

Table 4. Result of hyperparameter optimization for transfer learning from the synthetic dataset.

Hyperparameters for TL from Simulations	
Hyperparameter	Value
l2 regularizer	0.005
Drop out	0
Hidden Layer activation function	ReLu
Output Layer activation function	Sigmoid
Training epochs	500
Batch Size	4
Optimizer	Adam
Learning rate	E-3
Loss Function	Binary cross entropy

Table 5. Result of hyperparameter optimization for transfer learning from the synthetic dataset.

Hyperparameters for TL from Experimental Data	
Hyperparameter	Value
l2 regularizer	0.001
Drop out	0
Hidden Layer activation function	ReLu
Output Layer activation function	Sigmoid
Training epochs	750
Batch Size	2
Optimizer	Adam
Learning rate	E-3
Loss Function	Binary cross entropy

3.3. Transfer Learning from Experimental Data

In this paragraph, TL is implemented using experimental data relating to component 1 for the source dataset and experimental data relating to part 2 for the target dataset.

Figure 6 illustrates the progression of predictive metrics concerning the number of experimental data points associated with component 2 utilized during the fine-tuning phase. It is essential to note that these curves depict the trend of the mean metric values.

Analyzing the trend of the control network, it can be affirmed that its capability to achieve satisfactory predictive performance is evident when utilizing a large amount of data during the fine-tuning process. Consequently, the challenge of forecasting the weld line’s visibility based on process parameters is solvable through neural networks.

Figure 6 also presents the curves related to the TL models implemented with soft start and random initialization modes. In this case, pre-training was carried out using experimental data (obtained from machine tests) related to part 1. It is noticeable how the curves related to the transferred models were consistently positioned above the curve related to the CN. Specifically, the CN achieved predictive capacity similar to TL models only when 60 data points from the target dataset were utilized. Therefore, implementing TL in both forms allowed obtaining a predictive model for component 2 using fewer data points than a training mode without TL.

The trend of the recall curve is particularly significant. Maximizing recall entails minimizing false negatives, thus points erroneously classified as defect-free. Having a high recall value helps reduce the risk of producing and potentially introducing components with poor aesthetic characteristics into the market. If the number of transfer data points is less than 60, TL models consistently exhibit higher recall values than the CN. The CN reaches the recall value of TL models only when utilizing 60 data points from the target dataset.

When comparing the curves of the transferred models, it is observed that both transfer methods consistently produced accuracy and AUC values that closely resembled each other. Additionally, employing the soft start transfer method conferred notable advantages regarding recall, mainly when a limited number of experimental points were utilized for

fine-tuning. Nonetheless, compared with the CN, the advantage of TL models diminished as the number of data points used in the transfer phase increased.

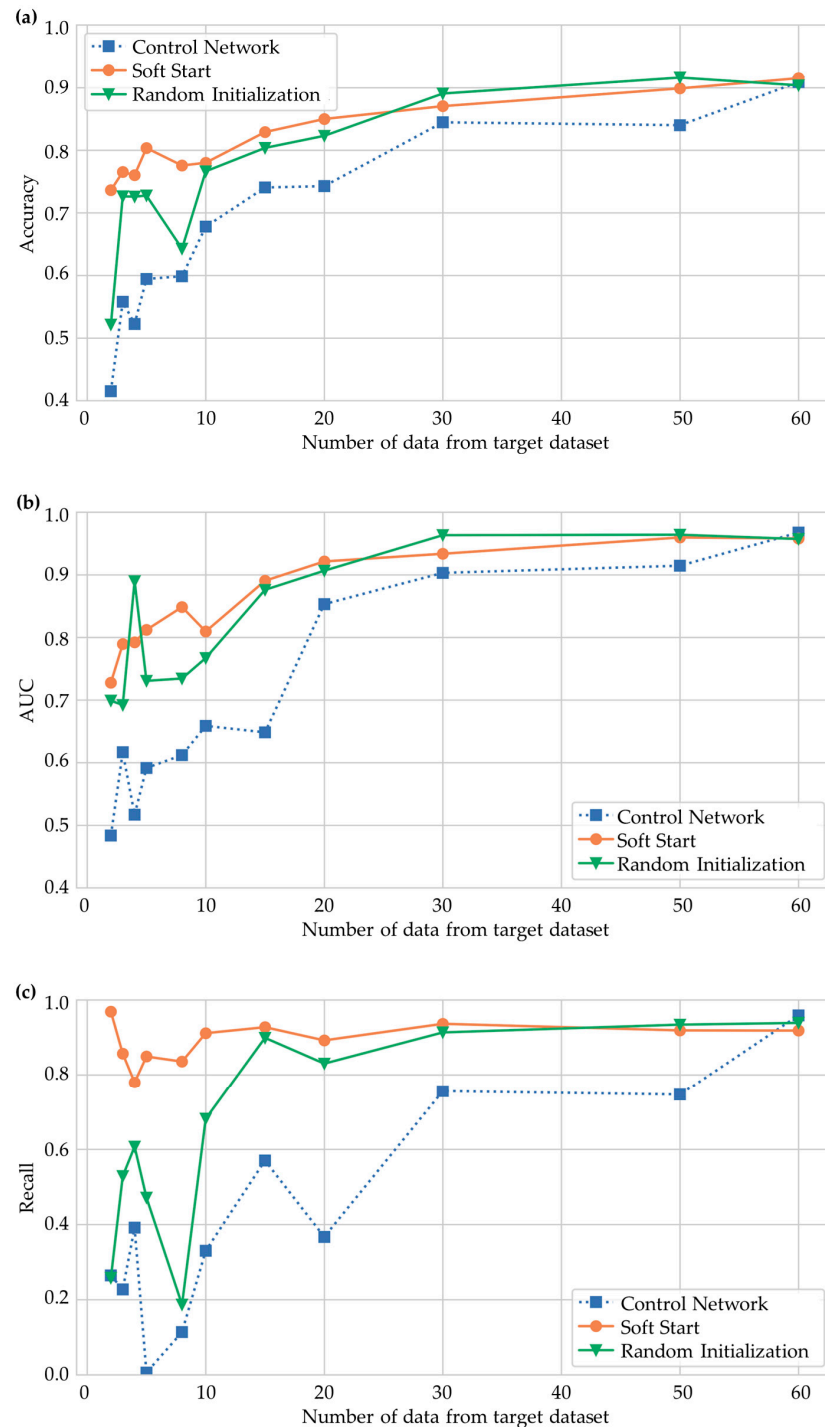


Figure 6. Result of transfer learning between experimental data of part 1 and experimental data of part 2. Evolution of (a) accuracy, (b) AUC and (c) recall as the number of experimental data used in fine-tuning training.

The results of this initial analysis thus confirmed the effectiveness of the TL soft start, as previously observed by Tercan et al. [39]. However, unlike the previous case, in this scenario, the RI-type TL proved superior to the CN (in the study by Tercan et al. [39], the RI model did not lead to any advantages over the non-pre-trained network).

From the previous observations, it is possible to notice how TL can reduce the data necessary to achieve an acceptable predictive capacity for the ANN related to component 2. The final performances achieved using 60 transfer data points for the three trained models (CN, RI, and SS) were similar, indicating that TL can effectively reduce the required data volume without compromising the ANN’s performance.

In conclusion of this section, the repeatability of the results will be observed. To this end, Figure 7 shows a box plot comparing the CN and the SS model (the box plot results from evaluating the metrics as described in Section 2.5).

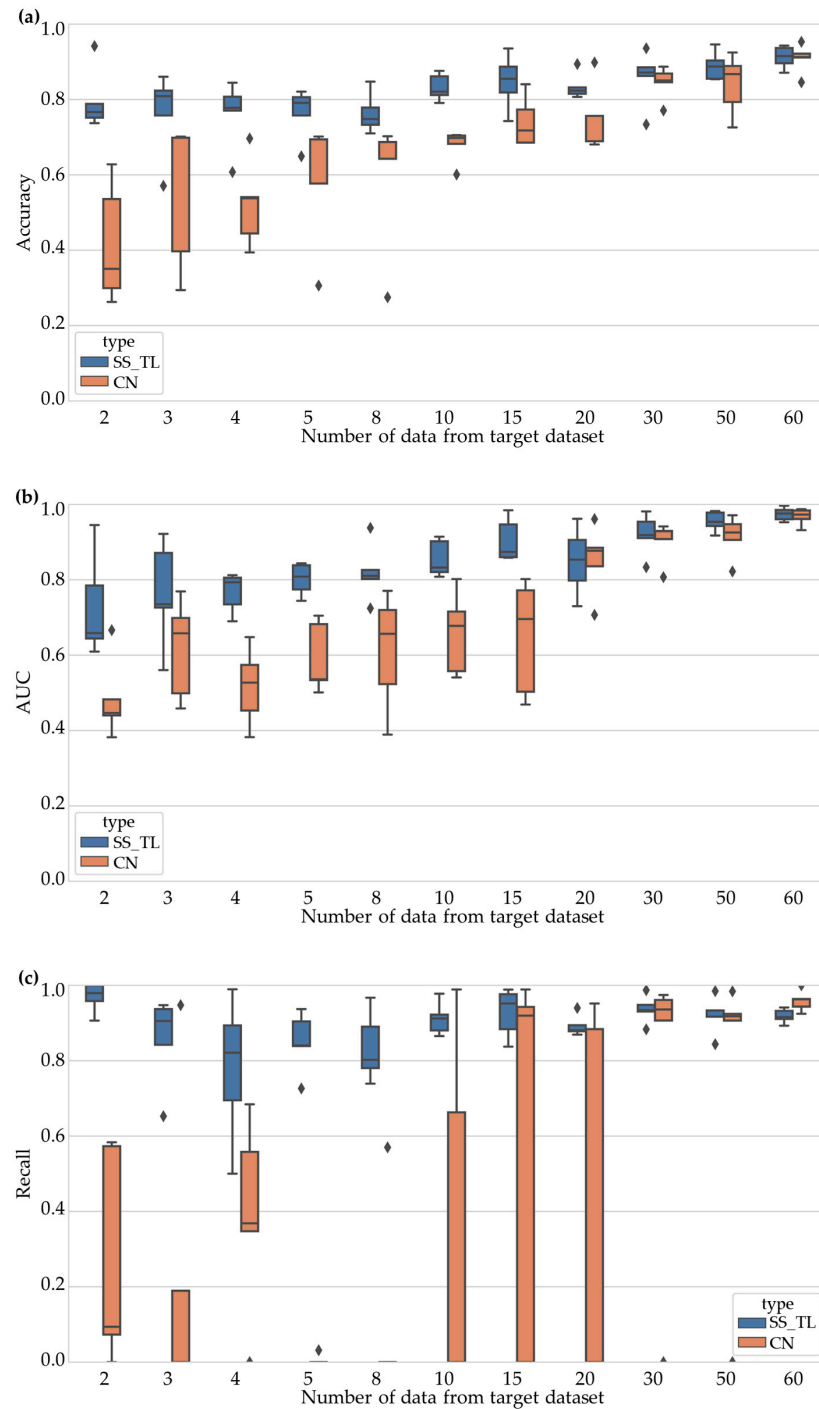


Figure 7. Comparison between CN and the SS model obtained from experimental data (SS_TL). Evolution of (a) accuracy, (b) AUC, and (c) recall as the number of experimental data used in fine-tuning training increased.

From Figure 7, it is evident that the average performance of the pre-trained model was superior to that of the CN, and the variability of the results was lower in the pre-trained model. This observation becomes particularly pronounced when considering the AUC and Recall metrics. These metrics showed high variability due to the inconsistent training outcomes of the control network. Specifically, when the number of transfer data points was limited, the control network produced a constant output equal to the majority class. This compromised the value of recall and AUC and underscored an additional advantage of TL application.

The present study confirmed the feasibility of employing TL to diminish the experimental data required for training predictive models in injection molding [39–42]. This introduces the potential to utilize ML and TL techniques to forecast the occurrence of aesthetic defects.

3.4. Transfer Learning from Simulation Data

The effect of replacing experimental data related to part 1 with simulations of the molding process of part 1 was also evaluated. The data from the target dataset remained unchanged from the previous case (experimental data related to part 2). In Figure 8, the average value of the metrics is again represented as a function of the number of data points used for knowledge transfer.

It is worth noting that all the observations made in the previous paragraph remain valid. The advantage of the SS model over the RI model became apparent when a low number of data were used in the transfer.

The best models derived from transfer learning using experimental data and synthetic data were compared. The comparison was conducted through a boxplot, depicted in Figure 9. The boxplot also describes the evolution of the metrics as the number of data points used in the transfer increased.

From Figure 9, it is evident that both SS models exhibited comparable performances. Nonetheless, it is worth highlighting that transferring from simulation results showed an overall improvement in the model's mean performance, particularly in terms of accuracy and AUC, while also reducing the variability of the metrics. The cause of this difference was presumed to stem from the greater generality of synthetic data. The imperfect classification conducted through the FLR (Section 2.3) rendered the synthetic dataset more easily adaptable than the experimental dataset. Additionally, the presence of outliers or points incorrectly classified by the operators in the experimental source dataset led the TL from experimental data to require more data points to adapt to the new domain. It is also important to point out that the slight difference between the hyperparameters of the two models could have contributed to this variation in performance.

The comparison between the two soft start TL methods, along with the results presented in Figure 7, allows for extending the comparison between CN and SS using experimental data to the comparison between CN and SS using simulations.

Finally, in Figure 6, it can be observed that the performance of the SS model pre-trained on synthetic data stabilized around 20 data points, whereas the control network achieved a comparable performance to this model only when approximately 50 data points from part 2 were utilized. On the other hand, from Figure 8, it can be seen that the SS model pre-trained on synthetic data stabilized its performance using only 10 data points for fine-tuning. In contrast, the control network matched the performance of this model only when employing 60 data points for knowledge transfer. The recent observations indicated a 60% reduction in the amount of part 2 data needed to train the neural network through transfer learning from simulation. Conversely, transfer learning from simulation extended this reduction to 83%. Transfer learning from simulation minimized the data points required for good forecasting capacity. Furthermore, this approach maximized cost efficiency compared to traditional training methods. The final performances of both SS models, achieved with 30 or more data points from part 2, were nearly identical. In Figures 10 and 11, the confusion matrices for the TL from experimental and artificial data, respectively, are reported. For each cell of

the matrix, an example of a component relating to the cell is shown. The matrices relate to fine-tuning with 60 data relating to part 2.

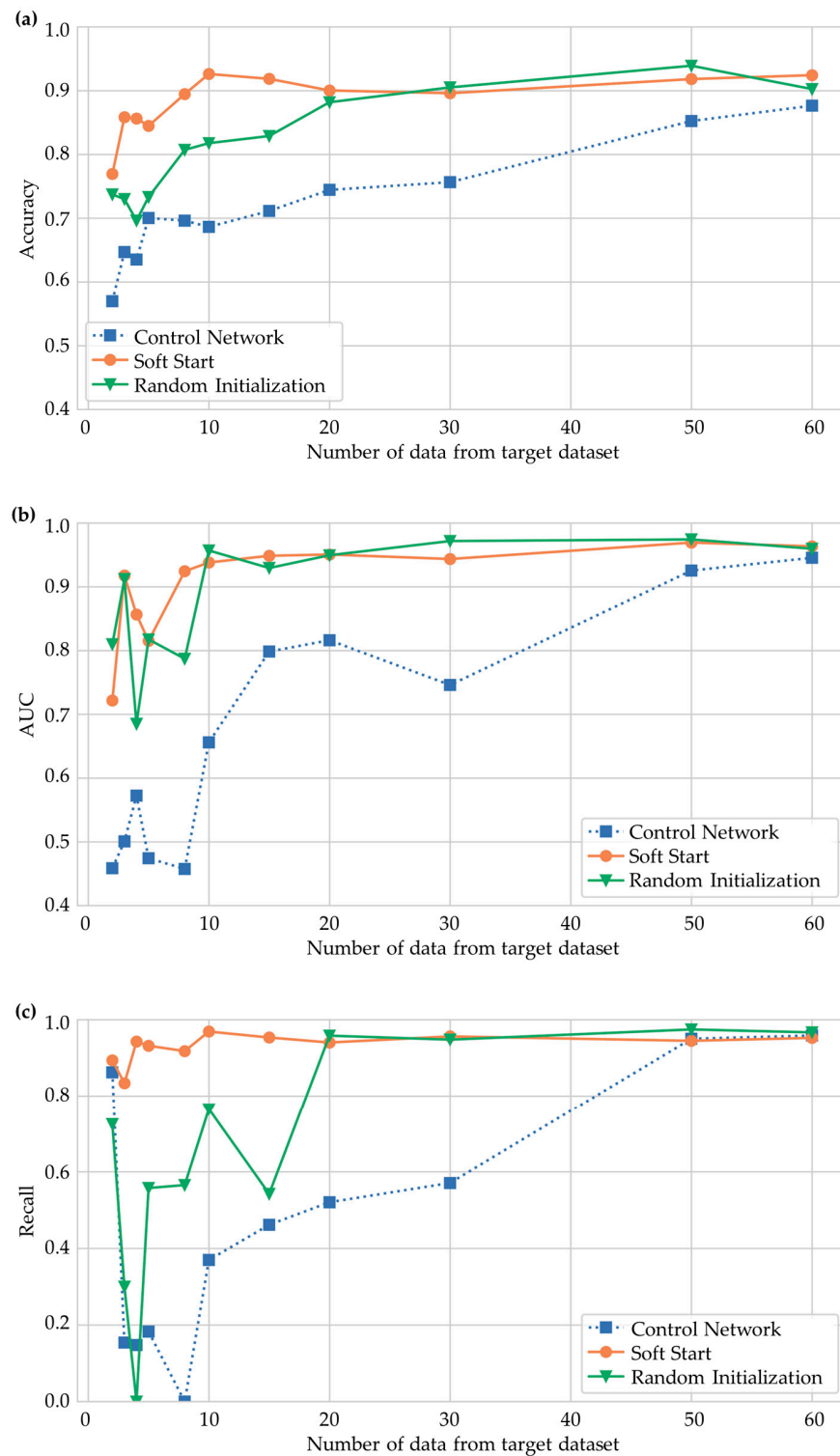


Figure 8. Result of transfer learning between synthetic data of part 1 and experimental data of part 2. Evolution of (a) accuracy, (b) AUC, and (c) recall as the number of experimental data used in fine-tuning training increased.

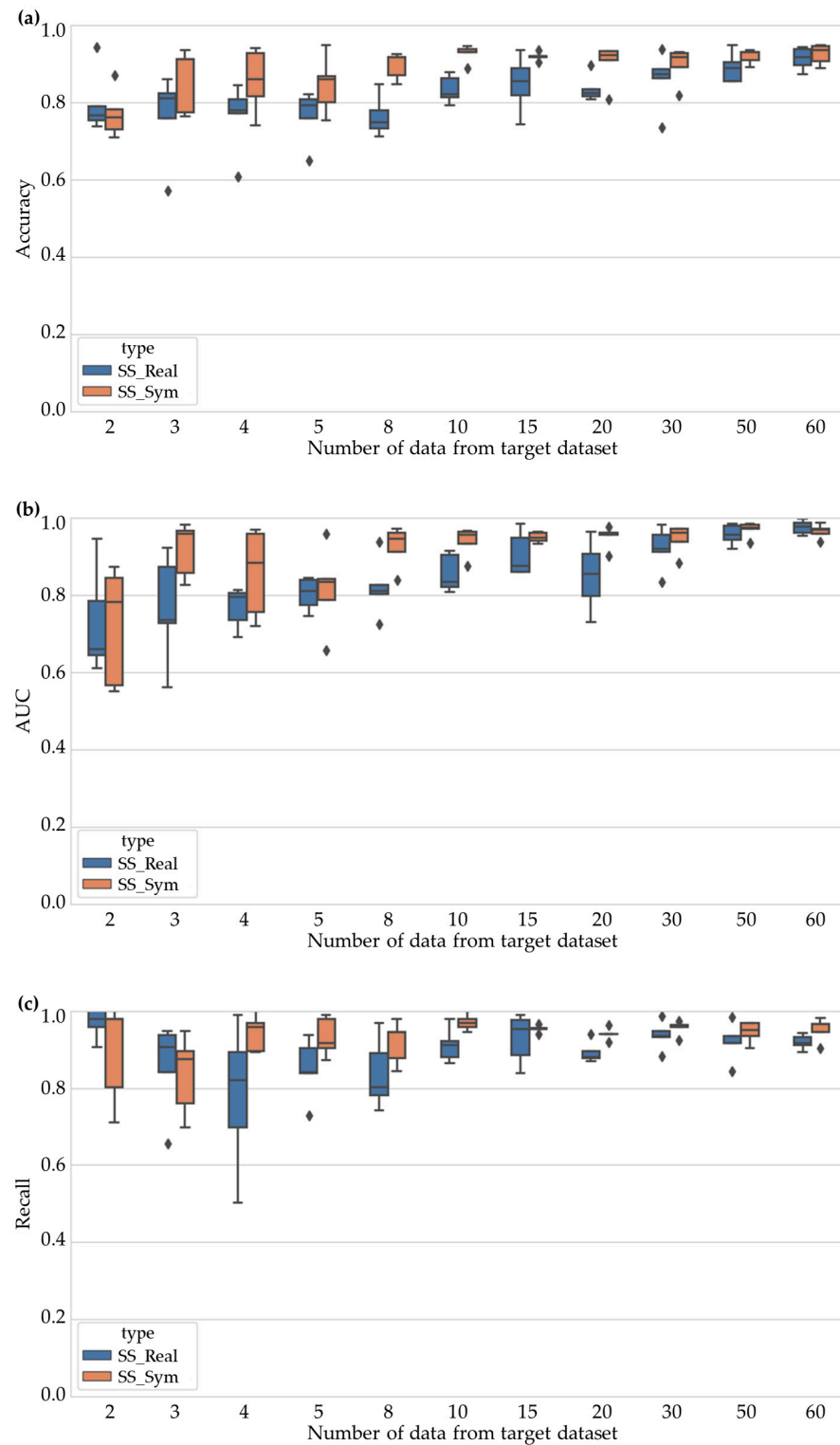


Figure 9. Comparison between the SS model obtained from simulations (SS_Sym) and the SS model obtained from experimental data (SS_Real). Evolution of (a) accuracy, (b) AUC, and (c) recall as the number of experimental data used in fine-tuning training increased.

Implementing transfer learning from a synthetic dataset has led to results comparable, if not superior, to TL from experimental data. This success is primarily attributed to the correlation established between simulation results and experimental observations during the data processing (Section 2.3). However, it is essential to acknowledge that correlating experimental data with simulations requires the evaluation of five samples

related to component 1, underscoring that generating synthetic datasets also involves experimental endeavors.

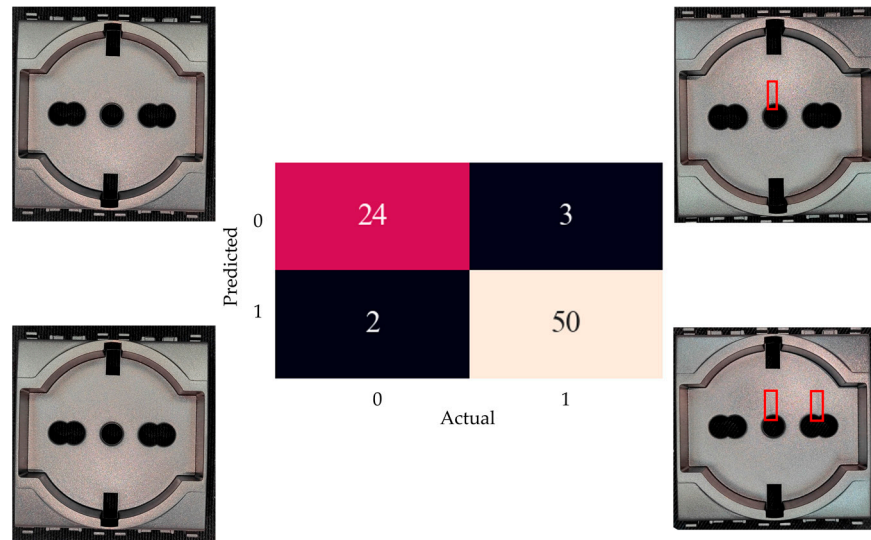


Figure 10. Confusion matrix related to transfer learning from experimental data (60 data points used in the fine-tuning phase).

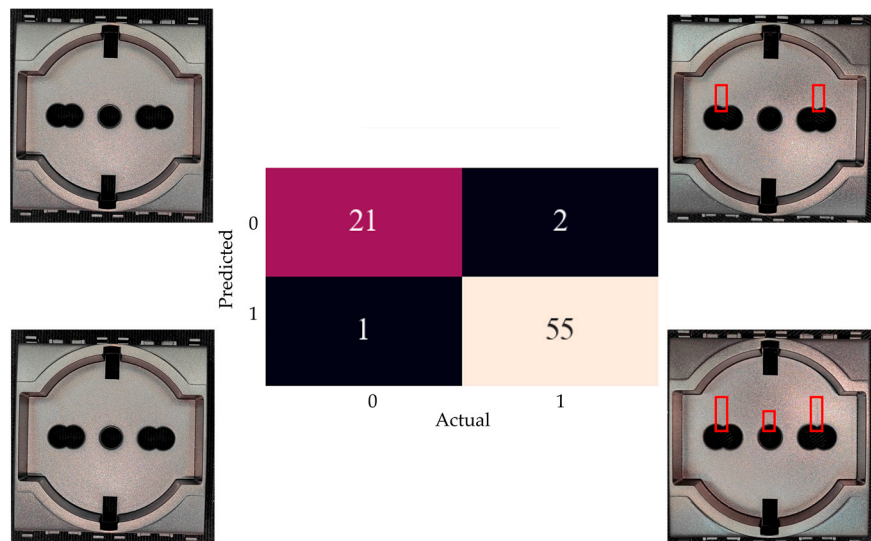


Figure 11. Confusion matrix related to transfer learning from synthetic data (60 data points used in the fine-tuning phase).

Therefore, this current work extends the findings of Gim et al. [42] by introducing the possibility of employing simulations to pre-train predictive models capable of assessing the occurrence of aesthetic defects in components manufactured through injection molding. This enables further reductions in the costs associated with obtaining predictive models.

3.5. Effect of Geometry on TL

Finally, the effect of geometry on transfer learning was evaluated. To do this, the TL between simulations of the molding process of component 1 and the experimental data of component 2 was compared with the TL between both simulations and experimental data from component 1. It is important to note that, for the subsequent comparison, the models were retrained from scratch. To mitigate the influence of the target dataset size (as the number of experimental data varied between part 1 and part 2), the comparison

was conducted by isolating 100 experimental points from each dataset. The results of this comparison are described by the boxplot shown in Figure 12. Again, the comparison was made by observing the evolution of the metrics as the number of data used in the transfer increased.

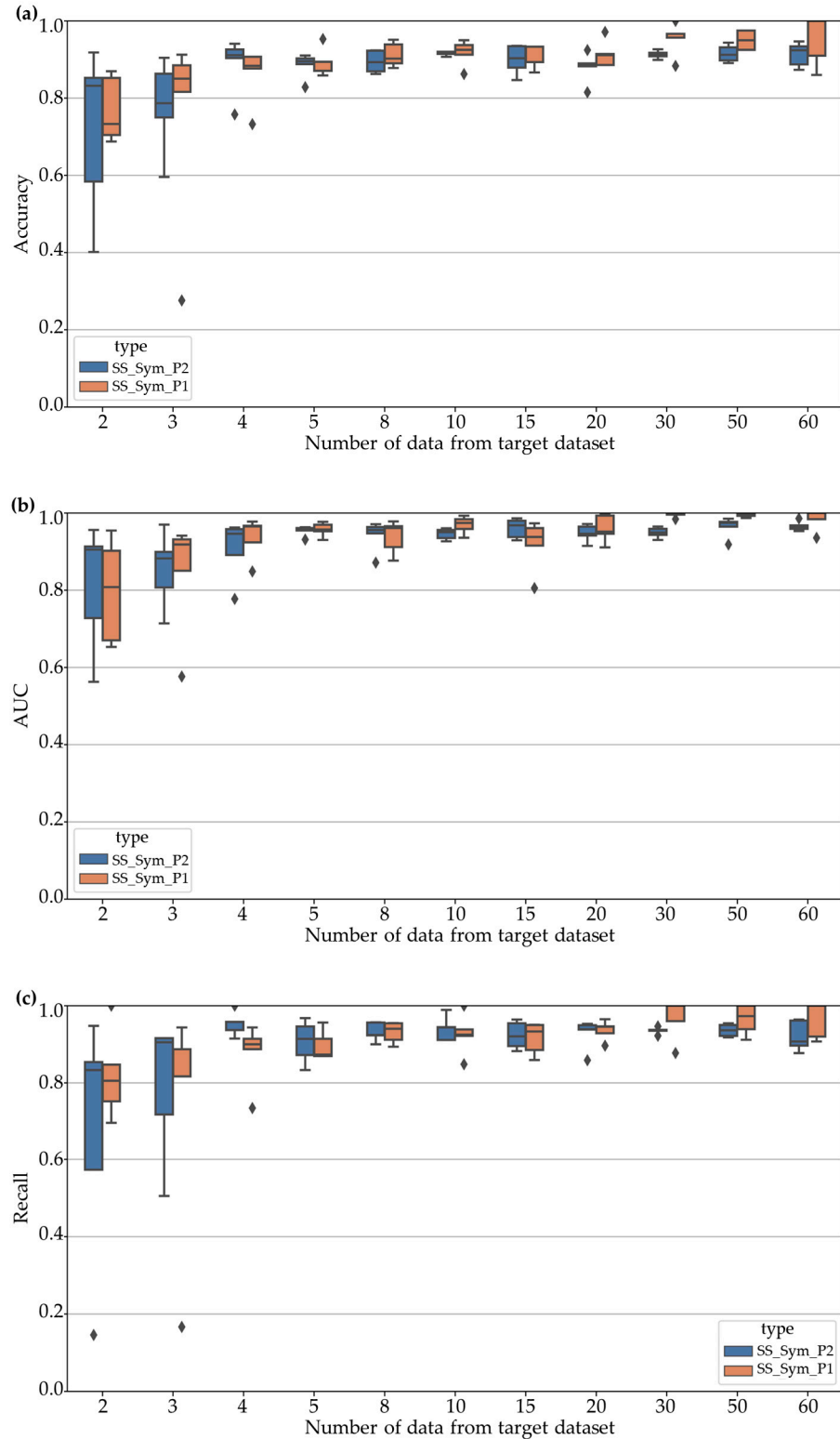


Figure 12. Comparison between the SS model related to part 1 (SS_sym_P1) and the SS model about part 2 (SS_sym_P2). Evolution of (a) accuracy, (b) AUC, and (c) recall as the number of experimental data used in fine-tuning training increased.

In Figure 12, it can be observed that the two transferred models (both using SS) exhibited very similar mean values and variability of metrics. It is noticeable that the model trained with simulations and data from part 1 achieved slightly higher mean values of the metrics than the model transferred to component 2. Additionally, the variability range of the metrics for the model related to component 2 was higher, especially when using limited data points in the transfer, compared to what was observed in the model related to part 1. This observation can be attributed to the lesser similarity between the experimental dataset of part 2 and the synthetic dataset, as opposed to the experimental dataset of component 1. Therefore, the SS related to part 2 will require more data to bridge the gap between the source and target datasets. The slight differences between these two transferred models are acceptable, given the significant reduction in experimental and computational burden associated with developing the classifier for part 2. In fact, in implementing the second model, there was no need to collect new simulations; instead, TL allowed the reuse of data collected for part 1. The application of TL, in all cases, reduced the number of experimental data required to achieve acceptable predictive metric values compared to non-pre-trained models.

4. Conclusions

In conclusion, this study delved into the realm of transfer learning (TL) applied to predictive modeling for injection molding processes, explicitly focusing on forecasting weld line visibility. Through extensive experimentation and analysis, several key findings have emerged, shedding light on the efficacy and implications of TL in this context.

First and foremost, the investigation into TL from experimental data showcased promising results. By leveraging knowledge from a source dataset relating to component 1 and transferring it to predict weld line presence on component 2, a significant improvement in predictive performance compared to traditional training methods was observed. Notably, TL models consistently outperformed the control network, mainly when fine-tuning was conducted with limited data points. It is also important to note how the TL from experimental data reached similar metrics values to the non-pre-trained model, but with 60% less data relating to part 2. This highlights the potential of TL to enhance predictive capabilities while minimizing the data required for training, thus presenting a valuable avenue for cost-effective and efficient modeling within the injection molding domain.

Furthermore, the exploration of TL from simulation data yielded equally compelling insights. By replacing experimental data with molding process simulations, comparable or even superior performance in predictive metrics was seen. Furthermore, the TL from simulations achieved metric values similar to those of the control model, but with 83% less data. The findings suggest that TL from simulation data could be a valuable tool for enhancing predictive capabilities, while mitigating the challenges associated with data collection and processing in real-world industrial settings.

Moreover, the effects of geometry on the model's predictive ability were evaluated. Compared to the model trained using only component 1 data, the model targeting part 2 had lower metrics mean values but higher variability. Nonetheless, TL facilitated the reuse of data collected for component 1, thereby reducing the experimental and computational burden of developing classifiers for subsequent components. This underscores the practical significance of TL in streamlining the modeling process and maximizing cost efficiency in industrial applications.

In summary, this study demonstrated the possibility of using transfer learning to predict the aesthetic characteristics of components made by injection molding. It also showed the possibility of using simulation data to improve the classifiers' performance.

It is reasonable to think that this work can be extended by considering transfer learning between components made of different materials and/or on different injection molding machines. All of this will progressively bring ML closer to application in an industrial context, making it an increasingly competitive alternative for determining the process parameters of injection molding machines.

Author Contributions: Conceptualization, A.P. and G.L.; methodology, A.P.; software, A.P. and G.B.; validation, A.P. and G.B.; formal analysis, A.P. and G.B.; investigation, A.P. and G.B.; resources, G.L.; data curation, A.P. and G.B.; writing—original draft preparation, G.B.; writing—review and editing, G.L., A.P. and M.S.; visualization, G.B. and M.S.; supervision, G.L.; project administration, G.L.; funding acquisition, G.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

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

References

1. Fernandes, C.; Pontes, A.J.; Viana, J.C.; Gaspar-Cunha, A. Modeling and Optimization of the Injection-Molding Process: A Review. *Adv. Polym. Technol.* **2018**, *37*, 429–449. [[CrossRef](#)]
2. Chen, Z.; Turng, L. A Review of Current Developments in Process and Quality Control for Injection Molding. *Adv. Polym. Technol. J. Polym. Process. Inst.* **2005**, *24*, 165–182. [[CrossRef](#)]
3. Ozelcik, B.; Erzurumlu, T. Comparison of the Warpage Optimization in the Plastic Injection Molding Using ANOVA, Neural Network Model and Genetic Algorithm. *J. Mater. Process Technol.* **2006**, *171*, 437–445. [[CrossRef](#)]
4. Steenkamp, J.-B.E.M. Conceptual Model of the Quality Perception Process. *J. Bus. Res.* **1990**, *21*, 309–333. [[CrossRef](#)]
5. Linke, B.; Das, J. Aesthetics and Gloss of Ground Surfaces: A Review on Measurement and Generation. *J. Manuf. Sci. Eng.* **2016**, *138*, 064501. [[CrossRef](#)]
6. Toensmeier, P. Auto Color Trends Reveal a Bright Future: Understanding Color Trends and How They Apply to Vehicle Design Gives Processors a Big Advantage in Automotive Supply. *Plast. Eng.* **2013**, *69*, 18–22. [[CrossRef](#)]
7. Xia, N.; Zhao, P.; Xie, J.; Zhang, C.; Fu, J.; Turng, L.-S. Defect Diagnosis for Polymeric Samples via Magnetic Levitation. *NDT E Int.* **2018**, *100*, 175–182. [[CrossRef](#)]
8. Zhang, J.; Zhao, P.; Zhao, Y.; Huang, J.; Xia, N.; Fu, J. On-Line Measurement of Cavity Pressure during Injection Molding via Ultrasonic Investigation of Tie Bar. *Sens. Actuators A Phys.* **2019**, *285*, 118–126. [[CrossRef](#)]
9. Fernandes, C.; Pontes, A.J.; Viana, J.C.; Nóbrega, J.M.; Gaspar-Cunha, A. Modeling of Plasticating Injection Molding—Experimental Assessment. *Int. Polym. Process.* **2014**, *29*, 558–569. [[CrossRef](#)]
10. Zhao, Y.; Zhao, P.; Zhang, J.; Huang, J.; Xia, N.; Fu, J. On-Line Measurement of Clamping Force for Injection Molding Machine Using Ultrasonic Technology. *Ultrasonics* **2019**, *91*, 170–179. [[CrossRef](#)]
11. Xia, N.; Zhao, P.; Kuang, T.; Zhao, Y.; Zhang, J.; Fu, J. Nondestructive Measurement of Layer Thickness in Water-assisted Coinjection-molded Product by Ultrasonic Technology. *J. Appl. Polym. Sci.* **2018**, *135*, 46540. [[CrossRef](#)]
12. Gao, S.; Qiu, Z.; Ma, Z.; Yang, Y. Flow Properties of Polymer Melt in Longitudinal Ultrasonic-assisted Microinjection Molding. *Polym. Eng. Sci.* **2017**, *57*, 797–805. [[CrossRef](#)]
13. Tsai, K.-M.; Hsieh, C.-Y.; Lo, W.-C. A Study of the Effects of Process Parameters for Injection Molding on Surface Quality of Optical Lenses. *J. Mater. Process Technol.* **2009**, *209*, 3469–3477. [[CrossRef](#)]
14. Zhao, P.; Dong, Z.; Zhang, J.; Zhang, Y.; Cao, M.; Zhu, Z.; Zhou, H.; Fu, J. Optimization of Injection-Molding Process Parameters for Weight Control: Converting Optimization Problem to Classification Problem. *Adv. Polym. Technol.* **2020**, *2020*, 7654249. [[CrossRef](#)]
15. Gao, H.; Zhang, Y.; Zhou, X.; Li, D. Intelligent Methods for the Process Parameter Determination of Plastic Injection Molding. *Front. Mech. Eng.* **2018**, *13*, 85–95. [[CrossRef](#)]
16. Yang, D.; Danai, K.; Kazmer, D. A Knowledge-Based Tuning Method for Injection Molding Machines. *J. Manuf. Sci. Eng.* **2001**, *123*, 682–691. [[CrossRef](#)]
17. Chen, W.-C.; Kurniawan, D. Process Parameters Optimization for Multiple Quality Characteristics in Plastic Injection Molding Using Taguchi Method, BPNN, GA, and Hybrid PSO-GA. *Int. J. Precis. Eng. Manuf.* **2014**, *15*, 1583–1593. [[CrossRef](#)]
18. Zhou, H.; Zhao, P.; Feng, W. An Integrated Intelligent System for Injection Molding Process Determination. *Adv. Polym. Technol. J. Polym. Process. Inst.* **2007**, *26*, 191–205. [[CrossRef](#)]
19. Rai, R.; Tiwari, M.K.; Ivanov, D.; Dolgui, A. Machine Learning in Manufacturing and Industry 4.0 Applications. *Int. J. Prod. Res.* **2021**, *59*, 4773–4778. [[CrossRef](#)]
20. Zhang, J.; Wang, J.; Lin, J.; Guo, Q.; Chen, K.; Ma, L. Multiobjective Optimization of Injection Molding Process Parameters Based on Opt LHD, EBFNN, and MOPSO. *Int. J. Adv. Manuf. Technol.* **2016**, *85*, 2857–2872. [[CrossRef](#)]
21. Tsai, K.-M.; Luo, H.-J. An Inverse Model for Injection Molding of Optical Lens Using Artificial Neural Network Coupled with Genetic Algorithm. *J. Intell. Manuf.* **2017**, *28*, 473–487. [[CrossRef](#)]
22. Ke, K.-C.; Huang, M.-S. Quality Prediction for Injection Molding by Using a Multilayer Perceptron Neural Network. *Polymer* **2020**, *12*, 1812. [[CrossRef](#)] [[PubMed](#)]
23. Chen, W.-C.; Wang, M.-W.; Chen, C.-T.; Fu, G.-L. An Integrated Parameter Optimization System for MISO Plastic Injection Molding. *Int. J. Adv. Manuf. Technol.* **2009**, *44*, 501–511. [[CrossRef](#)]

24. Kenig, S.; Ben-David, A.; Omer, M.; Sadeh, A. Control of Properties in Injection Molding by Neural Networks. *Eng. Appl. Artif. Intell.* **2001**, *14*, 819–823. [[CrossRef](#)]
25. Trovalusci, F.; Ucciardello, N.; Baiocco, G.; Tagliaferri, F. Neural Network Approach to Quality Monitoring of Injection Molding of Photoluminescent Polymers. *Appl. Phys. A* **2019**, *125*, 781. [[CrossRef](#)]
26. Guo, W.; Deng, F.; Meng, Z.; Hua, L.; Mao, H.; Su, J. A Hybrid Back-Propagation Neural Network and Intelligent Algorithm Combined Algorithm for Optimizing Microcellular Foaming Injection Molding Process Parameters. *J. Manuf. Process* **2020**, *50*, 528–538. [[CrossRef](#)]
27. Spina, R. Optimisation of Injection Moulded Parts by Using ANN-PSO Approach. *J. Achiev. Mater. Manuf. Eng.* **2006**, *15*, 146–152.
28. Manjunath, P.G.C.; Krishna, P. Prediction and Optimization of Dimensional Shrinkage Variations in Injection Molded Parts Using Forward and Reverse Mapping of Artificial Neural Networks. *Adv. Mat. Res.* **2012**, *463*, 674–678. [[CrossRef](#)]
29. Sedighi, R.; Meiabadi, M.S.; Sedighi, M. Optimisation of Gate Location Based on Weld Line in Plastic Injection Moulding Using Computer-Aided Engineering, Artificial Neural Network, and Genetic Algorithm. *Int. J. Automot. Mech. Eng.* **2017**, *14*, 4419–4431. [[CrossRef](#)]
30. Bensingh, R.J.; Machavaram, R.; Boopathy, S.R.; Jebaraj, C. Injection Molding Process Optimization of a Bi-Aspheric Lens Using Hybrid Artificial Neural Networks (ANNs) and Particle Swarm Optimization (PSO). *Measurement* **2019**, *134*, 359–374. [[CrossRef](#)]
31. Nagorny, P.; Pillet, M.; Pairel, E.; Le Goff, R.; Loureaux, J.; Wali, M.; Kiener, P. Quality Prediction in Injection Molding. In Proceedings of the 2017 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), Annecy, France, 26–28 June 2017; IEEE: Piscataway, NY, USA, 2017; pp. 141–146.
32. Yarlagadda, P.K.D. V Prediction of Processing Parameters for Injection Moulding by Using a Hybrid Neural Network. *Proc. Inst. Mech. Eng. B J. Eng. Manuf.* **2001**, *215*, 1465–1470. [[CrossRef](#)]
33. Silva, B.; Marques, R.; Faustino, D.; Ilheu, P.; Santos, T.; Sousa, J.; Rocha, A.D. Enhance the Injection Molding Quality Prediction with Artificial Intelligence to Reach Zero-Defect Manufacturing. *Processes* **2022**, *11*, 62. [[CrossRef](#)]
34. Lee, J.; Lee, Y.C.; Kim, J.T. Fault Detection Based on One-Class Deep Learning for Manufacturing Applications Limited to an Imbalanced Database. *J. Manuf. Syst.* **2020**, *57*, 357–366. [[CrossRef](#)]
35. Liu, J.; Guo, F.; Gao, H.; Li, M.; Zhang, Y.; Zhou, H. Defect Detection of Injection Molding Products on Small Datasets Using Transfer Learning. *J. Manuf. Process* **2021**, *70*, 400–413. [[CrossRef](#)]
36. Pan, S.J.; Yang, Q. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2009**, *22*, 1345–1359. [[CrossRef](#)]
37. Bengio, Y. Deep Learning of Representations for Unsupervised and Transfer Learning. In Proceedings of the ICML Workshop on Unsupervised and Transfer Learning, Washington, DC, USA, 2 July 2012; pp. 17–36.
38. Rosenstein, M.T.; Marx, Z.; Kaelbling, L.P.; Dietterich, T.G. To Transfer or Not to Transfer. In Proceedings of the NIPS 2005 Workshop on Transfer Learning, Vancouver, BC, Canada, 9 December 2005; Volume 898.
39. Tercan, H.; Guajardo, A.; Heinisch, J.; Thiele, T.; Hopmann, C.; Meisen, T. Transfer-Learning: Bridging the Gap between Real and Simulation Data for Machine Learning in Injection Molding. *Procedia CIRP* **2018**, *72*, 185–190. [[CrossRef](#)]
40. Lockner, Y.; Hopmann, C.; Zhao, W. Transfer Learning with Artificial Neural Networks between Injection Molding Processes and Different Polymer Materials. *J. Manuf. Process* **2022**, *73*, 395–408. [[CrossRef](#)]
41. Lockner, Y.; Hopmann, C. Induced Network-Based Transfer Learning in Injection Molding for Process Modelling and Optimization with Artificial Neural Networks. *Int. J. Adv. Manuf. Technol.* **2021**, *112*, 3501–3513. [[CrossRef](#)]
42. Gim, J.; Yang, H.; Turng, L.-S. Transfer Learning of Machine Learning Models for Multi-Objective Process Optimization of a Transferred Mold to Ensure Efficient and Robust Injection Molding of High Surface Quality Parts. *J. Manuf. Process* **2023**, *87*, 11–24. [[CrossRef](#)]
43. Loss, A.C.E. *37 on T. Measurements. S.E. 01 on C. and M. Standard Test Method for Assignment of the Glass Transition Temperatures by Differential Scanning Calorimetry*; ASTM International: West Conshohocken, PA, USA, 2014.
44. Choi, Y.; Song, D.; Yoon, S.; Koo, J. Comparison of Factorial and Latin Hypercube Sampling Designs for Meta-Models of Building Heating and Cooling Loads. *Energies* **2021**, *14*, 512. [[CrossRef](#)]
45. Géron, A. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow*; O'Reilly Media, Inc.: Newton, MA, USA, 2022; ISBN 1098122461.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.