Design De-identification of Thermal History for Collaborative Process-defect Modeling of Directed Energy Deposition Processes

Durant Fullington
Department of Industrial and System Engineering, Mississippi State University, Mississippi State, MS 39762, United States
dhf45@msstate.edu

Linkan Bian
Department of Industrial and System Engineering,
Center for Advanced Vehicular Systems (CAVS), Mississippi State University, Mississippi State, MS 39762, United States
bian@ise.msstate.edu

Wenmeng Tian[1]
Department of Industrial and System Engineering,
Center for Advanced Vehicular Systems (CAVS), Mississippi State University, Mississippi State, MS 39762, United States
tian@ise.msstate.edu

Abstract

There is an urgent need for developing collaborative process-defect modeling in metal-based additive manufacturing (AM). This mainly stems from the high volume of training data needed to develop reliable machine learning models for anomaly detection. The requirements for large data are especially challenging for small-to-medium-manufacturers (SMMs), for whom collecting copious amounts of data is usually cost prohibitive. The objective of this research is to develop a secured data sharing mechanism for Directed Energy Deposition (DED) based AM without disclosing product design information, facilitating secured data aggregation for collaborative modeling. However, one major obstacle is the privacy concerns that arise from data sharing, since AM process data contains confidential design information. The proposed Adaptive Design De-identification for Additive Manufacturing (ADDAM) methodology integrates AM process knowledge into an adaptive de-identification procedure to mask the printing trajectory information in AM

[1] Corresponding author

thermal history, which otherwise discloses major product design information. This adaptive approach applies a flexible data privacy level to each thermal image based on its similarity with the other images, facilitating better data utility preservation while protecting data privacy. A real-world case study was used to validate the proposed method based on the fabrication of two cylindrical parts using a DED process. These results are expressed as a pareto optimal solution, demonstrating significant improvements in privacy gain and minimal utility loss. The proposed method can facilitate privacy improvements of up to 30% with as little as 0% losses in dataset usability after de-identification.

Keywords: Additive manufacturing, data privacy, data sharing, de-identification, directed energy deposition, in-situ anomaly detection, thermal history.

1. Introduction

One of the biggest limitations in the broader adoption of Directed Energy Deposition (DED) based additive manufacturing (AM) techniques is the *in-situ* defect detection for part certification. It is crucial for users to detect process anomalies in an effective and timely manner since the offline counterpart methods have proven costly and time-consuming [1]–[4]. Machine learning and artificial intelligence has played a crucial role in the development of *in-situ* anomaly detection models for AM [4]–[7]. However, due to the high part complexity, the highly variable part designs and printing parameters, building a robust machine learning model for *in-situ* process monitoring requires large amounts of training data, which can be prohibitively expensive [4], [5], [8]. Recently, the AM research community has identified these obstacles as a serious roadblock for the accelerated adoption of AM, especially for those small-to-medium sized manufacturers (SMMs) [4]–[6].

One potential solution is to facilitate data sharing through the direct aggregation of process data from multiple AM users [9], [10]. The idea of data sharing has been proposed as an important tool to expand AM technologies [7] and several publications also see it as a remedy to limited data availability plaguing SMMs [5], [6]. The aggregated training data can then be leveraged to develop a more accurate, robust, and generalizable machine learning model for anomaly detection. Furthermore, these models would require less training data from each user than traditional independent machine learning models [5], [11]. This is especially helpful for SMMs, as it will decrease the amount of data required from each user and tackles one of the discussed challenges for integrating machine learning with AM [4]–[6].

Unfortunately, the major obstacle in aggregating process data from multiple AM users is the data privacy concerns that arise from sharing process data outside of the user's organization. This key drawback is a highly discussed limitation and forms one of the major gaps in the development and implementation of AM data sharing frameworks [5], [7]. In AM, the process data contains critical product design information, which heavily involves the intellectual property (IP) of the individual user. Sharing these data outside the user's organization can potentially expose the AM users to the risk of IP theft. This could occur when a malicious third-party gains access to the shared data and can reverse engineer the AM design specifications, utilizing the printing path and other parameters derived from the AM process data. What is worse, AM is typically used in new product prototyping due to its toolless and flexible fabrication for accelerated design iterations. Therefore, the risk of IP theft in AM process data can be even more detrimental to AM practitioners, especially SMM users.

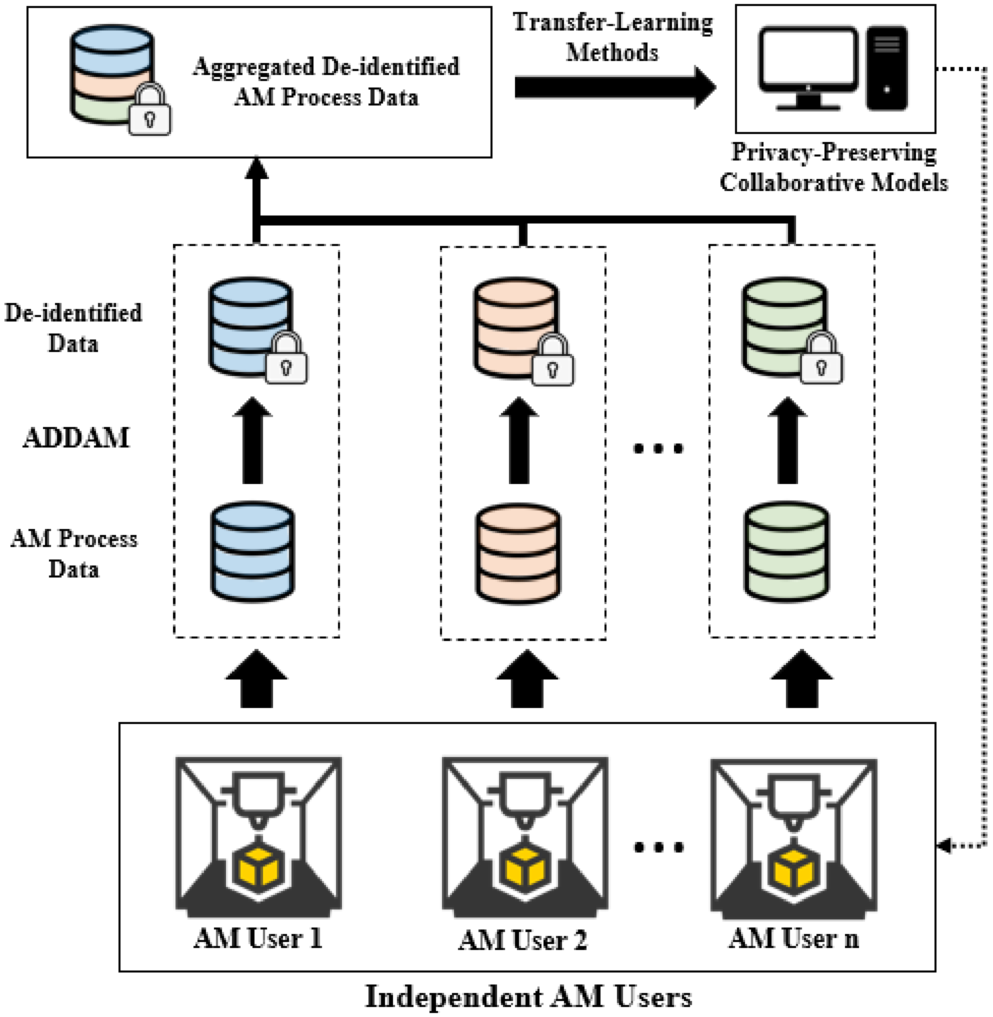

Aggregated De-identified AM Process Data
Transfer-Learning Methods
Privacy-Preserving Collaborative Models
De-identified Data
ADDAM
AM Process Data
AM User 1
AM User 2
AM User n
Independent AM Users

Figure 1: Overview of privacy-preserving collaborative process-defect modeling utilizing ADDAM. This process can be independently applied for multiple AM users, where the deidentified data can be aggregated and used to develop more robust privacy preserving collaborative models.

This paper proposes an Adaptive Design De-identification for Additive Manufacturing (ADDAM) methodology for masking the design information contained in AM thermal process data, while simultaneously retaining the quality related information for anomaly detection. This methodology will allow for the secure sharing of AM process data among multiple users, which establishes the foundation for data aggregation and transfer learning modeling, leading to the development of collaborative privacy-preserving anomaly detection models with improved IP security and model robustness (Figure 1). The technical contributions of this paper include: 1) the development of process data privacy and design de-identification framework for AM applications; and 2) the development of the new ADDAM algorithm with measurable privacy and utility for AM process data.

The reminder of the paper is organized as follows. The state-of-the-art studies are summarized in section 2. Section 3 discusses the data privacy problem and de-identification methods for AM applications. In section 4, the proposed ADDAM methodology is introduced, and section 5 introduces the case study to evaluate the effectiveness of the proposed method. Finally, the conclusion and future work are summarized in section 6.

2. Related Research

This section provides a survey of research related to the proposed method, which includes 1) Collaborative defect detection for metal-based AM; 2) AM process security and privacy concerns; and 3) a brief survey of the currently used anonymization techniques and their corresponding limitations.

2.1 Collaborative Defect Detection for Metal-based AM

This section focuses on the relevance of collaborative smart manufacturing in metal-based AM processes. Various *in-situ* process monitoring and defect detection methods have been proposed for identifying anomalies [12]. Among those methods, thermal imaging has been adopted to capture the AM thermal history under the premises that a stable thermal history will result in homogenous and thus defect-free structures. The high dimensional thermal history data are reduced to extract key process features that are then leveraged for anomaly detection [1]–[3], [13]–[16]. Moreover, layer-wise anomaly detection methods using thermal process data have been proposed for DED processes [3], [16]–[18], which provide an additional advantage compared to the defect detection models that only use local thermal features. However, the key limitation of this previous work is that these models were only evaluated using one set of design and printing parameters at a time. Changes in the process parameters can lead to deteriorated model accuracy, and the models would need to be re-trained and re-validated by newly collected data. This makes it potentially infeasible to develop accurate anomaly detection models for SMMs, who may print small batches of highly diverse parts [8], [10]. Transfer learning techniques can be leveraged to address the modeling limitations related to limited data availability. Transfer learning provides the user with the ability to apply learned knowledge or data from one domain to another related domain [19] This would allow the knowledge contained within multiple datasets to be leveraged in machine learning models, instead of completely discarding and re-collecting data to accommodate the change of AM process parameters. This can further the development of a collaborative data sharing framework. Currently, transfer learning has been proposed for transferring knowledge between different machines [10] and materials [20] for anomaly detection and distortion quantification [9], [21]. However, there are significant data privacy risks that may arise from sharing AM process data among different AM users. The AM process data contains confidential product information (e.g., design specifications and mechanical properties) that may jeopardize the product intellectual property (IP). By sharing AM process data outside their organization, AM users compromise their data privacy and are exposed to the risk for IP theft [22]–[24]. This is especially detrimental when using AM in the early phase of product prototyping and development. Lack of IP protection may lead to tremendous loss for the enterprise [25], [26]. Therefore, there is an urgent need in establishing a privacy-preserving data sharing framework to facilitate data sharing among multiple AM users for collaborative process-defect modeling, while not disclosing confidential product design information.

2.2 Privacy and Security Concerns in AM Systems

In the new era of industry 4.0, manufacturing systems are becoming more interconnected [27]. As AM systems have become increasingly prominent within industrial manufacturing applications, privacy and security have become significant issues that can affect a variety of different aspects of the AM process [24]. Traditionally, there are three fundamental concepts related to data security: *confidentiality*, *integrity,* and *availability*. This triad of security concerns encompasses vulnerabilities in manufacturing, including the overall data confidentiality, data reliability and consistency, and availability of equipment for service [26]. Most current data security and privacy concerns focus on preventing cyber-physical attacks that target on data *integrity* and *availability*, which can diminish the availability of the equipment or integrity of the printed parts and collected data [22], [24], [26], [28]–[30]. However, this leaves a significant gap for preventing cyber-domain attacks, which target on the product IP of the users [24], [25].

The main threats for AM IP protection are the attacks on data *confidentiality*. This type of attacks is commonly conducted by gaining malicious access to process data or related datasets and extracting key details to identify some confidential information [26]. This attack can be directly leveraged with AM process data to retrieve the product printing path information, and then reverse engineer the printed part design specifications [22], [24] (Figure 2). These attacks can be costly and detrimental to the AM users, as they directly attack the user's IP [26]. There are four specific tactics leveraged to preserve data privacy and prevent confidentiality attacks, including *anonymization*, *access control*, *encryption*, and *querying systems* [22]. From these different techniques, the most viable options for enhancing data security and facilitating transfer learning include anonymization and data encryption.

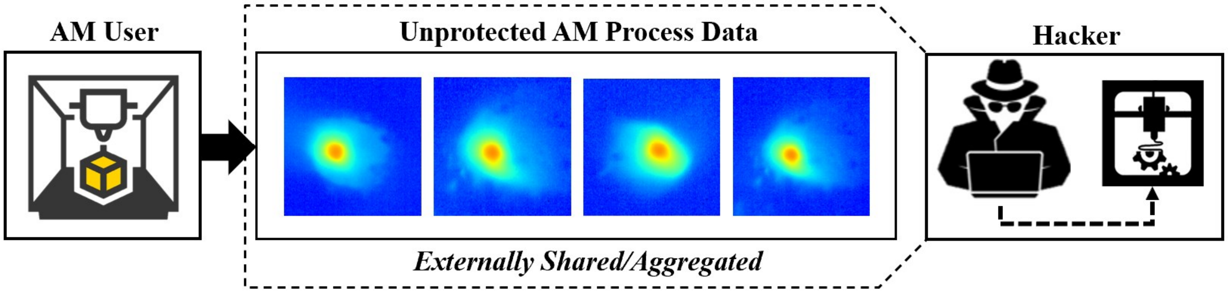


Figure 2: Current risk of externally shared or aggregated AM process data

The objective of anonymization is to remove or obscure the confidential information contained within the dataset, reducing the availability of specific, identifying characteristics available within the dataset [22], [31]. The privacy is enhanced by either suppressing or generalizing identifying features that can be used to collect sensitive information contained within the data. However, the biggest limitations facing anonymization revolve around ensuring that the data protection is strong enough to withstand re-identification attacks [22]. On the other hand, encryption is also a strong data security technique, which encodes the data so that it appears to be random, irrelevant data that is hard to understand without the proper encryption keys [32]. Despite the proven data protections, there are still reservations surrounding the overall usability of the post-encryption [22]. Specific forms of encryption, such as *homogeneous encryption,* are designed to allow computations to be performed once the data are encrypted [32], but the computational complexity is limited to only simple models [22], [33]. In addition, for both privacy measures, as the extent of the data protection increases, the overall usability of the protected data decreases [22], [31]. This means that achieving higher levels of data privacy traditionally leads to greater losses in data usability. Anonymization and encryption provide specific advantages to data privacy protection, but still face major challenges when balancing data privacy with data usability. Due to the additional computational restrictions associated with encryption, anonymization provides a potentially more effective framework for incorporating data privacy measures into collaborative, data-sharing AM applications.

2.3 The k-Anonymization Method and its Applications

This section details various anonymization methods, including the k-anonymization, k-same family of methodologies, and other de-identification models, which form the foundation for the proposed ADDAM methodology. Moreover, the major limitations of these methods when applying to AM design de-identification are summarized.

2.3.1 Traditional k-anonymization and Adaptations

k-anonymization is a specific form of de-identification for data privacy proposed in [34], and is an effective solution to guaranteeing data privacy, while still preserving some data usability. This method was originally designed for protecting individual sample identities and was primarily implemented for the tabular dataset applications. This includes data privacy protection for customer data [35], healthcare data [34], and public transportation data [36], as well as various other applications where the sample identity privacy is required. Tabular-structured datasets are defined as datasets that are minimally complex and contain independent (or weakly correlated) features, such as a person’s name, zip-code, social security number, health condition and others. These types of datasets provide an ideal application of k-anonymization, where the identity-compromising attributes are either generalized or suppressed to the point where there are at least $k-1$ identical samples for each sample in the dataset [34], [37]. However, for more complex applications, additional modifications are needed to improve the applicability of k-anonymization.

For example, the Mondrian multi-dimensional k-anonymization algorithm was formulated as an improved privacy-enhancing method to the traditional methodology [38]. The Mondrian method goes one step further to incorporate multidimensional partitioning to the anonymization procedure. This partitioning is used to achieve a more robust anonymization, as it factors in the relationship between different features during the generalization process [38]. Furthermore, clustering [39], [40] and p-sensitive anonymization algorithms [37] have also been proposed as other improvements to the traditional k-anonymization method. These updated methodologies still leverage the key generalization and suppression techniques used to ensure data privacy, but provide additional approaches to enhancing the process [23]. For all cases of k-anonymization, data protection techniques are applied to the identifying features, instead of applying anonymization to all features in the dataset. This helps to ensure the user-defined level of data privacy, while maintaining the usability of the non-identifying attributes.

However, k-anonymization methods face a few critical limitations. First, the de-identification approach is primarily applicable to tabular-structured datasets. Traditional applications of traditional k-anonymization and its variants (Mondrian [38], clustering [39], [40], p-sensitive [35], [37]) do not translate well to more complex data, such as image data or other multi-dimensional datasets. These datasets contain features that are highly correlated and highly nonlinear, which provides a new challenge for k-anonymization. Secondly, k-anonymization and most of its variants and enhancements cannot guarantee that there will be no data leakage [37]. These methods can provide enhanced data privacy, but do not provide complete protection, unless the dataset usability is extremely compromised. Finally, the proposed anonymization tactics of generalization and suppression are specific to the dataset application and can severely impact the interpretation of numerical attributes [34], [37]. This is primarily attributed to the generalization tactic, which in many cases converts the numerical attribute into a categorical variable (i.e., a person's numeric age into a categorical age range). This impacts the overall usability of the dataset and may potentially affect the applications. Because of the abovementioned limitations, several novel approaches to extending k-anonymization to the privacy preservation of more complex data structure have been proposed, as discussed in the next section.

2.3.2 k-Anonymization for Image Data

More recently, image data have become increasingly available, especially through the widespread implementation of security and surveillance monitoring systems. This has caused a drastic increase in the need for protecting individuals' privacy and identity [41]. Traditional naïve methods, such as blurring and pixilation, can mask the key identity information from images. However, they only serve the purpose of eliminating the identity of individuals within the images, and thus retain very little to no data utility [42]. Despite the alterations to the images, some of these methods only deter human recognition, as computer algorithms can be leveraged to reverse the distortions and re-identify those individuals [43]. To improve data privacy, several different techniques for facial de-identification algorithms have been developed [41]–[50]. These different approaches provide stronger protection guarantees and better overall data usability in de-identified images, pulling inspiration from the previous work of k-anonymization [34].

From the different approaches to facial de-identification, there are a few methods that provide robust de-identification capacities, which show potential for applications extending beyond facial image data. Firstly, the k-same approach takes the average of k similar images within a subset of facial images, and replaces the subset with an averaged, surrogate image [41]. This method is the most naïve scheme and extends the k-anonymization technique to complex image data, where these datasets can reach the same level of privacy as the k-anonymization algorithm (see [26] for proof). However, there are two main limitations of this methodology. The first is that the k-same method does not provide a satisfactory level of data utility [42]. This is because the image-space is highly non-linear and there is a steep utility loss when replacing the entire group of images with one single surrogate image. In addition, there is the threat of re-identification, since all the anonymization is performed using the original image dataset, meaning that some original information is contained within the published data [44]. From the k-same methodology, the k-same-select model was derived to improve the utility performance by providing prior knowledge about the dataset into the de-identification process, which further enhances the utility preservation [42]. Furthermore, the k-same-Model (k-same-M) approach also extends the k-same method to implement de-identification within the Active Appearance Models (AAM) [46], which are widely used in modeling and tracking facial image data. This produces a higher quality image, but there are still challenges in capturing key utility features, such as facial expressions during the anonymization process [49]. In summary, despite these enhancements, there were still significant gaps in applying data privacy to facial images to achieve a trade-off between privacy and utility.

To address the limitations of the k-same methods, the GARP-Face and APFD anonymization algorithms were developed for de-identifying facial images to achieve better balance between privacy and utility. Instead of replacing image groups with a surrogate image, both methods define the facial features, construct nearest neighborhoods, and use a separate utility specific subset of images to perform the anonymization. The GARP-Face (Gender-Age-Race) model [44] identifies useful features to preserve information (e.g., gender, age, and race) and develops classifiers to identify these features from the sample images. These features are then leveraged to identify k-similar images, which are then combined in the de-identification process to produce a surrogate image. The Attribute Preserved Face De-identification (APFD) method [45] follows a similar approach but leverages an additional optimization function that determines the optimal weights to be applied when averaging images. This weighted objective function is directly applied to the shape and appearance parameters, maximizing the number of common attributes the original and de-identified image share. Furthermore, both techniques also implement AAM to identify and characterize the shape and appearance parameters of the face. Overall, the results from this improved feature-targeting and preservation process show improved privacy and data utility preservation.

It is worth noting that these different facial de-identification methods apply the same level of data privacy to each image in the de-identified dataset, making them ***global de-identification*** approaches. However, the global de-identification approach is difficult to be directly applied to AM thermal images for the following reasons. Firstly, unlike the facial de-identification datasets, the AM thermal images suffer from limited data availability and a tendency to have repeating identities within the dataset. This can lead to compromised performance when directly applying a global de-identification model, as many of the nearest neighbor images may share the same identities, and the limited number of samples can degrade the overall dataset diversity. Secondly, the AM process data anomalies demonstrate high variations in their distributions, meaning that they are distinctly different from both the healthy distribution and each other. However, the facial image data do not encounter this problem, as most human faces will share a similar distribution of features. This creates another roadblock to directly implementing global de-identification methods, since directly averaging k nearest neighbors will blur the difference between healthy and abnormal melt pool images, leading to dramatically degraded data usability (i.e., anomaly detection performance).

3. De-identification and Data Privacy for AM

This section will introduce the various types of AM data, as well as the confidentiality and the vulnerability in these data. In addition, the role of data privacy in AM and the importance of maintaining the balance between data utility and privacy is explained. The formal definitions related to data privacy for AM applications set the foundation for the proposed ADDAM algorithm.

3.1 AM Data Description

As described in Figure 3, various types of AM data are generated in the four major steps of AM, i.e., *design*, *slicing*, *manufacturing*, and *inspection*. Together, these steps construct the cyber-physical AM systems [5].

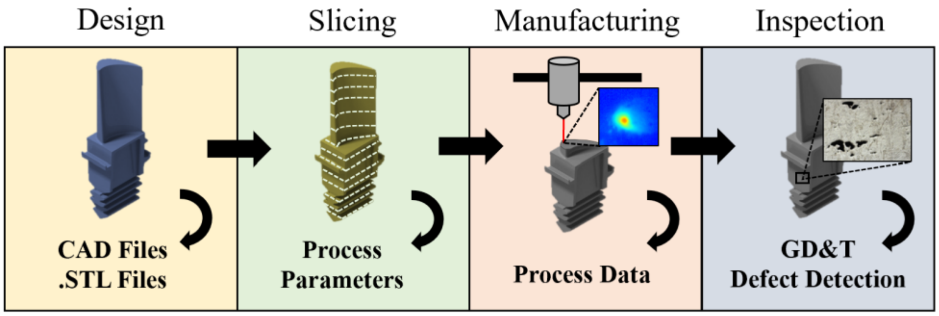


Figure 3: Key steps of AM and data generated

The *design* phase includes the generation of the CAD and STL files, which represent the detailed, three-dimensional part design. This information is highly confidential, especially for rapid prototyping applications. Because of this, the data generated during this phase (CAD and STL design files) should be maintained internally, and never shared for IP protection purposes.

The *slicing* phase takes the design file as the input and generates a g-code file, which contains several different process parameters, including the printing path, print speed, layer thickness, temperature settings, and many others. Like the design files, these process parameters also contain confidential design attributes, and should never be shared externally.

The *manufacturing* phase involves the physical printing process while generating a variety of process data, including thermal imaging data, acceleration, acoustics, and others. Recently, the process data play critical roles in *in-situ* process monitoring and anomaly detection. However, the process data contain confidential design information, particularly relating to the printing geometries and parameters. These embedded features can be extracted and linked back to the part design, compromising the product IP. Therefore, the implementation of data privacy measures is particularly important at this phase because the collected process data are expected to be externally shared and aggregated.

Finally, the *inspection* phase is where the final printed part is evaluated for quality assurance. This includes checking the Geometric, Dimension, and Tolerance features (GD&T) of the part, as well as detecting defects within the print part. Although this process also creates vulnerabilities for IP theft, most data collected during this phase will be stored internally and only accessed locally. The data from this phase that is shared externally for developing *in-situ* defect modeling (anomaly labels) usually does not contain confidential design information.

3.2 Key Definitions in AM Privacy

In this section, several important definitions in AM process data de-identification for process-defect modeling are introduced by integrating AM process knowledge into data privacy and anonymization related terminologies.

Definition I: AM data privacy is defined as the ability of the shared AM data that prevent a malicious third party from identifying critical product design specifications. For example, for AM thermal process data, specific privacy measures need to be applied directly to the melt pool images to properly de-identify/mask the printing trajectory information (Figure 4). This creates a safeguard for protecting against IP thefts through the AM process data.

When applying the de-identification framework, the AM data discussed in section 3.1 can be briefly categorized into three groups of attributes [23], [34], [37], [38], as summarized in Table 1.

1) **AM Sensitive attributes** are attributes that can directly identify the design information contained within the dataset. This includes design data (i.e., CAD files), attributes derived from the design data (e.g., g-codes and printing angular information), and the complete thermal history, all of which pose a significant IP privacy risk. Furthermore, *AM Design features* are embedded within the complete thermal history, which poses a significant risk of data privacy. These features can be directly extracted from the thermal process images themselves (as illustrated in Figure 4). This creates a major vulnerability for the product IP when sharing the data externally, where malicious third parties could gain access to the complete thermal image set and extract these critical design features. Thus, it is important that AM sensitive attributes are kept locally, or any relationship between the shared data and corresponding sensitive attributes needs to be de-identified.

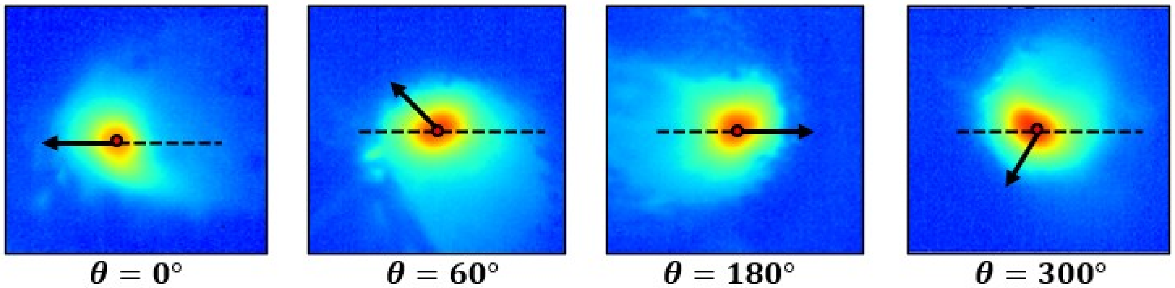


Figure 4: Melt pool angular orientation and printing path

2) **AM Quasi-identifiers** are attributes, that alone, do not directly give away the product design information. However, when used in conjunction with other AM quasi-identifiers, or sensitive attributes, they can be leveraged to further identify confidential design features. For example,

within the thermal process data, each melt pool image alone (or each pixel within the image) does not directly give away confidential design information. However, when a large enough set of thermal images are available, they can be directly used to re-identify the sensitive AM design features. Furthermore, features such as the layer-wise location of the melt pool, and the sequential image ID, can be used to enhance the identification of compromising trends and information within the process data. Ultimately, the AM quasi-identifier's relationship with the sensitive attributes should be removed or de-identified for secure data sharing.

3) **AM Insensitive attributes** are attributes that do not have any direct relationship with the design information. This includes the *AM Utility Features*, which represent the geometric and thermal features within the melt pools (e.g., melt pool area and eccentricity, and maximum temperature). Unlike the AM design features, these utility features are *insensitive* to design information, but informative for utility preservation (e.g., anomaly detection). Overall, they do not pose a security risk and are able to be leveraged for de-identification, or externally shared if desired.

Table 1: Breakdown for AM process data features and attributes

AM Steps	**AM Attributes Considered**	**Attribute Classification**	**Description**
Design	CAD/STL file	AM sensitive attribute	Detailed 3D geometric design and specifications file
Slicing	G-code file	AM sensitive attribute	Parameters related to the printing process and printing path information
Manufacturing	Thermal image	AM quasi-identifier	Independent Image matrix of thermal intensity readings
	Layer label	AM quasi-identifier	Layer-wise location of melt pool
	Image ID	AM quasi-identifier	Image ID, sequentially captured
	Angular label	AM sensitive attribute	Numeric label for angular information
Inspection	Anomaly label	AM insensitive attribute	Binary label for anomaly information

Definition II: AM data utility is defined as the overall usability of the dataset for specific modeling purpose (e.g., anomaly detection) after privacy-preserving measures [52]. For the AM process data de-identification, this means that sufficient information is retained in the de-identified data for the end-user to train defect detection models. This is measured by the ability of a machine learning model to accurately detect the presence of anomalies within the de-identified data.

4. The Proposed ADDAM Methodology

In this section, the ADDAM methodology is proposed for de-identifying design information from AM melt pool image data. This new methodology focuses on developing a secure aggregation mechanism for collaborative process-defect modeling by masking the design information in the thermal history while retaining the process quality information. This section starts with an overview of the proposed ADDAM methodology (Figure 5), followed by a subsequent breakdown for each of the main stages of the proposed method.

4.1 Proposed ADDAM Overview

The major advantage of the ADDAM algorithm is the introduction of the novel adaptive mechanism to determine the level of data privacy on a per-image basis. This deviates from the traditional forms of k-anonymization, which take a global approach to data privacy, de-identifying each image with the same, globally determined level of data privacy. The proposed adaptive approach is motivated by the following two reasons.

Firstly, the AM process data tends to be imbalanced and suffers from limited data availability, where there are vastly more cases of healthy melt pools as compared to abnormalities. This creates two major challenges. First, there is potentially a limited number of unique angular identities available to de-identify. This means that de-identifying a sample image with its k-closet images may not necessarily improve data privacy if its nearest neighbors contain the same angular identity. In addition, due to the rare and diverse nature of anomalies, the k-closest images of an abnormal image may include either healthy images or abnormal images with different abnormality categories, leading to reduced distinction between healthy and unhealthy melt pool images after de-identification. This will significantly jeopardize the data utility (i.e., anomaly detection). Secondly, during the printing process there is a noticeable thermal distribution change over time in the thermal history. As a result, the baseline for healthy melt pools observed at different layers would vary significantly, even though their process parameters are set the same. Implementing a global k

value completely neglects this drifting trend in the thermal distribution and will lead to de-identifying using images that are not actually neighbors in the printing process.

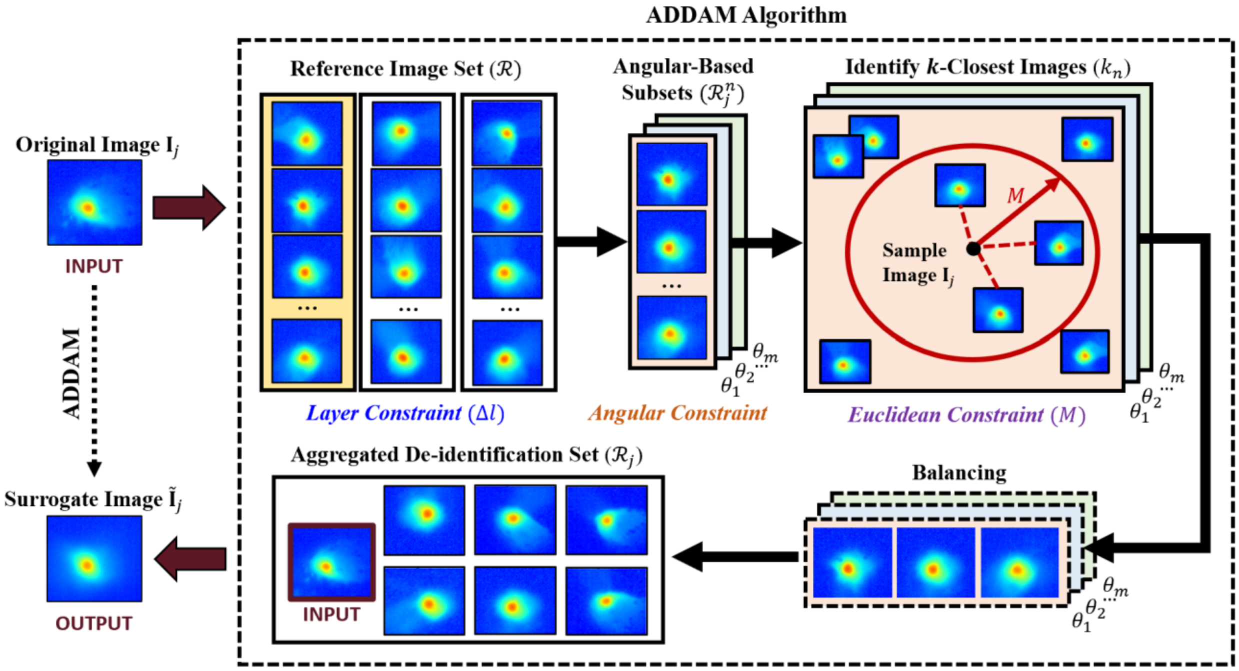

ADDAM Algorithm
Original Image $\mathbf{I}_j$
INPUT
ADDAM
Surrogate Image $\tilde{\mathbf{I}}_j$
OUTPUT
Reference Image Set ($\mathcal{R}$)
Angular-Based Subsets ($\mathcal{R}_j^n$)
Identify k-Closest Images (k_n)
M
Sample Image $\mathbf{I}_j$
θ_m
θ_2
θ_1
Layer Constraint (Δl)
Angular Constraint
Euclidean Constraint (M)
Aggregated De-identification Set ($\mathcal{R}_j$)
INPUT
Balancing

Figure 5: Overview of the ADDAM methodology

A reference or gallery set of s thermal images, with each image containing $r \times c$ pixel, can be denoted as $\mathcal{R} \coloneqq \{\mathbf{R}_i \in \mathbb{R}^{r \times c},\ i = 1, \dots, s\}$. The proposed ADDAM methodology defines a transformation function, $f : \mathbb{R}^{r \times c} \to \mathbb{R}^{r \times c}$, which generates a surrogate thermal image for each observed thermal image, $\mathbf{I}_j \in \mathbb{R}^{r \times c}$, as illustrated in Equation (1).

$$\tilde{\mathbf{I}}_j = f(\mathbf{I}_j) \tag{1}$$

where $\tilde{\mathbf{I}}_j \in \mathbb{R}^{r \times c}$ denotes the surrogate image for $\mathbf{I}_j$ with its angular identity $\varphi(\mathbf{I}_j)$ de-identified, where $\varphi(\cdot)$ denotes the instantaneous printing orientation of the thermal image. The transformation function f is implemented by pooling the observed thermal image $\mathbf{I}_j$ with a selective subset of $k_j - 1$ thermal images from the reference set $\mathcal{R}$, denoted as $\mathcal{R}_j \subseteq \mathcal{R}$ and $|\mathcal{R}_j| = k_j - 1$, where $|\cdot|$ denotes the cardinality of the image set.

The de-identification function, f, aims to improve data privacy by masking the design information (i.e., printing path information) from each image $\mathbf{I}_j$, while simultaneously retaining data utility for anomaly detection and part certification. The proposed ADDAM methodology can be divided into several stages, which are discussed in the following sub-sections.

4.2 Stage 1: Reference Set Selection

In real world applications, AM users have the ability to use their historical data, or data available from machine calibrations, to create a diverse and robust reference set $\mathcal{R}$ for de-identification. There are some key requirements to keep in mind when developing this independent reference set. *Firstly*, the reference set should have a high diversity of angular orientations. This is important as it will better facilitate proper de-identification, as more unique identities can lead to more variability in the de-identified images with respect to the angular identity. *Secondly*, the reference data needs to share a similar domain distribution the data to be de-identified. This is important for the similarity space construction and the preservation of the data utility, as the geometric and thermal features derived from each distribution are indicative to the overall characteristics of the distribution. If these features differ too much, it will drastically impact the adaptive procedure of the algorithm and lead to utility and/or privacy degradation. *Finally*, the reference set should not include any samples that are also within the set of images to be de-identified. This will lead to a degraded privacy gains, as these duplicate reference images would be guaranteed to be included in the adaptive-k samples used to de-identify the original image.

After selection of the reference images, the overall reference set quality can be evaluated in a couple of ways. The first is to evaluate the overall difference between the derived thermal and geometric features of the reference set and the de-identification set. These features play an important role in de-identification, and if their distribution in the reference set differs too much from the de-identification set, it will impact the overall algorithm performance. Secondly, the two domain distributions could be quantitatively evaluated using a distance metric, such as Maximum Mean Discrepancy (MMD) or Kullback-Leibler

Divergence. This allows a user to quantify the distance and difference between two distributions with metrics that are commonly used in transfer learning and domain adaption applications [53], [54].

4.3 Stage 2: Process Data Dimension Reduction

To reduce the dimensionality of the thermal images, the reference set, $\mathcal{R}$, is used to fit Vectorized Principal Component Analysis (vPCA) for low dimensional process feature extraction. The vPCA achieves dimensionality reduction by mapping the original melt pool images into a low-dimensional space, where each sample image, $\mathbf{I}_j$, is then transformed into this space, as illustrated in Equation (2).

$$\boldsymbol{v}_j = \mathrm{vec}(\mathbf{I}_j)\mathbf{W}_p \tag{2}$$

where $\mathbf{W}_p$ represents the projection matrix estimated from the reference image set, $\mathcal{R}$, and p denotes the percentage of the total variability explained by the extracted PCs, denoted as $\boldsymbol{v}_j$. In most cases, the value of p is set as 95% such that the major variability in the original melt pool image $\mathbf{I}_j$ can be retained in $\boldsymbol{v}_j$.

4.4 Stage 3: AM Utility Attribute Space Construction

The *Utility Attribute Space* (UAS) incorporates derived features to construct a vector space to evaluate the utility-aware similarity of sample images to images in the reference image set. The features used to construct this space include both the geometric features and the other insensitive, utility related features. These derived features can be directly indicative of the overall health status of the melt pool and play an important role in preserving the dataset utility and achieving adaptive de-identification. However, it is important to note that these features underperform compared to the features extracted using PCA for anomaly detection. For this reason, these features are not leveraged during classification. The UAS is leveraged to identify the abnormal and healthy melt pool images, based on how similar they are to their neighbors. This improves data privacy as it ensures that healthy melt pools, which tend to have a high number of neighbors, achieve a higher level of data privacy. Since healthy melt pools tend to make up the majority of data samples, this ensures better data set privacy. In addition, the UAS allows for abnormal melt pools to maintain a minimum level of de-identification, which in turn maintains dataset usability. This is due to the characteristic fact that the abnormal melt pools are dissimilar from healthy melt pools and each other, allowing these samples to maintain their distinct characteristics by using a lower adaptive k value. It is important to note that this will not compromise the overall dataset privacy, as with AM data not every image has to be de-identified to ensure data privacy. The main risks are exposed when a large set of images are available and can be used together, and these anomalies only make up a small subset of the data.

Multiple AM utility attributes are proposed to form the UAS. The first attribute is the L_2 norm of the reconstruction error denoted as g_j^1, which can be calculated in Eq. (3) for each $\mathbf{I}_j$,

$$g_j^1 = \left\| \mathbf{I}_j - \hat{\mathbf{I}}_j \right\|_2 \tag{3}$$

where $\hat{\mathbf{I}}_j \in \mathbb{R}^{r \times c}$ denotes the image reconstructed from $\boldsymbol{v}_j$. This feature is important as the vPCA algorithm is fit using healthy reference images, which provide a larger L_2 reconstruction error for melt pools that contain anomalies. Moreover, a few additional utility features can be extracted from each original melt pool image $\mathbf{I}_j$, including peak temperature and its row and column location in the field of view, as well as the area and eccentricity of the melt pool, which is segmented using the melting point of the feedstock material. These abovementioned features of $\mathbf{I}_j$ are denoted as g_j^w $(w = 1,2, \dots ,6)$. The 6-dimensional feature vector is denoted as $\boldsymbol{g}_j = \left(g_j^1, g_j^2, \dots , g_j^6\right)$, which forms the UAS to determine the similarity of each melt pool image $\mathbf{I}_j$ against the reference images.

A distance function is defined in the UAS, denoted as $d_g(\mathbf{X}, \mathbf{Y})$, which represents the Euclidean distance between two thermal images, i.e., **X** and **Y**, in the UAS. This distance function is used to identify the subset of images in $\mathcal{R}$ to be used to de-identify the observed image $\mathbf{I}_j$, and thus acts as one of the controlling mechanisms used to tune the sensitivity of the ADDAM algorithm when determining the adaptive k value.

4.5 Stage 4: Determination of the Adaptive k_j Value

This stage determines the adaptive k_j value for $\mathbf{I}_j$. The proposed method significantly departs from the traditional k-same, GARP, and APFD algorithms, which utilize a global k value to achieve image anonymization [44], [45]. There are two distinct and important operations within the ADDAM algorithm.

Firstly, the ADDAM algorithm implements a series of constraints when determining the k-closest reference images of $\mathbf{I}_j$. These constraints leverage characteristics of each melt pool, including the layer location and angular identity, and define the neighborhood size within the UAS. This plays a crucial role in the ADDAM algorithm, as it allows the user to adjust and control the sensitivity and tune the de-identification algorithm. *Secondly*, the adaptive algorithm employs an additional balancing mechanism, which ensures that the reference set, combined with the sample image $\mathbf{I}_j$, is equally diverse across all possible angular identities in the dataset. Both aspects are critical components that de-identify the angular identities while retaining the utility related information in the de-identified image

For each angular identity in $\mathcal{R}$, denoted as θ_n $(n = 1,2,\dots,m)$, the corresponding angular-reference set, used to de-identify $\mathbf{I}_j$, can be defined in equation (4),

$$\mathcal{R}_j^n = \left\{ \mathbf{R}_i \middle| \begin{array}{c} l(\mathbf{R}_i) \in \left[l(\mathbf{I}_j) - \Delta l, l(\mathbf{I}_j) + \Delta l\right] \\ \varphi(\mathbf{R}_i) = \theta_n \\ d_g(\mathbf{I}_j, \mathbf{R}_i) \leq M \end{array} \right\} \tag{4}$$

where the first constraint enforces the identified neighbors to be in proximity of $\mathbf{I}_j$ in terms of the build layers, where $l(\cdot)$ denotes the layer index where the thermal image is collected from, and Δl represents the pre-defined maximum allowable layer difference between the identified neighboring images and $\mathbf{I}_j$; the second constraint requires the elements in $\mathcal{R}_j^n$ to be of the angular identity θ_n; the last constraint forces that the Euclidean distance (denoted as d_g) between the identified neighboring images and $\mathbf{I}_j$ are no larger than a pre-defined threshold value M in the UAS defined in Stage 3. After applying these constraints, the number of closest reference images in $\mathcal{R}_j^n$ can be calculated as below.

$$k_j^n = \left|\mathcal{R}_j^n\right| \tag{5}$$

where $k_j^n \geq 0$, and k_j^n varies according to the similarity of $\mathbf{I}_j$ to the reference thermal images in $\mathcal{R}$ as well as the corresponding angular identity θ_n. For example, if $\mathbf{I}_j$ is a healthy thermal image, there will be many $\mathbf{R}_i$'s in proximity of $\mathbf{I}_j$ in terms of both build layers and within the UAS, and thus the value of k_j^n will be larger. However, if $\mathbf{I}_j$ is an unhealthy thermal image, there will be very few (or even none) neighboring thermal images in $\mathcal{R}$, and thus the k_j^n value will be very small (or even zero). In the case where one or more of the $k_j^n = 0$, $\mathbf{I}_j$ is probably extremely abnormal, and therefore will receive no de-identification to keep its significant deviation from the healthy group. This scenario is extremely rare within $\mathcal{J}$, and will not create any major privacy concerns as abnormal melt pools make up the minority. In addition, it is worth noting that the sample image $\mathbf{I}_j$ is the nearest neighbor to itself within the subset where $\varphi(\mathbf{I}_j) = \theta_n$. The sample image will be incorporated into the corresponding $\mathcal{R}_j^n$ of the same angular identity θ_n. This ensures that sample image angular identify will be accounted for when the algorithm undergoes a balancing procedure.

Subsequently, the adaptive algorithm involves a crucial balancing function that ensures that there is an equal representation of images within each reference subset $\mathcal{R}_j^n$. This prevents an overpopulation of one angular identity during the de-identification process, which can impact the amount of data privacy achieved. This step results from the major difference present between the ADDAM algorithm and traditional k-anonymization algorithms. Traditionally, when applying global anonymization techniques, each image within the dataset contained a unique identity, such as a human face. If this image is anonymized with any other identity in the dataset, there will be a resulting gain of privacy for that individual. However, with AM thermal process data, there are repeating identities within the dataset. Therefore, the de-identification with the same identity will not yield any privacy gains. Balancing the distribution of these angular identities within $\mathcal{R}_j$ guarantees that not one unique identity will be more prominent than the others during de-identification. This is accomplished by first ensuring that each angular-based subset previously determined is re-indexed into a monotonically increasing order, such that $d_g(\mathbf{I}_j, \mathbf{R}_{(1)}) \leq d_g(\mathbf{I}_j, \mathbf{R}_{(2)}) \leq \cdots \leq d_g\left(\mathbf{I}_j, \mathbf{R}_{\left(k_j^n\right)}\right) \leq \cdots \leq d_g(\mathbf{I}_j, \mathbf{R}_{(s)})$. Re-indexing ensures that the images with the shortest Euclidean distance to the sample image will be first in the order of the subsets.

From here, a fourth filter is applied, which limits the size of each subgroup to be equal to the smallest subgroup. This is the novel balancing procedure which ensures that each angular identity is equally represented within the closest k_n images to the sample image,

$$k_j^* = \min\left(k_j^n\right) \tag{6}$$

and the balanced identify subgroup $\mathcal{R}_j^{*n} = \left\{\mathbf{R}_i \middle| \mathbf{R}_i \leq d_g\left(\mathbf{I}_j, \mathbf{R}_{\left(k_j^*\right)}\right), \mathbf{R}_i \in \mathcal{R}_j^n\right\}$. Next, the aggregated de-identification set, $\mathcal{R}_j$, can be formed by directly merging $\mathcal{R}_j^{*n}$'s to form the larger and equally diverse de-identification dataset. This aggregated set $\mathcal{R}_j$ is directly used to de-identify sample image $\mathbf{I}_j$,

$$k_j = m \times k_j^* \tag{7}$$

$$\mathcal{R}_j = \bigcup_{n=1}^{m} \mathcal{R}_j^{*n} \tag{8}$$

where k_j is the number of aggregated, closest images used to de-identify $\mathbf{I}_j$. The aggregated de-identification set, $\mathcal{R}_j$, is a direct combination of all the balanced reference subgroups, $\mathcal{R}_j^{*n}$. This is the set of images (sample image and closest reference images) that will be directly used to de-identify $\mathbf{I}_j$.

4.6 Stage 5: Melt Pool Image De-identification

The final stage of the proposed methodology is AM process image de-identification, given the k_j neighboring images identified in stage 4. For each sample image $\mathbf{I}_j$, all the images in $\mathcal{R}_j$ are combined to form the anonymized image, $\tilde{\mathbf{I}}_j$, by directly averaging the dimensionally reduced images in $\mathcal{R}_j$ as below.

$$\widetilde{\boldsymbol{v}}_j = \frac{\sum_{\boldsymbol{v}_j \in \mathcal{R}_j} \boldsymbol{v}_j}{k_j} \tag{9}$$

where each image within the aggregated de-identification set is directly averaged to create a de-identified PC vector ($\widetilde{\boldsymbol{v}}_j$), which can then be reversely transformed into the original image space to obtain the surrogate image $\tilde{\mathbf{I}}_j$ to be published and aggregated with data from other AM users.

4.7 Evaluation of Design De-identification Performance

To evaluate the design de-identification performance for secured collaborative AM process-defect modeling, two novel anonymization performance metrics are introduced to meet the needs of AM applications. These metrics will allow for the measurable gain in privacy and loss in utility of the dataset shared, and then can be further evaluated using a Pareto Front [55], [56] to quantify the tradeoff between two conflicting objectives: 1) minimizing *Utility Loss* and 2) maximizing *Privacy Gain* [44]. These two metrics are derived from the traditional classification metrics, which have been previously leveraged to evaluate the performance of de-identification and *k*-anonymization algorithms [41], [43], [44], [46]. Traditionally, the data privacy performance can be gauged as the number of correct predictions before and after de-identification. This allows for a natural and easily implementable method for evaluating model performance using ML models by simply calculating the performance metrics before and after.

Definition III: Utility Loss, UL, is defined as the decrease in the anomaly detection performance (in percentage) due to de-identification.

$$\mathrm{UL} = -(X_{\mathrm{Base}} - X_{\mathrm{Anon}}) \tag{10}$$

where X_{Base} and X_{Anon} denote the anomaly detection performance metrics achieved by the original dataset and the de-identified dataset, respectively. It is worth noting that based on the definition, UL is usually a negative value. **Therefore, it is desirable to either *minimize* $|\mathrm{UL}|$ or *maximize*** UL. In addition, the UL metric is written in a general form of anomaly detection performance metrics above, while it relies on leveraging classification metrics, such as $\mathrm{F1}\ (13)$ or overall $\mathrm{Accuracy}\ (14)$. In general, the $\mathrm{F1}$ can be leveraged when evaluating UL, as AM process data is traditionally unbalanced with respect to the anomaly labels.

***Definition IV: Privacy Gain**,* PG, is evaluating the classifier model's ability to predict the printing path orientation between the baseline and de-identified datasets, ultimately evaluating the privacy gains from implementing de-identification algorithms.

$$\mathrm{PG} = Z_{\mathrm{Base}} - Z_{\mathrm{Anon}} \tag{11}$$

where Z_{Base} and Z_{Anon} denote the printing orientation classification performance metrics achieved by the original dataset and the de-identified dataset, respectively, and they are also written in general form and relies on the specific classification metric used $(12) - (13)$, which is determined heavily on the balanced or unbalanced characterization of the dataset. In general, the Accuracy can be leveraged when evaluating PG, if the datasets are balanced with respect to the print orientation labels. Had the angular class labeling been unbalanced, the F1 should be used.

Both PG and UL are plotted in a two-dimensional plot to find the pareto front of optimal solutions, determining the overall performance of ADDAM. The following equations describe the different classification metrics used to build the UL performance metric.

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FP}}; \text{Precision} = \frac{\text{TP}}{\text{TP+FN}} \tag{12}$$

$$\text{F1} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{13}$$

In these equations, TP represents the correct prediction that there is a defect present and the melt pool is abnormal, and TN represents the correct prediction that there are no defects present, and the melt pool is healthy. In addition, FP represents the incorrect prediction that there is a defect present, but the melt pool is healthy, and FN represents the incorrect prediction that there are no defects present, but the melt pool is abnormal. The metrics used depend on how balanced the data is with respect to class labels. For example, when the dataset is unbalanced, F1 should be used for X_{Base} and X_{Anon}. Otherwise, Accuracy would be a good choice [57]. Furthermore, accuracy is leveraged as the underlying metric behind the PG.

$$\text{Accuracy} = \frac{\text{Correctly Predicted Angular Identities}}{\text{Total Predictions}} \tag{14}$$

In summary, a pseudocode of the proposed ADDAM algorithm is detailed in Figure 6.

```
ADDAM De-identification Algorithm
Input: Sample Image Set $\mathcal{I}$
       Reference Subsets $\mathcal{R}$
       Derived feature set $g_j^w$ $(w = 1,2, \dots ,6)$ for both $\mathcal{I}$ and $\mathcal{R}$ image sets
Output: Anonymized Image Set $\tilde{\mathcal{I}}$

Dimensionality Reduction and UAS Construction:
    Fit vPCA with reference image set $\mathcal{R}$
    Transform each sample image $\mathbf{I}_j \in \mathcal{I}$ into $\mathcal{R}$ vPCA Space: $\boldsymbol{v}_j = \text{vec}(\mathbf{I}_j)\mathbf{W}_p$
    Calculate $g_j^1$ for each image and append to $\boldsymbol{g}_j$

ADDAM Algorithm:
    Generate the empty image set $\tilde{\mathcal{I}}$
    For $\boldsymbol{v}_j \in \mathcal{I}$:
        Generate the empty image set $\mathcal{R}_{reduced}$

    Applying the Layer Constraint
        For $\boldsymbol{v}_i \in \mathcal{R}$:
            Identify $l(\mathbf{R}_i)$
            if $l(\mathbf{R}_i) \in [l(\mathbf{I}_j) - \Delta l, l(\mathbf{I}_j) + \Delta l]$:
                Append $\boldsymbol{v}_i$ to $\mathcal{R}_{reduced}$
Applying the Angular Constraint
    Generate empty image set $\mathcal{R}_j^n$ for $\theta_n$ angular identities
    Append $\boldsymbol{v}_j$to corresponding $\mathcal{R}_j^n$ with $\varphi(\mathbf{I}_j) = \theta_n$
    For $\boldsymbol{v}_i \in \mathcal{R}_{reduced}$:
        if $\theta_n(\boldsymbol{v}_i) = \theta_n$
            Append $\boldsymbol{v}_i$ to $\mathcal{R}_j^n$
Applying the Euclidean Distance Constraint
    For $\boldsymbol{v}_i \in \mathcal{R}_j^n$:
        Calculate Euclidean distance $d_g(\mathbf{I}_j, \mathbf{R}_i)$ using the UAS features
        if $d_g(\mathbf{I}_j, \mathbf{R}_i) > M$,
            Remove $\boldsymbol{v}_i$
    $k_j^* = \min(\mathrm{len}(\mathcal{R}_j^n))$
Balancing the Angular Subsets
    For each $\mathcal{R}_j^n$:
        Re-index $\mathcal{R}_j^n$ as $\boldsymbol{v}_{(1)}, \boldsymbol{v}_{(2)}, \dots, \boldsymbol{v}_{(k_j^n)}$ in the increasing order based on $d_g$
        if $\mathrm{len}(\mathcal{R}_j^n) > k_j^*$
            Remove $\boldsymbol{v}_{(i^-)}$ if $i^- > k_j^*$
Forming the Aggregated De-Identification Set
    $\mathcal{R}_j = \mathrm{concatenate}(\mathcal{R}_j^n)$
    $k_j = m \times k_j^*$
De-Identification
    Average the remaining $k$-closest images $\tilde{\boldsymbol{v}}_j = \frac{\sum_{\boldsymbol{v}_i \in \mathcal{R}_j} \boldsymbol{v}_i}{k_j}$
    Append de-identified image $\tilde{\boldsymbol{v}}_j$ to $\tilde{\mathcal{I}}$ (PC matrix of Image)
Inverse Transform vPCA $\tilde{\mathcal{I}}$ to original image dimensions
Return Anonymized Image Set $\tilde{\mathcal{I}}$
```

Figure 6: The Pseudocode of the ADDAM Algorithm

5. Case Study

This section will discuss the case study used to validate the proposed ADDAM methodology with respect to both data privacy gain and data utility preservation.

5.1 Experimental Setup and Data Description

The experimental setup is visualized in Figure 7, which consists of an OPTOMEC LENS 750 Directed Energy Deposition (DED) machine equipped with a co-axial pyrometer camera (Stratonics Inc.) to capture the thermal images during the fabrication [2], [3], [17], [18]. The LENS DED machine leverages a 1.0 kW Nd:YAG laser, and the pyrometer is mounted above the DED machine, outside of the inert chamber, where it is aligned with a series of mirrors to obtain a co-axial view. The specifications of the pyrometer are as follows:

- Exposure time: 2.0274ms
- Image Size 752 × 480 and pixel pitch 6.45μm
- Captured temperature range: 1000–2500 °C
- Pixel clock: 5 MHz
- Image collection rate: 6.4 Hz

Two cylindrical specimens with different printing parameters and infill patterns were fabricated for data collection. The key printing parameters are summarized in Table 2.

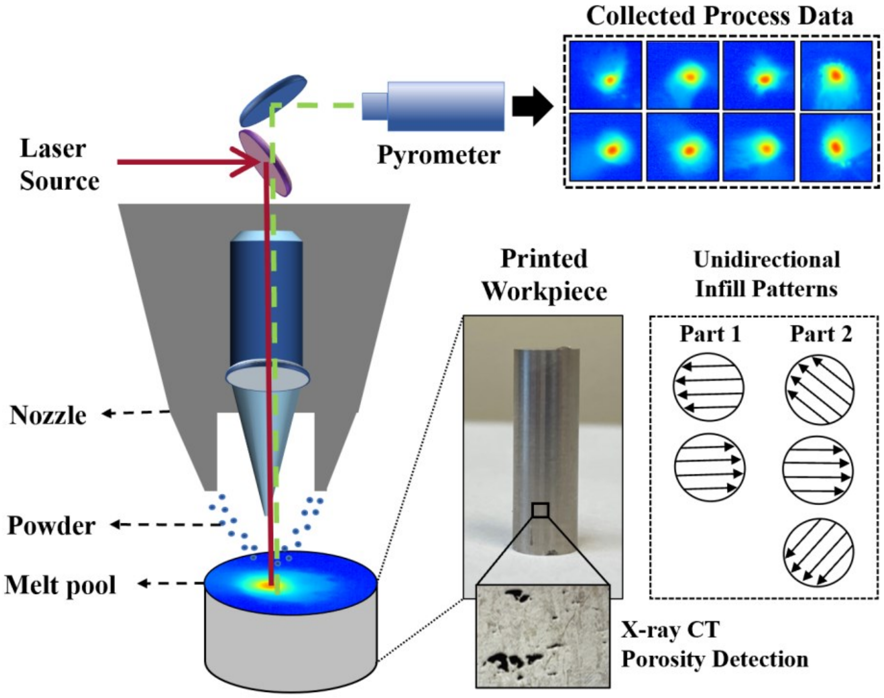


Figure 7: Experimental set-up and data collection methods

Table 2: Printing parameters for Part 1 and Part 2

Parameters	Part 1	Part 2
Scan speed	40 inch/min	50 inch/min
Powder feed rate	3 rpm	2.5 rpm
Hatch spacing	0.02 inch	0.025 inch
Power	300 W	350 W
Layer thickness	0.015 inch	0.015 inch
Number of thermal images utilized	1,616	842
Number of layers in the build	69	55
Number of anomalies	138 (6%)	N/A
Infill pattern	Unidirectional (0°/180°)	Unidirectional (60°/180/300°)

The specimen fabrication resulted in raw thermal images with 480 rows and 752 columns, in which each pixel represents a temperature reading at the corresponding location. First, these images are cropped into 201×201 to reduce the image dimensions and remove irrelevant regions that do not contain the melt pools. It is important to note that the initial cropping parameters were consistent across all the images. In addition, the instantaneous printing orientations of both datasets were determined by leveraging the g-codes of the two specimens post-processing. There are two unique angular identities in Part 1 (0°/180°), and three in Part 2 (60°/180°/300°). Furthermore, due to the existing trends in the AM thermal process data, only the data after layer 20 were leveraged for tuning and evaluating the performance of the different algorithms. This provides a better, more consistent evaluation of ADDAM performance. Overall, these two datasets will provide four unique angular identities and 2,458 thermal melt pool images for experimentation. This is a limited dataset that will allow more controlled experimentation and simulate the limited data availability faced by SMMs. The results are reflective and comparable to the application of ADDAM in a practical setting.

After part fabrication, the porosities were detected utilizing the XCT inspection and subsequently matched with the thermal images based on the porosity location and the g-code for Part 1 only. As a result, the thermal images were labeled as defect present (1) or defect absent (0). For Part 2, there is no post-process inspection data available for anomaly detection modeling.

5.2 Evaluation Procedure

5.2.1 Benchmark Method Selection

For benchmark comparison, a global k-anonymization approach was applied. This involves anonymizing each sample image with a constant number of k-closest neighbors, instead of allowing an adaptive k value to be applied to each image. This is indicative of the traditional global k-anonymization methods that have been used in the past, primarily in the k-same methods. The performance comparison will demonstrate the effectiveness of the proposed adaptive mechanism in de-identifying AM process data. It is worth noting that the global k value will be the only hyperparameter to tune for the benchmark method.

5.2.2 Two Testing Scenarios

Two different testing scenarios were designed to evaluate the performance of the ADDAM algorithm.

Scenario I: This scenario aimed at evaluating both the data utility and privacy by applying the ADDAM algorithm exclusively to Part 1, where there are both anomaly and theta labels. This scenario simulates a single, independent user who is applying the ADDAM algorithm to their dataset before data sharing.

Scenario II: This scenario was designed to evaluate the effect of additional instantaneous print orientations on the privacy preserving abilities of the ADDAM, as well as to evaluate the utility preservation abilities when aggregating two datasets. This is simulating the collaboration of two users, or a single user leveraging two datasets, to de-identify the thermal process data. Ultimately providing further validation to the results from the first scenario, as there were limited print orientations available within the first test, as well as providing an evaluation on the performance of ADDAM when aggregating multiple datasets.

5.2.3 Data Splitting for Evaluation

For both previously described scenarios, 30% of the sample images were used as the reference image set ($\mathcal{R}$) for the de-identification process, which simulates an independent reference or gallery set that shared a similar distribution to the de-identification data. The remaining images were used to as the sample images ($\mathcal{I}$). More specifically, for Part 1, 30% of the *healthy* melt pool images (*Class = 0*) were used to form $\mathcal{R}$. This is a similar tactic to those used in [3], where the distribution of the normal melt pools is leveraged to identify abnormal melt pools. However, for Part 2 there is no normal and abnormal class labels, so the reference data ($\mathcal{R}$) is taken by randomly sampling 30% of the original melt pool images. This data splitting method is described in detail in Figure 8.

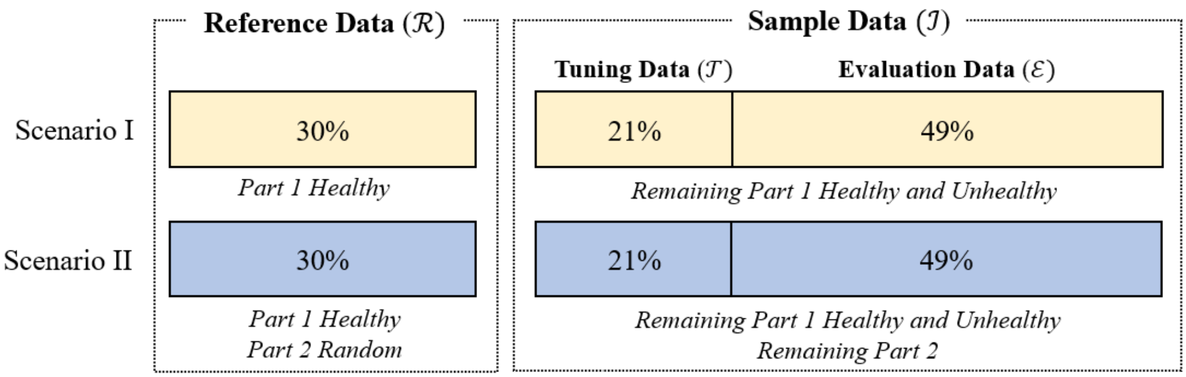

Reference Data ($\mathcal{R}$)
Sample Data ($\mathcal{J}$)
Tuning Data ($\mathcal{T}$)
Evaluation Data ($\mathcal{E}$)
Scenario I
30%
21%
49%
Part 1 Healthy
Remaining Part 1 Healthy and Unhealthy
Scenario II
30%
21%
49%
Part 1 Healthy
Part 2 Random
Remaining Part 1 Healthy and Unhealthy
Remaining Part 2

Figure 8: Data splitting for scenario I and scenario II

In addition, the Maximum Mean Discrepancy (MMD) [53] can be leveraged to verify the similarity of the distribution between the reference set and the de-identification set. The MMD is essentially defined as the distance between the feature mean of two distributions. This similarity metric has been commonly leveraged in transfer learning applications to determine the distance, or similarity, between the source and target domains [11], and can be used as a loss function in deep learning applications [58]. The calculated MMD scores between the reference set and the sample sets for both testing scenarios are summarized in Table 3. In general, the lower MMD score is, the smaller the distance between the feature means of the two datasets will be. It can be observed that the MMD scores for both testing scenarios are only 1.41% and 0.97% of the MMD score between the distributions of two fabricated parts.

Table 3: MMD evaluation of reference and sample sets for each scenario

Linear MMD	**Scenario I**	**Scenario II**	**Part 1 and Part 2**
	6.088	4.159	430.750

Furthermore, from the sample image set ($\mathcal{J}$), 30% of the images were randomly sampled and used as a tuning set ($\mathcal{T}$) to tune both the ADDAM user-defined hyperparameters (M and Δl) and the global k nearest neighbor parameter (k). This tuning data is first de-identified using different combinations of the user-defined hyperparameters and is then evaluated each time using an SVM classifier for anomaly detection and angular identity detection. The remaining 70% of the sample images were used as an evaluation set ($\mathcal{E}$) to gauge the performance of the optimal user-defined de-identification parameters identified from the tuning process. The evaluation set is de-identified using each of the parameter sets selected from the tuning data. After de-identification, the de-identified evaluation data was split into 80/20 training/testing sets and fed into SVM classifiers to predict anomalies and angular identities, producing the overall UL and PG performance of the de-identification algorithm. This final SVM performance evaluation was performed over 10 iterations and results in an averaged performance for the de-identification algorithm. This entire procedure was repeated for both scenarios, just with either Part 1 independent or Part 1 and Part 2 aggregated datasets, which also dictates if either anomaly-detecting and/or angular identity detecting SVM classifiers are leveraged.

To evaluate the algorithm performance in these two scenarios, an SVM classifier was chosen due to its ability to characterize the non-linear relationships within high-dimensional data. The SVM classifier was used during both the tuning stage and during the final evaluation stage, and the SVM hyperparameters were tuned using grid search cross-validation with a stratified shuffle splitting strategy. In addition, 10 replications were performed for each scenario test, and the average performance across these replications was reported and compared to evaluate model robustness.

5.3 Parameter Tuning

For each image within the sample dataset, there are several parameters to consider, these include the variability explained in the PCs (p) and the user-defined constraints related parameters, i.e., M and Δl. For the p value, the variability explained by the PCs was fixed at 95%. This value was chosen as an adequate level of variation that will reduce the high dimensionality of the data, while simultaneously capturing the explained variance within melt pools. This allows for less computational expensive experimentation while still retaining enough information to identify both the presence of abnormal melt pools and the detection of the print orientation angles. In addition, the user defined inputs, M and Δl, and the benchmark input, k, were evaluated over different ranges of values These ranges were designed to capture a variety of possible values and highlight how varying input values can affect the performance of the ADDAM algorithm and are depicted in Table 4.

Table 4: User defined input parameter ranges

User Defined Input	Candidate Values
Maximum Euclidean Distance (M)	0.25, 0.3, 0.4, 0.50, 0.60, 0.7, 0.8, 0.9, 1.0, 1.1, 1.25, 1.50
Maximum Allowable Layer Range (Δl)	1, 5, 10
Global k nearest neighbors	2, 5, 8, 10, 12, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 125, 150

The user-defined inputs were evaluated based on the tuning data set in terms of both PG and UL, and all the Pareto efficient solutions were found through evaluating the performance metrics on a mesh grid of the two de-identification hyperparameters. The pareto efficient solutions were chosen such that they maximized the increase in privacy, while minimizing the loss of utility. A visualization of the ADDAM tuning process is depicted in Figure 9. It is important to note that due to the limited number of unique angular identities, too high of a distance constraint (M) can lead to a decrease or stagnation in the privacy gain. In addition, larger Δl values can lead to higher privacy gains in some scenarios but can adversely impact usability.

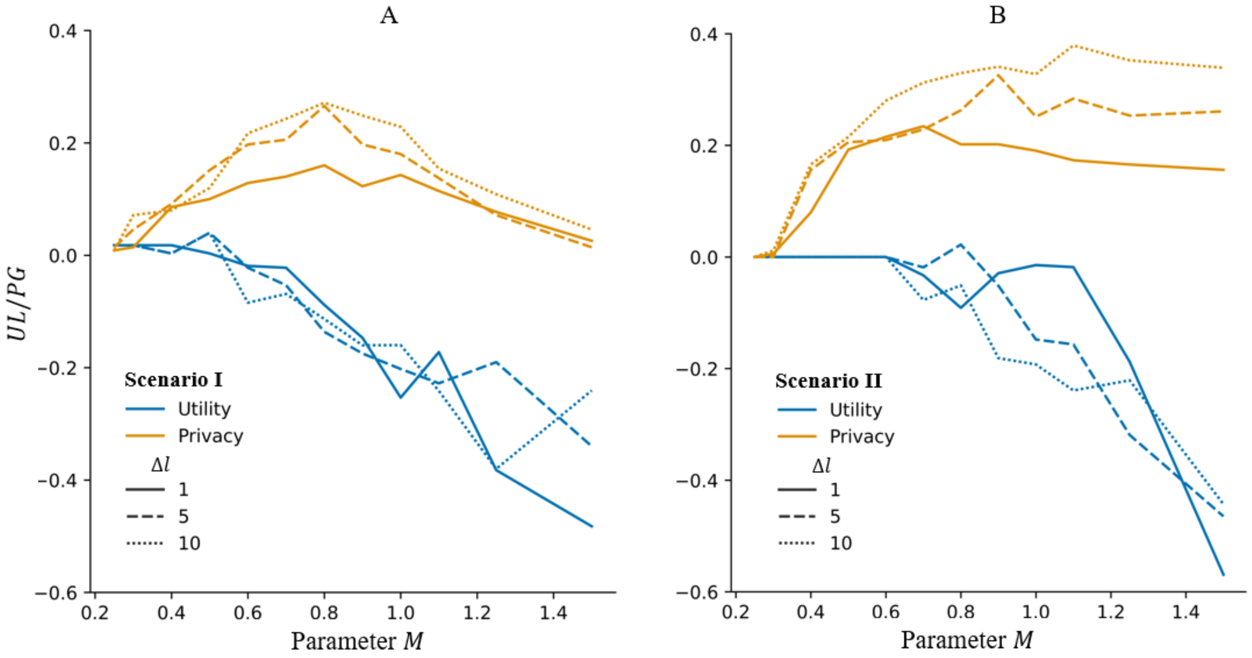


Figure 9: Hyperparameter tuning for distance constraint (M) and layer constraint (Δl) constraints for Scenario I (A) and Scenario II (B)

Furthermore, the benchmark methodology (global k-anonymization) was also tuned to provide comparable evaluations. This included using the same SVM classifier and tuning data split as the ADDAM algorithm. However, this method does not incorporate a balancing parameter, as it directly uses the $k-1$ nearest neighbors to de-identify the image. A visualization of the global k anonymization can be seen in Figure 10, and it is important to note that the general trend exists that increasing PG decreases the UL. This shows that there is a direct, inverse relationship in the privacy gain and utility preserving performance of global anonymization models. In addition, the variation in performance between k values can be attributed to the lack of unique angular identities available in each dataset and imbalanced nature of the dataset.

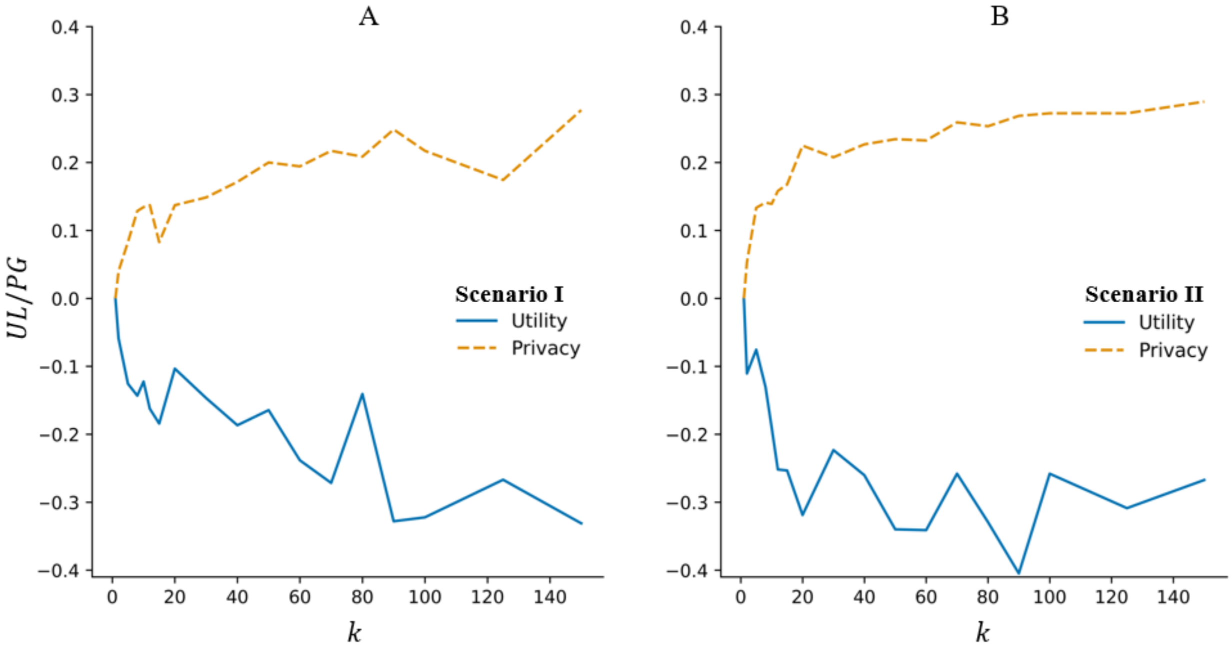


Figure 10: Global k benchmark method tuning for various k values for Scenario I (A) and Scenario II (B)

5.4 Results and Discussion

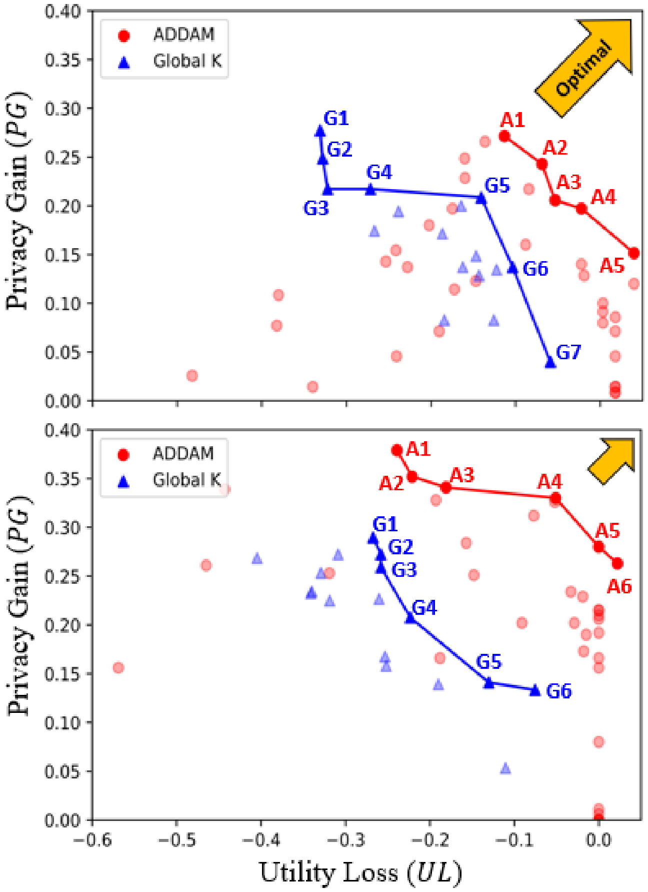
This section details the results from the experimentation described in the previous sections. All tests were evaluated using the same SVM model set-up described previously, to ensure comparability between the proposed and benchmark method.

Firstly, the baseline performance of the SVM model was determined for each of the two testing scenarios. This baseline test highlights the non-anonymized performance of the chosen SVM classifier, which is the maximum data utility that can be achieved. As noted previously, the F1 Score will be the primary metric to evaluate UL. The Accuracy metric will be leveraged when evaluating the angular classification performance, PG. The baseline results for both scenarios are listed in Table 5. In addition, it is important to note that the vPCA extracted features were chosen to evaluate our proposed ADDAM method due to their higher performance over the geometric and thermal features for anomaly detection.

Table 5: Baseline Results using vPCA to extract features and SVM for classification

Scenario	Anomaly Detection (F1-Score)	Angular Detection (Accuracy)
Scenario I	0.859	0.990
Scenario II	0.852	0.970

Secondly, the validation data ($\mathcal{T}$) was leveraged in the ADDAM algorithm and global k algorithm to determine which parameter(s) were optimal for each scenario. As illustrated in Figure 11, each point represents a combination of user-defined inputs (M and Δl) for ADDAM, or a global k level for the benchmark. From here, the pareto optimal points were identified (higher opacity) as the points that lie on the optimal front of the performance area for each scenario. The additional points (lower opacity) are the other combination of parameters which do not lie on the pareto optimal front. These points represent parameters that do not perform optimally using the datasets in Scenario I and II, and are not chosen to evaluate the final test performance. The specific performance and corresponding hyperparameter values are shown in Figure 11. It is important to note that the advantage of the ADDAM algorithm is its ability to preserve data usability, through a smaller $|\text{UL}|$, provided similar privacy gain, PG. From these optimal points, the corresponding hyperparameter sets were selected and then used to de-identify the testing dataset ($\mathcal{E}$) for the benchmark and ADDAM methods.



Scenario A Pareto Points

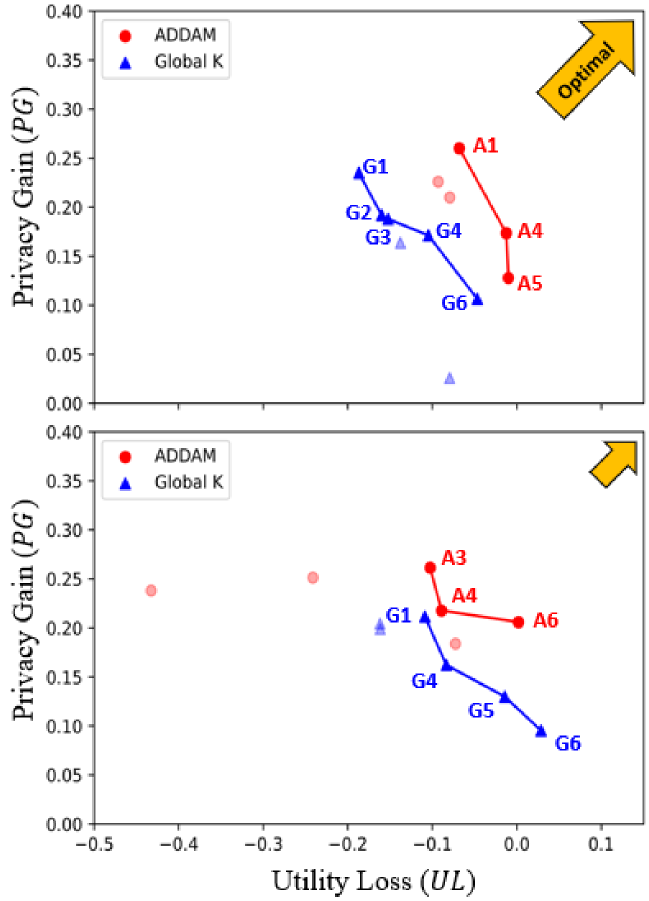
ADDAM					Global k			
#	M	Δl	UL	PG	#	k	UL	PG
A1	0.8	10	-0.11	0.27	G1	150	-0.33	0.28
A2	0.7	10	-0.07	0.24	G2	100	-0.32	0.22
A3	0.7	5	-0.05	0.21	G3	90	-0.33	0.25
A4	0.6	5	-0.02	0.20	G4	80	-0.14	0.21
A5	0.5	5	0.04	0.15	G5	70	-0.27	0.22
					G6	20	-0.10	0.14
					G7	2	-0.06	0.04

Scenario B Pareto Points

ADDAM					Global k			
#	M	Δl	UL	PG	#	k	UL	PG
A1	1.1	10	-0.24	0.38	G1	150	-0.27	0.29
A2	1.3	10	-0.22	0.35	G2	100	-0.26	0.27
A3	0.9	10	-0.18	0.34	G3	70	-0.26	0.26
A4	0.8	10	-0.05	0.33	G4	30	-0.22	0.21
A5	0.8	5	0.00	0.28	G5	8	-0.13	0.14
A6	0.6	10	0.02	0.26	G6	5	-0.08	0.13

Figure 11: Pareto front comparison during the parameter tuning (using $\mathcal{T}$) for the ADDAM and global k in Scenario I (A) and Scenario II (B). This details all possible solutions for the different combinations of tuning parameters for ADDAM, and highlights the pareto optimal solutions, which are detailed in the corresponding tables.

The final phase of experimentation takes the pareto optimal set of the hyperparameter values identified in the tuning stage and applies them to the held-out evaluation data $\mathcal{E}$ to determine an averaged performance in both PG and UL. This evaluation is similar to the tuning results depicted in Figure 11, however these represent the optimal combination of parameters used on the hold-out testing data, ultimately representing the final performance. The ADDAM algorithm again outperforms the benchmark method for both testing scenarios, which is detailed in Figure 12. The Pareto optimal values found from the evaluation data ($\mathcal{E}$) were better positioned to minimize |UL| and maximize PG for ADDAM, as compared to global k. These results show that the ADDAM algorithm uniformly outperforms the benchmark global k method.



Scenario A Pareto Points

ADDAM					Global *k*			
#	**M**	**Δl**	**UL**	**PG**	**#**	**k**	**UL**	**PG**
A1	0.8	10	-0.07	0.26	G1	150	-0.19	0.23
A4	0.6	5	-0.01	0.17	G2	100	-0.16	0.19
A5	0.5	5	-0.01	0.13	G3	90	-0.15	0.19
					G4	80	-0.10	0.17
					G6	20	-0.05	0.11

Scenario B Pareto Points

ADDAM					Global *k*			
#	**M**	**Δl**	**UL**	**PG**	**#**	**k**	**UL**	**PG**
A3	0.9	10	-0.10	0.26	G1	150	-0.11	0.21
A4	0.8	10	-0.09	0.22	G4	30	-0.08	0.16
A6	0.6	10	0.02	0.21	G5	8	-0.01	0.13
					G6	5	0.03	0.09

Figure 12: Pareto front comparison evaluating testing data ($\mathcal{E}$) using pareto efficient solutions identified in Figure 10, for ADDAM and global k in Scenario I (A) and Scenario II (B). This details the performance of the optimally selected tuning parameters from Figure 11 on the testing data, highlighting the overall performance of the proposed ADDAM algorithm.

Furthermore, in Scenario I, the ADDAM algorithm can achieve a comparable or slightly larger PG, without sacrificing as nearly as much data usability as the global k method. This trend is present when implementing the ADDAM algorithm in both the tuning and evaluation stages. In addition, for Scenario II the ADDAM algorithm was able to achieve a noticeable higher privacy gain value, i.e., PG, while maintaining a comparable, and even slightly better, utility loss than the benchmark method. This reinforces the effectiveness of the ADDAM algorithm in practical applications, where complex part geometries would be leveraged in the de-identification. This would provide more diverse angular identities, leading to more improved de-identification results. From both testing scenarios, the ADDAM algorithm was able to outperform the benchmark method in at least one or both optimization objectives. The better performance in utility preservation and increased data privacy of the ADDAM algorithm can be explained through the adaptive de-identification approach. With ADDAM, the user is maximizing the features preserved in the abnormal melt pools, because these images will receive lower, or even zero, level of de-identification. This effectively preserves the features that define the abnormalities. On the other hand, in the benchmark method with global k, the k-closest neighbors were chosen as a constant optimal value, which does not provide the de-identification flexibility to abnormal images. This, as a result, will blur the distinction between the healthy and abnormal melt pool images, sacrificing the AM data utility in anomaly detection.

In a practical application, these results would provide the AM user with the ability to leverage an optimal set of solutions and optimize a de-identification algorithm that best suits their needs. This can be primarily attributed to the pareto front evaluation technique, which provides an optimal set of solutions and allows the user to evaluate the tradeoff between utility preservation and data privacy. From here, a user can evaluate these optimal solutions and decide if they want to prioritize de-identification, utility preservation, or find a balance. This allows the user an additional level of customization to better meet their specific application needs.

6. Conclusion and Future Work

In conclusion, this paper proposes a novel, adaptive approach named the ADDAM methodology to achieving de-identification of design information for AM thermal process data, resulting in secure, de-identified AM process data that can be leveraged for the development of more robust *in-situ* defect detection models. This new adaptive de-identification approach outperforms the traditional global approaches to achieving dataset privacy. Ultimately improving overall dataset privacy (20-30% improvement), while sacrificing a limited amount of data utility (0-10% maximum loss in usability) on the controlled dataset. This creates a stronger defense against IP theft while still allowing AM users to aggregate data, overcoming some of the challenges posed by limited process data for robust process-defect modeling for SMM. Furthermore, the ADDAM algorithm was evaluated on thermal process data collected from a DED process, however, the adaptive framework can easily be expanded beyond DED systems. Many different metal-based AM systems could collect very similar thermal process data, and the adaptive approach itself provides a novel method for de-identifying AM process data, which tends to share the same characteristics of being unbalanced and containing a limited number of unique identities.

There are a few directions that remain open for future research. Firstly, the inclusion of additional angular identities provides a potentially effective improvement in the ADDAM algorithm performance. This includes evaluating the potential effects of using infill orientation angles that are not based on a unidirectional infill pattern or a free-formed component. In addition, leveraging larger datasets that reflect more complex part geometries will provide a more diverse reference set, which may result in stronger de-identification per image. This will ultimately translate into stronger dataset-level data privacy, and be reflective of practical applications. Furthermore, with an increased diversity of angular identities, a potential improvement for the evaluation method would be to apply a regression-based evaluation of the angular identities. This would provide a continuous-valued result, which could provide a more accurate evaluation of the angular identity detection. Secondly, the proposed ADDAM algorithm is aimed at providing a melt-pool wise data privacy, which will provide data privacy while achieving an elevated level of data utility preservation. Future research can potentially develop an additional, compounding privacy measures to further protect against re-identification attacks on a layer-wise level. This could involve incorporating additional image-augmentation measures and layer-wise anonymization techniques to the proposed adaptive de-identification method to achieve larger gains in data privacy. Finally, the adaptive approach to de-identification can be applied to other applications, outside of the AM domain. The ADDAM methodology implements a novel adaptive approach to de-identification that can be beneficial to achieving improved data privacy in different applications, especially where the traditional global k-anonymization approaches may not be as effective. This includes instances where the dataset may not have a large number of unique identities or instances there are additional features available that can be extracted and leveraged to enhance the data privacy through similarity space construction.

Acknowledgements

This work was partially sponsored by National Science Foundation CMMI-2046515.

References

[1] M. Khanzadeh, W. Tian, A. Yadollahi, H. R. Doude, M. A. Tschopp, and L. Bian, "Dual process monitoring of metal-based additive manufacturing using tensor decomposition of thermal image streams," *Addit Manuf*, vol. 23, no. July, pp. 443–456, 2018, doi: 10.1016/j.addma.2018.08.014.

[2] M. Khanzadeh, S. Chowdhury, M. Marufuzzaman, M. A. Tschopp, and L. Bian, "Porosity prediction: Supervised-learning of thermal history for direct laser deposition," *J Manuf Syst*, vol. 47, no. January, pp. 69–82, 2018, doi: 10.1016/j.jmsy.2018.04.001.

[3] S. H. Seifi, W. Tian, H. Doude, M. A. Tschopp, and L. Bian, "Layer-Wise Modeling and Anomaly Detection for Laser-Based Additive Manufacturing," *Journal of Manufacturing Science and Engineering, Transactions of the ASME*, vol. 141, no. 8, pp. 1–12, 2019, doi: 10.1115/1.4043898.

[4] J. Qin, F. Hu, Y. Liu, P. Witherell, C. C. L. Wang, D. W. Rosen, T. W. Simpson, *et al.*, "Research and application of machine learning for additive manufacturing," *Additive Manufacturing*, vol. 52. Elsevier B.V., Apr. 01, 2022. doi: 10.1016/j.addma.2022.102691.

[5] C. Liu, W. Tian, and C. Kan, "When AI meets additive manufacturing: Challenges and emerging opportunities for human-centered products development," *J Manuf Syst*, 2022, doi: 10.1016/j.jmsy.2022.04.010.

[6] J. Patel, Ed., *Data-Driven Modeling for Additive Manufacturing of Metals: Proceedings of a Workshop*. Washington, D.C.: National Academies Press, 2019. doi: 10.17226/25481.

[7] K. Aggour, R. Aman, T. Bell, C. Browne, R. Casukhela, M. Clemente, K. Cobb, *et al.*, "Strategic Guide: Additive Manufacturing Data Management and Schema."

[8] L. Cheng, F. Tsung, and A. Wang, "A statistical transfer learning perspective for modeling shape deviations in additive manufacturing," *IEEE Robot Autom Lett*, vol. 2, no. 4, pp. 1988–1993, Oct. 2017, doi: 10.1109/LRA.2017.2713238.

[9] X. Huang, T. Xie, Z. Wang, L. Chen, Q. Zhou, and Z. Hu, "A Transfer Learning-Based Multi-Fidelity Point-Cloud Neural Network Approach for Melt Pool Modeling in Additive Manufacturing," *ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg*, vol. 8, no. 1, Mar. 2022, doi: 10.1115/1.4051749.

[10] J. Ren, A. T. Wei, Z. Jiang, H. Wang, and X. Wang, "Improved Modeling of Kinematics-Induced Geometric Variations in Extrusion-Based Additive Manufacturing Through Between-Printer Transfer Learning," *IEEE Transactions on Automation Science and Engineering*, 2021, doi: 10.1109/TASE.2021.3063389.

[11] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, *et al.*, "A Comprehensive Survey on Transfer Learning," *Proceedings of the IEEE*, vol. 109, no. 1. Institute of Electrical and Electronics Engineers Inc., pp. 43–76, Jan. 01, 2021. doi: 10.1109/JPROC.2020.3004555.

[12] R. McCann, M. A. Obeidi, C. Hughes, É. McCarthy, D. S. Egan, R. K. Vijayaraghavan, A. M. Joshi, *et al.*, "In-situ sensing, process monitoring and machine control in Laser Powder Bed

Fusion: A review," *Additive Manufacturing*, vol. 45. Elsevier B.V., Sep. 01, 2021. doi: 10.1016/j.addma.2021.102058.

[13] M. A. Tschopp, "A Methodology for Predicting Porosity From Thermal Imaging of Melt Pools in Additive Manufacturing Thin Wall Sections," in *ASME 2017 12th international manufacturing science and engineering conference*, 2017, pp. 1–10. doi: 10.1115/MSEC2017-2909.

[14] M. Khanzadeh, S. Chowdhury, M. A. Tschopp, H. R. Doude, M. Marufuzzaman, and L. Bian, "In-situ monitoring of melt pool images for porosity prediction in directed energy deposition processes," *IISE Trans*, vol. 51, no. 5, pp. 437–455, 2019, doi: 10.1080/24725854.2017.1417656.

[15] Q. Tian, S. Guo, E. Melder, L. Bian, and W. "Grace" Guo, "Deep Learning-Based Data Fusion Method for In Situ Porosity Detection in Laser-Based Additive Manufacturing," *J Manuf Sci Eng*, vol. 143, no. 4, pp. 1–14, 2021, doi: 10.1115/1.4048957.

[16] L. Scime, D. Siddel, S. Baird, and V. Paquit, "Layer-wise anomaly detection and classification for powder bed additive manufacturing processes: A machine-agnostic algorithm for real-time pixel-wise semantic segmentation," *Addit Manuf*, vol. 36, Dec. 2020, doi: 10.1016/j.addma.2020.101453.

[17] M. Mahmoudi, A. A. Ezzat, and A. Elwany, "Layerwise Anomaly Detection in Laser Powder-Bed Fusion Metal Additive Manufacturing," *Journal of Manufacturing Science and Engineering, Transactions of the ASME*, vol. 141, no. 3, Mar. 2019, doi: 10.1115/1.4042108.

[18] M. N. Esfahani, M. M. Bappy, L. Bian, and W. Tian, "In-situ layer-wise certification for direct laser deposition processes based on thermal image series analysis," *J Manuf Process*, vol. 75, pp. 895–902, Mar. 2022, doi: 10.1016/j.jmapro.2021.12.041.

[19] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10. pp. 1345–1359, 2010. doi: 10.1109/TKDE.2009.191.

[20] S. Liu, A. P. Stebner, B. B. Kappes, and X. Zhang, "Machine learning for knowledge transfer across multiple metals additive manufacturing printers," *Addit Manuf*, vol. 39, Mar. 2021, doi: 10.1016/j.addma.2021.101877.

[21] J. Francis, A. Sabbaghi, M. Ravi Shankar, M. Ghasri-Khouzani, and L. Bian, "Efficient distortion prediction of additively manufactured parts using Bayesian model transfer between material systems," *Journal of Manufacturing Science and Engineering, Transactions of the ASME*, vol. 142, no. 5, May 2020, doi: 10.1115/1.4046408.

[22] Q. Hu, R. Chen, H. Yang, and S. Kumara, "Privacy-Preserving Data Mining for Smart Manufacturing," *Smart Sustain Manuf Syst*, vol. 4, no. 2, p. 20190043, 2020, doi: 10.1520/ssms20190043.

[23] P. Samarati and L. Sweeney, "Protecting Privacy when Disclosing Information: k-Anonymity and Its Enforcement through Generalization and Suppression," *IEEE Transaction on Knowledge and Data Engineering*, vol. 13, no. 6, pp. 1010–1027, 2001, doi: doi.org/10.1109/69.971193.

[24] M. N. Islam, Y. Tu, M. I. Hossen, S. Guo, and X. Hei, "A Survey on Limitation, Security and Privacy Issues on Additive Manufacturing," Mar. 2021, [Online]. Available: http://arxiv.org/abs/2103.06400

[25] M. Yampolskiy, T. R. Andel, J. T. McDonald, W. B. Glisson, and A. Yasinsac, “Intellectual property protection in Additive Layer Manufacturing: Requirements for secure outsourcing,” in *ACM International Conference Proceeding Series*, Dec. 2014, vol. 12-December-2014. doi: 10.1145/2689702.2689709.

[26] S. R. Chhetri, N. Rashid, S. Faezi, and M. A. al Faruque, “Security trends and advances in manufacturing systems in the era of industry 4.0,” in *IEEE/ACM International Conference on Computer-Aided Design, Digest of Technical Papers, ICCAD*, Dec. 2017, vol. 2017-November, pp. 1039–1046. doi: 10.1109/ICCAD.2017.8203896.

[27] F. Tao, Q. Qi, A. Liu, and A. Kusiak, “Data-driven smart manufacturing,” *J Manuf Syst*, vol. 48, pp. 157–169, Jul. 2018, doi: 10.1016/j.jmsy.2018.01.006.

[28] A. al Mamun, C. Liu, C. Kan, and W. Tian, “Real-time process authentication for additive manufacturing processes based on in-situ video analysis,” in *Procedia Manufacturing*, 2021, vol. 53, pp. 697–704. doi: 10.1016/j.promfg.2021.06.068.

[29] S. E. Zeltmann, N. Gupta, N. G. Tsoutsos, M. Maniatakos, J. Rajendran, and R. Karri, “Manufacturing and Security Challenges in 3D Printing,” *JOM*, vol. 68, no. 7, pp. 1872–1881, Jul. 2016, doi: 10.1007/s11837-016-1937-7.

[30] S. R. Chhetri, A. Canedo, and M. A. al Faruque, “KCAD: Kinetic cyber-attack detection method for cyber-physical additive manufacturing systems,” in *IEEE/ACM International Conference on Computer-Aided Design, Digest of Technical Papers, ICCAD*, Nov. 2016, vol. 07-10-November-2016. doi: 10.1145/2966986.2967050.

[31] S. Murthy, A. A. Bakar, F. A. Rahim, and R. Ramli, “A Comparative Study of Data Anonymization Techniques,” *IEEE 5th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performanceand Smart Computing (HPSC), and IEEE Intl Conference on Intelligent Data and Security*, 2019.

[32] C. Fontaine and F. Galand, “A survey of homomorphic encryption for nonspecialists,” *EURASIP J Inf Secur*, vol. 2007, 2007, doi: 10.1155/2007/13801.

[33] J. Gatlin, S. Belikovetsky, Y. Elovici, A. Skjellum, J. Lubell, P. Witherell, and M. Yampolskiy, “Encryption is futile: Reconstructing 3D-printed models using the power side-channel,” in *ACM International Conference Proceeding Series*, Oct. 2021, pp. 135–147. doi: 10.1145/3471621.3471850.

[34] L. Sweeney, “K-Anonymity: A Model for Protecting Privacy,” *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, vol. 10, no. 5, pp. 557–570, 2002, [Online]. Available: https://doi.org/10.1142/S0218488502001648

[35] S. Zhong, Z. Yang, and R. N. Wright, “Privacy-Enhancing k-Anonymization of Customer Data,” *Procedings of the twenty-fourth ACM SIGMOD-SIGACT-SIGART Symposium on principals of database system*, pp. 139–147, 2005, doi: doi.org/10.1145/1065167.1065185.

[36] B. S. Bhati, J. Ivanchev, I. Bojic, A. Datta, and D. Eckhoff, “Utility-Driven k-Anonymization of Public Transport User Data,” *IEEE Access*, vol. 9, pp. 23608–23623, 2021, doi: 10.1109/ACCESS.2021.3055505.

[37] J. Domingo-Ferrer and V. Torra, “A critique of k-anonymity and some of its enhancements,” *ARES 2008 - 3rd International Conference on Availability, Security, and Reliability, Proceedings*, pp. 990–993, 2008, doi: 10.1109/ARES.2008.97.

[38] K. LeFevre, D. J. DeWitt, and R. Ramakrishnan, “Mondrian multidimensional K-anonymity,” *Proc Int Conf Data Eng*, vol. 2006, p. 25, 2006, doi: 10.1109/ICDE.2006.101.

[39] J. L. Lin and M. C. Wei, “An efficient clustering method for k-anonymization,” *ACM International Conference Proceeding Series*, vol. 331, pp. 46–50, 2008, doi: 10.1145/1379287.1379297.

[40] S. Ni, M. Xie, and Q. Qian, “Clustering based k-anonymity algorithm for privacy preservation,” *International Journal of Network Security*, vol. 19, no. 6, pp. 1062–1071, 2017, doi: 10.6633/IJNS.201711.19(6).23.

[41] E. M. Newton, L. Sweeney, and B. Malin, “Preserving privacy by de-identifying face images,” *IEEE Trans Knowl Data Eng*, vol. 17, no. 2, pp. 232–243, Feb. 2005, doi: 10.1109/TKDE.2005.32.

[42] R. Gross, E. Airoldi, B. Malin, and L. Sweeney, “Integrating utility into face de-identification,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 3856 LNCS, pp. 227–242, 2006, doi: 10.1007/11767831_15.

[43] R. Gross, L. Sweeney, J. Cohn, F. de La Torre, and S. Baker, “Face De-Identification,” *In: Senior A. (eds) Protecting Privacy in Video Surveillance*, 2009, doi: 10.1007/978-1-84882-301-3_8.

[44] L. Du, M. Yi, E. Blasch, and H. Ling, “GARP-face: Balancing privacy protection and utility preservation in face de-identification,” *IJCB 2014 - 2014 IEEE/IAPR International Joint Conference on Biometrics*, 2014, doi: 10.1109/BTAS.2014.6996249.

[45] A. Jourabloo, X. Yin, and X. Liu, “Attribute preserved face de-identification,” *Proceedings of 2015 International Conference on Biometrics, ICB 2015*, pp. 278–285, 2015, doi: 10.1109/ICB.2015.7139096.

[46] R. Gross, L. Sweeney, F. de La Torre, and S. Baker, “Model-based face de-identification,” in *2006 Conference on Computer Vision and Pattern Recognition Workshops*, 2006, vol. 2006, p. 161. doi: 10.1109/CVPRW.2006.125.

[47] L. Meng and Z. Sun, “Face De-identification with perfect privacy protection,” *2014 37th International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2014 - Proceedings*, no. May, pp. 1234–1239, 2014, doi: 10.1109/MIPRO.2014.6859756.

[48] T. Li and L. Lin, “AnonymousNet: Natural face de-identification with measurable privacy,” *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, vol. 2019-June, pp. 56–65, 2019, doi: 10.1109/CVPRW.2019.00013.

[49] B. Meden, Z. Emersic, V. Struc, and P. Peer, “k -Same-Net : Neural-Network-Based Face Deidentification,” *2017 International Conference and Workshop on Bioinspired Intelligence (IWOBI)*, 2017, doi: 10.1109/IWOBI.2017.7985521.

[50] T. Nakamura, Y. Sakuma, and H. Nishi, “Face image anonymization as an application of multidimensional data K-anonymizer,” *Proceedings - 2019 7th International Symposium on Computing and Networking Workshops, CANDARW 2019*, pp. 155–161, 2019, doi: 10.1109/CANDARW.2019.00035.

[51] E. M. Newton, L. Sweeney, and B. Malin, “Preserving Privacy by De-Identifying Face Images.”

[52] J. Brickell and V. Shmatikov, “The Cost of Privacy: Destruction of Data-Mining Utility in Anonymized Data Publishing,” *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 70–78, 2008.

[53] B. Zhang, C. Chen, and L. Wang, “Privacy-preserving Transfer Learning via Secure Maximum Mean Discrepancy,” Sep. 2020, [Online]. Available: http://arxiv.org/abs/2009.11680

[54] S. Uguroglu and J. Carbonell, “Feature Selection for Transfer Learning,” 2011.

[55] H. A. Abbass, R. Sarker, and C. Newton, “PDE: A Pareto-frontier differential evolution approach for multi-objective optimization problems,” in *Proceedings of the IEEE Conference on Evolutionary Computation, ICEC*, 2001, vol. 2, pp. 971–978. doi: 10.1109/cec.2001.934295.

[56] W. Tian, J. Ma, and M. Alizadeh, “Energy consumption optimization with geometric accuracy consideration for fused filament fabrication processes,” *International Journal of Advanced Manufacturing Technology*, vol. 103, no. 5–8, pp. 3223–3233, Aug. 2019, doi: 10.1007/s00170-019-03683-5.

[57] G. Menardi and N. Torelli, “Training and assessing classification rules with imbalanced data,” *Data Min Knowl Discov*, vol. 28, no. 1, pp. 92–122, Jan. 2014, doi: 10.1007/s10618-012-0295-5.

[58] G. K. Dziugaite, D. M. Roy, and Z. Ghahramani, “Training generative neural networks via Maximum Mean Discrepancy optimization,” May 2015, [Online]. Available: http://arxiv.org/abs/1505.03906

List of Figures

List of Tables