

RESEARCH ARTICLE | FEBRUARY 14 2025

## Event-based vision in laser welding: An approach for process monitoring

Special Collection: [Proceedings of the International Congress of Applications of Lasers & Electro-Optics \(ICALEO 2024\)](#)

Patricia M. Dold ; Praveen Nadkarni ; Meiko Boley ; Valentin Schorb ; Lili Wu ;  
Fabian Steinberg ; Peter Burggräf ; Ralf Mikut 



*J. Laser Appl.* 37, 012040 (2025)

<https://doi.org/10.2351/7.0001526>




### Articles You May Be Interested In

Cell detection with convolutional spiking neural network for neuromorphic cytometry

*APL Mach. Learn.* (May 2024)


Neuromorphic MEMS sensor network

*Appl. Phys. Lett.* (April 2019)




**Journal of  
Laser Applications**


[Learn More](#)




**RAPID TIME  
TO ACCEPTANCE**




**COMMUNITY  
DRIVEN**



**EXPANSIVE  
COVERAGE**



**PRESTIGIOUS  
EDITORIAL BOARD**



**EXTENSIVE  
MARKETING**

# Event-based vision in laser welding: An approach for process monitoring

Cite as: J. Laser Appl. 37, 012040 (2025); doi: 10.2351/7.0001526

Submitted: 4 June 2024 · Accepted: 15 January 2025 ·

Published Online: 14 February 2025



Patricia M. Dold,<sup>1,2,a)</sup> Praveen Nadkarni,<sup>1,3</sup> Meiko Boley,<sup>1</sup> Valentin Schorb,<sup>1,2</sup> Lili Wu,<sup>3</sup>   
Fabian Steinberg,<sup>3</sup> Peter Burggräf,<sup>3</sup> and Ralf Mikut<sup>2</sup>

## AFFILIATIONS

<sup>1</sup>Bosch Research, Robert Bosch GmbH, Robert-Bosch-Campus 1, Renningen 71272, Germany

<sup>2</sup>Institute for Automation and Applied Informatics (IAI), Karlsruhe Institute of Technology (KIT), Hermann-von-Helmholtz-Platz 1, Eggenstein-Leopoldshafen 76344, Germany

<sup>3</sup>Chair for International Production Engineering and Management (IPEM), University of Siegen, Siegener Straße 152, Kreuztal 57223, Germany

**Note:** Paper published as part of the special topic on Proceedings of the International Congress of Applications of Lasers & Electro-Optics 2024.

<sup>a)</sup>Author to whom correspondence should be addressed; electronic mail: [patricia.dold@de.bosch.com](mailto:patricia.dold@de.bosch.com)

## ABSTRACT

Event cameras differ from conventional frame cameras: Rather than capturing images at a fixed rate, they asynchronously measure pixelwise brightness changes and stream an event sequence that includes time, location, and sign of the brightness changes. Event cameras provide attractive properties such as high temporal resolution and latency (both in the order of ms), very high dynamic range (120 dB), and low power consumptions because no routine operation needs to be recorded. Hence, event cameras have been successfully used, for instance, for object tracking, pose estimation, or image deblurring. Common in-process sensors for process monitoring in laser welding are photodiodes or high-speed frame cameras. Due to the described advantages of event cameras compared to these sensors, this paper for the first time investigates the potential of event cameras in laser welding. We present different promising ways to process the unconventional asynchronous output of event cameras for process monitoring, including time-series representations, as well as image representations. Our laser welding experiments show clear differences for different production scenarios in those event representations. Additionally, we propose a machine learning pipeline for the automatic detection of production errors with event camera.

Key words: event camera, laser welding, process monitoring, machine learning

© 2025 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.2351/7.0001526>

## I. INTRODUCTION

In production, laser welding processes—if not executed properly—are posing a risk of component failure. Consequently, to ensure the reliability of the components, stringent monitoring of the weld quality is needed. One promising option is the in-process monitoring of the welding keyhole, molten pool, or spatters. Monitoring sensors are acoustic sensors,<sup>1</sup> photodiodes,<sup>2,3</sup> x-ray sensors,<sup>4</sup> charge-coupled device (CCD) cameras,<sup>5</sup> and high-speed cameras.<sup>6</sup> The analysis of the sensor signals includes approaches based on machine

learning, such as support vector machines,<sup>7</sup> decision trees (DTs),<sup>8</sup> random forests,<sup>9</sup> or convolutional neural networks (CNNs).<sup>10–12</sup>

Event cameras (ECs), which are inspired by biological mechanisms, measure pixelwise changes in brightness, providing advantages such as high temporal resolution, high dynamic range, low power consumption because no routine operation needs to be recorded, and low latency<sup>13</sup> compared to standard cameras. These properties have led to promising applications in areas such as vibration measurement,<sup>14</sup> drowsiness detection systems,<sup>15</sup> counting of fast

10 March 2025 11:26:21

moving objects,<sup>16</sup> or image deblurring.<sup>17</sup> Depending on the further evaluation algorithm of an application, the asynchronous event data are processed and often transformed into alternative representation.<sup>13</sup> For instance, individual events can be directly handled by event-by-event methods like spiking neural networks.<sup>18,19</sup> However, those networks require special hardware and, therefore, are not considered further. Moreover, events can be represented as event packets, whereby a predefined amount of events is collectively processed.<sup>20</sup> Additionally, events can be accumulated into images that can be easily processed using computer vision algorithms.<sup>21</sup>

As the EC facilitates a more comprehensive understanding of exceptionally swift processes, the advantages of the EC compared to standard cameras for process monitoring in laser welding might be interesting. The adaptability of the EC in different domains emphasizes the potential to provide insights into the dynamic processes of laser welding processes. To the best of our knowledge, no prior publications have investigated the potential of EC for quality monitoring of laser welding processes. Therefore, the present paper includes the following contributions:

- Transformation of the asynchronous events of the EC data to different promising event presentations for in-process quality monitoring of laser welding defects.
- Proposal of different sampling strategies of the event camera data for quality monitoring of laser welding processes.
- Analysis and interpretation of the introduced event representations for different in production occurring laser welding defects.
- Usage of machine learning algorithms including decision trees and convolution neural networks to distinguish the introduced defects.

## II. EVENT CAMERAS

In contrast to standard cameras (e.g., Optronics Cyclone1HS-3500) that acquire images at a fixed rate (e.g., 30 fps), event cameras (ECs) respond to brightness changes asynchronously and independently for every pixel. Thus, the EC streams a data sequence of events. Each event represents a brightness change in a predefined magnitude at a pixel at a particular time. When the change exceeds a threshold, the camera sends an event. The event includes the  $x, y$ -location of the change in the EC observation file with  $x \in \{1, \dots, X\}$  and  $y \in \{1, \dots, Y\}$ , the time  $t$  at which the event occurred, and the 1-bit polarity  $p \in \{-1, 1\}$ . The polarity gives the type of change, being 1 for brightness increase and  $-1$  for brightness decrease. An event  $e$  is represented by the 4-tuple,

$$e = (x, y, t, p). \quad (1)$$

Therefore, the data sequence  $E$  delivered by an EC with  $N$  events is

$$E = \{e_i\}_{i=1}^N. \quad (2)$$

The explained EC encoding is inspired by the spiking nature of biological visual pathways.<sup>22</sup> Compared to standard cameras, the EC offers numerous potential advantages: ECs have an extremely high temporal resolution and low latency<sup>13</sup> (both in the order of

microseconds), which make these sensors react quickly to visual stimuli. Moreover, ECs have a very high dynamic range<sup>23</sup> (120 versus 60 dB of standard cameras) and low power consumption.<sup>24</sup> Therefore, ECs are currently used for feature detection and tracking like corners,<sup>25</sup> fastmoving objects,<sup>16</sup> optical flow estimation,<sup>26</sup> 3D reconstruction,<sup>27</sup> pose estimation,<sup>28</sup> visual-inertial odometry,<sup>29</sup> image reconstruction,<sup>20</sup> motion segmentation,<sup>30</sup> neuromorphic control,<sup>31</sup> or image deblurring.<sup>17</sup> However, the spatial resolution of the EC is generally lower than that of standard cameras.

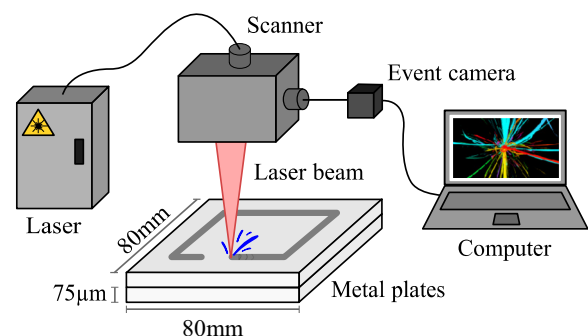
After having seen the advantages and applications of EC in diverse fields, we now look at how the EC can be used for the quality monitoring of laser welding processes.

## III. EXPERIMENTAL SETUP

Based on an industrial welding process, two metal plates were welded together. Therefore, in production occurring scenarios were provoked. The process is observed with the EC, and the aim is to recognize differences in its data. Figure 1 shows the experimental setup. Therefore, a laser beam was directed onto two thin metal plates by a 2D galvanometer scanner. The two metal plates have a dimension of  $80 \times 80 \text{ mm}^2$  and a thickness of  $75 \mu\text{m}$ . As a laser, an IPG YLR-2000-WC fiber laser with an infrared wavelength of  $1070 \text{ nm}$  at a power of  $300 \text{ W}$  (max. power  $2000 \text{ W}$ ) and a speed of  $0.5 \text{ m/s}$  was used. A scanner, a Raylase AS Fiber 30, focused the beam to a diameter of about  $50 \mu\text{m}$  on the work piece surface. An event camera (EC) observed the laser welding process coaxially and in-process. As the EC, a Prophesee Metavision Evaluation Kit 4HD (EVK4) was used. The EC data sequence then is transferred and processed on a computer.

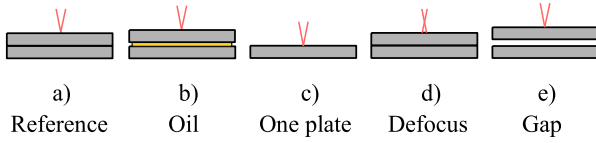
Anticipating scenarios that may arise during production, experiments were conducted. Figure 2 schematically shows the front view of the metal plates of five scenarios. In the *reference* scenario (a), the metal plates were positioned on top of each other. In (b), a thin film of oil was applied between the metal plates, which led to a spattering process. In (c), only *one* plate was welded. In (d), there is a *defocus* effect of the laser beam, and in (e) a *gap* was inserted between the two plates.

Having seen the experimental setup, we look at the resulting EC data. Figure 3 visualizes a sequence of events  $E$  [see Eqs. (1) and (2)]



**FIG. 1.** Experimental setup of the laser welding process of two thin metal plates. An event camera (EC) was used to observe the process *in situ*. The EC data sequence is transferred and processed on a computer.

10 March 2025 11:26:21



**FIG. 2.** Five different production scenarios have been considered: (a) reference, (b) oil, (c) one plate, (d) defocus, and (e) gap.

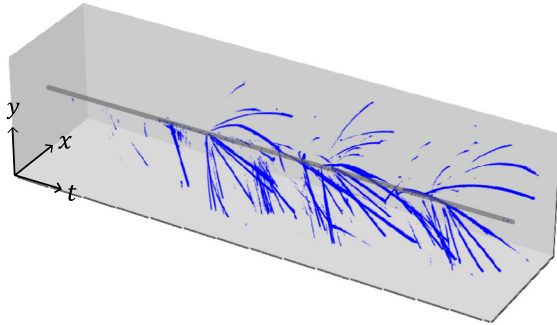
captured during the observation of the laser welding process. The events are shown in the three-dimensional space with the  $x$ ,  $y$ -location and the time  $t$  as axis. The individual events are marked in blue. The position of the laser is given by the gray line. While at the beginning less events occurred, more events were captured afterwards as the process spattered.

#### IV. EVENT REPRESENTATIONS

One of the key questions of event cameras is how to extract meaningful information from the event data. Therefore, the events are transferred to alternative representations that make it easier to observe changes in the laser welding process. In the following, we present three different promising event representations to monitor laser welding processes, namely, event rates (Sec. IV A), projection on the  $x$ ,  $y$ -plane (Sec. IV B), and projection along scanner data (Sec. IV C). Afterward, the resulting representations can be further analyzed, e.g., by classification algorithms.

##### A. Event rate

An event rate transfers a data sequence  $E$  into a time series. More precisely, an event rate gives the number of events that occur in equidistant time intervals per length of the intervals. Mathematically speaking, this means that assume having observed a laser process during the time interval  $T = [t_s, t_e]$ , which results in an EC data sequence  $E$ . Let us divide  $T$  in  $L$  equally long subintervals of length  $\Delta t$ , so  $L = (t_e - t_s)/\Delta t$ .



**FIG. 3.** Visualization of events (blue) during the observation of a laser welding process in the three-dimensional space, with the  $x$ ,  $y$ -location and the time  $t$  as the axis. The laser position is given by the gray line. The polarity  $p$  of the events is neglected.

Then,

$$t_k = t_s + (k - 1) \cdot \Delta t, \quad (3)$$

with  $k \in \{1, \dots, L\}$  defines the start times of the subintervals. Let  $f_i(e)$  be the time of event  $e$ . Then,  $G \subseteq E$  defined as

$$G(t_c) = \{e \mid c_1 \wedge c_2\}, \quad (4)$$

with  $c_1: e \in E$  and  $c_2: t_c \leq f_i(e) < t_c + \Delta t$  includes all events in an interval of length  $\Delta t$  beginning at  $t_c$ . Then, the event rate  $\varepsilon \in N^{1 \times L}$  is given by

$$\varepsilon = \left( \frac{|G(t_k)|}{\Delta t} \right)_{k=1}^L. \quad (5)$$

As the laser process is observed *in situ*, the laser beam always appears at the same position on the EC  $x$ ,  $y$ -plane. The  $x$ ,  $y$ -plane, therefore, lies in a local coordinate system in the center of which the laser beam hits the work piece. In our setup, the laser beam hits the piece in the center of the  $x$ ,  $y$ -plane. To get separate information near the laser beam, e.g., about the keyhole, and information further away from the laser beam, the  $x$ ,  $y$ -plane is divided into two areas: The first area is a circle with the radius  $r$  around the center in the  $x$ ,  $y$ -plane, and the second area is the  $x$ ,  $y$ -plane without the circle. So, an inside event rate  $\varepsilon_i$  and an outside event rate  $\varepsilon_o$  are defined in the following: Let  $d$  define the distance of the  $x$ ,  $y$ -location of an event  $e$  to the center in the  $x$ ,  $y$ -plane. Therefore,  $G_i \subseteq G$  defined as

$$G_i = \{e \mid c_1 \wedge c_2 \wedge d \leq r\} \quad (6)$$

includes all events whose  $x$ ,  $y$ -locations are inside a circle with radius  $r$  around the center in an interval of length  $\Delta t$ . Then, the inside event rate  $\varepsilon_i \in N^{1 \times L}$  is

$$\varepsilon_i = \left( \frac{|G_i(t_k)|}{\Delta t} \right)_{k=1}^L. \quad (7)$$

Similarly,  $G_o \subseteq G$  defined as

$$G_o = \{e \mid c_1 \wedge c_2 \wedge d > r\} \quad (8)$$

includes all events whose  $x$ ,  $y$ -locations are outside the circle.

Then, the outside event rate  $\varepsilon_o \in N^{1 \times L}$  is

$$\varepsilon_o = \left( \frac{|G_o(t_k)|}{\Delta t} \right)_{k=1}^L. \quad (9)$$

##### B. Projection on $x$ , $y$ -plane

This event representation projects events in the  $x$ ,  $y$ -plane, resulting in images of the dimension  $X \times Y$ . In the following, we present three different such  $x$ ,  $y$ -projections, namely, *binary* (Sec. IV B1), *density* (Sec. IV B2), and *temporal* (Sec. IV B3).

### 1. Binary

The binary projection marks all  $x, y$ -positions, where at least one event happens during a time interval. Let  $f_x(e)$  and  $f_y(e)$  be the  $x$  and  $y$  entries of an event  $e$ . Then,

$$C = \{(f_x(e), f_y(e)) \mid c_1\} \quad (10)$$

is the set of  $x, y$ -coordinates at which at least one event occurs. Thus, the entries of the binary projection  $P \in \{0, 1\}^{X \times Y}$  are

$$P_{xy} = \begin{cases} 1, & \text{if } (x, y) \in C, \\ 0, & \text{otherwise.} \end{cases} \quad (11)$$

So  $P$  has the entry 1 if at least one event occurs at position  $(x, y)$  and 0 if there is no event occurrence.

### 2. Density

Additionally to whether at least one event occurs at a  $x, y$ -position, the density projection provides information about how many events occur at a position. Let

$$D(x, y) = \{e \mid c_1 \wedge c_3\} \quad (12)$$

with  $c_3: (x, y) \in C$  being the set of events occurring at position  $(x, y)$ . Then, the entries of the density projection  $Q \in [0, 1]^{X \times Y}$  are

$$Q_{xy} = \begin{cases} \frac{|D(x, y)|}{\hat{D}}, & \text{if } D \neq \emptyset, \\ 0, & \text{otherwise,} \end{cases} \quad (13)$$

with  $\max(|D(x, y)|)$ . So the higher an entry in  $Q$ , the more events occur at its  $(x, y)$ -coordinates.

### 3. Temporal

In addition to whether at least one event occurs at a  $(x, y)$ -position, the temporal projection retains temporal information of events. Let

$$H(x, y) = \{f_t(e) \mid c_1 \wedge c_3\} \quad (14)$$

be the set of times of events occurring at position  $(x, y)$ . Then, the entries of the time projection  $R_1 \in [0, 1]^{X \times Y}$

$$R_{xy,1} = \begin{cases} \frac{\max(H(x, y)) - t_s}{t_e - t_s}, & \text{if } H \neq \emptyset, \\ 0, & \text{otherwise.} \end{cases} \quad (15)$$

So each  $x, y$ -position at which an event occurs is scaled depending on the time the event occurs. If several events occur at position  $(x, y)$ , the maximum time is taken. Similarly, using the minimum time leads to the time projection  $R_2 \in [0, 1]^{X \times Y}$ ,

$$R_{xy,2} = \begin{cases} \frac{\min(H(x, y)) - t_s}{t_e - t_s}, & \text{if } H \neq \emptyset, \\ 0, & \text{otherwise.} \end{cases} \quad (16)$$

### C. Projection along scanner data

This event representation additionally uses recorded scanner data during welding. These data include the  $(x_s, y_s)$ -location on the processing plane of the laser and its corresponding time  $t_s$ . Therefore, a scanner data entry  $s$  is represented by

$$s = (x_s, y_s, t_s). \quad (17)$$

Let  $f_t(s)$  be the time of an entry  $s$ . Then,

$$S = (s_i)_{i=1}^M, \quad (18)$$

with  $f_t(s_k) \leq f_t(s_{k+1})$ ,  $k \in \{1, \dots, M-1\}$  is the sequence of  $M$  scanner data entries with increasing time entries. Let

$$N(k) = \{e \mid c_1 \wedge c_4\}, \quad (19)$$

with  $c_4: f_t(s_k) \leq f_t(e) < f_t(s_{k+1})$  being the set of events occurring between two following time entries of the scanner data.

Let  $f_x(s)$  and  $f_y(s)$  be the  $x_s$  and  $y_s$  entries of  $s$ . Then, the center  $(c_x, c_y)$  of the EC observation field at time  $f_t(e)$  for an event  $e \in N(k)$  is at

$$c_x(e, k) = f_x(s_k) + \frac{f_t(e) - f_t(s_k)}{f_t(s_{k+1}) - f_t(s_k)} \cdot (f_x(s_{k+1}) - f_x(s_k)), \quad (20)$$

in the processing plane of the laser,  $c_y(e, k)$  results analogously. Let  $\gamma(\cdot)$  convert a distance in the laser processing plane into its corresponding amount pixels of the EC observation field. Then, the entries of the projection along scanner data  $V \in \{0, 1\}^{b_x \times b_y}$  with  $b_x = \gamma(\max(f_x(s)) - \min(f_x(s))) + X$  and  $b_y$  accordingly is

$$V_{jl} = \begin{cases} 1, & \text{if } j = \gamma(c_x(e, k) - \min(f_x(s))) + f_x(e) \text{ and} \\ & l = \gamma(c_y(e, k) - \min(f_y(s))) + f_y(e), \\ & \text{where } e \in N(k), k \in \{1, \dots, M-1\}, \\ 0, & \text{otherwise.} \end{cases} \quad (21)$$

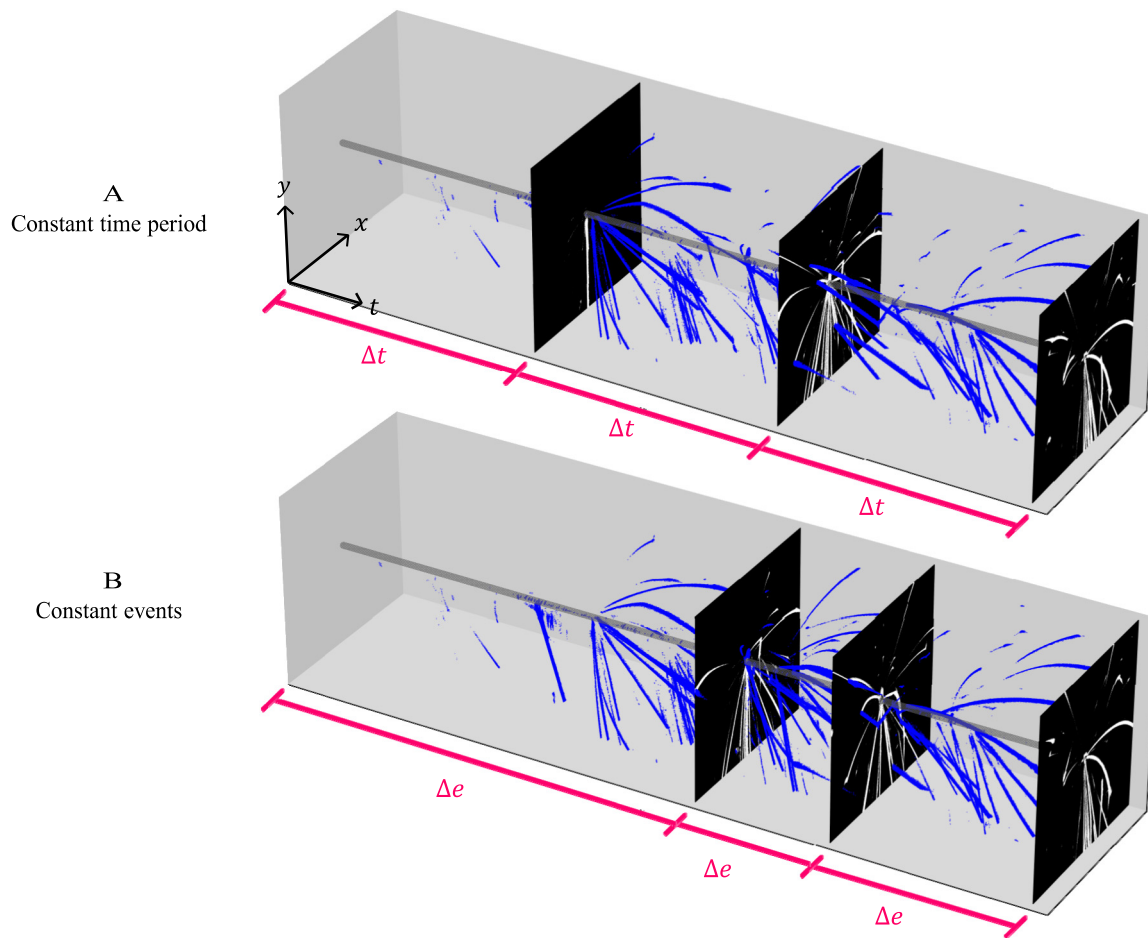
## V. SAMPLING STRATEGIES

After having seen different event representations, the question arises as to when and how often the representations should be generated during a welding process for quality monitoring purposes. Therefore, in the following, we present two different sampling strategies, namely, *constant time period* (Sec. V A) and *constant event rate* (Sec. V B).

### A. Constant time period

Assume having observed a laser process during the time interval  $T$ , which results in an EC data sequence  $E$ . Let us divide  $T$  in equally long subintervals of length  $\Delta t$ . The constant time period sampling strategy then generates the event representations for each subinterval of length  $\Delta t$ .

Figure 4(a) visualizes this sampling strategy. As in Fig. 3, the events (blue) of a sequence of events  $E$  are shown in the three-dimensional space with the  $x, y$ -location and the time  $t$  as



10 March 2025 11:26:21

**FIG. 4.** Sampling strategies: A visualizes the constant time period sampling strategy. So, when the time  $\Delta t$  has passed, an event representation is created resulting in three binary images. B visualizes the constant events sampling strategy. So, when an amount of  $\Delta e$  events have occurred, an event representation is created also resulting in three binary images. It is  $\Delta t = 16$  ms and  $\Delta e = 62000$ .

axis. When the time  $\Delta t$  has passed, a binary projection  $P$  is generated. In the figure, this results in three binary images. A drawback of considering a constant time period as sampling strategy is that even if few or no events occur—meaning that no or few changes are in the laser process—a representation is created. This can be seen in the figure: In the first of the three intervals, there are almost no changes in the process, yet an almost completely black image is created. This drawback is overcome in the next sampling strategy.

### B. Constant events

Assume having observed a laser process during the time interval  $T$ , which results in an EC data sequence  $E$ . Let us divide  $E$  in subsets with the same amount of events  $\Delta e$ . So each entry of the event rate  $e$  has the value  $\Delta e$ . This constant event sampling strategy then generates the event representations for each subset with  $\Delta e$  events.

Figure 4(b) visualizes this sampling strategy. When  $\Delta e$  events have occurred, a binary projection  $P$  is generated. In the figure, three binary images are shown. Thereby, projections are created after differently long time intervals. Compared to Fig. 4(a), the first of the three projections is created later, as fewer events occur at the beginning. To determine the time of an anomaly during quality monitoring, the time after each  $\Delta e$  has to be remembered.

## VI. RESULTS AND DISCUSSION

After gaining understanding about different sampling strategies and event representations, we now look at which differences can be seen in the different event representations for our in production occurring scenarios according to Fig. 1 like gap or spatter (Sec. VI A). Subsequently, to further show the potential of EC data for process monitoring of laser processes, different configurations in which the EC parameters, the process parameters or the

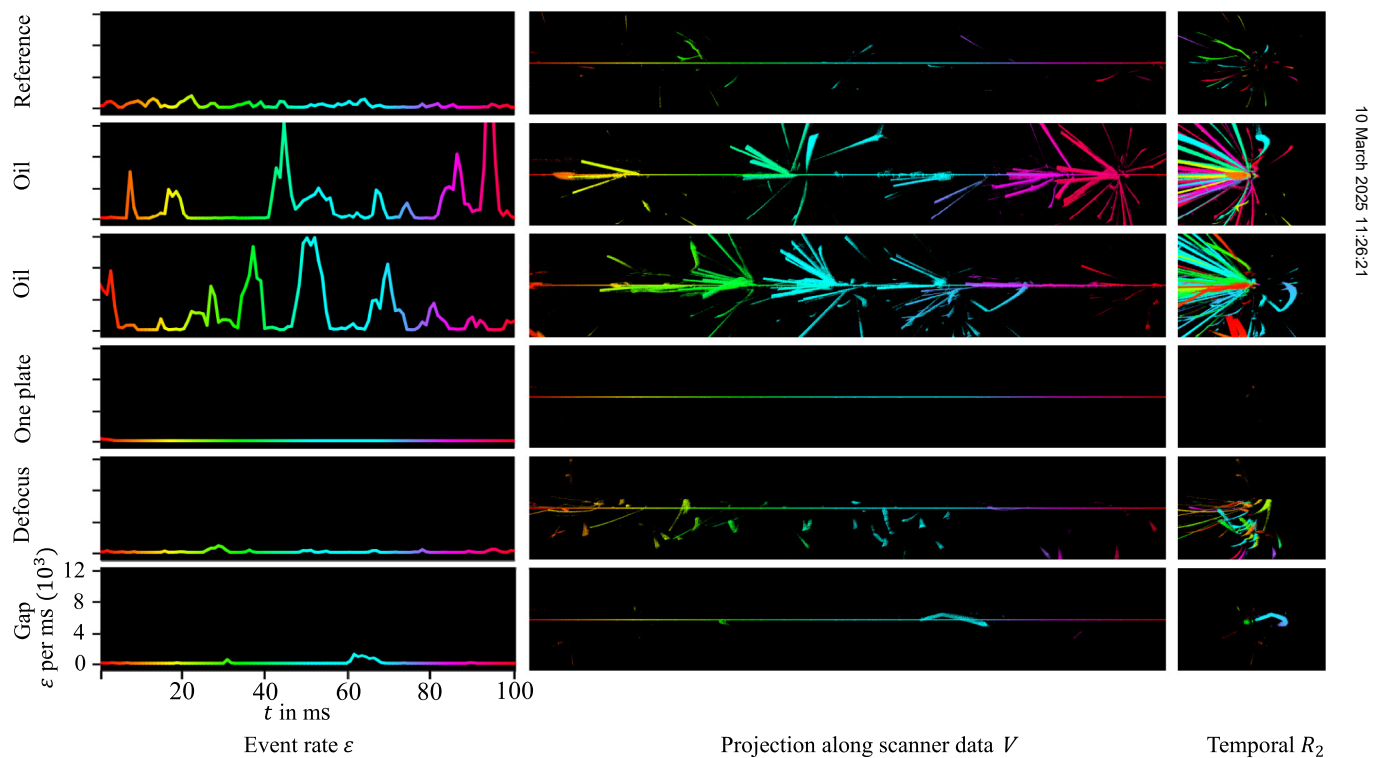
hardware was changed were realized (Sec. VI B). Afterward, we present a machine learning pipeline for quality monitoring including DTs and CNNs (Sec. VI C).

### A. Event representations of different scenarios

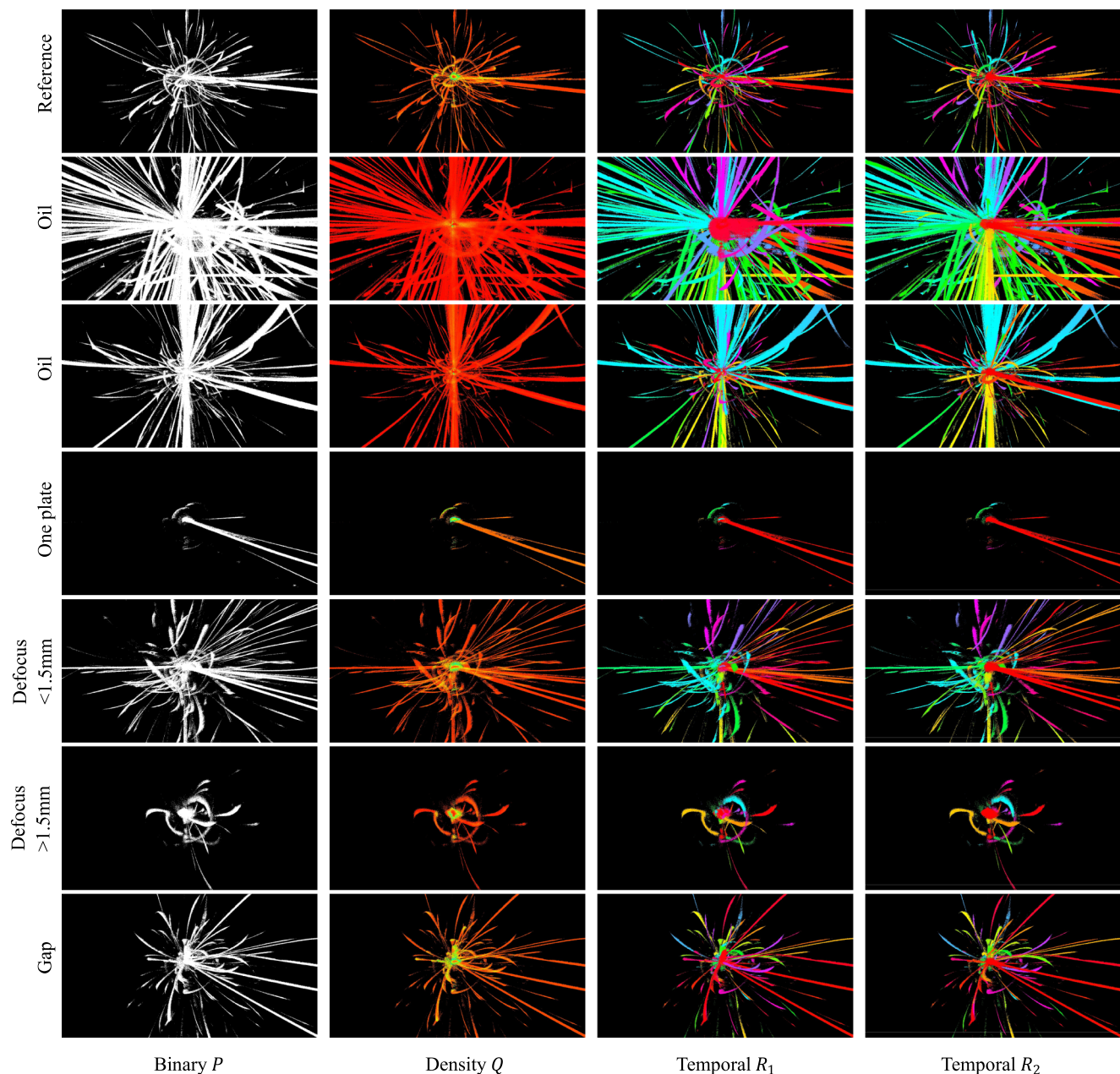
The first column of Fig. 5 shows the event rate  $\epsilon$  determined during the welding process of the introduced scenarios according to Fig. 2. The event rate  $\epsilon$  with  $\Delta t = 1$  ms is shown over the time in ms. The graph line is shown in rainbow colors. The color changes constantly over time. In the first row, the event rate of *reference* shows minor changes over time. The next two rows show *oily* processes. Compared to the reference, the event rate shows high peaks due to the spatters. The event rate of *one plate* shows nearly no events, the event rate of *defocus* shows tiny rises and falls, and the event rate of the *gap* shows few small peaks.

Figure 6 shows the projections on the  $x, y$ -plane determined during the welding process of the introduced scenarios. The first column of Fig. 6 shows the binary projection  $P$ , the second column the density projection  $Q$ , the third column the time projection  $R_1$ , and the last column the time projection  $R_2$ . The images include the

whole image field of the EC of  $1280 \times 720$  pixels. In the binary projection, differences in the images between the scenarios are visible. While in *reference*, a regular spatter pattern with most spatters being completely visible in the image, in *oil*, far more spatters are visible. In *one plate*, nearly no events occur. The fifth row results from *defocusing* smaller than 1.5 mm and the sixth row from *defocusing* greater than 1.5 mm. Therefore, the fifth row shows many thin spatters, whereas the sixth only shows few spatters. In the last row, a *gap* was introduced on one side of the rectangular welding geometry (see Fig. 1). As seen in the event rate (see Fig. 5 first column and last row), few events occur in that scenario. On closer inspection, one can see that few spatters occur on the left-hand side of the image, which is due to the introduced gap. In addition to the information whether or not an event occurred at a certain position in the  $x, y$ -plane, no further information like how many events or when the events occurred is available. This information is additionally shown in the next columns. The images of the density and temporal projections are shown in rainbow colors. So, for instance, for the density projection, the color map starts with red (no events), goes to yellow (medium number of events), and ends with green (most events). In the temporal projection, the



**FIG. 5.** Comparison of different event representations: The first column gives the event rate  $\epsilon$  with  $\Delta t = 1$  ms over the time  $t$  in ms. The second column gives the projections along scanner data  $V$ , and the last column gives the time projections  $R_2$ . For the event representations, the different scenarios (a)–(f) according to Fig. 2 are shown. So every row shows the different event representations of one scenario. Every process lasts 100 ms, and the rainbow coloring gives the time. So the coloring starts with red at 0 ms and changes to green at about 30 ms, to blue at about 60 ms, and finally back to red at the end of the process at 100 ms. Therefore, corresponding points between the event representations can be seen by matching colors.



10 March 2025 11:26:21

**FIG. 6.** Projections on the  $x, y$ -plane: The first column gives the binary projection  $P$ , the second column the density projection  $Q$ , the third column the time projection  $R_1$ , and the last column the time projection  $R_2$ . For those projections, the different categories (a)–(e) according to Fig. 2 are shown. Thereby, the fifth row results from defocusing smaller than 1.5 mm and the sixth row from defocusing greater than 1.5 mm. So every row shows the different projections on the  $x, y$ -plane of a category. The projections are calculated over the events occurring during the whole laser process of the two metal plates (see Fig. 1). The projections of the second, third, and last column have a rainbow coloring. So, for instance, in the density projections red coloring means few events per position, yellow means more, and green means even more events per position. In the temporal projections, the rainbow colors change constantly over time.

rainbow colors change constantly over time as in Fig. 5. With this additional temporal information, it can be seen that the spatters in the one plate scenario occur at the beginning of the process when the laser beam penetrates. Moreover, the other events in the one

plate scenario occur at four time intervals (pink, orange, green, light blue). Those events are due to reflections on a clamping device in the four corners of the welding geometry. At closer look, they are visible in all scenarios.

The second column of Fig. 5 shows the projections along scanner data  $V$  determined during the welding process of the introduced scenarios. The rainbow colors refer to the time, as in the event rate of Fig. 5. The horizontal, rainbow-colored lines give the position of the laser. While the event rate of  $A$  gives the amount of occurred events,  $B$  shows what happened spatially. For example, the event rate of reference shows small rises and falls. In the projection along the scanner data, it is visible that these events are caused by the regular occurrence of small spatters. Moreover, the high peaks in the event rate in the oil scenario are caused by strong ejections with several spatters that fly both forward and backward. So the projections along the scanner data  $V$  show the direction of the spatter relative to the global work piece position.

The third column of Fig. 5 shows the time projections  $R_2$  determined during the welding process of the introduced scenarios. The rainbow colors give the time as before. Corresponding points among the representations can be seen. For example, the light green serpentine flying spatter in the reference scenario in the projection along the scanner data is also visible in the temporal projection. However, the spatter path appears longer in the  $x$ -direction in the temporal projection. This is due to the fact that in the temporal projection, the events are shown relative to the laser position, whereas in the projection along the scanner data, the events are shown relative to the work piece. As the laser moves from left to right, the path of the backward flying light green spatter appears longer in the temporal projection than in the projection along the scanner data. If a spatter flies forward relative to the work piece, like the light blue spatter in the gap scenario in the projection along the scanner data, then its path appears shorter in the temporal projection. Moreover, it can be seen in the temporal projections of the oil scenario, that not all spatter paths are completely visible in the EC observation field. This means that some spatter path in the projection along the scanner data are cut off.

## B. Event representations of different configurations

To further show the potential of EC data for process monitoring of laser processes, four different configurations I–IV (see Table I) have been realized. Thereby, configuration I describes the previously used configuration according to Sec. III. Compared to configuration I, in configuration II parameters of the EC have been changed: The threshold for triggering an event has been reduced. This means that events are triggered for smaller intensity changes. This is indicated in the table as the sensitivity  $\sigma$ . In configuration

**TABLE I.** Configurations. As EC a Prophesee Metavision Evaluation Kit 4HD (EVK4) and a Prophesee Metavision evaluation kit 3HD (EVK3) were used. The sensitivity  $\sigma$  (Ref. 13) indicates how sensitive the EC parameters were set. As a laser, an IPG YLR-2000-WC (IPG 2000) and an IPG YLR-1000-WC-Y14 (IPG 1000) laser were used. As a scanner, a Raylase AS Fiber 30 (Raylase) and a Scanlab intelliWELD PR (Scanlab) were used. Moreover, the weld speed  $v$  in m/s and the weld power  $\eta$  in W are given.

Configuration	EC	$\sigma$	Laser	Scanner	$v$ (m/s)	$\eta$ (W)
I	EVK4	0	IPG 2000	Raylase	0.5	300
II	EVK4	1	IPG 2000	Raylase	0.5	300
III	EVK4	1	IPG 2000	Raylase	3.0	900
III	EVK3	0	IPG 1000	Scanlab	0.5	250

III, process parameters, namely, the weld speed  $v$  and the weld power  $\eta$ , have been increased. In configuration IV, experiments were carried out on a different system with a different EC and scanner as well as process parameters.

Figure 7 shows different event representations of the different scenarios for the four configurations. The first column shows the projections along the scanner data  $V$  of the reference scenario. A cut-out in which the welding path first is a curve and then continues in a straight line is shown. The other columns show the time projection  $R_1$  of the different categories. The first line shows configuration I. Due to the more sensitive EC parameters in configuration II compared to configuration I in each image, more events are visible. In addition to the visibility of spatters in configuration II, more events are generated along the welding, which give information about the keyhole. In contrast, more events are generated in uninteresting areas as well: Irrelevant events due to reflections on the clamping device in the curve but also on the straight line occur. In configuration III, more events occur, especially around the curve. In configuration IV, similar patterns like in configuration I are visible. In summary, configurations I and IV prove to be particularly suitable because they give the spatter behavior of the process and do not generate any additional irrelevant events.

## C. Event camera process monitoring with machine learning

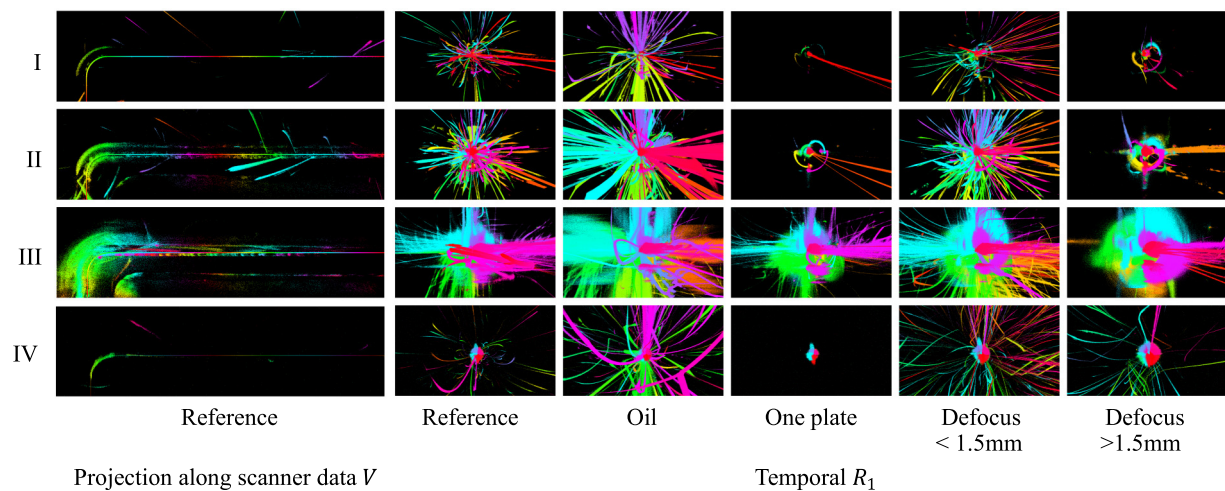
Figure 8 shows our proposed machine learning pipeline. After performing the experiments according to Sec. III, the received events are transformed either to event rates or to projections on the  $x, y$ -plane. A subset of transformed events is then used for the supervised training of a decision tree (DT) or a convolution neural network (CNN). DTs are well-suited due to their interpretable decision rules, while CNNs are explored due to their state-of-the-art performance in computer vision and pattern recognition tasks. Therefore, the label is given by the scenarios introduced in Fig. 2.

With  $T$  being the time interval in which a complete rectangular geometry according to Fig. 1 is welded, we select  $\Delta t = T$ . This means that for the welding of two metal plates, the event rate  $\epsilon$  contains one entry, namely, the number of events that occurred during  $T$ . The same applies for  $\epsilon_i$  and  $\epsilon_o$ . For  $\epsilon_i$  and  $\epsilon_o$ , visually a radius  $r = 60$  pixels was chosen. The training data instances of the DT, therefore, consist of three values, namely,  $\epsilon$ ,  $\epsilon_i$ , and  $\epsilon_o$ . Those training data instances are visualized by the dark boxes in Fig. 8. Instead of an event rate, for the CNN, one projection on the  $x, y$ -plane is created per time interval  $T$ . Therefore, CNNs were trained for the binary projection  $P$ , the density projection  $Q$ , and the time projection  $R_1$  as grayscale images.

Table I gives the number of metal plate pairs of the dataset. The dataset consists of 249 welded metal plate pairs. The number of metal plate pairs for training, validation, and testing is given by  $n_{\text{train}}$  (60%),  $n_{\text{val}}$  (20%), and  $n_{\text{test}}$  (20%). As given in Table II, the data are imbalanced between the scenarios. To compensate this, DT used Synthetic Minority Oversampling Technique (SMOTE). For the CNN, a data augmentation with balancing factors was incorporated.

The DTs were implemented with the library Scikit-learn.<sup>32</sup> The library uses an optimized version of the CART algorithm. To

10 March 2025 11:26:21



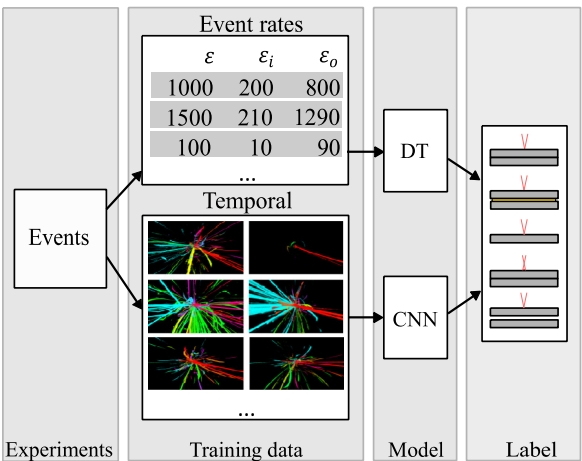
**FIG. 7.** Event representations of different configurations. I–IV give the configuration according to Table I. The first column gives the projections along the scanner data  $V$  and the second column the temporal projections  $R_1$ . Different scenarios according to Fig. 2 are given. Therefore, the second last column results from defocusing smaller than 1.5 mm and the last column from defocusing greater than 1.5 mm. The coloring between the projection along the scanner data and temporal are not correlated and are chosen due to illustration purposes: While in the projection along the scanner data  $V$ , the rainbow pattern is traversed twice, in the temporal projection  $R_1$  the rainbow color is traversed once.

measure the quality of a split, the Gini impurity was used as criterion. A grid search to find good hyperparameters, namely, the maximum depth of the tree, and the minimum number of chunks required to split an internal node was performed.

To teach the CNN invariances and robustness properties, a data augmentation including rotation, flipping, shifting, and shearing was used. The image size was resized from  $720 \times 1280$  pixels to

$100 \times 177$  pixels to accelerate the training process. The models were developed and trained using TensorFlow.<sup>53</sup> A wide range of CNNs were tried: Initially, the networks were kept shallow extended later to deeper networks. The filter size was changed between  $3 \times 3$  to  $7 \times 7$ . In addition to convolution layers, max pooling, global average pooling, batch normalization, or dropout have been used. Different activation functions such as ReLU, LeakyReLU, or tanh were tried. The final layer used a softmax activation. During the training of the CNN, the categorical cross-entropy was used as loss function. Moreover, an Adam optimizer with a learning rate  $l_r = 1 \cdot 10^{-4}$  was used.

Table III gives the weighted classification accuracies of the different event representations as input. Given the class imbalance in the data, standard accuracy can be misleading due to the underrepresentation of some classes. To overcome this issue, the weighted accuracy was used, which adjusts the overall accuracy by accounting for the number of instances in each class. The DT delivers an accuracy of 64.29% for the multiclass classification problem with five classes. In comparison, the CNN with binary projection  $P$  provides



**FIG. 8.** Machine learning pipeline. During welding, the received events are transformed either to event rates or projections on the  $x$ ,  $y$ -plane. The transformations are then used for the supervised training of a decision tree (DT) or a convolution neural network (CNN). Therefore, the label is given by the scenarios according in Fig. 2.

**TABLE II.** Number of metal plate pairs of the dataset.

Category	$n_{\text{train}}$	$n_{\text{val}}$	$n_{\text{test}}$	$n$
a) Reference	32	10	10	52
b) Oil	41	14	14	69
c) One plate	11	4	4	19
d) Defocus	22	7	7	36
e) Gap	43	15	15	73
Total	149	50	50	249

a higher accuracy of 72.96%. This makes sense, as the CNN processes a complete picture with information about the positions of the events, rather than relying on just three simple features as in the DT. Nevertheless, our results show that a density projection  $Q$  does not improve the classification accuracy and a time projection  $R_1$  significantly degrades the classification.

## VII. SUMMARY AND OUTLOOK

This paper, for the first time, investigates the potential of event cameras in laser welding. Event cameras (EC) differ from standard frame cameras: Rather than capturing images at a fixed rate, they asynchronously measure pixelwise brightness changes and output an event sequence that includes time, location, and sign of the brightness changes. To extract meaningful information from the asynchronous events that occur during a laser welding process, we present three promising event representations for in-process quality monitoring of laser welding processes: First, we propose the use of event rates that count the number of events occurring in a predefined time interval. Therefore, spattering processes can be quantified. Second, to also capture spatial information, we propose to transform the events into projections in the EC observation field. Three such projections, namely, binary, density, and temporal, have been analyzed. Third, to also incorporate scanner information, we propose to project the events along the scanner data. In our experiments, we observe the laser process of two metal plates coaxially and in process with an event camera and introduced different in production occurring scenarios. Those scenarios include contamination, leading to a spattering process or a gap between the two plates. Our laser welding experiments show clear differences for the scenarios in the event representations. To further encourage the potential of event cameras for process monitoring in laser welding, different configurations in which the event camera parameters, the process parameters, or the hardware parameters were changed are realized. After having analyzed the differences in the event representation when changing the laser welding process, we then present a machine learning pipeline for quality monitoring. Therefore, event rates were used to train a decision tree (DT) and projections of the events in the event camera observation field—so images—were used to train convolutional neural networks (CNNs). Therefore, the CNN outperformed the DT. Moreover, binary and density representations outperformed time representations across the CNN architectures.

**TABLE III.** Accuracies for training ( $A_{\text{train}}$ ) and testing ( $A_{\text{test}}$ ) for DT and CNN for the different event representations of configuration I. The CNN had the following architecture: for the binary projection  $P$ , the CNN consists of one convolution layer with a filter size of  $7 \times 7$  followed by max pooling and one dense layer. The time projection  $R_1$  results in the same CNN but with a filter size of  $3 \times 3$ . The density projection  $Q$  consists of three convolution layers with a filter size of  $3 \times 3$  with max pooling followed by two dense layers. The CNNs have ReLU as the activation function.

Method	Event representation	$A_{\text{train}}$ (%)	$A_{\text{test}}$ (%)
DT	$\mathcal{E}, \mathcal{E}_i, \mathcal{E}_o$ with $r = 60$	65.97	64.29
CNN	Binary projection $P$	79.47	72.96
CNN	Density projection $Q$	75.62	72.50
CNN	Time projection $R_1$	46.04	42.50

In the context of future research, to improve the classification accuracy, additional features such as spatter count, spatter angle, and spatter speed may be analyzed. In addition to additional features, further classification algorithms could be applied. Moreover, while our analysis classified the laser welding process after the weld of one experiment, other sampling strategies like analyzing the event data after a predefined amount of events has occurred could be applied.

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

### Author Contributions

**Patricia M. Dold:** Conceptualization (equal); Data curation (lead); Formal analysis (lead); Investigation (equal); Methodology (lead); Software (equal); Visualization (lead); Writing – original draft (lead); Writing – review & editing (equal). **Praveen Nadkarni:** Investigation (equal); Methodology (supporting); Software (equal); Visualization (supporting); Writing – original draft (supporting); Writing – review & editing (equal). **Meiko Boley:** Investigation (supporting); Project administration (equal); Supervision (equal); Writing – review & editing (equal). **Valentin Schorb:** Investigation (equal); Writing – review & editing (equal). **Lili Wu:** Supervision (supporting); Writing – review & editing (equal). **Fabian Steinberg:** Supervision (supporting); Writing – review & editing (equal). **Peter Burggräf:** Supervision (supporting); Writing – review & editing (equal). **Ralf Mikut:** Conceptualization (equal); Project administration (equal); Supervision (equal); Writing – review & editing (equal).

## REFERENCES

- W. Huang and R. Kovacevic, “Feasibility study of using acoustic signals for online monitoring of the depth of weld in the laser welding of high strength steels,” *Proc. Inst. Mech. Eng., Part B* **223**, 343–361 (2009).
- P. M. Dold, F. Bleier, M. Boley, and R. Mikut, “Two-stage quality monitoring of a laser welding process using machine learning: An approach for fast yet precise quality monitoring,” *at - Automatisierungstechnik* **71**, 878–890 (2023).
- I. Eriksson and A. F. Kaplan, “Evaluation of laser weld monitoring—A case study,” in *Proceedings of ICALEO, 28th International Congress on Applications of Lasers & Electro-Optics, Orlando, FL, 2–5 November 2009* (Laser Institute of America, Orlando, Florida), pp. 1419–1425.
- S. Shevchik, T. Le-Quang, B. Meylan, F. V. Farahani, M. P. Olbinado, A. Rack, G. Masinelli, C. Leinenbach, and K. Wasmer, “Supervised deep learning for real-time quality monitoring of laser welding with x-ray radiographic guidance,” *Sci. Rep.* **10**, 3389 (2020).
- A. Mayr, B. Lutz, M. Weigelt, T. Gläsel, D. Kießkalt, M. Masuch, A. Riedel, and J. Franke, “Evaluation of machine learning for quality monitoring of laser welding using the example of the contacting of hairpin windings,” in *2018 8th International Electric Drives Production Conference (EDPC), Schweinfurt, Germany, 4–5 December 2018* (IEEE, Piscataway, New Jersey, 2018), pp. 1–7.
- T. Ilar, I. Eriksson, J. Powell, and A. Kaplan, “Root humping in laser welding—An investigation based on high speed imaging,” *Phys. Proc.* **39**, 27–32 (2012).
- D. You, X. Gao, and S. Katayama, “Multisensor fusion system for monitoring high-power disk laser welding using support vector machine,” *IEEE Trans. Ind. Inf.* **10**, 1285–1295 (2014).

10 March 2025 11:26:21

- <sup>8</sup>P. M. Dold, M. Boley, F. Bleier, and R. Mikut, "An intelligent quality inspection system to detect laser welding defects," in *Proceedings of the Lasers in Manufacturing Conference, Munich, Germany, 26–29 June 2023* (Wissenschaftliche Gesellschaft Lasertechnik und Photonik e.V. (WLT), 2023).
- <sup>9</sup>C. Knaak, U. Thombansen, P. Abels, and M. Kröger, "Machine learning as a comparative tool to determine the relevance of signal features in laser welding," *Proc. CIRP* **74**, 623–627 (2018).
- <sup>10</sup>P. M. Dold, F. Bleier, M. Boley, and R. Mikut, "Multi-stage inspection of laser welding defects using machine learning," in *Proceedings 32. Workshop Computational Intelligence, Berlin, Germany* (KIT Scientific Publishing, Karlsruhe, Germany, 2022), Vol. 1, p. 31.
- <sup>11</sup>Y. Yang, L. Pan, J. Ma, R. Yang, Y. Zhu, Y. Yang, and L. Zhang, "A high performance deep learning algorithm for the automated optical inspection of laser welding," *Appl. Sci.* **10**, 933 (2020).
- <sup>12</sup>J. Hartung, A. Jahn, M. Stambke, O. Wehner, R. Thieringer, and M. Heizmann, "Camera-based spatter detection in laser welding with a deep learning approach," in *Forum Bildverarbeitung 2020, Karlsruhe, Germany, 26–27 November 2020* (KIT Scientific Publishing, Karlsruhe, Germany, 2020), p. 317.
- <sup>13</sup>G. Gallego, T. Delbruck, G. Orchard, C. Bartolozzi, B. Taba, A. Censi, S. Leutenegger, A. J. Davison, J. Conradt, K. Daniilidis, and D. Scaramuzza, "Event-based vision: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.* **44**, 154–180 (2022).
- <sup>14</sup>W.-j. Na, K. H. Sun, B. C. Jeon, J. Lee, and Y.-h. Shin, "Event-based micro vibration measurement using phase correlation template matching with event filter optimization," *Measurement* **215**, 112867 (2023).
- <sup>15</sup>G. Chen, L. Hong, J. Dong, P. Liu, J. Conradt, and A. Knoll, "EDDD: Event-based drowsiness driving detection through facial motion analysis with neuromorphic vision sensor," *IEEE Sens. J.* **20**, 6170–6181 (2020).
- <sup>16</sup>K. Bialik, M. Kowalczyk, K. Blachut, and T. Kryjak, "Fast-moving object counting with an event camera," *arXiv:2212.08384* (2022).
- <sup>17</sup>L. Pan, C. Scheerlinck, X. Yu, R. Hartley, M. Liu, and Y. Dai, "Bringing a blurry frame alive at high frame-rate with an event camera," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, 15–20 June 2019* (IEEE, Piscataway, New Jersey, 2019), pp. 6820–6829.
- <sup>18</sup>A. Kugele, T. Pfeil, M. Pfeiffer, and E. Chicca, "Efficient processing of spatio-temporal data streams with spiking neural networks," *Front. Neurosci.* (published online, 2020).
- <sup>19</sup>M. Osswald, S.-H. Ieng, R. Benosman, and G. Indiveri, "A spiking neural network model of 3D perception for event-based neuromorphic stereo vision systems," *Sci. Rep.* **7**, 40703 (2017).
- <sup>20</sup>H. Rebecq, G. Gallego, E. Mueggler, and D. Scaramuzza, "EMVS: Event-based multi-view stereo—3D reconstruction with an event camera in real time," *Int. J. Comput. Vision* **126**, 1394–1414 (2018).
- <sup>21</sup>M. Cook, L. Gugelmann, F. Jug, C. Krautz, and A. Steger, "Interacting maps for fast visual interpretation," in *The 2011 International Joint Conference on Neural Networks, San Jose, CA, 31 July–5 August 2011* (IEEE, Piscataway, New Jersey, 2011), pp. 770–776.
- <sup>22</sup>M. A. Mahowald and C. Mead, "The silicon retina," *Sci. Am.* **264**, 76–82 (1991).
- <sup>23</sup>H. Rebecq, R. Ranftl, V. Koltun, and D. Scaramuzza, "High speed and high dynamic range video with an event camera," *IEEE Trans. Pattern Anal. Mach. Intell.* **43**, 1964–1980 (2021).
- <sup>24</sup>A. Amir, B. Taba, D. Berg, T. Melano, J. McKinstry, C. Di Nolfo, T. Nayak, A. Andreopoulos, G. Garreau, M. Mendoza, J. Kusnitz, M. Debole, S. Esser, T. Delbruck, M. Flickner, and D. Modha, "A low power, fully event-based gesture recognition system," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, 21–26 July 2017* (IEEE, Piscataway, New Jersey, 2017), pp. 7243–7252.
- <sup>25</sup>X. Clady, S.-H. Ieng, and R. Benosman, "Asynchronous event-based corner detection and matching," *Neural Netw.* **66**, 91–106 (2015).
- <sup>26</sup>A. Z. Zhu, L. Yuan, K. Chaney, and K. Daniilidis, "EV-FlowNet: Selfsupervised optical flow estimation for event-based cameras," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, 18–22 June 2018* (IEEE, Piscataway, New Jersey, 2018), pp. 5578–5587.
- <sup>27</sup>J. Kogler, C. Sulzbachner, M. Humenberger, and F. Eibensteiner, "Address event based stereo vision with bio-inspired silicon retina imagers," *Adv. Theory Appl. Stereo Vision* (published online, 2019).
- <sup>28</sup>G. Gallego, J. E. A. Lund, E. Mueggler, H. Rebecq, T. Delbruck, and D. Scaramuzza, "Event-based, 6-DOF camera tracking from photometric depth maps," *IEEE Trans. Pattern Anal. Mach. Intell.* **40**, 2402–2412 (2018).
- <sup>29</sup>C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, "On-manifold preintegration for real-time visual-inertial odometry," *IEEE Trans. Robotics* **33**, 1–21 (2017).
- <sup>30</sup>M. Litzenberger, C. Posch, D. Bauer, A. N. Belbachir, P. Schon, B. Kohn, and H. Garn, "Embedded vision system for real-time object tracking using an asynchronous transient vision sensor," in *2006 IEEE 12th Digital Signal Processing Workshop & 4th IEEE Signal Processing Education Workshop, Teton National Park, Wyoming, 24–27 September 2006* (IEEE, Piscataway, New Jersey, 2006), pp. 173–178.
- <sup>31</sup>T. Delbruck and P. Lichtsteiner, "Fast sensory motor control based on event-based hybrid neuromorphic-procedural system," in *2007 IEEE International Symposium on Circuits and Systems, New Orleans, Louisiana, 27–30 May 2007* (IEEE, Piscataway, New Jersey, 2007), pp. 845–848.
- <sup>32</sup>F. Pedregosa, A. Varoquaux, G. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, "Scikit-learn: Machine learning in python," *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
- <sup>33</sup>M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: A system for large-scale machine learning," in *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16), Savannah, GA, 2–4 November 2016* (USENIX Association, 2016), pp. 265–283.

10 March 2025 11:26:21