

Robust Impurity Detection and Tracking for Tokamaks

C. Cowley,* P. Fuller, Y. Andrew, L. James, and L. Simons
Blackett Laboratory, Imperial College, London SW7 2AZ United Kingdom

M. Sertoli, S. Silburn, A. Widdowson, and JET contributors†
*Culham Centre for Fusion Energy, Culham Science Centre,
 Abingdon, Oxfordshire, OX14 3DB, United Kingdom*

I Bykov and D. Rudakov
General Atomics, San Diego, California 92186, USA

T. Morgan and Magnum-PSI team
*FOM Institute DIFFER, Dutch Institute For Fundamental energy Research,
 Association EURATOM-FOM, P.O. Box 6336, 5612 AJ Eindhoven, The Netherlands*

P. Bryant and B. Harris
*Department of Electrical Engineering and Electronics,
 University of Liverpool, Liverpool, L69 3GJ, United Kingdom*

(Dated: August 6, 2020)

A robust impurity detection and tracking code, able to generate large sets of dust tracks from tokamak camera footage, is presented. This machine learning based code is tested with cameras from the Joint European Torus, DIII-D, and Magnum-PSI, and is able to generate dust tracks with a 65 – 100% classification accuracy. Moreover, the number dust particles detected from a single camera shot can be up to the order 1000. Several areas of improvement for the code are highlighted, such as generating more significant training data sets, and accounting for selection biases. Although the code is tested with dust in single 2D camera views, it could easily be applied to multiple-camera stereoscopic reconstruction, or non-dust impurities.

I. INTRODUCTION

Magnetic confinement fusion in deuterium-tritium plasmas has the potential to be a clean, sustainable, and efficient alternative to current power generation methods. Across the globe numerous tokamaks have been built to achieve sustained nuclear fusion, with the Joint European Torus (JET) [2] and Doublet-III-D (DIII-D) [3] being two of the machines. To achieve ignition in a fusion plasma, high densities, temperatures, and confinement times are required [4]. The central challenge of fusion energy is in achieving these conditions, and attempts to do so have been through a multitude of physics and engineering optimisations.

In every operational tokamak exists impurities, which can either exist intrinsically or are extrinsically added for experimental purposes. One class of intrinsic impurity is dust, which is typically a nm-mm particulate created from the tokamak Plasma Facing Components (PFCs). These PFCs include the tokamak walls and divertor, which in current experimental devices are typically comprised of carbon, in the form of graphite or carbon-fibre composites, or metals such as tungsten, beryllium, and molybdenum [5, 6]. During tokamak operation, high en-

ergy events such as Edge Localised Modes (ELMs), disruptions, and vertical displacement events can load these PFCs with powers up to the order GWm^{-2} . These events can potentially liberate intrinsic dust, and even create more dust via processes such as spallation, vaporisation, and erosion of the PFC melt layer [7–9]. In addition to these plasma interactions, mechanical machining of PFCs can also contribute to the creation of dust.

Dust in tokamaks can pose both operational and safety concerns. In particular, dust particles can retain tritium fuel leading to an increase in the in-vessel tritium inventory over time. This is of concern both for safety and for the tritium fuel cycle, and as a consequence future machines have limited the tritium inventory to 700g [10]. In addition, when dust particles composed of high mass elements such as tungsten enter the plasma, this can cause sudden cooling of the plasma by increased radiation and a drop in the plasma fusion performance, or in extreme cases lead to plasma disruptions [11, 12].

In an attempt to understand and mitigate the risks posed by tokamak impurities, the physics of dust in plasmas become a highly researched area in recent years. At the forefront of this research are dust in plasma simulation codes such as MIGRATION of GRAINS in fusion devices (MIGRANe) [13] Dust in TOKamaks (DTOKS) [14], DUST-TRACKing (DUSTTRACK) [15] and DUST Transport (DUSTT) [16]. These codes solve differential equations governing the evolution of temperature, charge, and velocity of a dust grain in a plasma back-

* cyd.cowley15@ukaea.uk

† See the author list of [1]

ground [17, 18]. One key use for these codes is in determining which physical forces and phenomena play important roles in dust transport in tokamaks. In theory, one of these codes could be run with different forces active, to see what phenomena best replicate experiment. This, however, requires large statistically significant dust track datasets to compare against the models. Such datasets are currently sparse and manually selected. Here, the Robust Impurity Detector and Tracker (RIDAT) code is presented, that allows for automatic generation of such datasets from tokamak camera footage.

II. METHOD

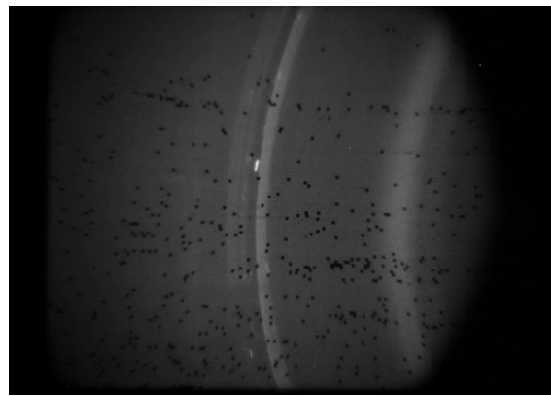
A. RIDAT code

The code for RIDAT was developed in *Python*, with aims of being able to detect statistically significant sets of dust tracks automatically. RIDAT was also built to be robust, such that it can be applied to various cameras and impurities. It can be separated into image processing and dust tracking modules. The image processing module receives a list of image frames from tokamak video footage as an argument, and outputs a list of dust grains detected in each frame - along with the dust grain properties. The dust tracking module uses these isolated detected grains as arguments, and outputs a list of connected dust tracks spanning multiple frames [19][20].

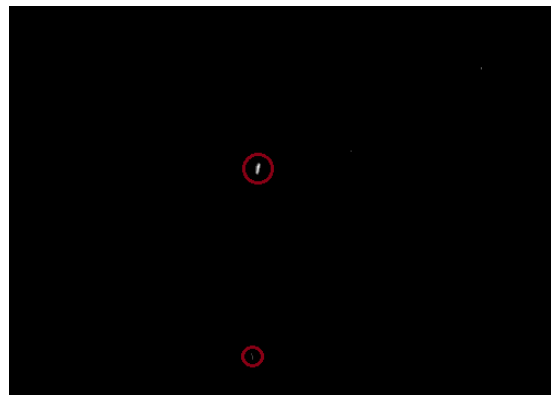
The image processing module detects and characterises dust grains, which can be seen as small bright ellipses or streaks in camera footage. To achieve this, the code works frame by frame, determining a background average for each frame. This is done by averaging the pixel intensities in temporally adjacent frames. This background average is then subtracted from the intensity profile of a frame, and all pixels over a user selected brightness threshold T_B are detected. These bright spots are collected into dust grains by grouping spatially adjacent bright pixels. Two successfully detected grains can be seen in Figure 1.

Finally, basic properties of the grains are determined, such as position, brightness, length, and width. Position and brightness are found by averaging the x-y pixel coordinates, and normalised intensities of a grain respectively. The length of a grain is found by taking the longest distance (px) between any two pixels on the grain. The width is then simply the grain area (px²) divided by length. It is important to note that - depending on camera frame rate - the length of a dust grain can either be indicative of it's physical dimensions, or how far it has travelled over the camera exposure period. Because of this, a user can manually select whether RIDAT is in Streaking or Non-Streaking mode. In the former, RIDAT treats the length as the distance traveled, and stores two positions of a grain: one at the start and one at the end of its motion.

The dust tracking module takes the unordered dust



(a)



(b)

FIG. 1: a) A still unedited image taken from the DiMES camera in DIII-D. b) The same still, but with RIDAT's background subtraction applied with a T_B of 35. Detected dust grains are indicated by red circles.

grain lists of each frame outputted by the image processing module as an argument. It returns an ordered list of fully connected dust tracks spanning multiple frames. In other words, the purpose of this module is to connect the temporally isolated dust grains found by the image processing module. The core of this module is an algorithm that works frame by frame analysing all possible tracks for a dust grain in frame N across frames $N+1$ and $N+2$. Then, the algorithm determines which of these tracks is the most likely real track using a Naive Bayesian machine learning classifier. Finally, it repeats the process on the next set of three frames ($N+1$, $N+2$, and $N+3$), joining the new data to any previously detected tracks.

A Naive Bayesian classifier is a supervised machine learning classifier that determines the probability of each classifier C_i fitting to data with features f , $P(C_i|f)$, using Bayes' theorem [21]. In the case of RIDAT, the classifier assumes a Gaussian probability density function, and classifies a dust track as either *True* (corresponding to a likely real track) or *False* (corresponding to an erroneous joining of unrelated dust grains). The features then consist of the following properties of a potential dust track:

- Mean change in position, $\mu_{\delta R}$ (px).
- Standard deviation of change in position, $\sigma_{\delta R}$ (px).
- Mean change in brightness, $\mu_{\delta B}$ (normalised).
- Mean change in width, $\mu_{\delta w}$ (px).
- Mean velocity angle, μ_{θ} ($^{\circ}$).
- Mean change in velocity angle, $\mu_{\delta\theta}$ ($^{\circ}$).

This type of classifier was selected due to its simplicity, and ability to determine not only if a track was correct (*True*), but what the likelihood of correctness was. Thus, for a given dust grain in frame N, if there are candidates for the corresponding grain in frame N+1 and N+2, this classifier can determine the most probable correct candidates with a Bayesian probability over a user defined threshold T_P . The listed features have been chosen as they are the physical properties intuitively thought to discriminate dust tracks from other image features. If a dust grain is traveling at a constant speed, for example, then the mean change in position should be relatively constant. It is important to note, however, that this example of a constant speed is not assumed by the user. In fact, even if all grains were travelling with a constant 3D speed, the projection onto a 2D camera would not be constant. The only assumption made is that there is some distribution of particle speed; the machine learning algorithm then determines the characteristics of this distribution.

Creation of the training data sets was achieved by a training function, that displays three consecutive video frames to the user. The user can select the correct path in frames 2 and 3 for a given grain in frame 1. The dust track features are saved, with a classification of *True*. Because a given grain only takes one *True* path, all other potential tracks stemming from the given frame 1 grain are classified *False*, which allowed the creation of large training data sets quickly. Although this method creates a significant disparity between classifications, it is a disparity mirrored by reality, as the number of possible paths for a grain greatly outweighs the one *True* path. An example of one *True* and one *False* dust track from a training data set from the linear Magnum-PSI machine is shown in Table 1.

$\mu_{\delta R}$	$\sigma_{\delta R}$	$\mu_{\delta B}$	$\mu_{\delta w}$	μ_{θ}	$\mu_{\delta\theta}$	C
9.5	2.5	0.4	0.4	0.0	0.0	<i>True</i>
26.6	20.5	0.7	0.0	-52.9	48.5	<i>False</i>

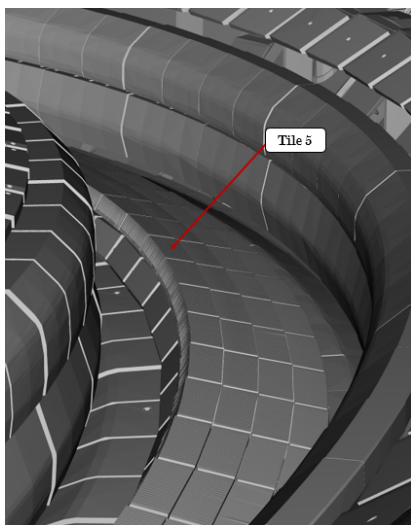
TABLE I: A table displaying the features of one correct and one incorrect dust track labeled by a user using RIDAT for the a shot in Magnum-PSI.

To test the efficacy of the code, RIDAT was applied to footage from three plasma machines: JET, DIII-D, and Magnum-PSI (M-PSI). The general process of application was first to determine parameters to be used in the

code, then to manually create a training data set for each machine, and finally to run the tracking code on one shot at a time. Due to the high variation in dust abundance, computational time for tracking varied across machines. The shortest computation time on a 4 core machine was less than 1 minute for a 10 frame shot in JET, with ~ 1 grain per frame. The most computationally expensive run was on a single 1000 frame shot in M-PSI; taking more than 20 hours for ~ 20 grains per frame.

B. Diagnostic Setup

Footage from the JET tokamak was captured with a Near Infrared (NIR) Hitachi KP-M1AP protection camera, with a frame rate of $50Hz$. This camera was part of a safety monitoring system, situated on the top half of JET's inner wall, with an angled view towards the divertor [22]. Footage collected was from a 2018 campaign, during which W dust was occasionally mobilised by the restart of the plasma. Footage from the NIR camera is shown in Figure 2.



(a)

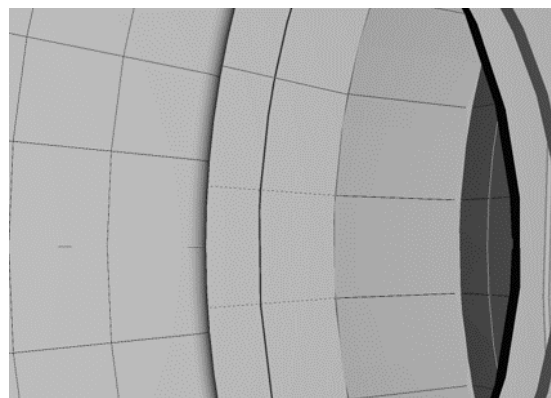


(b)

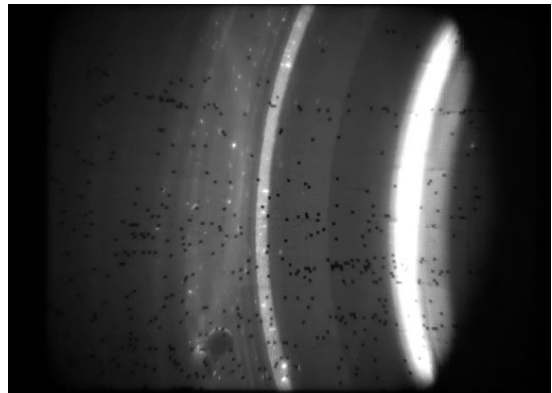
FIG. 2: a) A 3D reconstruction of JET's NIR camera view using the tokamak viewing software Calacam [23]. b) A still image taken from the same NIR camera.

In addition to JET, data was also collected from the tokamak DIII-D [3]. The camera used for this experiment was CIDTEK 3710D camera with a frame rate of $5kHz$, and a spectral width dependent upon the filter applied [24]. The camera was part of a Divertor Material Evaluation Station (DiMES), which has a vertical view of the divertor [25]. This top-down divertor view can be seen in Figure 3. Shots from DIII-D were captured during an experimental campaign studying turbulent transport and ELM control [26]. Dust grains observed in these videos were most likely comprised of intrinsic graphite, as this is the main material used for PFCs in DIII-D.

The final machine used for experimental data was Magnum-PSI (M-PSI), a linear plasma generator located in the Dutch Institute for Fundamental Energy Research. Camera footage collected was from a Phantom visible fast



(a)



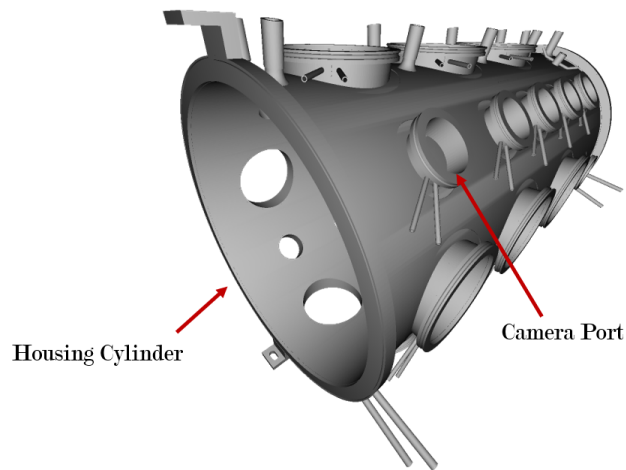
(b)

FIG. 3: a) A 3D Calcam reconstruction of the DiMES view in DIII-D [23]. The gray tiles indicate the divertor viewed from above. b) A still image taken from the DiMES visible camera.

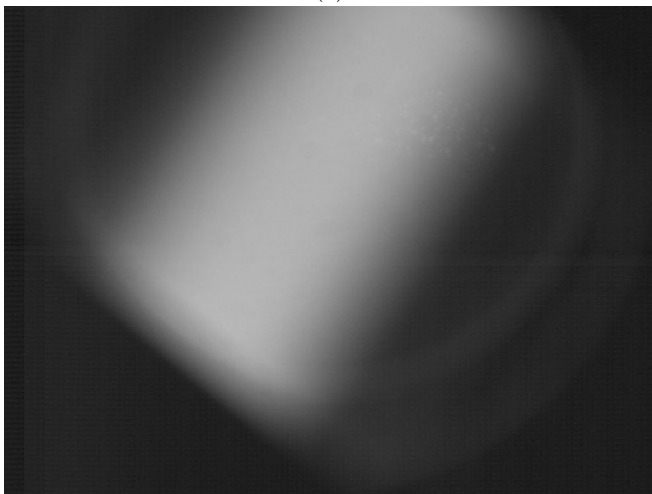
imaging camera located in one of the side camera ports of the cylindrical housing, shown in Figure 4. Footage was collected from over a dozen shots with a frame rate spanning from $400Hz - 1kHz$ [27]. The footage for this research was captured during an experiment in which dust was dropped vertically into the machine with varying columns of plasma flowing through it. The dust was comprised of W, made monodisperse by being horizontally shaken through a sieve with holes of radius $2.5\mu m$. The dust was initially dropped with no plasma present, then a cylindrical plasma column was generated with magnetic field values of 0.1, 0.2, 0.3, and $0.4T$.

III. RESULTS & ANALYSIS

RIDAT's machine learning algorithm was trained separately with JET, DIII-D, and M-PSI footage. The training datasets were randomly segmented from the image data, and the training set sizes were 365 for JET, 597×10^3 for DIII-D and 230×10^3 for M-PSI. The settings used for training and tracking were optimised in a number of preliminary runs. During these runs the change in



(a)



(b)

FIG. 4: a) A Calcam reconstruction of the plasma housing cylinder located in the Magnum-PSI machine [23]. b) A still image taken by the Phantom visible fast imaging camera in Magnum-PSI. The illuminated section indicates where a plasma column would be.

brightness variable was removed from the machine learning feature set. This was because only the brightest and easiest to identify grains were used in training. Thus, a bias was created for brighter dust grains in training, which extended to a bias in tracking when the brightness feature was active. The settings with which RIDAT was run for each machine are displayed in Table 2.

Machine	T_B	T_P	Streak	Size
JET	15	0.84	<i>true</i>	363
DIII-D	30	0.97	<i>false</i>	597×10^3
M-PSI	8	0.97	<i>true</i>	230×10^3

TABLE II: A table displaying the run settings for RIDAT applied to each plasma machine, where T_B is the threshold brightness, T_P is the threshold Bayesian probability, Streak determines whether grains are treated as single grains or traveling streaks, and size is the number of datapoints used for training the machine learning classifier.

A. JET

Each of the three trained models were applied back to their corresponding machines to detect and track dust. For JET, over 30 dust tracks were identified over 34 different shots. An example track overlaid on a still image from the divertor camera is shown in Figure 5.

From Figure 5 it is clear to see that the RIDAT predicted track follows the bright streak of the physical dust grain well. This demonstrates RIDAT's ability to accurately and automatically track a dust grain across multiple frames in video footage. However, the track shown in Figure 5 is also incomplete, as the code has failed to track the first and last frames of the grain's path. This may be due to the significant variation in track direction at the start (top) and tail (bottom left) end of the path. This deviation in track features, coupled with the small (363 track) training data set for JET could have been sufficient to incorrectly label the tail ends of the path as false. This issue may be resolved by using a larger, more varied training data set, or by more carefully optimising brightness and probability thresholds.

After the JET dust tracks were generated, the classification accuracy of the tracks was determined. This was done by overlaying detected tracks with camera footage, and manually determining whether the track positions follow their real counterparts (similar to the example in Figure 5). The classification accuracy of RIDAT applied to JET was calculated to be 65%, which is lower than the value predicted by the Bayesian probability cut-off of 84%. The reason this classification accuracy is so low is most likely due to the quality of the JET footage, as the 35% of incorrect identifications were mostly camera noise or plasma emission. Additionally, the rapid deviation in track features mentioned earlier can be partially attributed to the JET camera's long integration time. Thus, for footage similar to the NIR camera in JET, RIDAT may be used as a first step in track generation; followed by more rigorous manual vetting.

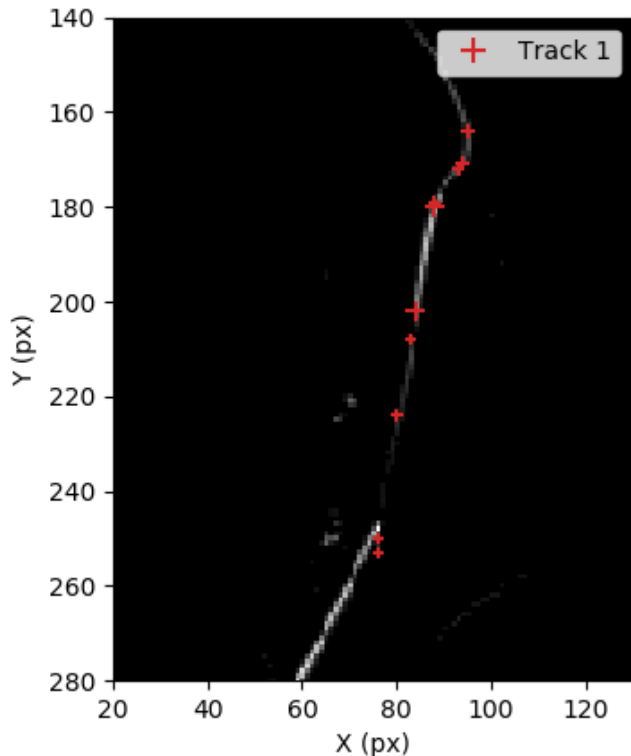


FIG. 5: The pixel positions of a RIDAT-generated dust track from JET. Positions were overlaid onto a combined still from JET’s NIR camera, found by summing the image data of every frame in this shot.

B. DIII-D

RIDAT’s application to DIII-D footage generated 925 dust tracks across two shots. Five example tracks are shown overlaid on an image from the DiMES camera in Figure 6.

The RIDAT generated dust tracks were again manually compared against camera data, and a classification accuracy of 99% was found. This high classification accuracy further demonstrates RIDAT’s ability to generate high quality track data from camera footage. Moreover, database of 925 tracks indicates RIDAT can successfully produce statistically significant datasets automatically. As such, it seems RIDAT is best suited to the type of data collected from DIII-D, in which there were large quantities of separated, bright, and consistent dust grains.

C. M-PSI

When applied to M-PSI, RIDAT produced over 1,500 dust tracks across 18 shots. Five example tracks are shown in Figure 7, overlaid on of a still image from the phantom camera.

Upon manual inspection, the classification accuracy of the RIDAT tracks was determined to be 74%. This clas-

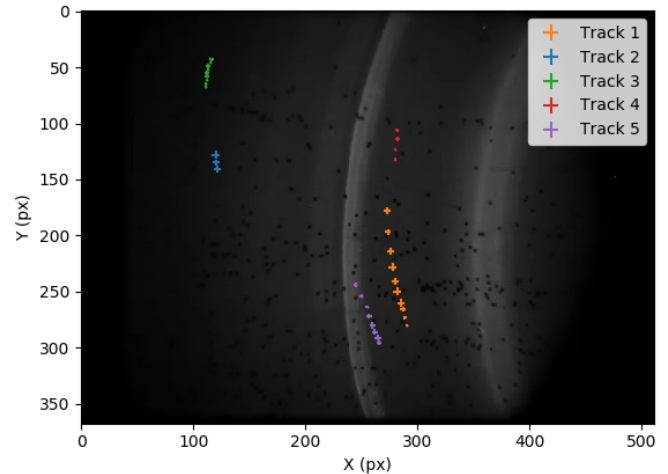


FIG. 6: The pixel positions of a random set of five RIDAT-generated tracks from shot 167345-11198 in DIII-D, overlaid on a still image from the same shot.

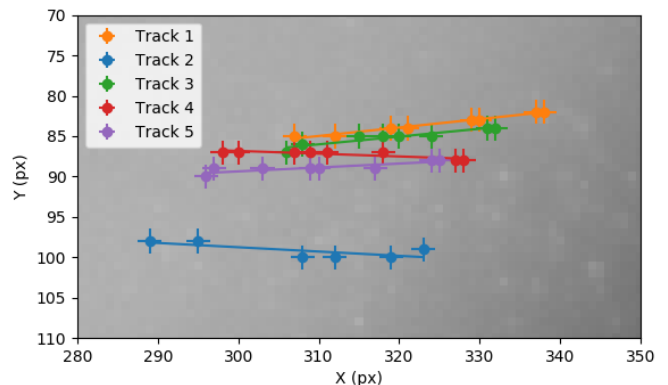


FIG. 7: The pixel positions of a random set of five RIDAT-generated tracks from the 0.4T Plasma data set of Magnum-PSI, overlaid on a still image from the same camera. To distinguish the trajectories of overlapping tracks, a linear fit has been applied to all tracks.

sification accuracy initially appears low. However, the majority of RIDAT’s inaccurate classifications for the M-PSI dust grains was because RIDAT could not distinguish between two very similar dust tracks. This was exacerbated by M-PSI’s particularly bright background seen in Figure 7. Thus, the performance of RIDAT could be improved by more carefully selecting image processing parameters; parameters such as the threshold brightness, and the number of frames over which a background is averaged.

IV. CONCLUSIONS

In conclusion, the RIDAT code was applied to camera footage from JET, DIII-D, and the linear plasma gen-

erator Magnum-PSI with varying levels of success. In JET, RIDAT detected over 30 dust tracks across 34 camera shots with a classification accuracy of 65%, and incorrectly discounted sections of some correct tracks. In DIII-D, 925 dust tracks were generated from two camera shots, with an adequate classification accuracy of 99%. For Magnum-PSI over 1,500 tracks were created with a classification accuracy of 74%. The issues of low classification accuracy and the discounting of correct track sections highlight the drawbacks of RIDAT. However, these may be improved upon by creating more robust training data, accounting for biases, and by more carefully selecting RIDAT parameters.

RIDAT could be used for a number of applications in the future. These include grain lifetime and splitting analysis, as well as 3D track reconstruction using dual stereoscopic camera systems such as those currently employed in the Mega Ampere Spherical Tokamak and TEXTOR [28, 29]. For effective use, however, the code should be improved, specifically by applying more optimised image processing techniques, such as those used in particle tracking for image microscopy [30]. In addition to significant algorithm improvements for the code, fine tuning of the input parameters should also be performed for each future scenario. For example, increasing the al-

gorithms rolling frame window from three to four could result in more accurate classifications.

V. ACKNOWLEDGMENTS

We thank Michael Coppins for his expert supervision and advice, that helped shape the direction for the RIDAT code.

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

This work was supported in part by the US Department of Energy under DE-FC02-04ER54698 and DE-FG02-07ER5491.

This work was supported in part by the Engineering and Physical Sciences Research Council, under grant EP/M001709/1. For the contribution to the M-PSI dust experiment we thank Prof. S. Ratynskaia (Space & Plasma Physics, School of Electrical Engineering Royal Institute of Technology (KTH) Stockholm, Sweden) and Marco De Angeli (Institute for Plasma Science and Technology (ISTP) Milan Italy).

-
- [1] J. Contributors, E. Joffrin, T. Donné, G. J. van Rooij, *et al.*, *Nuclear Fusion* **59**, 112021 (2019).
 - [2] F. Sartori, *IEEE Control Systems* **26** (2006), 10.1109/mcs.2006.1615273.
 - [3] J. Lohr, *DIII-D Research Operations annual report to the US Department of Energy* (1995).
 - [4] J. D. Lawson, *Proceedings of the Physical Society. Section B* **70**, 6–10 (1957).
 - [5] J. Roth, E. Tsitrone, T. Loarer, V. Philipps, S. Brezinsek, A. Loarte, G. F. Counsell, R. P. Doerner, K. Schmid, O. V. Ogorodnikova, and R. A. Causey, *Plasma Physics and Controlled Fusion* **50** (2008).
 - [6] T. Hirai, V. Barabash, F. Escourbiac, A. Durocher, L. Ferrand, V. Komarov, and M. Merola, *Fusion Engineering and Design* **125**, 250–255 (2017).
 - [7] A. Hassanein, G. Federici, I. Konkashbaev, A. Zhitlukhin, and V. Litunovsky, *Fusion engineering and design* **39**, 201 (1998).
 - [8] J. Winter, *Physics of Plasmas* **7** (2000).
 - [9] A. Hassanein, V. Belan, I. Konkashbaev, L. Nikandrov, V. Safronov, A. Zhitlukhin, and V. Litunovsky, *Journal of nuclear materials* **241**, 288 (1997).
 - [10] *Summary of the ITER Final Design Report*, ITER EDA Documentation Series No. 22 (INTERNATIONAL ATOMIC ENERGY AGENCY, Vienna, 2001).
 - [11] J. C. Flanagan, M. Sertoli, M. Bacharis, G. F. Matthews, P. C. de Vries, A. Widdowson, I. H. Coffey, G. Arnoux, B. Sieglin, S. Brezinsek, J. W. Coenen, S. Marsen, T. Craciunescu, A. Murari, D. Harting, A. Cackett, and E. Hodille, *Plasma Physics and Controlled Fusion* **57** (2015).
 - [12] M. Sertoli, J. Flanagan, M. Bacharis, O. Kardaun, A. Jarvinen, G. Matthews, S. Brezinsek, D. Harting, A. Cackett, E. Hodille, and et al., *Journal of Nuclear Materials* **467**, 394 (2015).
 - [13] L. Vignitchouk, P. Toliás, and S. Ratynskaia, *Plasma Physics and Controlled Fusion* **56**, 095005 (2014).
 - [14] M. Bacharis, M. Coppins, and J. E. Allen, *Phys. Rev. E* **82** (2010), 10.1103/PhysRevE.82.026403.
 - [15] G. Gervasini, E. Lazzaro, and A. Uccello, *Journal of Fusion Energy* **36**, 25 (2017).
 - [16] A. Y. Pigarov, S. I. Krasheninnikov, T. K. Soboleva, and T. D. Rognlien, *Physics of Plasmas* **12** (2005), 10.1063/1.2145157.
 - [17] J. Martin, M. Bacharis, M. Coppins, G. Counsell, and J. Allen, *EPL (Europhysics Letters)* **83**, 65001 (2008).
 - [18] J. D. Martin, Imperial College London (2006).
 - [19] <https://github.com/cydcowley/RIDAT>.
 - [20] cydcowley and pjf15, “cydcowley/ridat: Version 1.0.0 release,” (2019).
 - [21] A. Y. Ng and M. I. Jordan, in *Advances in neural information processing systems* (2002) pp. 841–848.
 - [22] A. Huber, S. Brezinsek, P. Mertens, B. Schweer, G. Sergienko, A. Terra, G. Arnoux, N. Balshaw, M. Clever, T. Edlingdon, *et al.*, *Review of scientific instruments* **83**, 10D511 (2012).
 - [23] S. Silburn, Zenodo (2018), doi: 10.1088/0370-1301/70/1/303.
 - [24] M. Groth, M. Fenstermacher, C. Lasnier, R. Hernandez, J. Moller, and R. Sturz, *Review of scientific instruments* **74**, 2064 (2003).
 - [25] D. Rudakov, A. Litnovsky, W. West, J. Yu, J. Boedo,

- B. Bray, S. Brezinsek, N. Brooks, M. Fenstermacher, M. Groth, *et al.*, *Nuclear fusion* **49**, 085022 (2009).
- [26] E. Davis, J. Rost, M. Porkolab, A. Marinoni, and M. Van Zeeland, *Review of Scientific Instruments* **89**, 10B106 (2018).
- [27] L. Simons, J. Holgate, C. Stavrou, B. Dudson, M. Thomas, P. Bryant, B. Harris, S. Ratynskaia, and T. Morgan, in *45th EPS Conference on Plasma Physics, EPS 2018, 2 July 2018 through 6 July 2018* (European Physical Society (EPS), 2018) pp. 1720–1723.
- [28] G. De Temmerman, M. Bacharis, J. Dowling, and S. Lisgo, *Nuclear Fusion* **50**, 105012 (2010).
- [29] A. Shalpegin, L. Vignitchouk, I. Erofeev, F. Brochard, A. Litnovsky, S. Bozhenkov, I. Bykov, N. den Harder, and G. Sergienko, *Plasma Physics and Controlled Fusion* **57**, 125017 (2015).
- [30] N. Chenouard, I. Smal, F. De Chaumont, M. Maška, I. F. Sbalzarini, Y. Gong, J. Cardinale, C. Carthel, S. Coraluppi, M. Winter, *et al.*, *Nature methods* **11**, 281 (2014).