Image segmentation and classification for Fission Track Analysis for Nuclear Forensics using U-Net model

Noam Elgad[1,2*], Rami Babayew[1,2], Mark Last[3], Aryeh Weiss[4], Erez Gilad[2], Galit Katarivas Levy[5], **Itzhak Halevy[1,2**]**

[1]*Engineering and Physics Departments, Nuclear Research Center Negev, Israel*

[2]*Unit of Nuclear Engineering, Faculty of Engineering Sciences, Ben-Gurion University of the Negev, Israel*

[3]*Department of Software and Data Engineering, Faculty of Engineering Sciences, Ben-Gurion University of the Negev, Israel*

[4]*Faculty of Engineering, Bar Ilan University, Israel*

[5]*Department of Biomedical Engineering, Faculty of Engineering Sciences, Ben-Gurion University of the Negev, Israel*

E-mails: [*]*n.elgad@gmail.com*; [**]*halevy.itzhak.dr@gmail.com*

Abstract

This study introduces a novel methodology for the detection and classification of star shapes in microscopic images, employing state-of-the-art deep learning techniques for segmentation and classification [1]. The U-Net model, a fully convolutional network, was utilized to carry out the segmentation of various star-like patterns in both single-class and multi-class scenarios.

Keywords

Nuclear Forensics; Fission Track Analysis; Holmeland Security; Safeguards Investigations; U-Net; Computer Vision

Introduction

Nuclear Forensics (NF) is a research discipline that focuses on the study of radioactive materials. For more than 50 years, international safeguards confirms that nuclear materials disclosed by a state to the IAEA are used for peaceful purposes [2]. One of safeguards main strengthening strategies is based on environmental sampling and analysis to discover radioactive fingerprints of undeclared activity. Environment sampling as a routine for safeguards inspections approved at 1995 [3]. Environmental samples analysis may increase the information about nuclear activitiy in terms of form, size and surface structure of uranium microparticles of industrial dust in nuclear plants using scanning electron microscopy [4].

Particle isotopic analysis in NF has grown quickly in the past two decades due to advances in particle isotopic composition determination [5-8]. Experiments in this field support the reaction against illicit use of materials, guiding nations to improve their nuclear defense capabilities using scientific methods, evidence, and additional sources, including law enforcement and intelligence agencies [9].

NF uses analytical techniques, including isotopic, chemical, and physical properties of nuclear material, as well as traditional forensic evidence like DNA samples, hair, fingerprints, tool marks, and gunpowder residue, to identify fissile radioactive material and establish connections to individuals, locations, and incidents. Fission Track Analysis (FTA) detects fissile compounds. Radioactive samples from nuclear incident sites are collected for research. After collecting samples on dedicated sampling paper, they are sandwiched between two polycarbonate Lexan resin detector sheets and bombarded with thermal neutron flux to create cracks. Traces form star-like formations (fission roses) that depict nuclear fission products' trajectories. Those remnants can reveal interesting details about the measured substance's origin. The requirement for advanced statistical modelling methods to extract additional information from radiological NF investigations [4] may make the effort of this work a step forward.

Machine Learning (ML) is a branch of AI that enables systems to learn from experience without explicit programming [10-12]. The software learn patterns in a sample set to decide how to handle new information. Recently, machine learning (ML) and its derivatives, particularly deep learning (DL), have evolved considerably due to the

exponential development of information capture, storage, and processing capacities. [13]. This approach is compatible with computer vision (CV) application [14,15], and integrates across various academic and engineering fields [16-19]. In recent years, star identification has moved from manual procedures on a database of thousands of images to automation, easing researchers' labor. This study aims to improve field automation and propose an innovative AI-based solution for FTA purposes using segmentation and classification of nuclear fission track shapes in microscopic images as part of the efforts being placed to further strength and improve safeguards realm.

Theory

This section provides a concise overview of the FTA process and presents a broad perspective on the selected technologies used for identification purposes in this study.

A. FTA

One advantage of the FTA method is its ability to identify fissile materials from background elements. One particle in 1,000,000 is fissile and significant. Laboratory research is done with necessary safety precautions like gloves. A particle-free sampling paper collects radioactive samples from items in the prescribed area. These particles are sealed in Lexan foil. To simplify collection, a mini-bulk model is built. This approach determines the average enrichment and concentration of the material, guiding the micro-bulk transfer quantity. A mean quantity of particles can indicate the existence of artificial chemicals, such as uranium isotopes, providing an initial forensic investigation trail. Mini-bulk sampling involves sampling a reference item and a suspected contaminated item. To dissolve the particles, the pieces are cut into 1 cm^2 squares and soaked in Heptane liquid (C_7H_{14}) in a test tube. The sample preparation method follows these steps: Three main steps comprise the experiment: (1) completing a mini-bulk preparation for ICP-MS testing, (2) preparing a 300–500 picogram sample, and (3) attaching the sample to a 31x30 mm Lexan surface to construct a Lexan detector. After sample preparation, a reactor irradiates the detector, which is a capture sandwich of several detectors. This irradiation exposes the detector to $5x10^{13}$ n/cm^2 s thermal neutrons for 20 minutes. The detector undergoes forced fissions due to irradiation. After irradiation, the detector sandwich cools for three days before being returned to the lab.

After cooling, markers are securely mounted to the top and lower detectors at three locations. This creates a coordinate system for transduction and navigation between the detectors. Then, fission traces from the examined material are identified in a lab. The particle's spatial placement determines its identification, and several chemical methods will test it. There are several star configurations that represent the orbits of nuclear fission products [20]. About 700 images are placed on a scanning stage-equipped optical microscope. The microscope has an automated vertical focusing mechanism. From 7,000 images, the computer chooses the sharpest image for each group. There is a list of landmarks. The shapes are cut to verify the precision of preparation, casting, screening, and cutting. If a star is visible in the configuration, the procedure was correct and the substance is real. The FTA method can identify fissile radioactive material via spontaneous or induced fissions. Thus, detector grooves show fission trail signatures and indicate the fission site. Since fission is isentropic, the fission product will be a star with a circular border. FTA offers exact location data on excised sections that can be sent for ICP-MS investigation. If uranium is present and the enrichment level is abnormal, forensic proof can be obtained. Generally accepted guidelines for enrichment levels are as follows: natural at 0.72%, commercial power reactors at 4-5%, research reactors at 20%, and nuclear weapons at 90% and higher. Figure 1 illustrates a schematic representation of the procedure.

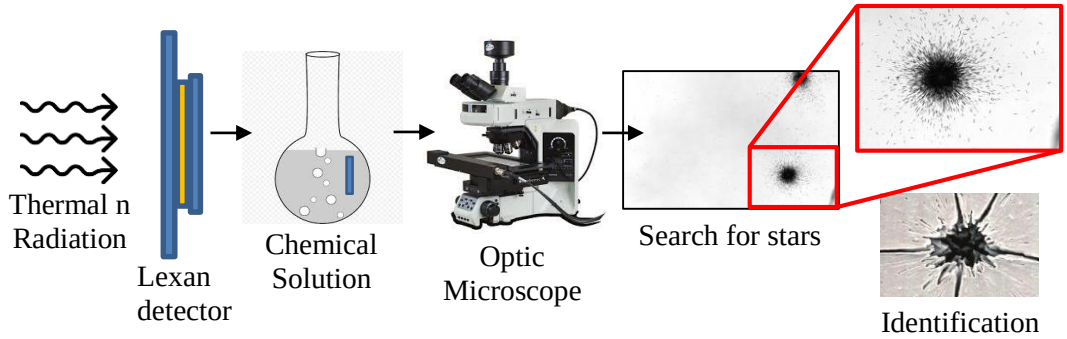


Fig. 1 Schematic flow of process of creating fission track stars

B. Identification processes

Until 2017, fission star detection was done manually, like "finding a needle in a haystack". Recent studies have developed automated algorithms for star identification [21,22]. Image processing software with a suitable user interface was used for identification. This technology hides manual-related difficulties. When people know the source of an example, they tend to work harder for great people. In a

study [23], a 10% increase in star identification (FP) was observed when the examiner understood the relevance of the example. Comparing testers with similar training shows a 15% difference. AI-based systems may have advantages over previous technology. Autonomous systems' objectivity is a major asset. Two staff members and two days are needed for manual testing. Modern image processing methods can discern stars from fractions of a micron to roughly 100 microns that are 20 to 300 micrometers in size. Depending on the model's interest, repeated testing by the same examiner yielded 90% accuracy. Using different testers reduced accuracy 85%. The false positive rate rose by 5% [23]. Figure 2 shows an example artifact and a variety of stars observed in our lab. The tested material is natural sand containing uranium [24]. The stars are located and the quantity of leaves indicates fissile material. This suggests that a large natural uranium particle can form a star like a small highly enriched particle. The shape of a star only shows the amount of fissile material present, not its enrichment.

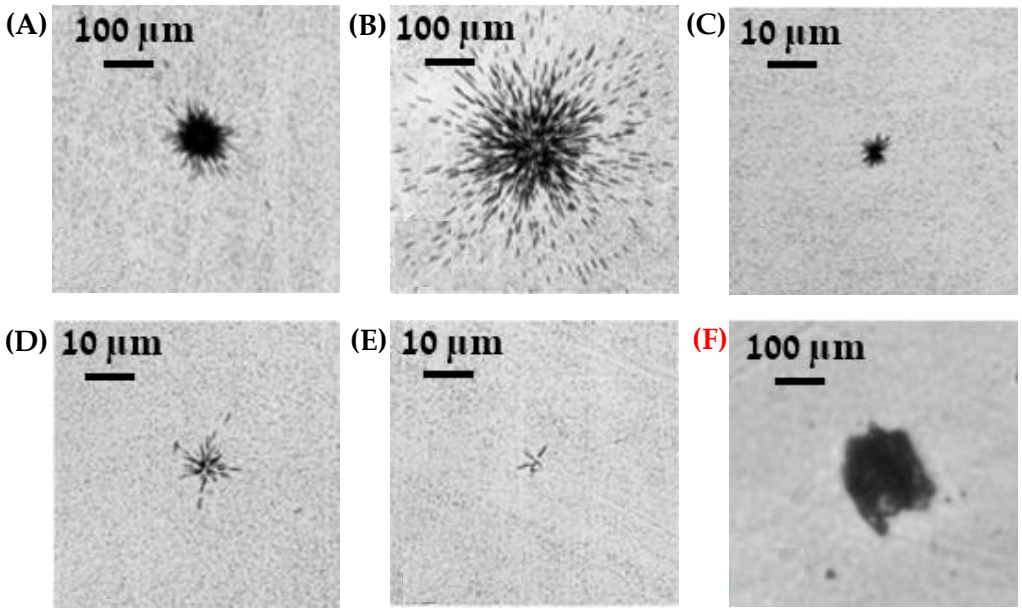


Fig. 2 Types of stars: (A) black center with a long halo (B) huge without a black center (C) small black center (D) not poor (E) poor (F) artifact

This study used U-Net neural networks, based on the Fully Convolutional Network (FCN) [25], to segment and classify fission stars in microscope images. Neural networks, also known as artificial neural networks, are computational mathematical models inspired by biological neural networks used in machine learning. This network has several interconnected information, input, and output components. A neural network has many primitive cells that do simple computations. The network's properties include neuron interconnections, synaptic weights, and activation function. It has three

layers: input, hidden, and output. Image recognition tasks typically exhibit a substantial quantity of hidden layers. The neuron aggregates the inputs by calculating a weighted total, and if the resultant value is over a specified threshold, the input is propagated to the output. A common learning algorithm is supervised learning, which propagates input across the network with random weights. This method compares received output to desired output and calculates error for each output neuron. After retrograde percolation, mistakes are updated and weights are modified. This strategy requires a lot of computer resources and processing time, but post-study network utilization is faster. The wide range of shapes and sizes they can take in an image in addition to the possibility of contaminants being mistaken for stars make textural star shapes difficult to recognize. Noise and background blurriness are common problems with microscope images. Therefore, typical image processing models are excessively sensitive to these areas. Stars have very varied shapes, and occasionally they overlap within the same regions, making standard segmentation and classification procedures much more complicated. Neural network-based semantic segmentation algorithms for image processing have advanced recently. Semantic segmentation labels each image pixel (e.g. person, road, sky, ocean, car).

Semantic segmentation is utilized in autonomous cars, industrial inspections, satellite image-based terrain classifications, and medical imaging analysis [26-29]. Currently, FCN algorithms are considered the 'State of The Art' in practical neural network performance [25]. These algorithms have yielded positive results in many segmentation challenges [30,31]. FCN architecture has only locally connected layers such as convolution, pooling, and un-sampling. Avoiding dense layers reduces parameters, making network training faster. U-Net networks [32] were developed by adapting the FCN architecture for biomedical image segmentation. This adaptation allows training with fewer images and improves segmentation accuracy. For segmentation and classification purposes, an algorithm based on U-Net network could be a viable option, since U-Net network captures the data's different shapes and textures, enabling accurate predictions and labeling of the stars. U-Net topology has a skip link between the Encoder and Decoder, two layers with ReLU activation functions and normalization, and a Maxpooling layer. Convolution creates image feature maps. Kernels, or quadratic two-dimensional filters, are applied to input data at fixed spatial locations during convolution. Convolution between the Kernel and the input produces the output.

The selection of the U-Net architecture was motivated by its shown efficacy and resilience in handling diverse spatial variations in object size, texture, and shape [33]. A model can be assessed using accuracy (precision) and recall (sensitivity). Accuracy measures the proportion of relevant examples among the recovered instances, while recall measures the proportion of relevant instances that were successfully retrieved, where:

$$Recall = \frac{TP}{TP+FN} \; ; \quad Precision = \frac{TP}{TP+FP} \tag{1}$$

Contemporary star identification methods have 80% precision. However, manual execution causes inconsistency. Thus, professional involvement is necessary and time-consuming. Star identification is difficult due to their varied forms and sizes. This intricacy makes it difficult to identify stars from other non-star phenomena like scratches, dirt, shadows, etc., preventing exact spatial coordinate determination. The study seeks a solution to these issues.

Our current focus in forensic segmentation and identification is demonstrating FTA segmentation capacity. An ImageJ Fiji distribution-based image processing method is used in an ongoing research collaboration between the International Atomic Energy Agency and the Israel Atomic Energy Commission. This algorithm finds three types of prominent stars. Additionally, the advanced approach requires redetermining the detection threshold for each system build [22,35]. To date, there has been a lack of utilization of neural networks. The primary challenge lies in identifying stars with few leaves. We use Semantic Segmentation to improve the existing approach by reclassifying sparse stars as a discrete category rather than noise. In the beginning of our inquiry, we found a rich star utilizing only 30 observations. Even with sparse stars, adding data improved performance. The ultimate goal is to recognize roughly 7 types of stars, including floating, asymmetric, and complex stars, using the skills gained in this work. The idea is to automate star identification for about ten categories. No methodology currently supports it.

Due to its compliance with a small training dataset, the U-Net technique was chosen for machine learning. Even with few categories and annotated images, this method allows accurate and reliable classification. Its goal was to progress the field by multi-

classifying several sorts of stars using a broad dataset for training. The dataset had over 100 labeled observations per categorization.

Experimental

The dataset acquisition process has two stages: Fission stars are first collected using a nuclear fission detector, Lexan. Second, an optical microscope generates 700 images with a 10% overlap. Each image is then divided into 250x250 micrometers.

A. Methodology

A binary division method was used to divide an image into foreground and background components to retrieve stars. High-intensity areas in the binary image were classified as stars. Measurements were based on star detection thresholds. Detection thresholds were set considering these characteristics: Star size, A criterion for a star vs. a scratch or dirt, ratio of pixels in the dark area to the bounding box, Circularity and Eccentricity.

B. Data Preparation

Star categorization was done to standardize communication. The categorization parameters were set by these criteria, which helped create a complete star catalog. The study had four classification criteria: size (large/small), richness (poor/not poor), center (black center/no black center), and halo length. Asymmetric, complex, and hovering stars have also been identified. The star categorization symbol is w, x, y, and z. The letters denote binary values of 0 or 1 for each condition. The magnitude, richness, core black region, and halo extent are represented by w, x, y, and z. Tables 1-3 presents Binary notation for star classification criteria, Typical images and color setting for labeling star types and Unique star types. All three tables are shown in Appendix A.

C. Machine Learning

Machine learning algorithms use mathematical models to learn and predict from input data. The model input data is split into different datasets. Model development often uses three datasets: training, validation, and test. The model is calibrated using the training set to optimize parameters using known cases. For each input vector in the training set, the model generates an output and compares it to the goal value. Model parameters are modified based on comparison results and learning technique. A validation set dataset uses the fitted model to estimate observation responses. The validation set impartially

evaluates a model's performance on the training data set, especially when modifying neural network hyperparameters like hidden units, layers, and layer width. Using the test set, a final model is impartially assessed against the training data.

The array data will not be used during model training to ensure an unbiased outcome image. Iteratively dividing a dataset into training and validation sets improves result stability and maximizes training data use. Cross-validation allows multiple model testing and training iterations. This method optimizes data use across arrays. In this study 5-fold Cross Validation technique was employed. This method partitioned the data into five subsets, with 60% for training, 20% for validation, and 20% for testing. A full data replacement process was designed to assure the independence of the validation and test sets. This was repeated five times, changing the training data composition.

Data must be meticulously prepared at the pixel level to support CV application training. The raw image has a light background and dark objects. The fission stars need identification. The image could feature shadows, stones, dirt, scrapes, and other factors. To train, each pixel must be classified correctly. Binarization inverts the shades to make the items white and the backdrop black. Blackening irrelevant information for identification additionally has to be done. Figure 3 shows a source image on the left and the image preparation results before training. All dataset images undergo this treatment. Source image resolution dropped from 20MB to 4MB during training data preparation. Due to information storage and training run session limits, the size was reduced. After adjusting the data, the training operation was completed and the results were satisfactory.

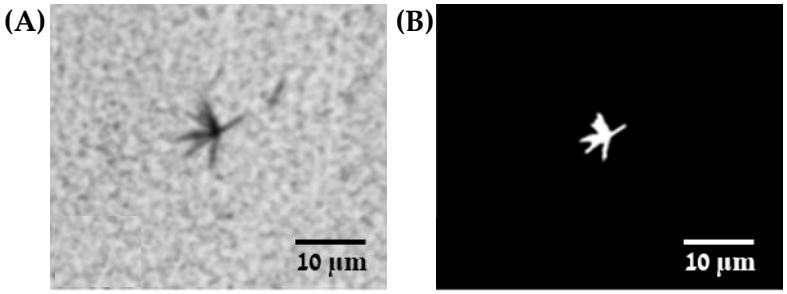


Fig. 3 (A) Original image containing a poor rose with dirty background. (B) The image after manual editing in preparation for training used as the tagged image

D. Architecture

The tasks of data preparation, running training runs, completing segmentation, and other operations were executed using the MATLAB software. As previously stated, this study's model uses U-Net neural network architecture. The name "U-Net" describes the network's architecture. The network has a connected path and a broad path, forming a U shape. Convolutional networks use iterative convolutions, ReLU activations, and pooling operations in the connection pathway. According to (2), the Rectified Linear Unit (ReLU) is a neural network activation function that mathematically defines its input as positive.

$$f(x) = x^{+} = \max(0, x) \tag{2}$$

Contraction reduces spatial information and increases feature information. A series of convolutions with high-resolution features from the shrinking path merges feature and spatial information in the expansive path. This study's U-Net architecture is shown in Figure 4. Table 4 details our model's network layers after this depiction. The model has three jump points and uses convolution.

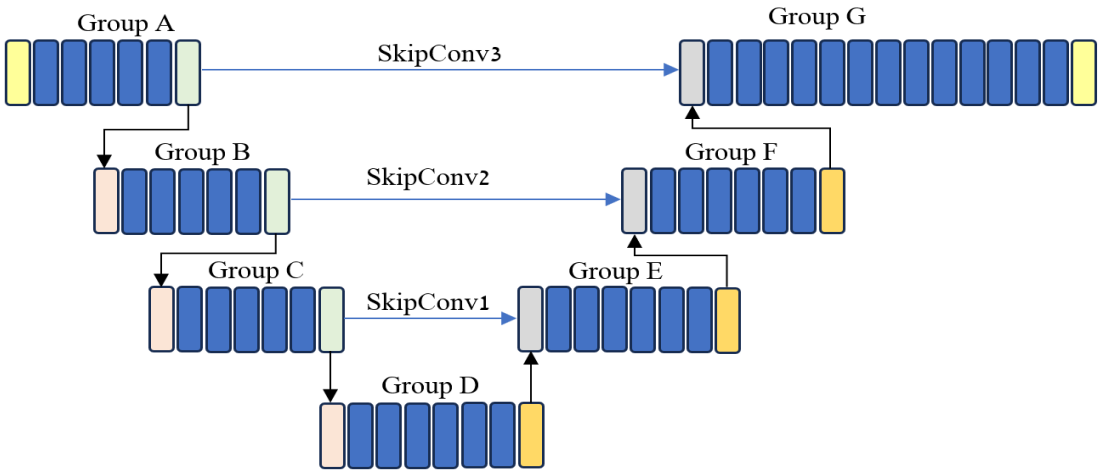


Fig. 4 Schematic diagram of the network layers in a U structure

Table 4 Network layers arranged by groups

Layer Group	A	B	C	D	E	F	G
Layers	Input image	Maxpool1	Maxpool2	Maxpool3	Add1	Add2	Add3
	Conv1	Conv3	Conv5	Conv7	Conv9	Conv11	Conv13
	Relu1	BN3	BN5	BN7	BN9	BN11	BN13
	Conv2	Relu3	Relu5	Relu7	Relu9	Relu11	Relu13
	BN2	Conv4	Conv6	Conv8	Conv10	Conv12	Conv14
	Relu2	BN4	BN6	BN8	BN10	BN12	BN14
		Relu4	Relu6	Relu8	Relu10	Deconv3	Relu14

				Deconv1	Deconv2		Conv15
							BN15
							Relu15
							Softmax
							Class Layer

One of the limitations associated with the utilization of deep learning techniques is the requirement for a substantial volume of data. For deep learning networks to achieve optimal performance, a substantial volume of data is required, typically in the order of millions of records. Furthermore, it is observed that the quality of training outcomes improves as the quantity of available information increases. As previously said, the U-Net model has an inherent advantage in its ability to effectively utilize small amounts of data. However, it is important to note that a substantial amount of data is still required. The process of generating data, particularly when it involves tagging information, presents a formidable undertaking that necessitates significant allocation of resources. The generation of artificial information, often known as Augmentation, has the potential to substantially enlarge the available data.

E. Image Labeling

As part of this work, we created a new classification system. A custom MATLAB application was used for image labeling. Sequential steps were taken: We start with source images for analysis. We then created labeling categories. Within these categories, we establish a reference label. This label can be used on stars of various categories and non-star objects like dirt or scratches. Notably, the label highlights star features. Finally, we compile and export all created data to a file for analysis.[1]

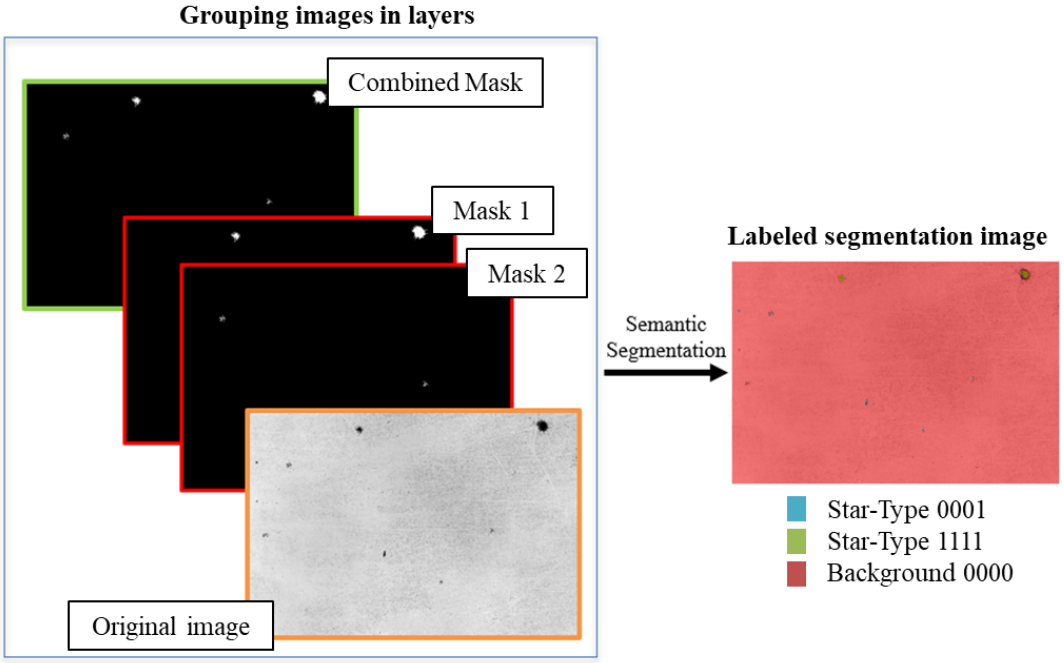
Experimental – Segmentation Processes

Once all data is ready, training starts. This study's system trained for single- and multi-class classification. Effective training produces a model that segment images and detect stars. The semantic segmentation procedure and evaluation are covered in this section.

A. Semantic Segmentation

[1] The thorough method development is not covered in this paper and is subject to separate publication.

Semantic segmentation labels each visual pixel as animal, human, road, sky, car, etc. The aim is to recognize certain features in an image, not categorize it. This approach distinguishes visual entities. In our work, segmentation requires labeling each pixel for different types of stars, non-star objects, and the background. A MATLAB function named semanticseg can be used. For segmentation, the function takes the trained neural network and an image. It then semantically segments the input image and assigns classification scores to each classified label. An array of scores is extracted and grouped by pixel or voxel in the supplied image. After this phase, the layers are blended using



Label Overlay to create a cohesive image with several labels. The Image Scaled Color technique generates a categorization results map. Figure 5 shows how semantic segmentation creates a multi-layer tagged image from a source image. This generation is done through training with an integrated mask image (Ground Truth). Each mask in the Ground Truth image represents a star type.

Fig. 5 An example of creating a tagged image using semantic segmentation

B. Segmentation Evaluation

We want to create a model that maximizes ROC curve area. We will explain the ROC curve, its ideas, and how to generate it in this chapter. We will also examine how to test our model against this curve.

The curve structure: The ROC curve shows binary classifier performance at different decision levels. A curve is created by plotting the true positive rate (TPR) and false positive rate (FPR) across acceptance levels. The sensitivity, or true positive rate, is a machine learning coverage measure. The false positive rate is 1 minus the specificity, or leakage or significance threshold. The receiver operating characteristic (ROC) curve shows how significance level affects sensitivity. Binary classification divides diagnose into positive and negative groups. True positive (TP), false positive (FP), true negative (TN), and false negative (FN) are the categorization outcomes. Table 5 uses the Confusion Matrix to display values. Our matrix is formed after training using segmentation (during the activation of the trained model on the test set) and can be used to calculate ROC values.

Table 5 Confusion matrix (in Parenthesis: our case)

		Actual Existence	
Classification		Exist (ST)	Not Exist (Dirt/ Stone etc.)
	Positive (ST)	True Positive	False Positive
	Negative (Dirt/ Stone etc.)	False Negative	True Negative

* In Parenthesis: confusion matrix related to our case

Visually, as depicted in Figure 6, the diagonal line serves as a demarcation between positive things on the left and negative items on the right. The classifier is represented by an egg-shaped boundary, encompassing positive objects inside its confines and excluding negative items outside of it.

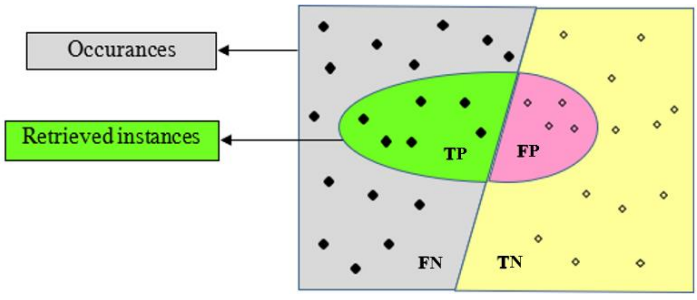


Fig. 6 True Positive (TP), False Positive (FP), True Negative (TN), False Negative (FN)

It should be noted that when considering misses of fission track analysis, it is crucial not to miss or miss a reduced number of actual stars, but marking non-stars as stars is permissible. meaning reducing false positives is much more important than reducing false negatives. This is due the fact that the meaning of FP is a radioactive material that might have been missed comparing to FN which means a non-radioactive item who may be marked falsely as radioactive.

C. Automation

Due to the large number of images involved, Automation is more efficient for segmentation and segmentation assessment. Therefore, within this study we developed a system that works using the following procedures: Users must select a source and destination folder for the image cluster. The applet systematically segments, combines layers, and generates a consistent image from all files. Next, create a map showing the classification results. Then, a folder with the same name as the current file is created to contain the consolidated image and the results map. Using this automation may enlarge the number of stars that can be analyzed all at once and get them all to be picked for analysis and squeeze the time it costs to perform this task manually. [2]

Results and discussion

This section provides an overview of the training and segmentation processes involved in both single class and multi class classification.

A. Single class training – poor & rich stars

Figure 7 shows two training outcomes from the trend for a low-quality star detection model. The primary training consisted of 20 epochs lasting 94 minutes, 50e 135 minutes, 100e 193 minutes, and 150e 492 minutes. Accuracy increases with repetitions and run duration. The figure shows two runs: 20 epochs and 1,580 iterations (top image) and 150 epochs and 11,850 iterations (bottom image). Both sets use 120 iterations for validation. The Accuracy chart at the top of each image shows the Mini-Batch Accuracy

[2] The in-depth automation development is not detailed in this paper and is planned to be published in a separate publication.

graph as a blue line and the Validation Accuracy as black dots. The Mini-Batch Loss (brown) and Validation Loss (black dots) are shown in the lower part of the figure.

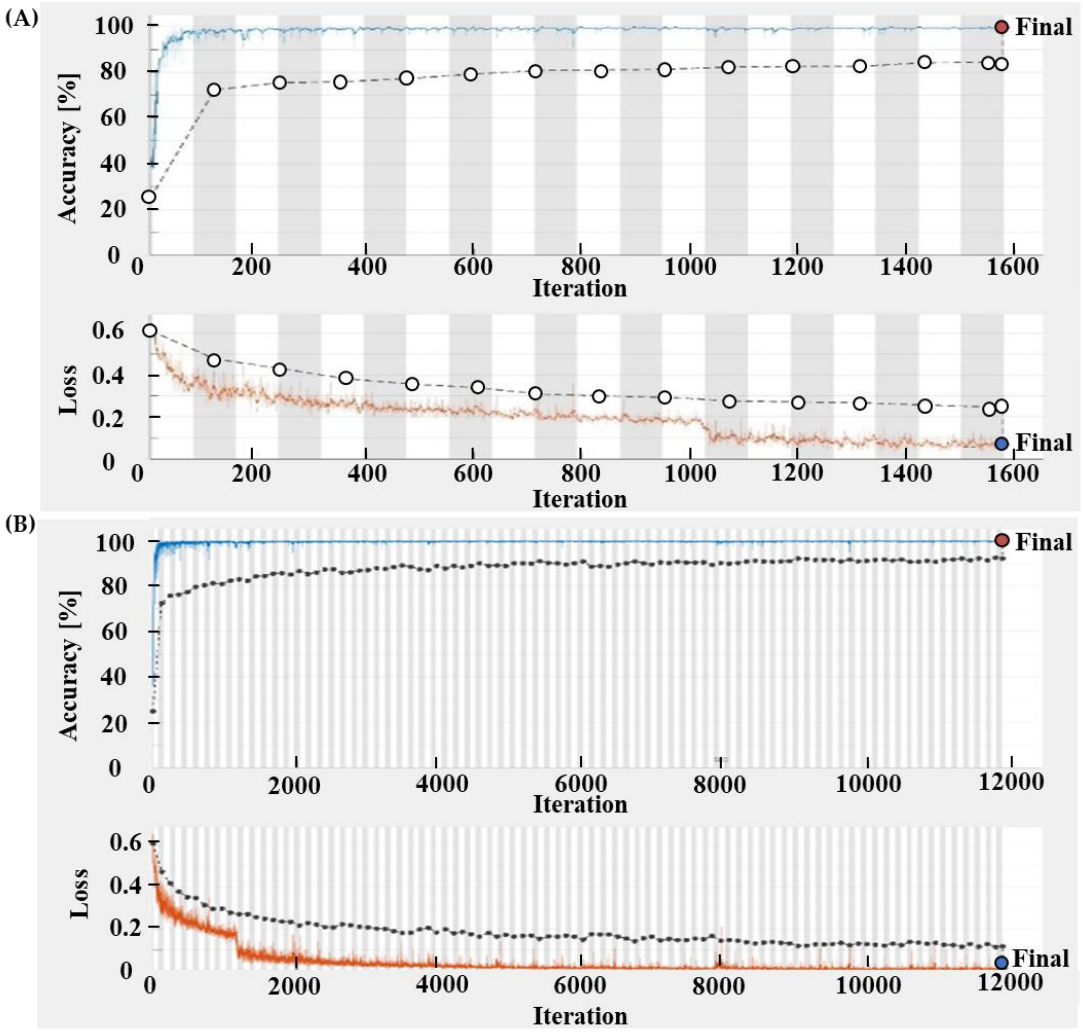


Fig. 7 Training results of single poor star model in 20 (A) and 150 (B) epochs

Table 6 displays the outcomes acquired from the four training sessions.

Table 6 Training results repetitions: 20, 50, 100, 150 epochs

epochs	Time Elapsed (hh:mm:ss)	Mini-Batch Accuracy	Validation Accuracy	Mini-Batch Loss	Validation Loss
20	01:34:02	NA	84.55%	NA	0.2470
50	02:15:04	99.79%	90.12%	0.0174	0.1627
100	03:13:33	99.83%	90.61%	0.0175	0.1238
150	07:52:39	99.76%	**92.04%**	0.0048	**0.1168**

In comparison to prior investigations [23], it has been observed that by carrying out training for a duration of 20 epochs, it is feasible to get similar outcomes of 84.55% accuracy (in comparison to about 85%) as reported by several human evaluators. The application of 50 epochs of training enhances the accuracy level, enabling comparable results to a single tester's test (90%). Subsequently, employing 150 epochs of training further enhances the accuracy level, resulting in an accuracy rate of 92.04%. This achievement surpasses the highest level attained by other methods thus far by approximately 2% (90%). Figure 8 displays a collection of segmented stars.

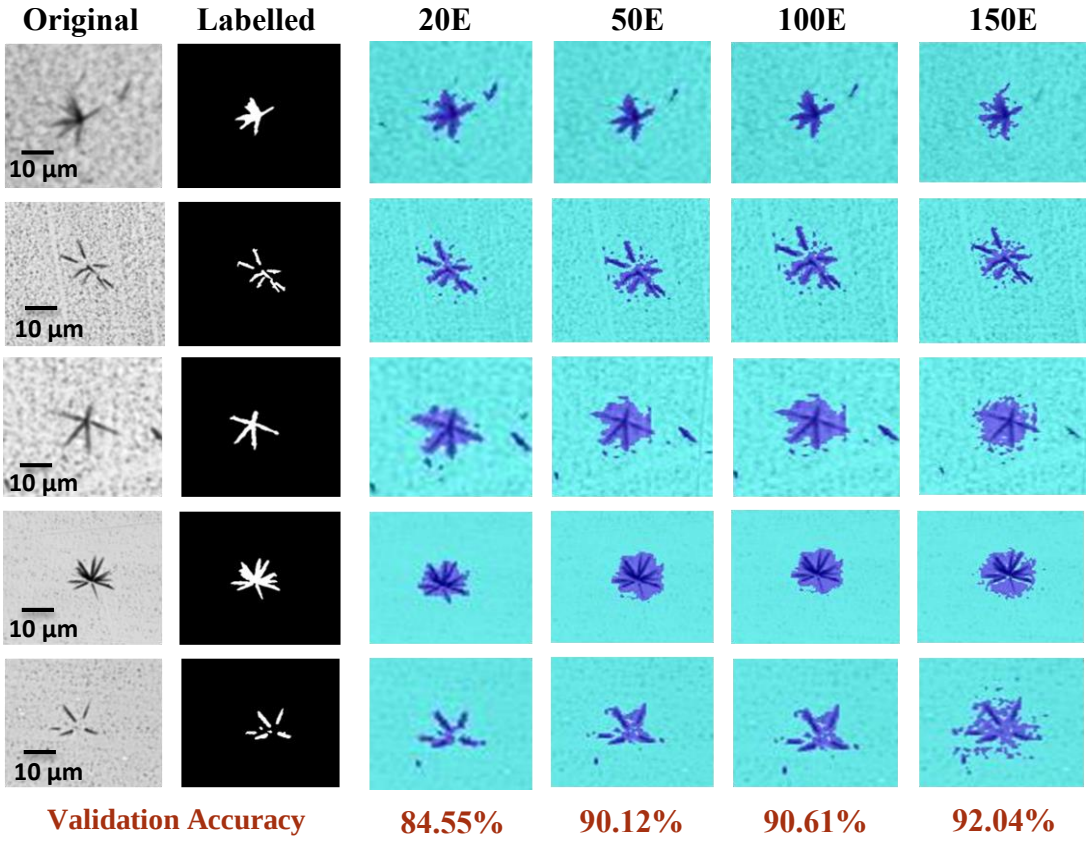


Fig. 8 Results of stars segmentation after training with 20, 50, 100 and 150 epochs

Some examples of a non-star objects source images are shown in Figure 9. The blue parts on the turquoise background symbolize segmentation with different epochs (20, 150). Five images were segmented after two classifiers were trained. The first three images show stones and dirt, the fourth an unclassified star object, and the fifth a scratch. The objective is for the blue marking (FP) to be minimal in all of these objects. As can be shown in the image, as the number of epochs increase, blue marking intensity decreases.

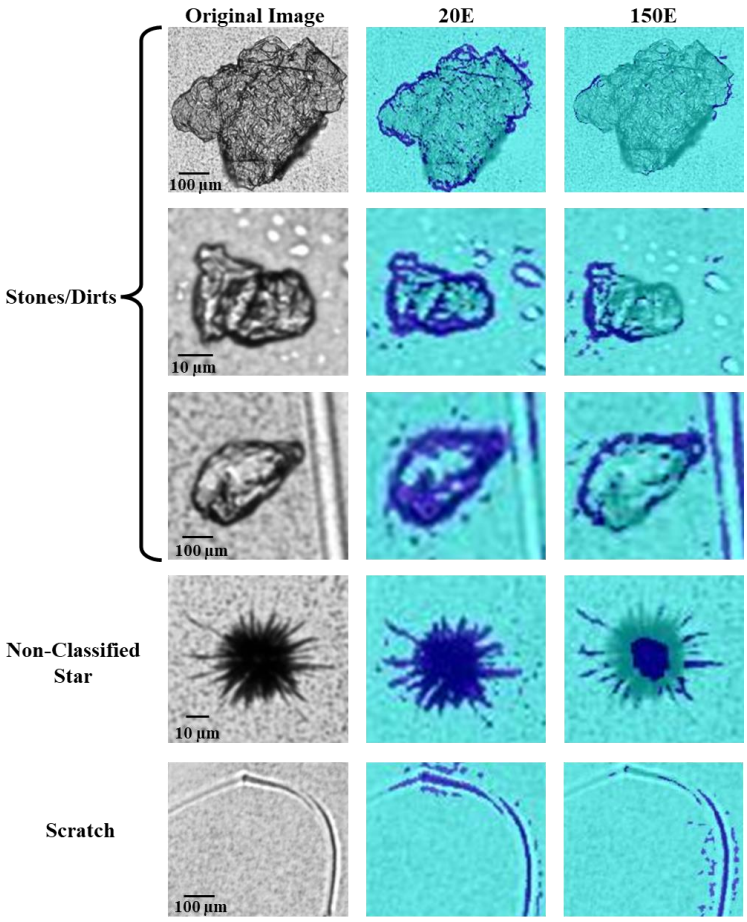


Fig. 9 Non-star objects segmentation results after training with 20 and 150 epochs

Table 7 presents a collection of focused segmentation images, illustrating a comparative analysis of outcomes achieved for objects classified as stars and those classified as non-stars. When considering stars, a higher degree of blue marking indicates a more favorable outcome. Conversely, for other objects, the objective is to minimize the blue marking and achieve a greater resemblance to the background hue.

Table 7 Comparison of segmentation between stars and non-star objects

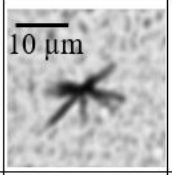
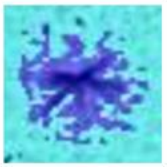
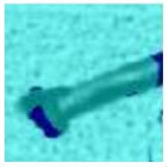
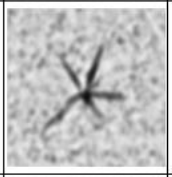
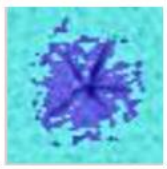
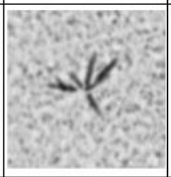
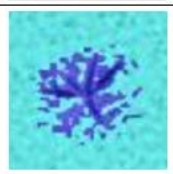
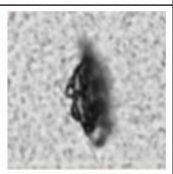
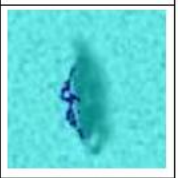
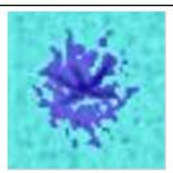
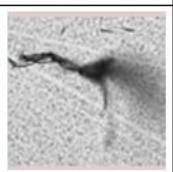
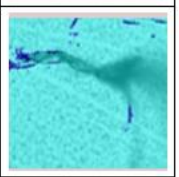
Sample ID	Stars (The bluer the merrier)			Non-Stars (The bluer the gloomier)	
	Original	Ground Truth	Segmented	Original	Segmented
tile_x004_y008					
tile_x005_y009					
tile_x004_y010					
tile_x004_y021					

Figure 10 depicts an image that has undergone segmentation, with the resulting segmentation image displayed on the left side, and a corresponding classification map on the right side. The visual attention is directed towards a star, which is distinctly highlighted in green, as well as an item that is not classified as a star, which is noted in red. The aforementioned table presents a shared view of these two objects.

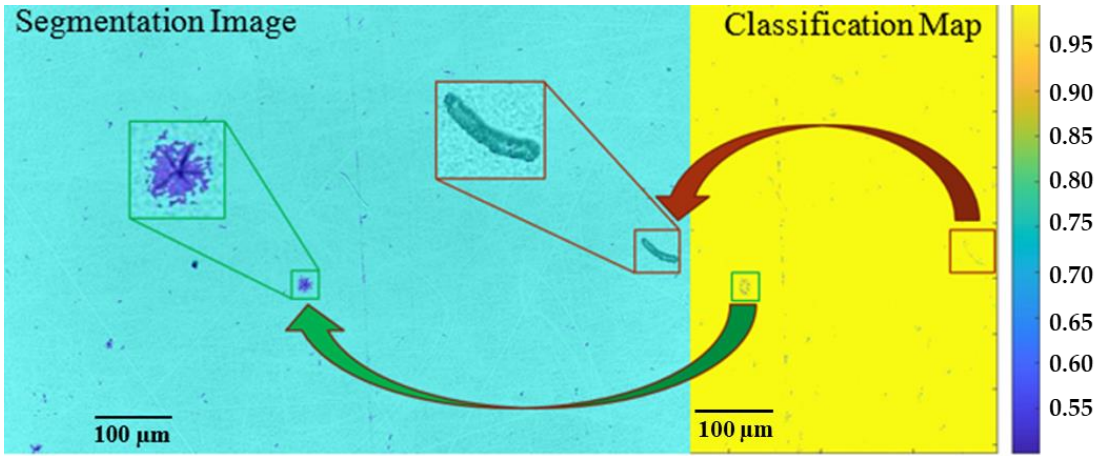


Fig. 10 Comparison between star and non-star object - a segmentation image and a classification map

This work produced two ROC curves using segmented images from a single-classifier model of a black-centered rich star, as shown in Figure 11. These curves also have AUC values calculated. The ROC curve shows how a binary classifier performs based on its decision threshold. At the middle, the graph shows the TPR-FPR relationship across acceptance criteria.

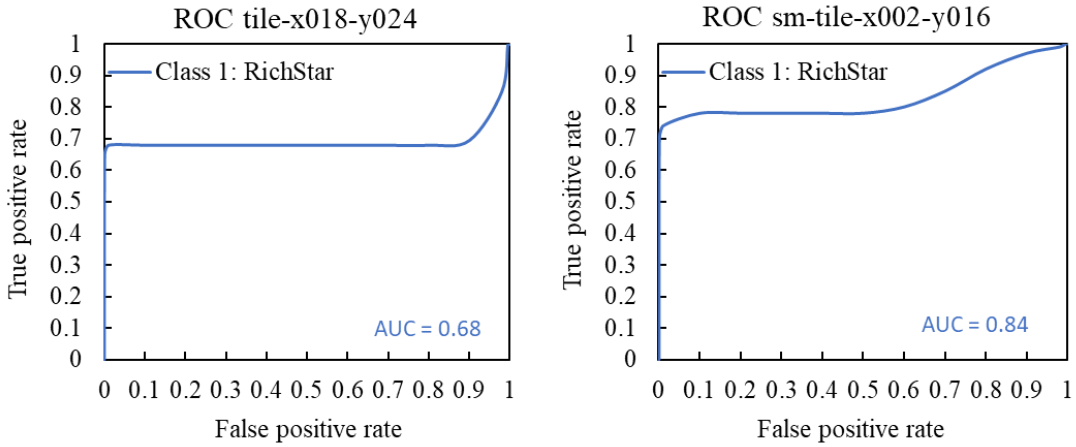


Fig. 11 ROC curves produced after completing model training for a single classifier

The AUC calculation shows the graph's area when a normal result is 0.5 to 1. Calculating AUC=0.5 yields a random prediction. Thus, the model does not improve star categorization. A model with a higher AUC value identifies and classifies better. Any outcome above 0.5 is better than random selection for maximizing value. The plots show a reasonable AUC=0.68 and a superior one at AUC=0.84 in our model.

Table 8 in the appendix summarizes one example of results of 4 images and shows pixel intensity, TP, FN, FP, recall and precision of each image. There should be noted that the model has a fair TP and FP and is biased toward zero FN in the specified images as these images went through an adaptive threshold algorithm which was developed for background cleaning.[3]

B. Multi class training – poor & rich stars

After training and segmenting rich and poor single star images, we train a multi-class model for both. one of the training outcomes where a model identified low (type 0001) and rich (type 1111) stars. The training has 20 epochs, 4880 iterations, and 250 validation iterations. Training results: Mini-Batch Accuracy: 96.43%, Validation

[3] This development is not detailed in this paper as it issued to be published separately.

Accuracy: 86.30%, Losses: 0.1461, 1.4974. Training outcomes are good for Validation Accuracy. Loss performance is good until epoch=12. However, findings oscillate after this. Several factors may explain this. One possibility is that pre-processing techniques, such as zero-meaning and normalization, were applied to the training set but not to the validation set, or vice versa. Another explanation could be that the observed discrepancy is due to overfitting, where the model has learned to perform well on the training data but fails to generalize effectively to the validation data. Due to the unchanged code execution mechanism, the first choice is less relevant. The second method is more sensible, thus models with more data were created. The dataset was enhanced via augmentation, and nodes were randomly excluded during training using dropout. Additionally, noise reduction and Cross Validation were used to improve data quality. This algorithm outperformed past techniques with an accuracy rate of 86.30%, compared to two persons' 85%. Compared to a manual tester's 90% recognition rate, the model's accuracy is still lower.

Figure 12 shows segmentation with classifiers: ST_0001, ST_1111, and background. The left half of the illustration shows images of ST_0001 stars with blue markings on yellow backgrounds. The right side shows ST_1111 type stars with green markings on a yellow background. The file names are centrally placed, indicating that the left and right images came from a single file.

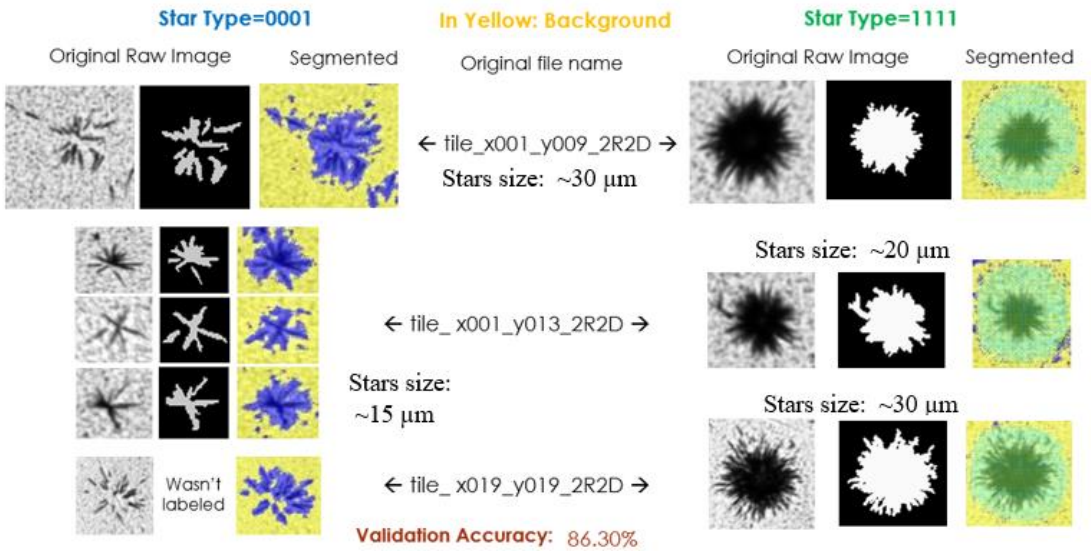


Fig. 12 Results of segmentation and identification of two types of stars after training

Two ROC charts show the performance of several classifiers created throughout the study, including AUC values, in Figure 13. Each chart shows multiple classifier graphs. Source and tagged images were used to generate segmentation images and charts. The chart below shows the results of two star categories in a trained model image. These categories include superior results with AUC values of 0.99 and 0.96. An alternate model is used to explore type 0001 and type 1111 stars. Type 1111 stars have an adequate AUC value of 0.66, as indicated in the chart. It shows less satisfactory results for type 0001 stars, with an AUC of 0.24. The value of 1111 is appropriate given its attributes. Its magnitude is less than 0.7. In contrast, it exceeds 0.5, the chance of a random guess. Trial 0001's results are below 0.5, indicating a technical problem.

According to examination, the threshold vector from the MATLAB perfcurv function is cut off at 0.33 instead of decreasing to 0, which appears to be a software fault. It is necessary to carry out new training with segmentation and ROC calculations after training and study the results to observe the root cause of a false results.

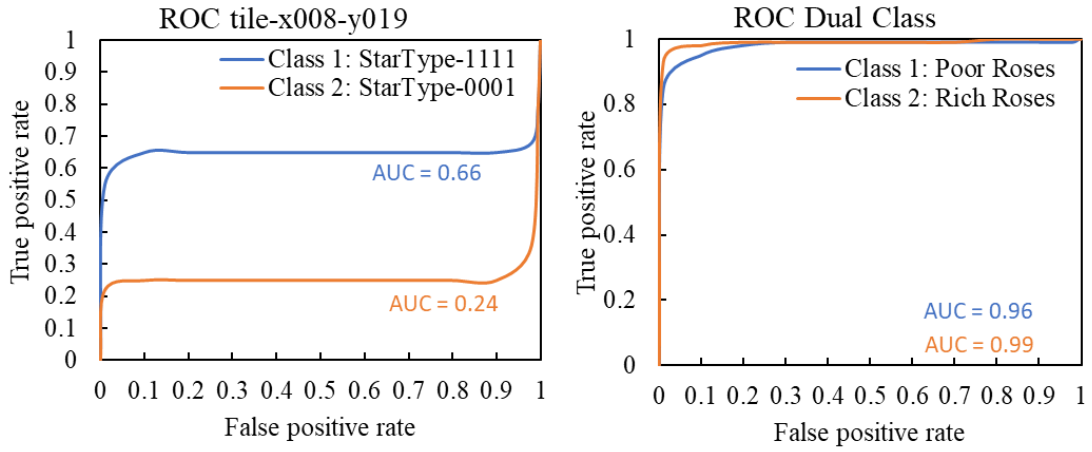


Fig. 13 ROC charts obtained after training different models with two classifiers

C. Synthesized stars

After performing multi class training and segmenting, we train a model with stars that have been generated using a specialized instrument designed for the purpose of creating synthetic stars 34. The process created 50 stars with 50 leaves, 50 stars with 100 leaves, and 15 stars with 300 leaves per rose. The results of synthesized star training model utilizing a retrain procedure consists of 40 epochs in the first stage, 10 in the second, and 40 in the third: Mini-Batch Accuracy of 74.04%, Validation Accuracy of 73.65%, Mini-Batch Loss of 0.1893, and Validation Loss of 1.9106.

Figure 14 shows ROC charts and AUC computations of segmentation images from multiple models. The top-charting model was trained on synthesized StarType_50, StarType_100, and StarType_200 stars. StarType_50 and StarType_200 performance is shown in the chart. AUC scores were 0.99 for the former and 0.94 for the latter. Another illustration of the same model with all three classes is on the left. A model trained with A50, A100, A300 stars is shown in the bottom plots. The right graph provides good AUC scores for the 3 classifiers: 0.94, 0.91, and 0.89. The left graphic provides AUC findings for the three classifications: 0.95, 0.96, and 0.61. All results are good, except for the A300 classification, which is only better than random guessing (better than 0.5).

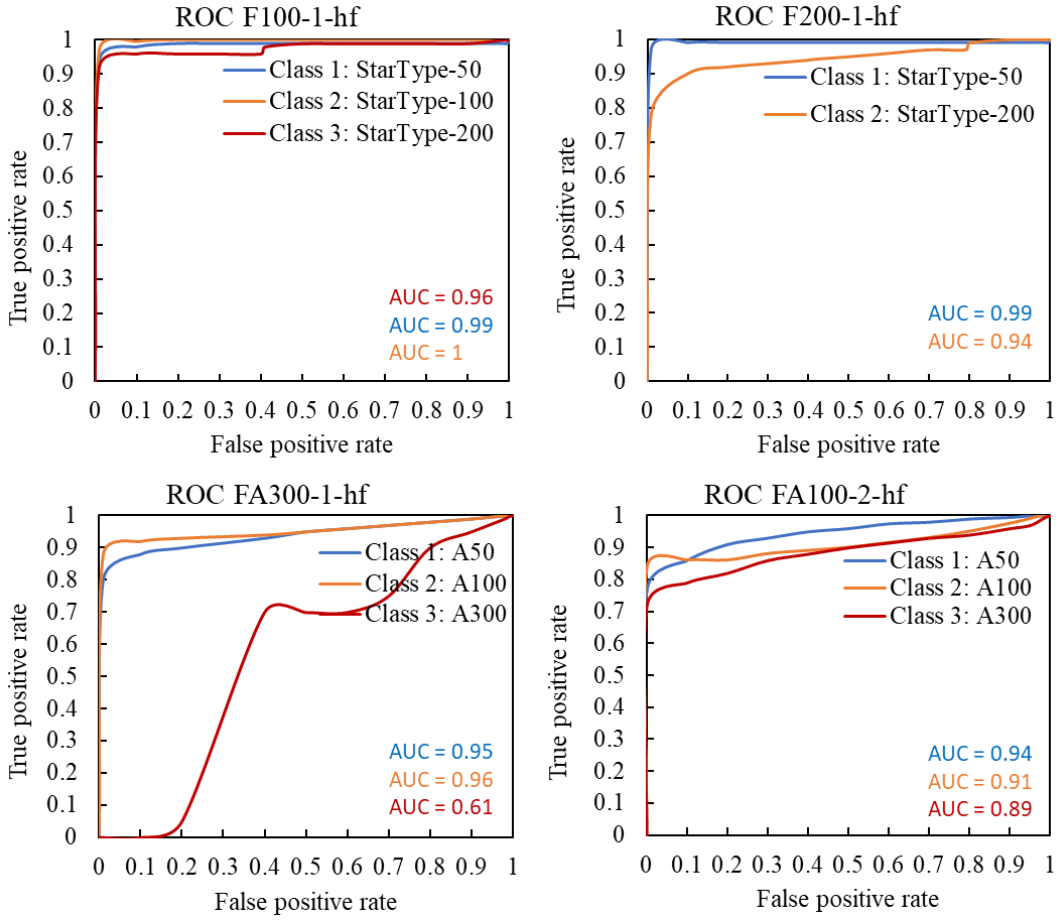


Fig. 14 ROC charts obtained for segmentation images from three-classifier models

Conclusions

This paper has introduced a novel methodology for detecting star shapes in microscopic images using advanced deep learning segmentation and classification. A U-Net FCN model was developed to accurately segment star-like patterns in single-class and multi-

class scenarios. The methodologies described in this study have been effectively employed as a novel analytical technique for the identification of fission track stars.

A model was created for stars with a diameter of less than 60 microns (200 pixels), less than 10 leaves, and no black center (type 0001). The model had a 0.84 ROC area. A computational model was also created to evaluate type 1111 stars of higher magnitude and brightness. This model had an outstanding ROC curve area of 0.90. Early models were also designed to diagnose star shape and size variations simultaneously. It has been observed that as epochs increase, model accuracy increases. After 20 epochs, the ML model attained ~85% accuracy, comparable to human effort. After 50 epochs, the model attained 90% accuracy, surpassing human effort. After 150, it reached 92%.

The research included creating a new star database, classifying star types, characterizing the model, designing an architecture, and optimizing. The research also established segmentation measures, optimized the number of epochs and validation frequency, and investigated background noise filtering thresholds. In addition, detection capabilities and adaptive threshold definitions were improved and Image Labeler tool was used to generate labeling information semi-automatically, followed by Auto Labeling automation, however these issues are not covered in this paper. A model with synthetic stars generated by a simultaneously built tool was created in the project's latter phases. The initial outcome of the Retrain approach in this collaboration yielded a classification accuracy of 73.65%. Three different types of stars were trained.

There are various directions in which we intend to continue expanding and deepening this work in subsequent studies. Improving filters and optimizing SNR to improve star signal and reduce background noise. Enhancing the algorithm and labeling automation. In the context of creating synthetic data, it is possible to expand in many directions, such as optimizing filters, improving and assessing the accuracy of simulation to produce stars from parallel work, or examining innovative AI-based data synthesis tools like GAN (Generative Adversarial Networks), whose use has grown in the last year[4]. This work began comparing model execution between SGDM, Adam, and RMSProp

[4] In the context of synthetic data, since this is an area that is particularly relevant for AI systems, for which there is an ever-increasing need for data, there is a rapid wild race in the academic and commercial world in which prominent universities 36, large and research-intensive companies such as Google 37 and young and fast start-up companies 38,39. Gartner predicts that by the year 2024 synthetic information will make up 60% of the data used for the development of AI and analytical projects 40,41.

optimizers. Add more solvers to complete the comparison. It is also possible to compare Softmax, Sigmoid, Cross Entropy, L1Loss, L2Loss, Huber, Mse (half mean square error), and CTC loss functions. Also comparing this work's U-Net network to novel sub-models like Deformable U-Net. It is also feasible to expand from a segmentation model to a detection model by adding bounding boxes and comparing findings. It is possible to extend the investigation to identify the precise coordinates of the detected stars, as well as investigate star types frequencies and star densities per unit area. A correlation test can be used to compare laboratory data with MS-ICP particle observations of particle size, isotopic composition, and star shape. Another expanding option is to combine arms with parallel study of Single Tracks detection in a high-flux gamma detector based on uranium foil that converts the gamma flux into fissions (fγ).

Acknowledgements

We would like to express our appreciation to all contributors to the success of this project. We wish to thank NRCN organization for funding the project.

All data for this project is available and can be reached by email request: n.elgad@gmail.com.

References

1. N. Elgad (2022) Nuclear Forensics – Fission Track Analysis – Star segmentation and classification using deep learning. M.Sc. Thesis. Ben Gurion University
2. Donohue, D.L., (1998) Strengthening IAEA safeguards through environmental sampling and analysis. Journal of Alloys and Compounds. 271-273: p. 11-18.
3. Bush, W., et al. (2001) IAEA experience with environmental sampling at gas centrifuge enrichment plants in the European Union. in Proceedings of the Symposium on International Safeguards: Verification and Nuclear Material Security.
4. Stebelkov, V. (2005) Informativeness of morphology analysis of uranium microparticles from industrial dust at nuclear plants. in GLOBAL 2005: Proceedings of the international conference on nuclear energy systems for future generation and global sustainability. Japan: Atomic Energy Society of Japan.

5. Lee, C.-G., et al. (2006) Improved Method of Fission Track Sample Preparation for Detecting Particles Containing Fissile Materials in Safeguards Environmental Samples. Japanese Journal of Applied Physics. 45(3L): p. L294.
6. Inglis, J.D., et al. (2023) An introduction to the significance of sample size in particle analyses for nuclear forensics and radiological investigations. Journal of Radioanalytical and Nuclear Chemistry. 3371-3381.
7. Salbu B, Lind OC (2020) Analytical techniques for charactering radioactive particles deposited in the environment. J Environ Radioact 211:106078
8. Park R et al (2022) Quantitative and isotopic analysis of single micrometer-sized uranium particles using multiple mass spectrometric techniques. J Radioanal Nucl Chem
9. Nuclear Forensics (Retrieved on July 14, 2020) IAEA org.
10. D. Michie et al. (1994) Machine Learning, Neural and Statistical Classification. Technometrics
11. T. Mitchel (1997) Machine Learning. McGraw-Hill, Stanford, Boston
12. A. Courville, Y. Bengio, I. Goodfellow (2016) Deep Learning. MIT Press
13. D. Amodei, D. Hernandez (2018) AI and Compute. OpenAI
14. Y. LeCun, Y. Bengio, G. Hinton (2015) Deep learning. Nature 521 (7553) pp. 436-444
15. A. Voulodimos, N. Doulamis, A. Doulamis, E. Protopapadakis (2018) Deep Learning for Computer Vision: A Brief Review. Computational Intelligence and Neuroscience. doi: 10.1155/2018/7068349
16. (Ret. 2020-07-15) Google AI algorithm masters ancient game of Go. Nature News & Comment
17. B. Zhang, S. Liu, Y. C. Shin (2019) In-Process monitoring of porosity during laser additive manufacturing process. Additive Manufacturing 28, 497-505
18. M. Rosanovsky, G. Oren, S. Ifergane, O. Beeri (2019) MLography: An Automated Quantitative Metallography Model for Impurities Anomaly Detection using Novel Data Mining and Deep Learning Appproach. arXiv:2003.04226v1 [eess.SP]
19. D. Shen, G. Wu, H. I. Suk (2016) Deep Learning in Medical Image Analysis. Annual Rev. Biomed. Engineer. Vol. 19. pp. 221-248
20. O. Stetzer, M. Betti, J. van Geel, N. Erdmann, J. V. Kratz, R. Schenkel, N. Trautmann (2004) Determination of the 235U content in uranium oxide particles by fission track analysis. Nuclear Instruments and Methods in Physics Research A 525. 582-592
21. I. Halevy, U. Admon, E. Chinea-Cano, A. M. Weiss, N. Dziga, E. Boblil, M. Dagan, I. Orion, R. Radus (2018) Advances in fission-track detection and analysis for nuclear forensics and safeguards investigations. Progress in Nuclear Science and Technology. Vol 5. pp. 175-178. DOI: 10.15669/pnst.5.175
22. A. M. Weiss, I. Halevy, N. Dziga, E. Chinea-Cano, U. Admon (2017) Fission Track Detection Using Automated Microscopy. Journal of Nuclear Engineering and Radiation Science. DOI: 10.1115/1.4036434
23. I. Halevy (2022) Private communication

24. A. Madbouly, M. H. Nassaf, A. M. Diab, S. A. El-Mongy (2009) A Comparative Analysis of Uranium Ore Using Laser Fluorimetric and Gamma Spectometry Techniques. Journal of Nuclear and Radiation Physics. Vol. 4. No. 2. pp.74-81
25. E. Shelhamer, J. Long, T. Darrel (2014) Fully Convolutional Networks for Semantic Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 39 (4):640-651
26. (Retrieved on May 13, 2022) Self-Driving Car in Carla. roboticsclub.mnnit.ac.in
27. T. Nakazawa, D. V. Kulkarni (2019) Anomaly Detection and Segmentation for Wafer Defect Patterns Using Deep Convolutional Encoder–Decoder Neural Network Architectures in Semiconductor Manufacturing. IEEE Transactions on Semiconductor Manufacturing. Vol. 32. No. 2
28. Z. Y-C. Liu, A. J. Chamberlin, K. Tallam, I. J. Jones, L. L. Lamore, J. Bauer, M. Bresciani, C. M. Wolfe, R. Casagrandi, L. Mari, M. Gatto, A. K. Diongue, L. Toure, J. R. Rohr, G. Riveau, N. Jouanard, C. L. Wood, S. H. Sokolow, L. Mandle, G. Daily, E. F. Lambin, G. A. De Leo (2022) Deep Learning Segmentation of Satellite Imagery Identifies Aquatic Vegetation Associated with Snail Intermediate Hosts of Schistosomiasis in Senegal, Africa. Remote Sensing
29. A. Valizadeh, M. Shariatee (2021) The Progress of Medical Image Semantic Segmentation Methods for Application in COVID-19 Detection. Computational Intelligence and Neuroscience
30. L. Yang, Y. Zhang, I.H. Guldner, S. Zhang, D.Z. Chen (2016) 3d segmentation of glial cells using fully convolutional networks and k-terminal cut. MICCAI. pp. 658-666
31. A.S. Aydin et al (2017) CNN Based Yeast Cell Segmentation in Mutli-Modal Fluorescent Microscopy Data. Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR). pp.753-759
32. O. Ronneberg, P. Fischer, T. Brox (2015) U-NET: Convolutional Networks for Biomedical Image Segmentation. arXiv:1505.04597
33. M. Zhang, X. Li, M. Xu, O. Li (2017) Image Segmentation and Classification for Sickle Cell Disease Using Deformable U-NET. arXiv:1710.08149v3
34. I. Halevy, S. Cerna-Maskova, R. Babayew, N. Elgad, J. Lorincik, D. Micheli (2022) Nuclear Forensics – Fission Track Analysis – Simulation and Image Analysis
35. U. Admon, E. Chinea-Cano, N. Dzigal, I. Halevy, E. Boblil, T. Elkayam, A. Weiss, and S. Vogt (2014) Advancements in Particle Analysis Procedures and Their Application in Characterizing Reference Materials for Safeguards. Symposium on International Safeguards: Linking Strategy, Implementation and People. Vienna, Austria. Oct. 20–24.
36. A. Jahanian, X. Puig, Y. Tian, P. Isola (2021) Generative Models as Data Source for Multiview Representation Learning. MIT Computer Science and Artificial Intelligence Laboratory
37. M. Bileschi (2022) Using Deep Leaning to Annotate the Protein Universe. Google AI Blog – The latest from Google Research
38. K. Wiggers (2022) Synthesis AI raises $17M to generate synthetic data for computer vision.
39. (2021) 7 Synthetic Data Companies for Training AI Models. nanalyze

40. A. White (2021) By 2024, 60% of the data used for the de-vel-op-ment of AI and analytics projects will be synthetically generated. Gartner
41. (2021) Maverick* Research: Forget About Your Real Data — Synthetic Data Is the Future of AI. Gartner Research

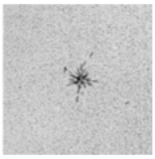
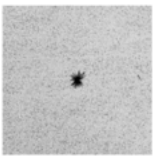
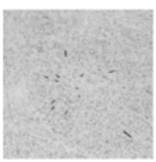
Appendix A

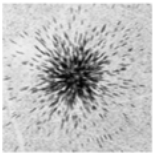
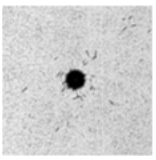
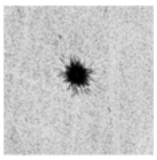
Table 1 Binary notation for star classification criteria

Attr.	Large	Small	Rich	Poor	With black center	No black center	Long halo	Short halo
Mark	1	0	1	0	1	0	1	0
Star Size	> 200px i.e. > 60μm	≤ 200px i.e. ≤ 60μm	> 15 leaves	≤ 15 leaves	>70px i.e. > 20μm	≤ 70px i.e. ≤ 20μm	> 30px i.e. > 10μm	≤ 30px i.e. ≤ 10μm

* px = pixels

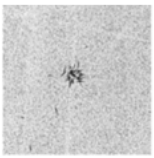
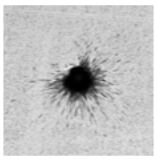
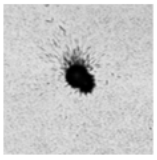
Table 2 Typical images and color setting for labeling star types

Star Description	Typical Shape	Code	Size
background			
non star		0000	
very poor star < 10 leaves small without black center		0001	< 200 pixels
non poor star without black center		0101	< 300 pixels
small with black center		0111	< 200 pixels
huge but poor star without black center		1001	> 600 pixels

Star Description	Typical Shape	Code	Size
huge and rich star with black center		1101	> 600 pixels
big star, with black center and short halo		1110	> 300 pixels
big star, with black center and long halo		1111	> 300 pixels

612

613 **Table 3** Unique star types

Star Name	Description	Typical Shape
Hovered star	A star composed of fission tracks in an undefined and ununiform direction	
Asymmetric star	A star composed of fission tracks in a defined asymmetrical direction	
Complex stars	Two stars covering each other	

614

615 **Table 8** An example of results table with Recall and Precision calculations for 4
616 images in a single Rich Star classifier model

Image Name	Class	Pixel Intensity	TP	FN	FP	Recall	Precision
tile_x001_y008	RichRose	255	2097	0	53	1.000	0.975
tile_x001_y009	RichRose	255	2079	0	72	1.000	0.967
tile_x001_y008	RichRose	255	2059	0	68	1.000	0.968
tile_x001_y009	RichRose	255	2007	0	232	1.000	0.896