Machine Learning Classifiers for Surface Crack Detection in Fracture Experiments

Adrien Müller, Nikos Karathanasopoulos*, Christian C. Roth, Dirk Mohr

Chair of Computational Modeling of Materials in Manufacturing, Department of Mechanical and Process Engineering, ETH Zurich, Switzerland

ABSTRACT

Correctly determining the onset of fracture is crucial when performing mechanical experiments. Commonly carried out by visual inspection, here an image-based machine learning approach is proposed to classify cracked and un-cracked specimens. It yields the potential to objectify and automate crack detection, thereby removing sources of uncertainty and error from the post-processing of experiments. More than 30'000 speckle-pattern images obtained from 77 experiments on three specimen geometries are evaluated. They comprise uniaxial tension, notched tension as well as axisymmetric V-bending experiments. Statistical texture features are extracted from all images. They include both first-order (variance, skewness, kurtosis) and higher-order statistical texture features, i.e. Haralick features. The discriminatory power of the texture information is evaluated based on the Fisher's Discriminant Ratio and feature correlations are identified and quantified. Image texture feature subsets of high discriminatory power are used to parse neural network architectures of different complexities from simple perceptron to feed-forward and cascade neural networks. It is found that a small subset of the investigated texture features is highly significant for all experiments. Using this feature subset in conjunction with multi-layer, non-linear and low complexity feed-forward network architectures classification accuracies in the order of 99% are obtained. At the same time, it is shown that linear classifiers are not sufficient to robustly distinguish the state of the specimens, even when high discriminatory power features are used.

Graphical Abstract:

1. Introduction

Correctly determining the instant of onset of fracture in material characterization experiments is of utmost importance when assessing a material's ductility. Especially when calibrating modern ductile fracture models [5, 6, 28] based on a minimal number of experiments only, any measurement error might have a significant impact on the final model predictions. In-plane and out-of-plane experiments are commonly used to investigate the fracture response of sheet metal. Examples are uniaxial tensile specimens with or without cut-outs (e.g. [14]), Nakazima experiments (e.g. [46]), tension-torsion of tubes [17, 24] or cruciform specimens [26]. In all these experiments, it is crucial to detect the onset of fracture to enable the determination of fracture strains directly through digital image correlation (e.g. [1]) or through hybrid experimental-numerical approaches (e.g. [3]).

In most mechanical experiments, the evolution of integral quantities such as actuator forces is monitored. It is common practice to assume a specimen to be crack free up to the load maximum, while a distinct drop in load is often considered as an indication for the onset of fracture. However, this assumption does not always hold true. For exam-

ple, Beerli et al. [7] found cracks with a length of up to 25% of the gage section width well before observing a change in the nature of the force-displacement response, which corresponds to a false negative result when using the maximum load criterion. Conversely, Noder et al. (2020)[33] reported a drop in force in V-bending experiments due to "punch lift-off" well before fracturing, which correspond to a false positive result.

In addition to actuator forces, images of speckle-painted specimen surfaces have also become an important default recording during mechanical experiments. Surface cracks may thus also be detected through user inspection of the recorded images. This local inspection approach is more sensitive as far as the early detection of surface cracks is concerned. However, this manual technique brings its own uncertainty due to possible human error and subjectivity. Several specimen geometries are usually tested with at least two repetitions and for several material orientations to assess the effects of stress state, anisotropy and material property variations. With the advent of robot-assisted automated testing methods [19], the amount of image data to be screened for emerging cracks has increased substantially, making reliable crack detection a time consuming effort in a fracture property characterization campaign.

---

* Corresponding author
E-mail address: nkaratha@ethz.ch (N. Karathanasopoulos).

Image-based information has been widely used as a basis for the analysis and classification of phenomena in different engineering applications (e.g. [43]). The quantitative analysis of the information contained in the images begins with the computation of first-order feature statistics such as the variance and skewness. However, they ignore information related to the spatial position which is not suitable for pattern recognition [27]. Information on the spatial distribution of the image characteristics can be captured with the help of higher-order statistics. A first example are Gabor filters which have been extensively used in the classification of textures [11]. An even more prominent method for the quantification of the spatial distribution of image features had been proposed by Haralick et al. (1973)[55] based on the so-called Gray-Level Co-occurrence Matrix (GLCM). In many applications, this method has been employed along with the intuitive interpretation of the computed features. Examples are the analysis of skin [35], the classification of landscapes [15] or leaf types [47], as well as the analysis of ultrasound images or MRI data in medicine [31].

Analogously to Digital Image Correlation (DIC), statistical features for texture analysis are computed for subparts of images, the so-called region of interest (ROI). Image types and sizes play an important role in determining a suitable ROI for the feature extraction process [45]. To obtain gray-scale invariance, normalizations with respect to the level of quantization can be applied thereby effectively removing any dependence of the Haralick features on the number of gray-levels in the images [27]. In selected applications, features obtained through the computation of so-called Zernike moments have been employed [50]. These higher-order moments are used to describe image shape features using orthogonal basis functions (e.g. [22]). This accurate identification of patterns of images has many applications, for example the recognition of alphabet characters and symbols [38]. However, as pointed out by Tahmasbi et al. (2011), the accuracy gained often comes at high computational costs.

Recently, machine learning techniques have been extensively used in engineering applications, including, but not limited to, classification problems [18], structural analysis [23], materials design [30], micromechanics [2], constitutive modeling [25, 53] and failure predictions [4, 16, 49]. In material science, a transition from physics-based to more data-driven modeling approaches is observed (e.g. [13, 21, 34]). The introduction of artificial intelligence allows design and analysis methodologies to become more objective, relying less on the subjective judgement of humans (e.g. [9, 32]). For example, basic perceptron models are already sufficient for the proper classification of different material types using ultrasonic data [36]. Artificial neural network models have been successfully applied to simulate structure-property related functions at different material modeling scales [20, 29, 40, 52]. Image analysis driven approaches such as convolutional neural networks (CNN) have also been employed to either identify functional relations [48, 51, 54] or classify structural attributes [42].

While image-analysis based, statistical parameters have been shown to provide a basis for the extraction of important information at different application domains [18], the material-failure characterization potential of features of the kind has not been yet investigated. As a result, the quality and potential of image-based statistical information to accurately identify material failure upon experimental testing remains to a great extend unknown. At the same time, the set of features that are of high significance has not yet been determined. Furthermore, a combined image-data and machine learning approach has not yet been established and the complexity of the machine learning formulation required for highly-accurate material failure characterization remains unknown for different experimental-testing cases.

In the present study, statistical image features are used in combination with machine learning classifiers to come up with a methodology for the automatic detection of surface cracks in mechanical experiments. Images obtained from tension and bending fracture experiments are used to compute first-order and higher-order Haralick texture features. A subset of highly significant image features is identified using

Fisher's Discriminant Ratio (FDR) along with a correlation matrix for each dataset. Different machine learning models are trained and tested on real data, before recommending the least complex classifier that provides reliable predictions of the instants of onset of fracture in mechanical experiments.

2. Crack detection method

A new method is proposed for detecting surface cracks in mechanical experiments. In a first step, the images of the deforming specimen surface are analyzed to compute statistical image characteristics. Subsequently, a neural network based classifier is introduced to discriminate between images with and without cracks.

2.1. Statistical feature extraction

For the unambiguous classification of images acquired during mechanical experiments, an extended set of statistical features is selected. In particular, first-order and higher-order image texture statistics are computed for all experimental images. The first-order statistics include the variance, skewness and kurtosis of the images (e.g. [44]), while the higher-order statistics include twenty Haralicks features as defined in [27]. No image normalizations or histogram equalizations are applied as the gray-level quantization invariant formulation of the Haralick features is used. A summary of all statistical features considered is given in Table 1. The calculation of the Haralick features is based on the square Gray-Level Co-occurrence Matrix (GLCM), calculated by counting the number of times that a gray-level value i prevails at a pixel a while the value j prevails at a pixel b, whereby the pixels a and b are separated by a constant offset (see Fig 1a). Iterating i and j over all possible gray-levels n_g, a complete $[n_g \times n_g]$ GLCM is obtained.

We consider the image function $I(l, k)$: $[l, k] \rightarrow g_{lk}$ that maps the position of a pixel at a position (l, m) for $l < N_x, k < N_y$ to a gray-level g, with N_x and N_y denoting the spatial dimensions of the image in horizontal and vertical direction. The GLCM can be written as

$$P_{ij}\left(\vec{d}, i, j\right) = \sum_{l=0}^{N_y} \sum_{k=0}^{N_x} \delta(I(l,k) - i)\delta\big(I\big(l + d_l, \ k + d_k\big) - j\big)... \quad (1)$$

Herein $\vec{d}$ is the offset vector with the components $[d_l, d_k]$ which describes the relative position of the pixels a and b. Additionally, the GLCM matrix of Eq. (1) is normalized such that:

$$p_{ij} = \frac{P_{ij}}{\sum_{i=0}^{n_g} \sum_{j=0}^{n_g} P_{ij}}... \quad (2)$$

Eq. (2) describes the joint probability mass function for the occurrence of the combination of gray-levels i and j in one image. The offset vector $\vec{d}$ between the pair of pixels is typically chosen to cover four independent directions, as schematically shown in Fig. 1a. Those four independent directions are used for the computation of the GLCM of each sub image, corresponding to the offset vectors [1,0], [1,1], [0,1], and [-1,1]. To achieve rotational invariance, the GLCM of the sub images described in Table 1 is obtained as the mean of the four GLCM matrices corresponding to the previous offset vectors [27]. The mean GLCM is used to compute the different Haralick features of Table 1. The corresponding mathematical definitions of the Haralick features are recalled in Appendix A.

The selection of a subset of features with high discriminatory capacity is made using Fisher's Discriminant Ratio (FDR) [44]. It incorporates the mean and variance of the feature attributes shown in Table 1 for the un-cracked (μ_{unc}, σ_{unc}) and cracked images (μ_c, σ_c) [39],

$$FDR = \frac{\left(\mu_{unc} - \mu_c\right)^2}{\sigma_{unc}^2 + \sigma_c^2}, \quad \mu = \frac{1}{N} \sum_{i=1}^{N} x_i, \quad \sigma = \frac{1}{N-1} \sum_{i=1}^{N} \left(x_i - \mu\right)^2 \quad (3)$$

The FDR values provide a quantitative measure of the potential of the different features to distinguish between cracked and un-cracked

Table 1
Summary of first and higher-order features computed for the set of experiments of Section 3.1

1st order	Higher order: Haralicks			
Variance	Autocorrelation	Diff. Average	Entropy	Max Prob.
Skewness	Cluster Prominence	Diff. Entropy	Homogeneity	Sum Average
Kurtosis	Cluster Shade	Diff. Variance	Inf. Corr. 1	Sum Entropy
	Contrast	Dissimilarity	Inf. Corr. 2	Sum Sq. Var.
	Correlation	Energy	Max Corr.	Sum Variance
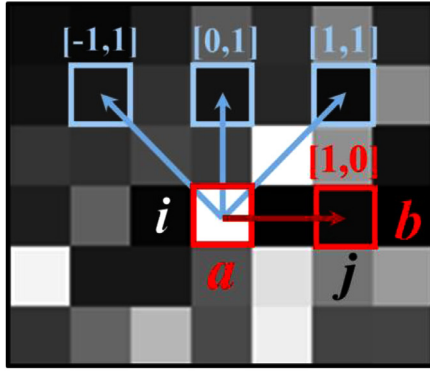
(a)
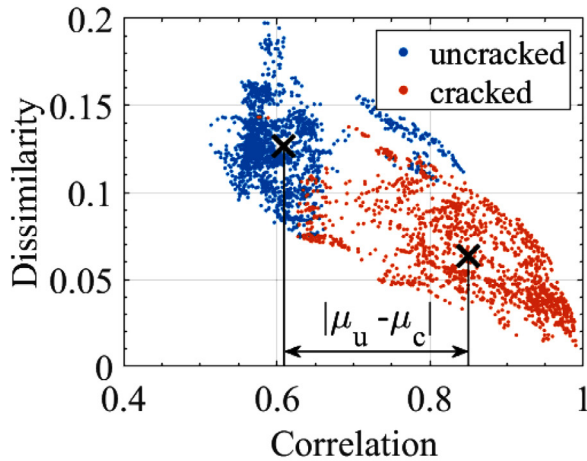
(b)

Fig. 1. Feature extraction and significance identification: (a) Determination of a Grey-Level Co-occurrence Matrix (GLCM): Illustration a pair of pixels a and b (red) containing grey-level values i and j and different offset directions (light blue). (b) Illustration of the Fisher Discriminant Ratio on a scatter-plot of the Haralick feature "dissimilarity" against the feature "correlation" for a set of un-cracked and cracked images extracted from the UT dataset.

images. High FDR values are obtained for classes with a substantial difference in their mean values and low interclass variance. An example is shown for the Haralick features "dissimilarity" and "correlation" in Fig. 1b. All features are then ranked in descending order according to the FDR value. It is noteworthy that the ranking methodology based on Eq. (3) disregards the correlation among different features. Apart from the FDR index, the standardized correlation matrix (with a dimension of 23 × 23) of all Haralick features included in Table 1 is computed for each experiment to quantify the correlation. The associated correlation coefficients are used as an additional selection metric for the identification of low-dimensional feature subsets. For the features vector X [N,

q], comprising the number of observations N and the number of available features q, the correlation matrix C is computed based on Pearson's correlation equation ("Classical Measures of Correlation," [10]):

$$C = \begin{bmatrix} C_{11} & \cdots & C_{1n} \\ & \ddots & \vdots \\ Sym & & C_{nn} \end{bmatrix}, \quad C_{ij} = \frac{\sum_{i=1}^{N} \left(X_i - \overline{X}_i \right)\left(X_j - \overline{X}_j \right)}{\sqrt{\sum_{i=1}^{N} \left(X_i - \overline{X}_i \right)^2} \sqrt{\sum_{i=1}^{N} \left(X_i - \overline{X}_j \right)^2}},$$

$$(i, j \in 1 \dots n) \tag{4}$$

The correlation matrix described by Eq. (4) is symmetric, with its trace equal to n, with the correlation coefficients to range in $C_{ij} \in [-1; 1]$. The marginal values denote a strong correlation between the corresponding features, while near-zero values indicate no relation at all.

2.2. Machine learning classifier

The machine learning classifier considered in this study can commonly be interpreted as a simple function $f : I \rightarrow \bar{y}$ mapping an input I to the output vector $\bar{y}$. Herein, classifier models of different complexity are considered including perceptron (linear model), multi-layer and cascading feed-forward neural networks. The mapping of the input vector I to the output $\bar{y}$ of the classifier model is given for the perceptron (P), multi-layer feed-forward (FF) and cascade feed-forward (FFC) network architectures as follows [12]:

$$\bar{y}^P = H(W^1 I + W_0^1)$$
$$\bar{y}^{FF} = f^n\left(W^n f^{n-1}\left(\dots W^2 f^2\left(f^1\left(W^1 I + W_0^1\right) + W_0^2\right)\dots\right) + W_0^n\right)$$
$$\bar{y}^{CFF} = f^n(W^n f^{n-1}(\dots W^2 f^2(\underbrace{f^1\left(W^1 I + W_0^1\right)}_{y^1} + W_0^2)\dots)..$$

$$+ W_0^n + W_1^n y^1 + \dots + W_n^n I) \tag{5}$$

Herein H describes the hard-limit activation function, while f^i represents the activation function of the different layers contained in the neural network design. The input feature vector is denoted with I, while the weight matrix and bias vector of layer i are denoted by W^i and W_0^i, respectively. The number of hidden layers (n) as well as the number of neurons per hidden layer (m) are hyper-parameters which need to be determined in addition to optimizing the weight parameters of the models using the backpropagation algorithm (see [37]).

A schematic of the different neural network architectures is provided in Fig. 2. The simple perceptron is shown in Fig. 2a, while the feed-forward neural network architectures are schematically shown in Fig. 2b. They comprise a minimum of two and a maximum of six hidden layers (Fig. 2b), which are parsed with a variable number of neurons per hidden layer (Fig. 2b), following a decreasing number of neurons per hidden layer. The outputs of the last layer are fed into a logsig function to yield its classification result. For the activation of the neurons contained in the hidden layers, linear, logarithmic sigmoid and hyperbolic tangent functions are considered. The mean square error is used as loss during the training of the networks, with $\bar{y}$ and y denoting the model predictions and the experimental classification labels (0=uncracked, 1=cracked) of the image, respectively,

$$MSE = (1/N) \sum_N (y - \bar{y})^2, \quad \bar{y} \in \left[\bar{y}^p, \bar{y}^{FF}, \bar{y}^{CFF}\right] \tag{6}$$
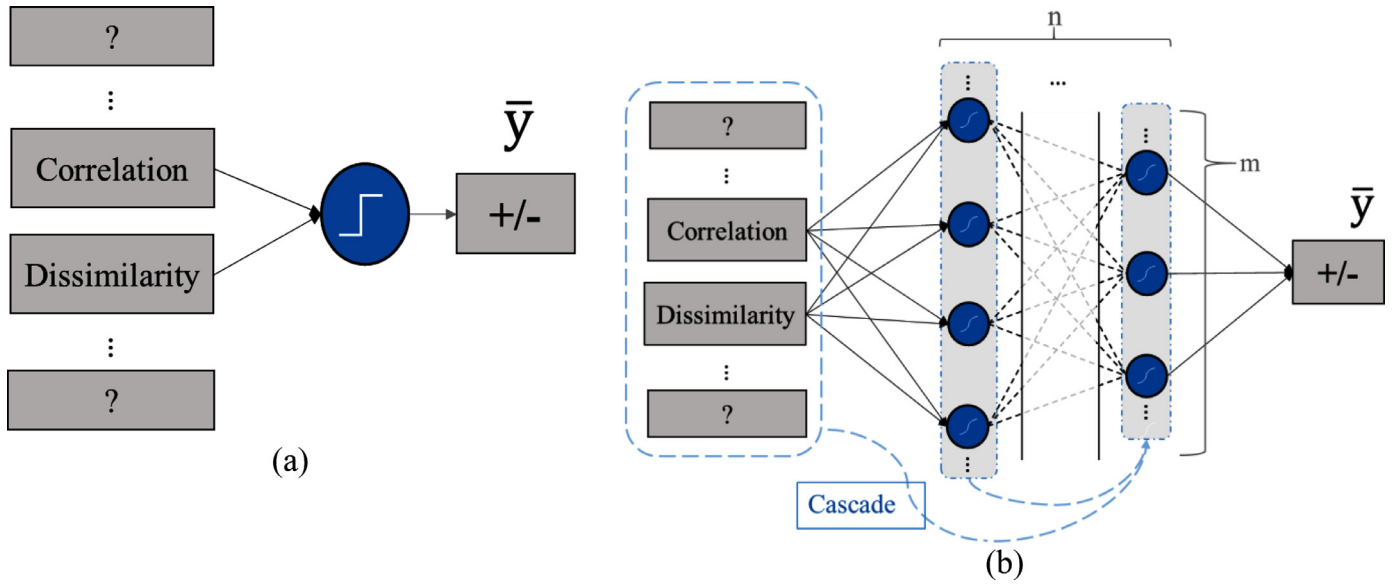
3

Fig. 2. Schematic representation of (a) a simple perceptron, hard-limit neural network architecture along with (b) a fully-connected, feed-forward network with a total of n hidden layers and variable number of neurons per layer with varying number of input features. A cascaded network architecture is indicated by a dashed blue line.

Different optimization algorithms are employed for training (Levenberg-Marquardt, Gradient descent, Conjugate Gradient Descent and Bayesian regularization).

The accuracy of the neural network classifier is determined by the ratio of the correctly classified cracked and un-cracked images over the total number of tested images. The correctly classified images are denoted as true-positive (TP) and true-negative (TN), while false network classifications are denoted as false negative (FN) and false positive (FP). The accuracy, sensitivity and specificity of the classifier are defined as

$$
\begin{aligned}
Accuracy &= \frac{TP + TN}{TP + TN + FN + FP}, \\
Sensitivity &= \frac{TP}{TP + FN}, \\
Specificity &= \frac{TN}{TN + FP}.
\end{aligned}
\tag{7}
$$

These three non-dimensional quantities fully characterize the quality and robustness of the classifier (e.g. [8]) with near-unity values describing the response of reliable models.

3. Experimental data

Various types of fracture experiments for sheet metal are performed to create a database for training, testing and validating of the proposed crack detection method. The experimental procedures are only briefly described as we focus on the exploration of the image data.

3.1. Image data base

Images acquired during material characterization experiments on sheet metal specimens are employed in this study. They are obtained from monotonic experiments on uniaxial tension (UT) and notched tension (NT) specimens with different cut-out radii (20mm for NT20- and 6 mm for NT6-specimens), see Figs. 3a-c. Images from tension experiments on 28 UT-, 21 NT20- and 23 NT6-specimens are used. The images were acquired at a frequency of 2 Hz which resulted in more than 500 images per experiment. In addition, images recorded at 1 Hz from five AxiSymmetric Bending (ASB) experiments with out-of-plane loading are also included in the data base. The engineering materials tested comprise aluminum alloys as well as uncoated advanced high strength steels of

various grades with different strain hardening and fracture properties. Prior to testing, a random speckle pattern is applied onto all specimens consisting of a very thin white background paint with black speckles of approximately 50μm diameter. This non-uniform optical texture is required for Digital Image Correlation (DIC) to measure the surface displacement field (e.g. Hildt; [41]). Fig. 4 gives an overview on the speckle pattern's gray level distribution for different specimens. It reveals that, due to the different paint jobs (and lighting conditions), a wide range of gray level distributions and thus input data is present.

Fig. 5 shows representative speckle patterns of a UT specimen at different stages of the experiment. The dataset creation for the training of the neural network models follows a three-step process. Initially, the full-size images of the different experiments are categorized in cracked and un-cracked by visual inspection. To accurately assess the onset of fracture and to include images with short cracks in the training datasets, cracked specimens (Fig. 5a) are traced backwards in time, up to the image where small cracks first become visible by eye. Subsequently, quadratic sub images of size 128×128 pixel, which corresponds to approximately one third of the gage section width, are extracted from the cracked (Fig. 5b) and un-cracked (Fig. 5e) specimen pictures. Sub-images are obtained from locations shifted to the left and right of the specimen length symmetry axis (Fig. 5a-c) or the middle radius of the ASB specimens, respectively (Fig. 5d). The offsets are chosen such that the sub-images remain within the region of interest (ROI). In particular, uncracked sub-images are extracted with a uniform distribution from the area in which cracks are expected to occur (Fig. 5d-f), whereas cracked sub-images were extracted from manually selected image positions where a crack was present as illustrated in Figs. 5a-c. The datasets are chosen to contain images with both small and large cracks. The images of the un-cracked (Figs. 5e, f) and cracked specimen (Figs. 5b, c) indicate that a wide range of texture characteristics and gray-level quantization apply to the extracted sub-images, which elucidates the diversity of the patterns encountered in mechanical experiments.

3.2. Training and testing data

For training and testing of classifier models, sub-images of the NT20- and NT6-specimens are grouped into a single dataset, while images of the UT- and ASB-specimens are treated independently. Table 2 summarizes the number of images and experiments used for each case. It pro-
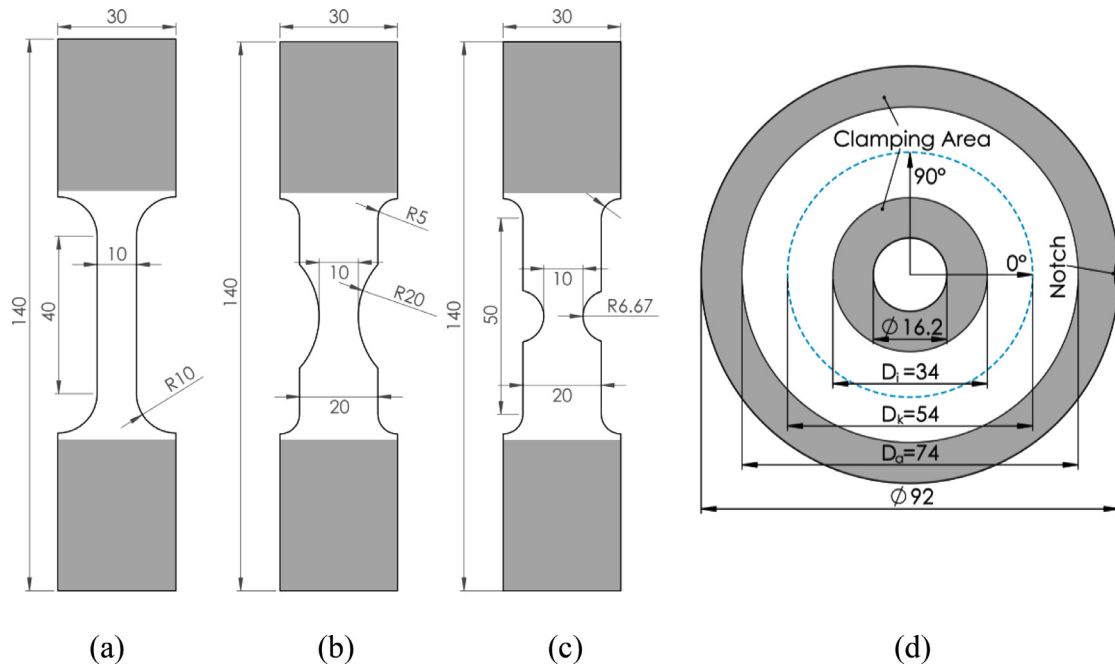
Fig. 3. Technical drawings of the specimen geometries used: (a) uniaxial tension (UT), notched tension with (b) R=20mm cut-out (NT20) and (c) R=6.67mm cut-out (NT6) and (d) axisymmetric V-bending.
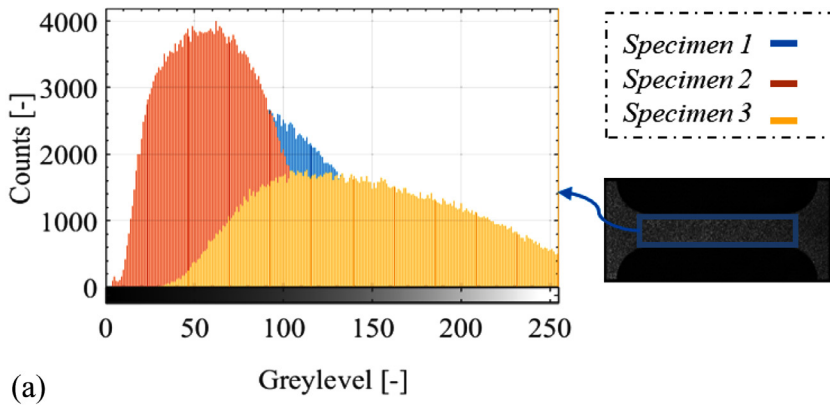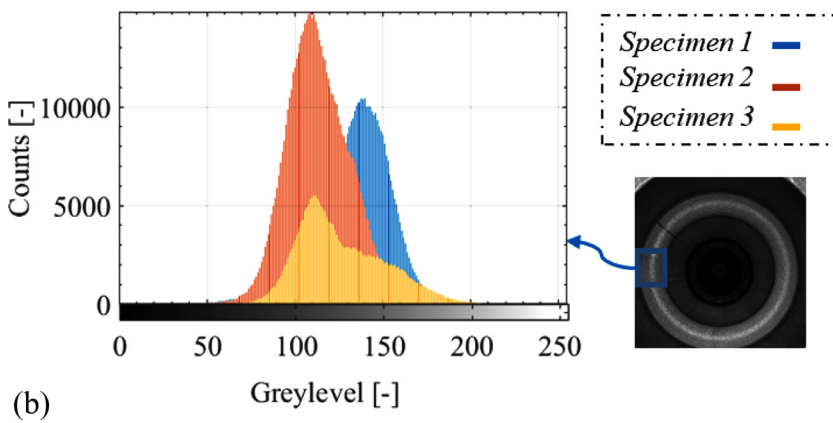

(a)

(b)

Fig. 4. Examples of the gray-level distributions for (a) UT and (b) axisymmetric bending specimens.
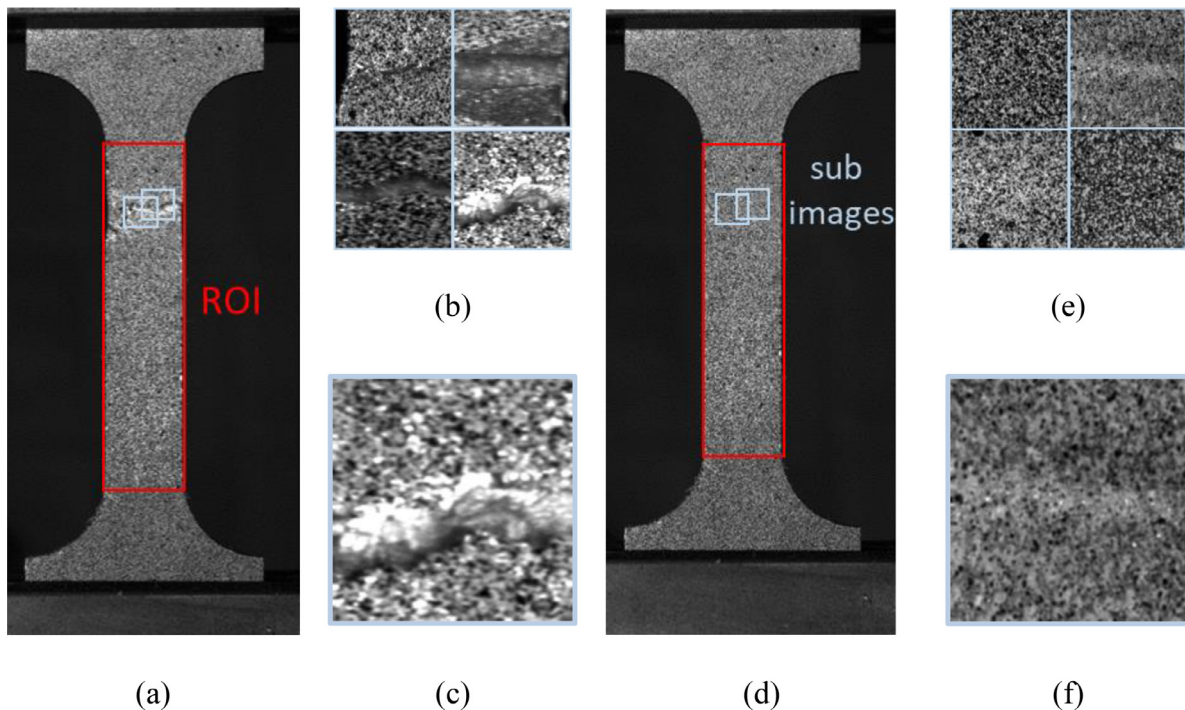
Fig. 5. Images of a uniaxial tensile experiment as used in for the machine learning based crack detection methodology: (a) fully cracked UT sample with area of interest (AOI, red box) and two selected locations of sub images (blue box), (b) selected sub images as used for the classification and (c) close-up of a cracked sub image. The corresponding images of the un-cracked specimen are shown in (d)-(f).

Table 2
Summary of the experimental data employed for the UT, NT6-NT20 and ASB experimental case.

Type	Nr. Images (cracked/ uncracked)	Nr. of Experiments	Size (Pixels)
NT6 & NT20	3644 / 3644	43	128 × 128
UT	2452 / 2452	28	128 × 128
ASB	10'955 / 11'692	5	128 × 128

vides the number of distinct experiments used to generate the data sets, the sub-image dimensions in pixels as well as the number of cracked and un-cracked images in each set. To investigate the accuracy of the proposed crack detection method, data for testing needs to be provided in addition to training data. For this, the full data set of one specific specimen geometry is divided into training and test sets so that the two sets are mutually exclusive. This avoids any duplications, ensuring that the test sets contain only data unknown to the learning algorithm. The complete set of test images for the NT6-/NT20- and for the UT-specimens is provided as Supplementary Material.

4. Results

4.1. Texture features

To identify the most discriminating first-order as well as second-order Haralick texture features, Fisher's Discriminant Ratio (FDR) index is evaluated. Fig. 6a summarizes the FDR indices for the ten most significant features for the notched tensile geometries (NT20 and NT6) along with the corresponding correlation matrix (Fig. 6b). All remaining features are omitted due to their low discriminatory power (*FDR* < 1). It is remarkable that none of the first-order image statistics plays a significant role in separating cracked from un-cracked images for the NT-specimens. Solely higher-order Haralick texture features have FDR values with significant discriminatory power. The feature correlation matrix according to Eq. 4 is used as an additional guide to identify linearly correlated features and to choose a reduced subset of independent

feature attributes (Fig. 6b). A threshold of $|C_{ij}| > 0.98$ is selected and plotted with an elliptic indicator that changes its contour from circular to needle-shaped as the correlation changes from fully-independent to perfectly linear. For the NT specimens, the feature pairs {1-2} and {3-4} ("Correlation" - "Inf. Corr. 2" and "Difference Average" - "Dissimilarity") show strong positive linear correlation, while negative linear correlation is obtained for the feature pair {6-7} ("Homogeneity" - "Contrast"). All linearly correlated feature sets are indicated by a dashed box in Fig. 6b. These observations allow for a subsequent removal of any redundant information from the input vector of the machine learning classifier, thereby significantly reducing the computational cost. The selected minimum number of features required for high accuracy classification results is detailed in Section 4.2.

Fig. 7a provides the FDR index values for the most significant features for the uniaxial tension (UT) training set, with the corresponding feature correlation matrix shown in Fig. 7b. Similar to the presentation of the results from experiments on notched specimens (Fig. 6), ten highly significant components are chosen based on their FDR values. However, here the previously chosen constraint of *FDR* > 1 needs to be slightly relaxed for the feature "Difference Variance". Overall, the same ten higher-order Haralick features as for notched tension (Fig. 6a) are found to be most significant for the images from UT experiments (Fig. 7a), while their feature values and ranking are slightly different. Moreover, the same linear correlation between subsets as for the NT is observed: ("Correlation" - "Inf. Corr 2" {1-2}) and ("Difference Average" - "Dissimilarity" {3-4}) show a positive linear correlation, while a negative linear correlation exists for the feature set ("Homogeneity" -
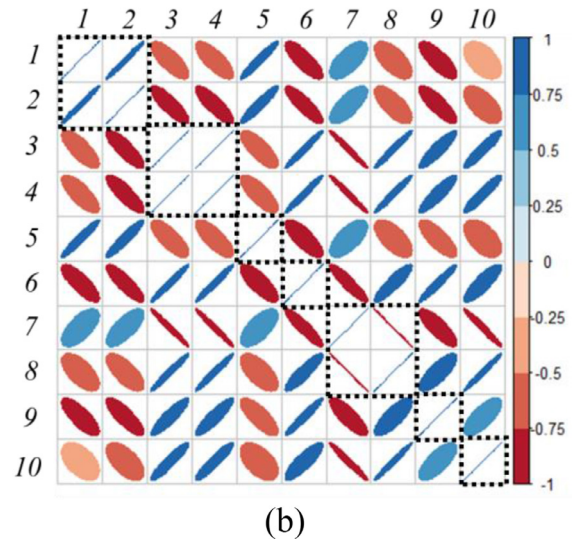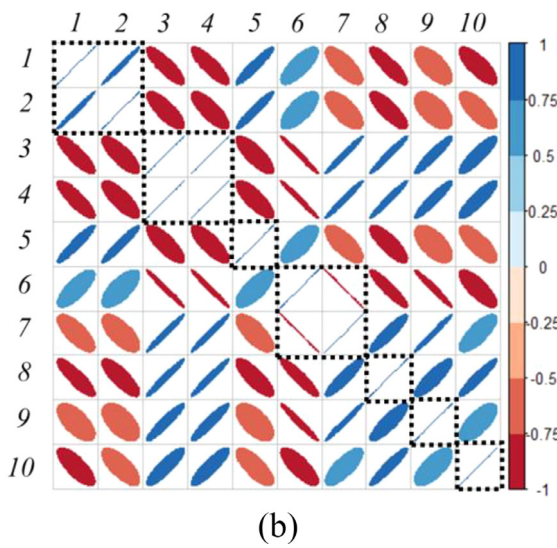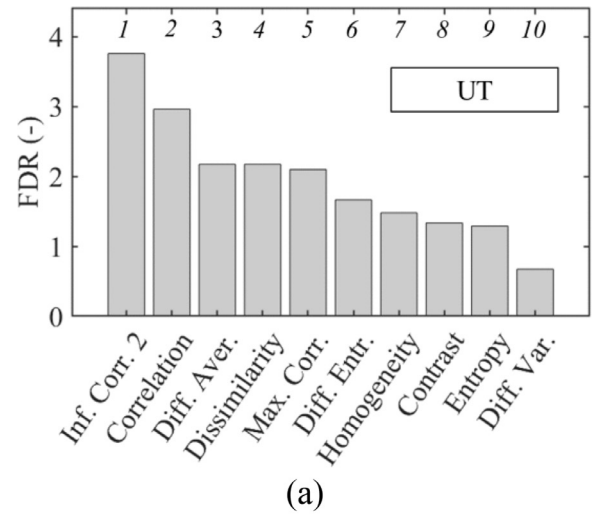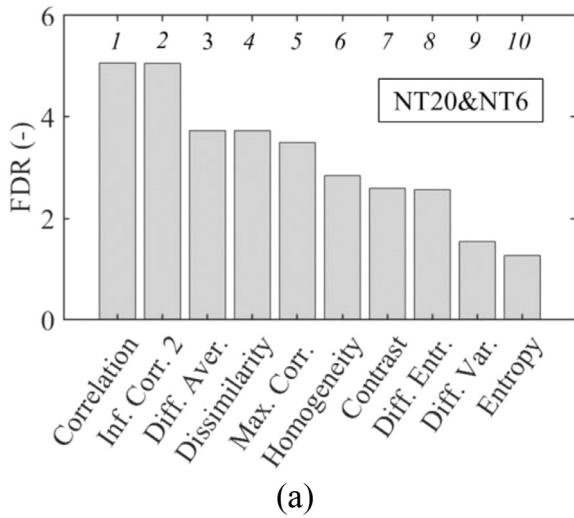
(a)

(b)

Fig. 6. Results of the texture feature analysis: (a) Fisher Discriminant Ratio (FDR) values for the features with the highest discriminatory power for the combined NT20-NT6 dataset, (b) corresponding feature correlation matrix. The dashed boxes indicate linearly dependent features.

(a)

(b)

Fig. 7. Results of the texture feature analysis: (a) Fisher Discriminant Ratio (FDR) values for the features with the highest discriminatory power for the UT dataset, (b) corresponding feature correlation matrix. The dashed boxes indicate linearly dependent features.

"Contrast" {7-8}). Overall, a maximum of six linearly independent features are discerned with an FDR index above unity.

The same assessment is carried out for the images obtained from axisymmetric bending experiments. Fig. 8a shows the most important features with $FDR > 1$, while Fig. 8b provides the corresponding feature correlation matrix with linearly dependent features marked by dashed boxes. As for the images from the other experiments, a linear correlation is observed for all six features. However, the variance (a first-order statistical feature) for the ASB images also yields FDR values well above unity and could hence be used for discriminatory purposes. For all geometries assessed, an excellent separation ability is obtained for several Haralick features, e.g. "correlation", "dissimilarity" and "difference in the entropy". These features appear to be equally significant for all type of experiments considered and well-suited as input variables for classifier models.

4.2. Identification of model hyper-parameters

To identify a reliable neural network model to distinguish between un-cracked and cracked images, a comparative study is carried out using three distinct neural network architectures. A minimum of two and a

maximum of seven linearly independent texture features are selected as input information. They are chosen by their FDR value according to the ranking provided in Figs. 6-8. For example, for a network with two inputs, the pair of independent texture features with the highest FDR value is selected. An overview of the best performing neural networks for up to five texture features is given in Table 3.

The results suggest that poor training and testing scores are obtained when using only two significant Haralick features as model input. When using the simple perceptron model, it can already be deduced from the range of classification scores that these pairs are not necessarily linearly independent. For the feature pair "correlation" and "dissimilarity", the partial interweaving of the feature values for cracked and un-cracked image data is exemplarily shown in Fig. 4b. When using more than four significant features as input, accuracy scores (Eq. (7)) of the order of 95% are already possible with a simple perceptron. With training and testing errors of the order of 1-2%, non-linear feed-forward neural networks allow for significantly more accurate results (Table 3). Even though a range of network configurations is assessed, all results of Table 3 correspond to the network designs with the lowest computational cost with a hyperbolic tangent activation function. The optimal configuration chosen for the feed-forward architecture for all specimen

Table 3

Performance of different neural network architectures for up to five linearly independent features in descending discriminatory power order (Figs. 6-8). Best training results along with the test-set accuracy (Eq. 7) for the complete data set for all specimen types.

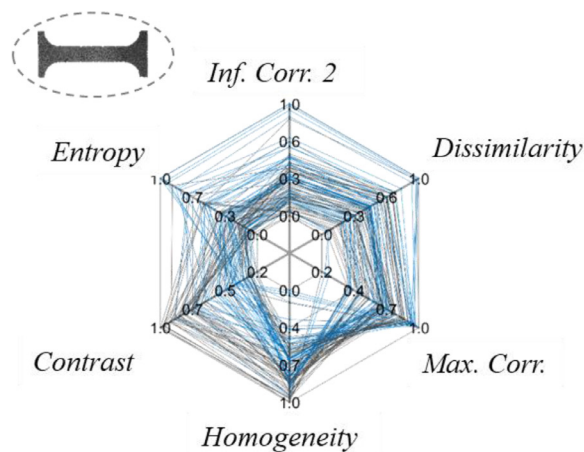| Nr. Features	Set	Perceptron	Feed-Forward	Cascade Feed-Forward
2	UT	87.9%/87.0%	98.1%/98.0%	98.5%/98.4%
	NT	90.7%/90.5%	96.5%/93.7%	97.0%/94.3%
	ASL	92.0%/91.0%	96.0%/97.0%	95.8%/96.2%
3	UT	93.1%/92.5%	99.7%/99.5%	99.4%/99.3%
	NT	92.1%/91.0%	98.5%/95.6%	98.2%/95.2%
	ASL	92.9%/92.7%	97.1%/97.0%	97.0%/97.3%
4	UT	94.4%/93.5%	99.7%/99.7%	99.5%/99.5%
	NT	93.0%/92.0%	98.5%/96.4%	98.4%/95.3%
	ASL	93.7%/93.3%	98.4%/98.2%	98.5%/98.3%
5	UT	94.7%/94.3%	99.8%/99.9%	99.5%/99.8%
	NT	94.5%/94.0%	99.2%/98.8%	99.4%/98.7%
	ASL	95.0%/94.7%	98.6%/98.9%	98.7%/98.4%
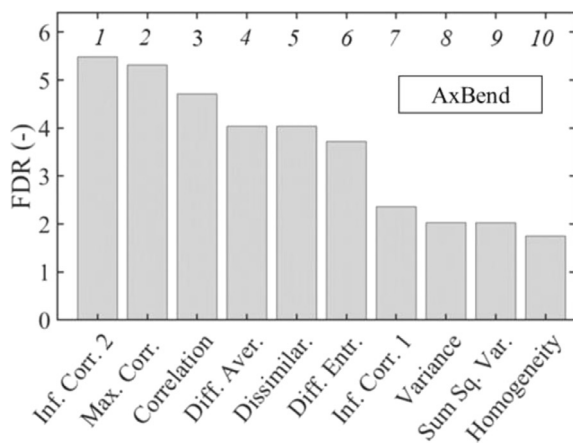
(a)
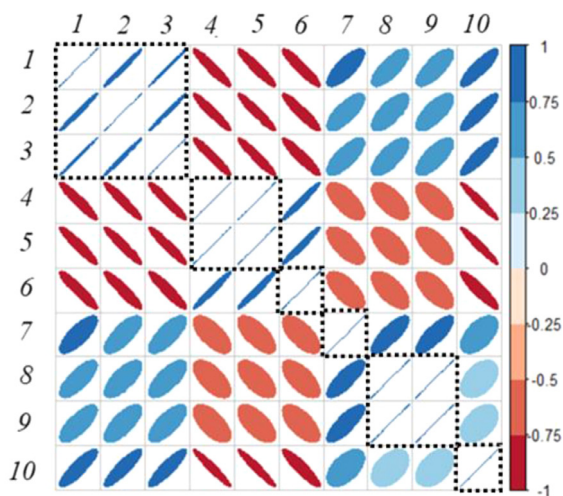
(b)

Fig. 8. Results of the texture feature analysis: (a) Fisher Discriminant Ratio (FDR) values for the features with the highest discriminatory power for the axisymmetric V-bending dataset, (b) corresponding feature correlation matrix. The dashed boxes indicate linearly dependent features.

geometries comprises three hidden layers with a $15 \times 9 \times 3$ node pattern and seven input features.

The required number of training epochs is computationally robust (Levenberg-Marquardt optimization algorithm) and kept below 70 for all input sizes and network architectures considered. In spite of the limited gain in accuracy for more than five input features, all seven independent input features are included in the training process, as of the low additional computational cost. It is noted that a whole training set is completed on average in under four minutes on a standard four-core laptop PC (CPU i7-8550U 1.8 GHz). The use of cascade network architectures does not provide any significant accuracy improvements over the feed-forward network architectures. Similar training and test scores are obtained for all testing cases and number of features employed (relative differences below 1.5% in Table 3). The number of misclassified images lies within 1% for all cases when either a feed-forward or a cascade feed-forward architecture is used.

In an attempt to visualize the effect of cracked (blue) and un-cracked (black) sub-images on the significant texture features, spider plots of the feature values for 100 randomly selected sub-images from all test sets are shown in Fig 9. The corresponding confusion matrices are provided in Figs. 9d-f. The results in Fig. 9a-c provide information on the feature values and on the separability of the cracked and un-cracked image classes for the highly significant features in Figs. 6-8. The weaving of the lines in Figs. 9a-c indicate that a complete separation of the cracked and un-cracked feature values cannot be obtained for any of the features. More specifically, similar feature values are obtained for the cracked and un-cracked sub-images, which becomes visible in the resulting overlap of parts of the spider plot. This observation is in agreement with the inability of linear classifiers (i.e. perceptron) to yield high accuracy scores, irrespective of the number of features employed (as summarized in Table 3).
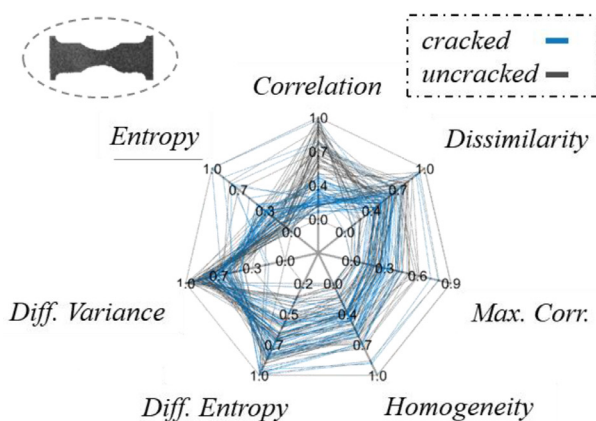
The confusion matrices in Figs. 9d-f provide more detailed information, not only on the accuracy (Eq. (7)) of the models, but also on their sensitivity and specificity. For the UT case (Fig. 9d), a 99.9% testing accuracy is obtained, with a total of 4736 out of 4740 images correctly classified using the multi-layer feed forward model with seven input features. We note that the confusion matrices of Figs. 9d-f test data that are unseen for the machine learning based algorithm and are not included in the training process. The specificity and sensitivity of the model leads to an accuracy error of 0.001. For the NT images, a 98.8% testing accuracy is obtained (Fig. 9e), with the number of false positives being approximately three times the number of false negatives, a marginally higher sensitivity than specificity (Eq. (7)). In the case of the axisymmetric bending, the accuracy amounts to 98.9% with similar sensitivity and specificity.

Fig. 9. Spider plots of the most important feature values for a total of 100 randomly selected cracked (blue) and un-cracked (grey) sub images extracted from the (a) NT, (b) UT and (c) ASB test sets. (d) (e) and (f) show the corresponding confusion matrices of the feed-forward neural network architecture with [15, 9, 3] neurons for UT and NT, and [15, 10, 5] neurons for ASB.

4.3. Automated crack detection in uniaxial and notched tension experiments

The fully-trained multi-layer feed-forward neural network model is used to automatically detect the onset of fracture in the test set of previously unseen images of UT- and NT-experiments. Fig. 10 shows representative results for a UT-test of an aluminum alloy carried out with

a crosshead speed of 2.4mm/min. The force-time curves (Figs. 10a, b) show an initial linear segment followed by an elasto-plastic regime in which a force maximum (~11.5kN) is reached. After the onset of necking, the force level decreases monotonically until an abrupt drop in the force is observed at 172s and approximately 10kN. This corresponds to the instant of catastrophic specimen failure and full specimen separation (blue cross, Fig. 10a,b). This event is detected by the neural net-
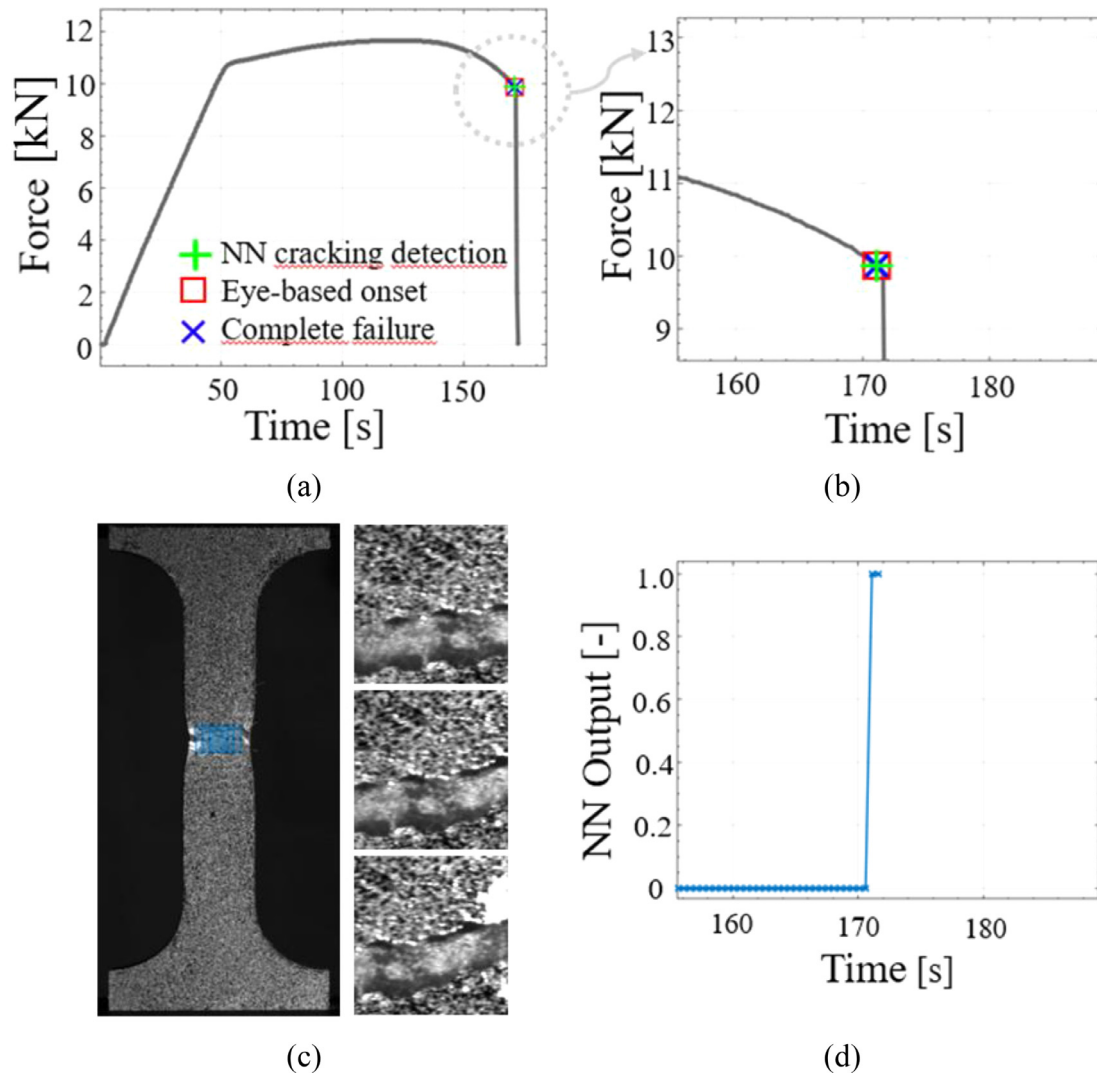
Fig. 10. Results of the automated NN-based crack-detection in a representative UT experiment of an aluminum alloy: (a) Force-time response, (b) close-up of the time interval where fracture occurs, (c) image of the specimen with the positions at which cracks have been detected by the trained NN (blue boxes) and (d) the NN output "failure probability" over time for the interval shown in (b).

work model (green plus, Fig 10a,b) and in excellent agreement with the instant of fracture detected through visual inspection (red box). The location of the onset of fracture is detected correctly by the neural network model (Fig. 10c), together with the three corresponding sub-images. Fig. 10d depicts the failure probability as a function of the experiment time. Being constantly zero for the un-cracked specimen, at 172 s the classifier output jumps to unity (cracked), which correlates with the abrupt and catastrophic manner of the fracture process.

Fig. 11 exemplarily presents the results of an experiment performed on a NT20-specimen of an advanced high strength steel at a crosshead speed of 0.5mm/min. Similar to the UT- experiment, a force maximum (~25 kN) is obtained after an elasto-plastic response. This is followed by a monotonic decrease in force during the necking phase of the specimen. However, in contrast to the UT-specimen, initial cracks on the specimen surface are visually observed (red box) approximately 20 s prior to final fracture (blue cross). The trained neural network with a threshold of 0.5 on the network output detects a crack even earlier than that (~720s, green plus Fig 11b) at the correct location (Fig. 11c). Different from the aluminum alloy, the network output (Fig. 11d) does not show any jump-like change from 0 to 1, but rather a smooth transition, which also reflects the rather slow tearing of the material after the onset of fracture. For example, a threshold close to unity would lead to the detection of

the instant of fracture initiation corresponding the result obtained by eye.

Overall 43 experiments are assessed and classified correctly. The texture features-based neural-network model accurately classifies failure, both for abrupt fracture as well as for failure with slow crack growth. For the earlier, a step-type increase in failure probability is observed (Fig. 10d), while the latter shows a gradual exponential increase of the failure probability (Fig. 11d).

4.4. Automated crack detection in axisymmetric bending experiments

In close analogy with the UT- and NT-experiments, the crack detection methodology is applied to axisymmetric bending experiments that previously have not been used for the training of the network. Fig. 12 shows representative results for an advanced high-strength steel loaded in an axisymmetric bending device at a crosshead speed of 2 mm/min. The force-time curve (Fig. 12a,b) reveals a monotonic increase of the force level through an elastic and a plastic range until an abrupt drop in force (blue cross) is observed after the force maximum (136kN). Prior to that point (and as already pointed out by Beerli et al. [7]), the visually determined onset of fracture, i.e. the first visual crack, occurs approximately 15 seconds earlier at 199s (red box). In contrast, the trained neu-
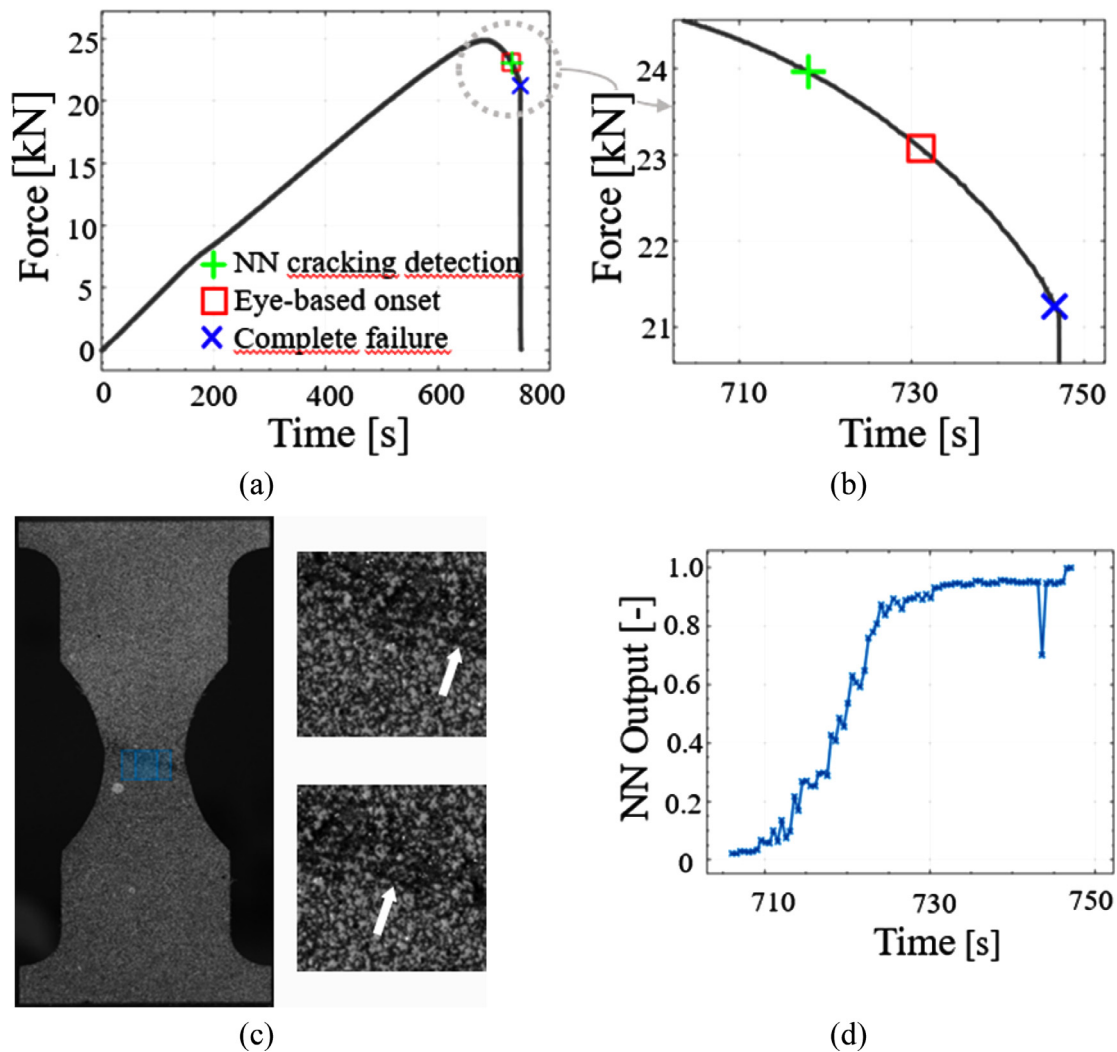
(a)

(b)

(c)

(d)

Fig. 11. Results of the automated NN-based crack-detection in a representative NT20 experiment on an advanced high strength steel: (a) Force-time response, (b) close-up of the time interval where fracture occurs, (c) image of the specimen with the positions at which cracks have been detected by the trained NN (blue boxes) and (d) the NN output "failure probability" over time for the interval shown in (b). White arrows indicate the location of the onset of fracture.

ral network detects a first crack approximately six images later (green plus). Its output, the failure probability, is calculated as the average of the failure probability for all sub images. It yields a constant zero value (un-cracked) for all images prior to the onset of fracture (203s), after which it increases to unity (cracked) within two time points. Fig. 12d shows the onset of fracture as classified by the neural network model at time point 203s as well as a subsequent time point 208s. The green square marks the location of the first detected crack, while the blue squares denote other locations where cracks have been detected. It is noteworthy that already at time step 203, well before any decrease in the force level is visible, cracks are found on both sides of the disc. They continue to grow as loading progresses which is correctly tracked by the machine learning classifier (time point 208s), thereby enabling a characterization of the cracking process. The testing datasets of Fig. 12 and the final neural network model used for the classification are provided as Supplementary Material.

5. Discussion

The results demonstrate that the onset of fracture in mechanical experiments can be successfully detected using a machine learning classifier. However, the success relies on the appropriate selection of the input features and the architecture of the underlying neural network model.

The analysis of the experimental data reveals that none of the computed Haralick features is linearly separable, which made it impossible for simple perceptron classifiers to accurately distinguish among cracked and un-cracked images. A small subset of statistics (Table 1) is found to be highly informative, thus rendering a robust classification feasible. Seven features with linearly-independent Fisher Discriminant Ratio (FDR) values of high significance are identified in Figs. 6-8. The most important features are second-order Haralick features, while first-order statistics do not provide enough relevant information.

Already when using the first two most significant image features, the use of a non-linear feed-forward neural network model allows for training and testing accuracies of the order of 97% and 95% for the images from UT- and NT-experiments, respectively (Table 3). When using two features as model input, a cascade network model provides more accurate predictions than a standard multi-layer feed-forward neural network, at slightly higher computational cost. However, when employing more than four linearly-independent features (Figs. 6-8), differences between the cascade and feed-forward architectures vanish, with the classification accuracy of the neural network approaching and partially exceeding 99% (Figs. 9-12). It is noteworthy that this high-accuracy classification is already obtained solely by using image-based information, while the time-sequence of the images is not used as an input parameter. This shows that neural network designs based on highly informa-
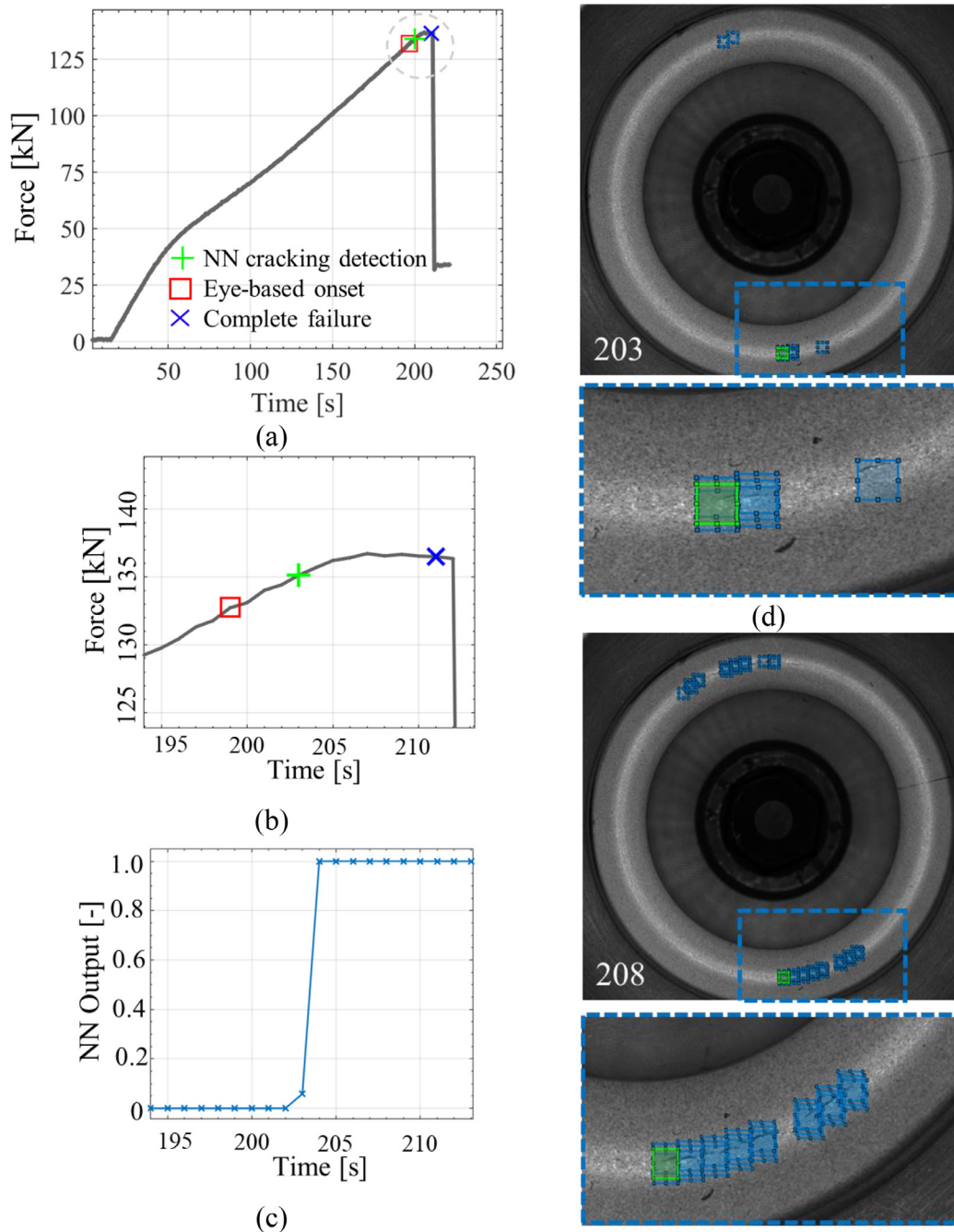
Fig. 12. Results of the automated NN-based crack-detection in a representative Axisymmetric Bending experiment on an advanced high strength steel: (a) Force-time response, (b) close-up of the time interval where fracture occurs, (c) the NN output "failure probability" over time for the interval shown and (d) images of the specimen with the positions at which cracks have been detected by the trained NN (blue boxes) Blue squares denote regions identified as cracked, while a green square indicates the location of initial onset of fracture.

tive statistical information can robustly classify material failure in images obtained from material characterization experiments. The developed methodology automatically detects both the time, as well as the location of the onset of fracture, such that even the simultaneous crack initiation at different positions can be processed and crack progression can be monitored (Fig. 12).

It should be noted that the proposed method is not based on the tracking of displacement fields in time (e.g. digital image correlation (DIC) techniques), making it insensitive to any loss of correlation issue. However, the accuracy and specificity limits of the presented method directly depend on the quality of the provided image data. In particular,

out-of-focus images, low-quality or failed speckle patterns can deteriorate the algorithm's performance, leading to an increased number of misclassified images. Moreover, surface images do not contain any information on the material state within the specimen investigated. Therefore, the elaborated methodology can be used to complement and not substitute, information extracted from specimen-scale, finite element based analysis. Only for specimen geometries with proportional loading histories and onset of fracture on the surface, a hybrid experimental-numerical analysis might be omitted.

It is noted that the speckle pattern itself could lead to a false positive classification of images. However, the analysis of the dataset (>30'000

images) has revealed that instances of this kind are rare, their frequency of appearance being smaller than the misclassification of the neural network model reported. Moreover, a false classification of an image as cracked because of the speckle pattern, usually occurs from the very beginning of the experiment. As such, it remains classified as falsely positive throughout the duration of the test, so that it can be well discerned from surface cracks due to fracture at a later loading state.

6. Conclusion

A machine learning based methodology is developed to automatically detect the onset of fracture in images from material characterization experiments. Overall more than 30'000 images from uniaxial, notched tension and axisymmetric bending experiments are assessed. First- and second-order (Haralick) image features are considered as input for the training of perceptron, multi-layer and cascade feedforward neural network models. It is found that:

1) A small subset of second-order Haralick features including the correlation, dissimilarity, homogeneity and entropy provide high discriminatory power information for crack detection.
2) The accuracy of basic perceptron classifiers is not sufficient for detecting the onset of fracture in mechanical experiments.
3) Computationally-efficient, multi-layer feed-forward neural network models with a small number of second-order image statistics as input, deliver a classification accuracy of up to 99.9%.

The obtained classifier model can be employed to detect both the initiation and propagation of cracks. Even though the present work focusses on the analysis of ductile fracture images, the proposed methodology is readily applicable to brittle fracture problems as well. The machine learning classifier outperforms existing automated fracture detection techniques, such as maximum force criteria. It therefore is expected to become an important element of robot-assisted material testing systems for fracture analysis.

Author Statement

A.M. and N.K. conceived the presented idea, carried out the theoretical and computational analysis and lead the writing of the manuscript. C.R. contributed to the acquisition and processing of the experimental data, as well as to the writing of the manuscript. D.M. supervised the project, provided critical feedback and contributed to the final manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijmecsci.2021.106698.

Appendix A

In Table A1, the mathematical definitions of the different Haralick features (Table 2) are summarized. More elaborate explanations of the underlying physical interpretations can be found in [27]. The mathematical abbreviations needed to compute the Haralick features can be

Table A1
Haralick texture features definitions as derived in [27] for all feature components defined in Table 1.

Abbreviations		Features	
Δ	$\frac{1}{N}$	Autocorrelation	$\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{i}{N}\cdot\frac{j}{N}\right)\tilde{p}(i,j)\Delta_{ij}$
Δ_{x+y}	$\frac{1}{2N-1}$	Cluster prominence	$\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{i}{N}+\frac{j}{N}-2\tilde{\mu}\right)^3\tilde{p}(i,j)\Delta_{ij}$
Δ_{ij}	$\frac{1}{N^2}$	Cluster shade	$\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{i}{N}+\frac{j}{N}-2\tilde{\mu}\right)^4\tilde{p}(i,j)\Delta_{ij}$
$\tilde{p}(i,j)$	$\frac{x(i,j)}{\sum_{i=1}^{N}\sum_{j=1}^{N}x(i,j)\Delta_{ij}}$	Contrast	$\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{i}{N}-\frac{j}{N}\right)^2\tilde{p}(i,j)\Delta_{ij}$
$\tilde{p}_x(i)$	$\sum_{j=1}^{N}\tilde{p}(i,j)\Delta$	Correlation	$\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{\frac{i}{N}-\tilde{\mu}_x}{\tilde{\sigma}_x}\right)\left(\frac{\frac{j}{N}-\tilde{\mu}_y}{\tilde{\sigma}_y}\right)\tilde{p}(i,j)\Delta_{ij}$
$\tilde{p}_y(j)$	$\sum_{i=1}^{N}\tilde{p}(i,j)\Delta$	Difference average	$\sum_{k=0}^{N-1}\frac{k+1}{N}\cdot\tilde{p}_{x-y}(k)\Delta$
$\tilde{\mu}_x$	$\sum_{i=1}^{N}\frac{i}{N}\cdot\tilde{p}_x(i)\Delta$	Difference entropy	$-\sum_{k=0}^{N-1}\tilde{p}_{x-y}(k)\log\tilde{p}_{x-y}(k)\Delta$
$\tilde{\mu}_y$	$\sum_{j=1}^{N}\frac{j}{N}\cdot\tilde{p}_y(j)\Delta$	Difference variance	$\sum_{k=0}^{N-1}\left(\frac{k+1}{N}-\tilde{\mu}_{x-y}\right)^2\tilde{p}_{x-y}(k)\Delta$
$\tilde{\sigma}_x^2$	$\sum_{i=1}^{N}\left(\frac{i}{N}-\tilde{\mu}_x\right)^2\cdot\tilde{p}_x(i)\Delta$	Dissimilarity	$\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\frac{i}{N}-\frac{j}{N}\right|\cdot\tilde{p}(i,j)\Delta_{ij}$
$\tilde{\sigma}_y^2$	$\sum_{j=1}^{N}\left(\frac{j}{N}-\tilde{\mu}_y\right)^2\cdot\tilde{p}_y(j)\Delta$	Energy	$\sum_{i=1}^{N}\sum_{j=1}^{N}\tilde{p}(i,j)^2\Delta_{ij}$
$\tilde{p}_{x+y}(k)$	$\sum_{i=1}^{N}\sum_{j=1,|i+j=k|}^{N}\tilde{p}(i,j)\Delta$	Entropy	$-\sum_{i=1}^{N}\sum_{j=1}^{N}\tilde{p}(i,j)\log\tilde{p}(i,j)\Delta_{ij}$
$\tilde{p}_{x-y}(k)$	$\sum_{i=1}^{N}\sum_{j=1,|i-j|=k}^{N}\tilde{p}(i,j)\Delta$	Homogeneity	$\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\tilde{p}(i,j)}{1+\left(\frac{i}{N}-\frac{j}{N}\right)^2}\Delta_{ij}$
$\tilde{\mu}_{x+y}$	$\sum_{k=2}^{2N}\frac{2(k-1)}{2N-1}\cdot\tilde{p}_{x+y}(k)\Delta_{x+y}$	Information measure of correlation 1	$\frac{\widetilde{HXY}-\widetilde{HXY}1}{max(\widetilde{HX},\widetilde{HY})}$
$\tilde{\mu}_{x-y}$	$\sum_{k=0}^{N-1}\frac{k+1}{N}\cdot\tilde{p}_{x-y}(k)\Delta$	Information measure of correlation 2	$\sqrt{1-\exp[-2(\widetilde{HXY}2-\widetilde{HXY})]}$
$\widetilde{HX}$	$-\sum_{i=1}^{N}\tilde{p}_x(i)\cdot\log\tilde{p}_x(i)\Delta$	Maximal Correlation	$\sqrt{\lambda_2(\tilde{Q}(i,j))}$
$\widetilde{HY}$	$-\sum_{j=1}^{N}\tilde{p}_y(j)\cdot\log\tilde{p}_y(j)\Delta$	Maximum probability	$\max_{i,j}\tilde{p}(i,j)$
$\widetilde{HXY}$	$-\sum_{i=1}^{N}\sum_{j=1}^{N}\tilde{p}(i,j)\cdot\log\tilde{p}(i,j)\Delta_{ij}$	Sum average	$\sum_{k=2}^{2N}\frac{2(k-1)}{2N-1}\tilde{p}_{x+\gamma}(k)\Delta_{i+j}$
$H\tilde{Y}X1$	$-\sum_{i=1}^{N}\sum_{j=1}^{N}\tilde{p}(i,j)\cdot\log[\tilde{p}_x(i)\cdot\tilde{p}_y(j)]\Delta_{ij}$	Sum entropy	$-\sum_{k=2}^{2N}\tilde{p}_{x+y}(k)\log\tilde{p}_{x+y}(k)\Delta_{i+j}$
$\widetilde{HXY}2$	$-\sum_{i=1}^{N}\sum_{j=1}^{N}\tilde{p}_x(i)\cdot\tilde{p}_y(j)\cdot\log[\tilde{p}_x(i)\cdot\tilde{p}_y(j)]\Delta_{ij}$	Sum squares var.	$\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{i}{N}-\tilde{\mu}\right)^2\tilde{p}(i,j)\Delta_{ij}$
$\tilde{Q}(i,j)$	$\sum_{k=1}^{N}\frac{\tilde{p}(i,k)\tilde{p}(j,k)}{\tilde{p}_x(i)\tilde{p}_y(k)}\Delta_{ij}$	Sum variance	$\sum_{k=2}^{2N}\left(\frac{2(k-1)}{2N-1}-\tilde{\mu}_{x+y}\right)^2\tilde{p}_{x+y}(k)\Delta_{i+j}$
λ_2	Eigen$_2(\tilde{Q}(i,j))$	Sum variance	$\sum_{k=2}^{2N}\left(\frac{2(k-1)}{2N-1}-\tilde{\mu}_{x+y}\right)^2\tilde{p}_{x+y}(k)\Delta_{i+j}$

found on the left part of Table A1. The gray-level invariant texture features are presented in the right part of the Table. For the abbreviations and definitions employed, N denotes the number of gray-levels at which the image was quantized, while $x(i, j)$ represents the non-normalized form of the GLCM, which is equivalent to the definition of Eq. (1).

References

[1] Abedini A, Butcher C, Worswick MJ. Fracture Characterization of Rolled Sheet Alloys in Shear Loading: Studies of Specimen Geometry, Anisotropy, and Rate Sensitivity. Exp. Mech. 2017;57:75–88. doi:10.1007/s11340-016-0211-9.
[2] Ali U, Muhammad W, Brahme A, Skiba O, Inal K. Application of artificial neural networks in micromechanics for polycrystalline metals. Int. J. Plast. 2019;120:205–19. doi:10.1016/j.ijplas.2019.05.001.
[3] Anderson D, Butcher C, Pathak N, Worswick MJ. Failure parameter identification and validation for a dual-phase 780 steel sheet. Int. J. Solids Struct. 2017;124:89–107. doi:10.1016/j.ijsolstr.2017.06.018.

[4] Ang JY, Abdul Majid MS, Mohd Nor A, Yaacob S, Ridzuan MJM. First-ply failure prediction of glass/epoxy composite pipes using an artificial neural network model. Compos. Struct. 2018;200:579–88. doi:10.1016/j.compstruct.2018.05.139.

[5] Bai Y, Wierzbicki T. A new model of metal plasticity and fracture with pressure and Lode dependence. Int. J. Plast. 2008;24:1071–96. doi:10.1016/j.ijplas.2007.09.004.

[6] Barsoum I, Faleskog J. Rupture mechanisms in combined tension and shear-Micromechanics. Int. J. Solids Struct. 2007;44:5481–98. doi:10.1016/j.ijsolstr.2007.01.010.

[7] Beerli T, Grolleau V, Mohr D, Roth CC. Axisymmetric V-Bending: a Single Experiment to Determine the Fracture Strain and Weakest Sheet Material Direction for Plane Strain Tension; 2021. under revision.

[8] Bozek J, Mustra M, Delac K, Grgic M. A Survey of Image Processing Algorithms in Digital Mammography. In: Grgic M, Delac K, Ghanbari M, editors. Recent Advances in Multimedia Signal Processing and Communications. Berlin, Heidelberg: Springer Berlin Heidelberg; 2009. p. 631–57. doi:10.1007/978-3-642-02900-4_24.

[9] Butler KT, Davies DW, Cartwright H, Isayev O, Walsh A. Machine learning for molecular and materials science. Nature 2018;559:547–55. doi:10.1038/s41586-018-0337-2.

[10] . In: in: Robust Correlation. John Wiley & Sons, Ltd; 2016. p. 10–32. doi:10.1002/9781119264507.ch2.

[11] Clausi DA, Ed Jernigan M. Designing Gabor filters for optimal texture separability. Pattern Recognit 2000;33:1835–49. doi:10.1016/S0031-3203(99)00181-8.

[12] Demuth HB, Beale MH, De Jess O, Hagan MT. Neural Network Design. 2nd ed. Stillwater, OK, USA: Martin Hagan; 2014.

[13] Draxl C, Scheffler M. NOMAD: The FAIR concept for big data-driven materials science. MRS Bull 2018;43:676–82. doi:10.1557/mrs.2018.208.

[14] Dunand M, Mohr D. Hybrid experimental-numerical analysis of basic ductile fracture experiments for sheet metals. Int. J. Solids Struct. 2010. doi:10.1016/j.ijsolstr.2009.12.011.

[15] Ulaby FT, Kouyate FB.B. and T.H.L.W.. Textural Information in SAR Images. EEE Trans. Geosci. Remote Sens 1986;GE-24:235–45. doi:10.1109/TGRS.1986.289643.

[16] Greve L, Schneider B, Eller T, Andres M, Martinez JD, van de Weg B. Necking-induced fracture prediction using an artificial neural network trained on virtual test data. Eng. Fract. Mech. 2019;219:106642. doi:10.1016/j.engfracmech.2019.106642.

[17] Haltom SS, Kyriakides S, Ravi-Chandar K. Ductile failure under combined shear and tension. Int. J. Solids Struct. 2013. doi:10.1016/j.ijsolstr.2012.12.009.

[18] Huo H, Rong Z, Kononova O, Sun W, Botari T, He T, Tshitoyan V, Ceder G. Semi-supervised machine-learning classification of materials synthesis procedures. npj Comput. Mater. 2019;5:62. doi:10.1038/s41524-019-0204-1.

[19] Jordan B, Gorji MB, Mohr D. Neural network model describing the temperature- and rate-dependent stress-strain response of polypropylene. Int. J. Plast. 2020;135:102811. doi:10.1016/j.ijplas.2020.102811.

[20] Jose KVJ, Artrith N, Behler J. Construction of high-dimensional neural network potentials using environment-dependent atom pairs. J. Chem. Phys. 2012;136:194111. doi:10.1063/1.4712397.

[21] Kalinin SV, Sumpter BG, Archibald RK. Big–deep–smart data in imaging for guiding materials design. Nat. Mater. 2015;14:973–80. doi:10.1038/nmat4395.

[22] Khotanzad A, Hong YH. Invariant image recognition by Zernike moments. IEEE Trans. Pattern Anal. Mach. Intell. 1990;12:489–97. doi:10.1109/34.55109.

[23] Kohar CP, Zhumagulov A, Brahme A, Worswick MJ, Mishra RK, Inal K. Development of high crush efficient, extrudable aluminium front rails for vehicle lightweighting. Int. J. Impact Eng. 2016;95:17–34. doi:10.1016/j.ijimpeng.2016.04.004.

[24] Korkolis YP, Kyriakides S, Giagmouris T, Lee L-H. Constitutive modeling and rupture predictions of Al-6061-T6 tubes under biaxial loading paths. J. Appl. Mech. 2010;77:64501. doi:10.1115/1.4001940.

[25] Kotha S, Ozturk D, Ghosh S. Parametrically homogenized constitutive models (PHCMs) from micromechanical crystal plasticity FE simulations, part I: Sensitivity analysis and parameter identification for Titanium alloys. Int. J. Plast. 2019;120:296–319. doi:10.1016/j.ijplas.2019.05.008.

[26] Kuwabara T, Ikeda S, Kuroda K. Measurement and analysis of differential work hardening in cold-rolled steel sheet under biaxial tension. J. Mater. Process. Technol. 1998;80–81:517–23. doi:10.1016/S0924-0136(98)00155-1.

[27] Löfstedt T, Brynolfsson P, Asklund T, Nyholm T, Garpebring A. Gray-level invariant Haralick texture features. PLoS One 2019;14:e0212110. doi:10.1371/journal.pone.0212110.

[28] Lou Y, Huh H, Lim S, Pack K. New ductile fracture criterion for prediction of fracture forming limit diagrams of sheet metals. Int. J. Solids Struct. 2012;49:3605–15. doi:10.1016/j.ijsolstr.2012.02.016.

[29] Majid A, Khan A, Javed G, Mirza AM. Lattice constant prediction of cubic and monoclinic perovskites using neural networks and support vector regression. Comput. Mater. Sci. 2010;50:363–72. doi:10.1016/j.commatsci.2010.08.028.

[30] Mao Y, He Q, Zhao X. Designing complex architectured materials with generative adversarial networks. Sci. Adv. 2020;6. doi:10.1126/sciadv.aaz4169.

[31] Mayerhoefer ME, Schima W, Trattnig S, Pinker K, Berger-Kulemann V, Ba-Ssalamah A. Texture-based classification of focal liver lesions on MRI at 3.0 Tesla: A feasibility study in cysts and hemangiomas. J. Magn. Reson. Imaging 2010;32:352–9. doi:10.1002/jmri.22268.

[32] Nikolaev P, Hooper D, Webber F, Rao R, Decker K, Krein M, Poleski J, Barto R, Maruyama B. Autonomy in materials research: a case study in carbon nanotube growth. npj Comput. Mater. 2016;2:16031. doi:10.1038/npjcompumats.2016.31.

[33] Noder, J., Dykeman, J., & Butcher, C., n.d. New Methodologies for Fracture Detection of Automotive Steels in Tight Radius Bending: Application to the VDA 238-100 V-Bend Test. https://doi.org/10.1007/s11340-020-00627-z/Published.

[34] Oses C, Toher C, Curtarolo S. Data-driven design of inorganic materials with the Automatic Flow Framework for Materials Discovery. MRS Bull 2018;43:670–5. doi:10.1557/mrs.2018.207.

[35] Ou X, Pan W, Xiao P. In vivo skin capacitive imaging analysis by using grey level co-occurrence matrix (GLCM). Int. J. Pharm. 2014;460:28–32. doi:10.1016/j.ijpharm.2013.10.024.

[36] Roy A, Barat P, De SK. Material classification through neural networks. Ultrasonics 1995;33:175–80. doi:10.1016/0041-624X(94)00022-H.

[37] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature 1986;323:533–6. doi:10.1038/323533a0.

[38] Sadanand AK, Prashant LB, Ramesh RM, Pravin LY. In: Impact of zoning on Zernike moments for handwritten MODI character recognition; 2015. p. 1–6. in: 2015 International Conference on Computer, Communication and Control (IC4). doi:10.1109/IC4.2015.7375516.

[39] Sharma L, Yadav DK, Singh A. Fisher's linear discriminant ratio based threshold for moving human detection in thermal video. Infrared Phys. Technol. 2016;78:118–28. doi:10.1016/j.infrared.2016.07.012.

[40] Singh K, Rajput SK, Mehta Y. Modeling of the hot deformation behavior of a high phosphorus steel using artificial neural networks. Mater. Discov. 2016;6:1–8. doi:10.1016/j.md.2017.03.001.

[41] Sutton M, Orteu J-J, Schreier H. Image Correlation for Shape, Motion and Deformation Measurements. Basic Concepts, Theory and Applications, Image Correlation for Shape, Motion and Deformation Measurements: Basic Concepts. Theory and Applications 2009. doi:10.1007/978-0-387-78747-3.

[42] Swanson Kirk, Trivedi S, Lequieu J, Swanson Kyle, Kondor R. Deep learning for automated classification and characterization of amorphous materials. Soft Matter 2020;16:435–46. doi:10.1039/C9SM01903K.

[43] . Preface. In: Theodoridis S, Koutroumbas K, editors. Pattern Recognition. Boston: Academic Press; 2009. p. 15–17. doi:10.1016/B978-1-59749-272-0.50002-5.

[44] Theodoridis S, Koutroumbas K. Pattern Recognition, Fourth Edition. 4th ed. USA: Academic Press, Inc.; 2008.

[45] Torheim T, Malinen E, Kvaal K, Lyng H, Indahl UG, Andersen EKF, Futsaether CM. Classification of dynamic contrast enhanced MR images of cervical cancers using texture analysis and support vector machines. IEEE Trans. Med. Imaging 2014;33:1648–56. doi:10.1109/TMI.2014.2321024.

[46] Uthaisangsuk V, Prahl U, Münstermann S, Bleck W. Experimental and numerical failure criterion for formability prediction in sheet metal forming. Comput. Mater. Sci. 2008;43:43–50. doi:10.1016/j.commatsci.2007.07.036.

[47] VijayaLakshmi B, Mohan V. Kernel-based PSO and FRVM: An automatic plant leaf type detection using texture, shape, and color features. Comput. Electron. Agric. 2016;125:99–112. doi:10.1016/j.compag.2016.04.033.

[48] Wang K, Sun W. A multiscale multi-permeability poroplasticity model linked by recursive homogenizations and deep learning. Comput. Methods Appl. Mech. Eng. 2018;334:337–80. doi:10.1016/j.cma.2018.01.036.

[49] Wang L, Zhang Z, Long H, Xu J, Liu R. Wind Turbine Gearbox Failure Identification with Deep Neural Networks. IEEE Trans. Ind. Informatics 2017;13:1360–8. doi:10.1109/TII.2016.2607179.

[50] Wang W, Mottershead JE, Mares C. Mode-shape recognition and finite element model updating using the Zernike moment descriptor. Mech. Syst. Signal Process. 2009;23:2088–112. doi:10.1016/j.ymssp.2009.03.015.

[51] Wei H, Zhao S, Rong Q, Bao H. Predicting the effective thermal conductivities of composite materials and porous media by machine learning methods. Int. J. Heat Mass Transf. 2018;127:908–16. doi:10.1016/j.ijheatmasstransfer.2018.08.082.

[52] Yang Z, Yabansu YC, Al-Bahrani R, Liao W, Choudhary AN, Kalidindi SR, Agrawal A. Deep learning approaches for mining structure-property linkages in high contrast composites from simulation datasets. Comput. Mater. Sci. 2018;151:278–87. doi:10.1016/j.commatsci.2018.05.014.

[53] Zhang A, Mohr D. Using neural networks to represent von Mises plasticity with isotropic hardening. Int. J. Plast. 2020;132:102732. doi:10.1016/j.ijplas.2020.102732.

[54] Zheng X, Zheng P, Zhang R-Z. Machine learning material properties from the periodic table using convolutional neural networks. Chem. Sci. 2018;9:8426–32. doi:10.1039/C8SC02648C.

[55] Haralick R, Dinstein I, Shanmugam K. Textural Features for Image Classification. IEEE Transactions on Systems, Man and Cybernetics 1973;SMC-3(6). doi:10.1109/TSMC.1973.4309314.