**Real-time Burn Classification using Ultrasound Imaging**

Sangrock Lee, Rahul, Hanglin Ye, Deepak Chittajallu, Uwe Kruger, Tatiana Boyko, James K. Lukan, Andinet Enquobahrie, Jack Norfleetand Suvranu De

**Supplementary information  
S-1. Grey-level co-occurrence matrix features**

Haralick *et al*.1 introduced various GLCM features to characterize textures of images. Supplementary Table S.1 lists the GLCM features.

**Supplementary Table S1.** GLCM texture features1-3.

|  |  |
| --- | --- |
| **GLCM texture features** | **Formula** |
| Autocorrelation |  |
| Cluster prominence |  |
| Cluster shade |  |
| Contrast\* |  |
| Correlation\* |  |
| Difference entropy |  |
| Difference variance\* |  |
| Dissimilarity |  |
| Energy |  |
| Entropy |  |
| Homogeneity\* |  |
| Information measure of correlation I |  |
| Information measure of correlation II\* |  |
| Inverse difference\* |  |
| Maximum probability\* |  |
| Sum average |  |
| Sum entropy\* |  |
| Sum of squares |  |
| Sum variance |  |

where is the number of quantized grey-levels in the image; is the entry in the normalized GLCM such that where is defined in equation (1); is the marginal probability with respect to rows and is the marginal probability with respect to columns; , is the marginal probability with respect to anti-diagonal components and , is the marginal probability with respect to diagonal components; and , respectively, are the average pixel intensity pairs about and ; and are the standard deviation of pixel intensity pairs about and , respectively; and , respectively, are average pixel intensity pairs about and ; , , , , and are the entropies of corresponding probabilities, where is defined as 0. The features with an asterisk are the selected features using sequential backward selection.

**S-2. Kernel Fisher discriminant analysis**

Kernel Fisher discriminant analysis (KFDA) is a widely used method in machine learning, pattern recognition, and statistics as a tool of dimensionality reduction, classification and data visualization.

KFDA finds an optimal vector which maximizes the between-class variance and minimizes within-class variance by projecting data onto the optimal vector4,5. If the data is subdivided into classes, we can find at most vectors, by solving an eigenvalue problem, associated with between-class variance and within-class variance, and the eigenmatrix is given by

|  |  |
| --- | --- |
|  | (S2.1) |

where are arranged in descent order according to the magnitude of the eigenvalues. From equation (S2.1) the score of a new data point can be defined as

|  |  |
| --- | --- |
|  | (S2.2) |

where is the data point and is the nonlinear mapping. Note that has the most discriminatory information among the scores, *i.e.*, maximizes the between-class variance and minimizes the within-class variance, and provides the second-best result. In this study, corresponds to four burn groups. We use , and for data visualization in three-dimensional space.

**S-3. Support vector machine with radial basis function kernel**

In this section, SVM with RBF kernel is introduced. SVM finds an optimal hyperplane which separates data with the largest margin, where the margin is defined by the Euclidian distance to the closest point from the hyperplane. Finding the optimal hyperplane with soft margin is equivalent to solving an optimization problem given by 6

|  |  |
| --- | --- |
| subject to: | (S3.1) |

where is the normal vector of the hyperplane and is the bias, resulting in a hyperplane . are the training data, are the labels where can be either of +1 or -1, is the slack variable and is a constant which controls the shape of decision boundary such that too small results in underfitting while too large may lead to overfitting. The dual form of equation (S3.1) is given by 6

|  |  |
| --- | --- |
| subject to: , | (S3.2) |

where is the Karush-Kuhn-Tucker (KKT) coefficient vector, is the vector whose components are , is the vector whose components are 1’s and is the Gram matrix whose component . In equation (S3.2), the inequality in the constraint is applied elementwise. By minimizing equation (S3.2) with respect to, we obtain the classifier6

|  |  |
| --- | --- |
|  | (S3.3) |

If a sample data is above the hyperplane  it is classified to +1 group. In a similar way, if a sample is below the hyperplane , it is classified to -1 group. By introducing nonlinear transformation kernel the Gram matrix is transformed into RBF kernel is adopted which is written by , yielding the kernel matrix where is the kernel scale parameter. are crucial factors determining classification performance so the two factors are chosen by leave-one-out cross-validation to ensure the optimal performance. and used in the classification are listed in Supplimentary Table S2. The classifier is rewritten as

|  |  |
| --- | --- |
|  | (S3.4) |

where is one of any support vectors, and is the corresponding label of . Also, we can define a score as follows:

|  |  |
| --- | --- |
|  | (S3.5) |

where the score amounts to the distance from the data point to the optimal hyperplane. Depending on the sign of the score, the group to which the designated data belongs is determined. Time complexity of prediction is where is the number of support vectors and is the number of features7. MATLAB library was used to implement SVM. The CPU time for classification using equation (S3.4) is measured to be of the order of *μs*.

The values of hyper parameters are listed in Supplementary Table S2. The parameters were chosen so that leave-one-out cross-validation can yield the minimum error in pairwise binary classification. Search range of is [0.1, 1000] and that of is [0.1, 100].

**Supplementary Table S2.** KKT vector components () and kernel parameters () used for burn classification.

|  |  |  |
| --- | --- | --- |
| Burn groups |  |  |
| 200ºF for 10s - 200ºF for 30s | 0.1 | 2 |
| 200ºF for 10s - 450ºF for 10s | 100 | 10 |
| 200ºF for 10s - 450ºF for 30s | 0.6 | 1 |
| 200ºF for 30s - 450ºF for 10s | 1000 | 10 |
| 200ºF for 30s - 450ºF for 30s | 1 | 1 |
| 450ºF for 10s - 450ºF for 30s | 1 | 2 |

**References**

1 Haralick, R. M., Shanmugam, K. & Dinstein, I. Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics* **SMC-3**, 610-621, <https://doi.org/10.1109/TSMC.1973.4309314> (1973).

2 Soh, L. K. & Tsatsoulis, C. Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. *IEEE Transactions on Geoscience and Remote Sensing* **37**, 780-795, <https://doi.org/10.1109/36.752194> (1999).

3 Clausi, D. A. An analysis of co-occurrence texture statistics as a function of grey level quantization. *Canadian Journal of Remote Sensing* **28**, 45-62, <https://doi.org/10.5589/m02-004> (2002).

4 Mika, S., Ratsch, G., Weston, J., Scholkopf, B. & Mullers, K. R. in *Neural Networks for Signal Processing IX: Proceedings of the 1999 IEEE Signal Processing Society Workshop (Cat. No.98TH8468).* 41-48.

5 Bishop, C. M., Bishop, P. N. C. C. M., Hinton, G. & Press, O. U. *Neural Networks for Pattern Recognition*. (Clarendon Press, 1995).

6 Cortes, C. & Vapnik, V. Support-Vector Networks. *Machine Learning* **20**, 273-297, <https://doi.org/10.1023/a:1022627411411> (1995).

7 Claesen, M., De Smet, F., Suykens, J. A. K. & De Moor, B. Fast Prediction with SVM Models Containing RBF Kernels. *arXiv e-prints* (2014).