Material Classification with a 7th Sense; Boosting Recognition Ability with Thermal Imagery

Willi Großmann (grossmann@hsu-hh.de)
Helmut Schmidt University

Helena Horn
Helmut Schmidt University

Oliver Niggemann
Helmut Schmidt University

Research Article

Keywords: Material Recognition, Remote Sensing, Thermal Imagery, Transfer Learning, Sensor Fusion

Posted Date: June 10th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1743102/v1

Material Classification with a 7th Sense

Boosting Recognition Ability with Thermal Imagery

Willi Großmann · Helena Horn · Oliver Niggemann

Received: date / Accepted: date

Abstract Material recognition using optical sensors is a key enabler technology in the field of automation. Nowadays, in the age of deep learning, the challenge shifted from (manual) feature engineering to collecting big data. State of the art recognition approaches are based on deep neural networks employing huge databases.

But still, it is difficult to transfer these latest recognition results into the wild – various lighting conditions, a changing image quality, or different and new material classes are challenging complications. Evaluating a larger electromagnetic spectrum is one way to solve these challenges.

In this study, the infrared (IR) emissivity as a material specific property is investigated regarding its suitability for increasing the material classification reliability. Predictions of a deep learning model are combined with engineered features from IR data.

This approach increases the overall accuracy and helps to differentiate between materials that visually appear similar. The solution is verified using real data from the field of automatized disinfection processes.

Keywords Material Recognition · Remote Sensing · Thermal Imagery · Transfer Learning · Sensor Fusion

1 Introduction

Humans recognize materials based on spectral, texture, and context data [1]. Machine vision simulates this cognitive process in many industrial applications. Application areas for material recognition are, for example, sorting processes such as waste separation [2], the monitoring of construction progress [3], and urban or botanical investigations with remote sensing [4]. The knowledge about material properties is of great importance for the interaction of robots with everyday objects [5] or for the ongoing automation of manufacturing and other industrial processes using modern smart technology, also known as Industry 4.0 [6].

Due to the coronavirus pandemic, the development of automatized disinfection processes has recently become a field of special interest. The recognition of materials is highly important in these applications, because material properties influence the persistence of pathogens [7] as well as the effectiveness of the disinfectant [8]. An incorrect material classification could lead to an incorrect application of disinfectant and thus to an insufficient disinfection process.

One contemporary approach of material recognition uses Convolutional Neural Networks (CNN) to base identification not only on the consideration of different kind of visual data but also on created and learned context between given information. Huge existing material databases [9, 10, 11] enable the training of deep CNNs and allow investigations of material recognition possibilities on the basis of images. Additionally, a broad range of pretrained CNNs is available for transfer learning, giving the advantage of less effort to build databases and to train CNNs for new applications. By using such pretrained CNNs, less data is needed to train a solution for specific application scenarios [12].

However, the material classes of existing databases are defined too general for many technical applications. In disinfection or sorting processes, for example, the detection of various metallic materials or the distinction between wood and wood imitation, is a basic requirement. This distinction is much more difficult because visually similar materials such as aluminum and stainless steel have to be considered.

Willi Großmann (corresponding author)
Helmut-Schmidt-University, University of the Bundeswehr
Holstenhofweg 85
D-22043 Hamburg
Germany
Tel.: +49 4421 7584-6311
E-mail: grossmann@hsu-hh.de

Therefore, using only CNNs to evaluate visual appearance of materials no longer seems to be the solution. Even CNNs cannot distinguish between similar visual data and thus between materials with the same colors and textures.

A possible solution to solve this problem could be the evaluation of a larger electromagnetic spectrum, as the material specific information in the data increases. In this case, the identification could be almost 100 % precise, but this would require relatively expensive measurement equipment.

As an alternative solution, the usage of the IR range seems to be a good compromise because it requires rather inexpensive cameras for the detection of thermal radiation [13]. Based on these circumstances, this study follows three research hypotheses (RH):

RH 1: Evaluating the IR range additionally to the VIS range is a cost-effective option to significantly improve the reliability of remote material recognition for industrial processes.

RH 2: In a controlled environment, thermal imaging helps to differentiate between materials that visually appear to be similar. According to the Stefan-Boltzmann law, the emissivity of technical bodies influences their thermal radiation performance. Therefore, the apparent temperature in relation to the true body temperature leads to the infrared emissivity as a material specific property.

RH 3: The median and the variance of the relativized material temperatures are characteristic material features.

Recognition accuracy can be increased when these engineered features are combined with learned VIS features from a CNN. A support vector machine (SVM) is suitable for this feature fusion. This approach outperforms other methods from the field of remote sensing such as data level fusion or image fusion.

The contributions of this paper are briefly described as follows.

i) Material classes with a high technical range of application are differentiated. A distinction is made between different metals and visually similar materials.
ii) For this purpose, the material specific apparent temperature is fused with learned VIS features to increase the recognition accuracy for industrial remote sensing applications.
iii) The proposed fusion algorithm outperforms visual material recognition approaches and comparable sensor fusion algorithms from the field of remote sensing.

The rest of this paper is organized as follows: At first, related literature regarding material recognition approaches, thermal imaging, and electro optical sensor fusion is presented (section 2). Next, details for the proposed architecture (section 3) and for the experimental design of the new material recognition approach (section 4) are provided. Finally, concluding remarks and further investigations (section 5) are proposed.

2 Related Work

Early material recognition approaches evaluated color distributions and patterns. [14] started to study the general perception of materials and suggested differentiating materials based on their reflectivity and shape. [15] suggested an extended rating to cover more aspects of appearance. They found that reflectance, texture, color, shape, and environment illumination are suitable parameters for material classification.

Experimenting on the Flickr Material Database (FMD) [9], [16] found a SVM is better suited to classify these features than previous methods. FMD therefore covers many aspects of the appearance of different materials.

[17] confirmed these results with their investigations and showed that local image information such as color, texture, or shape are not sufficient for material recognition. For better classification results they suggested taking context information (like the object category) into account.

Convolutional neural networks (CNNs) do not require any local image information to be given for material classification. Instead, the algorithm learns the features required for recognition. Indeed, modern methods using CNNs showed significantly better results [10, 18].

The disadvantage of learning these networks from scratch is the high demand for labeled data. Possibilities to reduce this high demand for data are the employment of pretrained models [19, 20].

Recent studies applied ensemble learning approaches to combine multiple pretrained classifiers [21, 11] and achieve better classification results. However, higher accuracies can also be achieved by evaluating more specific material data with less computationally intensive algorithms. The latter could be addressed by the evaluation of IR data in addition to the visual information of the materials.

So far, only a few studies took the IR range into account. Its suitability for material classification has been shown in laboratory environment [22]. Based on a thermodynamic model, they examined the material specific heat conduction for classification using a Nearest Neighbor algorithm. Additional parameters were based on a water permeation experiment. However, the experimental setup was not designed for industrial use. The material samples were heated, moistened, and recorded from a short distance for several minutes.

Another approach also considered changing environmental conditions [13]. Thermal images were taken into the wild and classified using a CNN. However, the recording distance stayed constant at approximately 0.5 m and the materials were chosen with quite characteristic patterns.

A more recent approach exploited thermal conductivities by making physical contact and observing temperatures during heat transfer [23]. This research confirmed that eval-

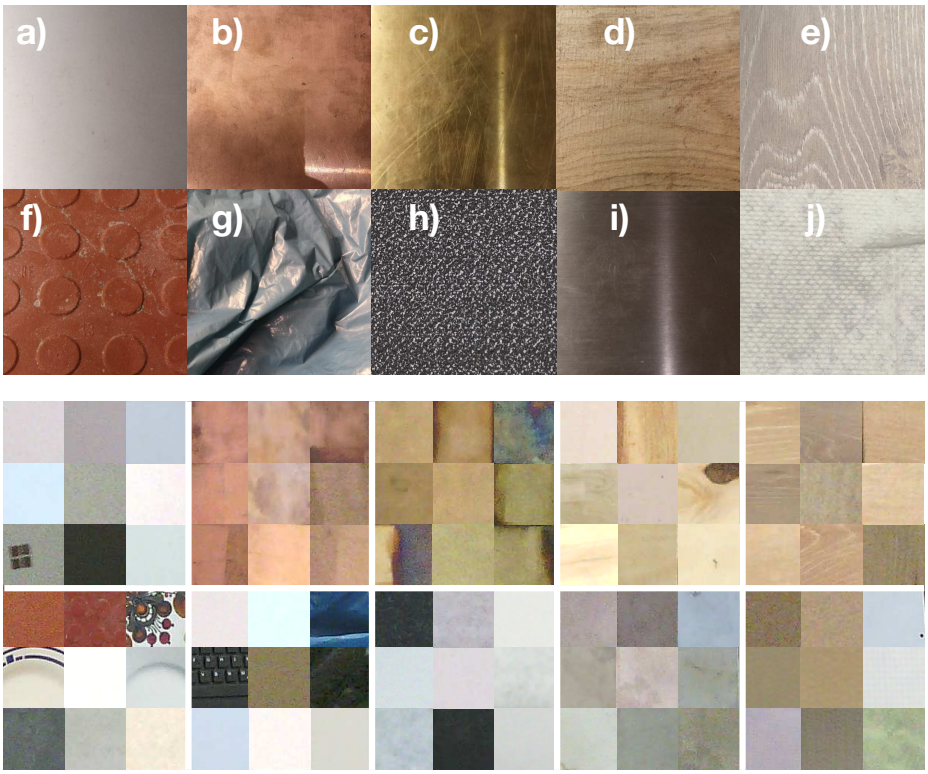

Fig. 1 Material classes and samples of the database used. The first two rows show some material samples: a) Aluminum, b) Copper, c) Brass, d) Wood, e) Wood Imitations, f) Ceramics, g) Plastics, h) Textiles, i) Stainless Steel, j) Cellulose. Some cropped images are shown in the same order in the lower two rows. These images were used for training and evaluation of the CNN.

uating thermal data improves the accuracy of material identification.

While a IR- and VIS data fusion was not applied in most cases for material recognition approaches, it is quite commonly needed for electro-optical systems in different areas of remote sensing. State of the art approaches applied image fusions on different levels [24]. For example, [25] proposed a VGG 19 framework for an image fusion of RGB and IR images on feature level, while a data-level fusion was applied by [26] within a VGG 16 CNN.

However, since images were fused, these approaches did not consider multimodal data. Remote sensing applications required for autonomous driving successfully applied SVMs for multimodal sensor fusions [27]. These applications needed to fuse visual images with multiple disparate data such as radar or vibration information [28].

Therefore, in order to use material recognition algorithms for industrial applications, several investigations have to be done in addition. Based on the research hypotheses, this study examines the following circumstances that have not been taken into account yet:

RH 1) The previous material recognition approaches focused on the investigation of general visual phenomena. The material classes were selected based on specific patterns or general appearance.
The material data of this study focuses on high technical benefit and a wide range of applications. In addition, different lighting conditions and shooting distances are considered to take realistic environmental conditions into account for industrial applications.

RH 2) The basic feasibility of using thermal data for material recognition has already been proven. However, the recognition of visually similar materials or imitations has not been investigated yet.
Everyday materials often appear similar. For example, plastics can be used as an imitation for metals or wood. In order to evaluate this, a distinction is made in this study between visually similar material classes such as aluminum and stainless steel or wood and wood imitation.

RH 3) SVMs are a proven algorithm for classifying engineered and learned features. They are also used successfully to evaluate various sensor data from electro-optical systems.
However, it must first be evaluated whether the measured temperature distributions have material specific properties. From this distributions, features can be engineered that help with material identification.

3 Solution

Since there is no known database that combines IR and VIS data so far, a new database was specially created for this study. The VIS feature extraction and evaluation is done with a CNN, while the final sensor fusion is realized using a SVM. The proposed algorithm fuses learned VIS features with engineered IR features. This enables the identification of visually similar materials.

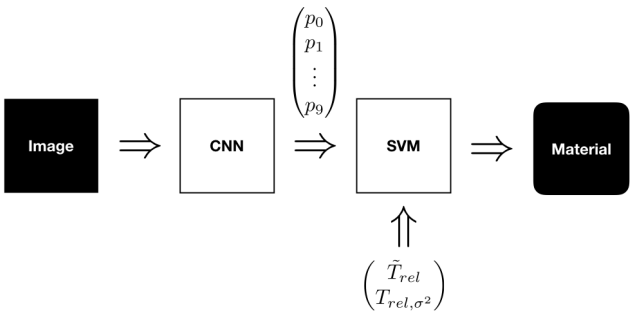

Fig. 2 Algorithm pipeline: The CNN extracts and classifies VIS features from a RGB Image. The Softmax probabilities $\boldsymbol{p}$, the median relative temperature $\tilde{T}_{\text{rel}}$ and the variance T_{rel,σ^2} are part of the feature vector. A SVM merges these features for classification and outputs the resulting material.

This study focuses on raw temperature data instead of false color images. According to the Stefan-Boltzmann law, the temperature data appear comparatively characteristic. This enables a description with statistical parameters. The material specific expected temperature value is estimated by calculating the median of a temperature field. The variance is taken as a scatter parameter.

The visual data appear significantly less characteristic. This data is evaluated and classified here using a CNN. For the proposed solution, the results from the CNN are combined with the engineered statistical parameters.

3.1 Database

Aluminum, copper, brass, wood, wood imitations, ceramics, plastics, textiles, stainless steel, and paper were chosen for evaluation. These materials are broadly applied in industry and have certain similarities in texture and color. The images of the materials are presented in Fig. 1, showing the resemblance between brass (c), wood (d), wood-imitations (e) and cellulose (j) as well as between aluminum (a), stainless steel (i), and some of the textiles (h). The material samples were photographed from distances $d \in \{1, 1.5, 2, 2.5\}$ m with an angle $\alpha < 90\,^{\circ}$, as presented in Fig. 3. The samples were placed on cardboard to facilitate segmentation.

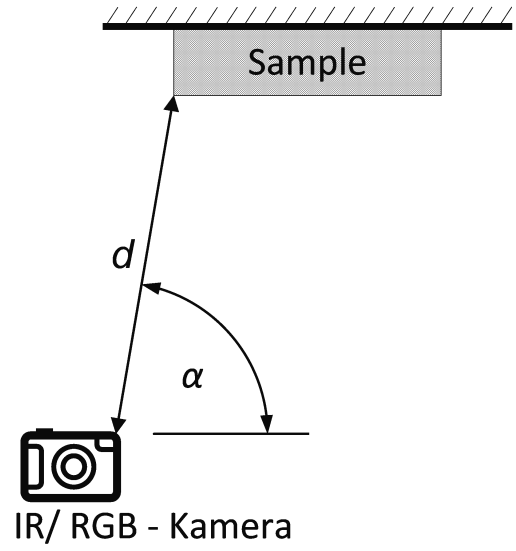

Fig. 3 Experimental design of collecting images from material samples: The IR and RGB images were taken with an angle $\alpha < 90\,^{\circ}$ and a distance d between 1 m and 2.5 m.

The pictures were taken with the FLIR T540 (Flir Systems Inc., Wilsonville, OR, US). This uncooled thermal camera has an integrated digital camera which enables taking IR and RGB images simultaneously. The database includes 1112 cropped RGB images with corresponding temperature distributions. The images were taken indoor in closed rooms avoiding thermal radiation reflections and ensuring the sample temperature corresponded with the ambient temperature. On this account, the material specific IR emissivity leads to material specific temperature distributions, when measured with the thermal camera.

Hyperparams	
Epochs:	20
Batch size:	10
Optimizer:	Adam Algorithm
Learning rate:	$1 * 10^{-4}$
Criterion:	Cross Entropy Loss
Split:	Stratified Tenfold
Random state:	42

Table 1 Table of hyperparameters used to implement the transfer learning.

3.2 Feature extraction with relearned CNN

Cropped images from the database were used to relearn a VGG 16 [29]. This CNN was pretrained on the Imagenet [30] database and loaded from the PyTorch Model Zoo. The VGG 16 architecture itself has proven suitable as a starting model for material recognition [3, 18, 10].

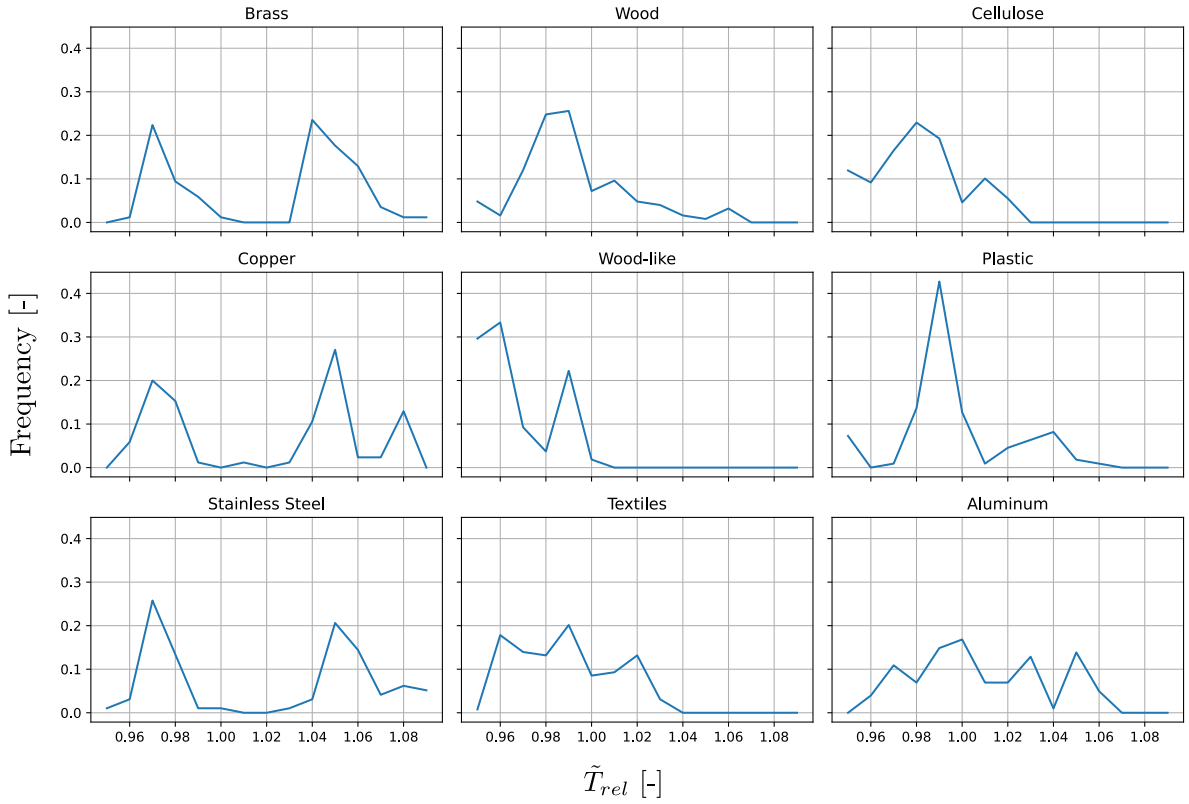

Fig. 4 Histograms of the material specific median relative temperature $\tilde{T}_{rel}$. As described in Alg. 1, the temperature fields of each material sample were put into perspective with the respective ambient temperature. The median was then calculated for each temperature field. The distribution across the database is presented in the histograms above.

For transfer learning, the prediction layer was replaced with an adapted fully connected layer. While the original VGG 16 architecture can classify 1000 different labels, only 10 are required here. The relearned CNN was then used to classify the RGB images from the database. The applied hyperparameters are presented in Tab. 1.

If the image $I_{\mathrm{RGB}} \in \mathbb{R}^{83\times83\times3}$ is the input and the Softmax probabilities $\boldsymbol{p} \in \mathbb{R}^{10}$ are the output of the classification model $CNN(\cdot)$, the above relation can be formally written as $\boldsymbol{p} = CNN(I_{\mathrm{RGB}})$. The Softmax probabilities are used here instead of the raw CNN features due to the much higher information density.

3.3 Sensor fusion with SVM

The temperatures of each temperature field T were relativized by dividing them with the ambient temperature ϑ_{a}. In order to consider IR data for material recognition, the median $Med(\cdot)$ and variance $Var(\cdot)$ of the relative sample temperatures were used as features and linked with the Softmax probabilities $\boldsymbol{p}$ of the VGG 16 as shown in Fig. 2. Thus, the feature vector $\boldsymbol{x}$ consists of probabilities per material class (VIS features) and the median as well as the variance of the relative sample temperatures (IR features), what is formally presented in Alg. 1.

Algorithm 1 Sensor fusion

Input: Temperature field $T \in \mathbb{R}^{30\times30}$, ambient temperature $\vartheta_{\mathrm{a}}[^{\circ}\mathrm{C}]$, Softmax probabilities $\boldsymbol{p}$
Model: Classifier SVM_{RBF}
Output: Material prediction $\hat{y}$

1: $T_{\mathrm{rel}} = \frac{T}{\vartheta_{\mathrm{a}}}$
2: $\tilde{T}_{\mathrm{rel}} = Med(T_{\mathrm{rel}})$
3: $T_{\mathrm{rel},\sigma^2} = Var(T_{\mathrm{rel}})$
4: $\boldsymbol{x} = concatenate(\boldsymbol{p}, \tilde{T}_{\mathrm{rel}}, T_{\mathrm{rel},\sigma^2})$
5: $\hat{y} = SVM_{\mathrm{RBF}}(\boldsymbol{x})$.
6: **return** $\hat{y}$

A SVM with Radial Basis Function (RBF) kernel and the one–versus-one approach for the multi-class classification $SVM_{\mathrm{RBF}}(\cdot)$ were chosen to classify the materials based on the combined feature vector. Based on the algorithm, SVMs find the optimal solution of a classification problem. In addition, the influence of IR features can be examined objectively.

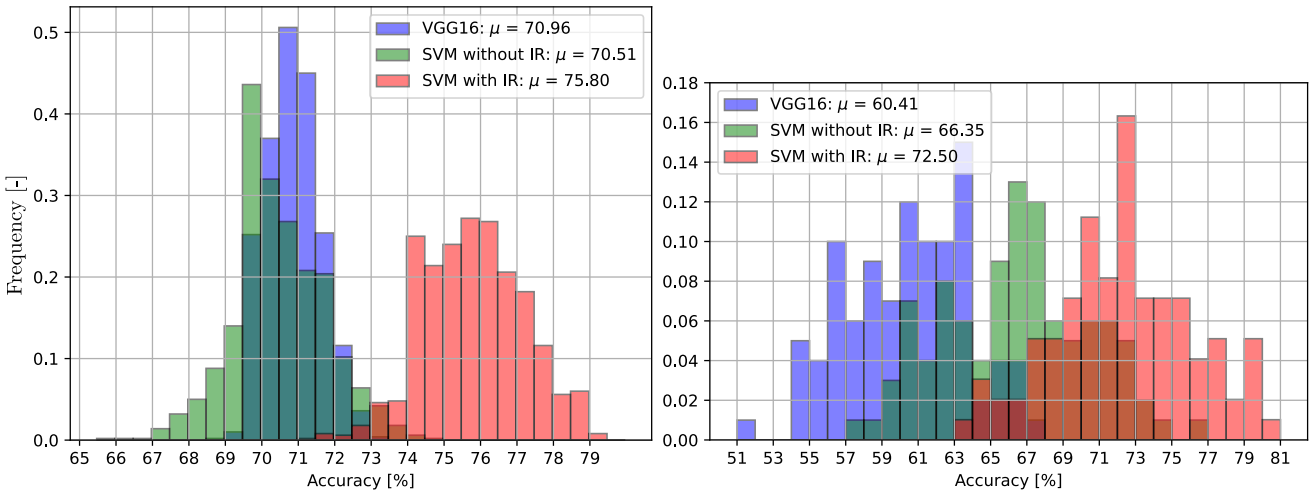

Fig. 5 Histograms of the convergence analysis: To evaluate whether the algorithm converges, one experiment was done 1,000 times with a constant train test split (left) and another with 10 times in a stratifold tenfold split (right).

4 Experiments and Results

Subsequently, the median relative temperature $\tilde{T}_{\text{rel}}$ was calculated and histograms were created, describing the frequency of the median values per class(see Fig. 4). These histograms show characteristic distributions for each material. So, in principle, the classification of materials must benefit from this additional IR data. Especially the metals shown in the first column have a very specific frequency distribution, which clearly differentiates them from non-metals and aluminum. Aluminum is covered by a characteristic oxide layer, which influences the emissivity.

4.1 Empirical Tests

For a convergence analysis, 80 % of the samples were randomly chosen for the training and 20 % for the validation. The training set was oversampled to balance the numbers of samples per class. The balanced dataset was used to relearn the pretrained VGG 16 1,000 times applying the hyperparameters shown in Tab. 1.

Finally, the SVM was fitted with the Softmax material probabilities from the CNN and the IR features. For image processing and the subsequent evaluation, Python 3.7 specifically the packages Scikit-Image [31], PyTorch [32] and Scikit-Learn [33] were used.

To investigate the influence of the IR features, an additional classification was done with dropped IR features as baseline. The resulting probability distribution of the accuracies is presented in Fig. 5, on the left side.

The VGG 16 (blue) clearly converges to the arithmetic mean $\mu = 70.96$. The mean of the SVM without IR features (green) is nearly the same as the VGG 16 while the SVM with IR features (red) is showing significant better results as the SVM with dropped IR features. A one-sided paired t-test confirms this with a p-value of one.

To further examine the algorithm on the database, a cross validation was done ten times in a second test with a stratifold tenfold split of the database. The probabilities of the accuracies are shown in Fig. 5, on the right side. Except of the split ratio, here nine to one, all other model settings stayed the same. Even when the distributions are not as separated as in the previous test, the improvement in accuracy of material recognition when using IR features is still significant, as another one-sided paired t-test confirms with a p-value of one.

4.2 Comparison to previous work

State of the art material recognition approaches use CNNs to classify visual images. The VGG 16 has proven itself and is taken here to extract the visual features. To enable the most meaningful comparison possible, all experiments apply the same classification algorithm (see sec. 3.2) with the same hyperparameters (see Tab. 1). The only difference is the proposed sensor fusion.

While there is no comparable approach in the field of material detection, optical sensor fusions are often performed in similar remote sensing applications. Two approaches are presented here for comparison. The first approach proposes a fusion of image features using a VGG 19 framework [25]. The merged image contains key features of both input images.

Since this study uses temperature arrays, a transformation is realized by normalizing and multiplying by 255. The result therefore corresponds to an 8-bit grayscale image.

Fig. 6 Confusion matrices from the cross validation test. The mean accuracies are shown when the feature vector is classified with dropped IR data (left) and with included IR data (right).

The second approach applies an image fusion [26]. They suggest to set the grayscale IR image as an additional color channel. The fused image is classified using an relearned VGG 16 with extended first layer kernels. The new initial weights are set as the mean RGB weights.

Here, the temperature array is rescaled and standardized to implement this data fusion approach. The mean results of a tenfold split are presented in Tab. 2.

Method	Accuracy [%]
Visual (VGG 16):	60.41
Feature Fusion:	50.52
Image Fusion:	36.89
Proposed Fusion:	**72.50**

Table 2 Comparison of different sensor fusion approaches.

4.3 Evaluation of Experiments

For cross validation, the material specific benefit of IR data as an additional feature is demonstrated in Fig. 7. The median improvement of the accuracy from aluminum, brass, and cellulose is about ten percent points (pp). Wood shows the best improvement of more than 15 pp. The median of the other materials is about zero, to be discussed next in Fig. 6.

It shows the arithmetic mean accuracy of each predicted vs. true label, comparing the classification without (left) and with (right) IR data. The SVM, using a feature vector with dropped thermal features, reaches an overall mean accuracy of 66.4 % in the sample recognition. The overall mean accuracy increases by 6.1 pp when the IR features are added.

In principle, regarding the mean accuracies per material, the combined model benefits from IR features. This is when VIS features are similar and IR features are different. The

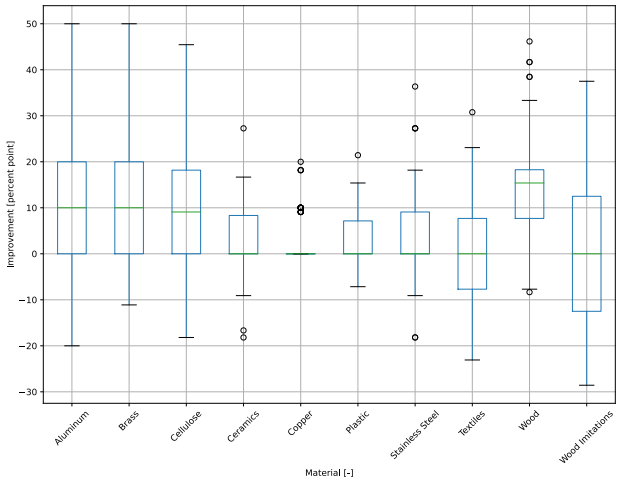

Fig. 7 Boxplots of the material specific improvement when using IR features for ten times in a tenfold cross validation.

combined model can distinguish, for example, much better between wood and wood imitations by using IR data. The mean accuracy of the class wood imitations increases just by 3 pp, because more samples are classified as plastic, being the actual material.

The accuracy of aluminum recognition increases by 21 pp because the miss classification as stainless steel or as textiles decreases. The IR data helps to distinguish between these two metals, which visually appear to be similar. As shown in Fig. 4, aluminums emissivity significantly differs from the emissivity of other metals because of its characteristic oxide layer. Many of the cropped textile images visually appear similar like stainless steel or aluminum, as Fig. 1 shows. About 20 % of the textile samples are false classified as these materials. By using IR data, the mean accuracy of textiles increases by 14 pp.

The recognition accuracy of brass increases by 12 pp, because the differentiation from wood-imitations, wood, and textiles is enhanced by IR data. On the other hand, 8 pp more brass samples were classified as aluminum although VIS- and IR features should be different. This could be an indication to train the CNN within further epochs.

Plastics and cellulose are the least recognized material classes with an accuracy of 33 % without IR data. Their recognition increases by 4 pp when IR data is considered. It seems like the CNN was not able to learn the characteristic visual properties. Additionally, 10 % of the plastics were classified as wood imitate, which actually is plastic.

Copper shows no improvement, see Fig. 7. That is because the accuracy with dropped IR features is already at 97 %. However, the mean accuracy can still be improved when IR features are added, as Fig. 6 shows.

IR data enhances differentiation between stainless steel and non-metals. The miss classification of cellulose and textiles decreased. But unlike above, adding IR features did not help differentiating between aluminum and stainless steel.

4.4 Discussion

The asymptotic behavior of the solution is examined as part of a convergence analysis and by a repeated cross validation. Based on these experiments, the material specific improvement is assessed using confusion matrices and boxplots.

In summary, IR features as an additional feature enhance differentiation between materials and boost recognition ability. The accuracy of the solution increases significantly when including IR data.

Materials which visually appear to be similar, such as wood and wood imitate, or aluminum and stainless steel were classified and could be differentiated more precisely.

This study uses a SVM to combine learned visual features from a CNN with engineered IR features from thermal imaging. To evaluate the results, the classification is done with and without these features. The aim of this study is not to reach the best possible classification results but to examine whether infrared data helps to increase the recognition accuracy.

Nevertheless the proposed method clearly outperforms previous approaches. It is remarkable that other fusion algorithms leads to poorer classification results than evaluating the data without fusion.

One possibility could be the homogeneity of the respective temperature fields. These appear more material specifically constant and without texture. Feature and image fusion approaches, on the other hand, try to combine characteristic textures.

5 Conclusion and Outlook

In the age of deep learning, the main challenge in material recognition is not feature engineering but data collection. Nevertheless, it is shown that engineered features based on a physical model can still help to improve the recognition accuracy.

When extending the evaluated electromagnetic spectrum to the IR range, a significant improvement of recognition is possible. With the engineered features, the overall mean accuracy increases by 6 pp. Also, the additional features help to classify materials which visually appear to be the same.

Therefore, over 1100 VIS and IR images were taken from ten material classes in controlled indoor environments. Even if the evaluated temperature distributions, based on the IR emissivity, do not lead to material specific fingerprints, they help differentiating between certain materials when used as an additional feature.

Based on the three research hypotheses, this article provides the following answers:

RH 1) Evaluating the IR range additionally to the VIS range has proven to be an effective option to significantly boost the reliability of material recognition for industrial processes.

RH 2) Therefore, materials which are broadly applied in industry and have certain similarities in texture and color are evaluated. Additionally, different lighting conditions and recording distances are taken into account.

RH 3) The IR features used are material-specific and increase classification accuracy. The proposed sensor fusion algorithm is realized with a SVM which has proven to be a suitable option for this material recognition application.

The results show significant improvements for material recognition. However, it is not examined whether these are the best possible accuracy results. Therefore, an extended training by comparing different pipelines is necessary.

In addition, the database should be extended with more material samples to further investigate the reliability of the results and to increase the generalization ability. While the data of metals seems to appear characteristic, the identification of cellulose and plastic in particular must be backed with additional training samples in order to obtain sound classification results.

Declarations

Funding No funds, grants, or other support was received.

Conflict of interest The authors have no financial or proprietary interests in any material discussed in this article.

Availability of data and material The database can be requested from the corresponding author.

Code availability The code can be requested from the corresponding author.

References

1. R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6):610 – 621, 1973. URL https://doi.org/10.1109/TSMC.1973.4309314.
2. K. Balakrishnan, R. Swathy, and T. D. Subha. Automatic waste segregator and monitoring system. *Journal of Microcontroller Engineering and Applications*, 3, 2016. URL www.researchgate.net/publication/317720527.
3. N. Ghassemi, H. Mahami, M. T. Darbandi, A. S., S. Hussain, F. Nasirzadeh, R. Alizadehsani, D. Nahavandi, A. Khosravi, and S. Nahavandi. Material recognition for automated progress monitoring using deep learning methods (preprint submitted). *Journal of Advanced Engineering Informatics*, 2020. URL https://arxiv.org/abs/2006.16344.
4. A. Ghosh, M. Ehrlich, Larry D., and Rama C. Unsupervised super-resolution of satellite imagery for high fidelity material label transfer. In *IGARSS 2019 - IEEE International Geoscience and Remote Sensing Symposium*, pages 5144–5147, 2019. URL https://doi.org/10.1109/IGARSS.2019.8900639.
5. Z. Erickson, S. Chernova, and C. Kemp. Semi-supervised haptic material recognition for robots using generative adversarial networks. In *Proceedings of the 1st Annual Conference on Robot Learning*, volume 78 of *Proceedings of Machine Learning Research*, pages 157–166. PMLR, 13–15 Nov 2017. URL https://proceedings.mlr.press/v78/erickson17a.html.
6. D. P. Penumuru, S. Muthuswamy, and P. Karumbu. Identification and classification of materials using machine vision and machine learning in the context of industry 4.0. volume 31, pages 1229–1241, 2020. URL https://doi.org/10.1007/s10845-019-01508-6.
7. V. S. Springthorpe and S. A. Sattar. Carrier Tests to Assess Microbicidal Activities of Chemical Disinfectants for Use on Medical Devices and Environmental Surfaces. *Journal of AOAC INTERNATIONAL*, 88 (1):182–201, 11 2019. URL https://doi.org/10.1093/jaoac/88.1.182.
8. H. Horn and B. Niemeyer. Aerosol disinfection of bacterial spores by peracetic acid on antibacterial surfaces and other technical materials. *American Journal of Infection Control*, 48(10):1200–1203, 2020. URL https://doi.org/10.1016/j.ajic.2020.01.019.
9. L. Sharan, R. Rosenholtz, and E. H. Adelson. Accuracy and speed of material categorization in real-world images. *Journal of Vision*, 14(10), 2014. URL https://people.csail.mit.edu/lavanya/fmd.html.
10. S. Bell, P. Upchurch, N. Snavely, and K. Bala. Material recognition in the wild with the materials in context database. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3479 – 3487, 2015. URL https://doi.org/10.1109/CVPR.2015.7298970.
11. J. Xue, H. Zhang, K. Nishino, and K. J. Dana. Differential viewpoints for ground terrain material recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 44(3):1205–1218, 2022. doi: https://doi.org/10.1109/TPAMI.2020.3025121.
12. F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He. A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1):43 – 76, 2021. URL https://doi.org/10.1109/JPROC.2020.3004555.
13. Y. Cho, N. Bianchi-Berthouze, N. Marquardt, and S. J. Julier. Deep thermal imaging. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018. URL https://doi.org/10.1145/3173574.3173576.
14. E. H. Adelson. On seeing stuff: the perception of materials by humans and machines. In Bernice E. Rogowitz and Thrasyvoulos N. Pappas, editors, *Human Vision and Electronic Imaging VI*, volume 4299, pages 1 – 12. International Society for Optics and Photonics, SPIE, 2001. URL https://doi.org/10.1117/12.429489.
15. C. Liu, L. Sharan, E. H. Adelson, and R. Rosenholtz. Exploring features in a bayesian framework

for material recognition. In *IEEE, 2010 Computer Vision and Pattern Recognition (CVPR)*, pages 239–246, 2010. URL http://dx.doi.org/10.1109/CVPR.2010.5540207.

16. I. Badami. Material recognition: Bayesian inference or svms? In *The 16th Central European Seminar on Computer Graphics*, CESCG 2012, 2012. URL https://old.cescg.org/CESCG-2012/papers/Badami-Material_Recognition_Bayesian_Inference_or_SVMs.pdf.
17. L. Sharan, C. Liu, R. Rosenholtz, and E. Adelson. Recognizing materials using perceptually inspired features. *International journal of computer vision*, 103: 348 – 371, 2013. URL https://doi.org/10.1007/s11263-013-0609-0.
18. P. Bian, W. Li, Y. Jin, and R. Zhi. Ensemble feature learning for material recognition with convolutional neural networks. *International Journal of Computer Vision*, 64:348 – 371, 2018. URL https://doi.org/10.1186/s13640-018-0300-z.
19. K. S. Younis, W. Ayyad, and A. Al-Ajlony. Embedded system implementation for material recognition using deep learning. In *2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT)*, pages 1–6, 2017. URL https://doi.org/10.1109/AEECT.2017.8257769.
20. H. Zhang, Z. Jiang, Q. Xiong, J. Wu, T. Yuan, G. Li, Y. Huang, and D. Ji. Gathering effective information for real-time material recognition. *IEEE Access*, 8:159511–159529, 2020. doi: https://doi.org/10.1109/ACCESS.2020.3020382.
21. S. J., S. T, G. K., and I. R. A study on ensemble feature learning for material recognition. In *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, pages 1184–1187, 2021. doi: https://doi.org/10.1109/ICCMC51019.2021.9418399.
22. P. Saponaro, S. Sorensen, A. Kolagunda, and C. Kambhamettu. Material classification with thermal imagery. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015. URL https://doi.org/10.1109/CVPR.2015.7299096.
23. T. Bhattacharjee, H. M. Clever, J. Wade, and C. C. Kemp. Material recognition via heat transfer given ambiguous initial conditions. *IEEE Trans. Haptics*, 14(4): 885–896, 2021. doi: https://doi.org/10.1109/TOH.2021.3089990.
24. K. Gadzicki, R. Khamsehashari, and C. Zetzsche. Early vs late fusion in multimodal convolutional neural networks. In *2020 IEEE 23rd International Conference on Information Fusion (FUSION)*, pages 1–6, 2020. doi: https://doi.org/10.23919/FUSION45008.2020.9190246.
25. H. Li, X. Wu, and J. Kittler. Infrared and visible image fusion using a deep learning framework. In *2018 24th International Conference on Pattern Recognition (ICPR)*, pages 2705–2710, 2018. doi: https://doi.org/10.1109/ICPR.2018.8546006.
26. Z. Liu, J. Wu, L. Fu, Y. Majeed, Y. Feng, R. Li, and Y. Cui. Improved kiwifruit detection using pre-trained vgg16 with rgb and nir information fusion. *IEEE Access*, 8:2327–2336, 2020. doi: https://doi.org/10.1109/ACCESS.2019.2962513.
27. A. Vakil, J. Liu, P. Zulch, E. Blasch, R. Ewing, and J. Li. A survey of multimodal sensor fusion for passive rf and eo information integration. *IEEE Aerospace and Electronic Systems Magazine*, 36(7):44–61, 2021. URL https://doi.org/10.1109/MAES.2020.3006410.
28. A. Kurup, S. Kysar, and J. Bos. Svm-based sensor fusion for improved terrain classification. In Michael C. Dudzik and Stephen M. Jameson, editors, *Autonomous Systems: Sensors, Processing, and Security for Vehicles and Infrastructure 2020*, volume 11415, pages 121 – 128. International Society for Optics and Photonics, SPIE, 2020. URL https://doi.org/10.1117/12.2558960.
29. K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In *3rd International Conference on Learning Representations*, ICLR 2015, pages 1 – 14, 2015. URL https://arxiv.org/abs/1409.1556v4.
30. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3):211–252, 2015. URL https://doi.org/10.1007/s11263-015-0816-y.
31. S. van der Walt, J. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. Warner, N. Yager, E. Gouillart, and T. Yu. scikit-image: Image processing in python. *PeerJ*, 2, 2014. URL https://doi.org/10.7717/peerj.453.
32. A. Paszke, S. Gross, S. Chintala, G. Chananand E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer. Automatic differentiation in pytorch. In *31st Conference on Neural Information Processing System*, 2017. URL https://openreview.net/forum?id=BJJsrmfCZ.
33. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.