

# Deep Thermal Imaging: Proximate Material Type Recognition in the Wild through Deep Learning of Spatial Surface Temperature Patterns

Youngjun Cho

Nadia Bianchi-Berthouze

Nicolai Marquardt

Simon J. Julier

University College London, London, United Kingdom

{youngjun.cho.15, nadia.berthouze, n.marquardt, s.julier}@ucl.ac.uk

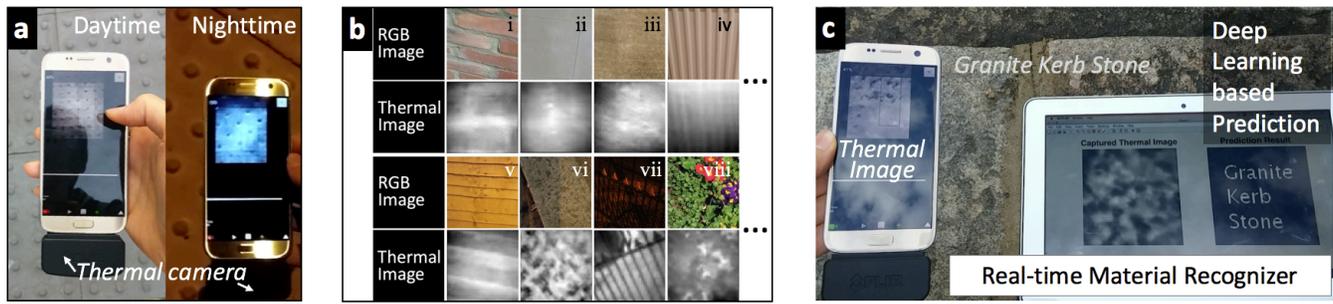


Figure 1. Deep Thermal Imaging for in-the-wild material recognition: (a) spatial temperature patterns of material during day/nighttime, (b) examples of RGB and quantized thermal images of materials (i: brick; ii: marble; iii: carpet; iv: radiator; v: wood fence; vi: granite; vii: iron fence; viii: flower bed), (c) the real-time material recognizer.

## ABSTRACT

We introduce *Deep Thermal Imaging*, a new approach for close-range automatic recognition of materials to enhance the understanding of people and ubiquitous technologies of their proximal environment. Our approach uses a low-cost mobile thermal camera integrated into a smartphone to capture thermal textures. A deep neural network classifies these textures into material types. This approach works effectively without the need for ambient light sources or direct contact with materials. Furthermore, the use of a deep learning network removes the need to handcraft the set of features for different materials. We evaluated the performance of the system by training it to recognize 32 material types in both indoor and outdoor environments. Our approach produced recognition accuracies above 98% in 14,860 images of 15 indoor materials and above 89% in 26,584 images of 17 outdoor materials. We conclude by discussing its potentials for real-time use in HCI applications and future directions.

## Author Keywords

Material recognition; in the wild; deep learning; sensing; context-aware mobile computing; thermal imaging.

## ACM Classification Keywords

H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces - Input devices and strategies

## INTRODUCTION

Context awareness, including material type awareness, is of critical importance in many situations for both humans and ubiquitous technologies [13,15,49]. Methods that contribute to build context awareness can support decision making, increase safety and facilitate interaction with the environment. A compelling use case is Microsoft's *Seeing AI* [58] for visually impaired people. Using a smartphone camera and microphone, *Seeing AI* recognizes the environment and provides an audio description of the immediate context (e.g., the scene and people nearby). While such systems are useful, they do not provide all information the users may need. For example, in the case of physical manipulation, it is not sufficient to know what the object is, but also what material it is made of.

Creating the ability to recognize materials is the focus of our paper. A variety of sensing technology have been investigated in HCI to support material awareness. SpecTrans [48], for instance, is an image sensor-based material recognizer that analyzes reflected light patterns produced by multi-spectra lighting sources. Among its use cases, the authors show how it could support embedded marker applications (e.g., encoded reality [59]) without the need to physically modify the objects. Another example is RadarCat [55], a contact-based recognizer which uses a lightweight radar system to develop a "material dictionary", allowing users to search for the name of the material an object is made of by simply touching it with the recognizer. Whilst these applications enable new ways to understand and

interact with materials, they require controlled illumination and direct contact with those materials, which is not always possible or desirable in real life situations.

To overcome the above limitations, we propose a new proximate (10cm-50cm range) material recognition system which uses a smartphone-based mobile thermal camera and a deep learning technique to classify material types. This approach, illustrated in Figure 1, requires no physical contact between the sensor and materials and is robust to environmental light and temperature changes. In addition, by using deep learning, we remove the need to hand-engineer features for recognition. Such techniques make it easier to extend the recognition capabilities to new types of materials by simply adding sufficient new training data [18,57]. This is critical for real-world applications where the materials to be recognized are not always all known in advance.

**Contributions** - This paper makes four main contributions:

- *Deep Thermal Imaging*: a novel material recognizer system based on a low-cost, mobile thermal camera which can be integrated into a smartphone and is robust to changes in illumination and temperature conditions.
- Deep learning-based software tools for feature learning on spatial thermal textures of materials, enabling easy extension to a wider set of materials. The implementation of our recognition engine is open-sourced.
- In-the-wild studies with 32 types of indoor and outdoor materials (for a total of 41,444 thermal images) to demonstrate the capability and robustness of Deep Thermal Imaging to discriminate between materials under different (i) illumination settings, (ii) environment temperatures, (iii) point of views and proximate distances.
- Discussion of practical HCI use cases benefitting from *Deep Thermal Imaging* material recognition capabilities.

## RELATED WORK

In this section, we discuss prior work on *proximity sensing and its use in context-aware HCI applications*, review technical methods for *material classification* and synthesize studies about *thermal imaging*.

### Proximity Sensing in HCI Applications

Proximity is defined as “physical distance” in social contexts [12]. The term *proximity sensing* is generally used to indicate the ability to detect the presence of close-range objects [6]. Various types of sensors have been used for proximity sensing, including ultrasonic sensors [33], capacitive sensors [54], laser range finders [42] and photo-detectors [21]. Proximity sensing is useful in many HCI situations. For example, Eric and Jochen [16] explored how detecting the presence of nearby objects can help construction workers avoid potential hazards. Boyd et al. [3] proposed a necklace-type wearable prototype using photo-detector to help people with autism gain awareness of their surroundings. Computer agents have also exploited proximity sensing to interact with people, e.g., to understand people’s in-air gesture and

hovering over a mouse [22] or a display screen [9,60]. Whilst this technology has been used to increase context awareness, its focus has been mainly on a) simply signaling that an object is present or b) detecting the presence of well pre-defined object (often human skin). Wiese et al. [54] have started to enhance proximity sensing capability by combining multiple sensing modalities to sense the type of material a mobile phone is placed on. However, their range of materials is limited to carry out predefined tasks such as whether the phone is in a pocket.

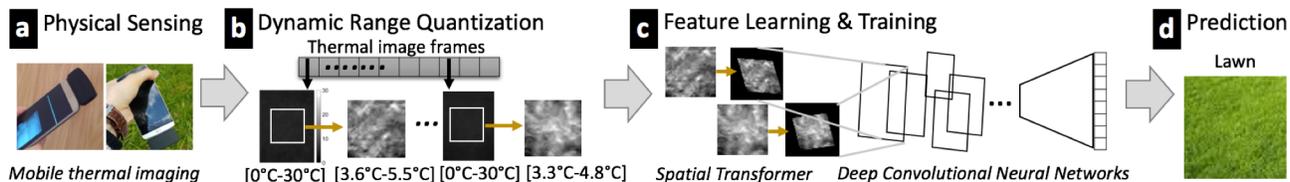
### Material Classification using Sensing Devices

The ability to automatically recognize materials has been recently explored by HCI and computer vision researchers. The aims have been to enable novel ubiquitous interaction and richer context awareness [13,15,49]. For instance, Harrison and Hudson [20] have prototyped a photoresistor based material recognizer. Sato et al. [48] have proposed a real-time transparent and specular material type recognizer using a mouse image sensor and light emitters. Yeo et al. [55] have implemented a low-cost radar system-based material recognizer, and have used a smartphone front-camera and screen flash light [56] for better scalability. Researchers have also explored how friction and acceleration signals can be used to understand or recognize materials [7,31,35].

Whilst these approaches support material awareness, they have some limitations. The most important is that many methods require the sensors to be in contact with the material or very close to it (with 10cm of). In many situations, this is not feasible (e.g., hazardous, or is not socially acceptable).

Visible spectrum-based imaging approaches (conventional RGB) [2,14,34,53], on the other hand, can provide touchless ways to discriminate between different materials. However, they are susceptible to *illumination effects*: for instance, extremes of darkness and brightness incapacitate these algorithms due to sensor saturation or sensitivity [1,34].

This limitation can be addressed through the use of other imaging modalities that read electromagnetic radiation outside the visible range (e.g., radar imaging [5], thermal imaging [36,52]). In fact, the human eyes and RGB cameras that mimic the eyes are only sensitive to a narrow range of electromagnetic spectrum called “visible light”. They capture objects through the radiations in the visible spectrum their surface emits or reflects when hit by a source of light [45]. This means that without sources emitting visible light, such as the sun or incandescent bulbs, this type of vision system is generally unable to “see” the world. However, at other wavelengths, objects are able to illuminate themselves. According to *Planck’s law*, every object above absolute zero emits thermal radiation [39,52]. Most of this radiation falls in the infrared spectrum range (between 0.9 and 14  $\mu\text{m}$ ) rather than the visible spectrum range (between 380 and 780 nm). It can be captured by thermal cameras but not by RGB cameras. Therefore, we propose to use thermal imaging as an inexpensive, contactless sensing modality to recognize



**Figure 2. The overall process for proximate material type recognition with Deep Thermal Imaging: a) physical sensing via mobile thermal imaging, b) dynamic range quantization on every thermal image frame, c) feature learning and training process based on spatial transformer and deep convolutional neural networks, d) prediction results.**

material types independently of the level of illumination (e.g., working in dark environment).

### Thermal Imaging & its HCI Applications

Thermal imaging is a key method for measuring spatial temperature of materials, objects and scenes [36]. It has been used in applications that range from inflammation diagnosis [41] to vehicle occupant detection [39], and to breathing and stress monitoring [8,10]. Thermal imaging has also been investigated for HCI applications. Larson et al. [32], for instance, have proposed a thermal imaging-based surface gesture recognition system which uses residual heat traces on a surface. Kurz [30] extended this concept to support touch detection in 3D augmented space. Shirazi et al. [47] have shown how to use thermal reflection characteristics of materials for gestural interaction. However, these approaches have had limited use in the wild because they used high cost, heavy thermal cameras and were sensitive to environmental temperature changes.

Recently, a new group of low-cost, lightweight (<35gram) and mobile thermal cameras have been developed (see Table 1). They have the potential to deploy and power thermal imaging based HCI applications (a smartphone embedding a thermal camera has been already launched [61]). Indeed,

Product	Spatial Resolution	Frame Rate	Thermal Sensitivity (NETD)	Price (in July 2017)
<sup>+</sup> FLIR One 2G	160x120	<8.7Hz (unsteady)	<0.10°C	Not applicable
<sup>+</sup> FLIR One 3G	80x60	<8.7Hz (unsteady)	0.15°C	\$199.99
<sup>+</sup> FLIR One Pro	160x120	<8.7Hz (unsteady)	0.15°C	\$399.99
<sup>+</sup> Seek Thermal Compact	206x156	<9Hz (unsteady)	0.50°C	\$249.00
<sup>+</sup> Seek Thermal Compact PRO	320x240	>15Hz (unsteady)	Unknown	\$499.00
<sup>+</sup> Therm-App Hz	384x288	25Hz	<0.07°C	\$999.00
*FLIR A35	320x256	60Hz	<0.05°C	\$5,900.00
*OPTRIS PI-640	640x480	32Hz	0.075°C	\$11,310.04
*FLIR A325sc	320x240	60Hz	<0.05°C	\$12,322.46
*FLIR SC5000mb	640x512	100Hz	0.017°C	Over \$13,000.00

**Table 1. Specification and price of commercial thermal cameras: <sup>+</sup>Low-cost, Mobile Thermal Cameras (< 34x14x67mm<sup>3</sup> Dimension), \*High-end thermal imaging systems (> 90x46x56 mm<sup>3</sup> dimension).**

Naik et al. [37] have shown that, in controlled environments, a low-cost thermal camera has sufficient resolution and accuracy to grade the maturity of a fruit. Furthermore, Cho et. al [10] have recently investigated how to address the effect of high temperature changes on low cost, low resolution thermal cameras. They proposed an optimal quantization method to compensate for high thermal dynamic range scenes (e.g., outdoor environments with highly varying temperature). The method, together with novel machine learning techniques, may enable the use of low cost thermal camera in more ambitious and unconstrained material recognition applications.

In the next section, we introduce our new mobile thermal imaging-based proximate material type recognizer that can work in the wild in a wide range of environmental illumination, temperature, and view point.

### MATERIAL RECOGNITION STRATEGY

Deep Thermal Imaging is a novel approach for sensing and recognizing the material of proximate objects. Rather than focusing on the overall material temperature, we propose to use thermal cameras to capture the characteristic spatial thermal pattern of the compounds or elements forming a material [4]. Figure 2 illustrates our four-stage approach: a) physical sensing of proximate surfaces through mobile thermal imaging; b) dynamic range quantization to decrease environmental temperature effects and enhance thermal texture; c) deep-learning-based feature learning and training; and d) real-time recognition of materials.

#### Physical Sensing: Mobile Thermal Imaging

In our study, we used a mobile, low-cost thermal camera (FLIR One 2G, see Table 1) which converts 8-14μm radiation spectral density to temperature. When attached to a smartphone, it can be used to continuously sense and record thermal image sequences of different types of surface materials. Its portability means that it is easy to gather images of the same object from many different perspectives (e.g., shape, geometry, illumination) and distances.

#### Dynamic Range Quantization

Various factors affect the thermal texture of materials captured by the camera. Each material's compound combinations, surface structure (e.g., roughness) and geometry (e.g., cavities) influence both the amount of radiation emitted from the surface and the emissivity [52]. Given this, thermal imaging detects the thermal radiation and *estimates* the temperature by using an emissivity value, so-

called *apparent temperature* (not *true temperature*). Instead of attempting to measure the true temperature of the material, we aim to capture the thermal radiation differences over material surfaces to discriminate them.

To maximize the differences between materials and compensate for environmental temperature changes, we introduce the *Dynamic Range Quantization* method, simplified from the *Optimal Quantization* method proposed in [10]. Quantization, the process of mapping input data to output data within a specific range, is crucial for handling thermal image sequences in the wild [10]. Our method amplifies the *relative* difference between neighboring temperature elements by taking into account the different amount of radiation on every single thermal image of a material surface (Figure 2b).

First, it crops an  $N \times N$  square at the center of the captured image. In our implementation, we use a  $75 \times 75$  square given the low spatial resolution ( $160 \times 120$ ) of the thermal camera. Cropping is important because low-cost systems, designed for mobile thermal imaging, suffer from lens-inducing calibration errors that can lead to extremely low or high temperature at the edge of the image [10,61]. The second step is to adaptively quantize the thermal measurements using the expression  $u = \Gamma(T)$ , where  $[T_1, T_2]$  is the range of temperature and  $[u_1, u_2]$  is the range of a pixel value on the output thermal image. The temperature range is given by

$$T_1(k) = \min_{\forall i,j \in ROI} S_{ij}(k), \quad T_2(k) = \max_{\forall i,j \in ROI} S_{ij}(k) \quad (1)$$

where  $S_{ij}(k)$  is the element  $(i, j)$  of the thermal matrix  $S$  with frame number  $k$ . The final quantized value  $I_{ij}(k)$  on the processed thermal image is:

$$I_{ij}(k) = (u_2 - u_1) \cdot \left( \frac{S_{ij}(k) - T_1(k)}{T_2(k) - T_1(k)} \right). \quad (2)$$

This conversion has two effects. First, it produces an output image with more distinct spatial thermal patterns of material (Figure 2b), which help to discriminate between the different types. Second, these patterns are more invariant to changes in ambient temperature.

### Feature Learning Process through Deep CNN

Given the parallel to image processing with RGB images, our feature learning process uses the deep Convolutional Neural Network (CNN) illustrated in Figure 2c. Many CNNs have been proposed for RGB-vision analysis (e.g., [11,26,40,53]). In this paper, we use the Spatial Transformer-based CNN [26]. This architecture gets its name from the fact that the Spatial Transformer artificially manipulates spatial data information within CNNs. This helps to handle high-level variances of spatial thermal image patterns due to different shooting perspectives and geometries. This technique also performs well on lower resolution images, which is ideal for our thermal image. For more details on the implementation,

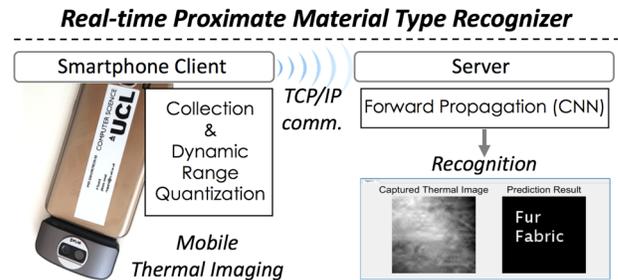


Figure 3. Implementation of Deep Thermal Imaging, the real-time proximate material type recognizer.

see [26]. Once the implemented deep learning network has been trained, it can be used to automatically classify material types (Figure 2d). Detailed parameters about the CNN layers are introduced in each study section.

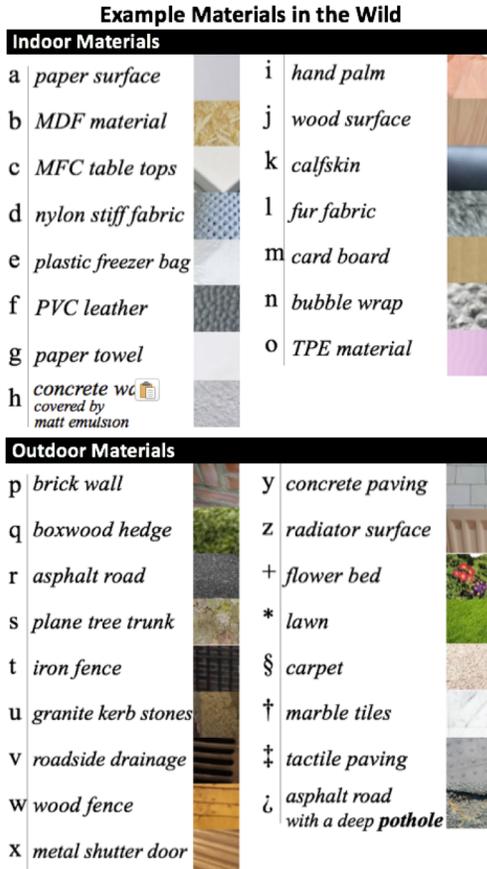
### Implementation of Real-Time Proximate Material Type Recognizer

Figure 3 illustrates the architecture of real-time material recognizer. The smartphone client application (i) controls the mobile thermal imaging sensor (FLIR One and its SDK) to collect a temperature matrix of material, (ii) quantizes the matrix, and then (iii) wirelessly sends it to a server using the TCP/IP protocol. The received data is fed to the trained deep CNN for classification. This transformer network was implemented using the *MatConvNet* framework [51]. Alternatives based on embedded platforms with a GPU could be also used, such as *NVIDIA Jetson TK1* [62] with *Caffe* [63]. A smartphone could also be used for classification by using *TensorFlow Mobile* [64].

Finally, the easy-to-use API for the quantization and the deep learning process and example codes for the raw thermal data collection from the thermal camera are made available at <http://youngjuncho.com/2018/DeepThermalImagingAPI>.

### MATERIALS IN THE WILD: OUR DATASETS

There are many possible ways to categorize materials; in our study, we grouped them into *indoor materials* and *outdoor materials*. Figure 4 illustrates the example classes of material surfaces we have targeted. Indoor materials are encountered in the context of a house or an office where the environment is typically more stable and controlled. The selection of indoor materials was mainly based on recent works [48,56], but was extended to include new classes to understand the limits of the approach (e.g., bubble wrap and paper towel given similarity with freezer polyethylene bag and paper, respectively). Outdoor materials are found in conditions which are very variable and are exposed to dirt and weathering effects. The choice of materials was linked to one of our use cases on wheelchair users discussed later. Caretakers were informally interviewed, and prior studies on wheelchair navigation were used (e.g., [27]). Materials which frequently appear on the ground (e.g., granite) or obstacles (e.g., fences) were considered important to facilitate navigation. Such materials are challenging as they are less likely to have constant patterns of spatial temperature, texture and geometry. Two materials were also



**Figure 4. Example materials collected in the wild: (a) materials generally found in indoor places (less influenced by environmental thermal dynamics), (b) outdoor materials which have richer surface patterns as they are more likely to be affected by surrounding environment temperature, weather and dirt.**

added to the outdoor collection - *radiator in building corridor (denoted as z)* and *building entrance carpet (denoted §)* even though they are physically found indoors. The reason is that they tend to be more variable – their surfaces may be exposed to high thermal dynamics and dirt.

To evaluate the performances of *Deep Thermal Imaging*, we conducted two studies *in the wild* [43] outside of controlled lab settings (e.g., temperature, light and dirt control). The recording of images was carried out by sampling while walking or running around an existing environment and taking thermal videos of material around from different perspective and repeating such recording in various places, at different times of the day, temperature and weather conditions. This process was facilitated by the low and unsteady frame rate of the thermal camera (3-8 thermal images every second), and ensured a collection of non-similar images of the same materials to maximize the generalization capability of the trained recognition model.

### STUDY I: INDOOR MATERIALS

In this section, we report the experimental protocol, our collected thermal image dataset and the recognition results

for indoor materials.

### Experimental Protocol & Dataset (DeepTherm I)

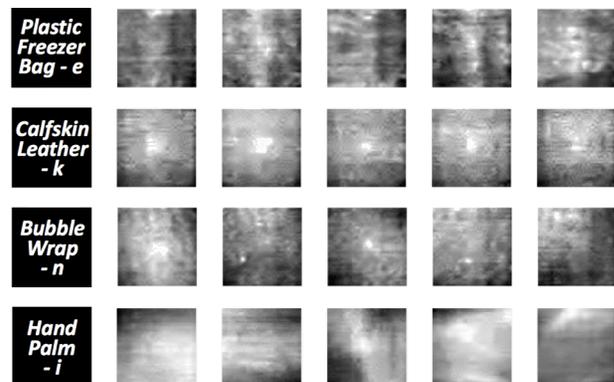
For Study I, the 15 indoor materials shown in Figure 4 were used to create the dataset *DeepTherm I* which consists of 14,860 processed thermal images (average count of data for each individual class: 990.7, SD=425.9; 400-600 images of each material per each variable). The dataset was created by recording thermal image sequences in a room with different lighting levels (bright / dark), with/without air-conditioning, different places (on a floor or a desk) and from different perspectives (Figure 5). The spatial temperature patterns were collected from different angles and different distances (between 10 and 50 cm, from the camera lens to the material). The data was collected five times in about 3 weeks. For *i. hand palm* (Figure 4), we collected thermal images of both hand palms from 4 people. Figure 6 shows some of example thermal images collected and quantized for this study. The dataset was divided into a training set and a test set along with a cross-validation method (Set size: 80%-20% in 5-fold; 90%-10% in 10-fold).

### Implemented CNN and Training

The CNN architecture for Study I used the architecture in [26] and consisted of one average pooling layer, two convolutional layers, two max pooling layers and two fully connected layers. We used a Rectified Linear Unit (ReLU) as an activation function to speed up the feature learning



**Figure 5. Data collection for Study I: thermal images of indoor materials were recorded under different conditions (lighting, place, air-conditioning, direction and distance).**

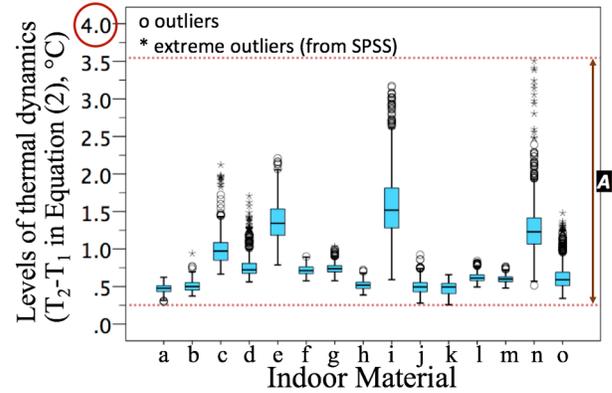


**Figure 6. Examples of collected and quantized thermal images of e. plastic freezer bag, k. calfskin leather, n. bubble wrap and i. hand palm.**

process [38]. Quantized images of size 75x75 were resized to 60x60 to be feed-forwarded into the first layer to use the layer composition in [26]. In the training process, a learning rate of 0.001 and a batch size of 256 were used for 350 epochs. The CNN was implemented on the *MatConvNet* framework and ran on a 64-bit Windows desktop (Intel Core i7-3770 CPU @ 3.40GHz 20GB RAM).

**Results**

Figure 7 shows thermal dynamic levels of the collected thermal images of the 15 indoor materials. The thermal dynamics were examined by using the temperature range ( $T_2-T_1$ ) in Equation (1). Some materials, such as *i. hand palms* and *n. bubble wrap*, showed higher thermal dynamics than materials that have more regular surface patterns including *a. paper surface* and *b. MDF materials*. The thermal dynamics ranged between 0.259°C and 3.507°C (denoted as *A* in Figure 7) with a mean of 0.889°C (SD=0.4608). We evaluated the performance of the proposed approach on the *DeepTherm I* dataset. We tested using two types of cross-validation (5-fold and 10-fold) within the dataset. The reason was to understand the effects of training with smaller datasets. For the creation of random partitions (i.e., k-fold), the MATLAB function *cvpartition* was used. As results, the CNN classifier achieved a mean accuracy of 97.899% (SD=2.105) and 98.076% (SD=2.075) for the 5-fold and the 10-fold validations respectively. Figure 8 summarizes the mean class accuracy results from 10-fold cross-validation. As it can be seen here, performances are very high for each material type. The worst performing materials were materials *e. plastic freezer bag*, *n. bubble wrap* and *o. TPE material*. Even in these cases, the materials were correctly classified in more than 90% of the cases. One explanation possible for these slightly lower performances is that these three types of material are transparent, shiny and reflective and hence their thermal radiation patterns are likely to be more influenced by surroundings [52].



**Figure 7. Levels of thermal dynamics of the total 14860 thermal images across the 15 indoor materials (a-o). The thermal dynamics were estimated by using ( $T_2-T_1$ ) in Equation (2). [A] indicates the range of those values.**

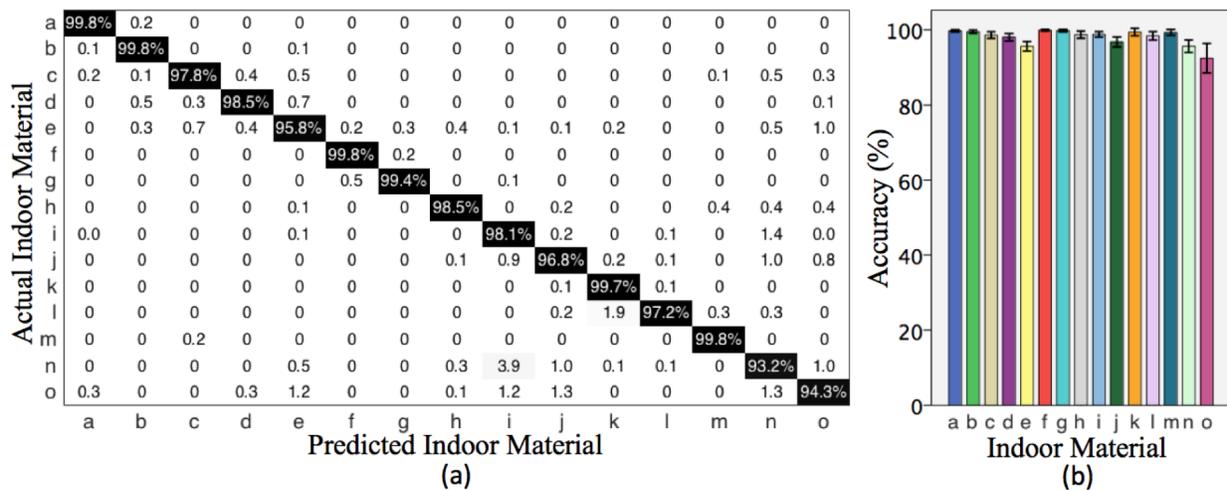
**STUDY II: OUTDOOR MATERIALS**

This study consists of two parts. The first part follows the design of Study I to investigate the possibility of discriminating between different types of *outdoor* material. The second part investigated the ability of the already trained model to discriminate between the same type of outdoor materials but gathered at real-time and under completely different conditions from the data in the training set (i.e., similar to the *zero-shot learning* [44] testing concept).

**PART I: Training and Testing within the Dataset**

**Experimental Protocol & Dataset (DeepTherm II)**

The 17 outdoor materials (illustrated in Figure 4) were targeted. The data collection process produced the *DeepTherm II* dataset which includes 26,584 labelled thermal images. The average number of collected spatial thermal patterns from each material was 1563.8 (SD=295.3; about 300-500 images of each material per each condition). The accompanying preview video shows examples of data.



**Figure 8. 10-fold cross validation results from Study I (low dynamic proximate objects, labels for materials shown in Figure 4): (a) confusion matrix (in percentage; color range: white (0%) - black (100%)), (b) mean class accuracy distribution (error bar: 95% confidence interval).**

Data Collection Study II: Outdoor Materials	Daytime		Nighttime		
	In April	300-500 thermal image frames per each material		300-500 thermal image frames per each material	
	In August	300-500 thermal image frames per each material		300-500 thermal image frames per each material	

Figure 9. Data collection for Study II: thermal images of outdoor materials were recorded by two experimenters while walking/running outside both at daytime and nighttime in different seasons.

As illustrated in Figure 9, two experimenters collected thermal images in two different months (April and August) both during the day and at night. As in the protocol of Study I, experimenters collected spatial material temperature patterns while moving or running slowly around material during shooting to exploit the low frame rate of the thermal camera and recorded materials from different perspectives. As for Study I, imaging materials from different perspectives, distances, times of days, and seasons was critical to evaluate the generalization capability of the modelling approach. However, we avoided rainy days and materials which were significantly covered by clutter such as sand or big trash bags. Figure 10 shows examples of thermal images collected and quantized for this study. A higher number of recordings were collected for materials – such as iron fences – with high degree of variability due to thermal reflection [52].

### Implemented CNN and Training

A pilot test showed that the CNN architecture used for the first study performed worse on the *DeepTherm II* dataset (less than 70%). A possible reason is that the trained neural network overfitted the training set given noise in the recordings of the outdoor materials. An overfit model is more likely to produce a lower recognition accuracy on new unseen data [50]. Hence, we added “dropout” to the end of the neural network layers in the CNN, which is a simple but powerful way to reduce overfitting issues [29]. The architecture was trained and tested on the *DeepTherm II* dataset (see parameters in Table 2).

### Results

Figure 11 shows the thermal dynamic ranges of the collected images of the 17 outdoor materials. The temperatures ranged between 0.341°C and 28.744°C with a mean 2.959°C (SD=1.963). Using 5-fold and 10-fold cross validation methods, the implemented CNN classifier (Table 2) achieved a mean accuracy of 84.526% (SD=9.875) and 89.356% (SD=6.833), respectively. Figure 12 summarizes the mean class accuracy results from 10-fold cross-validation. Surprisingly, *z* (*radiator in building corridor*) and *§*

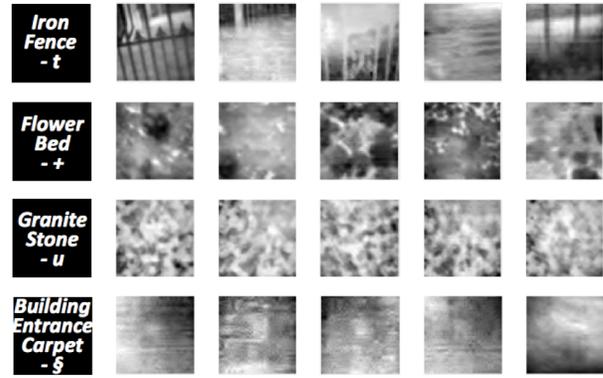


Figure 10. Examples of collected and quantized thermal images of *t.* iron fence, *+* flower bed, *u.* granite stones (pavement edge) and *§.* building entrance carpet. These images contain some degree noise, e.g., dirt and rubbish.

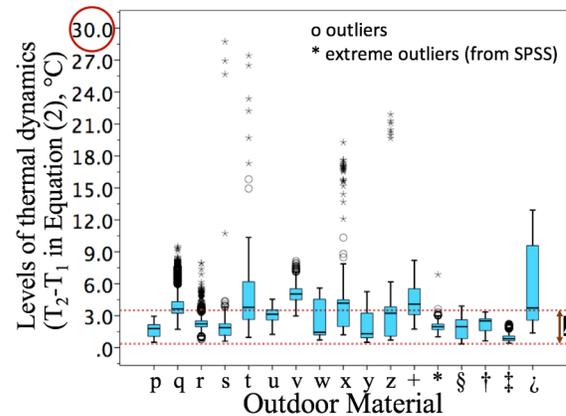


Figure 11. Thermal dynamics levels for the total 26584 thermal images across all 17 outdoor materials (p-*z*). The thermal dynamics were estimated by using  $(T_2-T_1)$  in Equation (2). The thermal dynamic range of indoor materials are marked in red dot lines (*A*) for comparison.

(*building entrance carpet*), which are on the border between outdoor and indoor areas, showed highest accuracies (97.9% and 98.1%, respectively). On the other hand, *+* (*flower bed*), *\** (*lawn*) and *ζ* (*asphalt road with a deep pothole*) showed worst performances just below or around 80%. Those results can be explained by the fact that there is a greater variability in the composition of these materials (e.g., soil, flowers).

### PART II: Testing under Different Conditions in Real-Time

#### Experimental Protocol

Finally, we assessed the scalability of the Deep Thermal Imaging and its generalization to new unseen data with different conditions, i.e., training on the outdoor dataset of study II and testing on a new dataset containing the same type of material but collected under different conditions, e.g., iron fences *in front of windows* (transparent materials like glass could influence temperature patterns [47]), *rubber covered tactile paving*. In addition, we tested how well the system could recognize material that was partially altered and covered by other material, i.e., pavement covered by sand,

Layer	0	1	2	3	4	5	6	7	8	9	10	11	12
Type	Input	Aver. Pool.	Conv.	ReLU	Max Pool.	Conv.	ReLU	Max Pool.	Fully conn.	ReLU	Fully conn.	<b>Drop out</b>	Soft max
Kernel size	n/a	2x2	7x7	n/a	2x2	5x5	n/a	2x2	1x1	n/a	1x1	(0.3)	n/a
Kernel Num.	n/a	n/a	12	n/a	n/a	24	n/a	n/a	48	n/a	17	n/a	n/a

Table 2. CNN architecture used for Study II: based on [26] and dropout technique [29] to handle overfitting issues.

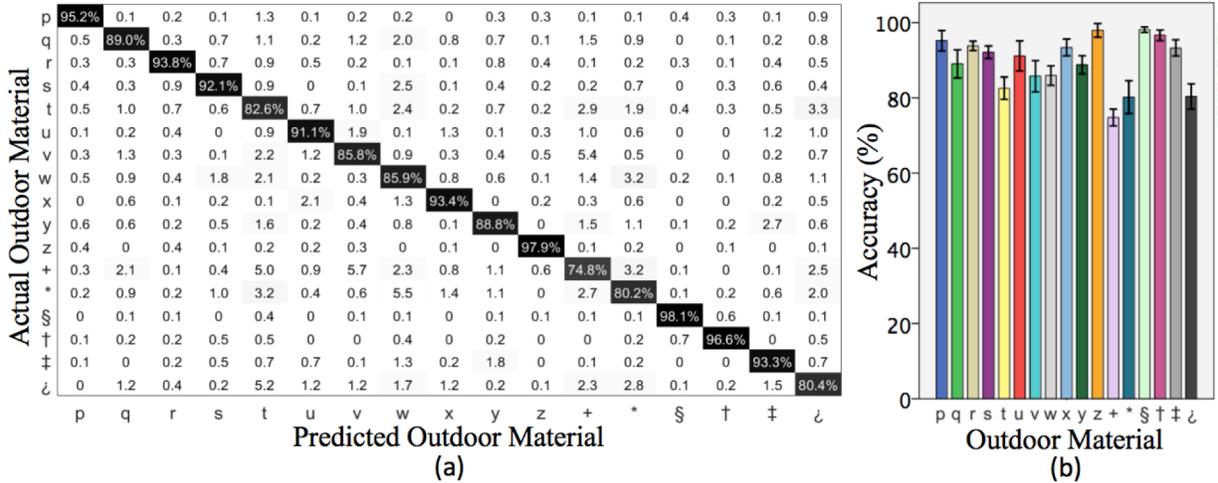


Figure 12. 10-fold cross validation results from Study II (high dynamic temperature range proximate material, labels for materials shown in Figure 4): (a) confusion matrix (in percentage; color range: white(0%)-black(100%)), (b) mean class accuracy recognition distribution (error bar: 95% confidence interval).

and wet because of rain. This part of the study was conducted in real-time to simulate its use in real-life applications: shooting a thermal image of material and immediately testing the classifier pre-trained from the *DeepTherm II* dataset. The real-time proximate material type recognizer (shown in Figures 1c and 3) was designed to make a prediction when a decision probability was higher than a set threshold value (i.e., a score produced from the softmax function in Table 2).

### Results

Figure 13 shows the mean class accuracy results for the thermal images of the new data (in unseen conditions) with respect to the 17 classes from the training set (in Part I of Study II). Prediction results were well above chance level (classifying material among 17 material types) except for the wet cases in comparison with the related dry conditions (e.g., dried asphalt road, concrete paving with clear borders). Performances in recognizing the new iron fence and the rubber-covered tactile paving among 17 outdoor materials were relatively high with 63.16% and 69.23% accuracies, respectively. Given that even the state-of-the-art deep learning techniques for RGB-vision-based material recognition tasks produced about a 50% mean accuracy in discriminating materials across 10 categories from cross-validated dataset experiments (i.e., when training on a dataset and testing on new dataset) [2], the obtained performances can still be considered very good. On the other hand, the classifier trained on *DeepTherm II* was unable to recognize wet materials, particularly tactile and concrete paving in the 6th and 7th columns in Figure 13, suggesting the need to include such weather condition in the training set.

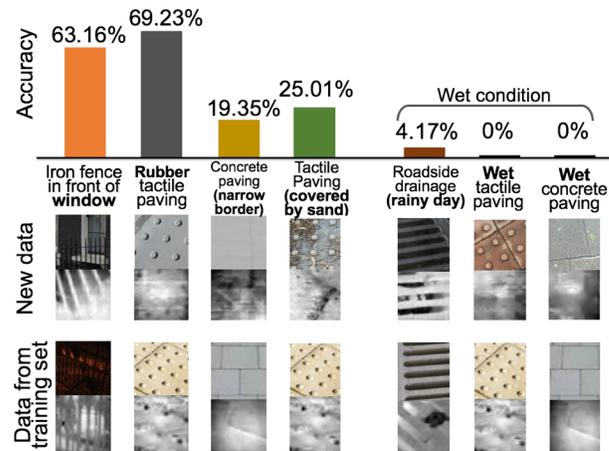


Figure 13. Mean class accuracy results (17 classes) from real-time recognition test on new data in different conditions (Study II, Part II): prediction results were by and large lower, but we note that this task is highly challenging in image classification (e.g. about 50% mean accuracy from 10 classes cross-dataset experiments in [2]).

### EXAMPLE USE CASES

Before discussing the lesson learnt from the experimental studies, we discuss some of the possible HCI uses of our *Deep Thermal Imaging* system and its value in comparison with other material-sensing approaches.

#### Extension of Indoor Use Cases

As addressed in [48], a material recognizer helps lessen the need for special markers and tags in Augmented Reality

(AR), context-awareness applications. Earlier recognizers, however, limit possible scenarios due to the need to be in contact with the material to be recognized or to use special surfaces (e.g., adding transparent material to a showcase as a tag to provide video instruction on a user’s smartphone [48]). *Deep Thermal Imaging* does not have such requirements allowing more seamless HCI applications. For example, by integrating a thermal camera into a AR headset (e.g., Microsoft HoloLens), a home redecoration application could help its user select the right type of painting for the walls or for any other object in the home by simply moving around the home. The application could show how that surfaces would appear when painted with a particular color of a certain type of paint. Indeed, the resulting color depends on the material properties and not just on the paint used. A simpler version could be obtained by using a smartphone rather than the AR headset by pointing the smartphone to the objects as the person moves around.

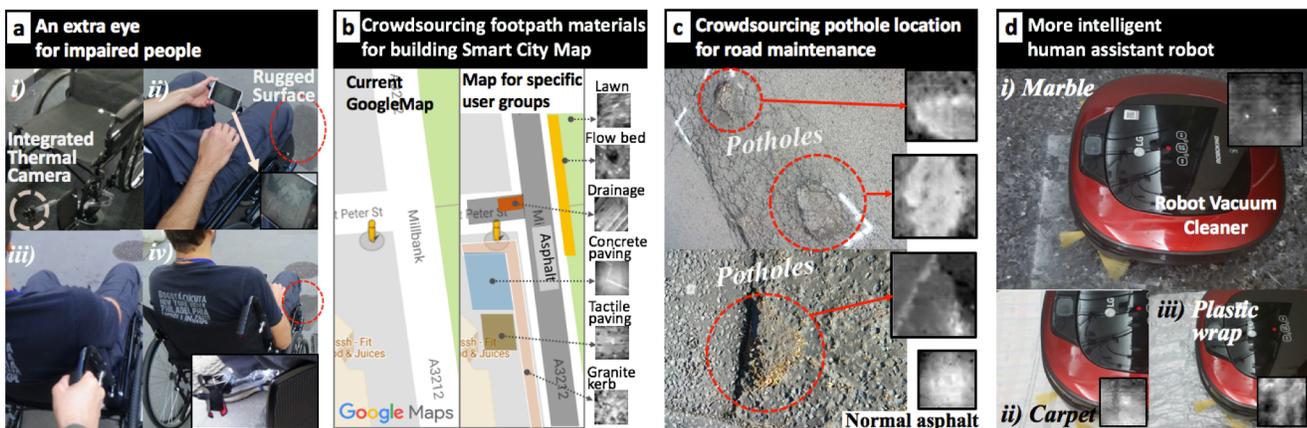
Another interesting example could be the improvement of human assistant robots, such as, vacuum cleaning robots (e.g., Dyson 360 Eye robot [65], LG Hom-Bot square [66]). These systems use ultrasonic sensors [33] and photo-detectors [21] to detect the presence of obstacles. However, these systems are limited in their recognition of floor types. Users must intervene to adjust parameters and tooling for maximizing the task performance when changing floor type (e.g., hard marble floor to deep pile carpet as shown in Figure 14(d)). Given that our system is less affected by ambient illumination and can differentiate between material types in a contact-less manner, it can facilitate automatic tools and parameters adjustment without the continuous intervention of a human operator. The capability could be extended to other floor types beyond the tested floor materials (e.g., marble, carpet) by increasing the training set.

### Outdoor Use Cases

Deep Thermal Imaging could become an artificial eye to increase a person’s ability to discern outdoor surroundings without touching them. Visually impaired people could benefit from this approach. As discussed earlier, the Microsoft *Seeing AI* [58] can be of a huge help for the community with its function to describe the scene around a user. Beyond typical situations with moderate light levels, a smartphone RGB camera used for the Seeing AI may not work properly. The performance could be improved by complementing it with our system that is illumination-independent.

A wheelchair user could also exploit this system to avoid more unpleasant footpaths when navigating at night or backwards and in crowded areas with the setup described in Figure 14(a). Informal interviews with wheelchair users highlight the difficulty that footpaths may present. Deep Thermal Imaging integrated on the wheelchair could detect poor surfaces as the wheelchair approaches it (e.g., uneven surface, manhole cover [27]) and alert the user (e.g., through sound, vibration) or inform the automatic engine driver. It can also help caretakers when pushing a wheelchair given their limited visibility of the footpath (as per interviews).

Another important application, extending the previous one, is road material and state crowdsourcing [19,24,25]. This is fundamental to improve route planning for pedestrians, cyclists including the user groups described above. This is particularly important in a novel environment or in non-industrialized countries where streets are not yet well mapped or maintained [19,23–25]. Crowdsourcing, through the setup (with a smartphone GPS) in Figure 14(a), could help build a smart city map (e.g., Figure 14(b)) by improving the information provided by current maps (e.g., Google Accessibility Map [67]). Similarly, the crowdsourcing



**Figure 14. Example applications: (a) an additional eye for people (visually impaired people or wheelchair users could benefit by being alerted that a footpath is wet or in poor conditions even at night), (b) crowdsourcing materials on footpath for building smart city maps, this in turn can help bicycle users, runners, the low vision community to navigate using route most suitable to their means and needs, (c) crowdsourcing pothole locations for road maintenance (potholes cause a serious problem in cities, maintaining them is challenging), and (d) more intelligent human assistant robot (Deep Thermal Imaging can provide the ability to recognize the floor types and conditions, leading to more efficient and adaptive smart-home robots).**

application could help detect potholes on roads which cause serious problems in metropolitan cities like London and New York [17,46]. Lots of money and time are required to find and repair potholes. Beyond voluntary tagging, automatic sensing using accelerometer and GPS is being currently explored for this purpose [17]. However, the information extracted from such approach contains vibration noise. Our system could be used to complement the information provided by the accelerometers to further confirm the presence of a pothole and also indicate the pavement material for preliminary road maintenance planning (Figure 14(c)).

## DISCUSSION

Novel technological sensing approaches are central to many HCI innovations. Recent work [2,48,53,55] investigated material recognition to support material awareness. Building on them, our paper contributes to this body of work by proposing mobile thermal imaging as a new recognition modality to supplement some of the limitations of existing approaches. Our approach, in particular, offers higher flexibility when used in the wild where there are a variety of materials and environmental factors. It does not require contact between the sensor and the material, reducing safety concerns, and social issues and creating more seamless types of interactions and applications. It is accessible to general CHI audiences, hobbyists and end users (e.g., impaired people) as it works with low-cost hardware and uses a deep learning-based approach which can be easily trained to recognize new materials without requiring any special skills by the end user (i.e., feature hand-crafting).

Two main studies were conducted to investigate the efficacy and potential of the Deep Thermal Imaging system in recognizing typical indoor and outdoor material under different illumination, temperature conditions and point of views. Overall, Deep Thermal Imaging showed highly good performance in discriminating between 15 indoor materials and between 17 outdoor surfaces types. The mean class accuracies were slightly higher in the case of indoor materials than outdoor ones. The slightly lower performance for the outdoor material can be explained by higher thermal dynamics (see Figures 7 and 11) and lower homogeneity between objects of similar material. For example, different parts of a pedestrian path could have different amounts of dirt or footprints, altering surface thermal radiation patterns.

### Limitations and Future Directions

Despite the high performances shown by our results, there is space for improvement. First, thermal imaging is affected by thermal reflection of wet and glossy materials [28,47,52]. For instance, the reflective bubble wrap and plastic freezer bag in Figure 6 (material *n* and *e*) had some sporadic (high apparent temperature) white spots on their thermal images. This is due to their high thermal reflectivity [28,47,52] leading to the higher thermal dynamic ranges in Figure 7. In the case of completely unseen contexts and their interaction (e.g., Study II/Part II), the performances also decreased (as expected) but they are still well above chance level. This

suggests that various environmental factors such as weather and adjacent materials could also influence its ability to generalize. Given that reflectance and transmission are influenced by thermal radiations received from other materials, material with high reflectivity and transmissivity may require a larger dataset as they may be more affected by the environments and material they are placed on or closed to. Given the nature of deep learning, growing a dataset and optimizing parameters of the model [26] could help generalize to the variety [18,57]. A possible way to continuously increase the dataset is to build a crowdsourcing site as discussed for the various HCI use cases.

Lastly, constraints arise from the resolution and imaging quality of the thermal imaging system itself. In particular, low spatial resolution limits the distance at which materials can be classified to less than 0.5 meter. In addition, the thermal sensitivity ( $<0.1^{\circ}\text{C}$ ) was also limited. Some quantized thermal images in Figure 6 (material *k*) and Figure 10 (material *§*) have a white round spot at the center which were possibly induced by the limited sensitivity. When we relate it to their narrow range of temperature distribution (lower than  $1^{\circ}\text{C}$ , see Figure 7 and 11), the bulk of thermal radiation differences within the spot might not be captured. When we first undertook this Deep Thermal Imaging project, only the *FLIR One (2G)* was available as a mobile thermographic channel. However, now we can see more options coming to the market such as *Therm-App Hz* (see Table 1). Given the high performances produced by our method, we can expect that by using the new low cost but more sensitive cameras that are emerging, these performances will get even better. Indeed, our approach is applicable to any types of thermal cameras. Beyond this unimodal sensing, combination of multiple low-cost mobile sensors could be the next step to build better material and context awareness (e.g., LIDAR rangefinder to easily detect the depth of potholes and our system to do material types).

## CONCLUSION

Deep Thermal Imaging is a new technique to recognize materials in proximity which is based on the use of mobile thermal imaging and deep learning. To the best of our knowledge, this paper provides the first demonstration that spatial thermal patterns of materials can be used to enable automatic material recognition tasks with good performances (above 89% mean accuracy from 15 indoor and 17 outdoor materials). We leveraged a mobile, low-frame rate thermal camera to build a large dataset (41,444 images) of variant thermal textures of materials. Potential use cases of the systems compared with the state of the art are also discussed highlighting the potential of the system. The collected DeepTherm I and II datasets and API are available for download at (<http://youngjuncho.com/datasets>).

## ACKNOWLEDGMENTS

Youngjun Cho was supported by University College London Overseas Research Scholarship (UCL-ORS) awarded to top quality international postgraduate students.

## REFERENCES

1. Miika Aittala, Tim Weyrich, and Jaakko Lehtinen. 2015. Two-shot SVBRDF capture for stationary materials. *ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH 2015* 34, 4: 110–122. <https://doi.org/10.1145/2766967>
2. Sean Bell, Paul Upchurch, Noah Snavely, and Kavita Bala. 2015. Material Recognition in the Wild With the Materials in Context Database. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3479–3487.
3. LouAnne E. Boyd, Xinlong Jiang, and Gillian R. Hayes. 2017. ProCom: Designing and Evaluating a Mobile and Wearable System to Support Proximity Awareness for People with Autism. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*, 2865–2877. <https://doi.org/10.1145/3025453.3026014>
4. William D. Callister and David G. Rethwisch. 2011. *Materials science and engineering: an introduction*. John Wiley & Sons NY.
5. Victor C. Chen and Hao Ling. 2002. *Time-frequency transforms for radar imaging and signal analysis*. Artech House.
6. E. Cheung and V. J. Lumelsky. 1989. Proximity sensing in robot manipulator motion planning: system and implementation issues. *IEEE Transactions on Robotics and Automation* 5, 6: 740–751. <https://doi.org/10.1109/70.88096>
7. Youngjun Cho, Andrea Bianchi, Nicolai Marquardt, and Nadia Bianchi-Berthouze. 2016. RealPen: Providing Realism in Handwriting Tasks on Touch Surfaces Using Auditory-Tactile Feedback. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*, 195–205. <https://doi.org/10.1145/2984511.2984550>
8. Youngjun Cho, Nadia Bianchi-Berthouze, and Simon J. Julier. 2017. DeepBreath: Deep Learning of Breathing Patterns for Automatic Stress Recognition using Low-Cost Thermal Imaging in Unconstrained Settings. In *the 7th International Conference on Affective Computing and Intelligent Interaction, ACII 2017*, 456–463.
9. Youngjun Cho, Munchae Jung, and Sunuk Kim. 2015. Electronic device having proximity touch function and control method thereof, US Patent, Publication Number: US 2015/0062087.
10. Youngjun Cho, Simon J. Julier, Nicolai Marquardt, and Nadia Bianchi-Berthouze. 2017. Robust tracking of respiratory rate in high-dynamic range scenes using mobile thermal imaging. *Biomedical Optics Express* 8, 10: 4480–4503. <https://doi.org/10.1364/BOE.8.004480>
11. Mircea Cimpoi, Subhransu Maji, and Andrea Vedaldi. 2015. Deep Filter Banks for Texture Recognition and Segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3828–3836.
12. Mike Crang and Nigel Thrift. 2002. *Thinking Space*. Routledge.
13. Anind K. Dey and Jonna Häkkinä. 2008. Context-Awareness and Mobile Devices. *Handbook of research on user interface design and evaluation for mobile technology*: 205–217.
14. Andrey Dimitrov and Mani Golparvar-Fard. 2014. Vision-based material recognition for automated monitoring of construction progress and generating building information modeling from unordered site image collections. *Advanced Engineering Informatics* 28, 1: 37–49. <https://doi.org/10.1016/j.aei.2013.11.002>
15. Paul Dourish. 2004. What we talk about when we talk about context. *Personal and Ubiquitous Computing* 8, 1: 19–30. <https://doi.org/10.1007/s00779-003-0253-8>
16. Marks Eric and Teizer Jochen. Proximity Sensing and Warning Technology for Heavy Construction Equipment Operation. *Construction Research Congress 2012*: 981–990. <https://doi.org/10.1061/9780784412329.099>
17. Jakob Eriksson, Lewis Girod, Bret Hull, Ryan Newton, Samuel Madden, and Hari Balakrishnan. 2008. The pothole patrol: using a mobile sensor network for road surface monitoring. In *Proceedings of the 6th international conference on Mobile systems, applications, and services*, 29–39. <https://doi.org/10.1145/1378600.1378605>
18. Ross Girshick. 2015. Fast R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision*, 1440–1448.
19. Kotaro Hara, Vicki Le, and Jon Froehlich. 2013. Combining Crowdsourcing and Google Street View to Identify Street-level Accessibility Problems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*, 631–640. <https://doi.org/10.1145/2470654.2470744>
20. Chris Harrison and Scott E. Hudson. 2008. Lightweight Material Detection for Placement-aware Mobile Computing. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology (UIST '08)*, 279–282. <https://doi.org/10.1145/1449715.1449761>
21. Antonius Hendriks, Damian M. Lyons, and Frank Guida. 2001. Vacuum cleaner with obstacle avoidance, US Patent, US6226830 B1.
22. Ken Hinckley and Mike Sinclair. 1999. Touch-sensing Input Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*

- (CHI '99), 223–230.  
<https://doi.org/10.1145/302979.303045>
23. Catherine Holloway and Nick Tyler. 2013. A micro-level approach to measuring the accessibility of footways for wheelchair users using the Capability Model. *Transportation planning and technology* 36, 7: 636–649.
  24. H. Holone, G. Misund, and H. Holmstedt. 2007. Users Are Doing It For Themselves: Pedestrian Navigation With User Generated Content. In *The 2007 International Conference on Next Generation Mobile Applications, Services and Technologies (NGMAST 2007)*, 91–99.  
<https://doi.org/10.1109/NGMAST.2007.4343406>
  25. Harald Holone, Gunnar Misund, Håkon Tolsby, and Steinar Kristoffersen. 2008. Aspects of Personal Navigation with Collaborative User Feedback. In *Proceedings of the 5th Nordic Conference on Human-computer Interaction: Building Bridges (NordCHI '08)*, 182–191.  
<https://doi.org/10.1145/1463160.1463180>
  26. Max Jaderberg, Karen Simonyan, Andrew Zisserman, and koray kavukcuoglu. 2015. Spatial Transformer Networks. In *Advances in Neural Information Processing Systems 28*, C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama and R. Garnett (eds.). Curran Associates, Inc., 2017–2025.
  27. Piyawan Kasemsuppakorn and Hassan A. Karimi. 2009. Personalised routing for wheelchair navigation. *Journal of Location Based Services* 3, 1: 24–54.  
<https://doi.org/10.1080/17489720902837936>
  28. J. Kölzer, E. Oesterschulze, and G. Deboy. 1996. Thermal imaging and measurement techniques for electronic materials and devices. *Microelectronic Engineering* 31, 1: 251–270.  
[https://doi.org/10.1016/0167-9317\(95\)00348-7](https://doi.org/10.1016/0167-9317(95)00348-7)
  29. Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. 2012. ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou and K. Q. Weinberger (eds.). Curran Associates, Inc., 1097–1105.
  30. Daniel Kurz. 2014. Thermal touch: Thermography-enabled everywhere touch interfaces for mobile augmented reality applications. In *Mixed and Augmented Reality (ISMAR), 2014 IEEE International Symposium on*, 9–16.  
<https://doi.org/10.1109/ISMAR.2014.6948403>
  31. Gierad Laput, Robert Xiao, and Chris Harrison. 2016. ViBand: High-Fidelity Bio-Acoustic Sensing Using Commodity Smartwatch Accelerometers. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*, 321–333. <https://doi.org/10.1145/2984511.2984582>
  32. Eric Larson, Gabe Cohn, Sidhant Gupta, Xiaofeng Ren, Beverly Harrison, Dieter Fox, and Shwetak Patel. 2011. HeatWave: Thermal Imaging for Surface User Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*, 2565–2574. <https://doi.org/10.1145/1978942.1979317>
  33. John J. Leonard and Hugh F. Durrant-Whyte. 2012. *Directed Sonar Sensing for Mobile Robot Navigation*. Springer Science & Business Media.
  34. C. Liu, L. Sharan, E. H. Adelson, and R. Rosenholtz. 2010. Exploring features in a Bayesian framework for material recognition. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 239–246.  
<https://doi.org/10.1109/CVPR.2010.5540207>
  35. H. Liu, X. Song, J. Bimbo, L. Seneviratne, and K. Althoefer. 2012. Surface material recognition through haptic exploration using an intelligent contact sensing finger. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 52–57.  
<https://doi.org/10.1109/IROS.2012.6385815>
  36. J. M. Lloyd. 2013. *Thermal Imaging Systems*. Springer Science & Business Media.
  37. Sapan Naik and Bankim Patel. 2017. Thermal imaging with fuzzy classifier for maturity and size based non-destructive mango (*Mangifera Indica L.*) grading. In *Emerging Trends & Innovation in ICT (ICEI), 2017 International Conference on*, 15–20.  
<https://doi.org/10.1109/ETIICT.2017.7977003>
  38. Vinod Nair and Geoffrey E. Hinton. 2010. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th international conference on machine learning (ICML-10)*, 807–814. Retrieved from [http://machinelearning.wustl.edu/mlpapers/paper\\_files/icml2010\\_NairH10.pdf](http://machinelearning.wustl.edu/mlpapers/paper_files/icml2010_NairH10.pdf)
  39. I. Pavlidis, Peter Symosek, B. Fritz, Mike Bazakos, and Nikolaos Papanikolopoulos. 2000. Automatic detection of vehicle occupants: the imaging problem and its solution. *Machine Vision and Applications* 11, 6: 313–320.
  40. Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. 2015. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In *Advances in Neural Information Processing Systems 28*, C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama and R. Garnett (eds.). Curran Associates, Inc., 91–99.
  41. E. F. J. Ring and K. Ammer. 2012. Infrared thermal imaging in medicine. *Physiological Measurement* 33, 3: R33. <https://doi.org/10.1088/0967-3334/33/3/R33>
  42. Marc Rioux. 1984. Laser range finder based on synchronized scanners. *Applied Optics* 23, 21: 3837–3844. <https://doi.org/10.1364/AO.23.003837>

43. Yvonne Rogers. 2011. Interaction design gone wild: striving for wild theory. *Interactions* 18, 4: 58–62. <https://doi.org/10.1145/1978822.1978834>
44. Bernardino Romera-Paredes and Philip Torr. 2015. An embarrassingly simple approach to zero-shot learning. In *International Conference on Machine Learning*, 2152–2161.
45. Chandra Roychoudhuri, Al F. Kracklauer, and Kathy Creath. 2008. *The nature of light: what is a photon?* CRC Press.
46. Fauzia Saeed, Siti Qamariatul, Mujib Rahman, and Alan Woodside. 2015. The state of pothole management in UK local authority. *Bituminous Mixtures and Pavements VI*: 153–159.
47. Alireza Sahami Shirazi, Yomna Abdelrahman, Niels Henze, Stefan Schneegass, Mohammadreza Khalilbeigi, and Albrecht Schmidt. 2014. Exploiting Thermal Reflection for Interactive Systems. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*, 3483–3492. <https://doi.org/10.1145/2556288.2557208>
48. Munehiko Sato, Shigeo Yoshida, Alex Olwal, Boxin Shi, Atsushi Hiyama, Tomohiro Tanikawa, Michitaka Hirose, and Ramesh Raskar. 2015. SpecTrans: Versatile Material Classification for Interaction with Textureless, Specular and Transparent Surfaces. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*, 2191–2200. <https://doi.org/10.1145/2702123.2702169>
49. Albrecht Schmidt, Michael Beigl, and Hans-W. Gellersen. 1999. There is more to context than location. *Computers & Graphics* 23, 6: 893–901. [https://doi.org/10.1016/S0097-8493\(99\)00120-X](https://doi.org/10.1016/S0097-8493(99)00120-X)
50. Igor V. Tetko, David J. Livingstone, and Alexander I. Luik. 1995. Neural network studies. 1. Comparison of overfitting and overtraining. *Journal of Chemical Information and Computer Sciences* 35, 5: 826–833. <https://doi.org/10.1021/ci00027a006>
51. Andrea Vedaldi and Karel Lenc. 2015. MatConvNet: Convolutional Neural Networks for MATLAB. In *Proceedings of the 23rd ACM International Conference on Multimedia (MM '15)*, 689–692. <https://doi.org/10.1145/2733373.2807412>
52. Michael Vollmer and MÃ Klaus-Peter. 2017. *Infrared thermal imaging: fundamentals, research and applications*. John Wiley & Sons.
53. Ting-Chun Wang, Jun-Yan Zhu, Ebi Hiroaki, Manmohan Chandraker, Alexei A. Efros, and Ravi Ramamoorthi. 2016. A 4D Light-Field Dataset and CNN Architectures for Material Recognition. In *Computer Vision – ECCV 2016 (Lecture Notes in Computer Science)*, 121–138. [https://doi.org/10.1007/978-3-319-46487-9\\_8](https://doi.org/10.1007/978-3-319-46487-9_8)
54. Jason Wiese, T. Scott Saponas, and A.J. Bernheim Brush. 2013. Phoneprioception: Enabling Mobile Phones to Infer Where They Are Kept. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*, 2157–2166. <https://doi.org/10.1145/2470654.2481296>
55. Hui-Shyong Yeo, Gergely Flamich, Patrick Schrempf, David Harris-Birtill, and Aaron Quigley. 2016. RadarCat: Radar Categorization for Input & Interaction. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*, 833–841. <https://doi.org/10.1145/2984511.2984515>
56. Hui-Shyong Yeo, Juyoung Lee, Andrea Bianchi, David Harris-Birtill, and Aaron Quigley. 2017. SpeCam: Sensing Surface Color and Material with the Front-facing Camera of a Mobile Device. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '17)*, 25:1–25:9. <https://doi.org/10.1145/3098279.3098541>
57. Xiangxin Zhu, Carl Vondrick, Deva Ramanan, and Charless C. Fowlkes. 2012. Do we need more training data or better models for object detection. In *BMVC*.
58. Seeing AI. Retrieved from <https://www.microsoft.com/en-us/seeing-ai/>
59. Encoded Reality. Retrieved from [http://viral.media.mit.edu/projects/encoded\\_reality/](http://viral.media.mit.edu/projects/encoded_reality/)
60. Porsche Panamera 2017. Retrieved from [http://www.motorauthority.com/news/1105162\\_2017-porsche-panamera-deep-dive/page-3#](http://www.motorauthority.com/news/1105162_2017-porsche-panamera-deep-dive/page-3#)
61. CAT S60: The World's First Thermal Imaging Smartphone. Retrieved from <https://www.catphones.com/?product=cat-s60-smartphone>
62. JETSON TK1. Retrieved from <http://www.nvidia.com/object/jetson-tk1-embedded-dev-kit.html>
63. Caffe deep learning framework. Retrieved from <http://caffe.berkeleyvision.org/>
64. TensorFlow Mobile. Retrieved from <https://www.tensorflow.org/mobile/>
65. Dyson 360 Eye robot. Retrieved from <https://www.dyson.co.uk/robot-vacuums/dyson-360-eye-overview.html>
66. LG HOM-BOT Square. Retrieved from <http://www.lg.com/uk/hom-bot>
67. Google Maps now lets users add wheelchair accessibility details for locations. Retrieved from <https://techcrunch.com/2017/07/08/google-maps-now-lets-users-add-wheelchair-accessibility-details-for-locations/>