Hot or Not? Robust and Accurate Continuous Thermal Imaging on FLIR cameras

Titti Malmivirta*, Jonatan Hamberg*, Eemil Lagerspetz*, Xin Li*,
Ella Peltonen†‡, Huber Flores* and Petteri Nurmi*§
*Department of Computer Science, University of Helsinki, Helsinki, Finland
†Insight Centre for Data Analytics, University College Cork, Cork, Ireland
‡University of Oulu, Oulu, Finland
§Lancaster University, Lancaster, United Kingdom
titti.malmivirta@helsinki.fi, jonatan.hamberg@helsinki.fi, eemil.lagerspetz@cs.helsinki.fi,
xin.li@helsinki.fi, ella.peltonen@oulu.fi, huber.flores@helsinki.fi, petteri.nurmi@cs.helsinki.fi

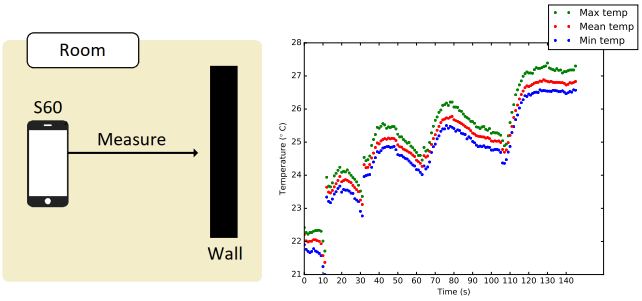
***Abstract*—Wearable thermal imaging is emerging as a powerful and increasingly affordable sensing technology. Current thermal imaging solutions are mostly based on uncooled forward looking infrared (FLIR), which is susceptible to errors resulting from warming of the camera and the device casing it. To mitigate these errors, a blackbody calibration technique where a shutter whose thermal parameters are known is periodically used to calibrate the measurements. This technique, however, is only accurate when the shutter's temperature remains constant over time, which rarely is the case. In this paper, we contribute by developing a novel deep learning based calibration technique that uses battery temperature measurements to learn a model that allows adapting to changes in the internal thermal calibration parameters. Our method is particularly effective in *continuous* sensing where the device casing the camera is prone to heating. We demonstrate the effectiveness of our technique through controlled benchmark experiments which show significant improvements in thermal monitoring accuracy and robustness.**

***Index Terms*—thermal sensing, thermal imaging, sensor calibration, deep learning, mobile computing, sensing, IoT, pervasive computing**

I. Introduction

Thermal Imaging is increasingly available thanks to affordable off-the-shelf thermal cameras. Indeed, attachable micro-USB thermal cameras costing around $100 (e.g., FLIR ONE) are already available on the market, and also some high-end smartphones integrate thermal cameras (e.g., Caterpillar CAT S60 and CAT S61[1]). Thanks to this development, new types of application areas and studies that utilize thermal sensing are emerging. Examples of these application areas include energy auditing of buildings [1], [2], diverse medical applications [3], [4], [5], continuous monitoring of animals [6], [7], search and rescue operations [8], psychological sensing applications such as detection of cognitive load or affective states [9], [10], and energy modeling of IoT devices [11].

Off-the-shelf thermal cameras predominantly rely on uncooled *forward looking infrared (FLIR)* which measures variations in radiation reflected at infrared wavelengths [12]. A fundamental challenge with FLIR cameras is that measurements are affected by temperature of the camera and the device casing it [13]. This is particularly problematic in continuous monitoring where camera and device temperature can fluctuate significantly over time. To mitigate these effects, the thermal sensor needs to be periodically recalibrated. A common technique for recalibration is to cover the thermal sensor with a shutter whose thermal properties are known (typically measured in laboratory environments), and to estimate the required calibration parameters from the difference between current and known values of the shutter [13]. While this method can mitigate errors, unfortunately it is effectively only when the shutter's temperature remains stable, which rarely is the case. This is particularly the case for smartphones and other wearables where CPU, internal temperature, and use of the camera all influence the shutter's temperature; see Sec. II.



(a) (b)

Fig. 1. Temperature measurements of a target in ambient room temperature (see (a)) fluctuate over time as seen in (b).

To highlight and illustrate the severity of errors in continuous thermal monitoring, Fig. 1 shows temperature measurements from a stable target (cardboard box in ambient room temperature) obtained using the thermal camera of a Caterpillar CAT S60 smartphone. From the figure we can clearly see how the temperature continues to increase even though the camera is pointed at an object with constant temperature. In Figure 2 we have highlighted points where the device performs internal calibration to show how it is unable to mitigate errors in the measurements. Indeed, from the plot

[1]https://www.catphones.com/

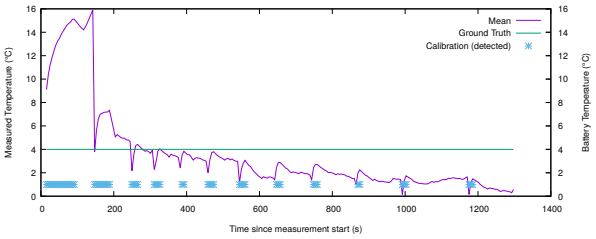


Fig. 2. When the thermal camera calibrates, measurement values drop suddenly, and take tens of seconds to rise back to normal levels.

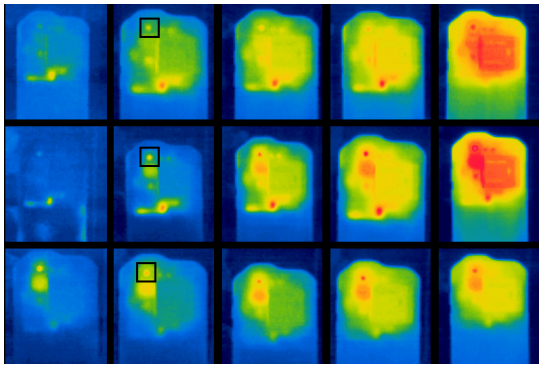


Fig. 3. Left to right: thermal images at 1, 2, 3, 5, and 10 min when running a stress test. When the camera and CPU are both active (middle row), the heat signature is very different from when only the thermal camera is active (bottom row). When only CPU is active for a longer period, we get higher heat saturation (top row). The small hot point indicated by the boxes in the second column images is the thermal camera aperture on the device. The larger hot area below it in the middle and bottom row of images is the flash LED aperture.

it would appear that calibration actually exacerbates the errors instead of mitigating it. The reason for this is that operating the camera, running CPU intensive applications, and other factors result in heat seeping through to the shutter, and invalidating the device internal thermal parameters.

In this paper, we contribute by developing a novel *deep learning based calibration technique* for improving the quality of thermal imaging, particularly for applications that require continuous monitoring. In our approach, battery temperature is used as a proxy for estimating the heating of the device, and a calibration function that captures the offset between thermal camera measurements and the actual temperature is learned. By compensating values of the FLIR camera with offset estimates, the performance of the thermal imaging is significantly improved. We demonstrate the effectiveness of our approach through benchmark experiments carried out in carefully controlled hot and cold conditions. Results of our experiments demonstrate significant improvements in both the accuracy and robustness of thermal monitoring.

II. Motivation and Background

The focus of our work is on improving the accuracy and robustness of FLIR measurements. As shown in Fig. 1, current off-the-shelf FLIR cameras are prone to measurement errors, which limit the usefulness of measurements obtained from them. In this section we investigate how different factors affect the error. We first examine the effectiveness of the blackbody calibration technique used by current off-the-shelf cameras and demonstrate how even that is a source of error, due to incorrect assumptions of the thermal characteristics of the shutter used as blackbody. Secondly, we examine how CPU and camera use result in heating of the device, and how this heat affects the temperature of the shutter. Finally, we demonstrate that these errors can be, to a large extent, identified from changes in the battery temperature of the device casing the camera.

A. Effect of Calibration on Thermal Camera Error

Fig. 2 shows values of thermal camera measurements taken over a 20 minute period by monitoring the surface of water inside a Smart Fridge (see Sec. IV). From the figure we can observe that each calibration cycle results in a clearly identifiable peak in temperature measurements. In the plot, the first two calibration cycles have a high error, after which the error starts to stabilize. During tests of the thermal camera we have seen this to be a relatively common occurrence, which suggests that the first $1-2$ calibration cycles should be omitted. Initially the calibration cycles are more frequent, but the frequency converges to around 180 seconds, which seems to be a device internal parameter on the Caterpillar CAT S60 smartphones used in our experiments. Finally, while the magnitude of changes resulting from the calibration algorithm seems to converge, the error in temperature measurements increases over time. As we demonstrate later in this section, this increase in error mirrors changes in the internal temperature of the device, which in this scenario cools down from around $30°$C to around $24°$C; see Fig. 4.

B. Effect of CPU and Device Casing

We next examine the extent of heat seeping through into the shutter area. To estimate this, we carried out experiments where we monitored the temperature of the back cover of a CAT S60 smartphone while running a stress test that maximized device CPU use. We repeated the experiment by running the stress test while having the thermal camera active, as well as having the thermal camera active without running the CPU stress test. We did this to explore the effects of heat conduction inside the device while the thermal camera is in use. The temperature of the device was captured using a FLIR TG167 hand-held thermal camera, which produced heat images of the back cover. The results are shown in Fig. 3. From the images we can observe clear thermal hotspots, i.e., areas that warm up most. From running the stress test for $1, 2, 3, 5$, and 10 minutes with and without operating the thermal camera, we observe distinguishable heat signatures, indicating that heat resulting from system load is independent from heat resulting from use of the thermal sensor; see Fig. 3.

In all experiment conditions we can clearly see that the area around the thermal camera aperture and the shutter used in the blackbody calibration heats up. The effect of this heating is further exacerbated by the CAT S60 being dust and waterproof,

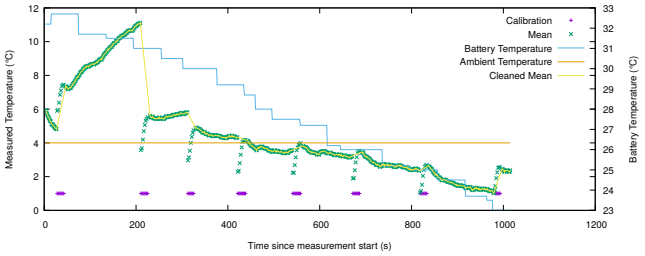


Fig. 4. Thermal calibration while measuring cold water in a smart refrigerator. The temperature measurement is affected by the temperature of the device, as evidenced by the internal battery temperature sensor.

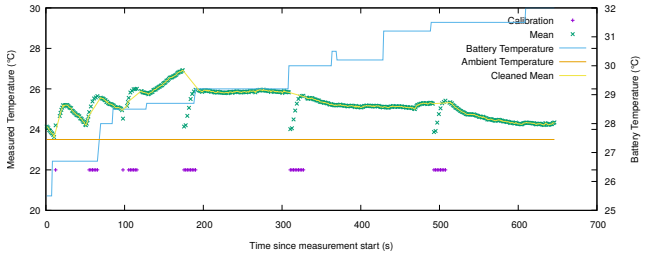


Fig. 5. Temperature measurements of a wall in an ambient temperature of 24°C. The temperature measurement is affected by the temperature of the device, resulting in higher than expected values.

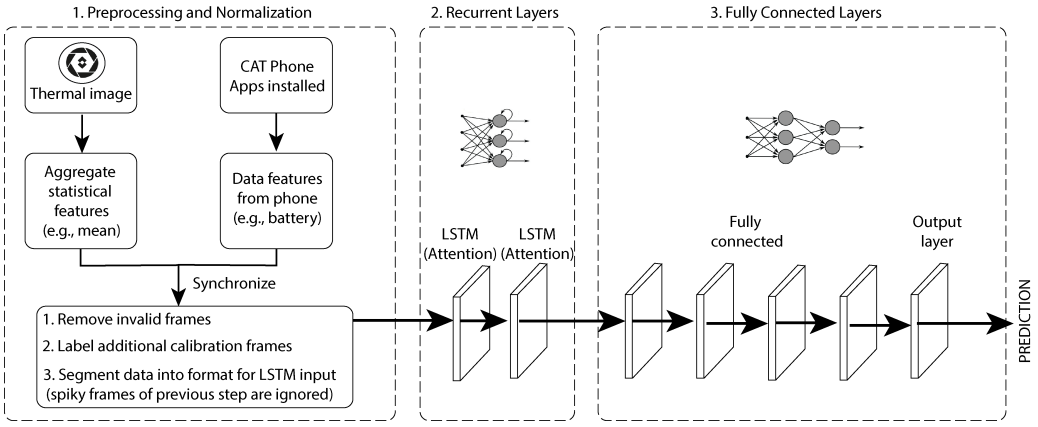


Fig. 6. The deep neural network structure of our deep calibration approach for thermal cameras.

which means most of the heat is released through the opening of the thermal camera. We can see that the hottest areas on the right are the hinge of the SIM slot cover, a larger area under the SIM slot door, and the opening of the thermal camera. With the CPU active (middle and top rows), a large area around the opening of the camera flash LEDs also heats up. In summary, besides seeing a clear increase in thermal radiation in the thermal camera during operating the camera, we can see that also other operations of the device have an influence on the temperature around the thermal camera.

C. Relationship between Battery Temperature and Thermal Camera Measurements

We next investigate how the error in FLIR measurements changes when the temperature of the device casing it changes. We consider battery temperature, available through most smart battery interfaces, as a proxy for change in device temperature. We consider two experiment setups, measuring cold water inside a smart fridge, and measuring a painted wall in ambient temperature; see Sec. IV.

Fig. 4 shows how the measured mean temperature of the water changes over time in the smart fridge scenario, and Fig. 5 shows results for ambient temperature scenario. The pattern of measurement error is similar in both cases. From both images we can see an error caused by calibration, which stabilizes after two calibration cycles. For monitoring a wall in ambient temperature, the values are slightly higher than the ground truth, with the values initially rising as battery temperature increases. For the smart fridge case, the reverse occurs, i.e., the thermal camera values decrease below ground truth as the device cools down. In both experiments, the thermal camera values drift to a state where they consistently differ from the ground truth. Result of these experiments highlight how changes in battery temperature mirror errors in the thermal camera values, and thus serve as a proxy for assessing effect of the device casing the camera.

III. Adaptive Thermal Calibration

The previous section demonstrated that the temperature measurements recorded on the CAT S60 thermal camera suffer from several problems. Firstly, the data fluctuates as the device self-calibrates, and this can occur at any time when operating the camera. Secondly, the values before and after these fluctuations are further away from the true temperature of the object than desired, and seem to mirror changes in the internal temperature of the device. To mitigate the effects of these error sources, in this section we develop a novel deep learning based calibration technique for improving the accuracy and robustness of thermal camera measurements.

A. Overview and Implementation

An overview of our model is shown in Fig. 6. In the first stage (Preprocessing and Normalization), we extract five aggregate statistical features from the thermal image: difference

in temperature, minimum, maximum, average, and variance of the target object (as given by the image shown in the screen). Difference in temperature measures changes in target object over time, while the other features characterize the most recent thermal image. In our experiments, we assume the input image has been cropped to match with the target object. When this is not the case, image segmentation techniques can be used to identify the appropriate target (see, e.g., [14]). Parallel to this, we retrieve battery temperature and latest CPU usage that are closely correlated with the heating/cooling process from the smartphone operating the thermal camera. Both sets of features, together with the internal calibration label (In Progress or Tuned) returned by the device are then synchronized into a single (8 dimensional) representation, which is transformed to match input for an LSTM layer. In the transformation process we first remove noise introduced by the internal calibration by removing all data from the calibration period. The internal calibration labels returned by the CAT S60 smartphone seem to correspond to the end of the calibration period and reflect the end of the period where measurements have clearly identifiable drift. To remove all noisy measurements, we apply a reverse peak detection using the frame labeled as "calibrating" as the end of the peak and remove all preceding frames matching the same peak. On average, this process removes 13 frames preceding the end of the calibration period while preparing segmented training data for our model.

Our current implementation uses the Keras[2] deep learning library with TensorFlow as backend. In our experiments we run the model on a commodity laptop (HP EliteBook 820). However, the model can be converted to TensorFlow Lite[3] based implementation, which can be employed on Android mobile devices such as the CAT S60. In total, measurements of 315 minutes at 1 fps and 3 MB size after preprocessing and feature extraction, are used for model training and testing, in which the amount of data varies depending on different evaluation settings (see Sec. IV). The training is fast, as only 8 numerical features and a relatively small step size (see Sec. III-A) are employed in our LSTM based model. The model is trained for 10 epochs and each epoch takes on average 10 seconds, depending on the evaluation scenario.

B. Deep Learning Model

Our deep learning model consists of two further stages, stacked LSTM layers (Recurrent layers) with attention mechanism [15] followed by 5 fully-connected layers. The Recurrent Layers use 10s as the time step for LSTM to take temporal dependencies into account. As heating or cooling of the device is a gradual process whose impact changes over time, we utilized LSTM layers to capture these temporal patterns. The 10 seconds time step causes the model to predict drift error based on 10 seconds of data prior to the current frame. The attention mechanism enables finer tuning and integration of the input and output at each time step. It calculates a weight array that gives different importance to each intermediate output, and calculates the dot product of the weight and intermediate output array as the output of LSTM layers. In addition, LSTM layers can mitigate fluctuations caused by noisy sensor data because they are robust against individual erroneous frames that show spikes in temperature measurements. To ensure the 10s time step setting is consistent, we only consider data of 1Hz in our main experiments. Since amount of data from each experiment scenario is limited, we employ L2 (0.01) kernel regularization on the weights of the LSTM layers as well as recurrent dropout (rate = 0.1) to prevent overfitting. In the final stage, Fully Connected Layers, the 5 fully-connected layers aggregate the high-dimensional activation output from LSTM layers and match it with a final response. The 5 layers all take RELU as activation function except for the output layer, which uses a linear function. The layers have 64, 128, 32, and 1 units, respectively.

[2]https://keras.io/

[3]https://www.tensorflow.org/lite/

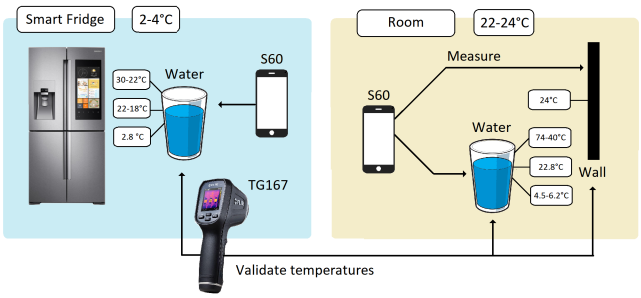


Fig. 7. Experimental setup with the CAT S60 devices measuring water containers and a wall at two different ambient temperatures (smart fridge and room). Temperatures of the containers are validated with the FLIR TG167.

TABLE I
SUMMARY OF EXPERIMENT SETUPS

Environment	**Hot**	**Cold**	**Room Temp**
Room	74 - 40 °C	4.5 - 6.2 °C	22.8 °C
Fridge	30 - 22 °C	2.8 °C	22 - 18 °C

IV. EXPERIMENTAL SETUP

We rigorously evaluate our calibration technique using data collected from two controlled environments with differing ambient temperatures (ambient and cold) and target objects with three different temperatures (hot, ambient, cold). In the following we detail our experimental setup and measurements.

Evaluation Environments: We consider thermal measurements captured with the FLIR camera of a Caterpillar CAT S60 smartphone. We collect measurements from two environments, a Samsung RB38M7998S4 smart fridge as cold environment and an indoor office space as ambient. Fig. 7 illustrates the measurement setups, and Table I summarizes the temperature values in our test environment. In the table, the last column of first row and middle column of second row correspond to static setups, others are dynamic. In both environments, measurements from the FLIR camera were captured at 1 fps.

Cold, Smart Fridge: To acquire measurements from a controlled cold environment, we placed the CAT S60 camera inside a Samsung RB38M7998S4 smart fridge and used the FLIR to monitor surface of water placed in a plastic container. The internal temperature of the fridge was set to 4 °C. As the smartphone operated inside the fridge, its operating temperature cooled over time.

Ambient, Room: As our second environment, we consider a room with an ambient temperature of 23.5 - 25 °C. The temperature of the room was measured using a digital thermometer placed approximately one meter apart from the phone.

Target Objects: The camera was pointed at water containers with controlled temperatures. Water temperature was validated using a TM-947SD Thermometer which measured temperature every two seconds. We use water containers as target objects because water reflects a high fraction of thermal radiation. Specifically, since FLIR cameras measure thermal radiation instead of absolute temperature, the accuracy of the measurements is sensitive to the fraction of thermal energy that the target object emits. The energy reflected by an object is referred to as its emissivity, denoted ϵ. Emissivity ranges from full absorption (i.e., no thermal energy is reflected, $\epsilon = 0$) to full reflectance (i.e., all energy is reflected $\epsilon = 1$). For water, emissivity is approximately 0.98, i.e., most of the thermal radiation is reflected back and visible in the FLIR measurements. Target objects had three different temperatures: hot, cold, and ambient, all measured in both environments. Temperature of the objects was either static or dynamic: cold object in cold environment and ambient temperature object in ambient environment had static temperatures, and in other combinations temperatures were dynamic.

Cold: When measuring in the cold environment, the cold container was placed in the fridge well in advance to ensure that the temperature of the water matched the internal temperature of the fridge as closely as possible. According to the a TM-947SD Thermometer, water temperature was 2.3 - 2.5 °C after 3 hours in the refrigerator. In the ambient environment the water warmed during the measurement, starting at approximately 4.5 °C and being 6.2 °C at the end.

Ambient: When measuring an ambient temperature object in ambient temperature the water container was placed in room temperature for a few hours before the measurement. Temperature of the water was measured to be 22.8 °C using the thermometer. The thermal camera was used to measure water surface for 15 minutes, repeated three times. To regulate the temperature of the phone there was a 15 minute break between the repetitions. Object temperature was stable during the measurement. In the cold environment the temperature of the water cooled down during the measurements, starting at about 24 °C. The phone was kept in the refrigerator 15 minutes in first two measurements and 30 minutes in the third. Water temperature was 22 °C and 18 °C at the end of the second and third measurement periods, respectively.

Hot: In both environments the temperature of the water cooled down during the measurements. In the ambient temperature environment temperature of the hot water was approximately 70 °C at the beginning of the measurement and at the end it was 40 to 45 °C. In the cold environment the hot water was colder than in the ambient environment to prevent it warming up the refrigerator. The warm water bowl was kept in the refrigerator for 45 minutes and the temperature measured by the thermometer cooled from 33.7 °C to 22.9 °C in the first measurement and from 27.2 °C to 19.1 °C in the second.

Additional frame rates: Since the frame rate of the camera can be assumed to affect the temperature of the phone, 3 additional frame rates were also used. This was done in the ambient temperature environment with the water bowl as the measured object. The frame rates used were 0.05 fps, 0.2 fps and 3 fps, each measured twice. Duration of the measurement was 10 minutes at 3 fps and 15 minutes at other frame rates.

Thermometer synchronization: Temperature measurements were verified with a TM-947SD Thermometer. The thermometer allows only manual adjustment of time and hence no precise time synchronization was possible. To align timestamps as closely as possible, we performed a two phase synchronization. First, we manually aligned times between the thermometer and the CAT S60 smartphone to match as closely as possible. While this is sufficient for cases where the temperature remains (approximately) static, inaccuracies in the time synchronization may introduce some errors for dynamically changing temperatures. In the dynamic temperature experiments the synchronization was done by warming up the thermometer sensor at the end of each measurement period. The warming up was done with objects clearly warmer than the measured water to cause the temperatures measured by the thermometer to increase rapidly. To get the equivalent time point from the CAT phone, the phone's thermal camera was pointed at the thermometer and the object warming it. In the refrigerator environment the warming up was done by holding thermometer sensor in hand. At the end of the experiment in the room environment the water was still warmer than human skin, so the warming up was done with a bowl of just boiled water (about 85 degrees). Since the temperature of the water was either decreasing or increasing slowly during all dynamic measurements, rapid upward change in temperature was detectable and easy to match to the measurements from the camera. Matching was done by both observing where the maximum temperature measured by the camera increased, and by looking at the visible images taken by the camera.

Thermal Measurements: We capture radiometric thermal images as a series of Kelvin matrices (320×460) at varying frame rates. For this purpose, we developed an Android application using the FLIR One SDK[4] that allows us to capture images and collect various metrics automatically without requiring user interaction. We additionally collect the sensor *tuning state*, which describes the current calibration phase of the device (Tuned or In Progress). Data points are stored with their

[4]https://developer.flir.com/sdk-documentation/

respective timestamps to the device storage and extracted for analysis after measurements.

Other Measurements: During the thermal image capturing, we also periodically collect system information from the CAT phone. Since we suspect that the automatic calibration performed by the thermal sensor does not account for the changing properties of the shutter, we first tried to model the heat conducted from other components as a result of system load. However, since the internal sensor temperatures used in the shutter-based calibration do not seem to be exposed by any publicly known programmable interface, we had to resort to *battery temperature* values reported by the BatteryManager API in Android instead. This provides a coarse estimate of the overall temperature of the casing on most smart devices.

V. Results

We rigorously evaluate our approach using the measurements described in the previous section. We report only results for 1 fps frame rate as the performance of our model was comparable for other frame rates. We focus on generalization capability of the deep learning model across environments and objects with differing target temperatures. We also compare our approach to a SVR-based predictor to demonstrate the need for deep learning. We also assess how inclusion of training data from similar environments affects deep learning performance. We separately assessed performance using the entire image or parts of it. Best results were consistently obtained by cropping input images to match with the target object. As a consequence, all results, including the original error, have been computed from cropped images.

A. Generalization Across Environments and Objects

We begin our evaluation by examining the performance of our deep learning model in situations where the test environment or test object have differing temperatures than those used to train our model. Specifically, we run leave-one-environment-out (i.e., 2-fold cross-validation considering fridge and room as environments) and leave-one-object-out (i.e., 3-fold cross-validation considering hot, ambient and cold targets) cross-validation experiments.

Table II shows the results of our experiments and compares our deep learning approach against a SVR baseline. The values in the table are in $^{\circ}C$. We separately consider mean, max, standard deviation, and offset of (1) raw data; (2) stabilized data; and data corrected with (3) SVR or (4) our deep learning approach (mean and standard deviation only). The overall error improved from 3.25 to 3.12, and 1.62 to 1.07, respectively. Then we split the data based on the temperature of the experiment object, hot, cold, or room temperature, to perform *Leave One Object Out Validation*. Our approach improves the overall accuracy of the measurement from 1.88 to 1.52 for the ambient temperature object and from 2.50 to 1.69 for the hot object, but it fails for the cold object validation where the overall error grows to 5.63 from 4.15.

B. Performance across Static and Dynamic Objects

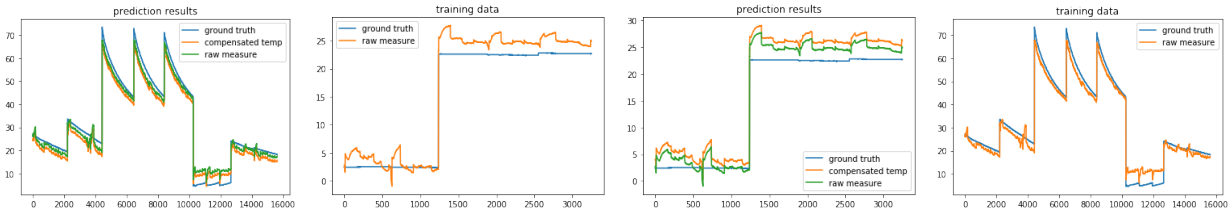
As the next step of evaluation, we consider how well the model generalizes across static and dynamic target objects. We perform this using a 2-fold cross-validation where all data with static objects is considered in one fold, and all data with dynamically changing temperatures in the other.

The results in Table II show that our model fails in both cases as the measurements follow different distributions. The main problem for the model is that the patterns between battery temperature and target object are inconsistent across the evaluation setups. For example, in Fig. V-B, the measured temperature in training data is almost always higher than the ground truth, so it is intuitive for the model to learn this pattern, which is the other way around in most of the validation data. Therefore our approach fails to predict the error.

To improve on performance, we separately assessed how incorporating training data from dynamic environments would help our deep learning model. In this case the overall error decreased from 2.55 to 2.05, however our approach still fails on some experiments and pushes it further away from the ground truth, as in Fig.V-B. To summarize, our model is capable of recalibrating thermal camera measurements and to improve their accuracy when the target objects are approximately stationary. When the device is heating, the performance consistently improves. However, when device is cooling down, performance gains remain smaller. In the case of dynamic objects, the performance suffers unless data from similar environments is incorporated into training. This suggests that a reasonable calibration model could be trained with a number of conditions, incorporating both static and dynamic measurements.

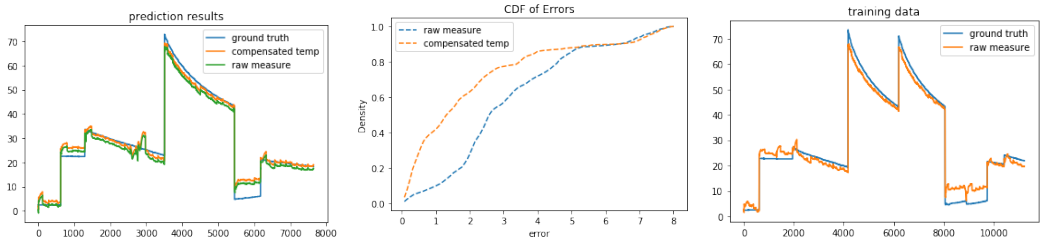
C. SVR Baseline Comparison

As a baseline for our deep learning model, we run the support vector regression (SVR) prediction model with a linear kernel function where battery temperature is used as a predictor, and consider the offset as the response variable. We chose SVR as it has been successfully used in a wide range of sensor-based machine learning tasks [16], [17]. For cross validation, we use the similar setting than with deep learning model. We compare the predicted offset and the true offset and analyze the difference by calculating mean absolute error, a standard procedure for prediction performance analysis. The results in Table II show the mean, maximum, and standard deviation of the error for (1) original data, (2) stabilized data, (3) error for SVR-corrected estimates and (4) deep learning based estimates, respectively. For both methods we used leave-one-environment-out cross-validation (see Sections V.A and V.B) to estimate the errors. We can see that SVR improves the errors for some of the experiments, but fails to improve from the cleaned mean in the Room vs Fridge, Dynamic vs Static and Cold+Room temperature vs Hot experiment cases. The main issue with SVR is that it tends to overfit on the difference between battery temperature and target object temperature. As a result, SVR provides good results when a consistent pattern can be identified, but fails in other cases. While error could be



(a) (b) (c) (d)

Fig. 8. When we leave dynamic or static temperature objects out of training, the y axis shows the temperature in Celsius degree and the x axis represents timestamps where several experiments are concatenated in the plot, (a) shows the compensated temperature given by our approach, compared with raw measure and ground truth for leave dynamic out validation, (b) illustrates training data of leave dynamic out, (c) shows the compensated temperature for leave static out, and (d) illustrates training data of leave static out.



(a) (b) (c)

Fig. 9. (a) compensated temperature given by our approach, compared with raw measure and ground truth, (b) CDF of error for compensated temp and raw measure, (c) Illustration of training data, when static and dynamic data were included in training and validation. In (a) and (c) x axis shows timestamps while y axis shows temperature in Celsius degree.

reduced in some cases through the use of alternative Kernels or temporal smoothing (e.g., HMM or Kalman filter), the key issue, lack of consistent pattern in input features, would persist. The use of LSTM with attention mechanism allows more complex relationships between input features and output features to be captured, thus providing increased robustness and reducing overfitting, motivating the use of deep learning.

VI. Discussion

Software-based Model: Our experiments were carried out on a CAT S60 smartphone with integrated FLIR camera. The thermal camera is a self-contained subsystem within the device, and for example the calibration process is not exposed to applications running on the device. Our method requires no additional hardware, and can be used without access to the internals of the thermal camera subsystem and hence we expect similar data cleaning and correction procedures to be beneficial for other wearable thermal imaging systems. Note that other devices may require using different input features as proxy for device temperature changes (e.g., voltage changes) instead of relying on battery temperature and CPU usage.

Generality: Our method has been tested with two CAT S60 thermal camera smartphones in hot, cold and ambient conditions. Our method does not rely on any special features of these devices, and it can be applied on any Android device with a thermal camera. Initial experiments with a Caterpillar CAT S61 smartphone in cold environment have indicated similar errors. However, further research exploring errors on a wider range of devices is needed.

Operative Temperature Range: The CAT S60 smartphone specifications mention an operative range of 15–35°C and a scene temperature range of 10–120°C. In some of our measurements, the battery temperature indicated values higher than 35°C, and in the refrigerator measurement, the average temperature of the scene was below 10°C. Despite this, With our error correction method, we were able to correct the performance of the thermal camera to that within the specified ranges. In other words, our method can potentially be used to increase the operative range of FLIR cameras.

Physical Aspects: Our method considers battery temperature as the main input for predicting erroneous offset within thermal images. However, the model does not directly incorporate the various underlying theoretical relations in thermophysics. For example, Stefan-Boltzmann's law explains how the total energy emitted by a blackbody is directly proportional to the fourth power of its temperature, whereas Wien's displacement law states that the blackbody radiation curve peaks at a wavelength inversely proportional to its temperature. Detailed exploration and application of these thermophysical grounds could potentially help further improve the model.

Applications: While the S60 thermal camera can detect temperature contrasts without calibration, the measurement values are not accurate enough without out method. Our method enables accurate (error reduced by up to 61%), continuous monitoring of object temperature without the need for cal-

TABLE II
RESULTS OF PERFORMANCE COMPARISON BETWEEN OUR DEEP LEARNING APPROACH AND SVR

Model and Test	Fridge, Room	Room, Fridge	Static, Dynamic	Dynamic, Static	Hot+Room Temp, Cold	Cold+Room Temp, Hot	Cold+Hot, Room Temp
Orig. error mean	3.83	1.41	2.98	2.98	4.37	2.60	1.84
Orig. error max	12.75	10.38	11.49	11.49	21.03	16.14	4.23
Orig. error std	1.76	1.07	1.55	1.55	2.61	1.50	0.72
Stabilized mean	3.52	1.39	2.76	2.76	3.86	2.38	1.86
Stabilized max	5.99	3.62	4.68	4.68	6.07	5.99	3.87
Stabilized std	0.87	0.76	0.76	0.76	0.88	1.00	0.71
SVR Corrected mean	3.79	9.41	1.85	3.77	5.39	7.10	2.37
SVR Corrected max	6.18	13.67	3.78	5.95	7.69	8.79	4.42
SVR Corrected std	0.84	2.67	0.67	0.81	0.96	0.47	0.76
DL Corrected mean	3.12	1.07	3.69	3.42	5.40	1.65	1.51
DL Corrected std	2.50	1.17	1.10	1.29	2.51	1.04	1.56

ibration hardware at the measurement location. Besides the scenarios covered in the Introduction, there are several other domains that would benefit from our approach. For example, with our method, portable and compact thermal cameras such as the Flir One and the S60 can be used in the professional domain to, e.g., monitor temperatures of goods in transit in the cold chain (temperature-controlled supply chain of, e.g., food), detect problems in cooling of a data center, and monitor cognitive load. In addition, consumer applications include using the device as a thermometer, as a warning system for hot surfaces in the kitchen, mapping the airflow within a house, and checking the condition of door and window insulation through prolonged monitoring as the outside temperature changes during the day and night.

VII. RELATED WORK

This paper focuses on calibration and processing of the FLIR thermal images in order to eliminate errors caused by measurement devices themselves or any ambient background heat. Mobile FLIR cameras can utilize periodic calibration by using the camera's shutter as an external blackbody equivalent, as presented by Nugent et al. [13], but in the case of mobile devices, ambient temperatures also have to be considered as a part of the calibration process. Lin et al. [18] take into account changes in ambient temperature as a reason for fast changes in FPA temperatures, and present a method where the camera output is corrected based on FPA temperature instead of using the shutter. This however requires access to FPA temperatures. Shutter-less calibration of uncooled thermal cameras is also explored by Bieszczad et al. [19] who leverage the mean FPA temperature for offset calculation and correction. Cao et. al [20] present an online, in the wild image processing method for estimating the current FPA temperature utilizing offline per-pixel calibrations, in which they find a temperature minimizing the intensity variance within a thermal image. However, this relies on the assumption that effects of ambient heat are always displayed as non-uniformity in the image.

Mobile thermal sensing: Mobile devices are known to be heat sources as well as temperature sensors. Gurrum et al. [21] study temperature measurements for different smartphone hardware components, such as battery and display. Xie et al. [22] mention especially CPU and battery as the major heat generating components. Therminator [23] simulates the relationship between temperature of hardware components and skin layers of smartphones. To summarize, hardware components, such as CPU, GPU, and battery, can produce a temperature varying from 30 to 50 Celsius degrees in normal room conditions. When considering a mobile device as a thermal sensor, one must take into account the heat generated by the device itself in addition to the ambient temperature.

Background temperature control: Certain background-subtraction techniques have been used to distinguish areas of interest; these may involve, for example, fusion with visual images and thermal images, or object recognition [24], [9]. Also, background temperature have been controlled by using a heated plate [25]. Problem of these techniques relate that they cannot dismiss the thermal effect of the background heat to the actual interest area. By considering how thermal images, especially ones taken by FLIR cameras, should be adjusted and processed, we can produce reliable images where variation of the background heat can be taken into account.

VIII. SUMMARY AND CONCLUSION

We contributed by developing a novel calibration technique for improving the measurement quality of FLIR cameras. In our approach, battery temperature, CPU usage, and other internal parameters were used to learn a mapping that captures the effect of device temperature change on errors in the thermal measurements. By compensating the values of the FLIR camera with error estimates, the values returned by the camera can be corrected and the performance of the thermal imaging significantly improved. We demonstrated the effectiveness of our approach through benchmark experiments carried out in carefully controlled hot and cold conditions. Results of our experiments demonstrated significant improvements in both the accuracy and robustness of thermal monitoring, reducing the maximum error by 61.0% on average.

ACKNOWLEDGMENT

This research has been financially supported by Academy of Finland grants 317875, 297741, 296139, and 303825, and 6Genesis Flagship (grant 318927). This article only reflects the authors' views.

REFERENCES

[1] M. L. Mauriello, L. Norooz, and J. E. Froehlich, "Understanding the role of thermography in energy auditing: Current practices and the potential for automated solutions," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ser. CHI '15. New York, NY, USA: ACM, 2015, pp. 1993–2002. [Online]. Available: http://doi.acm.org/10.1145/2702123.2702528

[2] M. L. Mauriello, M. Saha, E. B. Brown, and J. E. Froehlich, "Exploring novice approaches to smartphone-based thermographic energy auditing: A field study," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ser. CHI '17. New York, NY, USA: ACM, 2017, pp. 1768–1780. [Online]. Available: http://doi.acm.org/10.1145/3025453.3025471

[3] O. Faust, U. R. Acharya, E. Ng, T. J. Hong, and W. Yu, "Application of infrared thermography in computer aided diagnosis," *Infrared Physics & Technology*, vol. 66, pp. 160–175, 2014.

[4] M. Jaspers, M. Carrière, A. Meij-de Vries, J. Klaessens, and P. van Zuijlen, "The flir one thermal imager for the assessment of burn wounds: Reliability and validity study," *Burns*, vol. 43, no. 7, pp. 1516–1523, 2017.

[5] J. Alametsä, M. Oikarainen, J. Viik, and J. Perttunen, "Improving nursing methods by using thermal imaging: observations by cat s60 mobile phone," *Finnish Journal of eHealth and eWelfare*, vol. 9, no. 2-3, pp. 74–81, 2017.

[6] P. Christiansen, K. A. Steen, R. N. Jørgensen, and H. Karstoft, "Automated detection and recognition of wildlife using thermal cameras," *Sensors*, vol. 14, no. 8, pp. 13 778–13 793, 2014.

[7] A. Kumar and G. P. Hancke, "A zigbee-based animal health monitoring system," *IEEE Sens. J*, vol. 15, no. 1, pp. 610–617, 2015.

[8] P. Nurmi and S. Tarkoma, "Low-cost support for search and rescue operations using off-the-shelf sensor technologies," *Electronics Letters*, vol. 53, no. 15, pp. 1011–1013, 2017.

[9] Y. Abdelrahman, E. Velloso, T. Dingler, A. Schmidt, and F. Vetere, "Cognitive heat: Exploring the usage of thermal imaging to unobtrusively estimate cognitive load," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 1, no. 3, pp. 33:1–33:20, Sep. 2017. [Online]. Available: http://doi.acm.org/10.1145/3130898

[10] M. M. Khan, R. D. Ward, and M. Ingleby, "Infrared thermal sensing of positive and negative affective states," in *Robotics, Automation and Mechatronics, 2006 IEEE Conference on*. IEEE, 2006, pp. 1–6.

[11] H. Flores, J. Hamberg, X. Li, T. Malmivirta, A. Zuniga, E. Lagerspetz, and P. Nurmi, "Evaluating energy-efficiency using thermal imaging," in *Proceedings of The 20th Annual International Workshop on Mobile Computing Systems and Applications*, ser. HotMobile '19. ACM, 2019.

[12] F. A. Rosell, G. L. Harvey, J. B. Goodell, W. R. Lawson, and J. A. Ratches, "The fundamentals of thermal imaging systems." NAVAL RESEARCH LAB WASHINGTON DC, Tech. Rep., 1979.

[13] P. W. Nugent, J. A. Shaw, and N. J. Pust, "Radiometric calibration of infrared imagers using an internal shutter as an equivalent external blackbody," *Optical Engineering*, vol. 53, no. 12, p. 123106, 2014.

[14] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern recognition*, vol. 26, no. 9, pp. 1277–1294, 1993.

[15] F. Karim, S. Majumdar, H. Darabi, and S. Chen, "Lstm fully convolutional networks for time series classification," *IEEE Access*, vol. 6, pp. 1662–1669, 2018.

[16] J. He and A. Arora, "A regression-based radar-mote system for people counting," in *2014 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, March 2014, pp. 95–102.

[17] Y. Yang, J. Luo, J. Hao, and S. J. Pan, "Counting via led sensing: Inferring occupancy using lighting infrastructure," *Pervasive and Mobile Computing*, vol. 45, pp. 35–54, 2018.

[18] D. Lin, P. Westfeld, and H.-G. Maas, "Shutter-less temperature-dependent correction for uncooled thermal camera under fast changing fpa temperature," *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, p. 619, 2017.

[19] G. Bieszczad, T. Orżanowski, T. Sosnowski, and M. Kastek, "Method of detectors offset correction in thermovision camera with uncooled microbolometric focal plane array," in *Electro-Optical and Infrared Systems: Technology and Applications VI*, vol. 7481. International Society for Optics and Photonics, 2009, p. 74810O.

[20] Y. Cao and C.-L. Tisse, "Shutterless solution for simultaneous focal plane array temperature estimation and nonuniformity correction in uncooled long-wave infrared camera," *Appl. Opt.*, vol. 52, no. 25, pp. 6266–6271, Sep 2013. [Online]. Available: http://ao.osa.org/abstract.cfm?URI=ao-52-25-6266

[21] S. P. Gurrum, D. R. Edwards, T. Marchand-Golder, J. Akiyama, S. Yokoya, J. F. Drouard, and F. Dahan, "Generic thermal analysis for phone and tablet systems," in *2012 IEEE 62nd Electronic Components and Technology Conference*, May 2012, pp. 1488–1492.

[22] Q. Xie, J. Kim, Y. Wang, D. Shin, N. Chang, and M. Pedram, "Dynamic thermal management in mobile devices considering the thermal coupling between battery and application processor," in *2013 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, Nov 2013, pp. 242–247.

[23] Q. Xie, M. J. Dousti, and M. Pedram, "Therminator: A thermal simulator for smartphones producing accurate chip and skin temperature maps," in *2014 IEEE/ACM International Symposium on Low Power Electronics and Design (ISLPED)*, Aug 2014, pp. 117–122.

[24] J. W. Davis and V. Sharma, "Background-subtraction using contour-based fusion of thermal and visible imagery," *Computer Vision and Image Understanding*, vol. 106, no. 2, pp. 162 – 182, 2007, special issue on Advances in Vision Algorithms and Systems beyond the Visible Spectrum. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1077314206001834

[25] P. K. Matkowski, "Comparative thermal analysis of commercial and novel hybrid thermal greases," in *Proceedings of the 2014 37th International Spring Seminar on Electronics Technology*, May 2014, pp. 64–69.