HeatSight: Wearable Low-power Omni Thermal Sensing

Rawan Alharbi
Northwestern University
Evanston, IL, USA

Chunlin Feng
Northwestern University
Chicago, IL, USA

Sougata Sen
BITS Pilani, Goa Campus
Goa, India

Jayalakshmi Jain
Northwestern University
Chicago, IL, USA

Josiah Hester
Northwestern University
Evanston, IL, USA

Nabil Alshurafa
Northwestern University
Chicago, IL, USA

ABSTRACT

Thermal information surrounding a person is a rich source for understanding and identifying personal activities. Different daily activities naturally emit distinct thermal signatures from both the human body and surrounding objects; these signatures exhibit both spatial and temporal components as objects move and thermal energy dissipates, for example, when drinking a cold beverage or smoking a cigarette. We present HeatSight, a wearable system that captures the thermal environment of the wearer and uses machine learning to infer human activity from thermal, spatial, and temporal information in that environment. We achieve this by embedding five low-power thermal sensors in a pentahedron configuration, which captures a wide view of the wearer's body and the objects they interact with. We also design a battery life-saving mechanism that selectively powers only those sensors necessary for detection. With HeatSight, we unlock thermal as an egocentric modality for future interaction research.

CCS CONCEPTS

• **Human-centered computing → Ubiquitous and mobile computing systems and tools**.

KEYWORDS

human activity detection; thermal sensing; wearable; low-power

ACM Reference Format:
Rawan Alharbi, Chunlin Feng, Sougata Sen, Jayalakshmi Jain, Josiah Hester, and Nabil Alshurafa. 2021. HeatSight: Wearable Low-power Omni Thermal Sensing. In *2021 International Symposium on Wearable Computers (ISWC '21), September 21–26, 2021, Virtual, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3460421.3478811

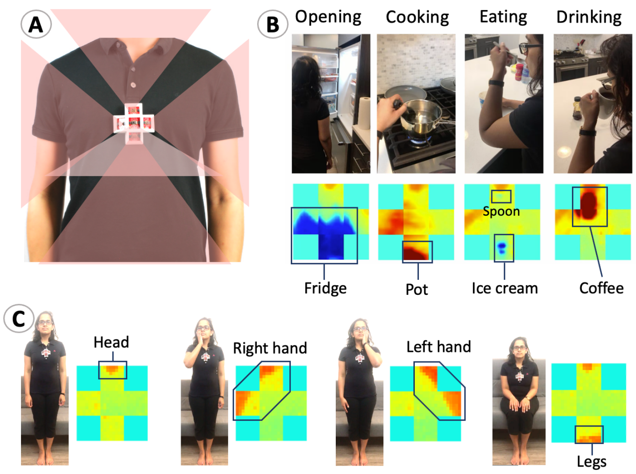


Figure 1: HeatSight is (A) a chest worn wearable system consisting of five independent low-power thermal sensing arrays allowing a thermal omni capture (180°) of the wearer's surroundings including: (B) the thermal radiation emitted from objects around the wearer, and (C) the human thermal signature emitted from the wearer's body over time. These thermal human-object interaction signatures can be used for human activity detection.

1 INTRODUCTION

Thermal sensors provide physical temperature estimates, generating unique heat signatures of objects across time and space. As a wearable modality, they enable tracking multiple parts of the human body and its interaction with objects in the environment without the need for direct contact. Unlike light-based cameras, thermal sensors do not require scene illumination. Moreover, low-resolution thermal sensors are more privacy-preserving as they capture less information. Thermal sensors can also be covered with opaque thermal passing materials and integrated into existing accessories, increasing the device's wearability [4]. However, thermal sensors' field of view (FoV) is usually too small to capture objects in the scene reliably, and changing the FoV of the thermal camera is an expensive process.

In this paper, we present HeatSight, a new wearable system for human activity detection (HAR) based on thermal sensing. The system utilizes passive infrared (thermal) radiation, which naturally radiates from humans and objects they interact with. HeatSight comprises five low-power thermal sensors configured in a pentahedron on the chest, offering a 180° , large, egocentric (frontal body) FoV of the wearer. HeatSight can capture the wearer's frontal body along with their surrounding objects, providing a rich stream of thermal, temporal, and spatial information for machine-learning-driven human activity recognition models. HeatSight balances between two competing system requirements: (1) the need for a large FoV to

ACM ISBN 978-1-4503-8462-9/21/09…$15.00
https://doi.org/10.1145/3460421.3478811

recognize hand-related activities; and (2) the need to save energy to obtain all-day battery life. To address this FoV-energy tradeoff, we developed a triggering mechanism that dynamically turns on and off thermal sensors as the hand and objects appear in the FoV.

To assess HeatSight's feasibility in classifying human activity, we conducted an in-lab experiment with participants performing 14 daily life activities that are of interest to the wearable research community. Our results show that HeatSight achieves a classification accuracy of 85.21%, and when compared with keeping all sensors on, our triggering mechanism doubles battery lifetime.

Overall, we believe this work demonstrates how thermal can be used to detect human activities from an egocentric point of view, creating new possibilities for human activity context-aware applications.

2 RELATED WORK

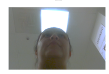
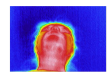
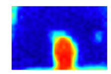
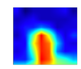
Thermal sensors are used in many applications. Larson et al. captured heat traces of human hands using an *overhead* thermal camera to understand surface interactions [10]. Kawashima et al. detected high-level daily and abnormal activities (e.g., walking or falling) by placing low-resolution thermal cameras on the ceiling [9]. To increase the FoV of the thermal camera, Shirazi et al. [18] and Abdelrahman et al. [1] utilized thermal reflection to capture mid-air gesture interactions that take place outside the FoV of the thermal cameras. High-resolution thermal cameras with a *frontal* view of the human face have been used to measure cognitive load [2] and stress [3]. In some applications, a *handheld* thermal camera is more useful as it allows a close-up or targeted capture of an object. For example, Cho et al. demonstrated that thermal cameras attached to smartphones could improve automatic recognition of indoor materials (e.g., brick, carpet) [5]. Handheld thermal cameras can also assist novice users in house inspection and thermographic-based energy auditing [11–13]. GlimpseData utilized a low-resolution thermal camera to filter out unwanted scenes (e.g., scenes with human presence) from an RGB life-logging camera [7]. Despite the many applications of thermal sensors, multiple thermal sensing arrays have not been thoroughly investigated as a *wearable* modality for activity detection and interaction monitoring, primarily due to high energy requirements.

3 HEATSIGHT IMPLEMENTATION

3.1 Thermal Sensor Selection

There are several types of thermal sensors available in the market; we explored three high- to low-power/resolution thermal sensors including: (1) high-power/resolution FLIR camera [6], (2) medium-power/resolution MLX camera [14], and (3) low-power/resolution Grid-EYE camera [16]. Table 1 compares the parameters of the three thermal sensors along with a sample image. The images were obtained by wearing each of the sensors on the chest with the lens pointing up. We also show the RGB image for reference. In all the thermal images, the outline of the human head is apparent, even in the case of the low-resolution Grid-EYE thermal sensor, making all three viable candidates for HAR. However, upon assessing the parameters tradeoff space presented in Table 1, we eliminated FLIR due to its high current draw, which makes it unsuitable for capturing data for an entire waking day with a small-sized battery. Given our desire to capture wide FoV thermal information surrounding the wearer, we opted to use the Grid-EYE thermal sensor, given wearability and battery life considerations.

Table 1: Comparison of parameters of three thermal sensors. HeatSight uses an array of Grid-EYE thermal sensors.

	FLIR	MLX	Grid-EYE
Resolution (p)	160x120	32x24	8x8
Range (μm)	8-14	8-14	8-13
FOV ($^\circ$)	57	110x75	60x60
Sensitivity ($^\circ C$)	0.05	0.1	0.16
Current (mA)	45.5	23	4.5
Sample Image			
			

3.2 System

Sensors: We connect five low-resolution thermal sensors (Grid-EYEs) to a development board (Artemis Nano) using a custom PCB shield. Each thermal sensor captures an 8×8 thermal image, which we then combine to produce a 24×24 thermal image. The microcontroller controls the sensors' data acquisition and stores time-stamped sensor data in a micro SD card. The average sampling rate is 10 Hz.

Encapsulation: We designed, and 3D printed a case to encapsulate HeatSight. The case contains a base that holds the sensors at the desired angle (140° to achieve a 180° FoV). The system is worn on the body using a neck strap and strong magnets clasping onto the wearer's garments to ensure correct positioning.

Calibration: Since we collect data from five thermal sensors, it is crucial to calibrate all sensors (i.e., they should produce the same output when facing the same object). Thermal cameras are calibrated using a camera shutter or cover. We designed a calibration cover with a known temperature. We place this cover over each HeatSight for 1 minute to obtain the offset error and correct each sensor using the offset. This calibration process is performed once for each HeatSight device.

Current Consumption: We used an ultra-low-power development board (Artemis Nano). The current draw of the system in *full power mode* (all sensors activated) is 27 mA, while in *trigger mode* (some sensors in sleep mode), it ranges from 7 mA (4 sensors in sleep mode) to 22 mA (1 sensor in sleep mode). Currently, the system is connected to a 3.7-V LiPo battery with a capacity of 500 mAh that enables the system prototype to remain functional for 18 hours in a single charge.

3.3 Human Activity Detection

We utilize Convolutional Neural Networks [17] to learn discriminative features from each activity in our dataset. Our network comprises two convolutional layers (depth = 8 and 16) and two fully connected layers (n = 784 and 100, dropout = 0.3). In each convolutional layer, we perform the convolution operator (kernel

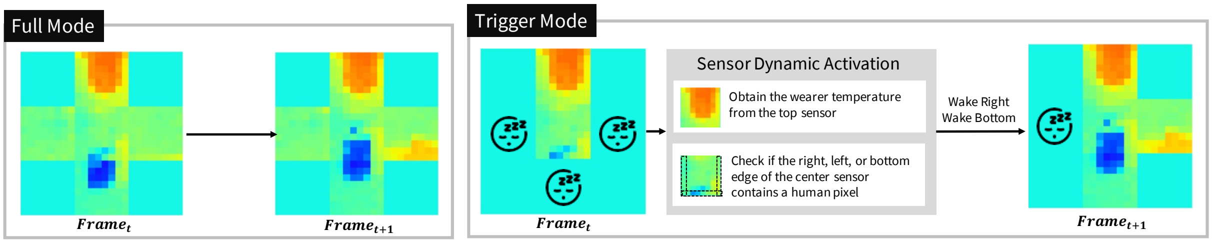


Figure 2: HeatSight modes of operation: (1) Full Mode when all of the sensors are activated and (2) Trigger Mode when we use our sensor selective activation algorithm to reduce power consumption. (Video figure attached)

size = 1 and 3, stride = 1), batch normalization [8], rectified linear unit activation [15], and max pooling [19]. The input to our model is a $24 \times 24 \times 35$ thermal image. For each time-step, we obtain a 24×24 image by stitching all of the 8×8 images obtained from each thermal sensor. This enables the model to learn spatial and thermal information across all sensors. To capture temporal information, we use a 3.5-second window by stacking 35 24×24 consecutive thermal frames (as channels). We normalize the thermal image to speed up the training of the neural network. We built our model using PyTorch and trained it from scratch using 100 epochs.

3.4 Reducing Energy Draw via Trigger Mode

Not all five sensors will provide useful information about an activity. Each activity relies on data coming from a subset of the sensors (e.g., the prediction of eating might only require information from the right and top sensors). Putting the non-essential sensors to sleep will reduce HeatSight's energy consumption. We consider a selective activation strategy in which a pre-defined favored sensor is always activated (e.g., top and center) and triggers other sensors as needed to predict activities more accurately. We developed a triggering heuristic to selectively control the sensors' wake and sleep state.

Our selective sensor activation (SSA) algorithm predicts the sensors needed to create $Frame_{t+1}$ using information from the top and center sensors at time t. First, we obtain the wearer's human temperature range from the top sensor as the head is always in the FoV of the top sensor by design. We then define *human pixels* to be a range within ∓ 6 of the wearer temperature. Then we check the right, left, and bottom edge (8 pixels each) of the center sensor to see if a human pixel exists at an edge. If a single human pixel exists, we activate the nearest sensor to that edge. In Figure 2, we show an example of consecutive images in Full Mode and in Trigger Mode (i.e., using the SSA algorithm). The image shows a person about to drink a cold beverage. $Frame_t$ in the Trigger Mode mode has two sensors activated (awake) while other sensors are asleep. In $Frame_t$, we determine what sensors to activate in $Frame_{t+1}$. The right and the bottom senor in $Frame_{t+1}$ is activated because the right and bottom edges of the center sensor in $Frame_t$ contained at least one human (wearer) pixel. In contrast, the left sensor remains asleep in $Frame_{t+1}$ as there is no human pixel on the left edge of the center sensor in $Frame_t$. Our SSA algorithm is simple enough to be implemented on a microcontroller. Moreover, the time required to run our algorithm and wake the selected sensors is less than the sampling interval of sensor data acquisition (0.1 seconds), preventing data loss.

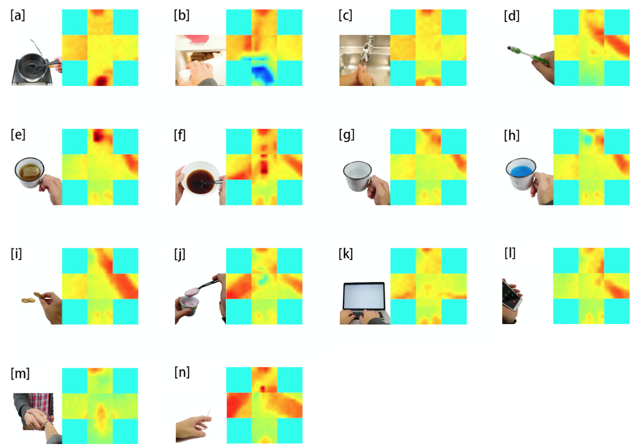


Figure 3: Activities performed by volunteers wearing HeatSight (see video): [a] stirring hot water in a pot, [b] opening a fridge, [c] washing hands, [d] brushing teeth, [e] drinking hot beverage, [f] eating soup, [g] drinking room temperature beverage, [h] drinking cold beverage, [i] eating room temperature crackers, [j] eating cold yogurt, [k] typing laptop, [l] answering a phone call, [m] meeting and greeting someone while standing, and [n] smoking a cigarette.

4 EVALUATION

4.1 Data Collection

We recruited six volunteers (three women and three men) and asked them to wear the devices while performing a set of hand-related activities that are part of daily human activities and that are of interest to the wearables research community (see Figure 3 for the list of activities). Participants performed the complete set of activities in the lab once per round. They were asked to perform each activity for at least 15 seconds for each round. We used a timer with audible sounds to remind the participants when to start or stop an activity. In total, participants performed three rounds and took a 2-minute break after completion of each round. The experiment lasted 30 minutes per participant.

4.2 Activity Detection in Full Mode

We trained a model using data from the first and third rounds using all participants' data and then tested it on data collected from the second round for each participant. This type of evaluation can help in assessing if the system sensing modality provides enough information for the model to learn and classify the different activities performed at different times. We obtained a mean f-score of 85.21%. Table 2 shows the f-score of each activity using the Full Mode.

4.3 Activity Detection in Trigger Mode

To quantify the effect of our sensor activation approach on HeatSight energy and human activity prediction, we simulate sensor activation using the same algorithm applied retrospectively on the data we collected from our volunteers to generate a new dataset. Simulation is necessary in our case because we want to compare both Full and Trigger Modes on the same dataset, allowing us to thoroughly assess the effect of triggering on activity detection. We then feed this dataset into the same activity detection model we built earlier for testing without retraining. Table 2 shows the f-score of each activity under Trigger mode.

Specifically, we run the SSA algorithm at each time t to determine the sensor status at time $t + 1$ and create $Frame_{t+1}$. We save the activation status for each sensor in a CSV file indexed by time. We then use this information to create a new dataset for each participant. We check the sensor activation state for each sensor in a time step. If it is 0 (non-active), we discard the frame for that sensor at that time step. Otherwise, we keep the data. CNN-based models, however, expect a constant input size. Therefore, we have to fill the missing frames of inactive sensors. We replace the missing sensor frames with a frame that has the median temperature value. We then use this newly generated dataset as an input to the HAR model that we generated previously under Full Mode to get the prediction results.

Table 2 shows the f-score for the Full and Trigger Modes. From the table, we observe that for most activities, the drop in f-score is less than 15%, with the overall average drop in f-score due to triggering being 9.73%.

4.4 Resource Usage

We use the generated CSV (as described above) to estimate the power cost of the entire system at each time-step, from which we can then estimate the Joules used and battery lifetime.

$$\sum_{t=0}^{n} N_t \times I \times V \times T$$

At each time interval t, we obtain the number of sensors that are activated, N_t, and multiply it by the current I, voltage V, and time elapsed in seconds T for one Grid-EYE (we obtain I's value from the datasheet, which is 0.0045 A).

In Table 3, we show the difference in energy consumption between the Full and Trigger Mode by running the same SSA algorithm on the whole dataset. We also present the effect of energy reduction on battery lifetime for each participant (including the participant's non-activity frames). Overall, we observed that energy consumption reduced by 49.35%, resulting in an average battery lifetime gain of 3.13 hours (SD=1.4 hours) in Trigger Mode.

Table 2: F-score in Full and Trigger Mode. Full Mode assumes that all five sensors are activated, while Trigger Mode activates the sensor based on our SSA algorithm

Activity	Full Mode	Trigger Mode	Delta
stir_hot	100.00	90.85	-9.15
open_fridge	97.92	89.07	-8.85
wash_hand	90.82	78.41	-12.41
social	99.46	90.73	-8.73
screen_laptop	91.96	67.83	-24.13
drink_hot	93.05	82.66	-10.39
eat_hot	94.92	88.33	-6.58
drink_room	41.01	48.20	7.19
drink_cold	91.67	77.64	-14.03
eat_room	52.99	50.99	-2.00
eat_cold	89.88	76.85	-13.03
brush_teeth	96.45	74.58	-21.87
call	69.10	61.45	-7.65
smoke	83.66	79.05	-4.61
Average	85.21	75.47	-9.73

Table 3: Reduction in power consumption in Trigger Mode

	Full Mode (J)	Trigger Mode (J)	Percentage reduction	Battery (+H)
P1	220.28	93.10	57.74	+3.81
P2	136.80	68.56	49.89	+2.95
P3	140.92	48.84	65.35	+4.87
P4	116.76	77.00	34.06	+1.66
P5	174.27	125.36	28.07	+1.29
P6	122.35	47.71	61.01	+4.24

5 CONCLUSION

Wearable low-power omni thermal sensing can unobtrusively capture the movement of the wearer's body parts as they interact with objects, providing rich information to model many everyday human activities. Here we present HeatSight, a low-powered chest-worn system that utilizes multiple arrays of low-power thermal sensors to detect human activity. HeatSight can capture a 180° egocentric thermal view of the wearer's body along with any heat-emitting object around the wearer. The system produces a 24×24 thermal image at 10Hz, which allows the capture of thermal, spatial, and temporal information to detect human activity. We deploy our prototype on multiple participants who wore HeatSight and performed fourteen activities of everyday living. Our model shows 85.21% classification accuracy across all activities. We design a selective sensor activation algorithm that uses triggering heuristics to reduce energy. We believe that our work highlights an underutilized sensing modality that enables a wide range of applications. HeatSight provides an invisible, and private mechanism for recognizing human activity, further enabling context-aware applications.

REFERENCES

[1] Yomna Abdelrahman, Alireza Sahami Shirazi, Niels Henze, and Albrecht Schmidt. 2015. Investigation of material properties for thermal imaging-based interaction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 15–18.

[2] Yomna Abdelrahman, Eduardo Velloso, Tilman Dingler, Albrecht Schmidt, and Frank Vetere. 2017. Cognitive heat: exploring the usage of thermal imaging to unobtrusively estimate cognitive load. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 33.

[3] Fatema Akbar, Ayse Elvan Bayraktaroglu, Pradeep Buddharaju, Dennis Rodrigo Da Cunha Silva, Ge Gao, Ted Grover, Ricardo Gutierrez-Osuna, Nathan Cooper Jones, Gloria Mark, Ioannis Pavlidis, et al. 2019. Email Makes You Sweat: Examining Email Interruptions and Stress Using Thermal Imaging. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 668.

[4] Rawan Alharbi, Tammy Stump, Nilofar Vafaie, Angela Pfammatter, Bonnie Spring, and Nabil Alshurafa. 2018. I Can't Be Myself: Effects of Wearable Cameras on the Capture of Authentic Behavior in the Wild. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3 (2018), 90.

[5] Youngjun Cho, Nadia Bianchi-Berthouze, Nicolai Marquardt, and Simon J Julier. 2018. Deep Thermal Imaging: Proximate Material Type Recognition in the Wild through Deep Learning of Spatial Surface Temperature Patterns. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 2.

[6] Flir. 2021. High definition thermal cameras. https://www.flir.com/. Last Accessed: 01/15/2021.

[7] Seungyeop Han, Rajalakshmi Nandakumar, Matthai Philipose, Arvind Krishnamurthy, and David Wetherall. 2014. GlimpseData: Towards continuous vision-based personal analytics. In *Proceedings of the 2014 workshop on physical analytics*. ACM, 31–36.

[8] Sergey Ioffe and Christian Szegedy. 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167* (2015).

[9] Takayuki Kawashima, Yasutomo Kawanishi, Ichiro Ide, Hiroshi Murase, Daisuke Deguchi, Tomoyoshi Aizawa, and Masato Kawade. 2017. Action recognition from extremely low-resolution thermal image sequence. In *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*. IEEE, 1–6.

[10] 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*. ACM, 2565–2574.

[11] Matthew Louis Mauriello, Brenna McNally, and Jon E Froehlich. 2019. Thermporal: An Easy-to-Deploy Temporal Thermographic Sensor System to Support Residential Energy Audits. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 113.

[12] Matthew Louis Mauriello, Leyla Norooz, and Jon E Froehlich. 2015. 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*. ACM, 1993–2002.

[13] Matthew Louis Mauriello, Manaswi Saha, Erica Brown Brown, and Jon E Froehlich. 2017. 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*. ACM, 1768–1780.

[14] melexis. 2020. Far infrared thermal sensor array (32x24 RES). https://www.melexis.com/en/product/MLX90640/Far-Infrared-Thermal-Sensor-Array/. Last Accessed: 01/15/2021.

[15] 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.

[16] Panasonic. 2020. Grid-eye Infrared Array Sensor. https://na.industrial.panasonic.com/products/sensors/sensors-automotive-industrial-applications/lineup/grid-eye-infrared-array-sensor/. Last Accessed: 01/15/2021.

[17] Waseem Rawat and Zenghui Wang. 2017. Deep convolutional neural networks for image classification: A comprehensive review. *Neural computation* 29, 9 (2017), 2352–2449.

[18] 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 SIGCHI Conference on Human Factors in Computing Systems*. ACM, 3483–3492.

[19] Dominik Scherer, Andreas Müller, and Sven Behnke. 2010. Evaluation of pooling operations in convolutional architectures for object recognition. In *International conference on artificial neural networks*. Springer, 92–101.