

Efficient Material Selection for Training Occluded mmWave Radar-Based Gesture Recognizers

Nuwan T. Attygalle^{1,*,\dagger}, Luis A. Leiva², Matjaž Kljun³, and Klen Čopić Pucihar^{3,*}

¹Université catholique de Louvain, Belgium

²University of Luxembourg, Luxembourg

³University of Primorska, Slovenia

*Corresponding authors: nuwan.attygalle@uclouvain.be, klen.copic@famnit.upr.si

^{\dagger}Work done while affiliated with University of Primorska, Slovenia.

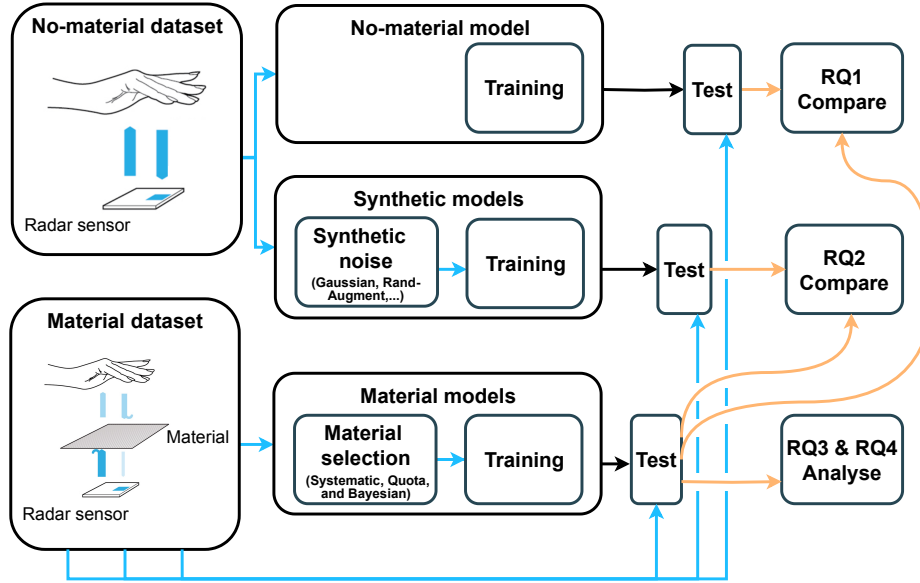


Figure 1. Methodology overview of radar-based gesture recognition through materials.

ABSTRACT

Radar-based gesture recognition has emerged as a promising approach for unobtrusive interaction. Unlike camera-based systems, radar sensors can detect gestures through opaque materials, enabling seamless embedding into various everyday objects. However, it remains unclear how to train models efficiently for robust gesture recognition through diverse materials. To investigate this, we collected a dataset of 17,520 gesture recordings performed through 73 everyday materials. By comparing several material-sampling and data-augmentation strategies, we found that a small carefully selected representative subset of training materials was sufficient to match performance of a classifier trained on the full material dataset. Our results showed that the models trained on 14 quota-sampled materials achieved accuracies of 95.8% and 91.2%, comparable to training on all 73 materials (96.8% and 91.6%) and significantly better than training without material data (66.8% and 65.8%). Among the evaluated sampling approaches, Quota sampling also provided the best overall trade-off between performance and practicality. In contrast, classifiers trained on augmented data performed worse than those trained on actual material-specific data. Taken together, these findings indicate that, for the tested sensor, gesture set, and material collection, carefully selected real-material data offer a practical route to reducing material-specific data collection in radar-based gesture recognition while preserving generalisation. Code, models, and data are available in the public repository, with additional details provided in the supplementary materials: <https://gitlab.com/hicuplab/seeing-through>.

Keywords: millimetre-wave radar; gesture recognition; material selection; sampling; synthetic noise.

1 Introduction

While traditional interfaces such as keyboards, mice, and touchscreens remain prevalent, alternative modalities are broadening the Human-Computer Interaction (HCI) landscape. Among them, mid-air gesture recognition is gaining popularity due to its intuitive nature, contactless control, and potential for seamless integration into pervasive computing environments. However, gesture sensing systems face challenges with feedback, contextual awareness, and robustness under real-world conditions [1–3], especially when the sensing hardware is embedded in everyday objects [4–6].

Embedding radar sensors introduces added complexity, as the radar signal must propagate through the object’s material twice (on transmission and on reflection) before reaching the receiving antenna. Signal attenuation depends on the material’s electromagnetic properties, particularly its transmission coefficient (T_c). Although prior work has shown that material properties affect signal quality and recognition accuracy [7–9], these studies often examined few materials or overlooked the effect of training with material-specific data.

Deep learning classifiers typically improve with exposure to a larger and more diverse training set, until a saturation point is reached [10–13]. Accordingly, a gesture classifier trained on gesture recordings of a radar signal captured through a wide range of materials (material-specific data) is expected to outperform a classifier trained with limited or no materials. However, re-training a classifier each time the gesture set changes is costly, as it requires re-capturing data through all materials. This highlights the need for more efficient training strategies, which is an open research challenge.

To this end, we conducted a systematic study on radar-based gesture recognition through materials, aiming to optimise the training process. Specifically, we investigate how to train deep learning models that are both accurate and resource-efficient by identifying a minimal yet representative set of training materials. Prior work [7, 9] has shown that the transmission coefficient is a strong and practical descriptor for predicting performance degradation through materials and is easier to acquire than alternatives such as Signal to Noise Ratio (SNR) (see Section 2.3; [7]). Building on these findings, we hypothesise that T_c can guide material selection.

Using a millimetre-wave (mmWave) radar sensor (Google Soli), we collected a comprehensive dataset of 17,520 gesture recordings through 73 everyday materials. Six gestures were each repeated 40 times per material (i.e., *Material dataset*). We evaluated several training strategies, including synthetic noise injection and three sampling techniques (systematic, quota, and Bayesian) each applied with two objective functions (i.e., *Significance-based* and *Multiple-criteria-based*) for determining the optimal number of training materials. These experiments address the following research questions:

RQ1: Does training with *Material dataset* improve performance?

RQ2: Can training with synthetic noise injected into *No-material dataset* achieve comparable performance to training with the *Material dataset*?

RQ3: How can training materials be intelligently selected?

RQ4: How few training materials are needed to achieve comparable performance to training on all available materials in the *Material dataset*?

Our results show that training with material-specific data yields significantly better performance than training without it. These gains cannot be replicated through synthetic noise injection, which highlights the importance of training with material data. Notably, under our experimental conditions, models trained on data from 14 materials, selected using quota sampling based on their T_c values, can match the performance of models trained on all 73 materials. Overall, our results show that material-property guided data selection reduce training-data requirements, suggesting scalable deployment of radar-based gesture recognition systems.

2 Background and Related Work

2.1 Radar sensing technologies

Radar sensing uses radio-frequency (RF) waves to detect motion by transmitting pulses via a Tx antenna and analysing the echoes received by an Rx antenna to extract object properties such as velocity and radar cross-section [14]. Unlike RGB camera or infrared (IR) based systems, radar is unaffected by lighting conditions [15] and can penetrate non-metallic materials [9], making it suitable for sensing through occlusions. It has thus been applied in gait analysis [16, 17], movement recognition [16, 18–20], sleep and respiration monitoring [21, 22], and gesture-based interaction [20, 23–29], across frequencies from 2.4 GHz [22, 29] to 24 GHz [21, 25].

Millimetre-wave radars operating at 50–70 GHz offer high spatial resolution, making them ideal for detecting fine-grained finger movements and complex hand gestures [30, 31]. Their compact size, low power consumption, and scalable manufacturing

further support integration into miniaturised devices [30]. These sensors can capture subtle, nonrigid 3D gestures, including overlapping finger motions, with high precision [32].

Prior work has demonstrated effective gesture recognition for micro-gestures [32, 33], object manipulation [31], hybrid gesture interaction [34], and applications ranging from augmented reality [35] to musical expression [36, 37]. Some studies have also shown that mmWave radars can differentiate materials placed directly over them [38, 39]. While these studies often assume clear line-of-sight or minimal occlusion, real-world embedded applications require robust recognition through diverse materials. However, the impact of sensing *through* everyday materials on gesture recognition performance has not been systematically investigated. This paper addresses this gap by examining how occluding materials affect mmWave radar gesture recognition and by proposing strategies for selecting a minimal set of training materials to reduce data collection while preserving accuracy.

2.2 Sensing through materials

When sensing through materials, the radar signal must propagate through an occluding layer before capturing motion. While technologies like acoustic and capacitive sensors support such interactions [40–42], radar offers distinct advantages, including higher spatial resolution [43], touchless interaction (unlike capacitive sensors), and longer-range detection (unlike acoustic sensors). The emergence of low-cost on-chip sensors such as Google Soli,¹ and UltraSense TouchPoint,² together with low-power embedded microcontrollers such as TI’s MSP430³ and advancements in deep learning have enabled rich gesture interactions [33, 44], often outperforming acoustic [45–49] and capacitive approaches [50–52]. Despite these advances, only a few studies systematically explore radar-based gesture sensing through materials.

Existing work on RF-based through-material sensing focuses largely on detecting poses, objects and people behind walls [53–57] or common materials [28, 58, 59], and coarse activity recognition on surfaces [60, 61]. Some efforts aim to correct signal distortions via denoising autoencoders [62–64] and Convolutional Neural Networks (CNNs) to counteract radar interference [65–67]. However, fine-grained gesture recognition through materials remains under explored. Only a few studies attempt this [7, 8, 58], and none systematically use material-specific training data, which our work addresses.

2.3 The effect of materials on radar signal quality

When sensing through materials, the radar signal must propagate through an occluding material twice. Since different materials have different characteristics, the signal will be affected in different ways. The effect can range from signal attenuation (low energy) [68] to signal distortion (shape change) [69, 70]. These changes can affect gesture recognition performance.

Čopić Pucihar et al. [7] investigated how to characterise materials with respect to radar-based gesture recognition performance and, in doing so, addressed the attenuation-versus-distortion question. They quantitatively compared several candidate descriptors (e.g., transmission coefficient T_c , SNR, received energy, and amplitude) and found that T_c (capturing attenuation) is the strongest predictor of recognition performance degradation, substantially outperforming SNR (Spearman’s $\rho = 0.86$ for Transmission Coefficient vs. $\rho = 0.62$ for SNR).

Moreover, obtaining reliable, repeatable measurements of signal distortion characteristics for FMCW mmWave radars, such as phase/frequency-dependent effects across the sensor bandwidth (57–64 GHz) typically requires specialized instrumentation (high frequency network vector analysers) and calibration procedures, which are expensive and not widely available. Building a training method that depends on such hard-to-acquire measurements would substantially limit the practical applicability of the approach in real-world deployments.

Based on the above reasons, we selected T_c as a practical proxy of material descriptor for guiding the material selection and training design.

2.4 Material selection

Since training a gesture classifier using every possible material is impractical, researchers typically rely on sampling or optimisation methods to choose a representative set. The effectiveness and generalisability of the model depend on how well this subset reflects the broader material population.

2.4.1 Sampling techniques

We studied two sampling approaches to select an optimum set of materials per sampling approach for training gesture recognisers: Systematic (or Random) and Quota sampling. *Systematic sampling* is a probabilistic technique that ensures a non-biased, representative selection of a larger population [71] and is valued for its simplicity and effectiveness [71, 72]. While *Quota sampling* is a non-probabilistic technique enabling a more targeted selection of materials based on specific characteristics [73], such as the T_c in our case. By clustering materials according to their T_c and selecting samples from these

¹<https://www.norgram.co/work/google-soli-website>

²<https://ultrasensesys.com/touchpoint/>

³<https://www.ti.com/microcontrollers-mcus-processors/msp430-microcontrollers/overview.html>

clusters (e.g., choosing the centroid or the nearest material), Quota sampling ensures that our subset of materials covers the full spectrum of T_c values in the training dataset.

2.4.2 Optimisation algorithms

Optimisation algorithms are essential for solving problems where an objective function can be defined. Some of these algorithms include Grid Search [74], Random Search [75], Genetic Algorithms (GAs) [76], Tree-structured Parzen Estimator (TPE) [77], and Bayesian Optimisation (BO) [18, 76, 78, 79].

A comparative study by Liashchynskyi et al. [80] indicated that GAs often outperform Grid and Random Search in terms of efficiency and effectiveness. Moreover, BO has been noted for its superior performance over GAs [81] and TPE [82] due to its ability to efficiently explore and exploit the search space [18], and in particular under limited resources or high-dimensional parameter spaces [83, 84]. We therefore use BO as a strong optimisation-based baseline for material selection in our study (hereafter referred to as Bayesian optimisation sampling), and we use it to sample materials and compare its performance against the two sampling techniques described above.

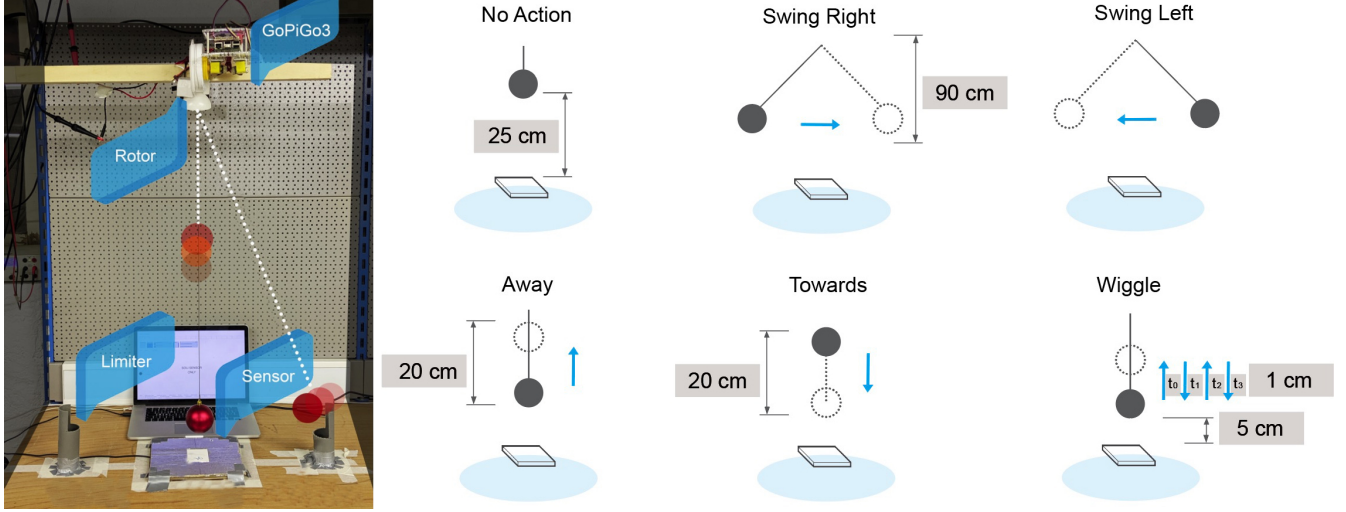


Figure 2. Experimental setup and the gesture set. Left figure shows the experiment setup and the right figure shows the executed gestures with the robot. On the right figure filled and dashed circles denote the initial and final position of the ball respectively. The same gesture set was used as in [7].

3 Experiments

We chose six distinct gestures (Figure 2) informed by a survey of previous studies using mmWave radar sensing for interaction [7, 33]. Further, since it is crucial to execute gestures with minimal articulation variation, we replicated the mechanical system proposed in prior work [7].

We experimented with two different deep learning architectures: a 3D convolutional neural network (referred to as *Conv3D*) and a 2D convolutional neural network using spectrograms (referred to as *Spectrogram*) derived from previous work [7, 85] with one key difference when preprocessing data. Here, Range Doppler [43, 86] images from 4 Rx antennas are not averaged, but concatenated into one sequence. Refer to Section 1 of Supplementary Materials for more details about architectures.

We trained all classifiers with the Adam optimiser with learning rate $\eta = 0.001$ and decay rates $\beta_1 = 0.9$, $\beta_2 = 0.999$. We set the maximum number of epochs to 400, and also set an early stopping criteria of 100 epochs, which means that training stops if the validation loss does not improve after 100 epochs, and the best classifier weights are retained.

The early stopping threshold was set to a high number of epochs due to the grokking phenomenon observed during training [87], where the model initially overfit, showing a large gap between training and validation accuracy, before eventually converging to near-perfect performance on both. This transition from overfitting to generalization required extended training time.

3.1 Data collection

We collected the radar signals from 40 repetitions of six gestures sensed through 73 everyday materials [7], resulting in a total of 17,520 recordings. The materials used were identical to those reported in the material catalogue⁴ by Čopič Pucihar et al.

⁴<https://solidsonsoli.famnit.upr.si/>

[7]. We primarily focused on non-conductive materials such as fabrics, plastics, and wood, as radar waves do not effectively penetrate conductive materials like metals. For instance, the KAPA plate, which contains a thin aluminium layer (0.02 mm thick), exhibits a transmission coefficient of just 0.14, indicating that 86% of the signal amplitude is attenuated. Such materials are therefore unsuitable for radar sensing applications and were not included in material training. Additionally, we recorded 200 repetitions of each gesture without any material occluding the sensor, totalling 1,200 recordings.

Previous research has demonstrated that Range Doppler [43] images can be used for robust gesture recognition [2, 8, 29, 88], so we opted to use them in this experiment as well. Each gesture is represented as a sequence of Range Doppler images, where each image depicts the reflected energy intensity as a function of target range (radial distance from the radar) and velocity (direction and magnitude of the target). Positive velocity corresponds to motion away from the radar (increasing range), while negative velocity indicates motion toward the radar (decreasing range). One Range Doppler image was computed for each transmitter-receiver antenna pair.

We used the Google Soli sensor, a Frequency-Modulated Continuous-Wave (FMCW) radar operating in the 57–64 GHz range. The sensor was configured following our prior work [4], with a sampling rate of 1000 Hz, Doppler range sensitivity set to $[-16, 0]$, and the built-in adaptive clutter filter disabled. Experiments were conducted using Soli’s default mode, which activates one transmit (Tx) and four receive (Rx) antennas.

3.2 Statistical analysis

We first consider parametric statistical tests (e.g., paired t-tests, repeated measure ANOVA) and verify their assumptions by testing normality with the Shapiro–Wilk test and homogeneity of variance with Levene’s test. When assumptions are violated, we use non-parametric alternatives, such as the Friedman test followed by post-hoc Wilcoxon signed-rank tests. We apply Bonferroni correction to control for multiple comparisons. For pairwise comparisons, we use two-tailed tests by default, and one-tailed tests only when the hypothesis is directional. In particular, one-tailed tests are used in subsection 3.6, where we estimate the optimal number of training materials using a significance-based procedure.

3.3 Data partitioning

We aimed to evaluate several training strategies, including how incorporating material-specific data affects classification performance. To ensure that the effects observed were due to the introduction of material-specific data, it was crucial to control for other confounding variables, such as the size of a training partition. Therefore, we needed a sophisticated data partitioning method that introduced material-specific data without increasing the size of the training partitions. This was achieved by reducing the number of repetitions for each material.

To accomplish this, we first limited the data partition size to 630 repetitions across the six gestures. Given the amount of data available for each material, this was the maximum partition size that allowed us to build a classifier using a minimum of three materials. We consider this size appropriate because it is comparable to the 840 repetitions used in previous work for training and validating classifiers with similar architectures and gesture sets [4, 7]. This decision was further supported by pilot experiments, where classifiers trained and evaluated with 840 and 630 repetitions, respectively, showed comparable performance levels despite the difference in partition size.

The data partitioning process was conducted in two stages. In Stage 1, we split the data to create the test partition. In Stage 2, the remaining data was used to generate training and validation partitions. These partitions were created randomly, but the specific configuration depended on the number of training materials and the method used for material selection. Despite varying the number of training materials, the sizes of the training and validation partitions remained consistent. This was achieved by adjusting the number of repetitions per gesture we included in each partition. This process is illustrated in detail in Figure 3.

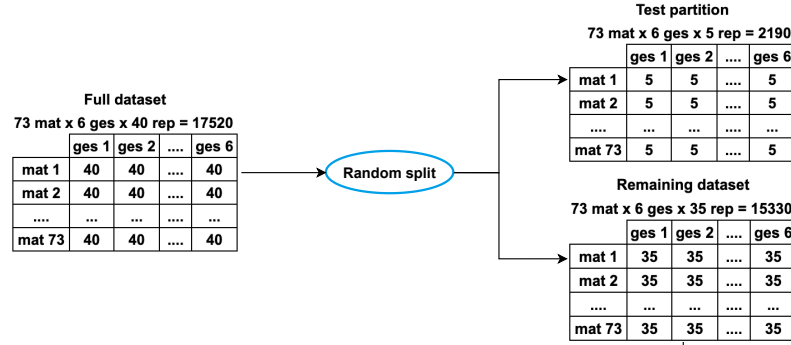
3.4 Training with synthetically generated noise (RQ1 & RQ2)

To determine whether training with materials improve performance (**RQ1**) and whether similar performance improvements can be achieved through synthetic noise [89] generation (**RQ2**), we trained and compared the classifiers under six conditions:

- *No-material* classifier trained only on data where the sensor was not occluded by materials (i.e., *No-material dataset*).
- *All-materials* classifier trained with all 73 materials (i.e., *Material dataset*).
- *Gaussian-noise* classifier, where Gaussian noise was added to training partition (i.e., *No-material dataset* + *Gaussian-noise*).
- *RandAugment* classifier where training partition was modified through stochastic data augmentation routine specially designed for image classification problem (i.e., *No-material dataset* + *RandAugment*).
- *RandAugment with Gaussian-noise* classifier, where training partition was first transformed with Gaussian-noise and then augmented through stochastic data augmentation (i.e., *No-material dataset* + *RandAugment with Gaussian-noise*).

Stage 1

Repeat once



Stage 2

Repeat once per model (22 times)

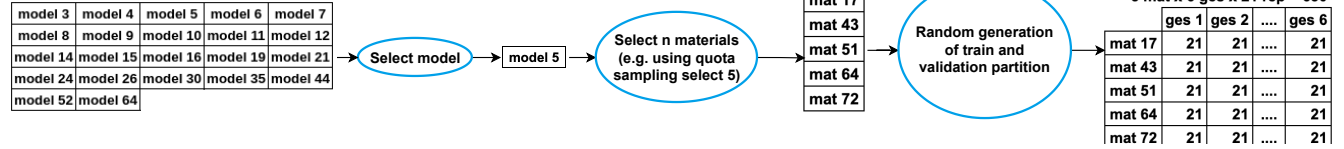


Figure 3. Data Split. In the Stage 1, the entire dataset (73 materials $\times$ 6 gestures $\times$ 40 repetitions) was randomly divided to create a test partition, which consisted of 73 materials $\times$ 6 gestures $\times$ 5 repetitions. All trained classifiers were evaluated on this unseen test partition. In Stage 2, the remaining data was used to generate training and validation partitions. These partitions were created randomly, but the specific configuration depended on the number of training materials and the method used for material selection. For example, the figure illustrates partitioning for a classifier trained with 5 training materials selected using *Quota sampling*. Despite varying the number of training materials, the sizes of the training and validation partitions remained consistent. This was achieved by adjusting the number of repetitions per gesture we include in each partition.

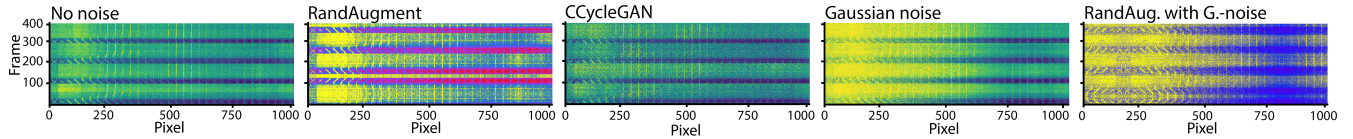


Figure 4. Spectrogram images with injected synthetic noise using different methods. The x-axis represents 1024 px which are flattened from a 32×32 Range Doppler image. The y-axis represents frames as a time series: We stack 4 time series images on top of each other, where each representing one receiving antenna.

- *CCycleGAN* classifier, where training partition was first transformed with Conditional CycleGAN we specifically trained for this purpose (i.e., *No-material dataset* + *CCycleGAN*).

The partition sizes were kept consistent though all conditions and aforementioned training procedures were followed. To make our evaluation more robust, we retrained our classifiers three times, each time generating new data partitions according to a different random seed. We report average classification performance [7], a combination of AUC (Area Under the ROC Curve) and Weighted Accuracy, defined as $\text{AvgPerf} = (\text{Accuracy} + \text{AUC})/2$. This combined metric captures both how often predictions are correct and how well the model's confidence scores distinguish the correct gesture from the others. Both Accuracy and AUC scores are weighted by class support. In supplementary materials (section 6) we additionally report the individual Accuracy, AUC, and Micro-F1, to provide further insights into the performance results. Details regarding synthetic noise generation are provided in the next section.

3.4.1 Synthetic noise generation

All data augmentations in this study were performed on spectrogram images, created by flattening 32×32 Range Doppler images into vectors of 1,024 values and stacking them, which resulted in a 400×1024 px image (Figure 4 top-left). For the *Gaussian-noise* classifier, we generated Gaussian noise with a mean of 0.001 and a randomly selected noise level, choosing a standard deviation within the $[0,1]$ range. This approach simulated the varying noise level for materials with varying T_c (i.e., the lower the T_c , the higher the noise level and vice versa). Next, we applied a linear gradient mask starting at a random pixel value and ending at the 512th pixel (to target the approximate volume around the sensor where gestures were performed). Finally, the

generated noise was added to the spectrogram image, and all values were clipped to the $[0,1]$ range. An example of a modified image can be seen in [Figure 4](#).

When preparing data for the *RandAugment* classifier, we used a stochastic data augmentations [90]: value range bounds were set between 0 and 1, with 5 augmentations per image and a magnitude of 0.5 [90]. We considered 10 transformations: AutoContrast, Equalize, Solarize, Color, Contrast, Brightness, ShearX, ShearY, TranslateX and TranslateY. An example of a modified image can be seen in [Figure 4](#).

For the *RandAugment with Gaussian-noise* classifier, we first applied Gaussian noise to the image and then used *RandAugment* on the noisy image (seen [Figure 4](#)). The procedures for both steps followed the methods described earlier.

Finally, for the *CCycleGAN* classifier, we trained a Conditional CycleGAN (CCycleGAN) [91] to convert *No-material dataset* samples to *Material dataset* samples. CCycleGAN is an extension of the popular CycleGAN [92] architecture, where the generation process is conditioned on class labels. Specifically, we introduced as labels the gesture type and the material type to the learned (image-based) latent space, so that the decoder could generate spectrograms of specific gestures performed either with or without materials, using unpaired input and output images. The model hyperparameters were maintained consistent with those outlined in the original CCycleGAN paper [91] to ensure a fair alignment with established benchmarks. An example of the modified image can be seen in [Figure 4](#).

3.4.2 Results

The results for the Conv3D architecture demonstrate that the *All-materials* classifier achieves the best classification performance across the entire range of T_c (see the green line in [Figure 5](#) top-left). For the Spectrogram architecture, the *All-materials* classifier performs best over most of the T_c range (see the green line in [Figure 5](#) bottom-left).

The box plots on the right of [Figure 5](#) illustrate the classifier performance for both Conv3D and Spectrogram architectures, including: *No-material*, *All-materials*, *Gaussian-noise*, *RandAugment*, *RandAugment with Gaussian-noise* and *CCycleGAN* classifiers. We performed normality and equal variance assumption analyses and identified violations of these assumptions (see [Section 2](#) in Supplementary Materials for details). Consequently, we conducted non-parametric tests, including the Friedman test and post-hoc Wilcoxon signed-rank tests with Bonferroni-corrected p-values. The results of these tests indicate that the *All-materials* classifier significantly outperformed all other classifiers for both architectures ([Figure 5](#) right), suggesting that synthetic data augmentation did not achieve the performance levels of a classifier trained with noise induced by real materials.

Furthermore, in the Conv3D architecture, post-hoc tests revealed that none of the classifiers trained with synthetically augmented data outperformed the *No-material* classifier. In contrast, for the Spectrogram architecture, synthetic data augmentation led to significant performance improvements. The *Gaussian noise* classifier ($W = 320, r = -0.73, p < .001$) *RandAugment* classifier ($W = 529, r = -0.55, p < .01$) and the *CCycleGAN* classifier ($W = 456, r = -0.65, p < .001$) all showed significant improvements over the *No-material* classifier.

These results show that diversifying our data by training with material-specific data improves classification performance, making our classifiers more robust to material occlusion. However, it also indicates that the tested augmentation methods failed to approximate the performance of a classifier trained with materials, especially when aiming to approach the performance of the *All-materials* classifier. However, training with material-specific data requires a lot of resources and effort for acquiring the data and retraining classifiers (i) every time the gesture set changes, or (ii) when we want to sense though a new set of materials. Therefore, in the following section, we focus on optimising this process by minimising the amount of material-specific data required while trying to match the performance of our best classifier (i.e., *All-materials* classifier).

3.5 Analysis of sampling techniques (RQ3)

We experimented with three techniques defined in [subsubsection 2.4.1](#): *Systematic*, *Quota*, and *Bayesian* optimisation sampling. All 73 materials were characterised based on their T_c values reported in the materials catalogue published by Čopić Pucihar et al. [7].

3.5.1 Systematic random sampling

All materials were first sorted in ascending order according to their T_c values. Next, we randomly selected the starting point and then every n^{th} element, where n depended on the number of materials we needed to sample from the set. For example, to select 100 elements from a population of 1000, we would randomly choose a starting point in the range between 1 and 10 and then select every 10th element. In our case, this selection process was repeated 22 times for each unique number of materials we experimented with.

3.5.2 Quota sampling

We selected materials based on a budget. For example, to select 100 elements from a population of 1000, we would divide the population into 100 element categories using clustering and then select one element from each category. In our study, we applied K-means clustering to group the 73 materials based on their T_c values into n clusters, where n corresponded to the

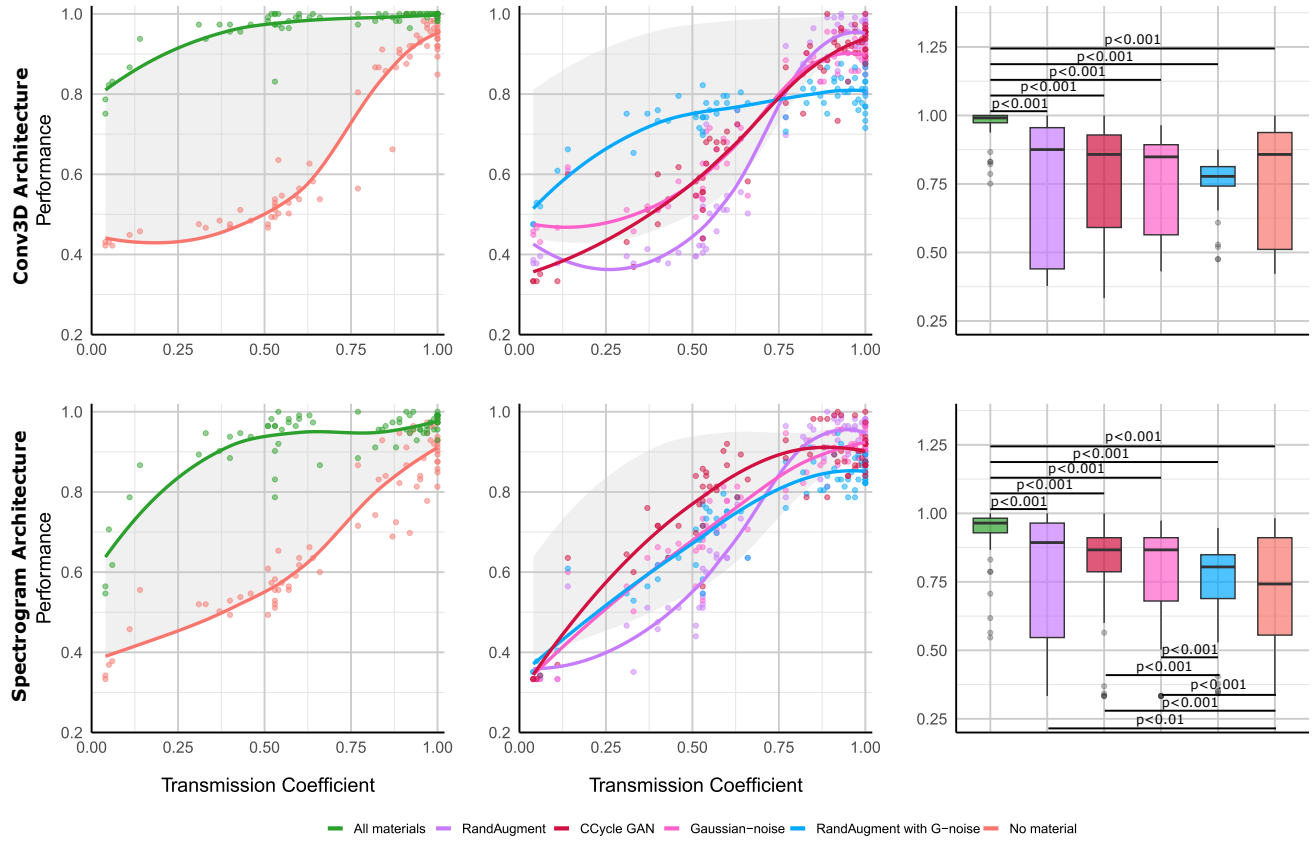


Figure 5. Comparing classification performance when training with synthetically generated noise. We compare *All-materials*, *No-material*, *Gaussian-noise*, *RandAugment*, *RandAugment with Gaussian-noise* and *CCycleGAN* classifiers for Conv3D (top) and Spectrogram (bottom) architectures. For Friedman and Wilcoxon signed-rank post-hoc statistics refer to *Section 3* of Supplementary Materials.

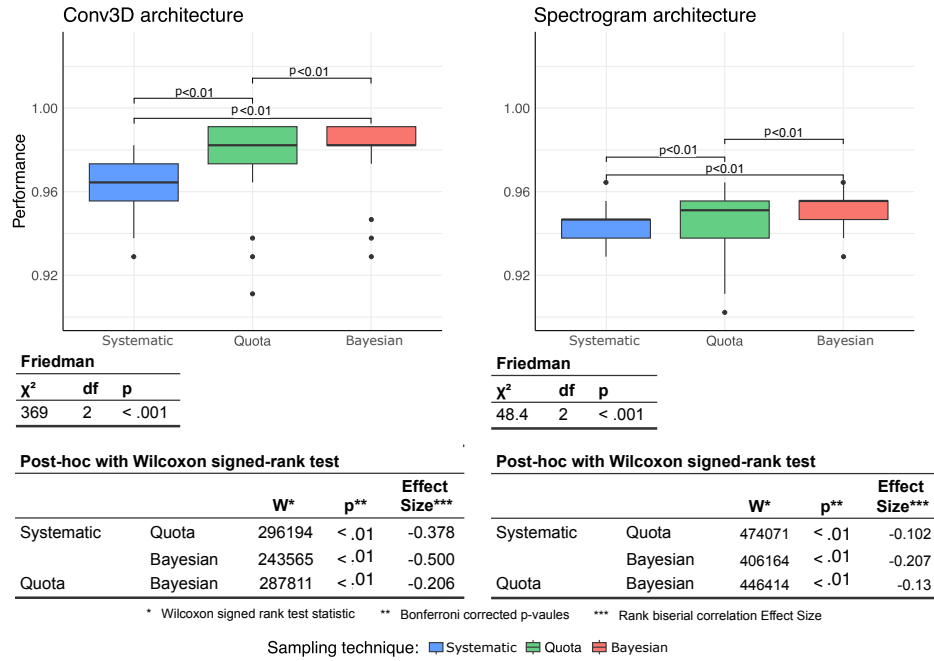


Figure 6. Comparing classification performance using *Systematic*, *Quota*, and *Bayesian* optimisation sampling for both Conv3D and Spectrogram architectures. Each distribution is represented by the average performance scores of 22 classifiers across 73 materials, totalling 1606 data points. The top graphs display box plots of performance, while the bottom tables presents Friedman statistics and post-hoc Wilcoxon signed-rank test results.

number of materials we needed to sample. Once the clusters were formed, we selected the material closest to the centroid of each cluster. This selection process was repeated 22 times, once for each unique number of materials we experimented with.

3.5.3 Bayesian optimisation sampling

We treated material selection as an optimisation problem. We used T_c values to define the search space, aiming to identify a set of training materials with T_c values that will maximise classification performance. We implemented BO with the GPyOpt package [93], adhering to its default hyperparameters. Due to the high computational cost (i.e., ~ 45 minutes per Conv3D classifier on Quadro RTX 6000 GPU) we limited, based on the literature, the maximum number of iterations to 15. Because Bayesian optimization requires repeated model training and evaluation, and requires access to a full material dataset, it is computationally expensive and complicates practical deployment. We therefore treat it as a theoretical upper bound rather than a recommended practical strategy.

Bayesian optimisation is known for its effectiveness in converging to near-optimal solutions with appropriately chosen priors and acquisition functions. For example, Snoek et al. [79] demonstrated that the Branin-Hoo function, a standard benchmark in Bayesian optimisation, achieved near-optimal convergence within the first 15 iterations. They also observed rapid convergence when tuning nine hyperparameters of a convolutional neural network (CNN). While Wang et al. [94] reported near-optimal solutions for a body part classifier within 15-20 iterations. Our experiments also showed that performance typically plateaued at around 15 iterations, thus we chose this hyperparameter.

3.5.4 Evaluation method

Each sampling technique was evaluated by training 22 classifiers, each varying in the number of materials used for training. The choice of materials varied depending on the sampling technique. We maintained a consistent data partition size and used the same test partition for all experiments, as detailed in Figure 3. To ensure robustness in our evaluation, we retrained each classifier three times, generating new data partitions for each run. For each iteration, we recorded the classification performance, defined as $\text{Performance} = (\text{Accuracy} + \text{AUC})/2$, across the 73 materials. We conducted normality and equal variance assumption analyses and identified assumption violations (see Section 2 in Supplementary Materials for details). Consequently, we assessed differences among the sampling techniques using non-parametric tests, including the Friedman test and post-hoc Wilcoxon signed-rank tests, with Bonferroni-corrected p-values.

3.5.5 Results

In Figure 6 we present the overall classification performance of classifiers trained with a reduced set of materials (e.g., 3, 4, 5, 6, ..., 64), with material selection done using *Systematic*, *Quota*, and *Bayesian optimisation sampling* for both Conv3D and Spectrogram architectures. The results show noticeable differences in classification performance between the sampling techniques (see Figure 6 top graphs). The Friedman test revealed significant differences for both Conv3D ($\chi^2(2) = 369$, $p < .001$) and Spectrogram ($\chi^2(2) = 48.4$, $p < .001$) architectures. Subsequent pairwise comparisons indicated that *Bayesian optimisation sampling* consistently outperformed the other techniques, while *Systematic sampling* performed the worst. This trend was observed for both architectures (Figure 6 bottom tables), although the effect sizes were smaller for the Spectrogram architecture, suggesting less pronounced differences. Furthermore, the effect sizes for the differences between *Bayesian optimisation* and *Quota sampling* were also relatively small ($r = -0.13$ and $r = -0.206$ for Conv3D and Spectrogram respectively), indicating smaller differences compared to other sampling pairs.

3.6 Selecting the optimal number of materials (RQ4)

We anticipated that different sampling techniques will exhibit varying saturation points, with higher-quality sampling methods leading to faster saturation and superior overall performance. Thus, we started our experimentation by identifying saturation points for different sampling techniques. Next, we sought to better understand which materials were selected and whether there exists a set of influential materials that best improves the performance of the classifier.

3.6.1 Objective functions

We considered two objective functions: *Significance-based* and *Multiple-criteria-based*. In *Significance-based* function the optimal number of training materials is the smallest number at which no significant difference in performance exists between the selected classifier and the *All-materials* classifier trained with all materials (i.e., *All-materials* classifier). The only downside of this approach is that it relies on a single parameter. In the *Multiple-criteria-based* function the optimal number of training materials is the number of materials included in the best-performing classifier, considering four selection parameters: (1) median performance (higher is better); (2) variance of performance (lower is better); (3) number of materials (lower is better); and (4) significance level p (higher is better).

3.6.2 Evaluation method

To determine the optimal number of materials according to *Significance-based* objective function, we first conducted assumption tests to select the appropriate statistical methods. We then performed a single-tailed test, adjusted with Bonferroni correction for multiple comparisons, to compare the classification performance of the *All-materials* classifier with 22 classifiers trained on reduced sets of materials (*N-materials*, where $N = 3, 4, 5, 6, \dots, 64$). This test reveals whether the *All-materials* classifier significantly outperforms the *N-materials* classifiers, where the alternative hypotheses is $H_a : \text{diff} > 0$, with $\text{diff} = \text{Perf}_{\text{All-materials}} - \text{Perf}_{\text{N-Materials}}$. As soon as this is not the case, there is no performance gain of adding more material-specific data, and we can assume that the saturation point was reached.

To determine the optimal number of materials according to *Multiple-criteria-based* objective function, we used a Multi-Criteria Decision-Making (MCDM) algorithm. Among various MCDM techniques we chose the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS), since it is a widely used method [95] that is transparent and easy to understand [96]. In this context, the ideal solution would involve the fewest materials (e.g., 3), the highest median performance (e.g., 100%), and the minimum variance in performance (e.g., 0). Conversely, the negative ideal solution would consist of the maximum number of materials (e.g., 64), the lowest performance, and the highest variance (e.g., 1).

3.6.3 Results: Optimal number of materials

The results in Figure 7 illustrate how classification performance evolves with an increasing number of materials for various sampling techniques and classifier architectures (Conv3D and Spectrogram). These results validate our assumption that classification performance improves with an increasing number of training materials, but eventually plateaus as the classifier saturates. Additionally, we observe a reduction in performance variance as the number of materials increases. Notably, the *Bayesian optimisation sampling* (Figure 7 right column) shows a smoother performance increase compared to *Systematic* and *Quota sampling* (Figure 7 left and middle columns, respectively). This trend is consistent across both Conv3D and Spectrogram architectures.

At the top of each plot in Figure 7 we present Bonferroni-corrected p-values from a Wilcoxon signed-rank one-tailed test. These p-values indicate whether the *All-materials* classifier significantly outperforms classifiers trained with a reduced set of materials, referred to as *N-materials* classifiers, where $N = 3, 4, 5, 6, \dots, 64$. Non-parametric statistics were employed due to violations of normality and equal variance assumptions (see Section 2 of Supplementary Materials for details).

By analysing the reported p-values, we can determine that the saturation point varies depending on the sampling technique. For example, the results of the Conv3D architecture shown in the top row of Figure 7 reveal that in the case of *Systematic sampling*, the *All-materials* classifier significantly outperformed all classifiers until 64-materials classifier; in the case of *Quota*

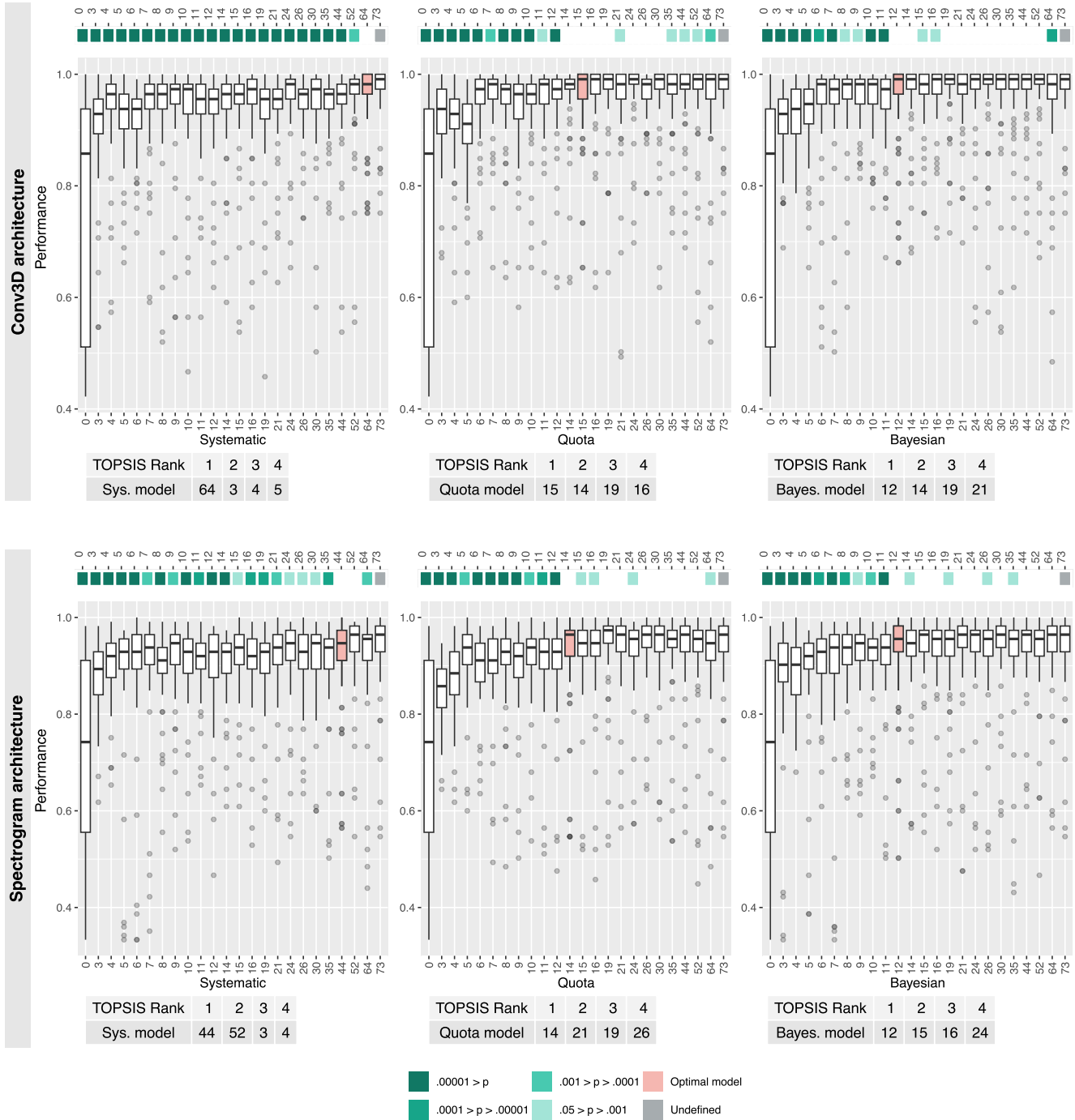


Figure 7. Change of gesture-classification performance as the number of training materials increases. Columns compare the three material-selection methods (*Systematic sampling*, *Quota sampling*, *Bayesian optimisation sampling*) and rows compare the two model types (Conv3D, Spectrogram). At the top part of each figure, we present the results of a Wilcoxon signed-rank one-tailed test comparing the *All-materials* classifier (trained with 73 materials) against classifiers trained with reduced sets of materials (N -materials: $N = 3, 4, 5, \dots, 72$). This test evaluates whether the reduced-material model performs significantly worse. The orange-highlighted point marks the smallest number of materials whose performance is not significantly worse than the 73-material baseline (our selected “best practical” model). The table below each plot lists the TOPSIS ranks of the four best reduced-material models to show which selections perform best.

sampling 14-materials classifier; and in the case of *Bayesian optimisation sampling* the 12-materials classifier. Therefore, we can assume the saturation points are 64, 14, and 12 materials for *Systematic*, *Quota* and *Bayesian optimisation sampling* respectively. Similarly, we can see the saturation points for Spectrogram architecture which are: 44, 14, and 12 for *Systematic*, *Quota* and *Bayesian optimisation sampling* respectively. Again, the results are very similar for both architectures.

The reported p-values further highlight the superiority of *Bayesian optimisation* and *Quota* over *Systematic sampling*, as the former two techniques reached the saturation point with fewer materials. Classifiers at these saturation points can be considered optimal because they achieve the performance of the *All-materials* classifier while using the smallest possible number of training materials.

The results from the MCDM using the TOPSIS method (shown in the tables at the bottom of each plot in Figure 7) align well with the previous significance-based selection approach. The optimal classifier, as determined by the highest TOPSIS rank, is the same in all but one case: *Quota sampling* for the Conv3D architecture (row 1, column 2 in Figure 7). In this instance, the 15-materials classifier is identified as optimal, rather than the 14-materials classifier, which is ranked second. Despite this small discrepancy, the results are consistent overall, confirming the robustness of both the significance-based and MCDM approaches.

3.6.4 Results: Chosen materials

The results for Conv3D and Spectrogram architectures in Figure 8 compare the optimal classifiers of sampling techniques (determined in subsection 3.6.3) with corresponding classifiers trained with the same number of materials but a different sampling technique. For example, *Bayesian 12-materials* classifier is considered optimal classifier for *Bayesian optimisation sampling* and is compared with *Systematic 12-materials* classifier and *Quota 12-materials* classifier. We provide the same analysis for the Conv3D and Spectrogram architectures in Section 4 of Supplementary Materials with Wilcoxon signed-rank test results comparing classifier performances.

The top histograms in Figure 8 show the distribution of selected materials across transmission coefficients (T_c) for various sampling techniques. *Systematic sampling* shows a notable gap in high transmission coefficient ranges, particularly in the 12-materials and 14-materials classifiers, indicating potential under-representation of materials with high T_c . Conversely, *Quota sampling* achieves a balanced selection across the full T_c spectrum, ensuring comprehensive coverage and contributing to a more robust performance. *Bayesian optimisation sampling* exhibits a tendency to favour materials with higher T_c , which may enhance classifier performance by emphasising materials with significant influence. These patterns highlight how the choice of sampling technique affects the representation of transmission coefficients and, consequently, the overall effectiveness of the classifiers.

The middle plots of trend lines in Figure 8 illustrate how classifier performance varies with the transmission coefficient of materials. In most cases, the optimal classifier consistently outperforms the others, except for the *Systematic 44-materials* classifier with Spectrogram architecture (see Section 4 of Supplementary Materials). Statistical analysis, including Friedman and post-hoc Wilcoxon signed-rank tests with Bonferroni-adjusted p-values (see Section 4 of Supplementary Materials), confirms that the optimal classifier generally surpasses the others, except for the one mentioned above where *Bayesian 44-materials* classifier outperformed *Systematic 44-materials* classifier ($W = 593, p < 0.05, r = -0.412$). Additionally, for classifiers with a high number of materials (e.g., 64-materials), performance variance is minimal because these classifiers include a large portion of materials, making the choice of sampling technique less critical.

4 Discussion, limitations, and future work

Using a mmWave radar sensor we collected a comprehensive dataset comprising 17,520 gesture recordings through 73 everyday materials. We evaluated several training strategies, including synthetic noise injection and three sampling techniques (systematic, quota, and Bayesian) each applied with two objective functions for determining the optimal number of training materials (i.e., Significance-based and Multiple-criteria-based functions). These experiments address four research questions we discuss hereafter.

4.1 Training with real vs. synthetic material data (RQ1, RQ2)

The results presented in Figure 5 show strong evidence that training with materials improves performance. Our *All-materials* classifier, trained on data collected through 73 different materials, significantly outperformed the *No-materials* classifier addressing RQ1. This result is also in line with preliminary work of Leiva et al. [8]. However, did the classifier improve because it learnt the specific characteristics of different materials, or did it simply become more resilient to noise by being exposed to noisy data? If the latter is true, we should be able to observe similar improvements using synthetic data augmentation. The results shown in Figure 5 indicate that this approach falls short, as none of the data augmentation techniques tested (i.e., *Gaussian-noise*, *RandAugment*, *RandAugment with Gaussian-noise*, or *CCycleGAN*), could match the performance of the *All-materials* classifier, thus answering our RQ2. However, training with materials is resource-intensive, as it requires collecting

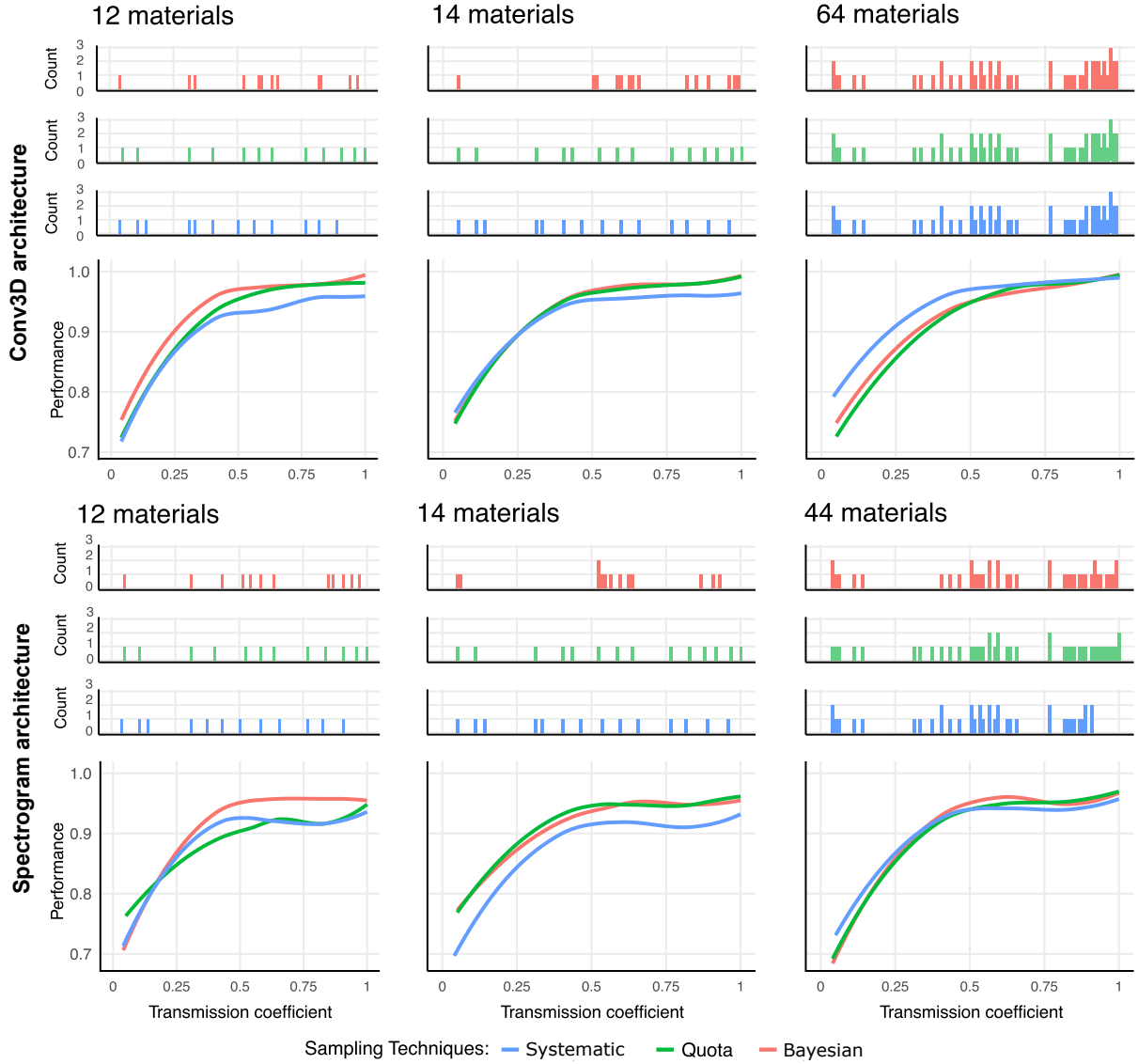


Figure 8. Comparison of optimal classifiers of sampling techniques for Conv3D and Spectrogram architectures. Columns show the sampling method whose optimal classifier is being highlighted (*Bayesian sampling*: left, *Quota sampling*: middle, *Systematic sampling*: right). For each column, we compare that method’s optimal classifier against classifiers trained with the same number of materials but selected using the other sampling methods (e.g., the *Bayesian 12-materials* model is compared to the *Quota 12-materials* and *Systematic 12-materials* models). The histograms indicate which materials were selected, positioned by their transmission coefficient T_c . The middle plots show how classification performance varies as a function of T_c . Statistical comparisons (Wilcoxon signed-rank tests) and full performance distributions (box/violin plots) are provided in the supplementary material.

data through each material. This highlights the need to optimise the training process, particularly in determining which materials to select and how many, as discussed next.

4.2 Which materials to select (RQ3)?

The results presented in Figure 6 highlight that different sampling techniques achieve different results; however, we were unable to identify a specific influential set of materials that would improve the classifier performance (Figure 8). The performance gain observed for *Bayesian optimisation sampling*, which, despite being designed to iteratively search for an optimal set of materials, only slightly outperformed *Quota sampling*. Our analysis also reveals that *Bayesian optimisation sampling* tends to select fewer materials with low T_c values, particularly when the number of training materials is limited. However, the underlying reasons for the exclusion of low T_c materials remain unclear and are beyond the scope of this study. To support future research, we have released the complete list of selected materials for all experimental configurations in Section 5 of the Supplementary Materials.

The key takeaway is that the selection of training materials should ensure comprehensive coverage of the T_c range in which the system is expected to operate. We recommend *Quota sampling* as a practical strategy because it provides broad coverage of the T_c range, and, unlike *Bayesian optimisation sampling*, does not require any material-specific training data during the sampling process while it can be effectively employed as long as the T_c of the encountered materials are known.

4.3 How many materials to select (RQ4)?

The findings in Figure 7 demonstrate a strong alignment between the *Significance-based* and *Multiple-criteria-based* objective functions. The results also confirmed our hypothesis that increasing the number of training materials enhances classification performance, but this improvement diminishes at the saturation point. Different sampling techniques reached this saturation point at different stages. *Bayesian optimisation sampling* reached saturation with 12 materials, followed by *Quota sampling* with 14 materials, and *Systematic sampling* with 44 and 64 materials for the Spectrogram and Conv3D architectures, respectively. Considering that higher-quality sampling methods will lead to faster saturation, *Bayesian optimisation* reached saturation with the fewest materials, but its advantage over *Quota sampling* was marginal. Because *Bayesian optimisation* requires repeated model training and evaluation, while *Quota sampling* provides a more practical trade-off between performance, simplicity, and deployment cost.

4.4 Limitations and future work

Using a single scalar T_c to approximate recognition degradation works best when the material mainly introduces uniform attenuation. However, it can break down for thicker barriers (20–50 mm) with a large dielectric mismatch to air (e.g., acrylic or drywall materials), because the mismatch produces multiple internal reflections in a form of internal scattering and multipath effects. These effects can raise the effective clutter level and introduce phase-dependent distortions that degrade gesture features, while the average transmission (and thus T_c) may remain relatively high. As a result, a T_c -based model can become overly optimistic for such materials. In contrast, when the dielectric mismatch is small (e.g., Styrofoam / Styrodur), interface reflections are weak and the barrier behaves closer to air, so signal structure is preserved even in thicker materials thus recognition performance is largely maintained even at larger thicknesses. Finally, for thin sheets that appear “highly attenuating” the T_c -based prediction can be too pessimistic: although some energy is lost, the barrier may introduce little reflection-driven clutter or waveform distortion, so the classifier can still reliably extract discriminative gesture signatures.

In the above scenarios T_c is likely not the best descriptor of material properties. However, for material selection to be practical, we require a systematic descriptor that can be measured reliably and inexpensively in real-world settings. We therefore use the transmission coefficient T_c as a proxy for overall material-induced sensing degradation: it captures the dominant through-material attenuation and, in practice, correlates strongly with recognition degradation across materials, as was shown by Čopič Pucihar et al [7].

We do not include difference-map visualizations (i.e. subtracting the no-material range–Doppler response from the material response) in this work. While such maps can help localize attenuation or ghosting in the range and velocity dimensions, producing physically meaningful and comparable difference maps across materials is non-trivial. Simple subtraction assumes that the no-material and material measurements are perfectly aligned in time, range, and phase, and that the barrier is the only source of change. But in practice, even small registration differences and barrier-induced shifts can cause subtraction to mix true attenuation with misalignment artifacts and may obscure genuine ghosting. A robust difference-map analysis would therefore require additional calibration and alignment steps and a dedicated processing pipeline, which we leave as an important direction for future work.

The main trends were observed for both deep learning architectures (i.e., Conv3D and Spectrogram), suggesting that the findings are likely generalisable to other CNN network architectures. However, in our study, we only used Google’s Soli sensor, an FMCW radar operating in the 57–64 GHz range, limiting the generalisation of results to other FMCW sensors with a similar frequency bands and front-end configuration. However, a recent literature review of 50 radar-based HCI studies found that 80% employed FMCW radars, and 63% operated within the 57–81 GHz range, suggesting that our setup is representative of the

majority of current research [97]. Nonetheless, a comparative evaluation across different radar types (e.g., FMCW, CW, UWB, pulsed), frequency bands (e.g., beyond the mmWave range), and antenna configurations (e.g., varying Tx/Rx arrangements) would provide additional insights and remains an important direction for future work.

One limitation is the use of Range Doppler images as the primary signal representation method. While this method is quite common [7, 32, 33], it may not be the most optimal for sensing gestures through materials. Future research should explore alternative signal representation methods to determine which method offers the best performance in this context.

Another limitation is the use of a robotic system for gesture performance instead of a human hand. While this approach provides a stable dataset, it does not reflect the inherent variability of human gestures (e.g., gesture speed, articulation, hand shape, and trajectory). Consequently, our conclusions are primarily derived from a low-variance gesture execution scenario, and greater execution variability may affect the absolute performance levels. That said, prior work by Čopić Pucihar et al. [7] evaluated T_c -based recognition degradation on both robot-executed and human datasets and reported consistent trends between T_c and performance degradation, suggesting that our main observations are likely to generalize. Nevertheless, human variability could still shift the absolute performance levels reported here, so future work should validate these findings with human-subject studies.

Gaussian noise and RandAugment are widely used data augmentation techniques that can improve classifier performance even when applied to existing datasets [90, 98–100]. However, more generalisable but computationally expensive deep learning approaches exist, such as Generative Adversarial Networks (GANs) for image reconstruction and cross-domain image mapping. We experimented with CCycleGAN and observed only marginal improvements over Gaussian noise and RandAugment. In contrast, Tran et al. [101] trained a GAN on over one million real images to model the noise distribution, achieving state-of-the-art performance on real-world noisy images. The limited gains in our case likely stem from training CCycleGAN on a relatively small and less diverse radar dataset, which lacked the volume needed to model complex noise patterns effectively. This comparison underscores the importance of large-scale data and domain-specific architectures when applying generative models for realistic data augmentation.

Other advanced, computationally intensive deep learning methods for synthetic data generation include semi-supervised Variational Autoencoders (VAEs), which have been used to generate noisy data for training image restoration networks [102], and Autoencoders (AEs) for denoising image sequences rendered via Monte Carlo methods [103]. Transformer architectures have also gained prominence for their effectiveness in sequence modelling tasks [104], and were leveraged to design a multi-person 3D pose estimation model capable of sensing through walls by mapping camera inputs to radar data [105]. These sophisticated techniques for synthetic data generation and noise reduction present promising opportunities for advancing radar-based gesture recognition through materials.

5 Conclusion

When radar sensors are embedded in everyday objects for interaction purposes, the radar signal needs to penetrate the occluding material twice, when emitting the signal and when receiving the signal bounced from the gesture performed. Due to a large variety of materials, training a general gesture classifier is costly. In this work we investigated different approaches, including synthetically generated noise and sampling techniques, for optimising selection of training materials while maintaining high gesture classification accuracy. Under our experimental conditions, using one FMCW mmWave radar sensor, a fixed gesture set executed by a robotic system, and 73 everyday materials, models trained on 14 materials selected by quota sampling achieved accuracies of 95.8% and 91.2% for Conv3D and Spectrogram models, respectively. These results were comparable to those obtained when training on all 73 materials, which yielded accuracies of 96.8% and 91.6%. In contrast, models trained without material data performed significantly worse, with accuracies of only 66.8% and 65.8%. Data augmentation alone did not close this gap, yielding performance comparable to that of models trained without material examples.

Indeed, we have demonstrated that training with only 14 strategically selected materials using quota sampling achieves comparable performance to using the full dataset of 73 materials. Among the evaluated selection approaches (i.e. Systematic, Quota, and Bayesian optimization sampling), Quota sampling provided the best overall trade-off between performance and practicality. Our work contributes to the field of radar sensing through materials by offering a scalable method for building efficient and cost-effective gesture recognition systems with minimal data overhead. Our code, models, and data are publicly available at <https://gitlab.com/hicuplab/seeing-through>.

References

1. Norman, D. A. Natural user interfaces are not natural. *Interactions* **17**, DOI: [10.1145/1744161.1744163](https://doi.org/10.1145/1744161.1744163) (2010).
2. Čopić Pucihar, K., Vatavu, R.-D., Salami, D. & Attygalle, N. T. (eds.) *Radar-Based Human-Computer Interaction*. Human-Computer Interaction Series (Springer Nature Switzerland, Cham, 2026), 1 edn.

3. Attygalle, N. T., Vanderdonckt, J. & Parthiban, V. Turning gesture sequences into gesture composition by macro-instructions: Definition, framework, and application in smart home. *Proc. ACM Hum.-Comput. Interact.* **10**, DOI: [10.1145/3816768](https://doi.org/10.1145/3816768) (2026).
4. Attygalle, N. T. *et al.* No interface, no problem: gesture recognition on physical objects using radar sensing. *Sensors* **21**, DOI: [10.3390/s21175771](https://doi.org/10.3390/s21175771) (2021).
5. Attygalle, N. T., Kljun, M. & Čopič Pucihar, K. Gesture recognition on deformable objects using millimeter-wave radar. In *Companion Proceedings of the 17th ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, EICS '25 Companion, 50–58, DOI: [10.1145/3731406.3734979](https://doi.org/10.1145/3731406.3734979) (Association for Computing Machinery, New York, NY, USA, 2025).
6. Čopič Pucihar, K., Kljun, M. & Attygalle, N. T. *Radar-Based Gesture Recognition on Deformable Objects*, 107–122 (Springer Nature Switzerland, Cham, 2026).
7. Čopič Pucihar, K., Attygalle, N. T., Kljun, M., Sandor, C. & Leiva, L. A. Solids on soli: Millimetre-wave radar sensing through materials. *Proc. ACM Hum. Comput. Interact.* **6**, 1–19, DOI: [10.1145/3532212](https://doi.org/10.1145/3532212) (2022).
8. Leiva, L. A., Kljun, M., Sandor, C. & Čopič Pucihar, K. The wearable radar: Sensing gestures through fabrics. In *Proc. MobileHCI*, DOI: [10.1145/3406324.3410720](https://doi.org/10.1145/3406324.3410720) (2021).
9. Attygalle, N. T., Kljun, M., Sluÿters, A. & Čopič Pucihar, K. *Radar-Based Human-Computer Interaction When Sensing Through Materials*, 65–84 (Springer Nature Switzerland, Cham, 2026).
10. Cho, J., Lee, K., Shin, E., Choy, G. & Do, S. How much data is needed to train a medical image deep learning system to achieve necessary high accuracy?, DOI: [10.48550/arXiv.1511.06348](https://doi.org/10.48550/arXiv.1511.06348) (2016). [1511.06348](https://arxiv.org/abs/1511.06348).
11. Couch, J., Arnaout, R. & Arnaout, R. Beyond size and class balance: Alpha as a new dataset quality metric for deep learning, DOI: [10.48550/arXiv.2407.15724](https://doi.org/10.48550/arXiv.2407.15724) (2024). [2407.15724](https://arxiv.org/abs/2407.15724).
12. Liu, L. *et al.* Deep Neural Network Ensembles Against Deception: Ensemble Diversity, Accuracy and Robustness . In *Proc. MASS*, DOI: [10.1109/MASS.2019.00040](https://doi.org/10.1109/MASS.2019.00040) (2019).
13. Masisi, L., Nelwamondo, F. V. & Marwala, T. The effect of structural diversity of an ensemble of classifiers on classification accuracy, DOI: [10.48550/arXiv.0804.4741](https://doi.org/10.48550/arXiv.0804.4741) (2008). [0804.4741](https://arxiv.org/abs/0804.4741).
14. Skolnik, M. I. *et al.* *Introduction to Radar Systems* (McGraw-Hill Education, New York, NY, USA, 2002), 3rd edn.
15. Čopič Pucihar, K., Salami, D. & Attygalle, N. T. *Current State and Future Research Directions of Radar-Based Human-Computer Interaction*, 235–267 (Springer Nature Switzerland, Cham, 2026).
16. Otero, M. Application of a continuous wave radar for human gait recognition. In *Proc. SPIE*, DOI: [10.1117/12.607176](https://doi.org/10.1117/12.607176) (2005).
17. Chen, V. C., Li, F., Ho, S.-S. & Wechsler, H. Micro-doppler effect in radar: phenomenon, model, and simulation study. *IEEE Trans. Aerosp. Electron. Syst.* **42**, DOI: [10.1109/TAES.2006.1603402](https://doi.org/10.1109/TAES.2006.1603402) (2006).
18. Zhao, Z. *et al.* Object state optimization algorithm based on Bayesian random sampling for visual object tracking. *Sci. Reports* **15**, 37237, DOI: [10.1038/s41598-025-21033-2](https://doi.org/10.1038/s41598-025-21033-2) (2025).
19. Wang, Y. & Fathy, A. E. Micro-doppler signatures for intelligent human gait recognition using a uwb impulse radar. In *Proc. APSURSI*, DOI: [10.1109/APS.2011.5996925](https://doi.org/10.1109/APS.2011.5996925) (2011).
20. Chen, H. *et al.* RaGeoSense for smart home gesture recognition using sparse millimeter wave radar point clouds. *Sci. Reports* **15**, 15267, DOI: [10.1038/s41598-025-00065-8](https://doi.org/10.1038/s41598-025-00065-8) (2025).
21. Rahman, T. *et al.* Dopplesleep: a contactless unobtrusive sleep sensing system using short-range doppler radar. In *Proc. UbiComp*, DOI: [10.1145/2750858.2804280](https://doi.org/10.1145/2750858.2804280) (2015).
22. Zhuang, Y. *et al.* Sleepsense: Non-invasive sleep event recognition using an electromagnetic probe. In *Proc. BSN*, DOI: [10.1109/BSN.2015.7299364](https://doi.org/10.1109/BSN.2015.7299364) (2015).

23. Kellogg, B., Talla, V. & Gollakota, S. Bringing gesture recognition to all devices. In *Proc. NSDI*, DOI: [10.5555/2616448.2616477](https://doi.org/10.5555/2616448.2616477) (2014).
24. McIntosh, J., Fraser, M., Worgan, P. & Marzo, A. Deskwave: Desktop interactions using low-cost microwave doppler arrays. In *Proc. CHI*, DOI: [10.1145/3027063.3053152](https://doi.org/10.1145/3027063.3053152) (2017).
25. Molchanov, P., Gupta, S., Kim, K. & Pulli, K. Short-range fmcw monopulse radar for hand-gesture sensing. In *Proc. RadarCon*, DOI: [10.1109/RADAR.2015.7131232](https://doi.org/10.1109/RADAR.2015.7131232) (2015).
26. Paradiso, J., Abler, C., Hsiao, K.-y. & Reynolds, M. The magic carpet: physical sensing for immersive environments. In *Proc. CHI EA*, DOI: [10.1145/1120212.1120391](https://doi.org/10.1145/1120212.1120391) (1997).
27. Paradiso, J. A. The brain opera technology: New instruments and gestural sensors for musical interaction and performance. *J. New Music. Res.* **28**, DOI: [10.1076/jnmr.28.2.130.3119](https://doi.org/10.1076/jnmr.28.2.130.3119) (1999).
28. Pu, Q., Gupta, S., Gollakota, S. & Patel, S. Whole-home gesture recognition using wireless signals. In *Proc. MobiCom*, DOI: [10.1145/2500423.2500436](https://doi.org/10.1145/2500423.2500436) (2013).
29. Wan, Q., Li, Y., Li, C. & Pal, R. Gesture recognition for smart home applications using portable radar sensors. In *Proc. EMBC*, DOI: [10.1109/EMBC.2014.6945096](https://doi.org/10.1109/EMBC.2014.6945096) (2014).
30. Nasr, I., Karagozler, E., Poupyrev, I. & Trotta, S. A highly integrated 60-ghz 6-channel transceiver chip in 0.35 μm sige technology for smart sensing and short-range communications. In *Proc. CSICS*, DOI: [10.1109/CSICS.2015.7314475](https://doi.org/10.1109/CSICS.2015.7314475) (2015).
31. Wei, T. & Zhang, X. mtrack: High-precision passive tracking using millimeter wave radios. In *Proc. MobiCom*, DOI: [10.1145/2789168.2790113](https://doi.org/10.1145/2789168.2790113) (2015).
32. Lien, J. *et al.* Soli: Ubiquitous gesture sensing with millimeter wave radar. *ACM Trans. Graph.* **35**, DOI: [10.1145/2897824.2925953](https://doi.org/10.1145/2897824.2925953) (2016).
33. Wang, S., Song, J., Lien, J., Poupyrev, I. & Hilliges, O. Interacting with soli: Exploring fine-grained dynamic gesture recognition in the radio-frequency spectrum. In *Proc. UIST*, DOI: [10.1145/2984511.2984565](https://doi.org/10.1145/2984511.2984565) (2016).
34. Ens, B. *et al.* Exploring mixed-scale gesture interaction. In *Proc. SIGGRAPH Asia Posters*, DOI: [10.1145/3145690.3145740](https://doi.org/10.1145/3145690.3145740) (2017).
35. Čopič Pucihar, K. *et al.* The missing interface: Micro-gestures on augmented objects. In *Proc. CHI EA*, DOI: [10.1145/3290607.3312986](https://doi.org/10.1145/3290607.3312986) (2019).
36. Bernardo, F., Arner, N. & Batchelor, P. O soli mio: Exploring millimeter wave radar for musical interaction. In *Proc. NIME*, DOI: [10.5281/zenodo.1176252](https://doi.org/10.5281/zenodo.1176252) (2017).
37. Sandor, C. & Nakamura, H. Soliscratch: A radar interface for scratch djs. In *Proc. ISMAR Adjunct*, DOI: [10.1109/ISMAR-Adjunct.2018.00129](https://doi.org/10.1109/ISMAR-Adjunct.2018.00129) (2018).
38. McIntosh, J., Marzo, A., Fraser, M. & Phillips, C. Echoflex: Hand gesture recognition using ultrasound imaging. In *Proc. CHI*, DOI: [10.1145/3025453.3025807](https://doi.org/10.1145/3025453.3025807) (2017).
39. Yeo, H.-S., Flamich, G., Schrempf, P., Harris-Birtill, D. & Quigley, A. Radarcats: Radar categorization for input & interaction. In *Proc. UIST*, DOI: [10.1145/2984511.2984515](https://doi.org/10.1145/2984511.2984515) (2016).
40. Xu, W. *et al.* Acoustic sensing mechanisms, technologies, and applications: A survey. *CCF Trans. Pervasive Comput. Interact.* DOI: [10.1007/s42486-025-00193-0](https://doi.org/10.1007/s42486-025-00193-0) (2025).
41. Piyathilaka, L., Sooriyaarachchi, B., Kodagoda, S. & Thiyagarajan, K. Capacitive sensor based 2d subsurface imaging technology for non-destructive evaluation of building surfaces. In *Proc. CIS/RAM*, DOI: [10.1109/CIS-RAM47153.2019.9095810](https://doi.org/10.1109/CIS-RAM47153.2019.9095810) (2019).
42. Attygalle, N. T. ehenna: A topology-centered taxonomy of geometry-defined finger input on skin and glove conductive interfaces. *Proc. ACM Hum.-Comput. Interact.* **10**, DOI: [10.1145/3807968.3817087](https://doi.org/10.1145/3807968.3817087) (2026).

43. Hao, Z., Sun, Z., Li, F., Wang, R. & Peng, J. Millimeter wave gesture recognition using multi-feature fusion models in complex scenes. *Sci. Reports* **14**, 13758, DOI: [10.1038/s41598-024-64576-6](https://doi.org/10.1038/s41598-024-64576-6) (2024).
44. Palipana, S., Salami, D., Leiva, L. A. & Sigg, S. Pantomime: Mid-air gesture recognition with sparse millimeter-wave radar point clouds. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **5**, DOI: [10.1145/3448110](https://doi.org/10.1145/3448110) (2021).
45. Zhang, C. *et al.* Fingerping: Recognizing fine-grained hand poses using active acoustic on-body sensing. In *Proc. CHI*, DOI: [10.1145/3173574.3174011](https://doi.org/10.1145/3173574.3174011) (2018).
46. Harrison, C., Tan, D. & Morris, D. Skinput: appropriating the body as an input surface. In *Proc. CHI*, DOI: [10.1145/1753326.1753394](https://doi.org/10.1145/1753326.1753394) (2010).
47. Iravantchi, Y., Goel, M. & Harrison, C. Beamband: Hand gesture sensing with ultrasonic beamforming. In *Proc. CHI*, DOI: [10.1145/3290605.3300245](https://doi.org/10.1145/3290605.3300245) (2019).
48. Iravantchi, Y., Zhang, Y., Bernitsas, E., Goel, M. & Harrison, C. Interferi: Gesture sensing using on-body acoustic interferometry. In *Proc. CHI*, DOI: [10.1145/3290605.3300506](https://doi.org/10.1145/3290605.3300506) (2019).
49. Zhang, Z., Hao, Z. & Li, M. ULTRAWX: A ubiquitous realtime acoustic gesture information interaction system based on Tiou DODA. *Sci. Reports* **15**, 9654, DOI: [10.1038/s41598-025-93837-1](https://doi.org/10.1038/s41598-025-93837-1) (2025).
50. Tariq, O. B., Lazarescu, M. T., Iqbal, J. & Lavagno, L. Performance of machine learning classifiers for indoor person localization with capacitive sensors. *IEEE Access* **5**, DOI: [10.1109/ACCESS.2017.2721538](https://doi.org/10.1109/ACCESS.2017.2721538) (2017).
51. Grosse-Puppendahl, T., Berghoefer, Y., Braun, A., Wimmer, R. & Kuijper, A. Opencapsense: A rapid prototyping toolkit for pervasive interaction using capacitive sensing. In *Proc. PerCom*, DOI: [10.1109/PerCom.2013.6526726](https://doi.org/10.1109/PerCom.2013.6526726) (2013).
52. Wong, W., Juwono, F. H. & Khoo, B. T. T. Multi-features capacitive hand gesture recognition sensor: A machine learning approach. *IEEE Sensors* **21**, DOI: [10.1109/JSEN.2021.3049273](https://doi.org/10.1109/JSEN.2021.3049273) (2021).
53. Zheng, T. *et al.* SiWa: See into Walls via Deep UWB Radar. In *Proc. MobiCom*, DOI: [10.1145/3447993.3483258](https://doi.org/10.1145/3447993.3483258) (2021).
54. Wang, R. *et al.* Through-Wall Object Recognition and Pose Estimation. In *Proc. ISARC*, DOI: [10.22260/ISARC2019/0157](https://doi.org/10.22260/ISARC2019/0157) (2019).
55. Memon, Q., Wodajo, B., Tekleab, S. & Alshehi, E. Detection of Static and Moving Objects behind Walls and Surfaces – An Experimental Investigation. In *Proc. ICCAI*, DOI: [10.1145/3594315.3594317](https://doi.org/10.1145/3594315.3594317) (2023).
56. Adib, F., Hsu, C.-Y., Mao, H., Katabi, D. & Durand, F. Capturing the human figure through a wall. *ACM Trans. Graph.* **34**, DOI: [10.1145/2816795.2818072](https://doi.org/10.1145/2816795.2818072) (2015).
57. Chetty, K., Smith, G. E. & Woodbridge, K. Through-the-wall sensing of personnel using passive bistatic wifi radar at standoff distances. *IEEE Trans. Geosci. Remote. Sens.* **50**, DOI: [10.1109/TGRS.2011.2164411](https://doi.org/10.1109/TGRS.2011.2164411) (2011).
58. Sluÿters, A., Lambot, S., Vanderdonckt, J. & Villarreal-Narvaez, S. Analysis of User-Defined Radar-Based Hand Gestures Sensed Through Multiple Materials. *IEEE Access* **12**, DOI: [10.1109/ACCESS.2024.3366667](https://doi.org/10.1109/ACCESS.2024.3366667) (2024).
59. Attygalle, N. T., Vuletic, U., Kljun, M. & Pucihar, K. Č. Towards hand gesture recognition prototype using the iwr6843isk radar sensor and leap motion. In *Proceedings of the 8th Human-Computer Interaction Slovenia (HCI SI) Conference 2023*, 78–88 (2024).
60. Avrahami, D., Patel, M., Yamaura, Y. & Kratz, S. Below the surface: Unobtrusive activity recognition for work surfaces using rf-radar sensing. In *Proc. IUI*, DOI: [10.1145/3172944.3172962](https://doi.org/10.1145/3172944.3172962) (2018).
61. Avrahami, D., Patel, M., Yamaura, Y., Kratz, S. & Cooper, M. Unobtrusive Activity Recognition and Position Estimation for Work Surfaces Using RF-Radar Sensing. *ACM Trans. Interact. Intell. Syst.* **10**, DOI: [10.1145/3241383](https://doi.org/10.1145/3241383) (2020).
62. Vishwakarma, S., Ummalaneni, V., Iqbal, M. S., Majumdar, A. & Ram, S. S. Mitigation of through-wall interference in radar images using denoising autoencoders. In *Proc. RadarConf*, DOI: [10.1109/RADAR.2018.8378796](https://doi.org/10.1109/RADAR.2018.8378796) (2018).
63. Vishwakarma, S. & Ram, S. S. Mitigation of Through-Wall Distortions of Frontal Radar Images Using Denoising Autoencoders. *IEEE Trans. Geosci. Remote. Sens.* **58**, DOI: [10.1109/TGRS.2020.2978440](https://doi.org/10.1109/TGRS.2020.2978440) (2020).

64. Ram, S. S., Vishwakarma, S., Sneh, A. & Yasmeen, K. Sparsity-based autoencoders for denoising cluttered radar signatures. *IET Radar Sonar Navig.* **15**, DOI: [10.1049/rsn2.12065](https://doi.org/10.1049/rsn2.12065) (2021).
65. Rock, J., Toth, M., Meissner, P. & Pernkopf, F. Cnns for interference mitigation and denoising in automotive radar using real-world data. In *Proc. NeurIPS* (2019).
66. Rock, J., Toth, M., Meissner, P. & Pernkopf, F. Deep interference mitigation and denoising of real-world fmcw radar signals. In *Proc. RADAR*, DOI: [10.1109/radar42522.2020.9114627](https://doi.org/10.1109/radar42522.2020.9114627) (2020).
67. Ristea, N.-C., Anghel, A. & Ionescu, R. T. Fully convolutional neural networks for automotive radar interference mitigation. In *Proc. VTC*, DOI: [10.1109/VTC2020-Fall49728.2020.9348690](https://doi.org/10.1109/VTC2020-Fall49728.2020.9348690) (2020).
68. Muqaibel, A. & Safaai-Jazi, A. Characterization of wall dispersive and attenuative effects on uwb radar signals. *J. Frankl. Inst.* **345**, 640–658, DOI: <https://doi.org/10.1016/j.jfranklin.2008.01.002> (2008). Advances in Indoor Radar Imaging.
69. Vishwakarma, S. & Ram, S. S. Mitigation of through-wall distortions of frontal radar images using denoising autoencoders. *IEEE Transactions on Geosci. Remote. Sens.* **58**, 6650–6663, DOI: [10.1109/TGRS.2020.2978440](https://doi.org/10.1109/TGRS.2020.2978440) (2020).
70. Mahfouz, M., Fathy, A., Yang, Y., Ali, E. & Badawi, A. See-through-wall imaging using ultra wideband pulse systems. In *34th Applied Imagery and Pattern Recognition Workshop (AIPR'05)*, 6 pp.–53, DOI: [10.1109/AIPR.2005.40](https://doi.org/10.1109/AIPR.2005.40) (2005).
71. Valentine, H. T., Affleck, D. L. & Gregoire, T. G. Systematic sampling of discrete and continuous populations: sample selection and the choice of estimator. *Can. J. For. Res.* **39**, DOI: [10.1139/X09-019](https://doi.org/10.1139/X09-019) (2009).
72. Yates, F. & Thornton, H. G. Systematic sampling. *Philos. Trans. R. Soc. A* **241**, DOI: [10.1098/rsta.1948.0023](https://doi.org/10.1098/rsta.1948.0023) (1948).
73. Yang, K. & Banamah, A. Quota sampling as an alternative to probability sampling? an experimental study. *Sociol. Res. Online* **19**, DOI: [10.5153/sro.3199](https://doi.org/10.5153/sro.3199) (2014).
74. Hsu, C.-W., Chang, C.-C. & Lin, C.-J. A practical guide to support vector classification (2003).
75. Bergstra, J. & Bengio, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **13** (2012).
76. Bergstra, J., Bardenet, R., Bengio, Y. & Kégl, B. Algorithms for hyper-parameter optimization. In *Proceedings of the 25th International Conference on Neural Information Processing Systems, NIPS'11*, 2546–2554 (Curran Associates Inc., Red Hook, NY, USA, 2011).
77. Bergstra, J., Yamins, D. & Cox, D. Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In *Proc. ICML*, vol. 28 (2013).
78. Hutter, F., Hoos, H. H. & Leyton-Brown, K. Sequential model-based optimization for general algorithm configuration. In *Proc. LION*, DOI: [10.1007/978-3-642-25566-3_40](https://doi.org/10.1007/978-3-642-25566-3_40) (2011).
79. Snoek, J., Larochelle, H. & Adams, R. P. Practical bayesian optimization of machine learning algorithms. In *Proc. NeurIPS* (2012).
80. Liashchynskyi, P. & Liashchynskyi, P. Grid search, random search, genetic algorithm: A big comparison for nas (2019). [1912.06059](https://doi.org/10.1912.06059).
81. Pelikan, M., Goldberg, D. E. & Cantú-Paz, E. Boa: the bayesian optimization algorithm. In *GECCO'99: Proceedings of the 1st Annual Conference on Genetic and Evolutionary Computation - Volume 1*, GECCO'99, 525–532 (Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1999).
82. Turner, R. *et al.* Bayesian optimization is superior to random search for machine learning hyperparameter tuning: Analysis of the black-box optimization challenge 2020 (2021). [2104.10201](https://doi.org/10.2104.10201).
83. Kusne, A. G. *et al.* On-the-fly closed-loop materials discovery via bayesian active learning. *Nat. Commun.* **11**, DOI: [10.1038/s41467-020-19597-w](https://doi.org/10.1038/s41467-020-19597-w) (2020).
84. Kusne, A. G. *et al.* On-the-fly machine-learning for high-throughput experiments: search for rare-earth-free permanent magnets. *Sci. Rep.* **4**, DOI: [10.1038/srep06367](https://doi.org/10.1038/srep06367) (2014).

85. Attygalle, N. T., Kljun, M., Vuletić, U. & Čopič Pucihar, K. *Comparative Testing of Radar Signal Representations When Sensing Through Materials*, 47–61 (Springer Nature Switzerland, Cham, 2026).
86. Attygalle, N. T., Kljun, M. & Čopič Pucihar, K. *Digital Signal Processing Tools for Radar-Based Human-Computer Interaction*, 31–46 (Springer Nature Switzerland, Cham, 2026).
87. Power, A., Burda, Y., Edwards, H., Babuschkin, I. & Misra, V. Grokking: Generalization beyond overfitting on small algorithmic datasets, DOI: [10.48550/arXiv.2201.02177](https://doi.org/10.48550/arXiv.2201.02177) (2022). [2201.02177](https://doi.org/10.48550/arXiv.2201.02177).
88. Goswami, P., Rao, S., Bharadwaj, S. & Nguyen, A. Real-time multi-gesture recognition using 77 ghz fmcw mimo single chip radar. In *Proc. ICCE*, DOI: [10.1109/ICCE.2019.8662006](https://doi.org/10.1109/ICCE.2019.8662006) (2019).
89. Kannan, M. *et al.* An enhancement of machine learning model performance in disease prediction with synthetic data generation. *Sci. Reports* **15**, 33482, DOI: [10.1038/s41598-025-15019-3](https://doi.org/10.1038/s41598-025-15019-3) (2025).
90. Cubuk, E. D., Zoph, B., Shlens, J. & Le, Q. V. Randaugment: Practical automated data augmentation with a reduced search space. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 3008–3017, DOI: [10.1109/CVPRW50498.2020.00359](https://doi.org/10.1109/CVPRW50498.2020.00359) (2020).
91. Tesei, G. Generating realistic facial expressions through conditional cycle-consistent generative adversarial networks (ccyclegan). In *Proc. RIIAA* (2019).
92. Zhu, J.-Y., Park, T., Isola, P. & Efros, A. A. Unpaired image-to-image translation using cycle-consistent adversarial networks, DOI: [10.48550/arXiv.1703.10593](https://doi.org/10.48550/arXiv.1703.10593) (2020). [1703.10593](https://doi.org/10.48550/arXiv.1703.10593).
93. authors, T. G. GPyOpt: A bayesian optimization framework in python. <http://github.com/SheffieldML/GPyOpt> (2016).
94. Wang, Z., Hutter, F., Zoghi, M., Matheson, D. & De Freitas, N. Bayesian optimization in a billion dimensions via random embeddings. *J. Artif. Int. Res.* **55** (2016).
95. Basílio, M. P., Pereira, V., Costa, H. G., Santos, M. & Ghosh, A. A systematic review of the applications of multi-criteria decision aid methods (1977–2022). *Electronics* **11**, DOI: [10.3390/electronics11111720](https://doi.org/10.3390/electronics11111720) (2022).
96. Behzadian, M., Khanmohammadi Otaghsara, S., Yazdani, M. & Ignatius, J. A state-of-the-art survey of topsiis applications. *Expert. Syst. Appl.* **39**, DOI: [10.1016/j.eswa.2012.05.056](https://doi.org/10.1016/j.eswa.2012.05.056) (2012).
97. Pucihar, K. Č., Salami, D. & Attygalle, N. T. *Radar-Based Human-Computer Interaction*, chap. Current State and Future Research Directions of Radar-based Human-Computer Interaction, 257–288 (Springer International Publishing, 2025).
98. Silva, I. B. V. d. & Adeodato, P. J. L. Pca and gaussian noise in MLP neural network training improve generalization in problems with small and unbalanced data sets. In *Proc. IJCNN*, DOI: [10.1109/IJCNN.2011.6033567](https://doi.org/10.1109/IJCNN.2011.6033567) (2011).
99. Liu, L. *et al.* Boost ai power: Data augmentation strategies with unlabeled data and conformal prediction, a case in alternative herbal medicine discrimination with electronic nose. *IEEE Sensors* **21**, DOI: [10.1109/JSEN.2021.3102488](https://doi.org/10.1109/JSEN.2021.3102488) (2021).
100. Ottoni, A. L. C., de Amorim, R. M., Novo, M. S. & Costa, D. B. Tuning of data augmentation hyperparameters in deep learning to building construction image classification with small datasets. *Int. J. Mach. Learn. Cybern.* **14**, DOI: [10.1007/s13042-022-01555-1](https://doi.org/10.1007/s13042-022-01555-1) (2023).
101. Tran, L. D., Nguyen, S. M. & Arai, M. Gan-based noise model for denoising real images. In *Proc. ACCV*, DOI: [10.1007/978-3-030-69538-5_34](https://doi.org/10.1007/978-3-030-69538-5_34) (2021).
102. Zheng, D., Zou, Y., Zhang, X. & Bao, C. Senm-vae: Semi-supervised noise modeling with hierarchical variational autoencoder. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 25889–25899, DOI: [10.1109/CVPR52733.2024.02446](https://doi.org/10.1109/CVPR52733.2024.02446) (2024).
103. Chaitanya, C. R. A. *et al.* Interactive reconstruction of monte carlo image sequences using a recurrent denoising autoencoder. *ACM Trans. Graph.* **36**, DOI: [10.1145/3072959.3073601](https://doi.org/10.1145/3072959.3073601) (2017).
104. Vaswani, A. *et al.* Attention is all you need. In *Proc. NeurIPS*, DOI: [10.48550/arXiv.1706.03762](https://doi.org/10.48550/arXiv.1706.03762) (2017).
105. Kim, G. W., Lee, S. W., Son, H. Y. & Choi, K. W. A study on 3d human pose estimation using through-wall ir-uwb radar and transformer. *IEEE Access* DOI: [10.1109/ACCESS.2023.3244017](https://doi.org/10.1109/ACCESS.2023.3244017) (2023).

Funding

Research supported by the European Innovation Council (grant 101071147) and the European Commission through the InnoRenew CoE project (grant 739574) under the Horizon2020 Widespread-Teaming program and the Republic of Slovenia (investment funding of the Republic of Slovenia and the European Union of the European Regional Development Fund). We also acknowledge support from the Slovenian research agency ARRS (program no. P1-0383, J1-9186, J1-1715, J5-1796, and J1-1692).

Author contributions

N. T. Attygalle: Conceptualization, Methodology, Software, Validation, Writing - Original Draft, Reviewing and Editing. **L. A. Leiva:** Investigation, Methodology, Writing - Reviewing and Editing. **M. Kljun:** Investigation, Resources, Writing - Original Draft, Reviewing and Editing. **K. Čopič Pucihar:** Conceptualization, Visualization, Resources, Writing - Original Draft, Reviewing and Editing.

Additional information

Competing interests: The authors declare no competing interests.

Data availability: The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.