


RESEARCH

Open Access



Comparison of synthetic dataset generation methods for medical intervention rooms using medical clothing detection as an example

Patrick Schülein^{1*} , Hannah Teufel⁴, Ronja Vorpahl¹, Indira Emter¹, Yannick Bukschat¹, Marcus Pfister³, Nils Rathmann², Steffen Diehl² and Marcus Vetter^{1*}

*Correspondence:
p.schuelein@hs-mannheim.de;
m.vetter@hs-mannheim.de

¹ ESM-Institute, University
of Applied Sciences Mannheim,
Mannheim, Germany

² Clinic for Radiology and Nuclear
Medicine, Medical Faculty
Mannheim, Heidelberg

University, Mannheim, Germany

³ Siemens Healthineers AG,
Advanced Therapies, Forchheim,
Germany

⁴ ESM-Institute, Mannheim,
Germany

Abstract

Purpose: The availability of real data from areas with high privacy requirements, such as the medical intervention space is low and the acquisition complex in terms of data protection. To enable research for assistance systems in the medical intervention room, new methods for data generation for these areas must be researched. Therefore, this work presents a way to create a synthetic dataset for the medical context, using medical clothing object detection as an example. The goal is to close the reality gap between the synthetic and real data.

Methods: Methods of 3D-scanned clothing and designed clothing are compared in a Domain-Randomization and Structured-Domain-Randomization scenario using two different rendering engines. Additionally, a Mixed-Reality dataset in front of a greenscreen and a target domain dataset were used while the latter is used to evaluate the different datasets. The experiments conducted are to show whether scanned clothing or designed clothing produce better results in Domain Randomization and Structured Domain Randomization. Likewise, a baseline will be generated using the mixed reality data. In a further experiment it is investigated whether the combination of real, synthetic and mixed reality image data improves the accuracy compared to real data only.

Results: Our experiments show, that Structured-Domain-Randomization of designed clothing together with Mixed-Reality data provide a baseline achieving 72.0% mAP on the test dataset of the clinical target domain. When additionally using 15% (99 images) of available target domain train data, the gap towards 100% (660 images) target domain train data could be nearly closed 80.05% mAP (81.95% mAP). Finally, we show that when additionally using 100% target domain train data the accuracy could be increased to 83.35% mAP.

Conclusion: In conclusion, it can be stated that the presented modeling of health professionals is a promising methodology to address the challenge of missing datasets from medical intervention rooms. We will further investigate it on various tasks, like assistance systems, in the medical domain.

Keywords: Camera-based AI-methods, Medical clothing detection, Domain Randomization, Structured Domain Randomization, Synthetic dataset, Mixed Reality, Deformable objects, 3D scanning, 3D modeling

1 Introduction

For computer vision challenges in the medical intervention space such as object detection, person detection or more sophisticated challenges such as activity detection, few datasets exist [1, 19]. These datasets focus on 2D and 3D detection of the human pose. Additionally, although multiple cameras are used in the datasets, only around 700 frames per camera are annotated. Furthermore, datasets besides the aforementioned are only institution-related and not publicly available possibly due to data protection regulations and ethics requirements [18, 28]. Likewise, the data sets may lack the necessary variance for transferability to other localities. Moreover, other use cases besides the detection of the human pose exist. These include AI-based sterility detection of health professionals, whether and where certain medical devices are located or action recognition of health professionals [14]. From the works using camera-based systems during real interventions, to the best of our knowledge, no datasets besides the mentioned are available for the public.

The successes of deep learning in recent years are among others, due to the availability of large datasets such as Imagenet [16] for Image Classification, or MS COCO [8] for Object Bounding Box Detection. In addition, the research of new methods or architectures like [4, 9, 21, 2] and the research of high-performance hardware for parallel computing are to be mentioned. The authors of [20, 7] analyze the hardware topic in depth. In this work however, special focus is put on the availability of datasets and methods for dataset generation in order to reduce the necessary amount of real data from the target domain.

Several works already deal with the generation of synthetic data and with the goal of reducing the reality gap between synthetic and real data. Among these are the work on Domain Randomization (DR) and Structured Domain Randomization (SDR) [24, 25, 12]. In addition to other work, it has already been shown that the use of synthetic image data can decrease the required amount of real data [25]. Likewise, synthetic data of persons exist [5]. However, the challenge lies in the specific characteristics of a medical intervention space. Health professionals wear special clothing with sometimes multiple layers, wear sterile gloves, masks and hairnets. The differences between conventional human data and data from the medical field are large. Nevertheless, DR techniques sound promising for use in research questions around medical interventions.

This work presents a comparison in terms of detection accuracy and generalizability of different methods for synthetic clothing generation using either 3D clothing scans (SCANS) or designed CAD clothing (CAD) with the Skinned Multi-Person Liner Model (SMPL) [10]. The comparison is performed using the example of medical clothing object detection. To generate synthetic training data, both methods (SCANS, CAD) are incorporated into a DR environment called NVIDIA Deep Learning Dataset Synthesizer (NDDS) [23] and an SDR environment implemented in Unity, based on [22]. Likewise, the aim of the presented methodology is to explore a pipeline for the generation of synthetic data for the medical field, so that further research questions from the intervention

space can be explored. In addition to synthetic data, real data of different persons are recorded in front of a green screen which aims to close the reality gap between synthetic data and the target domain. To evaluate the different approaches, a data set from the target domain of the hospital is recorded. All data are split into train/val/test sets while the clinical dataset serves as a test set for all methods.

2 Related work

With the rise of synthetic data generation methods, for example DR [24], it has already been shown that synthetic data can reduce the amount of real data required [25]. However, one focus of research is the reduction of the reality gap between the synthetic data and the target domain.

Here, the aforementioned DR has turned out to be one way to reduce the gap. One idea of DR is that if enough variance can be generated in the synthetic data, reality represents another variance of the target domain [24].

The work of Tobin et al. [24] and Tremblay et al. [25] showed, that an object detection network for robot grasping or car detection can be trained from synthetic images with random positioning, random lighting, random backgrounds, distractor objects and non-realistic textures alone. In addition, the work of Tremblay et al. showed, that the necessary amount of real target domain data can be reduced while maintaining adequate accuracy, when pretrained with DR-generated images.

Also the work of Borkman et al. [3] showed that when using Unity Perception for synthetic data generation, the amount of real-world data could be reduced to 10% when used together with the synthetic data, while achieving better Average Precision (AP) score as with all real-world data alone.

DR has already been successfully applied in various fields. In addition to the mentioned areas of car detection and robot grasping, the work of Sadeghi et al. [17] for flying a quadcopter through indoor environments, Zhang et al. [31] for a table-top object reaching task through clutter or James et al. [6] for grasping a cube and placing it in a basket can be named.

This leads us to believe that DR is a suitable approach for the medical intervention room domain, where no real data are largely available and access to that domain is widely restricted.

Ablation studies of [25] and [24] showed, that high-resolution textures and higher numbers of unique textures in the scene improve performance. Also, [31] come to the conclusion, after testing their hypothesis, that using complex textures yields better performance than using random colors.

In contrast to the DR approach is the photorealistic rendering of the scene and objects. A number of datasets have been created for this purpose in recent years. Here the works of [26, 5, 29] or [27] are to be mentioned. Some of these works combine real image data with DR and photorealistic rendered image data.

In [26] a photorealistic rendered dataset was created for 21 objects of the YCB dataset. Here, the objects are rendered in different scenes with collision properties when falling down. The dataset is intended to accelerate progress in the area of object detection and pose estimation.

In [27], DR is combined with photo realistically rendered image data, for robotic grasping of household objects. Using the data generated in this way, the authors have managed to explore a real-time system for object detection and robot grasping with sufficient accuracy. They also showed that the combination of both domains improved performance as opposed to just one alone.

In the field of human pose estimation, the works of [5] and [29] need to be mentioned. Both works were able to show that the performance of networks can be increased by using synthetic and animated persons, respectively.

The work of [29] generates photorealistic synthetic image data and their ground truth for body part classifications.

In [5], animated persons are integrated into mixed reality environments. The movements were recorded by actors in a motion capture scenario and transferred to 3D scanned meshes. In their experiments, they were able to achieve a 20% increase in performance compared to the largest training set available in this domain.

State-of-the-art models for realistic human body shapes are the SMPL models introduced by [10] and improved by STAR in [11]. According to the authors, the SMPL model is a skinned vertex-based model which represents human shapes in a wide variety. In their work they learn male and female body shape from the CAESAR dataset [13]. Their model is compatible with a wide variety of rendering engines like Unity or Unreal and therefore highly suited to be used in synthetic data generation for humans. There also exist extensions to the SMPL model like MANO and SMPL-H which introduce a deformable hand model into the framework. MANO [15] is learned from 1000 high-resolution 3D scans of various hand poses.

3 Methods

As previously mentioned, real-world data collection in medical intervention rooms is complex, costly, and requires approval from an ethics board and the persons involved. As shown in the previous, DR/SDR can help train an object detection network with sufficient performance in real-world applications.

However, one challenge in dataset generation for the medical intervention space is domain-specific clothing. We argue, that randomizing the clothing textures with random textures would help improve detection rates of the clothing types, but when applied in real-world applications, for example a colored T-shirt would not be distinguishable from the targeted blue colored specific area clothing. For the general detection of cars as in [25] the randomization technique makes sense, but for the domain-specific use case presented here something else should be used in our opinion.

The questions we try to address in this work are:

1. How can health professionals be modeled for synthetic data generation?
2. Which techniques are best suited for SDR/DR clothing generation?
3. Can we close the reality gap further by including greenscreen data (Mixed Reality, MR)?
4. Can the required amount of real data be reduced by using SDR/DR/MR?
5. Can the accuracy be improved when combining real and synthetic data?

For point (1), we argue to use a deformable human shape model like the SMPL models. This provides sufficient variance of different human shapes and sizes. For point (2), we explore two different methods of clothing generation. First, we 3D scan various persons wearing medical clothing and generate a database of different medical clothing scans for each clothing type, which we call SCANS. Second, we commission a professional graphics designer to create assets based on the area clothing, which we call CAD. Regarding point (3), we take images in front of a greenscreen of different persons wearing medical clothing which we label by hand. For point (4), we investigate whether the required amount of real data can be reduced with consistent accuracy by mixing real and synthetic data. Finally, in point (5) we investigate whether the combination of synthetic image data and percentage of real data, improves the accuracy of real data alone.

To address the named questions further, we set up experiments where we want to detect the following classes with the help of the Scaled Yolov4 object detector [30].

The classes to be detected are:

- humans
- area clothing shirt
- area clothing pants
- sterile gown
- medical face mask
- medical hairnet
- medical gloves.

Examples of the medical clothing are given in Fig. 1.

The following section describes the character creation process and why specific tools and models are used.

3.1 Character creation

The medical characters we use in SDR/DR are built through a combination of SMPL body models, textures, animations and clothing assets. Within the following section each of the components used are presented and it is explained why they are used.

A body model is required for the creation of synthetic humans. As the base of our characters we use the male and female model of the SMPL+H model from [15]. The models cover a huge variety of realistic human shapes, which can be randomized through ten blend shapes. We decide to use the extended SMPL+H model instead of the original SMPL model [10]. This is because one of our clothing items are gloves and through the hand rig of the SMPL+H model, we will be able to create more deformations of the glove asset.

The SMPL models alone are surface models without texture. For the generation of humans a human texture is needed. To add more variation and realism to the appearance of the characters, the texture maps from [29] are used. Out of the 930 textures, only 138 (69 of every gender) have been used. This is, as we created our own cloth assets, only the textures of people in undergarments were relevant. Those texture maps were created out of 3D body scans from the CAESAR dataset [13] and cover a variety of skin colors and identities, however all of the faces have been anonymized [29].



Fig. 1 Medical clothing examples like area shirt, pants, mask and glove. These clothing types among others represent the target clothing for our object detection network

When working with synthetic humans in rendering engines, to create a variety of realistic humans, the human pose has to be modified. To provide a variety of realistic body poses, the models were animated through Motion Capture (MoCap) data, which has been captured within our laboratory. We track the movement of 74 joints down to the fingertips. We use an intrinsic Motion Capture suite with the Hand gloves Add-on called Perception Neuron Studio.¹ In order to keep the dataset simple, we only used one animation in our experiments. The potential to add more varying animations is given however.

After defining the body model, body textures and body poses, the medical clothing is needed. Two different approaches are investigated here. One is the generation of medical clothing using a 3D scanner and the other is the generation of designed clothing by a graphic designer. The 3D scanned clothing assets which we call SCANS are created with a 3D scanner called Artec Leo.² A 3D resolution of 0.2 mm was used to capture the medical cloths. For our synthetic training dataset we used clothing scans of 4 male and 4 female models. In this way, variations of the real-world textures, including reflections, wrinkles, colors and surface texture information are collected. After building an initial model from the 3D scanner, we adapt the cloths to fit the standard male and female

¹ Perception Neuron Studio suite and gloves addon: <https://neuronmocap.com/perception-neuron-studio-system>.

² 3D Scanner Artec Leo: <https://www.artec3d.com/portable-3d-scanners/artec-leo>.



Fig. 2 Examples of our 3D-scanned clothing assets with color augmentation

SMPL+H Character using 3D modeling techniques. According to our research, medical clothing usually come in the colors blue, green and light pink. To cover this variation in our dataset we augmented the texture maps. Examples of the scanned and rigged clothing assets can be seen in Fig. 2.

To evaluate the performance of 3D-scanned clothing assets, we compare them to hand designed clothing assets which we call CAD. Therefore, we have asked a designer³ on Fiver to model the clothes. Examples of those assets can be seen in Fig. 3. We first evaluated to what extent freely available assets from the assets stores can be used for this purpose. However, there are no assets available that match our specific clothing in total. Therefore, we have decided to have the assets designed. The designed assets have been processed in the same way as our scanned assets. They are also deformable and are bound to the same rig.

The creation of the synthetic persons is done by means of the rendering engines Unreal Engine 4 and Unity. The NDDS plugin for the Unreal Engine is used to generate the DR image data and a Unity plugin is used to generate the SDR image data.

For the synthetic data generation of DR, an Unreal Engine 4 plugin called NDDS [23] is used. This allows the generation of RGB images at rates similar to real cameras, as well as depth image data and segmentation masks of the scene within Unreal Engine 4. The plugin also creates bounding box labeling data for each object in the

³ azeemdesigns: https://www.fiverr.com/azeemdesigns?source=order_page_user_message_link.

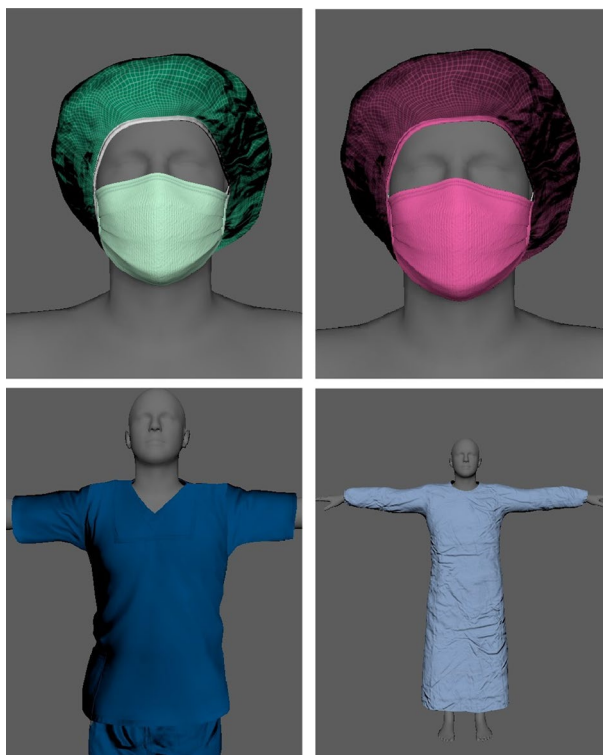


Fig. 3 Examples of the designed clothing assets with color augmentation

scene in 2D and 3D. The tool was specifically developed for DR and therefore provides tools for scene randomization like object or camera position, lighting and distractor objects, among others. Using a modular character blueprint, NDDS enables the generation of synthetic datasets for sterile clothing using 3D scanned clothing or designed clothing. Example images are given in Fig. 4 on the top row. We create two separate datasets, one with SCANS assets and another with CAD assets for DR. An activity diagram, which represents the blueprint for modular character creation in NDDS, is given in Fig. 5.

For dataset generation using SDR, we used a Unity plugin called ML-ImageSynthesis [22] as a base and adapted it to work with the universal rendering pipeline (URP) for quality improvement. Using Unity 2020.3.32f1, additional components have been added to enable an export of additional metadata regarding each generated image such as camera parameters, bounding boxes and world position. SDR is made possible by making use of a variety of custom-made components which allow the randomization of parameters such as lighting, material, texture, position. The plugin ProBuilder provided by Unity was used to build an intervention room based on the target domain of the real dataset (Klinikum). Scene randomization is achieved by utilizing the aforementioned randomization components. An activity diagram, which represents the blueprint for modular character creation in Unity, is given in Fig. 6



Fig. 4 Examples of synthetic RGB image data from Domain Randomization (DR) and Structured Domain Randomization (SDR) datasets (top: DR, bottom: SDR, left: SCANS, right: CAD)

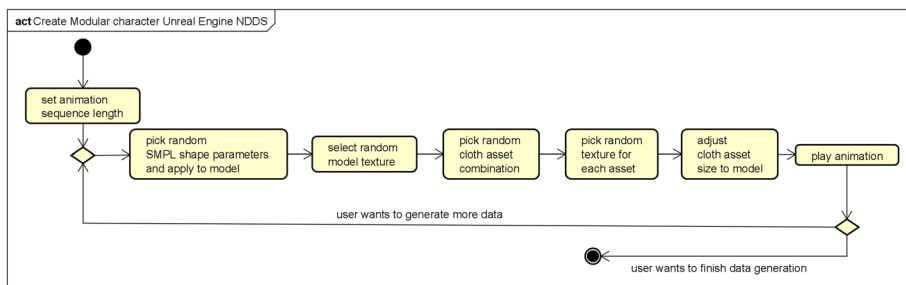


Fig. 5 Activity diagram of the modular character blueprint used in Unreal with NDDS

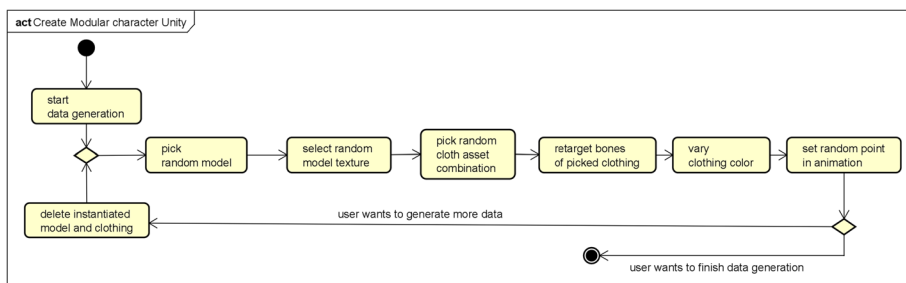


Fig. 6 Activity diagram of the character creation used in Unity for SDR

3.2 Datasets

To investigate the potential accuracy difference between SCANS, CAD and the combination with real data, different datasets were generated.

First, synthetic datasets of DR and SDR were generated for both SCANS and CAD clothing, using the presented pipelines in Unreal-Engine und Unity. These datasets are used to experiment to find out whether scanned clothing or designed clothing give better results.

Second, a dataset in front of a greenscreen was collected which we call Mixed-Reality (MR). It consists of 8 persons in the training dataset and 2 persons in the validation dataset. The recorded persons move in front of the green screen with a certain grasping motion, which is also used as motion animation for the synthetic data. This dataset aims to further close the reality gap between the synthetic image data and the real data by introducing real data in a mixed reality scenario without having to record data in the target domain.

Finally, a dataset of the target domain was recorded which we call Klinikum. It serves as a baseline comparison for all models and also represents the test data. This results in 331 labeled test data. In order to get a sufficient amount of testdata from the available, we decided to use a different split compared to the other datasets here.

In the following sections the lines Klinikum(100) or Klinikum(15) represent all available Klinikum train data, respectively, 15% randomly chosen train data. The lines real(100) and real(15) mean the same.

All datasets are divided into training and validation data.

Examples of real data in front of the green screen with exchanged background can be seen in Fig. 7. Examples of the synthetic data can be seen in Fig. 4 and finally examples from the clinical test data can be seen in Fig. 8.

Table 1 gives a breakdown of the sizes and distributions of the datasets.



Fig. 7 Examples of the greenscreen dataset with exchanged backgrounds



Fig. 8 Example images of the Klinikum dataset

Table 1 This table shows data distribution in the different datasets

| Dataset distributions | | | |
|-----------------------|--------|------------|------|
| Dataset name | Train | Validation | Test |
| DRscans | 10,000 | 1000 | / |
| DRcad | 10,000 | 1000 | / |
| SDRscans | 8000 | 2000 | / |
| SDRcad | 8000 | 2000 | / |
| MR-DR | 6443 | 1536 | / |
| Klinikum | 660 | 110 | 331 |

4 Experiments

Experiments were performed to investigate whether and how well SCANS compare to CAD clothing for detection in the medical environment. Additionally, experiments were carried out to determine if a percentage of real data together with synthetic data can achieve sufficient accuracy or even surpass real data alone. Finally, MR data were included in the experiments to determine whether they could further close the reality gap.

For our experiments, we used Scaled-Yolov4 [30] implementation from GitHub.⁴ At first, 6 different baseline networks were trained to show a basic comparison of the different methods and to determine if SCANS or CAD clothing provide better results. These baseline models include trainings with synthetic (DRscans, DRcad, SDRscans, SDRcad), mixed-reality (MR-DR) and real data from the clinic domain (Klinikum train).

Training was conducted with YOLOv4-p5 weights and default finetune parameters provided by Scaled Yolo-V4 GitHub repository. Only Mosaic Augmentation ratio parameters α and β were increased from 8.0 to 20.0. This is to weaken the blending images effect in the augmentation of the used implementation. Additionally a green-channel augmentation was used when MR data were present in the training dataset in order to reduce the greenscreen spill effect which we had troubles with in some classes. Here, we

⁴ <https://github.com/WongKinYiu/ScaledYOLOv4>.

Table 2 Results on Klinikum test-set for baseline trainings

| All | | | | |
|----------------|--------------|--------------|--------------|--------------|
| Experiment | <i>mAP</i> | <i>mAP50</i> | <i>P</i> | <i>R</i> |
| Klinikum train | 81.60 | 98.28 | 88.99 | 98.73 |
| DRscans | 46.06 | 75.51 | 76.89 | 77.83 |
| DRcad | 51.50 | 77.85 | 76.40 | 81.56 |
| SDRscans | 65.52 | 87.12 | 77.92 | 89.31 |
| SDRcad | 67.44 | 90.72 | 81.98 | 92.02 |
| MR-DR* | 60.94 | 82.29 | 79.54 | 84.04 |

The bold values generally represent the best results from the respective columns

*Additional green-channel augmentation used

Table 3 Results on Klinikum test-set for follow-up experiments

| All | | | | |
|------------------|--------------|--------------|--------------|--------------|
| Experiment | <i>mAP</i> | <i>mAP50</i> | <i>P</i> | <i>R</i> |
| Klinikum(100) | 81.95 | 98.57 | 88.88 | 98.85 |
| SDR+MR+real(100) | 83.35 | 98.14 | 88.41 | 98.59 |
| Klinikum(15) | 77.52 | 96.67 | 87.83 | 97.37 |
| SDR+MR+real(15) | 80.05 | 96.92 | 87.06 | 97.64 |
| SDR+MR | 72.00 | 92.27 | 83.81 | 94.07 |

The bold values generally represent the best results from the respective columns

For these experiments, we used the additional green-channel augmentation on all reported trainings

try to establish a baseline for the MR-DR data. We experimentally found out that using the green-channel augmentation helps the accuracy.

All networks were trained for 300 epochs and achieved convergence. All trained models were tested on the Klinikum test-set with IoU-threshold: 0.5 and confidence-threshold: 0.2. The used Yolov4 network was yolov4-p5, image size setting was 896 in training and test and the pretrained weights provided were used. The results of the baseline models are displayed in Table 2.

The results show, that CAD-based synthetic data generally give better results than SCAN based data on this experiment. This is why we use the SDRcad dataset for all follow-up experiments.

To investigate by how much the amount of real data can be reduced when used together with synthetic or MR data, while maintaining sufficient accuracy, experiments were conducted with a percentage distribution of the Klinikum training data. Here it is our main goal to find out whether using synthetic data together with MR data and percentages of real data surpasses the accuracy of real data alone. We choose 15% of real data because this results in 99 remaining training images which we argue is an adequate amount of image data which can be labeled by hand. We decided to use the mosaic augmentation during these experiments as well and use all datasets as training data instead of a finetune experiment. We argue that the network can better learn relevant features while maintaining the advantages of the additional synthetic data when seeing a variation of all used datasets mixed together with mosaic augmentation as when only finetuning. During these experiments, we decided

to include the aforementioned green-channel augmentation on all trainings. Additionally the real data runs were trained with the same number of optimization steps in order to ensure that the model converges while using less training data.

The follow-up results with SDRcad, MR-DR as well as real data are shown in Table 3.

5 Results

The results of the first experiment can be seen in Table 2. When comparing the SCANS clothing and the CAD clothing in the DR and SDR scenarios, both times the data sets with CAD clothing provide better results. This was surprising for us at this point. The possible reasons for this are discussed in the chapter discussion.

Similarly, when looking at the results of the individual classes in Table 4 it can be seen that, with the exception of the Mask class, the CAD clothing gives better results than the SCANS clothing in every case.

It is also clear from the results that SDR is superior to DR. This was to be expected based on previous work in this area, since in the presented experiments of SDR the environment is enriched with the objects present in the clinic and thus the network is better adapted to distractions.

While the MR-DR results are inferior to the SDR in many classes, they are superior to the DR except for the Gown class. The reasons for the poor performance of the Gown class in experiments with SDR have already been mentioned in the chapter 3.1. Here, the greenscreen spill was particularly negative, which is why the additional green-channel augmentation was applied.

The evaluation metric used is the mean Average Precision (mAP) with 2 different Intersection over Union (IoU) thresholds. The two thresholds are 0.5:0.95 for mAP and 0.5 for mAP50 as used in the Scaled-Yolov4 implementation [30].

For the Pants class, the MR-DR achieved the best results in this experiment for the mAP. For the mAP50, on the other hand, the situation is the same as for the other classes. This difference can probably be attributed to the inaccurate border of the pants under the shirt. Here, a possible different labeling of the real image data (Klinikum, MR-DR) compared to the automated labeling with the synthetic image data explains the difference.

In general, the two classes Mask and Glove deliver the worst results. This is also the case for the mAP50 category compared to the effect described for the Pants class. This can be attributed to the relatively small size of these classes. In the test data, difficult cases are included, which show the persons from the side. In these cases, the mask or the gloves are just visible from the side and the bounding box area covers a few pixels. This effect can be seen in the real data as well. Here, the Klinikum train dataset achieves an accuracy of 53.66% at mAP, whereas mAP50 is again at 95.85%. This effect can also be seen in the other training datasets but is less strong.

The results of the follow-up experiment, which examines the comparison of synthetic image data along with mixed reality data and a percentage distribution of real data, are shown in Table 3.

Here, the joint training dataset from SDR+MR improves the accuracy of the two individual datasets from the first experiment. However, the difference compared to 100% real data and even 15% real data is still present with the mAP and smaller in

mAP50. Nevertheless this result is of great interest for future work and experiments, as it displays a way to avoid using real data from the target domain altogether. The possibilities of mixed reality together with synthetic data should therefore be further investigated.

Furthermore, it can be seen that adding SDR+MR data to 15% and 100% real data increases the accuracy of the detections compared to real data alone. For 15% real data this is an increase of 2.53% and for 100% real data this is an increase of 1.4% for mAP.

Looking at the results of the individual classes, which are shown in Table 5, the dataset with SDR+MR+Klinikum(100) gives the best results for all classes except Mask.

Regarding the classes Mask and Glove which gave the worst results in the first experiment, the accuracy can be improved when merging SDR+MR data. This is another indication of the potential of synthetic and mixed reality data for applications in the medical field to significantly reduce the amount of real data required.

When looking at the results of the Pants class, which in the first experiment achieved 55.32% on SDR in mAP, this can be improved to 80.22% by combining with MR data. Likewise, the influence of the MR data with the greenscreen spill effect can be seen with the Gown class. The combination of SDR+MR data largely eliminates the negative influence of the MR data from the first experiment. This is an indicator for the noticeable low accuracy of this class, possibly due to the greenscreen spill.

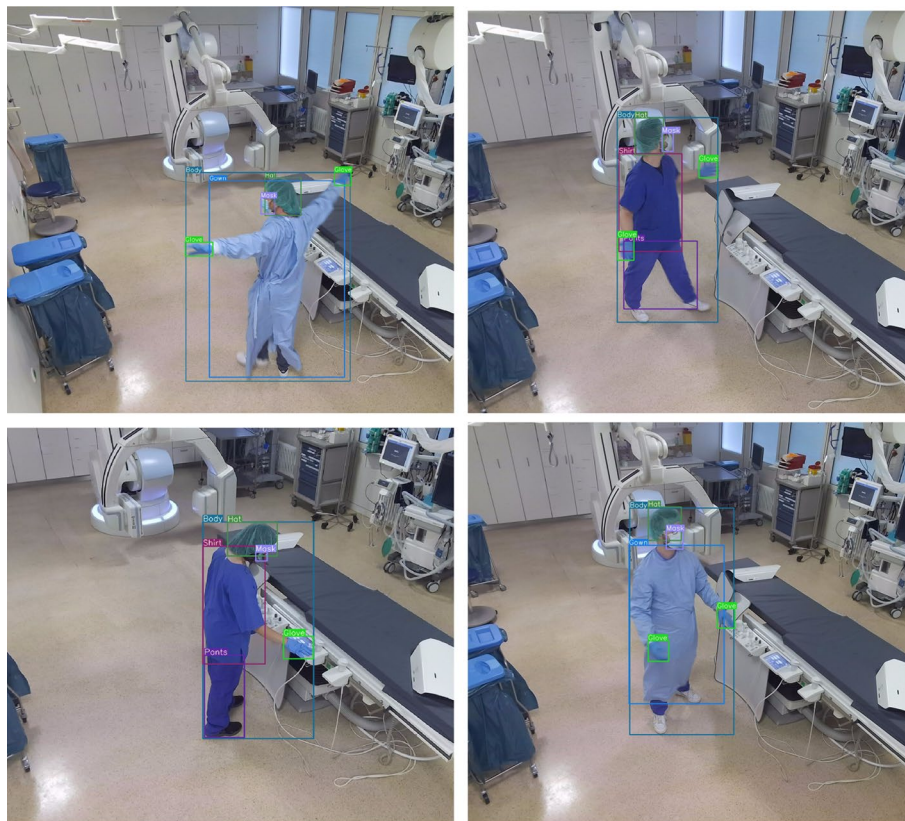


Fig. 9 Inference results with SDR+MR+real(100) data trained net. Only for the visualization of the image result here, we used a slightly higher confidence threshold of 0.4 and IoU-threshold of 0.5 compared to the presented results of all tables

Inference result images of the training with SDR+MR+real(100) data can be seen in Fig. 9. Only for the visualization of the image result here, we used a slightly higher confidence threshold of 0.4 compared to the presented results of all tables (0.2).

6 Conclusion

We were able to show that the use of SMPL models together with scanned or designed medical clothing is a suitable method for modeling healthcare professionals for artificial intelligence questions in the intervention space using the example of medical clothing detection. During our experiments we found out that the designed clothing generally performed better on our test dataset than the 3D scanned cloths. This result surprised us, as we expected the potentially more accurate textures of the 3D scan to have a positive impact on detection rates. However, according to the results, it cannot be ruled out that artifacts in the rendering pipeline or pre-processing pipeline that we did not detect have an influence on this. Additionally, further work can investigate whether scanned clothing should be designed to be more deformable, as this can combine the advantage of scanned textures along with realistic movement of the fabric. In order to make a final statement about the potential of 3D scanned clothing for the modeling of health professionals, further experiments should be conducted. Using Mixed-Reality data together with the synthetic data closed the gap further and while the margin is quite small, we could show that when using synthetic, mixed reality and 15% real data the remaining gap towards 100% real data could be nearly closed. Generally we could show that when using synthetic and mixed-reality together with a percentage of real data, it surpasses real data alone.

This is a good sign for the potential of synthetic and mixed-reality data in questions around medical interventions, as they contain enough information to close the reality gap. A trajectory with multiple percentage distributions of real data together with SDR+MR data is interesting for a larger test data set with multiple healthcare professionals and is the subject of future work. In the results shown, it has already been demonstrated that the fusion of SDR+MR data together with the real data improves the accuracy.

For questions in the intervention space, mixed-reality in particular allows data to be acquired outside the target domain to minimize privacy challenges. In future work, methods should be explored to reduce the greenscreen spill effect during data generation and to visualize the resulting data in more complex scenes similar to SDR. For this purpose, the use of deep learning networks for image enhancement is interesting to investigate.

In conclusion, the presented modeling of health professionals is a promising method to solve the problem of missing datasets from medical intervention rooms. We will further investigate it for various tasks in the medical field.

Appendix

In this appendix chapter, the full table results of the baseline experiment and the follow-up experiments are provided. Additionally to the all category, we report the results for all detection classes (Body, Gown, Shirt, Pants, Hat, Mask, Glove).

Full results

Table 4 shows the results of the baseline experiments and Table 5 shows the results of the follow-up experiments.

Table 4 Results on Klinikum test-set for baseline trainings

| All | | | | |
|-------------------|--------------|--------------|--------------|--------------|
| Experiment | mAP | mAP50 | P | R |
| Klinikum train | 81.60 | 98.28 | 88.99 | 98.73 |
| DRscans | 46.06 | 75.51 | 76.89 | 77.83 |
| DRcad | 51.50 | 77.85 | 76.40 | 81.56 |
| SDRscans | 65.52 | 87.12 | 77.92 | 89.31 |
| SDRcad | 67.44 | 90.72 | 81.98 | 92.02 |
| MR-DR* | 60.94 | 82.29 | 79.54 | 84.04 |
| Body | | | | |
| Klinikum train | 97.02 | 99.49 | 88.03 | 100 |
| DRscans | 75.24 | 96.26 | 84.57 | 96.07 |
| DRcad | 77.05 | 98.39 | 76.76 | 98.79 |
| SDRscans | 92.47 | 99.50 | 89.95 | 100 |
| SDRcad | 92.88 | 99.50 | 85.09 | 100 |
| MR-DR* | 80.88 | 95.29 | 82.55 | 95.77 |
| Gown | | | | |
| Klinikum train | 93.41 | 99.5 | 100 | 100 |
| DRscans | 66.16 | 87.5 | 100 | 87.22 |
| DRcad | 78.96 | 94.17 | 87.41 | 93.99 |
| SDRscans | 79.22 | 94.77 | 87.25 | 97.74 |
| SDRcad | 85.86 | 99.49 | 99.25 | 99.25 |
| MR-DR* | 50.18 | 57.5 | 100 | 57.14 |
| Shirt | | | | |
| Klinikum train | 94.28 | 99.47 | 55.03 | 99.49 |
| DRscans | 81.17 | 98.19 | 59.45 | 98.48 |
| DRcad | 82.82 | 97.97 | 56.85 | 98.48 |
| SDRscans | 84.72 | 95.28 | 56.76 | 95.45 |
| SDRcad | 92.22 | 99.41 | 55.65 | 99.49 |
| MR-DR* | 91.41 | 97.72 | 44.85 | 98.99 |
| Pants | | | | |
| Klinikum train | 86.81 | 98.5 | 98.98 | 98.48 |
| DRscans | 34.13 | 88.44 | 75.21 | 90.40 |
| DRcad | 48.14 | 92.22 | 92.89 | 92.42 |
| SDRscans | 54.45 | 91.31 | 90.05 | 91.41 |
| SDRcad | 55.32 | 95.61 | 91.83 | 96.46 |
| MR-DR* | 77.17 | 92.26 | 78.88 | 92.42 |
| Hat | | | | |
| Klinikum train | 85.62 | 99.27 | 98 | 100 |
| DRscans | 44.71 | 78.73 | 88.27 | 80.61 |
| DRcad | 54.99 | 94.41 | 97.89 | 94.90 |
| SDRscans | 78.44 | 95.59 | 67.94 | 99.49 |

Table 4 (continued)

| Hat | | | | |
|----------------|--------------|--------------|--------------|-----------|
| SDRcad | 79.42 | 97.46 | 97.45 | 97.45 |
| MR-DR* | 70.51 | 98.24 | 95.57 | 98.98 |
| Mask | | | | |
| Klinikum train | 53.66 | 95.85 | 88.8 | 96.52 |
| DRscans | 8.58 | 44.07 | 61.42 | 52.61 |
| DRcad | 6.09 | 29.59 | 50.95 | 46.52 |
| SDRscans | 28.33 | 63.54 | 78.77 | 66.09 |
| SDRcad | 25.24 | 71.91 | 69.29 | 76.52 |
| MR-DR* | 25.33 | 73.75 | 82.65 | 78.70 |
| Glove | | | | |
| Klinikum train | 60.43 | 95.88 | 94.10 | 96.59 |
| DRscans | 12.40 | 35.36 | 69.33 | 39.39 |
| DRcad | 12.44 | 38.17 | 72.02 | 45.83 |
| SDRscans | 40.99 | 69.88 | 74.72 | 75 |
| SDRcad | 41.10 | 71.63 | 75.29 | 75 |
| MR-DR* | 31.09 | 61.26 | 72.31 | 66.29 |

The bold values generally represent the best results from the respective columns

*Additional green-channel augmentation used

Abbreviations

Table 6 gives an overview of abbreviations used in the presented work. They are sorted in alphabetical order with sections for starting letters present in the work.

Table 5 Results on Klinikum test-set for follow-up experiments

| All | | | | |
|-------------------|--------------|--------------|--------------|--------------|
| Experiment | mAP | mAP50 | P | R |
| Klinikum(100) | 81.95 | 98.57 | 88.88 | 98.85 |
| SDR+MR+real(100) | 83.35 | 98.14 | 88.41 | 98.59 |
| Klinikum(15) | 77.52 | 96.67 | 87.83 | 97.37 |
| SDR+MR+real(15) | 80.05 | 96.92 | 87.06 | 97.64 |
| SDR+MR | 72.00 | 92.27 | 83.81 | 94.07 |
| Body | | | | |
| Klinikum(100) | 97.13 | 99.5 | 87.33 | 100 |
| SDR+MR+real(100) | 97.93 | 99.5 | 86.42 | 100 |
| Klinikum(15) | 94.24 | 99.5 | 92.46 | 100 |
| SDR+MR+real(15) | 96.84 | 99.5 | 86.2 | 100 |
| SDR+MR | 93.83 | 99.47 | 87.00 | 99.1 |
| Gown | | | | |
| Klinikum(100) | 93.29 | 99.5 | 99.25 | 100 |
| SDR+MR+real(100) | 96.30 | 99.5 | 100 | 100 |
| Klinikum(15) | 87.88 | 99.5 | 100 | 100 |
| SDR+MR+real(15) | 95.57 | 99.5 | 99.25 | 100 |
| SDR+MR | 85.11 | 95.47 | 99.22 | 95.49 |

Table 5 (continued)

| Shirt | | | | |
|------------------|--------------|--------------|--------------|--------------|
| Klinikum(100) | 94.51 | 99.46 | 55.49 | 99.5 |
| SDR+MR+real(100) | 96.4 | 99.37 | 57.10 | 99.5 |
| Klinikum(15) | 90.43 | 98.07 | 57.18 | 98.5 |
| SDR+MR+real(15) | 95.62 | 99.38 | 60.43 | 99.5 |
| SDR+MR | 94.17 | 99.31 | 54.42 | 99.5 |
| Pants | | | | |
| Klinikum(100) | 87.16 | 99.39 | 99.49 | 98.99 |
| SDR+MR+real(100) | 88.22 | 98.49 | 99.49 | 98.48 |
| Klinikum(15) | 81.70 | 97.40 | 95.12 | 98.48 |
| SDR+MR+real(15) | 85.78 | 97.5 | 97.97 | 97.47 |
| SDR+MR | 80.22 | 95.26 | 91.30 | 95.45 |
| Hat | | | | |
| Klinikum(100) | 84.77 | 99.29 | 97.51 | 100 |
| SDR+MR+real(100) | 86.90 | 99.3 | 98 | 100 |
| Klinikum(15) | 83.77 | 98.59 | 96.06 | 99.5 |
| SDR+MR+real(15) | 84.91 | 99.27 | 97.51 | 100 |
| SDR+MR | 77.6 | 99.22 | 97.51 | 100 |
| Mask | | | | |
| Klinikum(100) | 55.42 | 95.70 | 91.32 | 96.09 |
| SDR+MR+real(100) | 54.70 | 93.84 | 88.62 | 94.78 |
| Klinikum(15) | 50.30 | 93.17 | 85.66 | 93.48 |
| SDR+MR+real(15) | 44.28 | 90.95 | 82.81 | 92.17 |
| SDR+MR | 27.61 | 78.89 | 76.36 | 85.65 |
| Glove | | | | |
| Klinikum(100) | 61.39 | 97.11 | 91.79 | 97.35 |
| SDR+MR+real(100) | 62.98 | 97.0 | 89.24 | 97.35 |
| Klinikum(15) | 54.29 | 90.46 | 88.32 | 91.67 |
| SDR+MR+real(15) | 57.32 | 92.32 | 85.27 | 94.32 |
| SDR+MR | 45.46 | 78.29 | 80.88 | 83.33 |

The bold values generally represent the best results from the respective columns

Green-channel augmentation used on all experiments

Table 6 Abbreviations table in alphabetic order

| Abbreviations table | |
|----------------------------|-----------------------------------|
| A | |
| AP | Average-precision |
| C | |
| CAD | Designed medical clothing data |
| D | |
| DR | Domain-Randomization |
| I | |
| IoU | Intersection over Union |
| M | |
| mAP | Mean Average-Precision |
| mAP50 | Mean Average-Precision at 50% IoU |

Table 6 (continued)**Abbreviations table**

| | |
|--------|--|
| MR | Mixed-Reality |
| MR-DR | Mixed-Reality Domain-Randomization |
| N | |
| NDDS | NVIDIA Deep-Learning Dataset Synthesizer |
| P | |
| P | Precision |
| R | |
| R | Recall |
| S | |
| Scans | 3D scanned medical clothing data |
| SDR | Structured-Domain-Randomization |
| SMPL | Skinned-Person Linear Model |
| SMPL+H | Skinned-Person Linear Model + Hands |

Acknowledgements

We thank the ESM Institute, the Clinic for Radiology and Nuclear Medicine, and the project partners at SIEMENS Healthineers for their continuous and substantial support during the realization of this work. Hannah Teufel: During graduate student work at the ESM-Institute. Not affiliated with ESM-Institute anymore.

Author contributions

PS did the conception, selection of methodology, execution of experiments, and wrote most of the manuscript. HT, RV and IE labeled image data and generated the synthetic image data. They also contributed to the chapters Related Work and Methods. YB, MP, MV, SD, and NR made important contributions to the design, application, and performance of the experiments in discussions. All authors read and approved the final manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL. This research project is part of the Research Campus M2OLIE and funded by the German Federal Ministry of Education and Research (BMBF) within the Framework "Forschungscampus—public-private partnership for Innovations". The grant with code 13GW0389C was received by Prof. Dr. Marcus Vetter and Patrick Schülein is employed through it. Hannah Teufel, Ronja Vorpahl und Indira Emter are employed as research assistants through it. The grant with code 13GW0389B was received by Prof. Dr. Steffen Diehl and Prof. Dr. med. Nils Rathmann. The ESM Institute is supported with additional funds within a cooperation agreement with Siemens Healthcare GmbH. The funding was received by Prof. Dr. Marcus Vetter and Yannick Bukschat is employed through it.

Availability of data and materials

In addition to generated synthetic image data, the datasets used also consist of real image data which includes personal data. This does not allow us to provide the data online. However, it is possible to obtain the data from the author after a justified request through a bilaterally established data protection agreement.

Declarations**Informed consent**

This article does not contain patient data. Informed consent was obtained from all individual participants included in the study. The methods and information presented in this work are based on research and are not commercially available.

Competing interests

The authors declare that they have no conflict of interest.

Received: 14 February 2023 Accepted: 18 July 2023

Published: 2 August 2023

References

1. V. Belagiannis, X. Wang, H. Beny Ben Shitrit, K. Hashimoto, R. Stauder, Y. Aoki, M. Kranzfelder, A. Schneider, P. Fua, S. Ilic, H. Feussner, N. Navab, Parsing human skeletons in an operating room. *Mach. Vis. Appl.* (2016). <https://doi.org/10.1007/s00138-016-0792-4>
2. A. Bochkovskiy, C.-Y. Wang, H.-Y.M. Liao, Yolov4: Optimal speed and accuracy of object detection. <https://arxiv.org/pdf/2004.10934.pdf>. Accessed 24 Nov 2022

3. S. Borkman, A. Crespi, S. Dhakad, S. Ganguly, J. Hogins, Y.-C. Jhang, M. Kamalzadeh, B. Li, S. Leal, P. Parisi, C. Romero, W. Smith, A. Thaman, S. Warren, N. Yadav, Unity perception: generate synthetic data for computer vision. <http://arxiv.org/pdf/2107.04259v2.pdf>. Accessed 24 Nov 2022
4. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016), IEEE. <https://doi.org/10.1109/cvpr.2016.90>
5. C. Ionescu, D. Papava, V. Olaru, C. Sminchisescu, Human3.6m: large scale datasets and predictive methods for 3d human sensing in natural environments. *IEEE Trans. Pattern Anal. Mach. Intell.* **36**(7), 1325–1339 (2014). <https://doi.org/10.1109/TPAMI.2013.248>
6. S. James, A.J. Davison, E. Johns, Transferring end-to-end visuomotor control from simulation to real world for a multi-stage task. *CoRR abs/1707.02267* (2017). <http://arxiv.org/pdf/1707.02267.pdf>. Accessed 24 Nov 2022
7. Y. LeCun, 1.1 deep learning hardware: past, present, and future. In IEEE International Solid-State Circuits Conference—(ISSCC) (2019). IEEE (2019). <https://doi.org/10.1109/isscc.2019.8662396>
8. T. Lin, M. Maire, S.J. Belongie, L.D. Bourdev, R.B. Girshick, J. Hays, P. Perona, D. Ramanan, P. Dollár, C.L. Zitnick, Microsoft COCO: common objects in context. *CoRR abs/1405.0312* (2014). <http://arxiv.org/abs/1405.0312>. Accessed 24 Nov 2022
9. T.-Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollár, Focal loss for dense object detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **42**(2), 318–327 (2020). <https://doi.org/10.1109/tpami.2018.2858826>
10. M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, M.J. Black, SMPL: A skinned multi-person linear model. *ACM Trans. Graphics (Proc. SIGGRAPH Asia)* **34**(6), 24:81–24:816 (2015). <https://doi.org/10.1145/2816795.2818013>
11. A.A.A. Osman, T. Bolkart, M.J. Black, STAR: A sparse trained articulated human body regressor. In European Conference on Computer Vision (ECCV), pp. 598–613 (2020). https://doi.org/10.1007/978-3-030-58539-6_36
12. A. Prakash, S. Bochoon, M. Brophy, D. Acuna, E. Cameracci, G. State, O. Shapira, S. Birchfield, Structured domain randomization: Bridging the reality gap by context-aware synthetic data. In 2019 International Conference on Robotics and Automation (ICRA) (2019), IEEE. <https://doi.org/10.1109/icra.2019.8794443>
13. K. Robinette, S. Blackwell, H. Daanen, M. Boehmer, S. Fleming, Civilian American and European surface anthropometry resource (caesar), final report. volume 1. summary. 74. <https://www.humanics-es.com/CAESARvol1.pdf>. Accessed 24 Nov 2022
14. V.F. Rodrigues, R.S. Antunes, L.A. Seewald, R. Bazo, E.S. dos Reis, U.J. dos Santos, R. da R. Righi, L.G. da S., C.A. da Costa, F.L. Bertollo, A. Maier, B. Eskofier, T. Horz, M. Pfister, R. Fahrig, A multi-sensor architecture combining human pose estimation and real-time location systems for workflow monitoring on hybrid operating suites. *Future Gener. Comput. Syst.* **135**, 283–298 (2022). <https://doi.org/10.1016/j.future.2022.05.006>
15. J. Romero, D. Tzionas, M.J. Black, Embodied hands: Modeling and capturing hands and bodies together. *ACM Trans. Graph. (Proc. SIGGRAPH Asia)* **36**, 6 (2017). <https://doi.org/10.1145/3130800.3130883>
16. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A.C. Berg, L. Fei-Fei, ImageNet large scale visual recognition challenge. *Int. J. Comput. Vis. (IJCV)* **115**(3), 211–252 (2015). <https://doi.org/10.1007/s11263-015-0816-y>
17. F. Sadeghi, S. Levine, CAD2RL: Real single-image flight without a single real image. In Robotics: Science and Systems XIII (2017), Robotics: Science and Systems Foundation. <https://doi.org/10.15607/rss.2017.xiii.034>
18. A. Sharghi, H. Haugerud, D. Oh, O. Mohareri, Automatic operating room surgical activity recognition for robot-assisted surgery. *CoRR abs/2006.16166* (2020). https://doi.org/10.1007/978-3-030-59716-0_37
19. V. Srivastav, T. Issenhuht, K. Abdolrahim, M. de Mathelin, A. Gangi, N. Padoy, Mvor: A multi-view rgb-d operating room dataset for 2d and 3d human pose estimation
20. V. Sze, Y.-H. Chen, J. Emer, A. Suleiman, Z. Hardware. Zhang, for machine learning: Challenges and opportunities. In IEEE Custom Integrated Circuits Conference (CICC) (apr 2018). IEEE (2018). <https://doi.org/10.1109/cicc.2018.8357072>
21. M. Tan, R. Pang, Q.V. Le, EfficientDet: Scalable and efficient object detection. In 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2020), IEEE. <https://doi.org/10.1109/cvpr42600.2020.01079>
22. U. Technologies, ML-ImageSynthesis, 2017. <https://bitbucket.org/Unity-Technologies/ml-imagesynthesis/src/master/>. Accessed 24 Nov 2022
23. T. To, J. Tremblay, D. McKay, Y. Yamaguchi, K. Leung, A. Balanon, J. Cheng, W. Hodge, S. Birchfield, NDDS: NVIDIA Deep Learning Dataset Synthesizer (2018). https://github.com/NVIDIA/Dataset_Synthesizer. Accessed 24 Nov 2022
24. J. Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, P. Abbeel, Domain randomization for transferring deep neural networks from simulation to the real world. In 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (2017), IEEE. <https://doi.org/10.1109/iro.2017.8202133>
25. J. Tremblay, A. Prakash, D. Acuna, M. Brophy, V. Jampani, C. Anil, T. To, E. Cameracci, S. Bochoon, S. Birchfield, Training deep networks with synthetic data: Bridging the reality gap by domain randomization. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 1082–10828 (2018). https://openaccess.thecvf.com/content_cvpr_2018_workshops/papers/w14/Tremblay_Training_Deep_Networks_CVPR_2018_paper.pdf. Accessed: 24.11.2022
26. J. Tremblay, T. To, S. Birchfield, Falling things: A synthetic dataset for 3d object detection and pose estimation. In 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (2018), IEEE. <https://doi.org/10.1109/cvprw.2018.00275>
27. J. Tremblay, T. To, B. Sundaralingam, Y. Xiang, D. Fox, S. Birchfield, Deep object pose estimation for semantic robotic grasping of household objects. *arXiv preprint* <http://arxiv.org/abs/1809.10790> (2018). Accessed 24 Nov 2022
28. A.P. Twinanda, E.O. Alkan, A. Gangi, M. de Mathelin, N. Padoy, Data-driven spatio-temporal RGBD feature encoding for action recognition in operating rooms. *Int. J. Comput. Assist. Radiol. Surg.* **10**(6), 737–747 (2015). <https://doi.org/10.1007/s11548-015-1186-1>
29. G. Varol, J. Romero, X. Martin, N. Mahmood, M.J. Black, I. Laptev, C. Schmid, Learning from synthetic humans. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 4627–4635 (2017). https://openaccess.thecvf.com/content_cvpr_2017/papers/Varol_Learning_From_Synthetic_CVPR_2017_paper.pdf. Accessed 24 Nov 2022

30. C.-Y. Wang, A. Bochkovskiy, H.-Y.M. Liao, Scaled-YOLOv4: Scaling cross stage partial network. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 13029–13038 (2021). https://openaccess.thecvf.com/content/CVPR2021/papers/Wang_Scaled-YOLOv4_Scaling_Cross_Stage_Partial_Network_CVPR_2021_paper.pdf. Accessed 24 Nov 2022
31. F. Zhang, J. Leitner, M. Milford, P. Corke, Sim-to-real transfer of visuo-motor policies for reaching in clutter: Domain randomization and adaptation with modular networks. CoRR abs/1709.05746v1 (2017). <https://arxiv.org/pdf/1709.05746v1.pdf>. Accessed 24 Nov 2022

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
