Augmented Synthetic Dataset with Structured Light to Develop Ai-Based Methods for Breast Depth Estimation

Bruno Duarte
2Ai - School of Technology, IPCA, Barcelos, Portugal; Algoritmi Center, School of Engineering, University of Minho, Guimarães, Portugal
bduarte@ipca.pt

Bruno Oliveira
2Ai - School of Technology, IPCA, Barcelos, Portugal; Algoritmi Center, School of Engineering, University of Minho, Guimarães, Portugal; Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, Braga, Portugal; ICVS/3B's - PT Government Associate Laboratory, Braga/Guimarães, Portugal
boliveira@ipca.pt

Helena, R., Torres
2Ai - School of Technology, IPCA, Barcelos, Portugal; Algoritmi Center, School of Engineering, University of Minho, Guimarães, Portugal; Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, Braga, Portugal; ICVS/3B's - PT Government Associate Laboratory, Braga/Guimarães, Portugal
htorres@ipca.pt

Pedro Morais
2Ai - School of Technology, IPCA, Barcelos, Portugal
pmorais@ipca.pt

Jaime, C., Fonseca
Algoritmi Center, School of Engineering, University of Minho, Guimarães, Portugal
jaime@dei.uminho.pt

João, L., Vilaça*
2Ai - School of Technology, IPCA, Barcelos, Portugal
jvilaca@ipca.pt

ABSTRACT

Breast interventions are common healthcare procedures that normally require experienced professionals, expensive setups, and high execution times. With the evolution of robot-assisted technologies and image analysis algorithms, new methodologies can be implemented to facilitate the interventions in this area. To enable the introduction of robot-assisted approaches for breast procedures, strategies with real-time capacity and high precision for 3D breast shape estimation are required. In this paper, it is proposed to fuse the structured light (SL) and deep learning (DL) techniques to perform the depth estimation of the breast shape with high precision. First, multiple synthetic datasets of breasts with different printed patterns, resembling the SL technique, are created. Thus, it is possible to take advantage of the pattern's deformation induced by the breast surface in order to improve the quality of the depth information and to study the most suitable design. Then, distinct DL architectures, taken from the literature, were implemented to estimate the breast shape from the created datasets and study the DL architectures' influence on depth estimation. The results obtained with the introduction of a yellow grid pattern, composed of thin stripes, fused with the DenseNet-161 architecture achieved the best results. Overall, the current study demonstrated the potential of the proposed practice for breast depth estimation or other human body parts in the future when we rely exclusively on 2D images.

CCS CONCEPTS

• **Computing methodologies**; • **Artificial intelligence**; • **Computer vision**; • **Computer vision problems**; • **Reconstruction**;

KEYWORDS

architectures, breast interventions, deep learning, depth estimation, patterns, structured light, synthetic dataset

ACM Reference Format:
Bruno Duarte, Bruno Oliveira, Helena, R., Torres, Pedro Morais, Jaime, C., Fonseca, and João, L., Vilaça. 2022. Augmented Synthetic Dataset with Structured Light to Develop Ai-Based Methods for Breast Depth Estimation. In *2022 9th International Conference on Bioinformatics Research and Applications (ICBRA 2022), September 18–20, 2022, Berlin, Germany*. ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3569192.3569206

1 INTRODUCTION

Nowadays, there are several breast interventions performed with different objectives [1]. Most of them are related to medical problems, such as breast cancer, which demands a breast biopsy in case of cancer suspicion and, if confirmed and suitable, requires tumor removal. This procedure demands experienced medical staff and image-guided methodologies, being challenging to execute and expensive [2]. Other interventions are therapeutic, such as cyst aspiration and abscess drainage, or palliative, such as pleural effusion drainage or nerve block. Meanwhile, there are also some interventions related to aesthetic problems, such as breast reconstruction or breast augmentation. From medical to aesthetic issues, all of them can benefit from new advanced technologies, such as medical guidance through medical image analysis or robot-guided procedures

*2Ai - School of Technology, IPCA, Barcelos, Portugal; pmorais@ipca.pt

ICBRA 2022, September 18–20, 2022, Berlin, Germany

ACM ISBN 978-1-4503-9686-8/22/09…$15.00
https://doi.org/10.1145/3569192.3569206

[3]. Still, to include such strategies in breast interventions, the 3D shape estimation of the breast is required.

The structured light (SL) method is a well-validated computer vision technique that facilitates the information gathering of an object's volume [4]. This method lies in the use of a pattern, projected onto the object, to extract its depth information through the deformation it causes in the pattern. Traditionally, an image sensor acquires a 2D image of the scene with an overlapping pattern and, if the surface is nonplanar, the pattern will suffer variations, in comparison with the projected one, allowing to employ SL principles and algorithms to obtain the 3D surface shape [5]. More recently, with the expansion of Deep Learning (DL) methods in several areas and techniques, some implementations employed this methodology to extract the depth information about the captured scene. The combination of DL and SL methods is an interesting approach to achieving the 3D breast shape.

The traditional methods used for depth estimation are mainly based on the reflected echo captured after hitting an object with a controlled energy beam [6]. Although they offer an acceptable accuracy, they require a high cost, given that, besides an imaging device, its hardware involves a receiving and sensing device and following processing is still needed [7]. Alternatively, a strategy based on DL simply needs to acquire an image of the object, using a sensor, and then analyze this data in order to obtain the desired depth information. The obtained depth maps can then be used to deploy a 3D model of the object. However, DL-based strategies are highly dependent on their architecture [8] and rely on the coherence and variability of the dataset used to train them.

In this work, we aim to study the advantage of fusing DL and SL in breast shape estimation. Therefore, we initially created a toolchain to develop synthetic datasets, based on patterns. Here, a wide range of synthetic human models with different printed patterns were created. Then, multiple DL architectures were employed to solve the depth estimation problem. This strategy allows studying the influence of patterns on this technique and their design importance on depth estimation for 3D reconstruction problems.

The rest of the article is organized as follows: Section 2 describes the related works, Section 3 introduces the adopted methodology, Section 4 describes the experiments and provides the obtained results, Section 5 discusses the results of our search and Section 6 concludes on the significance of our work and its impact.

2 RELATED WORK

2.1 Deep learning on depth estimation

Depth estimation methods have recently reemerged, especially the ones that employ DL. Liu et al. [9], presented a way to extract the depth information from single images using deep convolutional neural networks (CNN). Furthermore, He et al. [10] introduced the concept of residuals into CNN and implemented the well-known ResNet, allowing models to have higher depth without suffering from an explosion of gradients, saturation, or a decrease in the accuracy. This contribution opened the way for Laina et al. [11], who created a network architecture built on top of the ResNet-50 to perform depth estimation from monocular images. Eventually, ResneXt [12] was created when a new dimension, called cardinality, was added. Besides the traditional dimensions of height and width, cardinality corresponds to the size of the set of transformations.

The DenseNet family of neural networks [13] is based on the existence of dense connections between layers, i.e each layer is connected to all the preceding and following layers, which increases the number of shorter connections between the input and output and enhances the feature maps propagation. In the end, it requires fewer parameters, memory, and computation to achieve state-of-art performances. Recent works using the DenseNet have been developed to perform depth estimation, namely DenseDepth [14], which was used as inspiration to deploy our DL model. It follows a standard encoder-decoder architecture, where the encoder corresponds to the DenseNet-169 network pre-trained on ImageNet. The loss function is a combination of point-wise L1 loss, L1 loss defined over the image gradient of the depth image, and structural similarity (SSIM) loss.

2.2 The structured light technique and its patterns

The SL technique has offered, over time, a suitable approach to resolve 3D reconstruction and depth estimation problems. It can be based on sequential projections, that acquire several shots of the scene. The patterns can be based on a binary or grey code and, in this case, the 3D coordinates are extracted based on triangulation principles. Besides that, the patterns can consist of sinusoidal projections and, in this case, the phase shift is the aspect taken into account [15]. Other variants were already implemented, such as continuous varying patterns, that use colored patterns and only require one shot of the scene. Stripe indexing methodology, which only requires one shot of the scene, can be based on color-coded stripes [16], segmented stripes, grayscale coded stripes, or De Bruijn sequence patterns. This method is adopted especially when facing occlusion problems and to mitigate the ambiguity that occurs when using patterns based on phase-shift. This stripes methodology is interesting when facing curvilinear surfaces [17], such as the breast, so is one of the types of pattern design to be analyzed in this work. Finally, instead of stripes, some works use patterns constituted by grids. Some of them are based on pseudo-random binary dots, mini-patterns as codewords, color-coded grids, or 2D color-coded dot arrays [18]. This approach will also be studied later, because of its ability to assign stripes in vertical and horizontal directions.

2.3 Deep learning on the structured light technique

In the past, all these patterns were analyzed using algorithms and mathematical formulas to extract the 3D coordinates. With the growth of DL, this technology was applied to SL problems, to help extract the 3D information present on the deformed patterns collected. A recent work [19] uses a generative adversarial network to fast and accurately estimate tissue oxygenation from single images containing a projected pattern on the surface. A depth map is obtained in [20], by combining SL and DL stereo matching techniques. In this work, they employed phase-shift patterns and the usual unwrapping step was improved by the DL network.

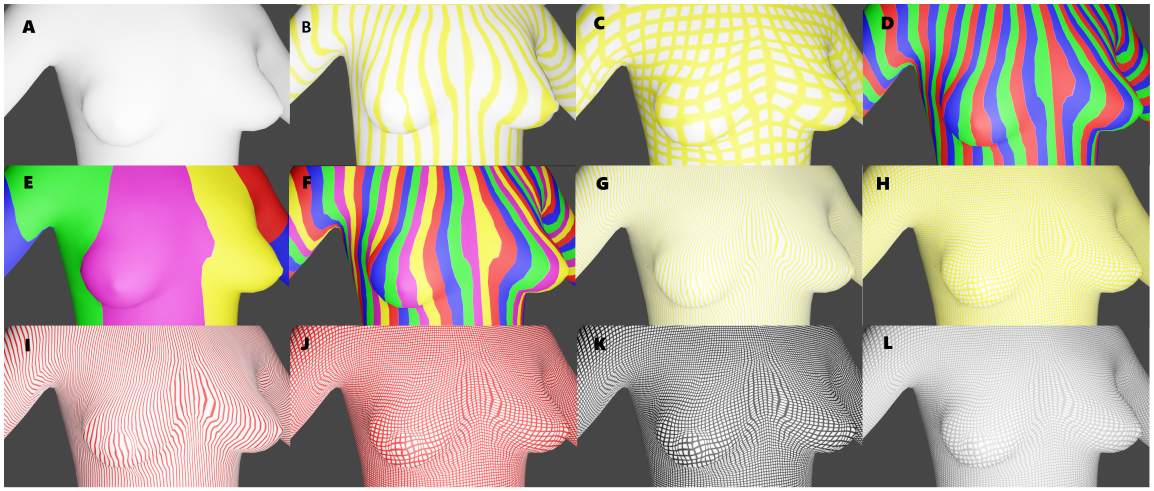

Figure 1: 12 Generated datasets where A has no pattern, B has a pattern made of thick yellow stripes, C has a pattern made of a thick yellow grid, D has a pattern made of a sequence of 3 colored stripes, E has a pattern made of a sequence of 5 colored stripes, F has a pattern made of a sequence of 5 colored thin stripes, G has a pattern made of thin yellow stripes, H has a pattern made of a thin yellow grid, I has a pattern made of thin red stripes, J has a pattern made of a thin red grid, K has a pattern made of a thin black grid and L has a pattern made of a thin grey grid.

3 METHODOLOGY

Given the need to create multiple datasets with different patterns, we developed a toolchain to generate human prototypes with distinct breast shapes. For each synthetic dataset, a different printed pattern was created. The datasets include images from different points of view that were fed to different DL architectures, implemented with the goal of performing breast depth estimation.

3.1 Dataset

To create each dataset, firstly, we generated human models using MakeHuman [21], an open-source tool for making 3D human models. This platform allows the creation of female prototypes with different breast types.

For it, we defined the human as a female and some breast parameters were randomly manipulated, such as size, volume, firmness, pointiness, the vertical and horizontal position of the breast, and the nipple's size and pointiness. However, the human's physiognomy is taken in account, so that the breasts do not appear disproportionate in comparison with reality. To simulate the projected pattern, characteristic of the SL technique, a pattern was printed on the prototype's skin by changing its skin texture. This approach is similar to the one used in the SL method because our pattern will be deformed by the breast volume as it would be if it was simply projected on the human breast surface.

After this, the human prototypes were imported to Blender [22] and multiple camera views were computed to rotate half a sphere around the body, maintaining a constant ray and focusing on the breast. In total, 20 surrounding horizontal positions were defined, and, for each of them, 10 positions were assigned in the vertical direction, corresponding to the shape of a half ellipse. The described mechanism permits the increase of the points of view and the variety of poses in the dataset since each camera captures images from its point of view. Simultaneously, depth maps were created. Finally, for each human prototype, we captured 200 RGB images and the respective 200 depth maps.

To compare the different patterns, 12 synthetic datasets were created. Each dataset is constituted of 35 human prototypes, which includes 7000 RGB images and 7000 respective depth maps. The first 25 humans are used to train the model, which corresponds to 5000 RGB images and, the rest of the humans, 2000 RGB images, are used for the validation stage. Since the prototype's breasts are randomly generated using the same seed, each human has the same breast shape as the respective human from the remaining datasets, which allows comparisons between datasets.

The first generated dataset does not include a pattern, since we intend to confirm if the presence of a pattern improves the results of depth estimation. The rest of the datasets have patterns that differ from each other. The patterns are all based on stripes, commonly used on the SL technique, and more simple than the ones based on phase-shifting or m-arrays. Regarding these types of designs, we studied the influence of distinct sequences of colors, and stripe thickness and also analyzed the behavior of a pattern made of vertical stripes when compared with a grid. An example of the generated datasets and corresponding patterns is shown in Figure 1. It is important to mention that the patterns that are composed of thin grids, were built with the same stripe's thickness.

3.2 Neural networks

The implemented network is based on the DenseDepth implementation. Our architecture is centered on the DenseNet-161, which, as stated before, presents dense connections between layers, forming Dense blocks, where each layer receives inputs from all preceding layers and passes its own to the following layers. The loss function

aggregates the point-wise L1 loss, L1 loss defined over the image gradient of the depth image, and SSIM loss.

Other architectures were used to understand which one is more suitable for this type of problem. They were chosen based on their performance in image classification on ImageNet [23]. Although this type of problem differs from ours, it is a way to compare some models used in the literature.

Firstly, the DenseNet-201 model was chosen because it offered, after DenseNet-161, the best results inside the DenseNet family. This model utilizes fewer parameters and its architecture is similar to the first implemented one. The only precaution was to modify the number of features utilized since it only requires 1920.

After this, the ResNet-152 was chosen to have one model from the ResNet architecture, since the ResNet family is commonly employed on depth estimation problems, especially ResNet-18 and ResNet-50. ResNet-152 was preferred because it offered better results for the image classification problem. The ResNet models learn residual functions concerning the layer inputs and are formed of residual blocks. This specific model combines 152 layers.

Finally, ResNeXt-101 was selected because it offered the best accuracy in the image classification problem. It is an adaptation of the ResNet network where a new dimension, cardinality, was incorporated, in addition to the dimensions of depth and width. ResNeXt architecture repeats a building block that aggregates a set of transformations with the same topology. When compared with other models, it demands a higher number of parameters.

After training, the models can receive an RGB image and output an image with a single channel, since it is a depth map and the distances are codified in the greyscale.

4 EXPERIMENTAL EVALUATION

We trained our networks, which integrates encoders pre-trained on ImageNet, during 20 epochs. In this procedure, we used a batch size of 2, a learning rate of 0.0001, and the Adam optimizer. The encoders receive an image of size 640x480 that is divided by 100, to normalize all the values to meters, since the maximum distance present in the ground truth depth maps is 1 meter. The ground truth is resized to 320x240. Some of the images used in the training set suffer transformations, such as horizontal flip and channel swap, with a probability of 50%.

After this, we resorted to the validation data to evaluate the model, by comparing the ground truth and the predicted depth maps. The predicted data was exported and, using Matlab, we calculated the mean error for each dataset. We also determined which images exhibit the minimum and maximum error for each dataset, in order to study which points of view are responsible for this error increase.

The DenseNet-201, ResNeXt-101, and Resnet-152 models were only trained using the dataset composed of a thin yellow grid pattern.

4.1 Evaluation metrics

To evaluate the model's performance we used the average relative error (rel), the root mean squared error (rms), the average error calculated using the logarithmic of base 10 (log10), and 3 different threshold accuracies, d_1, d_2 ,and d_3, whose thresholds correspond to 1.25, 1.25^2 and 1.25^3, as recurrently used in the literature. The later thresholds are calculated as presented below, where gt denotes the ground truth image and pred the predicted image:

$$\delta 1 = max\left(\frac{gt}{pred}, \frac{pred}{gt}\right)$$

As shown, these thresholds are dimensionless and the higher their value is, the more dissimilar the compared images can be. The result is presented in terms of the percentage of images from that dataset that respect the defined threshold. For further analyses, these thresholds can be modified.

4.2 Results

The obtained results for the 12 patterns, when using the DenseNet-161 encoder to perform the depth estimation, are shown in Table 1 and Figure 2 a).Table 2 and Figure 2 b) illustrate the results obtained for the dataset with the thin yellow grid pattern, the one with better results as shown in Figure 3 a), when estimated by the different architectures.

The colormaps created to analyze the images with the minimum and maximum error of the same dataset when predicted by the DenseNet-161 are presented in Figure 3. Two images with a medium error are also shown, to analyze different points of view. For further comparisons between the top 3 architectures, the thresholds d_1, d_2, and d_3 were decreased to 1.05, 1.05^2, and 1.05^3, respectively. The results are presented in Table 3.

5 DISCUSSION

This work aimed to study the impact of the incorporation of patterns, based on the SL technique, to improve the results of the depth estimation models. In fact, the dataset with no pattern presents the worst results, when compared with the others, as shown in Table 1, demonstrating the added-value of the described pattern-based technique. It is also visible that the dataset that presents better results, in terms of mean error, is the dataset composed of a thinner yellow grid pattern. The mean error present in this dataset corresponds to 4.9 millimeters in the virtual scene we created in Blender, which is lower than the initial value of 10 millimeters for the dataset with no pattern.

These results indicate that a pattern based on a grid, with thin stripes, facilitates the depth estimation of a surface. This could be explained by the fact that the presence of horizontal and vertical stripes, increases the volume of information that is interpreted by the DL model. When compared with the dataset with a pattern made of thin yellow stripes, that was designed with the same stripe's thickness, according to Table 1, we can see that the presence of vertical stripes exclusively does not offer the same good results. Nevertheless, other datasets are worth mentioning, such as the dataset composed of a red grid pattern, which showed great results too. It is possible to understand, by analyzing Figure 2 a), that this dataset also presented low outliers, which indicates good results.

The only thing distinguishing the aforementioned dataset from the others made of thin yellow, black, or grey grids is their color. Regarding the influence of colors, it is visible that the datasets made of a sequence of colored stripes offered better results than the yellow or red stripes superimposed on the white texture of the prototype, which might indicate that the DL model finds it easier

Table 1: The obtained results for the 12 patterns, when using the DenseNet-161 model. The datasets are named after the legend used in Figure 1.

Datasets	d 1	d 2	d 3	rel	Rms	log_10	Mean error	Min error	Max error
A	0.9046	0.9353	0.9529	0.0779	0.1004	0.0433	2.7164	0.419	16.4
B	0.9897	0.9967	0.9993	0.0216	0.0174	0.0084	1.5131	0.342	10.1
C	0.992	0.9982	0.9994	0.0235	0.0192	0.0096	2.7797	0.744	8.76
D	0.9895	0.9988	**0.9999**	0.0155	0.0152	0.0064	1.49	0.32	10.1
E	0.9868	0.9947	0.9994	0.0217	0.0161	0.0087	1.4349	0.270	15.6
F	0.9927	0.9995	**0.9999**	0.0183	0.0166	0.0076	1.4099	0.439	7.56
G	0.992	0.9996	**0.9999**	0.0173	0.0178	0.0074	1.6472	0.342	10.1
H	0.9905	0.9965	**0.9999**	0.0168	**0.015**	0.0068	**1.2717**	0.375	13.4
I	**0.9942**	**0.9997**	**0.9999**	**0.0136**	0.0154	**0.0057**	1.8356	0.37	6.6
J	0.9925	0.9996	**0.9999**	0.0241	0.0189	0.0101	1.3537	0.351	8.27
K	0.9894	0.9968	**0.9999**	0.0193	0.0162	0.0079	1.6936	0.526	11.8
L	0.9939	0.9997	**0.9999**	0.0147	0.0158	0.0062	1.9082	0.386	8.07

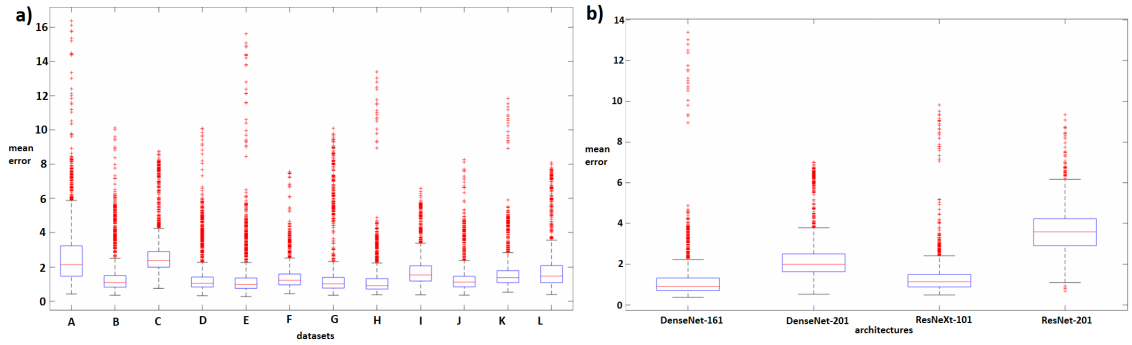

Figure 2: Boxplot (a) showing the mean error for the 12 patterns, measured in pixel's intensity, the datasets are named after the legend used in Figure 1, and (b) showing the mean error for the 4 architectures, measured in pixel's intensity.

Table 2: The obtained results for the dataset, composed of a thin yellow grid pattern, when using the 4 architectures.

Architectures	d 1	d 2	d 3	rel	rms	log_10	Mean error	Min error	Max error
DenseNet-161	0.9905	0.9965	**0.9999**	0.0168	0.015	0.0068	**1.2717**	0.375	13.4
DenseNet-201	**0.9943**	**0.9997**	**0.9999**	**0.013**	**0.0149**	**0.0055**	2.308	0.523	7
ReNeXt-101	0.9902	0.9975	0.9998	0.0147	0.0175	0.0061	1.409	0.488	9.83
ResNet-152	0.9908	0.9989	**0.9999**	0.025	0.02317	0.0107	3.6398	0.686	9.34

Table 3: The results obtained for each model when the thresholds d_1, d_2, and d_3 were decreased to 1.05, 1.052, and 1.053, respectively.

Architectures	d 1	d 2	d 3
DenseNet-161	0.963	0.9779	0.9852
DenseNet-201	0.9607	**0.9853**	**0.9893**
ReNeXt-101	**0.9653**	0.9801	0.9854

to understand volumes where different colors create boundaries over the object.

Besides that, we can conclude that the stripe's thickness also influences the results. Regarding the dataset made of a thick yellow grid pattern and the one constituted by a thin yellow grid pattern, it is possible to comprehend, by analyzing Table 1, that the thinner one offered better results, The same happened with the datasets made of a pattern with a sequence of 5 colors, where the one with thinner stripes gets better results. However, there is an exception to this coincidence, the dataset with thicker yellow stripes has better results than the dataset constituted by thinner yellow stripes.

Concerning the results of the different architectures presented in Table 2, it is possible to determine that the DenseNet-161 is the most suitable for this task because the mean error is lower than the others. However, from this model's outliers illustrated in Figure 2

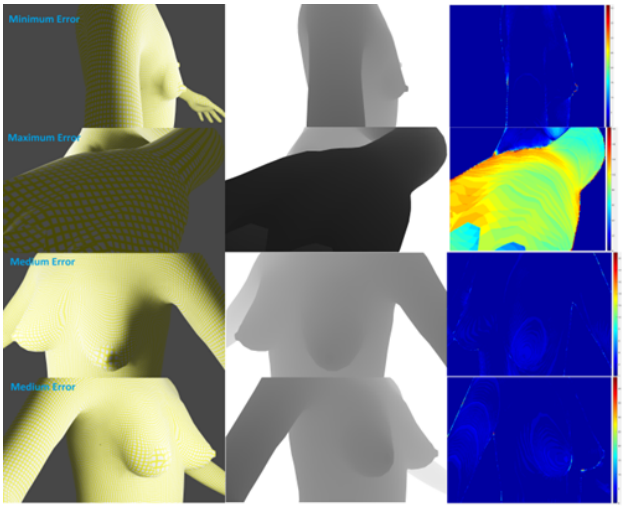

Figure 3: Colormaps showing the images with the minimum, maximum and medium error, for the dataset with the thin yellow grid pattern, when predicted by the DenseNet-161.

b), we can verify that when it fails, this model originates predictions with higher error. The dataset toolchain was conceived to include some occlusions, such as the prototype's arms or hands, similar to the problems that a robot-guided intervention could face in real life. The point of view that offers more complication is precisely one where the human's hand is in front of the camera, as shown in Figure 3, which demonstrates that the model might not be so versatile and able to adapt to occlusions like this, especially the ones that are rare in the training data. The colormap of the depth map with the minimum error is also illustrated and shows that the model can deal with the entire body easily. Relatively to other points of view, where the error is medium, we can observe that some of the errors are prevalent in the borders of the human body, the nipple zone, the abdomen, and the arms. Some zones of the breast show reduced error and it is visible that, depending on the point of view, the DL model has more difficulty to estimate the depth of the furthest or nearest breast.

When we decreased the d_1, d_2, and d_3 thresholds, our results, shown in Table 3, demonstrated which models can predict more images with low error. In this case, the ResNeXt-101 and DenseNet-161 showed the best results, contrary to the DenseNet-201, which had the best results for a higher value of the three thresholds. It means that the DenseNet-201 model has difficulty to predict depth maps with low error, but is excellent to make predictions with medium error and maintain this consistency. The DenseNet-161 model is able to predict depth maps with low error, as suggested by its percentage achieved for the d_1 threshold, but has some problems in some points of view, as stated before, which decreases the percentage obtained in the d_2 and d_3 metrics. The ResNeXt-101 does not show the best result for the mean error but proved to be a good model for depth estimation, being the one capable of performing more predictions with lower error.

In the future, there are some improvements that we intent to incorporate into our strategy. For example, to reduce the problems faced from some points of view, as shown in Figure 3, a solution would be adding a Long short-term memory segment to our network to introduce some temporal information and integrate the camera's trajectory to improve the spatial information. In order to promote a better application of our dataset and model to real problems, we can combine a real dataset and introduce other domain adaptation techniques.

6 CONCLUSION

In this work, we proved that it is possible and beneficial to fuse DL and SL techniques, in order to improve the results of the breast shape depth estimation. Taking advantage of the toolchain that we developed to produce synthetic datasets, we demonstrated that, not only do the employed patterns matter but also their color and stripe's thickness. We determined that the dataset composed of a thin yellow grid pattern is the most suitable to improve the depth information of a surface. Furthermore, we evaluated the capacity of four different models to perform this task and determined that the one based on the DenseNet-161 achieved better results. The architecture based on ResNeXt-101 and DenseNet-201 proved to be promising alternatives for depth estimation.

ACKNOWLEDGMENTS

The authors acknowledge the funding of the projects "NORTE-01-0145-FEDER000045" and "NORTE-01-0145-FEDER-000059", supported by Northern Portugal Regional Operational Programme (NORTE 2020), under the Portugal 2020 Partnership Agreement, through the European Regional Development Fund (FEDER). It was also funded by national funds, through the Fundação para a Ciência e a Tecnologia (FCT) and FCT/MCTES in the scope of the projects UIDB/05549/2020,UIDP/05549/2020 and LASI-LA/P/0104/2020. The authors also acknowledge FCT, Portugal and the European Social Found, European Union, for funding support through the "Programa Operacional Capital Humano" (POCH) in the scope of the PhD grants SFRH/BD/136721/2018 (B. Oliveira) and SFRH/BD/136670/2018 (H. Torres).

REFERENCES

[1] E. Dhamija, R. Singh, S. Mishra, and S. Hari, "Image-Guided Breast Interventions: Biopsy and beyond," Indian Journal of Radiology and Imaging, vol. 31, no. 2, pp. 391–399, Apr. 2021, doi: 10.1055/S-0041-1734223/ID/JR_70.

[2] K. L. Maughan, M. A. Lutterbie, and P. S. Ham, "Treatment of Breast Cancer," Am Fam Physician, vol. 81, no. 11, pp. 1339–1346, 2010, Accessed: May 20, 2022. [Online]. Available: www.aafp.org/afpAmericanFamilyPhysician1339

[3] Z. Xinran, D. Haiyan, L. Mingyue, and Z. Yongde, "Breast intervention surgery robot under image navigation: A review:," https://doi.org/10.1177/16878140211028113, vol. 13, no. 6, pp. 1–17, Jun. 2021, doi: 10.1177/16878140211028113.

[4] R. J. Valkenburg and A. M. McIvor, "Accurate 3D measurement using a structured light system," Image and Vision Computing, vol. 16, no. 2, pp. 99–110, Feb. 1998, doi: 10.1016/S0262-8856(97)00053-X.

[5] J. Geng, "Structured-light 3D surface imaging: a tutorial," Advances in Optics and Photonics, Vol. 3, Issue 2, pp. 128-160, vol. 3, no. 2, pp. 128–160, Jun. 2011, doi: 10.1364/AOP.3.000128.

[6] N. Jamwal, N. Jindal, and K. Singh, "A survey on depth map estimation strategies," pp. 14 (5 .)-14 (5 .), Mar. 2018, doi: 10.1049/CP.2016.1453.

[7] Y. Liu, J. Jiang, J. Sun, L. Bai, and Q. Wang, "A survey of depth estimation based on computer vision," Proceedings - 2020 IEEE 5th International Conference on Data Science in Cyberspace, DSC 2020, pp. 135–141, Jul. 2020, doi: 10.1109/DSC50466.2020.00028.

[8] A. Shrestha and A. Mahmood, "Review of Deep Learning Algorithms and Architectures", doi: 10.1109/ACCESS.2019.2912200.

[9] F. Liu, C. Shen, and G. Lin, "Deep Convolutional Neural Fields for Depth Estimation from a Single Image," Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 07-12-June-2015, pp. 5162–5170, Nov. 2014, doi: 10.1109/CVPR.2015.7299152.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2016-December, pp. 770–778, Dec. 2015, doi: 10.1109/CVPR.2016.90.

[11] I. Laina, C. Rupprecht, V. Belagiannis, F. Tombari, and N. Navab, "Deeper Depth Prediction with Fully Convolutional Residual Networks," Proceedings - 2016 4th International Conference on 3D Vision, 3DV 2016, pp. 239–248, Jun. 2016, doi: 10.1109/3DV.2016.32.

[12] S. Xie, R. Girshick, P. Dollár, Z. Tu, K. He, and U. San Diego, "Aggregated Residual Transformations for Deep Neural Networks", Accessed: Apr. 28, 2022. [Online]. Available: https://github.com/facebookresearch/ResNeXt

[13] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, vol. 2017-January, pp. 2261–2269, Aug. 2016, doi: 10.48550/arxiv.1608.06993.

[14] I. A. Kaust and P. Wonka, "High Quality Monocular Depth Estimation via Transfer Learning," Dec. 2018, Accessed: Jan. 31, 2022. [Online]. Available: https://arxiv.org/abs/1812.11941v2

[15] C. Sui, K. He, C. Lyu, Z. Wang, and Y. H. Liu, "3D surface reconstruction using a two-step stereo matching method assisted with five projected patterns," Proceedings - IEEE International Conference on Robotics and Automation, vol. 2019-May, pp. 6080–6086, May 2019, doi: 10.1109/ICRA.2019.8794063.

[16] C. Rocchini, P. Cignoni, C. Montani, P. Pingi, and R. Scopigno, "A low cost 3D scanner based on structured light," Computer Graphics Forum, vol. 20, no. 3, pp. 299–308, Sep. 2001, doi: 10.1111/1467-8659.00522.

[17] M. Alzuhiri, K. Farrag, E. Lever, and Y. Deng, "An Electronically Stabilized Multi-Color Multi-Ring Structured Light Sensor for Gas Pipelines Internal Surface Inspection," IEEE Sensors Journal, vol. 21, no. 17, pp. 19416–19426, Sep. 2021, doi: 10.1109/JSEN.2021.3086415.

[18] L. Song, S. Tang, and Z. Song, "A robust structured light pattern decoding method for single-shot 3D reconstruction," 2017 IEEE International Conference on Real-Time Computing and Robotics, RCAR 2017, vol. 2017-July, pp. 668–672, Mar. 2018, doi: 10.1109/RCAR.2017.8311940.

[19] M. T. Chen and N. J. Durr, "Rapid tissue oxygenation mapping from snapshot structured-light images with adversarial deep learning," https://doi.org/10.1117/1.JBO.25.11.112907, vol. 25, no. 11, p. 112907, Nov. 2020, doi: 10.1117/1.JBO.25.11.112907.

[20] F. Li, Q. Li, T. Zhang, Y. Niu, and G. Shi, "Depth acquisition with the combination of structured light and deep learning stereo matching," Signal Processing: Image Communication, vol. 75, pp. 111–117, Jul. 2019, doi: 10.1016/J.IMAGE.2019.04.001.

[21] "www.makehumancommunity.org." http://www.makehumancommunity.org/ (accessed May 02, 2022).

[22] "blender.org - Home of the Blender project - Free and Open 3D Creation Software." https://www.blender.org/ (accessed May 02, 2022).

[23] "DenseNet | Papers With Code." https://paperswithcode.com/model/densenet?variant=densenet-161 (accessed May 02, 2022).