Semantic Aware Data Augmentation for Cell Nuclei Microscopical Images with Artificial Neural Networks

Alireza Naghizadeh
Rutgers University
Newark, NJ
ar.naghizadeh@rutgers.edu

Hongye Xu
Rutgers University
Newark, NJ
hx123@njms.rutgers.edu

Mohab Mohamed
Rutgers University
Newark, NJ
mm2233@scarletmail.rutgers.edu

Dimitris N. Metaxas
Rutgers University
New Brunswick, NJ
dnm@cs.rutgers.edu

Dongfang Liu
Rutgers University
Newark, NJ
dongfang.liu@rutgers.edu

Abstract

In this section, we provide details of transformations of the image operations on datasets. We then further explain the datasets and the architectures found in related experiments. Finally, we look at the effect our method has on accuracy within segmentation models and discuss the results with provided samples of the generated images.

1. Details of the Effects of Image Operations on Datasets

The overall number of image operations to create the search space of policies is comprised of twenty operations which include 'FlipLR', 'FlipUD', 'AutoContrast', 'Equalize', 'Invert', 'Sharpness', 'Blur', 'Smooth', 'Rotate', 'Posterize', 'CropBilinear', 'Solarize', 'Color', 'Contrast', 'Brightness', 'ShearX', 'ShearY', 'TranslateX', 'TranslateY', 'Cutout'. The range of magnitude is ten for 'Rotate', 'Posterize', 'CropBilinear', 'Solarize', 'Color', 'Contrast', 'Brightness', 'ShearX', 'ShearY', 'TranslateX', 'TranslateY', 'Cutout'. Note there are some transformations, such as 'FlipLR', 'FlipUD', 'AutoContrast', 'Equalize', 'Invert', 'Sharpness', 'Blur', 'Smooth', that do not utilize a magnitude value. Figure 5 and Figure 6 show the outcomes of applying image operations on CAR-T and Kaggle datasets.

2. Datasets

Kaggle Dataset It was first published by Booz Allen Hamilton on Kaggle.com for the 2018 Data Science Bowl. This dataset contains a large number of segmented nuclei images (256x256). The images were acquired under a variety of conditions and vary in cell type, magnification, and imaging modality (brightfield vs. fluorescence). In order to ensure the consistency of the cell image background, we have excluded a very small number of images with a special background color. In the following experiments, we used 320 images for training, 134 images for validation and 134 images for testing.

Car-T Dataset This dataset was collected by Rutgers Cancer Institute of New Jersey. It contains 156 confocal images (1024 x 1024) obtained by microscope system A1R HD25 (Nikon, Japan), which provides a 25 mm field of view. In the following experiments, we used 93 images for training, 32 images for validation, and 31 images for testing.

3. Architectures

3.1. Wasserstein Auto-Encoder

The Wasserstein Auto-Encoder(WAE) minimizes a penalized form of the Wasserstein distance between the model distribution and the target distribution, which leads to a different regularizer than the one used within the Variational Auto-Encoder (VAE). This regularizer encourages the encoded training distribution to match the prior [6]. We use WAE to search the different policy combinations that yield the least reconstruction loss.

3.2. KG Instance Segmentation

KG Instance Segmentation is a box-based cell instance segmentation method. It first detects five pre-defined points of a cell via keypoints detection and then groups these points according to a keypoint graph that subsequently extracts the bounding box for each cell. Finally, cell segmentation is performed on feature maps within the bounding boxes [7]. This method is particularly effective for small sample segmentation tasks, and its performance exceeds DCAN, Mask R-CNN, and other methods.

4. Additional Experiment

For this experiment, we use data generated by our proposed method to improve cell segmentation accuracy and the robustness of model training results.

4.1. Setup

All of our experiments in supplementary are conducted on RTX Titan NVIDIA graphic cards. As for the embedded deep neural networks for our method, we used Pytorch [4]. For WAE [6], we used the implementation in [5] and for the KG Instance Segmentation [7] we used the implementation by the author.

4.2. Data Preparation

Mixed Car-T Dataset We mixed the images generated by our proposed method with the natural CAR-T images. The training set contains 240 generated images combined with 93 natural images. The validation set contains 80 generated images and 32 natural images.

Original Car-T Dataset This dataset has been introduced in the previous section and is used as a reference for the experiments.

4.3. Evaluation

In order to verify the effect of our method on real data, the training results of the two datasets (Mixed Car-T Dataset and Original Car-T Dataset) are both evaluated on the real CAR-T test set.

Evaluation Metrics We use the average precision (AP) at box-level IOU (intersection over union) [1] at thresholds between 0.5 to 0.8 to evaluate the bounding box detection performances. We use the AP at mask-level IOU [2, 3] at thresholds between 0.5 to 0.8 to evaluate the instance segmentation performances. We also report the mean mask-level IOU [8] between the predicted segmentation masks and the ground truth masks at thresholds between 0.5 to 0.8.

4.4. Results and Discussion

As shown in Table 1 and Table 2, for the BBox detection performances, the training results of the mixed dataset is at least 0.65 $\%$ and at most 4.02$\%$ better than the results of the real dataset on the AP. For semantic segmentation, the results of the mixed dataset is at least 0.39 $\%$ and up to 2.86$\%$ higher than results of the real data on AP. It also shows improvement of at least 0.65$\%$ and up to 1.18$\%$ when compared to the reference dataset at IOUs.

The results show that the method could easily improve the accuracy of object detection and semantic segmentation of cells. The reason is that augmentation generates a lot of new data that benefits the generalization of the training process, which reduces overfitting and improves the robustness of the model. In these experiments, we adopted a conservative approach and only generated 240 artificial images. The effects of the size of datasets need further investigation. More importantly, adopting different strategies to select augmentation policies and their effects on the training process also needs further investigation. We will conduct in-depth research in the future to explore the impact of data augmentation within model training.

5. Generated image samples

In the end, we generate additional image samples at higher resolutions.

References

[1] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The pascal visual object classes (voc) challenge. *IJCV*, 88(2):303–338, june 2010.

[2] Kaiming He, Georgia Gkioxari, Piotr Dollr, and Ross Girshick. Mask r-cnn, 2018.

[3] Yi Li, Haozhi Qi, Jifeng Dai, Xiangyang Ji, and Yichen Wei. Fully convolutional instance-aware semantic segmentation, 2017.

[4] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019.

[5] A.K Subramanian. Pytorch-vae. `https://github.com/AntixK/PyTorch-VAE`, 2020.

[6] Ilya Tolstikhin, Olivier Bousquet, Sylvain Gelly, and Bernhard Schoelkopf. Wasserstein auto-encoders, 2019.

[7] Jingru Yi, Pengxiang Wu, Qiaoying Huang, Hui Qu, Bo Liu, Daniel J. Hoeppner, and Dimitris N. Metaxas. Multi-scale cell instance segmentation with keypoint graph based bounding boxes, 2019.

[8] Jingru Yi, Pengxiang Wu, Menglin Jiang, Qiaoying Huang, Daniel J. Hoeppner, and Dimitris N. Metaxas. Attentive neural cell instance segmentation. *Medical Image Analysis*, 55:228–240, 2019.

BBox Evaluation	AP@0.5	AP@0.6	AP@0.7	AP@0.8
Mixed Car-T Dataset	77.35	71.13	63.97	45.35
Original Car-T Dataset	76.35	70.48	59.95	42.79

Table 1: The evaluation accuracy (%) for object detection/bounding box generation.

Sementic Evaluation	AP@0.5	IOU@0.5	AP@0.6	IOU@0.6	AP@0.7	IOU@0.7	AP@0.8	IOU@0.8
Mixed Car-T Dataset	78.26	81.60	72.02	83.45	63.04	85.53	46.28	88.28
Original Car-T Dataset	77.64	80.44	71.63	82.27	60.46	84.73	43.42	87.63

Table 2: The evaluation accuracy (%) for semantic segmentation.

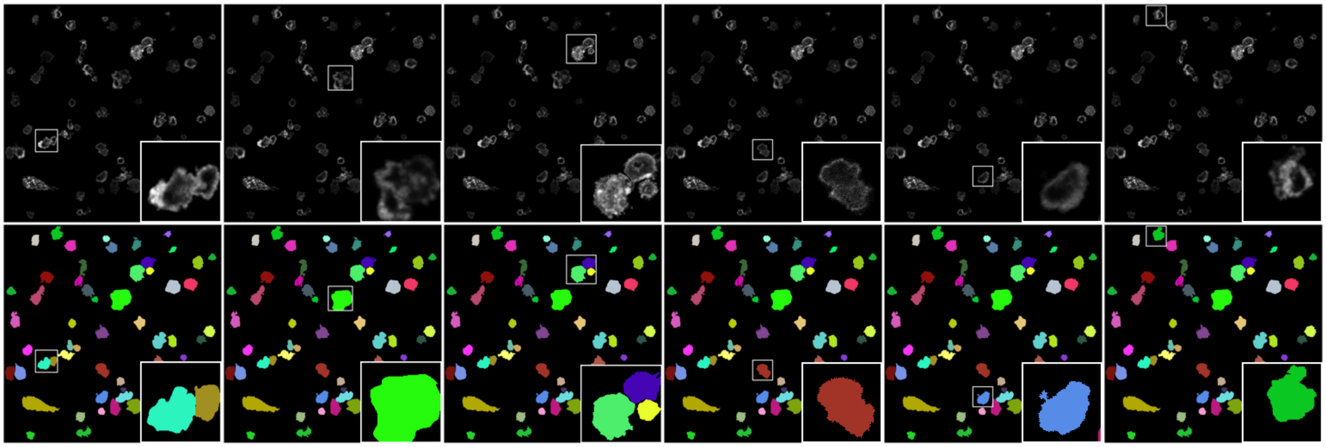

Figure 1: A sample of an artificially generated CAR-T dataset (first row) with its respective semantically segmented image (second row). In four different locations, the zoom-in functionality is used to provide better details of the generated cells and their respective masks.

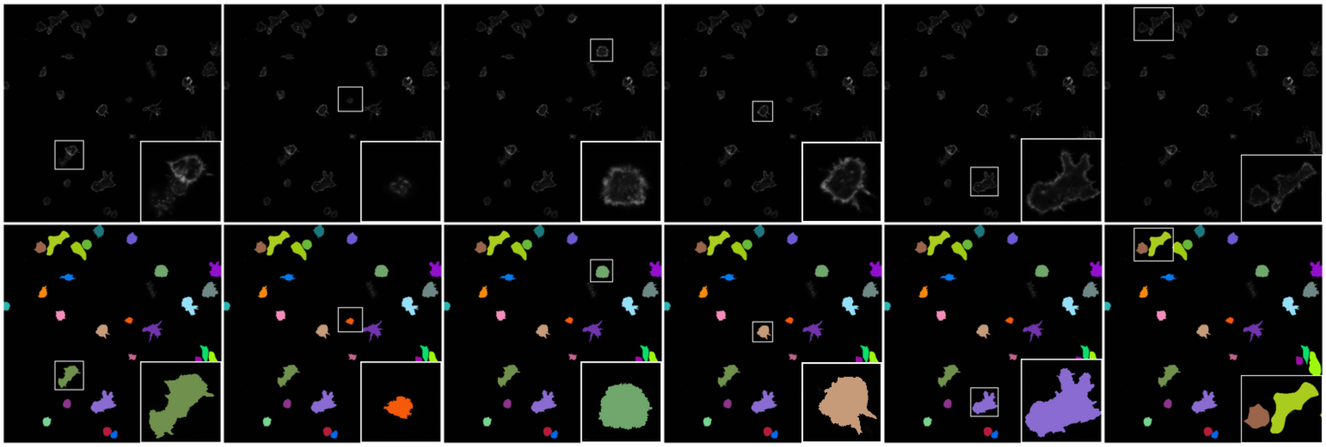

Figure 2: A real sample from the CAR-T dataset (first row) with its respective semantically segmented image (second row). In four different locations, the zoom-in functionality is used to provide better details of the real cells and their respective masks.

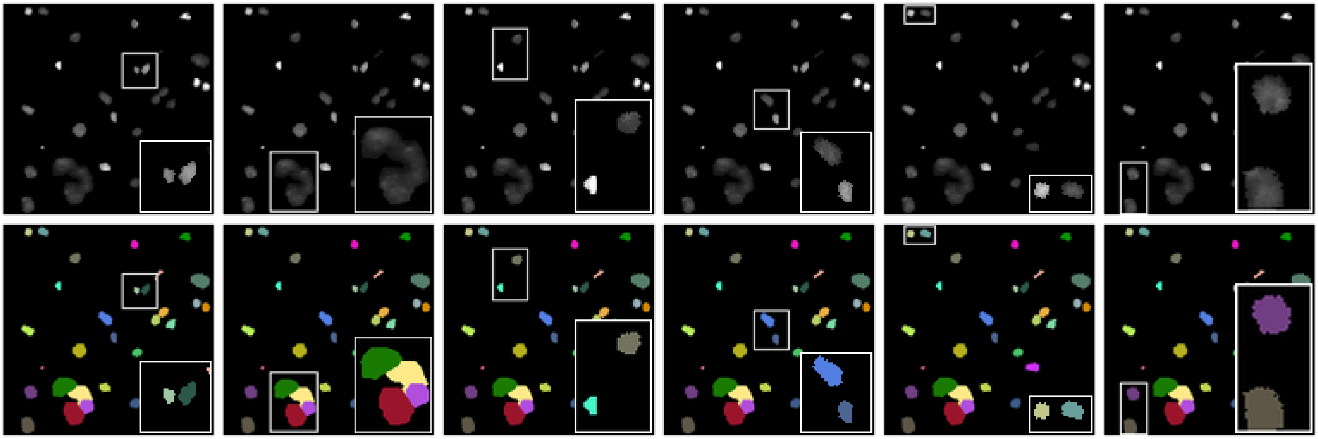

Figure 3: A sample of an artificially generated Kaggle dataset (first row) with its respective semantically segmented image (second row). In four different locations, the zoom-in functionality is used to provide better details of the generated cells and their respective masks.

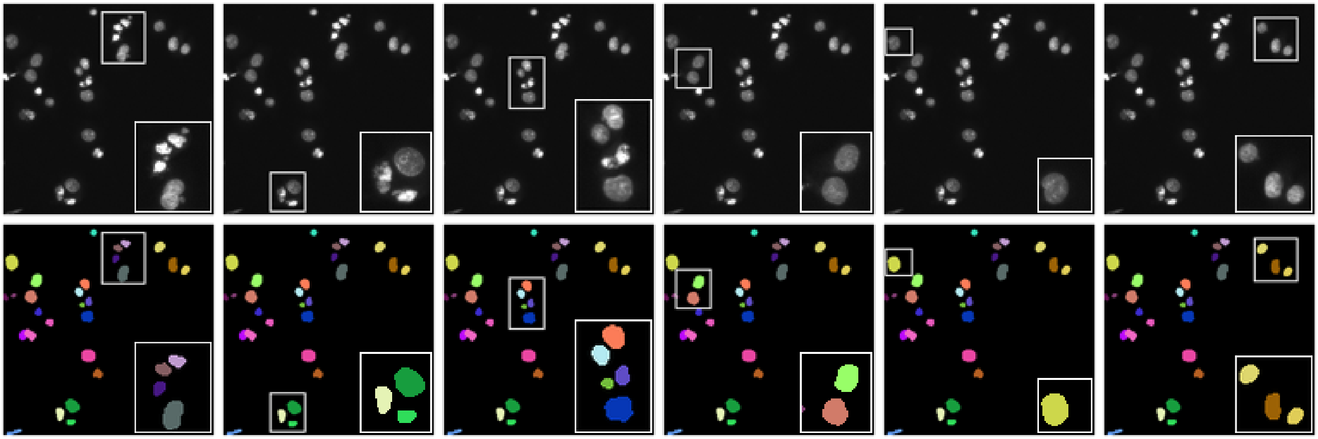

Figure 4: A real sample from the Kaggle dataset (first row) with its respective semantically segmented image (second row). In four different locations, the zoom-in functionality is used to provide better details of the real cells and their respective masks.

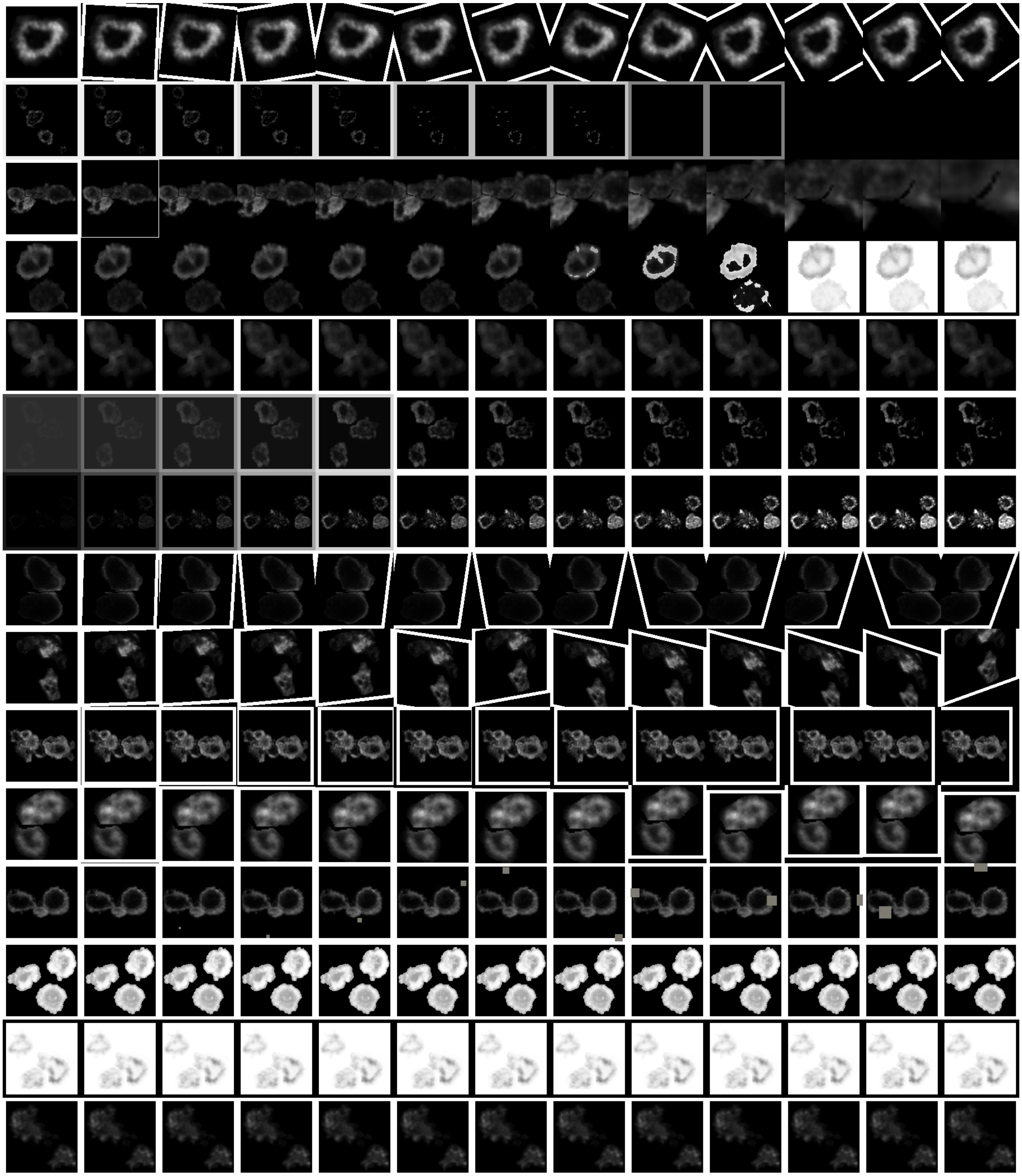

Figure 5: A sample using different forms of augmentation techniques on the cells. Each row represents a new augmentation technique applied on a multi-cell within the CAR-T dataset. The augmentations used in the order of rows are: 'Rotate','Posterize','CropBilinear','Solarize','Color','Contrast','Brightness','ShearX','ShearY', 'TranslateX','TranslateY','Cutout','Equalize','Invert','AutoContrast'.

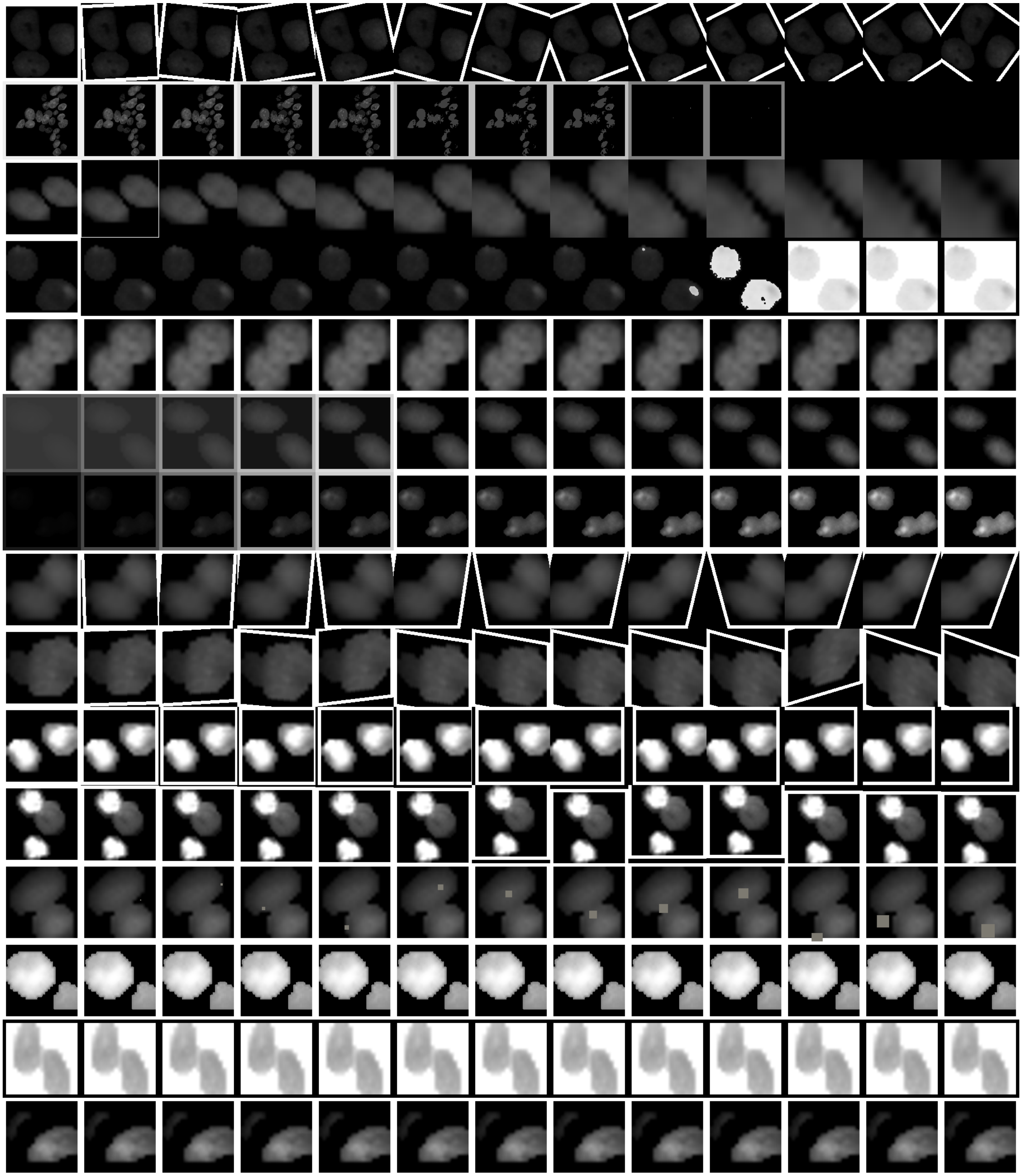

Figure 6: A sample using different forms of augmentation techniques on the cells. Each row represents a new augmentation technique applied on a multi-cell within the Kaggle dataset. The augmentations used in the order of rows are: 'Rotate','Posterize','CropBilinear','Solarize','Color','Contrast','Brightness','ShearX','ShearY', 'TranslateX','TranslateY','Cutout','Equalize','Invert','AutoContrast'.

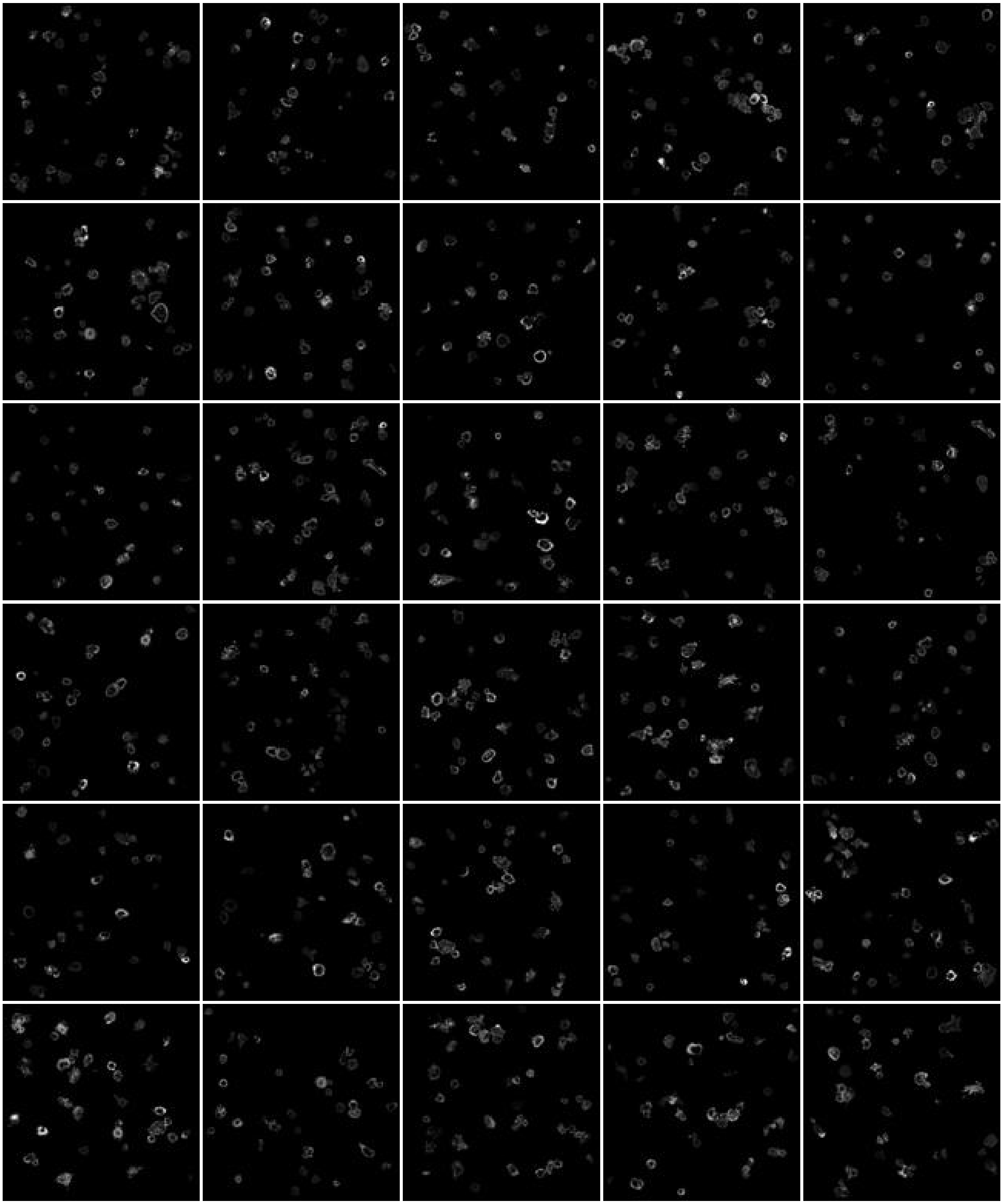

Figure 7: A sample of the final images created from the proposed method (128 × 128 for each sample) using the CAR-T dataset.

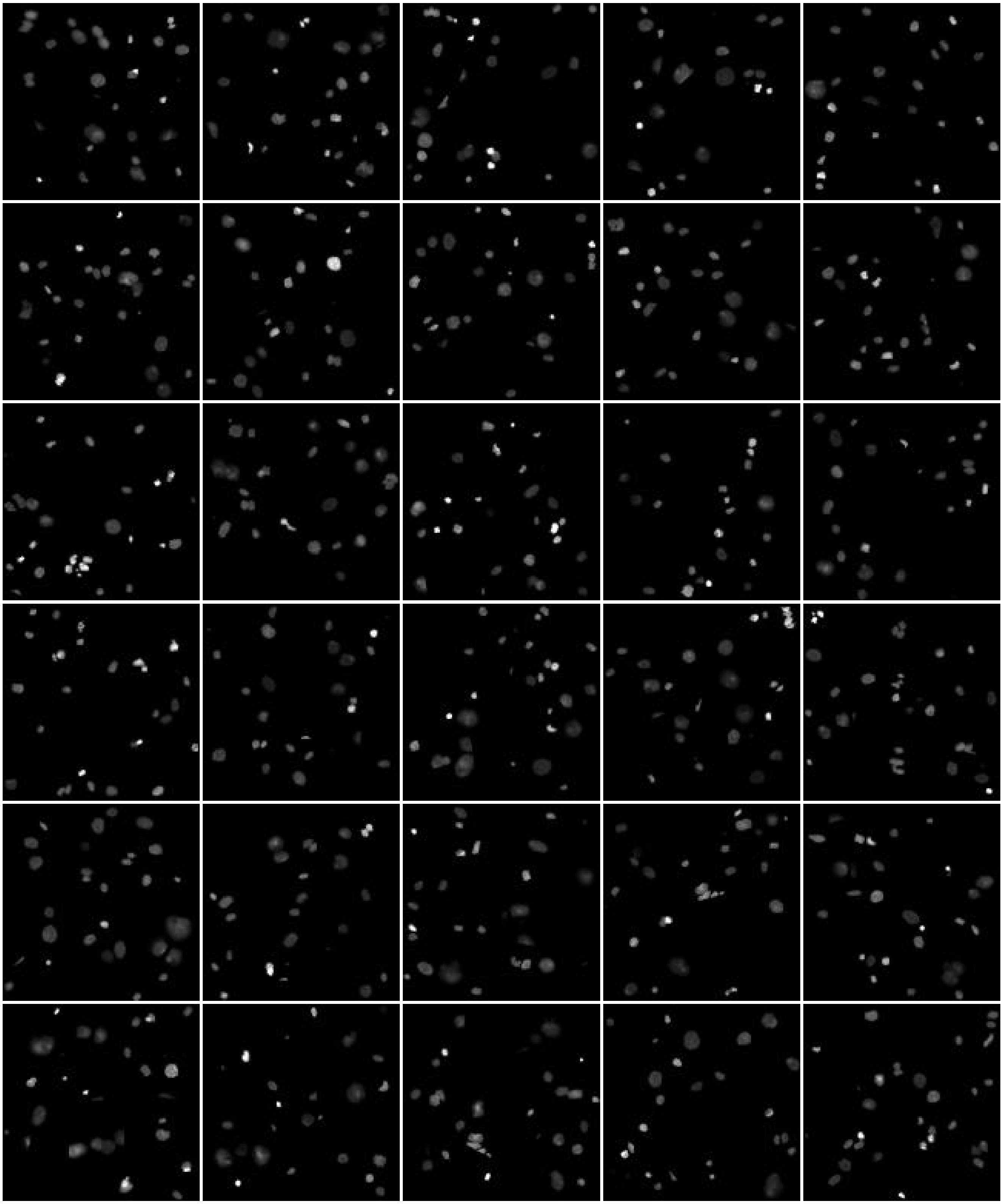

Figure 8: A sample of the final images created from the proposed method (128 × 128 for each sample) using the Kaggle dataset.

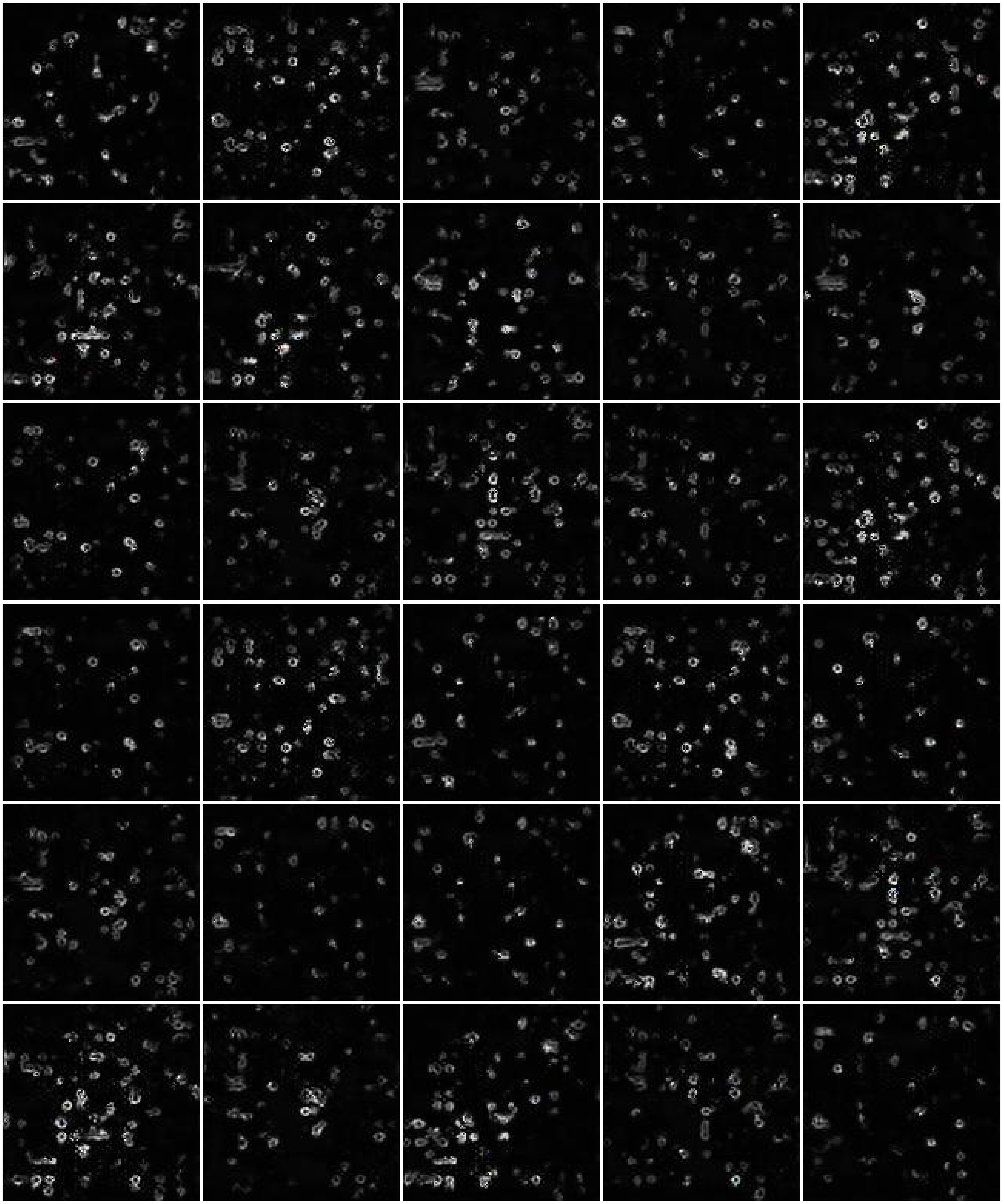

Figure 9: A sample of the final images created from the BigGAN-Diff method (128 × 128 for each sample) using the CAR-T dataset.

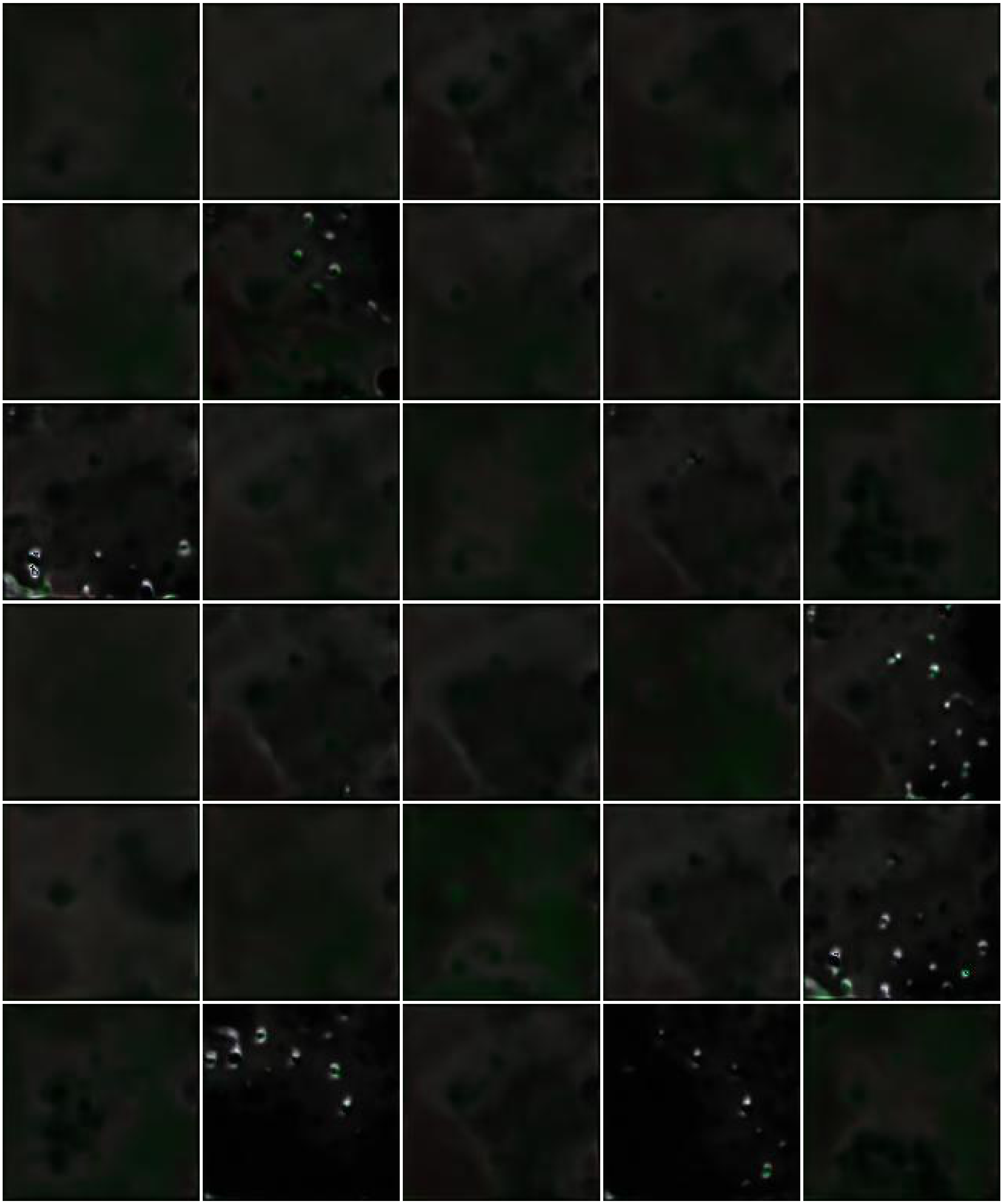

Figure 10: A sample of the final images created from the BigGAN-Diff method (128 × 128 for each sample) using the Kaggle dataset.

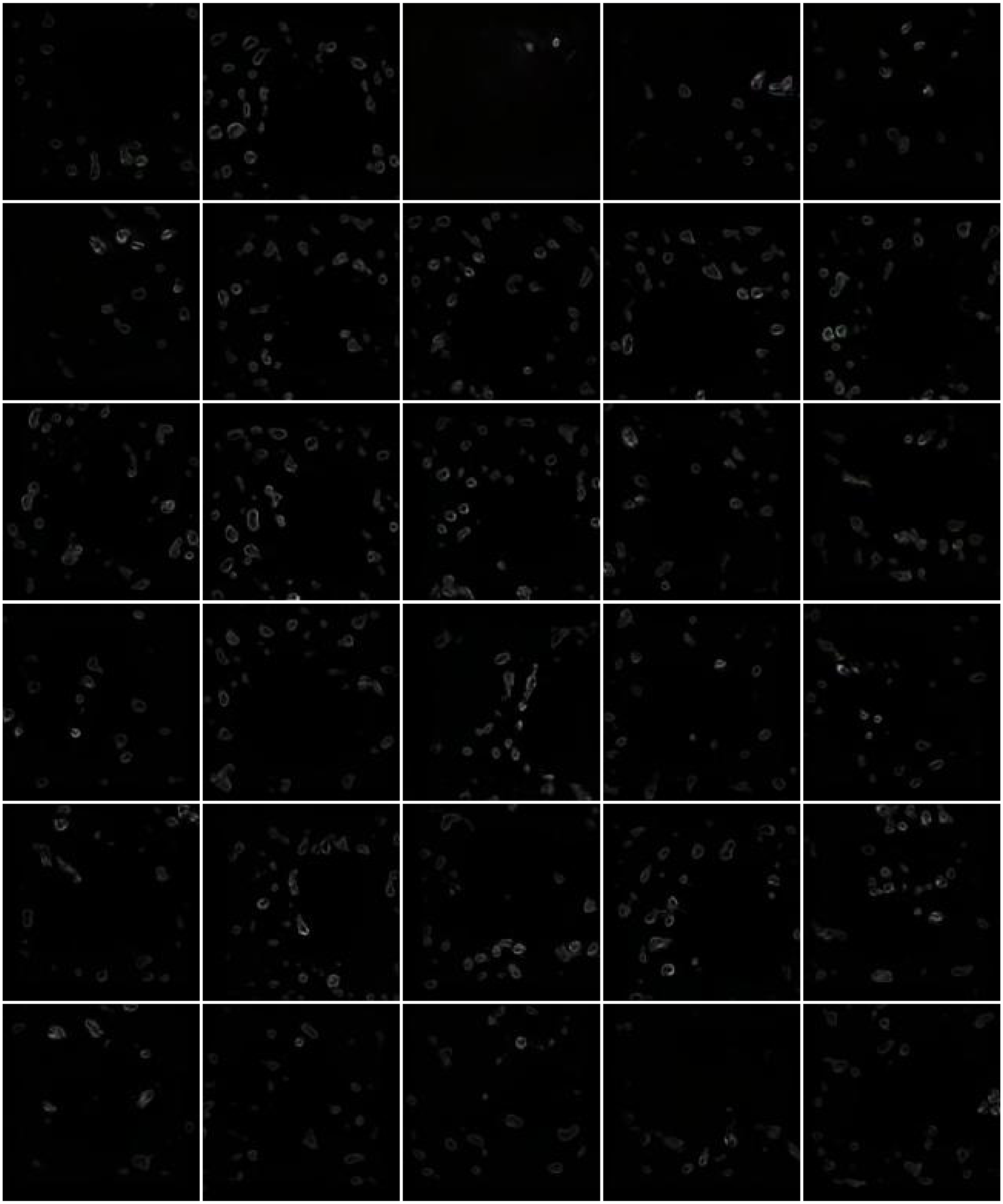

Figure 11: A sample of the final images created from the StyleGAN2-Diff method (128 × 128 for each sample) using the CAR-T dataset.

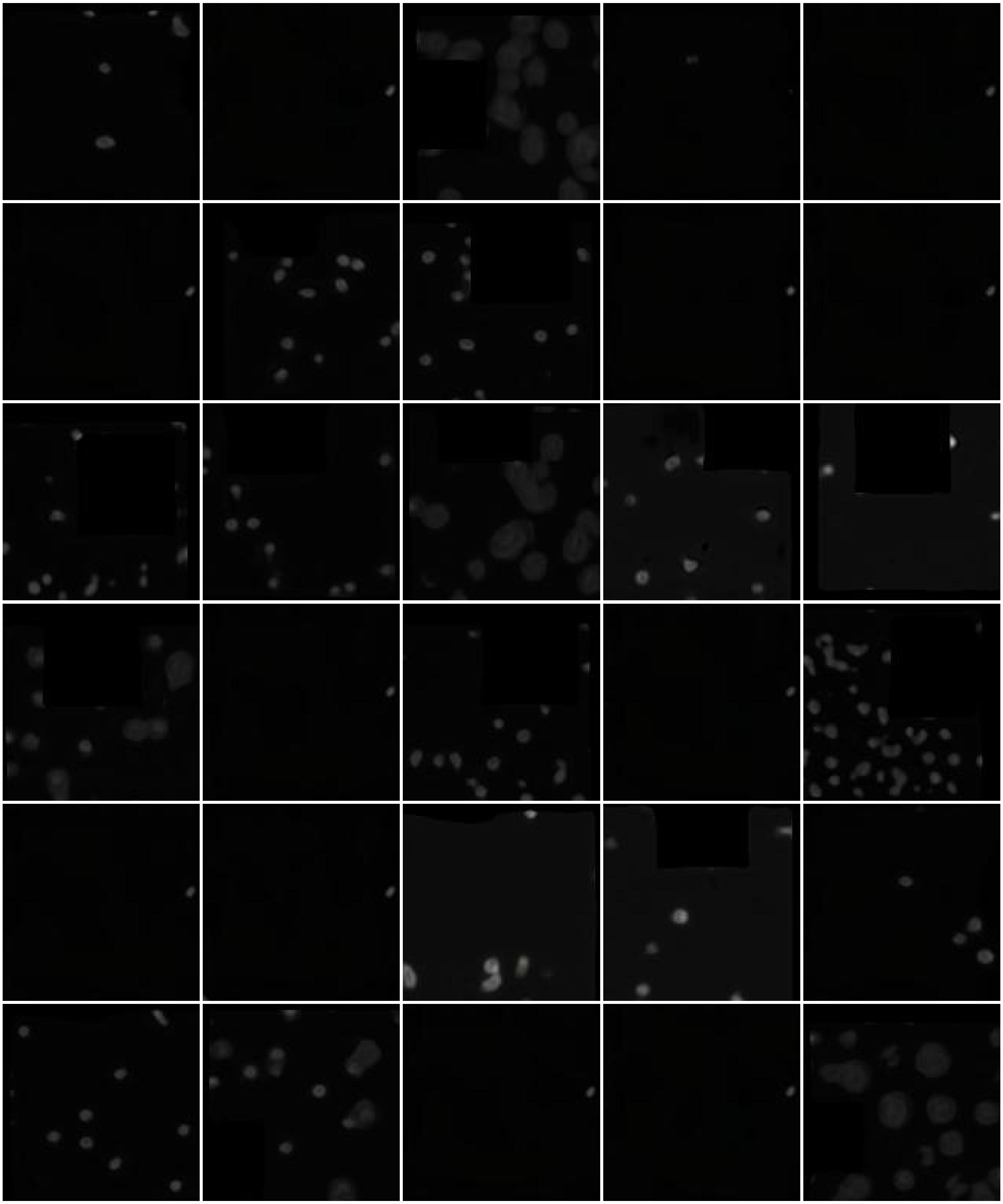

Figure 12: A sample of the final images created from the StyleGAN2-Diff method (128 × 128 for each sample) using the Kaggle dataset.

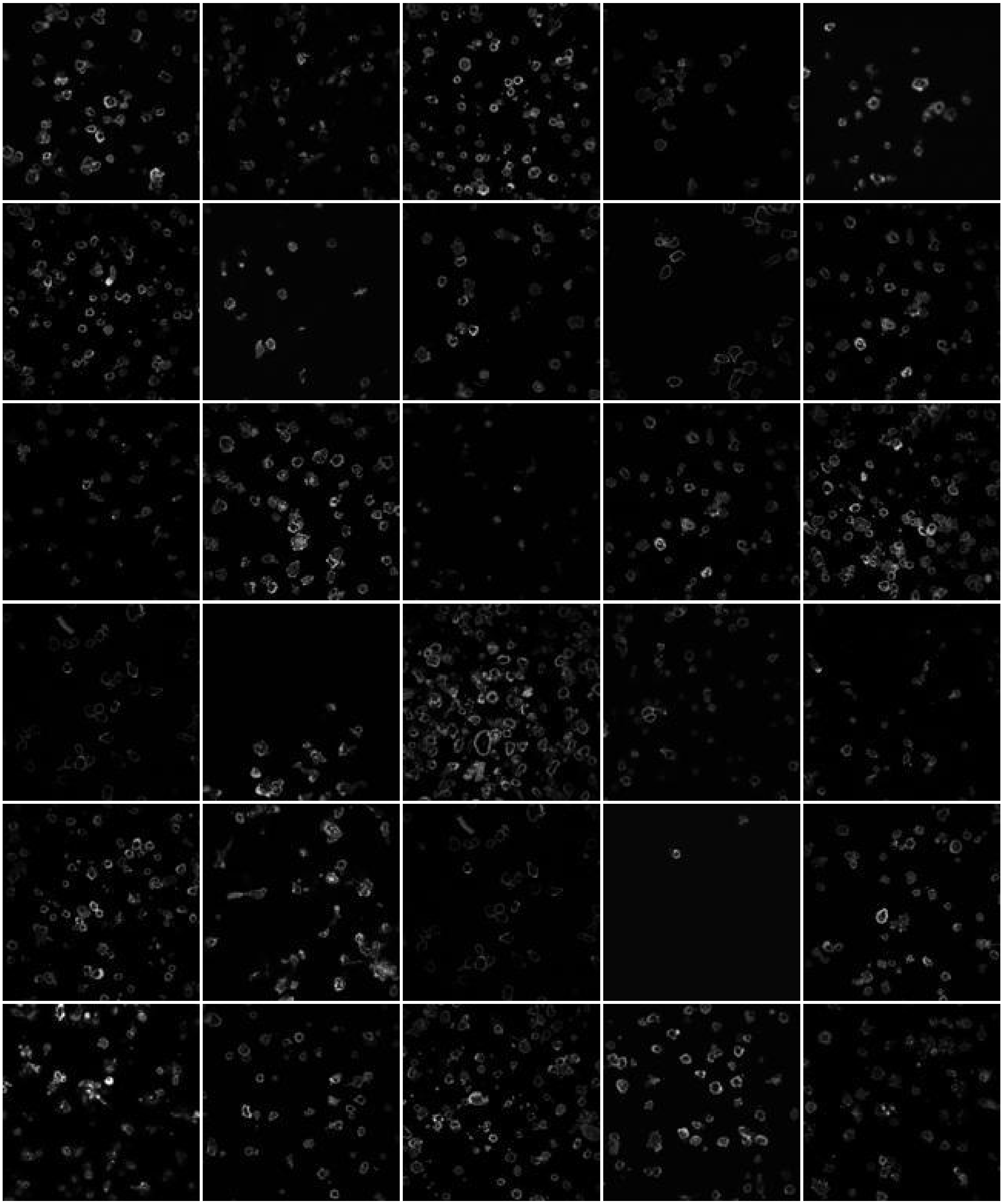

Figure 13: A sample of the final images created from the real images (128 × 128 for each sample) using the CAR-T dataset.

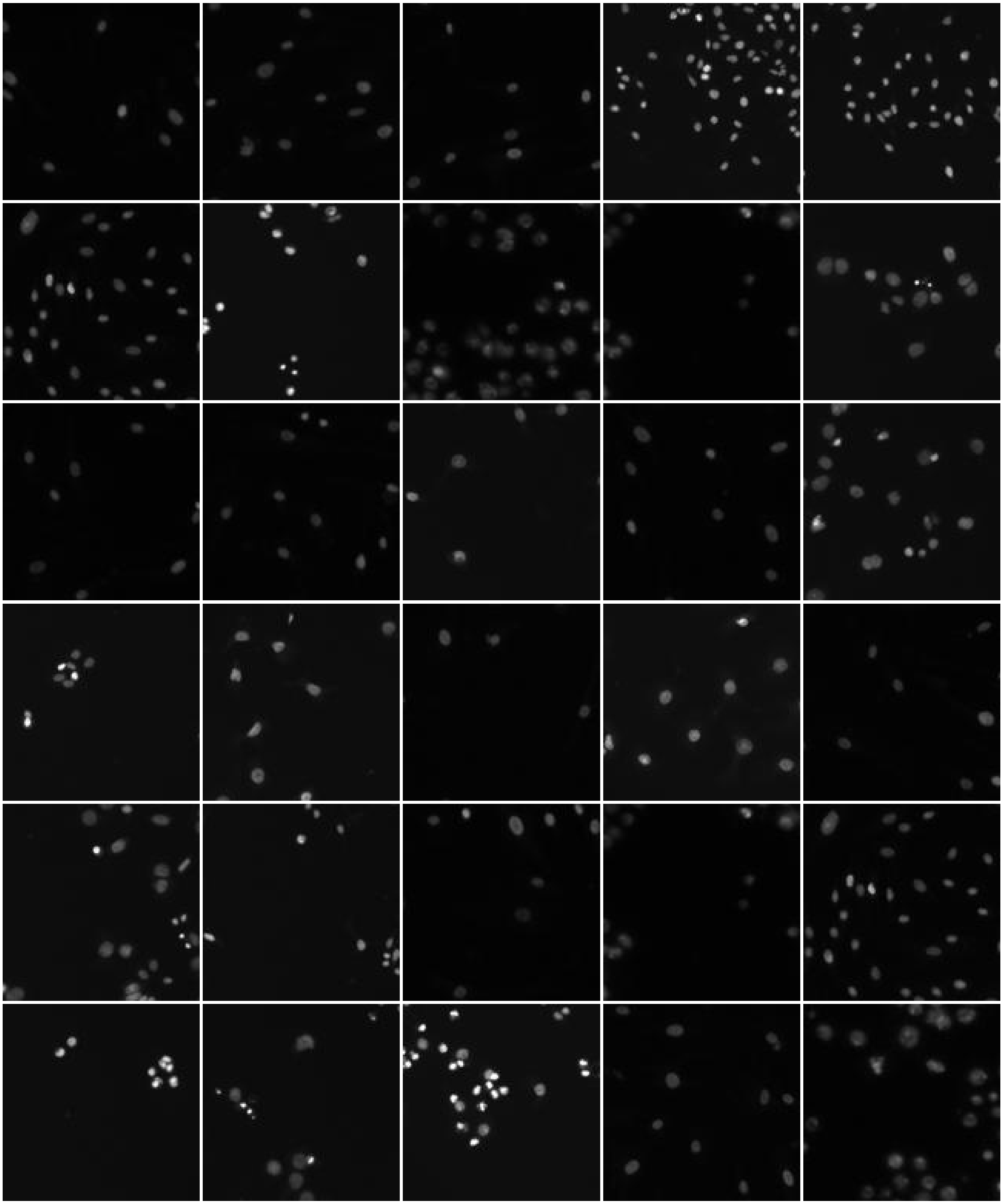

Figure 14: A sample of the final images created from the real images (128 × 128 for each sample) using the Kaggle dataset.