

FA4SANS-GAN: A Novel Machine Learning Generative Adversarial Network to Further Understand Ophthalmic Changes in Spaceflight Associated Neuro-Ocular Syndrome (SANS)

Sharif Amit Kamran, PhD,¹ Khondker Fariha Hossain, MS,¹ Joshua Ong, MD,² Ethan Waisberg, Mb BCh, BAO,³ Nasif Zaman, MS,¹ Salah A. Baker, PhD,⁴ Andrew G. Lee, MD,^{5,6,7,8,9,10,11,12} Alireza Tavakkoli, PhD¹

Purpose: To provide an automated system for synthesizing fluorescein angiography (FA) images from color fundus photographs for averting risks associated with fluorescein dye and extend its future application to spaceflight associated neuro-ocular syndrome (SANS) detection in spaceflight where resources are limited.

Design: Development and validation of a novel conditional generative adversarial network (GAN) trained on limited amount of FA and color fundus images with diabetic retinopathy and control cases.

Participants: Color fundus and FA paired images for unique patients were collected from a publicly available study.

Methods: FA4SANS-GAN was trained to generate FA images from color fundus photographs using 2 multiscale generators coupled with 2 patch-GAN discriminators. Eight hundred fifty color fundus and FA images were utilized for training by augmenting images from 17 unique patients. The model was evaluated on 56 fluorescein images collected from 14 unique patients. In addition, it was compared with 3 other GAN architectures trained on the same data set. Furthermore, we test the robustness of the models against acquisition noise and retaining structural information when introduced to artificially created biological markers.

Main Outcome Measures: For GAN synthesis, metric Fréchet Inception Distance (FID) and Kernel Inception Distance (KID). Also, two 1-sided tests (TOST) based on Welch's *t* test for measuring statistical significance.

Results: On test FA images, mean FID for FA4SANS-GAN was 39.8 (standard deviation, 9.9), which is better than GANgio model's mean of 43.2 (standard deviation, 13.7), Pix2PixHD's mean of 57.3 (standard deviation, 11.5) and Pix2Pix's mean of 67.5 (standard deviation, 11.7). Similarly for KID, FA4SANS-GAN achieved mean of 0.00278 (standard deviation, 0.00167) which is better than other 3 model's mean KID of 0.00303 (standard deviation, 0.00216), 0.00609 (standard deviation, 0.00238), 0.00784 (standard deviation, 0.00218). For TOST measurement, FA4SANS-GAN was proven to be statistically significant versus GANgio ($P = 0.006$); versus Pix2PixHD ($P < 0.00001$); and versus Pix2Pix ($P < 0.00001$).

Conclusions: Our study has shown FA4SANS-GAN to be statistically significant for 2 GAN synthesis metrics. Moreover, it is robust against acquisition noise, and can retain clear biological markers compared with the other 3 GAN architectures. This deployment of this model can be crucial in the International Space Station for detecting SANS.

Financial Disclosure(s): The authors have no proprietary or commercial interest in any materials discussed in this article. *Ophthalmology Science* 2024;4:100493 © 2024 by the American Academy of Ophthalmology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Since the early space shuttle missions, astronauts have reported subjective visual changes during spaceflight. These initial anecdotal reports included decreased focusing ability in the cabin and difficulty reading.^{1,2} These subjective visual changes led to an objective, longitudinal, stepwise investigation process that included preflight and postflight ophthalmic imaging (e.g., OCT orbital ultrasound and fundus photography). The initial description of neuro-ocular findings after long-duration spaceflight (LDSF)

included unilateral and bilateral optic disc edema, posterior globe flattening, choroidal folds, cotton wool spots, and hyperopic refractive error shifts.³ These microgravity-induced ophthalmic findings have been termed spaceflight associated neuro-ocular syndrome (SANS).⁴⁻⁶ Even after returning to Earth, LDSF astronauts have experienced persistent refractive errors that can last for multiple years,⁴ and residual optic disc edema, globe flattening, and choroidal folds have been observed in follow-up

imaging.^{4,7} The National Aeronautics and Space Administration has rated SANS with an elevated likelihood and consequence ratio and a “red” risk, indicating that it is a high priority based on the severity of human health and mission performance.⁸ Thus, further understanding and development of effective countermeasures for SANS is required for future spaceflight exploration.

Although the exact mechanism behind SANS is not well understood, neuro-ophthalmic imaging modalities have played an important role in defining SANS.⁹ The findings from multimodal ocular imaging have evolved over time and prompted substantive discussions on potential countermeasures for SANS.^{3,4,10,11} Current in-flight ocular imaging on the International Space Station (ISS) includes OCT, fundus photography, and orbital ultrasound (Fig 1).^{4,11} Postflight fluorescein angiography (FA) has also been utilized to analyze choroidal folds in SANS.¹ However, even terrestrial FA is technically difficult and requires intravenous fluorescein dye administration, and FA is not available onboard the ISS. Fluorescein angiography is an invasive procedure that utilizes fluorescent dye injected into the bloodstream to evaluate the vasculature and ocular circulation in the retina and choroid. Fluorescein angiography can also highlight optic disc edema (leakage) and choroidal folds. Fluorescein angiography has been used to characterize and stage chorioretinal folds-related maculopathy.^{12–14}

Although FA is not currently available on the ISS, deep learning techniques with ophthalmic imaging have introduced robust approaches to ophthalmic diseases.¹⁵ Deep learning-based architectures have reached state-of-the-art accuracy for recognizing various retinal diseases from OCT and fundus photography.^{16–18} In 2014, Goodfellow et al¹¹ proposed the revolutionary deep learning technique termed generative adversarial network (GAN),¹⁹ introducing a multitude of applications including synthetic image generation. Since then, GANs have been utilized for synthetic image generation from fundus and FA images.^{20–22} This GAN-based application is particularly useful in environments that do not have access to FA. Furthermore, by incorporating the available imaging such as OCT/fundus and functional visual assessments, including visual acuity and contrast sensitivity, we believe that we can eventually build a multimodal system that can become more robust for automated diagnosis and prognosis of SANS. This work is being completed concurrently with the development of a multimodal, head-mounted visual assessment system for rapid assessment of visual acuity, visual fields, contrast sensitivity, Amsler grid, and other functional parameters during spaceflight.^{23–25}

Generative adversarial networks may be the next pivotal technology to further advance our understanding of SANS and supplement our current imaging technology. In the following sections, we introduce “FA4SANS-GAN,” a novel GAN architecture to produce images of 1 modality from another (fundus to FA) to further analyze SANS. We elaborate the results of this deep learning framework and discuss future directions with these advances. To our knowledge, this is the first machine learning GAN-based

architecture to address the unique limitations in analyzing microgravity-induced ophthalmic changes during spaceflight.

Methods

GANs

Generative adversarial networks incorporate a min-max between 2 different architectures, Generator and Discriminator.^{20,21,26} The primary task of the generator is to synthesize realistic images for a new modality of images by learning representative features from a different kind of modality of images. In contrast, the discriminator’s task is to distinguish between the new modality’s real and synthesized images. Generative adversarial network architectures often incorporate multiple learning objective functions to synthesize high-quality visual results and learn discriminative features.

We illustrate such a GAN architecture in Figure 2, which takes color fundus images as input and synthesizes FA images. Combining a coarse-to-fine generator for image translation tasks results in very high-quality images. For this reason, we use a coarse generator G_c to extract and preserve global information, such as the structures of the optic disc, brightness, contrast, and a fine generator G_f for synthesizing smooth FA from fundus images by learning local information such as retinal venules, blood vessels, hemorrhages, and aneurysm. The discriminator distinguishes between real and fake images by first taking both real fundus and FA images and then real fundus and GAN-synthesized FA images. We pair 2 discriminators, D_c and D_f , with our coarse and fine generators. The generator comprises multiple convolutions, residual, attention, down-sampling, up-sampling blocks, and a feature fusion block between the fine and coarse generator. Contrarily, the discriminator consists of residual and down-sampling blocks. The convolution and residual layers are the main blocks for learning representative features to translate between 2 modalities of images. Additionally, the down-sampling blocks decrease the spatial size and increase the depth of the learned features. On the other hand, the up-sampling blocks increase spatial resolution and decrease the features’ depths. The main reason is to compress the spatial information and expand the manifold features to propagate more distinctive and representative features. The attention block is utilized for retaining and propagating rich spatial information that might be lost due to successive down-sampling and up-sampling operations. The fine generator, G_f has an input dimension of 512×512 whereas the coarse generator G_c takes an image with half the resolution, 256×256 , and both architectures synthesize FA images of the same size. The GAN architecture incorporates 4 learning objectives to extract and learn representative features, 2 for the generators and 2 for the discriminators. Both the generator and discriminator utilize multihinge loss for playing min-max game. This loss is given in equation (Eq. [1]). The generator also employs a reconstruction loss given in Eq. (2) to compare between real and synthesized FA images. Additionally, the discriminator utilizes a feature-matching loss, where we insert real FA and generated FA images along with real fundus images in 2 iterations and compute L1-loss for the features of the intermediate layers. This loss is given in Eq. (3).

$$\mathcal{L}_{adv}(G, D) = -E_{x,y}[\min(0, -1 + D(x, y))] - E_x[\min(0, -1 - (D(x, G(x))))] \quad (1)$$

$$\mathcal{L}_{recon}(G) = E_{x,y} \|G(x) - y\|^2 \quad (2)$$



Figure 1. OCT (A), orbital ultrasound (B), and fundus photography (C) on the International Space Station. Courtesy of the National Aeronautics and Space Administration (NASA) Image and Video Library. Permissions: NASA Media Usage Guideline (<https://www.nasa.gov/multimedia/guidelines/index.html>).

$$\mathcal{L}_{fm}(G, D_n) = E_{x,y} \sum_{i=1}^k \frac{1}{N} \left\| D_n^i(x, y) - D_n^i(x, G(x)) \right\| \quad (3)$$

In Eq. (1) the discriminators are first trained on the real fundus, x and real FA image, y and then trained on the real fundus, x and fake FA image, $G(x)$. We start with training the discriminators for 4 iterations on random batches of images. Next, we train the generator while keeping the weights of the discriminators frozen. In Eq. (2), the reconstruction loss is calculated between a real FA, y , and a generated FA, $G(x)$. In Eq. (3), the feature-matching loss is calculated between real fundus and FA image pairs (x, y) and real and generated FA image pairs $(x, G(x))$. We can formulate our final objective function in Eq. (4).

$$\min_G \max_D \mathcal{L}_{adv}(G, D) + \lambda_{recon} [\mathcal{L}_{recon}(G)] + \lambda_{fm} [\mathcal{L}_{fm}(G, D_n)] \quad (4)$$

Here, λ dictates we prioritize the generator over our discriminator. For this we pick a large λ value.

Human Subjects

We employ the fundus and FA pair provided in the Isfahan medical image and signal processing data set²⁰ for training. The data set consists of 30 image pairs of diabetic retinopathy with 3 severities (mild, moderate, and severe nonproliferative diabetic retinopathy) and 29 pairs of normal fundus and FA image pairs collected from 59 patients. Deidentified data was provided by the National Aeronautics and Space Administration. No ethics committee approval was required since all data used were deidentified.

Data Preprocessing

We select 17 pairs for training based on one-to-one alignment of images. The images have a resolution of 576×720 pixels and the images are in red, green, blue (fundus) and grayscale (FA) format. We take 50 random crops of size 512×512 from each image and augment to 850 images for training. For testing, we use 14 image pairs and crop them into 4 quadrants to create 56 image pairs. The detailed procedure is illustrated in Figure 3. We utilized Blur, Sharpening, and Gaussian Noise insertion for the spatial transformation. For radial transformation, we used Pinch and Whirl. All transformations were carried out in GNU Image Manipulation Program.²⁷

Hyperparameters

We use least squares GAN loss function for training our GAN model. We picked weight value of $\lambda_{recon} = 10$ (Eq. 2) and $\lambda_{fm} = 10$ (Eq. 3) for prioritizing the generator over the discriminator. For optimizer, we used Adam with learning rate $\alpha = 0.0002$, $\beta_1 = 0.5$ and $\beta_2 = 0.999$. We train with batch size, $b = 4$ for 100 epochs. It took approximately 8 hours to train our model on an NVIDIA Graphics Processing Unit A30 model.

Results

Image Synthesis Result for Spatial and Radial Transformations

We evaluated our architecture to test its adaptability and robustness to structural changes in the vascular patterns and structure of the eye. In our current experiments, we do not have access to in-flight SANS data, thus, we modeled the visual transformations ourselves and applied them to terrestrial images. This, in turn, will help us evaluate our model's robustness and adaptability. We visually evaluate both the original and transformed synthesized images. The applied transformations are: (1) blurring, to imitate out of focus fundoscopy or fundus photography in the presence of severe cataracts; (2) high contrast, to represent pupil dilation; (3) high peak signal-to-noise ratio (PSNR), to represent sensor and periodic noise; (4) radial distortion, representing the flattening of the retina resulting in the pulled/pushed retinal structure; and (5) rotational distortion, representing increased intraocular pressure. As shown in Figure 4, the proposed architecture produces images very similar to the ground truth under these global changes applied to the fundus image. This experiment shows the adaptability and reproducibility of the proposed network to uncover the changes in vascular structure.

As illustrated in Figure 4, our models produce realistic and convincing outputs. On the other hand, the visual result for Pix2Pix produces distorted local structures due to not learning the vascular information. Compared with that, GANgio and Pix2PixHD also produce impressive results. However, if we zoom in, we can see that the high PSNR and high contrast transformations contain fewer

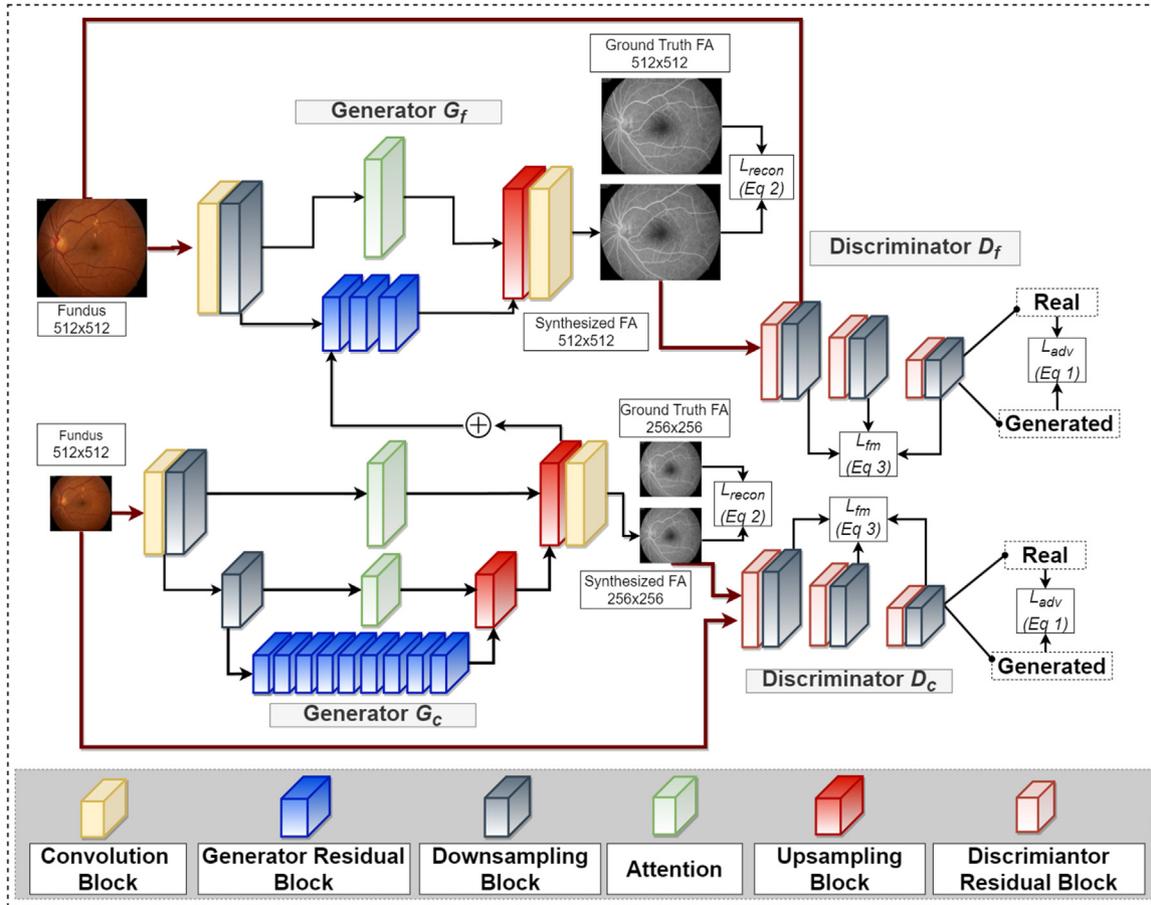


Figure 2. Fundus-to-fluorescein angiography (FA) synthesis generative adversarial network architecture. The architecture comprises 4 separate sub-architectures, 2 Generators and 2 Discriminators (for coarse and fine images). Each of the architectures comprises distinct building blocks, namely, Convolution, Generator Residual, Discriminator Residual, Downsampling, Attention, and Upsampling. The coarse generator and discriminator takes image of 256×256 resolution and fine generator and discriminators takes images of 512×512 resolution. The generators utilize reconstruction loss (equation [Eq] 2) and the discriminators utilize adversarial loss (Eq 1) and feature-matching (Eq 3) functions.

blood vessels than ours. These models, when exposed to radial and rotational distortions, also fail to generate rich venules, exudates, and aneurysms.

Descriptive for Synthesized FA Images

For quantitative evaluation, we evaluated the performance of the proposed method (FA4SANS-GAN) with that of 3 state-of-the-art techniques (GANgio,²⁶ Pix2Pix,²⁸ and Pix2PixHD²⁹) using Fréchet Inception Distance (FID)³⁰ and Kernel Inception Distance (KID).³¹ We computed the FID scores that were generated from the original angiogram, and those generated from the distorted fundus images by the 5 global and structural changes, i.e., blurring, high contrast manipulation, high PSNR, radial distortions, and rotational distortions. The quantitative results are illustrated in Tables 1 and 2. It should be noted that lower FID and KID scores indicate better results.

As it can be observed from Tables 1 and 2, our proposed method (FA4SANS-GAN) on average performs better compared with the state-of-the-art in all but 1 case. Only

for blurred images, the GANgio (also introduced by our team²⁶) has a slightly lower FID score on average compared with FA4SANS-GAN (Table 1).

Inferential Statistics for Synthesized FA Images

In the computer science and machine learning literature, descriptive statistics are quite frequently used to demonstrate how various machine learning methods compare to each other. Often times, many compared results are within a few percentage points of one another. As a result, comparisons drawn from the average accuracies across the test data do not specifically describe the statistical effect sizes or establish the statistical significance of the results. Although it might seem that performing null hypothesis significant testing would solve the issue, there are situations where this traditional approach will result in inappropriate statistical conclusions. According to Lakens,³² scientists should be able to provide support for both the null and alternative hypotheses in conducting an experiment. Traditional significance testing that is designed to merely reject the null hypothesis, although widely used in

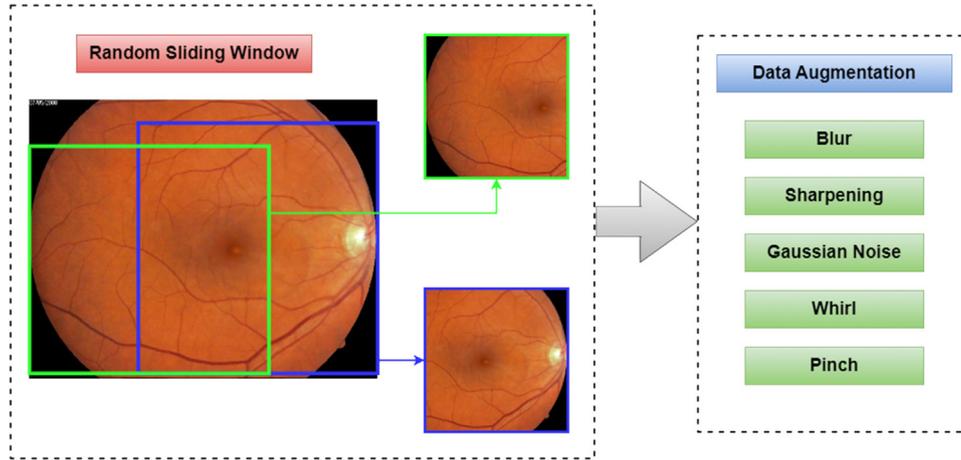


Figure 3. Random sliding window is utilized to take 512×512 cropping from original image and then different data augmentation such as Blur, Sharpening, Gaussian Noise, Whirl, and Pinch are used to further transform the image for training and testing.

social sciences, does not specifically allow us to conclude that meaningful effects are absent.³² Many studies, especially in clinical therapeutic settings, are focusing on determining the statistical superiority, equivalence, or noninferiority of newer treatments compared with standard of care.³³ Conducting equivalence testing can be done simply by using two 1-sided tests (TOST).³⁴

In many machine learning applications, new architectures are developed with the goal of improving the performance compared with state of the art. However, when specialized architectures and models are designed for specific applications, it is imperative to demonstrate that the new model is not inferior to the state of the art in a formal manner and with appropriate design. Therefore, we should consider the smallest effect size of interest representing upper and lower equivalence bounds $(-\Delta_L, \Delta_U)$. We then perform TOST on 2 null hypotheses; $H_{0_1} : \Delta \leq -\Delta_L$ and $H_{0_2} : \Delta \geq \Delta_U$.³² If both are rejected, we can conclude that the proposed method is statistically equivalent to standard methods. The new method could be tested to ensure it is not statistically inferior to the state of the art by setting the upper equivalence bound to ∞ (when higher values are better). To perform the TOST, first the significance bounds are established as either a standardized difference such as Cohen's d or raw score differences.³² The TOST will be performed on the observed data at an established significance level (e.g., $\alpha < .05$). This will result in $(1 - 2\alpha)$ confidence intervals that can be analyzed to conclude whether the compared methods are statistically equivalent, 1 is superior, or at least noninferior. In presenting the results, we could also include the significance level of our findings along with the confidence interval values.

Figures 5 and 6 show the results of our equivalence/noninferiority tests for comparing the FID and KID, respectively, of FA4SANS-GAN with that of GANgio (A), Pix2PixHD (B), and Pix2Pix (C). We have performed the comparisons on generating images from the original data (the bottom panel) and data transformed by blurring the image, sharpening the image to increase contrast, adding noise to increase the PSNR, and performing radial as well as

rotational distortions on the image. We consider a Cohen's $d = 0.4$ as the smallest effect size of interest to establish the upper and lower bounds of $\pm \delta = 5$ for the TOST procedures.

Looking at the overall performance section (the top section of Figs 5 and 6) we can conclude that the proposed FA4SANS-GAN is superior to all 3 traditional methods (A, B, and C). The TOST procedure based on Welch's t test indicated that the observed effect size was significantly within the superiority bounds of $-\infty$ and 0 scale points for (A) FA4SANS-GAN versus GANgio with $t(160.39) = -4.98$, $P = 0.006$; (B) FA4SANS-GAN versus Pix2PixHD with $t(165.49) = -8.76$, $P < 0.00001$; and (C) FA4SANS-GAN versus Pix2Pix with $t(165.67) = -14.49$, $P < 0.00001$. Similar observations were made with respect to KID measurements.

Discussion

To our knowledge, this is the first utilization of a GAN architecture to address the unique limitations in further analyzing SANS during spaceflight. The results showcase that our model is more robust for normal and transformed FA images than other state-of-the-art generative adversarial architectures. Moreover, for quantitative metrics for GAN evaluations such as FID and KID, the generated images preserve the vascular information accurately. Finally, our statistical hypothesis testing confirms that FA4SANS-GAN produces significantly better results compared with 3 state-of-the-art models. These results demonstrate FA4SANS-GAN's robust ability to produce an unavailable, relevant imaging modality on the ISS with current onboard modalities to further understand SANS.

There are several pertinent considerations for this deep learning technology in regard to our current understanding of SANS. Terrestrially, chronic choroidal folds have been observed to develop into a maculopathy termed chorioretinal fold-related maculopathy.^{12,14} Fluorescein angiography has been able to detect choroidal folds on Earth and following

| Architecture | Original | Blurred | High Contrast | High PSNR | Radial Distortion | Rotational Distortion |
|--------------|----------|---------|---------------|-----------|-------------------|-----------------------|
| Input | | | | | | |
| Ground Truth | | | | | | |
| FA4SANS-GAN | | | | | | |
| GANgio | | | | | | |
| Pix2PixHD | | | | | | |
| Pix2Pix | | | | | | |

Figure 4. Comparative comparison of fluorescein angiogram generated from normal and distorted fundus images with biological markers. FA = fluorescein angiography; GAN = generative adversarial network; PSNR = peak signal-to-noise ratio; SANS = spaceflight associated neuro-ocular syndrome.

injection the peaks are observed to become bright (hyperfluorescent) while the troughs of choroidal folds remain dark (hypofluorescent).¹³ Olsen et al¹² classified chorioretinal folds-related maculopathy into 3 stages with FA by observing areas of hypofluorescence and leakage areas. Occult choroidal neovascularizations were also observed in this study¹² and FA is a powerful tool to identify choroidal neovascularization³⁵ (e.g., neovascular age-related macular degeneration).³⁶

On prolonged exploration and future potential colonization missions, astronauts (and space colonists) will face potential ocular pathology from SANS but also the typical and age-related ocular pathologies (e.g., age-related macular degeneration). Thus, there may be a future need for GAN-based architectures to supplement available imaging

modalities. In 2018, OCT angiography (OCTA) was introduced to the ISS to allow for noninvasive posterior segment vascular imaging. OCT angiography and FA have different strengths in vascular imaging of the posterior segment,³⁷ but the invasive nature of FA makes OCTA an ideal choice for spaceflight. This noninvasive, GAN-based approach with fundus imaging addresses the invasive concerns of FA and in conjunction with in-flight OCTA may be complementary for assessing SANS. Future computational research is also being conducted in multimodal registration of these 2 angiography techniques to provide optimized angiographic analysis.³⁸

This GAN-based approach for imaging synthesis of unavailable modalities on the ISS may also be applied to magnetic resonance imaging (MRI). Postflight analysis MRI has documented posterior globe flattening with high

Table 1. Quantitative FID Score Comparisons between Our Method and 3 State-of-the-Art on Images with Different Transformations

| Image Transformation | FA4SANS-GAN | GANgio | Pix2PixHD | Pix2Pix |
|-----------------------|-------------|-------------|-------------|-------------|
| Original | 39.8 (9.9) | 43.2 (13.7) | 57.3 (11.5) | 67.5 (11.7) |
| High PSNR | 48.9 (12.0) | 55.6 (12.8) | 66.2 (15.3) | 63.8 (10.5) |
| Blurred | 47.4 (10.8) | 45.8 (11.8) | 59.1 (11.6) | 69.3 (12.2) |
| High contrast | 41.1 (14.3) | 49.3 (16.2) | 63.0 (12.4) | 64.9 (12.2) |
| Radial distortion | 39.9 (12.0) | 47.7 (16.0) | 54.4 (11.5) | 62.3 (12.5) |
| Rotational distortion | 46.5 (10.0) | 52.9 (13.9) | 62.9 (10.7) | 61.9 (8.7) |

Results are reported as $\mu(\sigma)$, where μ is the mean and σ is the standard deviation.

FA = fluorescein angiography; FID = Fréchet Inception Distance; GAN = generativ adversarial network; PSNR = peak signal-to-noise ratio; SANS = spaceflight associated neuro-ocular syndrome.

resolution and sensitivity;⁴ postflight MRI has also documented upward brain shift in LDSF astronauts which has served as a foundational finding for a hypothesis in SANS pathogenesis.^{10,39} Postflight MRI has shown an upward shift of the brain and optic chiasm in astronauts after LDSF.⁴⁰ Proposed by Shinojima et al,¹⁰ this upward shift may pull the connected optic nerve rearward and cause a restoration, anterior force of the dura of the optic nerve sheath onto the posterior eye. This restoration force would lead to the posterior globe flattening as seen in SANS. To further understand this hypothesis, quantification of the posterior globe flattening during spaceflight with GAN-based synthetic MRI may allow for better understanding of the effects of upward brain shift. Various studies have successfully incorporated GAN architectures with other imaging modalities for MRI image reconstruction,⁴¹ magnetic resonance-to-positron emission tomography image translation,⁴⁰ and synthetic image generation.⁴² Given the revolutionary applications involving GANs in medical imaging, this deep learning framework may be extended for translating and detecting SANS-like conditions without MRI images in spaceflight. Although orbital ultrasound can identify posterior globe flattening, MRI provides higher resolution and has been utilized to provide quantification of volumetric displacement of the posterior globe after spaceflight.^{43,44} The severity of these effects may be correlated with the duration of spaceflight, which can be determined with quantifiable, high-resolution images. Further understanding of whether these mechanisms play a role in SANS pathogenesis will

likely help aid in the development and application of effective countermeasures.

As SANS serves as one of the largest physiological barriers to planetary spaceflight, clinical countermeasures have been extensively studied and continue to be an area of prioritized investigation. A notable clinical application of this GAN technology is the ability to further evaluate vasculature changes/SANS progression in-flight that can help guide countermeasure titration and management.⁴⁵ Various and diverse SANS countermeasures have been investigated, many that can be titrated accordingly to this technology in conjunction with other in-flight modalities. It has been hypothesized that certain single nucleotide polymorphisms in 1-carbon metabolic pathways may lead to increased risk of SANS.^{46,47} Zwart et al⁴⁶ identified that increased number of G alleles in methionine synthase reductase 66 (MTRR 66) and C alleles serine hydroxymethyltransferase1 1420 (SHMT1 1420) led to increased odds of ophthalmic changes during spaceflight, and when combined with lower B-vitamin status (vitamin B2, B6, and B9), there is an increased risk of deterioration of vision. It has been investigated that B-vitamin status can modify polymorphism effects in 1-carbon pathways.⁴⁸ Thus, it has been proposed that B-vitamin supplementation may serve as an efficient, low-risk countermeasure for SANS^{47,49} and it has been reported that plans are in development to test B-vitamin supplementation on the ISS.⁴⁷ Modulation of the translaminal pressure gradient with goggles has also emerged as a potential countermeasure for SANS.⁵⁰ During spaceflight, it is thought that there is a greater increase in

Table 2. Quantitative KID score comparisons between Our Method and 3 State-of-the-Art on Images with Different Transformations

| Image Transformation | FA4SANS-GAN | GANgio | Pix2PixHD | Pix2Pix |
|-----------------------|-------------|-------------|-------------|-------------|
| Original | 27.8 (16.7) | 30.3 (21.6) | 61.0 (23.8) | 67.5 (11.7) |
| High PSNR | 40.4 (19.5) | 41.6 (24.7) | 61.0 (17.9) | 69.3 (12.2) |
| Blurred | 31.6 (19.8) | 40.9 (22.0) | 70.0 (25.6) | 64.9 (12.2) |
| High contrast | 35.6 (19.5) | 47.0 (23.3) | 61.7 (21.1) | 62.3 (12.5) |
| Radial distortion | 37.5 (23.6) | 46.7 (20.8) | 73.5 (20.7) | 61.9 (8.7) |
| Rotational distortion | 41.9 (18.1) | 53.7 (24.5) | 81.0 (18.1) | 63.8 (10.5) |

Results are reported as $\mu(\sigma)$, where μ is the mean and σ is the standard deviation

FA = fluorescein angiography; GAN = generativ adversarial network; KID = Kernal Inception Distance; PSNR = peak signal-to-noise ratio; SANS = spaceflight associated neuro-ocular syndrome.

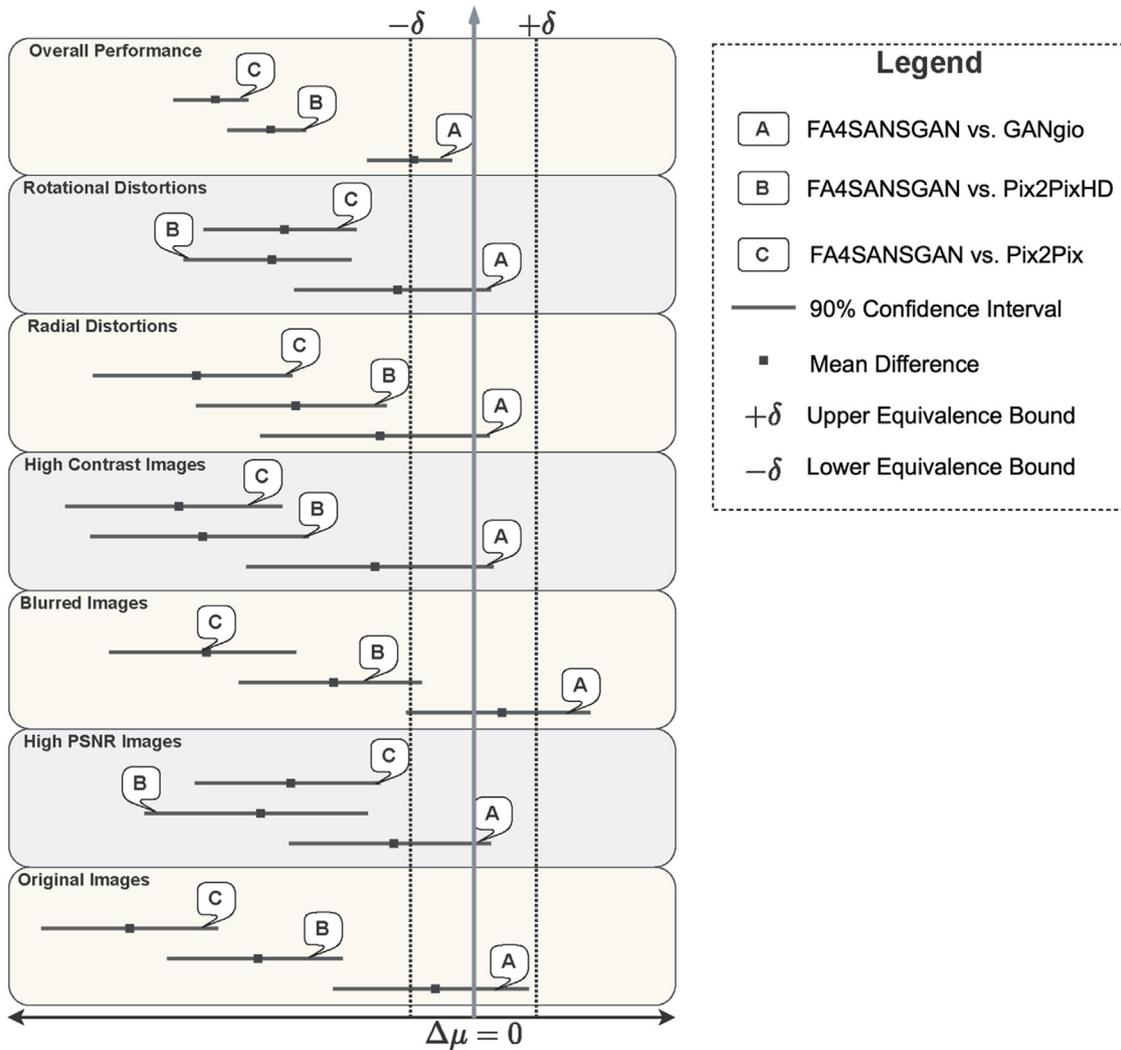


Figure 5. Inferential statistics based on two 1-sample tests procedure to compare Fréchet Inception Distance equivalence, superiority, and noninferiority of the proposed methods FA4SANS-GAN compared with GANgio (A), Pix2PixHD (B), and Pix2Pix (C). Results shown in the overall performance tab shows that the proposed method is statistically superior to all 3 state-of-the-art, as the 90% confidence interval is not crossing the 0 mean difference threshold, and the upper 90% confidence bound is lower than the upper equivalence bound ($+\delta$). FA = fluorescein angiography; GAN = generative adversarial network; PSNR = peak signal-to-noise ratio; SANS = spaceflight associated neuro-ocular syndrome.

intracranial pressure compared with intraocular pressure leading to a negative translaminar pressure gradient (TLPG) which may be a cause of SANS. Scott et al⁵⁰ observed that application of goggles during a SANS terrestrial analog led to an increase in TLPG, suggesting that goggles may be utilized as a SANS countermeasure by attenuating the negative TLPG shift during spaceflight. Lower body negative pressure (LBNP) has also been an area of high interest in SANS mitigation.^{51,52} This countermeasure aims to reduce cephalad fluid shifts that are observed in microgravity which have been proposed as a contributory factor to SANS pathogenesis.⁵² Given that LBNP has been utilized during spaceflight for other physiological systems,⁵³ this technique serves as a relatively translatable countermeasure for SANS in-flight. Hearon et al⁵² observed in a randomized crossover trial that nightly LBNP at -20 mmHg attenuated the increase in

choroidal volume and area normally seen when subjects laid supine. The authors suggested that nightly LBNP may be an effective countermeasure for SANS. B-vitamin supplementation, modulation of the TLPG with goggles, and LBNP can all be titrated in dosage/length of application accordingly. Thus, this GAN technology may provide additional insights into the ocular vasculature and SANS signs that can help determine whether additional countermeasures and/or an increase in titration of a countermeasure is necessary. Furthermore, this GAN technology may help to further elucidate the effects of specific countermeasures on ocular vasculature to further understand SANS pathogenesis.

Before in-flight integration of FA4SANS-GAN, this architecture will benefit from testing with a terrestrial analog for SANS. Given the nature of spaceflight and the limited number of astronauts, ground-based testing with available

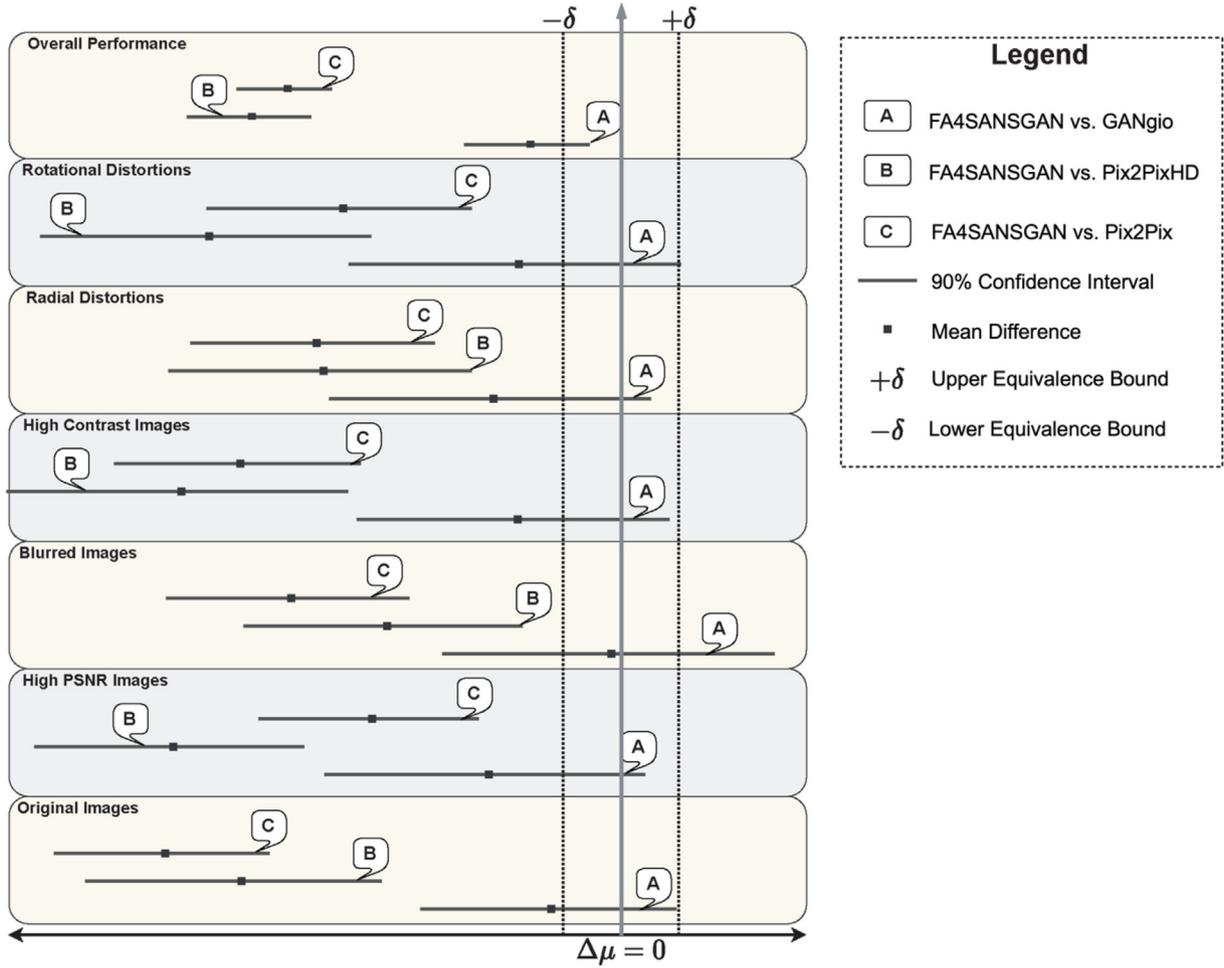


Figure 6. Inferential statistics based on two 1-sample tests procedure to compare Kernel Inception Distance equivalence, superiority, and noninferiority of the proposed method FA4SANS-GAN compared with GANgio (A), Pix2PixHD (B), and Pix2Pix (C). Results shown in the overall performance tab shows that the proposed method is statistically superior to all 3 state-of-the-art, as the 90% confidence interval is not crossing the 0 mean difference threshold, and the upper 90% confidence bound is lower than the upper equivalence bound ($+\delta$). FA = fluorescein angiography; GAN = generativ adversarial network; PSNR = peak signal-to-noise ratio; SANS = spaceflight associated neuro-ocular syndrome.

ground-truth imaging modalities will be a valuable future direction. Head-down tilt bed rest is a terrestrial analog that mimics the cephalad fluid shifts during microgravity and has become a promising analog for SANS.⁵⁴ Sixty days of strict head-down tilt bed rest has been observed to produce optic disc edema and chorioretinal folds, signs of SANS.⁵⁵ Thus, testing this architecture on healthy subjects in head-down tilt bed rest for extended periods of time will allow for further understanding of this architecture’s application in monitoring the development of pathological signs of SANS.^{56,57} Additionally, this GAN architecture will likely have terrestrial applications on Earth in under-resourced austere areas that do not have access to MRI or FA.^{58,59} By generating GAN-based synthetic images with cost-effective, less-invasive imaging modalities, individuals may be able to have increased accessibility with synthetic modalities in the future.

One of the limitations of our study is the missing SANS data for evaluating our model. By assessing the extraterrestrial SANS data, we can further improve our GAN model to accommodate changes needed to improve its image synthesis pipeline. Additionally, patient-specific information can also help synthesize FA images containing the progression of disease or degradation of macula over time. A future direction is to conduct experiments with such meta-information to extend our work further.

FA4SANS-GAN is a novel, first of its kind, machine learning GAN architecture designed to address the unique spaceflight challenges to further understand SANS. The results from FA4SANS-GAN showed that our model showcases an improved overall performance compared with other state-of-the-art GAN architecture for FA synthesis both quantitatively and qualitatively. Moreover, the testing demonstrates the statistical significance of the

results produced by our model and can robustly generate useful and relevant imaging for SANS that is unavailable on the ISS. This deep learning framework may help deepen humanity's understanding of this microgravity-induced

phenomenon and help to advance countermeasure development for optimal human safety and performance for spaceflight.

Footnotes and Disclosures

Originally received: December 10, 2022.

Final revision: January 11, 2024.

Accepted: February 5, 2024.

Available online: February 15, 2024. Manuscript no. XOPS-D-22-00263R1.

¹ Human-Machine Perception Laboratory, Department of Computer Science and Engineering, University of Nevada, Reno, Reno, Nevada.

² Department of Ophthalmology and Visual Sciences, University of Michigan Kellogg Eye Center, Ann Arbor, Michigan.

³ Department of Ophthalmology, University College Dublin School of Medicine, Belfield, Dublin, Ireland.

⁴ Department of Physiology and Cell Biology, University of Nevada School of Medicine, Reno, Nevada.

⁵ Center for Space Medicine, Baylor College of Medicine, Houston, Texas.

⁶ Department of Ophthalmology, Blanton Eye Institute, Houston Methodist Hospital, Houston, Texas.

⁷ Houston Methodist Research Institute, Houston Methodist Hospital, Houston, Texas.

⁸ Departments of Ophthalmology, Neurology, and Neurosurgery, Weill Cornell Medicine, New York, New York.

⁹ Department of Ophthalmology, University of Texas Medical Branch, Galveston, Texas.

¹⁰ Department of Ophthalmology, University of Texas MD Anderson Cancer Center, Houston, Texas.

¹¹ Department of Ophthalmology, Texas A&M College of Medicine, Texas.

¹² Department of Ophthalmology, The University of Iowa Hospitals and Clinics, Iowa City, Iowa.

Disclosure(s):

All authors have completed and submitted the ICMJE disclosures form.

The authors have no proprietary or commercial interest in any materials discussed in this article.

This research was supported by NASA Grant (80NSSC20K183): A Non-intrusive Ocular Monitoring Framework to Model Ocular Structure and Functional Changes due to Long-term Spaceflight.

HUMAN SUBJECTS: No identifiable human subjects or human subjects data were included in this study. Deidentified data was provided by the National Aeronautics and Space Administration. No ethics committee approval was required since all data used were deidentified.

No animal subjects were used in this study.

Author Contributions:

Conception and design: Kamran, Hossain, Ong, Tavakkoli

Data collection: Kamran, Hossain, Tavakkoli

Analysis and interpretation: Kamran, Hossain, Ong, Tavakkoli

Obtained funding: Tavakkoli

Overall responsibility: Kamran, Hossain, Ong, Waisberg, Zaman, Baker, Lee, Tavakkoli

Abbreviations and Acronyms:

Eq. = equation; **FA** = fluorescein angiography; **FID** = Fréchet Inception Distance; **GAN** = generative adversarial network; **KID** = Kernel Inception Distance; **ISS** = International Space Station; **LBNP** = lower body negative pressure; **LDSF** = long-duration spaceflight; **MRI** = magnetic resonance imaging; **OCTA** = OCT angiography; **PSNR** = peak signal-to-noise ratio; **SANS** = spaceflight associated neuro-ocular syndrome; **TLPG** = translaminal pressure gradient; **TOST** = two 1-sided tests.

Keywords:

Spaceflight-associated neuro-ocular syndrome, Generative adversarial networks, Ophthalmic imaging, Space medicine, Astronaut.

Correspondence:

Alireza Tavakkoli, PhD, Human-Machine Perception Laboratory, Department of Computer Science and Engineering, University of Nevada, Reno, Nevada, United States. E-mail: tavakkol@unr.edu.

References

1. Stenger MB, Tarver WJ, Brunstetter T, et al. Evidence Report: Risk of Spaceflight Associated Neuro-Ocular Syndrome (SANS). *Human Research Program Human Health Countermeasures Element*. National Aeronautics and Space Administration; 2017.
2. Jacob J, Bloomberg MFR, Gilles R, et al. Evidence Report: Risk of Impaired Control of Spacecraft/Associated Systems and Decreased Mobility Due to Vestibular/Sensorimotor Alterations Associated with Spaceflight. *Human Research Program Human Health Countermeasures Element*. National Aeronautics and Space Administration; 2016.
3. Mader TH, Gibson CR, Pass AF, et al. Optic disc edema, globe flattening, choroidal folds, and hyperopic shifts observed in astronauts after long-duration space flight. *Ophthalmology*. 2011;118:2058–2069.
4. Lee AG, Mader TH, Gibson CR, et al. Spaceflight associated neuro-ocular syndrome (SANS) and the neuro-ophthalmologic effects of microgravity: a review and an update. *NPJ Microgravity*. 2020;6:7. <https://doi.org/10.1038/s41526-020-0097-9>.
5. Ong J, Mader TH, Gibson CR, et al. Spaceflight associated neuro-ocular syndrome (SANS): an update on potential microgravity-based pathophysiology and mitigation development. *Eye (Lond)*. 2023;37:2609–2415.
6. Ong J, Lee AG. An introduction to space medicine and the physiological effects of spaceflight on the human body. In: Ong J, Lee AG, eds. *Spaceflight Associated Neuro-Ocular Syndrome*. Academic Press; 2022:1–7.
7. Mader TH, Gibson CR, Otto CA, et al. Persistent asymmetric optic disc swelling after long-duration space flight: implications for pathogenesis. *J Neuroophthalmol*. 2017;37:133–139.
8. Patel ZS, Brunstetter TJ, Tarver WJ, et al. Red risks for a journey to the red planet: the highest priority human health risks for a mission to Mars. *npj Microgravity*. 2020;6:33. <https://doi.org/10.1038/s41526-020-00124-6>.
9. Ong J, Tavakkoli A, Strangman G, et al. Neuro-ophthalmic imaging and visual assessment technology for spaceflight associated neuro-ocular syndrome (SANS). *Surv Ophthalmol*. 2022;67:1443–1466.

10. Shinjima A, Kakeya I, Tada S. Association of space flight with problems of the brain and eyes. *JAMA Ophthalmol.* 2018;136:1075–1076.
11. Ong J, Tarver W, Brunstetter T, et al. Spaceflight associated neuro-ocular syndrome: proposed pathogenesis, terrestrial analogues, and emerging countermeasures. *Br J Ophthalmol.* 2023;107:895–900.
12. Olsen TW, Palejwala NV, Lee LB, et al. Chorioretinal folds: associated disorders and a related maculopathy. *Am J Ophthalmol.* 2014;157:1038–1047.
13. Agrawal M, Tripathy K. *Choroidal Folds*. StatPearls Publishing; 2024.
14. Masalkhi M, Ong J, Waisberg E, Lee AG. Chorioretinal folds in astronauts: risk of chorioretinal fold-related maculopathy and terrestrial staging of disease. *Eye (Lond).* 2024;38:412–413.
15. Ting DSW, Pasquale LR, Peng L, et al. Artificial intelligence and deep learning in ophthalmology. *Br J Ophthalmol.* 2019;103:167–175.
16. Poplin R, Varadarajan AV, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat Biomed Eng.* 2018;2:158–164.
17. Kearney V, Ziemer BP, Perry A, et al. Attention-aware discrimination for MR-to-CT image translation using cycle-consistent generative adversarial networks. *Radiol Artif Intell.* 2020;2:e190027. <https://doi.org/10.1148/ryai.2020190027>.
18. Hossain KF, Kamran SA, Ong J, et al. Revolutionizing space health (Swin-FSR): advancing super-resolution of fundus images for SANS visual assessment technology. Posted online August 11, 2023. arXiv 2308.06332. <https://doi.org/10.48550/arXiv.2308.06332>
19. Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks. *Adv Neural Inf Process Syst.* 2014;27.
20. Kamran SA, Hossain KF, Tavakkoli A, et al. VTGAN: Semi-supervised retinal image synthesis and disease prediction using vision transformers. Posted online August 14, 2021. arXiv. 210406757. <https://doi.org/10.48550/arXiv.2104.06757>
21. Kamran SA, Hossain KF, Tavakkoli A, Zuckerbrod S. Attention2angiogan: synthesizing fluorescein angiography from retinal fundus images using generative adversarial networks. *25th International Conference on Pattern Recognition (ICPR)*. 2020:9122–9129.
22. Kamran SA, Hossain KF, Tavakkoli A, et al. A generative adversarial deep neural network to translate between ocular imaging modalities while maintaining anatomical fidelity. *J Vis.* 2022;22. <https://doi.org/10.1167/jov.22.3.3>.
23. Ong J, Zaman N, Kamran SA, et al. A multi-modal visual assessment system for monitoring spaceflight associated neuro-ocular syndrome (SANS) during long duration spaceflight. *J Vis.* 2022;22. <https://doi.org/10.1167/jov.22.3.6>.
24. Ong J, Tavakkoli A, Zaman N, et al. Terrestrial health applications of visual assessment technology and machine learning in spaceflight associated neuro-ocular syndrome. *npj Microgravity.* 2022;8:37. <https://doi.org/10.1038/s41526-022-00222-7>.
25. Ong J, Zaman N, Waisberg E, et al. Head-mounted digital metamorphopsia suppression as a countermeasure for macular-related visual distortions for prolonged spaceflight missions and terrestrial health. *Wearable Technol.* 2022;3. <https://doi.org/10.1017/wtc.2022.21>.
26. Tavakkoli A, Kamran SA, Hossain KF, Zuckerbrod SL. A novel deep learning conditional generative adversarial network for producing angiography images from retinal fundus photographs. *Sci Rep.* 2020;10:21580. <https://doi.org/10.1038/s41598-020-78696-2>.
27. GIMP: GNU Image Manipulation Program: GIMP Team;2019.
28. Isola P, Zhu J, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. 2017. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017: 5967–5976.
29. Wang T, Liu M, Zhu J, et al. High-resolution image synthesis and semantic manipulation with conditional GANs. 2018. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2018:8798–8807.
30. Heusel M, Ramsauer H, Unterthiner T, et al. GANs Trained by a two time-scale update rule converge to a local Nash equilibrium. Published online June 26, 2017. arXiv 1706.08500. <https://doi.org/10.48550/arXiv.1706.08500>
31. Bińkowski M, Sutherland DJ, Arbel M, Gretton A. Demystifying MMD GANs. 6th International Conference on Learning Representations. Posted online January 14, 2021. arXiv 1801.01401. <https://doi.org/10.48550/arXiv.1801.01401>
32. Lakens D. Equivalence tests: a practical primer for t tests, correlations, and meta-analyses. *Soc Psychol Personal Sci.* 2017;8:355–362.
33. Walker E, Nowacki AS. Understanding equivalence and noninferiority testing. *J Gen Intern Med.* 2011;26:192–196.
34. Schuurmann DJ. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *J Pharmacokinet Biopharm.* 1987;15:657–680.
35. Ruia S, Tripathy K. *Fluorescein Angiography*. StatPearls Publishing; 2022.
36. Shao J, Choudhary MM, Schachat AP. Neovascular age-related macular degeneration. *Dev Ophthalmol.* 2016;55:125–136.
37. La Mantia A, Kurt RA, Mejor S, et al. Comparing fundus fluorescein angiography and swept-source optical coherence tomography angiography in the evaluation of diabetic macular perfusion. *Retina.* 2019;39:926–937.
38. Martinez-Rio J, Carmona EJ, Cancelas D, et al. Robust multimodal registration of fluorescein angiography and optical coherence tomography angiography images using evolutionary algorithms. *Comput Biol Med.* 2021;134:104529. <https://doi.org/10.1016/j.compbiomed.2021.104529>.
39. Roberts DR, Albrecht MH, Collins HR, et al. Effects of spaceflight on astronaut brain structure as indicated on MRI. *N Engl J Med.* 2017;377:1746–1753.
40. Sikka A, Virk JS, Bathula DR. MRI to PET cross-modality translation using globally and locally aware GAN (GLA-GAN) for multi-modal diagnosis of Alzheimer’s disease. Published online August 4, 2021. arXiv 2108.02160. <https://doi.org/10.48550/arXiv.2108.02160>
41. Han C, Rundo L, Murao K, et al. MADGAN: Unsupervised medical anomaly detection GAN using multiple adjacent brain MRI slice reconstruction. *BMC Bioinformatics.* 2021;22. <https://doi.org/10.1186/s12859-020>.
42. Ran M, Hu J, Chen Y, et al. Denoising of 3D magnetic resonance images using a residual encoder–decoder Wasserstein generative adversarial network. *Med Image Anal.* 2019;55:165–180.
43. Sater SH, Sass AM, Rohr JJ, et al. Automated MRI-based quantification of posterior ocular globe flattening and recovery after long-duration spaceflight. *Eye (Lond)*. 2021;35:1869–1878.
44. Waisberg E, Ong J, Kamran SA, et al. Further characterizing the physiological process of posterior globe flattening in spaceflight associated neuro-ocular syndrome with generative adversarial networks. *J Appl Physiol.* 2023;134:150–151.
45. Ong J, Waisberg E, Masalkhi M, et al. Artificial intelligence frameworks to detect and investigate the pathophysiology of spaceflight associated neuro-ocular syndrome (SANS). *Brain Sci.* 2023;13. <https://doi.org/10.3390/brainsci13081148>.

46. Zwart SR, Gregory JF, Zeisel SH, et al. Genotype, B-vitamin status, and androgens affect spaceflight-induced ophthalmic changes. *FASEB J*. 2016;30:141–148.
47. Zwart SR, Smith SM. Genetics, vitamins, and spaceflight associated neuro-ocular syndrome. In: Lee AG, Ong J, eds. *Spaceflight Associated Neuro-Ocular Syndrome*. Academic Press; 2022:55–65.
48. García-Minguillán CJ, Fernández-Ballart JD, Ceruelo S, et al. Riboflavin status modifies the effects of methylenetetrahydrofolate reductase (MTHFR) and methionine synthase reductase (MTRR) polymorphisms on homocysteine. *Genes Nutr*. 2014;9. <https://doi.org/10.1007/s12263-014-0435-1>.
49. Smith SM, Zwart SR. Spaceflight-related ocular changes: the potential role of genetics, and the potential of B vitamins as a countermeasure. *Curr Opin Clin Nutr Metab Care*. 2018;21:481–488.
50. Scott JM, Tucker WJ, Martin D, et al. Association of exercise and swimming goggles with modulation of cerebro-ocular hemodynamics and pressures in a model of spaceflight-associated neuro-ocular syndrome. *JAMA Ophthalmol*. 2019;137:652–659.
51. Sung JK, Kassel R, Hargens A, Huang AS. *Potential countermeasures for Spaceflight Associated Neuro-Ocular Syndrome*. Spaceflight Assoc Neuro-Ocul Syndrome; 2022.
52. Hearon CM, Dias KA, Babu G, et al. Effect of nightly lower body negative pressure on choroid engorgement in a model of spaceflight-associated neuro-ocular syndrome: a Randomized Crossover Trial. *JAMA Ophthalmol*. 2022;140:59–65.
53. Charles JB, Lathers CM. Summary of lower body negative pressure experiments during space flight. *J Clin Pharmacol*. 1994;34:571–583.
54. Ong J, Lee AG, Moss HE. Head-down tilt bed rest studies as a terrestrial analog for spaceflight associated neuro-ocular syndrome. *Front Neurol*. 2021;12:648958. <https://doi.org/10.3389/fneur.2021.648958>.
55. Laurie SS, Greenwald SH, Marshall-Goebel K, et al. Optic disc edema and chorioretinal folds develop during strict 6 degrees head-down tilt bed rest with or without artificial gravity. *Physiol Rep*. 2021;9:e14977. <https://doi.org/10.14814/phy2.14977>.
56. Waisberg E, Ong J, Masalkhi M, et al. Anatomical considerations for reducing ocular emergencies during spaceflight. *Ir J Med Sci*. 2024;193:505–508.
57. Sarker P, Ong J, Zaman N, et al. Extended reality quantification of pupil reactivity as a non-invasive assessment for the pathogenesis of spaceflight associated neuro-ocular syndrome: a technology validation study for astronaut health. *Life Sci Space Res (Amst)*. 2023;38:79–86.
58. Waisberg E, Ong J, Paladugu P, et al. Advances in machine learning to detect preventable causes of blindness. *Eye*. 2023;37:2582–2583.
59. Waisberg E, Ong J, Kamran SA, et al. Transfer learning as an AI-based solution to address limited datasets in space medicine. *Life Sci Space Res (Amst)*. 2023;36:36–38.