

Image Styling Transfer

Presented by:
A J Shakthivelu (R21EK002)
Pavan S Raj (R21EK025)
Tarun P (R21ER075)

What is Image Style Transfer?

Style Transfer is the technique of blending style from one image into another image keeping its content intact. The only change is the style configurations of the image to give an artistic touch to your image.

The content image describes the layout or the sketch and Style being the painting or the colors. It is an application of Computer Vision related to image processing techniques and Deep Convolutional Neural Networks.



Why Deep Learning & CNN?

Automatic Feature Extraction:

Deep learning, especially CNNs, can automatically extract intricate features from images without the need for manual feature engineering. This capability is essential for capturing both content and style representations effectively in image style transfer tasks.

End-to-End Learning

Deep learning enables end-to-end learning, where the model learns to perform style transfer directly from raw image data. This holistic approach eliminates the need for intermediate processing steps, resulting in more streamlined and efficient workflows.

Transfer Learning

Deep learning facilitates transfer learning, where pre-trained models can be leveraged to bootstrap the style transfer process. By starting with models pre-trained on large-scale image datasets, such as ImageNet, practitioners can expedite training and improve performance on style transfer tasks with limited data.

Real-Time Processing

With advancements in hardware acceleration and optimization techniques, deep learning models, including CNNs, can achieve real-time performance for style transfer tasks. This capability enables interactive applications and seamless integration into production pipelines.

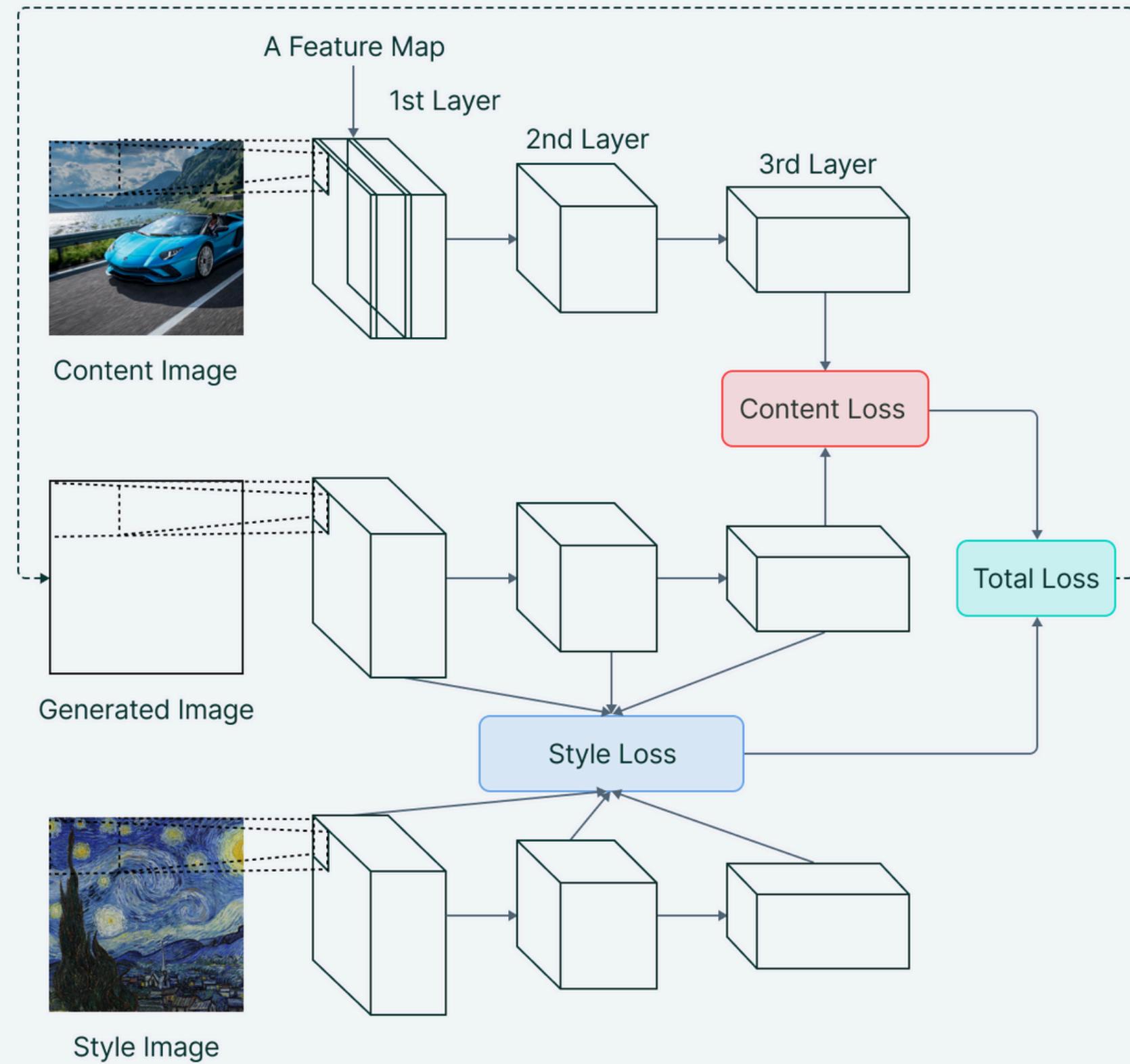
Neural Style Transfer

- Training a style transfer model requires two networks: a pre-trained feature extractor and a transfer network. NST uses a pre-trained model trained on ImageNet- VGG in TensorFlow.
- Images themselves make no sense to the model. These have to be converted into raw pixels and given to the model to transform it into a set of features, which is what Convolutional Neural Networks are responsible for.
- Thus, somewhere in between the layers, where the image is fed into the model, and the layer, which gives the output, the model serves as a complex feature extractor. All we need to leverage from the model is its intermediate layers, and then use them to describe the content and style of the input images.
- The input image is transformed into representations that have more information about the content of the image, rather than the detailed pixel value.

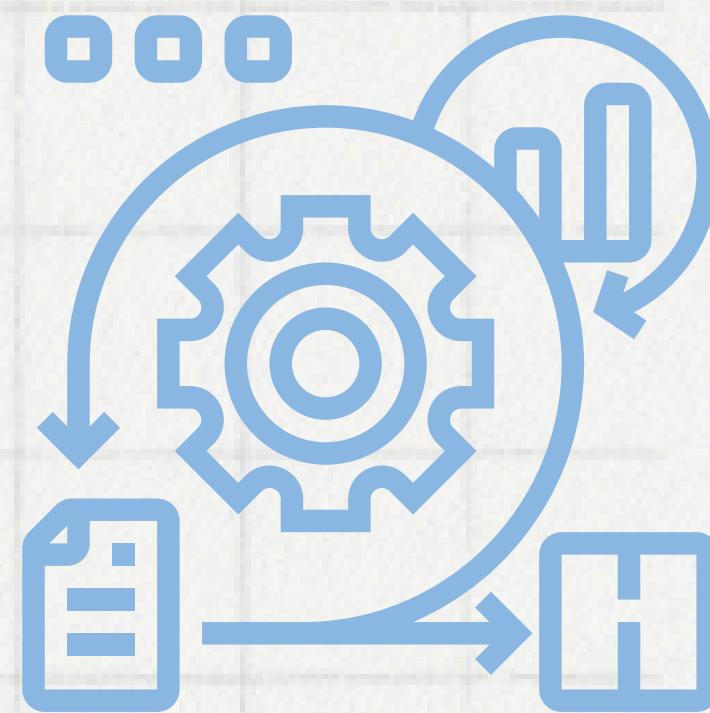


Neural Style Transfer Model Architecture

Optimize the Generated image w.r.t. the Total Loss



Methodology



01

Data collection & Pre-processing

Gather a diverse dataset of content and style images. The content images represent the target scenes or subjects, while the style images depict the desired artistic styles.

Preprocess the images to ensure consistency in size, aspect ratio, and color space. Common preprocessing steps may include resizing, normalization, and color adjustments.

02

Model & Architecture

We are using suitable deep learning architecture for image style transfer called **VGG**.

Design the architecture for the style transfer model, incorporating components for feature extraction, content representation, and style transfer operations.

Experiment with different hyperparameters to optimize performance and achieve desired style transfer results.

03

Training and Optimization

Define loss functions that capture both content and style similarities between the input and output images.

Content loss

It helps to establish similarities between the content image and the generated image.

Style loss

Gram matrix is a way to interpret style information in an image as it shows the overall distribution of features in a given layer. It is measured as the amount of correlation present between features maps in a given layer.

Set up the training pipeline, including data loading, model training, and evaluation.

04

Model evaluation & validation

The model underwent thorough evaluation on the validation set, using metrics like perceptual similarity and content/style reconstruction loss. Qualitative analysis included visual inspection and user feedback. The assessment validated the model's ability to generalize to new content and style combinations, showing promise for diverse real-world applications.

Tech Stack

- Python: language
- NumPy: library for numerical calculations
- Matplotlib: library for data visualisation
- Pytorch: a deep learning framework by Facebook AI Research Team for building neural networks
- torchvision: package consists of popular datasets, model architectures, and common image transformations for computer vision



CONTENT DATASET

The Content Data which is the [COCO Dataset](#) has been downloaded and extracted using wget command in the terminal. The script downloads the train2014.zip file from the official COCO website and saves it as coco.zip in the specified directory. It then extracts the contents of the zip file using the ZipFile function from the zipfile module and saves it in the content-data directory.

```
!wget --no-check-certificate \
"http://images.cocodataset.org/zips/train2014.zip" \
-O "/home/ec2-user/SageMaker/coco.zip"

local_zip = '/home/ec2-user/SageMaker/coco.zip'
zip_ref = zipfile.ZipFile(local_zip, 'r')
!mkdir /home/ec2-user/SageMaker/content-data
zip_ref.extractall('/home/ec2-user/SageMaker/content-
data')
zip_ref.close()
os.remove(local_zip)
print('The number of images present in COCO dataset
are:',len(os.listdir('/home/ec2-user/SageMaker/content-
data/train2014')))
```

STYLE DATASET

The Style Dataset has been downloaded and extracted from the Kaggle Competition [Painter by numbers](#). To ensure that the number of images in the style dataset matches the number of images in the content dataset, you can include the test dataset as well. By downloading the test dataset, you can randomly extract 3350 images from it to supplement the style dataset.

```
!ls -lha /home/ec2-user/SageMaker/kaggle.json
!pip install -q kaggle
!mkdir -p ~/.kaggle #Create the directory
!cp kaggle.json ~/.kaggle/
!chmod 600 /home/ec2-user/SageMaker/kaggle.json

!kaggle competitions download -f train.zip -p '/home/ec2-
user/SageMaker' -o painter-by-numbers
local_zip = '/home/ec2-user/SageMaker/train.zip'
zip_ref = zipfile.ZipFile(local_zip, 'r')
!mkdir /home/ec2-user/SageMaker/style-data
zip_ref.extractall('/home/ec2-user/SageMaker/style-data')
zip_ref.close()
os.remove(local_zip)
print('The number of images present in WikiArt dataset
are:',len(os.listdir('/home/ec2-user/SageMaker/train')))
```



Content



Style



$\alpha = 0.0$



$\alpha = 0.125$



$\alpha = 0.25$



$\alpha = 0.375$



$\alpha = 0.5$



$\alpha = 0.625$



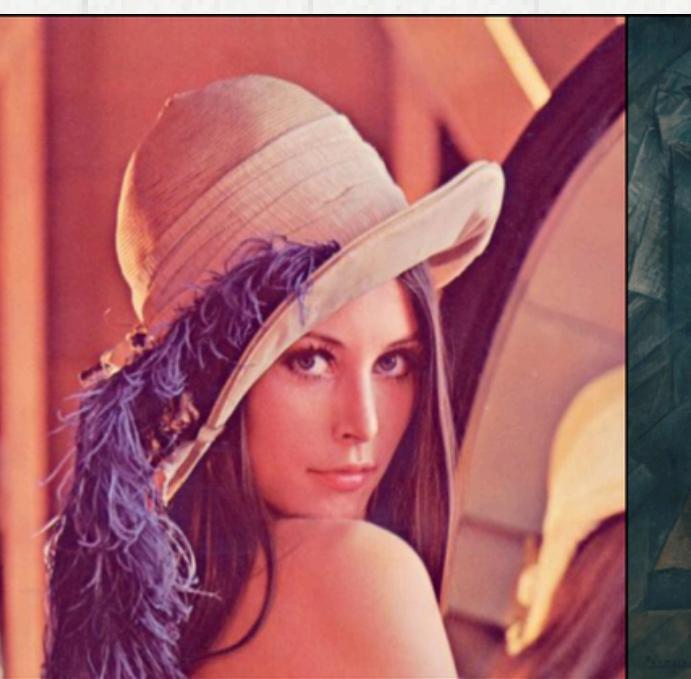
$\alpha = 0.75$



$\alpha = 0.875$



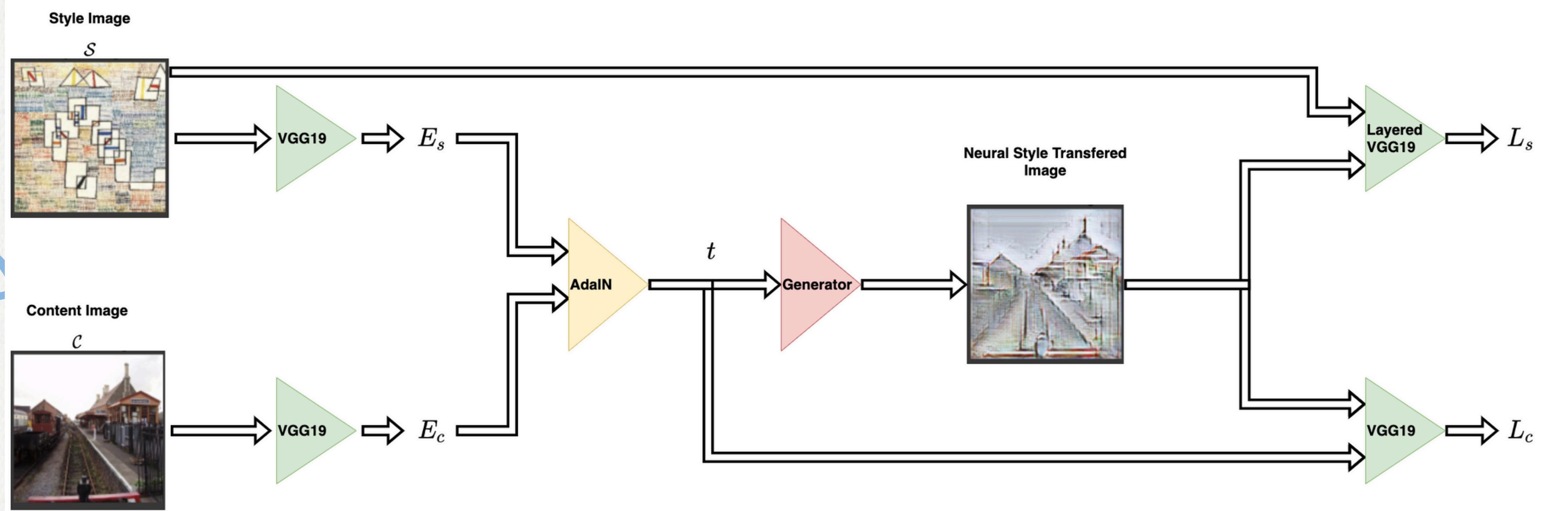
$\alpha = 1.0$



References

- <https://github.com/Surya-Prakash-Reddy/Image-Style-Transfer-Pytorch>
- <https://github.com/Moddy2024/AdaIN-Style-Transfer>
- <https://www.v7labs.com/blog/neural-style-transfer#h1>





Applications



Photo editors

Style transfer is extensively used in photo and video editing software.

These deep learning approaches and professional style transfer models can easily be applied to devices like mobile phones and give users a real-time ability to style images and videos.



Art and entertainment

Style transfer also provides new techniques that can change the way we look and deal with art.

It makes high-rated and over-priced artistic work reproducible for office and home decor, or for advertisements. Transfer models may help us commercialize art.



Gaming and Virtual reality

There are many cloud-powered video game streams that use image style transfer.

These models help the developers to provide interactive environments with customized artistic styles to users. This provides a 3D touch to the game and helps to enrich the artist inside every developer. Much similar to gaming, VR apps help to tell visual stories through their applications, games, films, and more.

