Report For Graduation Project

Created For Professor **Mahmoud Eid** By Team Members:

- Mohamed Hisham Abdelzaher
- Yousef Mohamed Abd El-Hay
- Abdelrahman Mostafa Mohamed
- Beshoy Madhat Sadeq
- Rina Khaled Ahmed
- Shereen Ashraf Ahmed

GitHub: https://github.com/MH0386/graduation_project

Email: alpha-AI@googlegroups.com

FPS Enhancer

Video frame rate enhancement by artificial intelligence is a technique that uses deep learning algorithms to generate new frames in between existing frames of a video, resulting in a higher frame rate and smoother motion. There are several approaches to video frame rate enhancement using AI, including video frame interpolation (VFI) and video super-resolution (VSR).

One recent study proposed a one-stage space-time video super-resolution framework that can directly reconstruct a high-resolution slow-motion video sequence from a low-resolution and low-frame-rate video. The proposed framework uses a feature temporal interpolation module to temporally interpolate LR frame features from the missing LR frames, capturing local temporal contexts. The results show that the proposed framework achieves better qualitative and quantitative performance on both clean and noisy LR frames and is several times faster than recent state-of-the-art two-stage networks.

Another study proposed an efficient face detection and recognition approach using machine learning and high-performance computing that can recognize faces with huge changes. The study used a convolutional neural network (CNN) and the recurrent neural network's subpart long-short-term memory technique to improve the reorganization procedure.

Video frame interpolation aims to synthesize nonexistent frames in between the original frames. While significant advances have been made in recent deep convolutional neural networks, the quality of interpolation is often reduced due to large object motion or occlusion. One study proposed a video frame interpolation method that explicitly detects occlusion by exploring depth information. The proposed model then warps the input frames, depth maps, and contextual features based on the optical flow and local interpolation kernels for synthesizing the output frame.

In summary, video frame rate enhancement by artificial intelligence is a promising technique that can generate new frames in between existing frames of a video, resulting in a higher frame rate and smoother motion. There are several approaches to video frame rate enhancement using AI, including video frame interpolation (VFI) and video super-resolution (VSR). These approaches use deep learning algorithms to improve the quality of video frames and can be applied to various applications, such as face detection and recognition and image and video quality enhancement.

It is difficult to rank the accuracy of each AI-based video frame interpolation model as it depends on several factors such as the complexity of the video, the amount of motion in the video, and the specific parameters used. However, here is a list of the models based on their popularity and general effectiveness:

- DVF (**D**eep **V**oxel **F**low)
- DAIN (**D**epth-**A**ware Video Frame **In**terpolation)
- RIFE (**R**eal-Time **I**ntermediate **F**low **E**stimation)
- Super SloMo
- RAFT (**R**ecurrent **A**ll-Pairs **F**ield **T**ransforms)
- TecoGAN (**Te**mporally **Co**herent **GAN**)
- EDVR (**E**nhanced **D**eep **V**ideo **R**estoration)
- CAIN (**C**ontent-**A**ware **In**painting **N**etwork)
- MEMC-Net (**M**otion **E**stimation and **M**otion **C**ompensation **Net**work)
- SepConv (**Sep**arable **Conv**olutional Networks)
- Softmax Splatting (**SS**)
- TOFlow (**T**ask-**O**riented **Flow**)

Comparisons

Table 2: **Quantitative comparisons on the UCF101, Vimeo90K, Middlebury-OTHER set, and HD benchmarks.** The images of each dataset are directly inputted to each model. Some models are unable to run on 1080p images due to exceeding the memory available on our graphics card (denoted as "OOM"). To report the runtime, we test all models for processing a pair of 640 × 480 images using the same device. **Bold** and underlined numbers represent the best and second-best performance. We use gray backgrounds to mark the methods that require pre-trained depth models or optical flow models.

Method	# Parameters (Million)	Runtime (ms)	UCF101 [41] PSNR	UCF101 [41] SSIM	Vimeo90K [48] PSNR	Vimeo90K [48] SSIM	M.B. [2] IE	HD [3] PSNR
DVF [48] (Vimeo90K)	1.6	80	34.92	0.968	34.56	0.973	2.47	31.47
Superslomo [17] (Vimeo90K)	19.8	62	35.15	0.968	34.64	0.974	2.21	31.55
SepConv [34]	21.6	51	34.78	0.967	33.79	0.970	2.27	30.87
TOFlow [2]	**1.1**	84	34.58	0.967	33.73	0.968	2.15	29.37
MEMC-Net [4]	70.3	401	35.01	0.968	34.29	0.970	2.12	31.39
DAIN [3]	24.0	436	35.00	0.968	34.71	0.976	2.04	31.64†
DSepConv [8]	21.8	236	35.08	0.969	34.73	0.974	2.03	OOM
CAIN [9]	42.8	38	34.98	0.969	34.65	0.973	2.28	31.77
SoftSplat [33]†	7.7	135	35.39	**0.970**	36.10	**0.980**	**1.81**	-
AdaCoF [21]	21.8	34	34.91	0.968	34.27	0.971	2.31	31.43
BMBC [36]	11.0	1580	35.15	0.969	35.01	0.976	2.04	OOM
CDFI [10]	5.0	198	35.21	0.969	35.17	0.977	1.98	OOM
EDSC [7]	8.9	46	35.13	0.968	34.84	0.975	2.02	31.59
RIFE	9.8	**16**	35.28	0.969	35.61	0.978	1.96	32.14
RIFE-Large (2T2R)	9.8	80	**35.41**	**0.970**	**36.13**	**0.980**	1.86	**32.32**

†: copy from the original papers.

As we can see, RIFE is relatively better than others. It has an excellent **P**eak **S**ignal-to-**N**oise **R**atio (PSNR) and **S**tructural **S**imilarity **I**ndex **M**easure (SSIM).

RIFE (Real-Time Intermediate Flow Estimation)

What is RIFE Video Frame Interpolation Accuracy?

RIFE (Real-time Intermediate Flow Estimation) is an intermediate flow estimation algorithm for Video Frame Interpolation (VFI). It uses a neural network named IFNet that can directly estimate the intermediate flows from coarse to fine with much better speed. RIFE training involves feeding two input frames into the IFNet to approximate intermediate flows and the fusion map. During the training phase, a privileged teacher refines the student's results based on ground truth to get more accurate approximations. Although the search results do not provide a specific metric for RIFE Video Frame Interpolation accuracy, they do mention that RIFE is a real-time algorithm that can process 720p videos at 30 frames per second. Additionally, RIFE is designed to directly estimate intermediate flows from coarse to fine, which can lead to better accuracy compared to other methods that first estimate bi-directional optical flows and then linearly combine them to approximate intermediate flows.

How does RIFE compare to other video frame interpolation algorithms in terms of speed?

RIFE (Real-Time Intermediate Flow Estimation) is a real-time intermediate flow estimation algorithm for video frame interpolation (VFI) that can process 720p videos at 30 frames per second. RIFE uses a neural network named IFNet that can directly estimate the intermediate flows from coarse to fine with much better speed compared to other methods that first estimate bi-directional optical flows and then linearly combine them to approximate intermediate flows. According to the search results, RIFE is 4–27 times faster than the popular SuperSlomo and DAIN methods and produces better results. Additionally, RIFE is the first flow-based and real-time VFI algorithm that processes 720p videos at 30 frames per second. Therefore, it can be concluded that RIFE is a fast and efficient algorithm for VFI compared to other methods.

Reference:

- RIFE Explained | Papers With Code
- RIFE AI interpolation - SmoothVideo Project (SVP)
- GitHub: megvii-research/ECCV2022-RIFE: ECCV2022 - Real-Time Intermediate Flow Estimation for Video Frame Interpolation
- RIFE, Real-Time Intermediate Flow Estimation for Video Frame Interpolation
- Real-Time Intermediate Flow Estimation for Video Frame Interpolation
- RIFE: Real-Time Intermediate Flow Estimation for Video Frame Interpolation | Papers With Code
- https://www.arxiv-vanity.com/papers/2011.06294
- RIFE AI interpolation - SmoothVideo Project (SVP)
- https://arxiv.org/pdf/2105.13482.pdf
- Video Frame Interpolation | Papers With Code
- Depth-Aware Video Frame Interpolation
- Transform Network Architectures for DL based End-to-End Image/Video Coding in Subsampled Color Spaces
- https://pubmed.ncbi.nlm.nih.gov/32225742
- IFRNet: Intermediate Feature Refine Network for Efficient Frame Interpolation
- https://pubmed.ncbi.nlm.nih.gov/34346905
- XVFI: eXtreme Video Frame Interpolation
- FILM: Frame Interpolation for Large Motion
- Asymmetric Bilateral Motion Estimation for Video Frame Interpolation
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9163436
- SuperFast: 200× Video Frame Interpolation via Event Camera
- Time Lens++: Event-based Frame Interpolation with Parametric Non-linear Flow and Multi-scale Fusion
- Multiple Video Frame Interpolation via Enhanced Deformable Separable Convolution
- TimeLens: Event-based Video Frame Interpolation
- Enhanced Motion-Compensated Video Coding With Deep Virtual Reference Frame Generation
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9480984
- Deep Learning for Content-based Personalized Viewport Prediction of 360-Degree VR Videos
- Many-to-many Splatting for Efficient Video Frame Interpolation
- PointINet: Point Cloud Frame Interpolation Network
- Learning Spatio-Temporal Downsampling for Effective Video Upscaling
- Unifying Motion Deblurring and Frame Interpolation with Events
- Power Efficient Video Super-Resolution on Mobile NPUs with DL, Mobile AI & AIM 2022 challenge
- ML Enhances Algorithms for Quantifying Non-Equilibrium Dynamics in Correlation Spectroscopy Experiments to Reach Frame-Rate-Limited Time Resolution
- Blurry Video Frame Interpolation
- Softmax Splatting for Video Frame Interpolation
- Inter-slice image augmentation based on frame interpolation for boosting medical image segmentation accuracy
- Artificial Intelligence based Video Codec (AIVC) for CLIC 2022
- Wyner-Ziv video coding with classified correlation noise estimation and key frame coding mode selection
- FISR: Deep Joint Frame Interpolation and Super-Resolution with a Multi-scale Temporal Loss

Data

- **Adobe240-FPS**: A dataset released by Adobe Systems that has high-frame-rate videos for training and evaluation
- **Middlebury**: The Middlebury dataset is a popular benchmark in computer vision, including video frame interpolation tasks. It provides a variety of high-quality video sequences with ground-truth frames.
- **Vimeo-90K**: A large-scale video dataset created by Vimeo with approximately 90,000 video clips with diverse content
- **UCF101**: The UCF101 dataset originally consisted of 101 action categories, but it can also be used for video frame interpolation tasks. It has a wide range of videos from various sources.

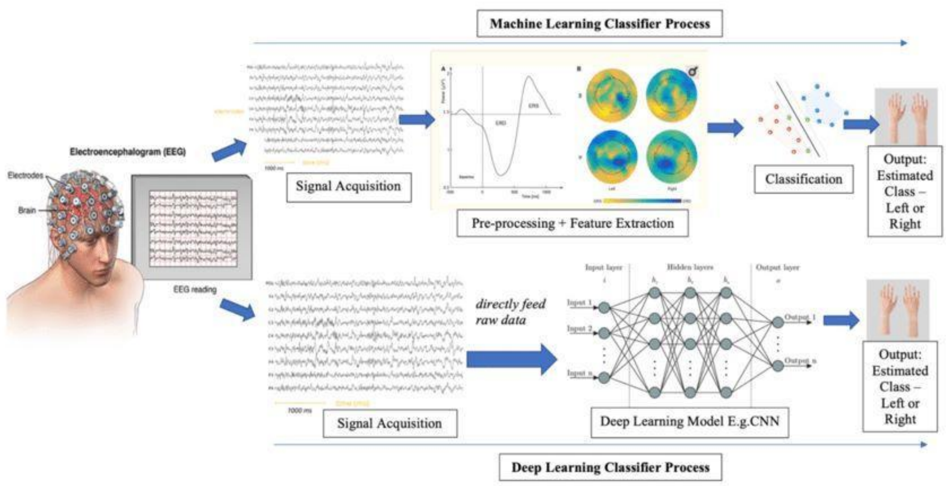
BCI with VR

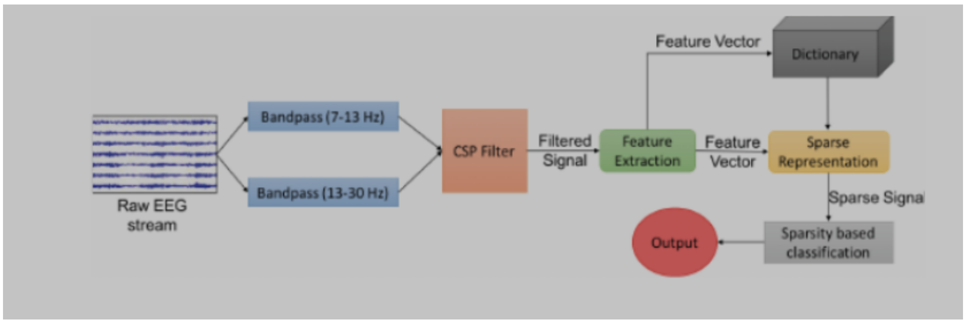
In recent years, there has been increasing interest in exploring potential applications of BCIs (Brain Computer Interfaces) combined with VR and AR in fields such as entertainment, education, healthcare, art, and design. Some examples include creating new forms of gameplay based on user intentions, enhancing learning outcomes through adaptive feedback loops, providing therapy for neurological conditions like stroke and spinal cord injuries, and developing creative tools for artists.

Challenges: Developing effective BCIs for use in VR and AR poses several technical challenges due to their limited resolution, bandwidth, and stability compared to traditional input methods like the keyboard, mouse, or joystick. Research efforts focus on optimizing signal quality, reducing latency, achieving important levels of precision and reliability, ensuring low energy consumption, minimizing invasiveness, mitigating sensory conflict between physical reality and VR illusions, and addressing ethical concerns related to privacy and consent.

Trends: Key trends in this emerging field involve advancing non-invasive EEG-based BCI solutions, developing machine learning algorithms for higher-level cognitive state detection, refining haptic feedback mechanisms to enhance the sense of presence, extending application domains beyond individual users toward collaborative group settings and social interactions, supporting accessibility for diverse populations, establishing industry standards, and fostering interdisciplinary collaboration between neuroscientists, engineers, designers, psychologists, clinicians, policymakers, and end-users

How will we use EEG signals?





How do I get the correct signal to the input, such as the signal responsible for tight arms or legs, and so on?

There are more works in the literature that have worked with EEG based on upper limb MI13–18 than on

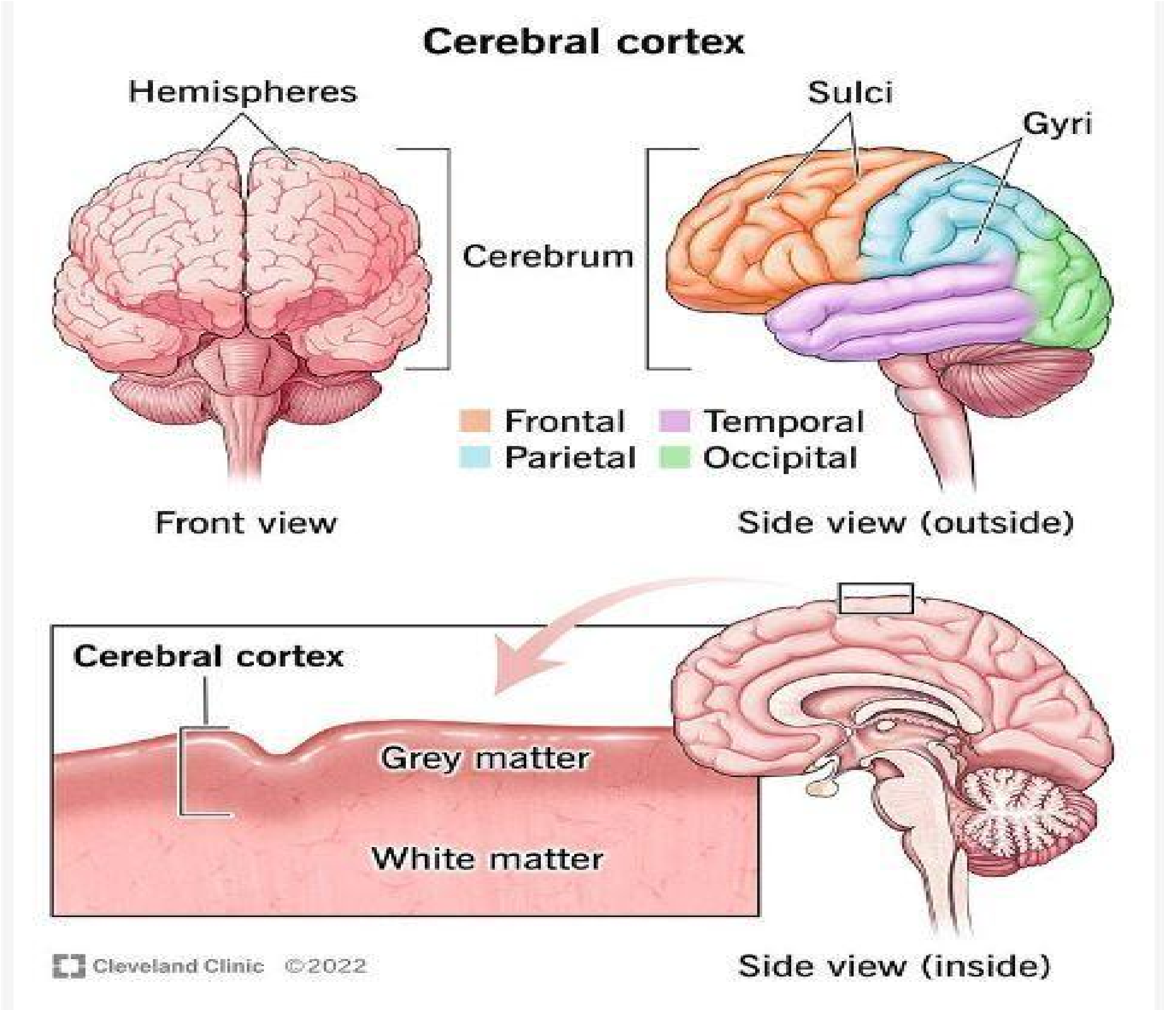
lower-limb MI. The main limitation of lower-limb MI is that the leg area of the motor cortex is

deeply located, around 1–4 cm from the surface. Therefore, EEG cannot accurately record this

activity, and it can be highly influenced by activity occurring near the surface of the skull. 19 Some authors have tried to discriminate leg MI from other types of MI that can be

more easily identified by EEG, such as left-hand, right-hand, and tongue MI, and the results obtained were promising. 20 However, considering a BCI for rehabilitation, it would be interesting to have a non-intention

condition in which subjects do not have to perform MI of another task to pause the system



What is the cerebral cortex?

Your cerebral cortex is the outermost layer of your brain. Its surface has many folds, giving it a wrinkled appearance. The folds consist of many deep grooves called sulci and raised areas called gyri. These folds add to the surface area of your cerebral cortex, allowing enormous amounts of information to be processed by more nerve cells. Your cerebral cortex makes up about half of your brain's total mass.

Your cerebral cortex consists of six layers of nerve cells that contain between 14 billion and 16 billion nerve cells. It is two millimeters (mm) to four mm (0.08 inches to 0.16 inches) thick.

Your cortex is divided into four lobes: frontal, parietal, temporal, and occipital. Each of these lobes is responsible for processing diverse types of information. Collectively, your cerebral cortex is responsible for the higher-level processes of the human brain, including language, memory, reasoning, thought, learning, decision-making, emotion, intelligence, and personality.

Why is the cerebral cortex called gray matter?

The gray matter found in the outer layer of your brain consists of nerve cell bodies, including the end portions of nerves called dendrites. Dendrites are the part of a nerve cell that receives a chemical message from another cell. Your cerebral cortex is gray because that section of the nerve lacks the fatty covering material called myelin.

The white matter in your brain is made up of bundles of axons, the long center section of a nerve cell that is wrapped in myelin. The myelin gives the tissue its whitish color.

What is the difference between the cerebral cortex and the cerebrum?

Your cerebral cortex is the outer layer that lies on top of your cerebrum. Your cerebrum is the largest area of your brain. Your cerebrum divides your brain into two halves called hemispheres. The hemispheres are attached by a bundle of nerve fibers called the corpus callosum. The corpus callosum allows your two hemispheres to communicate with each other.

What is the neocortex?

Most of your cerebral cortex is considered the neocortex. “Neo” means new. Your neocortex is so named because its appearance is thought to be new in vertebrate evolution. In humans, 90% of the cerebral cortex is the neocortex.

What are the functions of the cerebral cortex?

Your cerebral cortex is involved in many high-level functions, such as reasoning, emotion, thought, memory, language, and consciousness. Each lobe of your brain is associated with distinct functions.

Functions of the Frontal Lobe

Your frontal lobe is at the front of your brain, behind your forehead. Functions of your frontal lobe include:

- Decision-making, problem-solving
- Conscious thought
- Attention
- Emotional and behavioral control
- Speech production
- Personality
- Intelligence
- Body movement

Specific areas of note within this lobe are the motor cortex, the prefrontal cortex, and Broca's area. Your motor cortex is responsible for body movement. Your prefrontal cortex oversees "executive functions," such as thinking and problem-solving. It also supervises and directs other areas of your brain. Broca's area is a part of your frontal lobe that is involved with speech production.

Functions of the occipital lobe

Your occipital lobe is at the back of your brain. Functions of your occipital lobe include

- Visual processing and interpretation
- Visual data collection regarding color, motion, and orientation.
- Object and facial recognition.
- The depth and distance perception.
- Visual world mapping.

Functions of the parietal lobe

Your parietal lobe is between your frontal and occipital lobes and above your temporal lobe. Functions of your parietal lobe include

- Sensory information (touch, pressure, pain, position, vibration, temperature) processing.
- Spatial processing and spatial manipulation. This is the ability to understand where you are in three-dimensional space, such as how to navigate around your home or town.

Specific areas of note within this lobe are the somatosensory cortex. It receives sensory information ("feeling" information) from all over your body. Here is an example of how brain lobes work together:

> The motor cortex in your brain's frontal lobe sends the message that directs the muscles in your arm and hand to reach out toward a cup of soup on your kitchen table. The somatosensory cortex of your parietal lobe assesses the information

delivered through your touch of the cup, including the judgment of its temperature. Spatial processing in your parietal lobe allows you to grasp the cup, flawlessly navigating the hand-to-cup distance relative to the table and other objects.

Functions of the Temporal Lobe

Your temporal lobe is between your frontal and occipital lobes and below your parietal lobe. Functions of your temporal lobe include

- Language comprehension, speech formation, and learning.
- Memory.
- Hearing.
- Nonverbal interpretation.
- Sound-to-visual image conversion.

A specific area of note within this lobe is Wernicke's area. This area was more recently discovered to be involved in language comprehension based on speech tones and sounds, linking them to previously learned sounds.

What are the areas of the cerebral cortex?

Some researchers look at the brain in another way and classify the areas of the cerebral cortex by their three main types of functions: sensory, motor, and association areas.

Sensory areas: These areas of your cerebral cortex receive sensory information from your senses and your environment. Functions include:

- Making sense of visual information and object recognition. An area of your occipital lobe processes these functions called the visual cortex.
- Assessing touch, temperature, position, vibration, pressure, and pain information from your body. These functions are processed by an area of your parietal lobe called the somatosensory cortex.
- Processing hearing information. This function is processed by an area of your temporal lobe called the auditory cortex.
- Processing taste and flavor. An area of your frontal lobe processes these functions called the gustatory cortex.

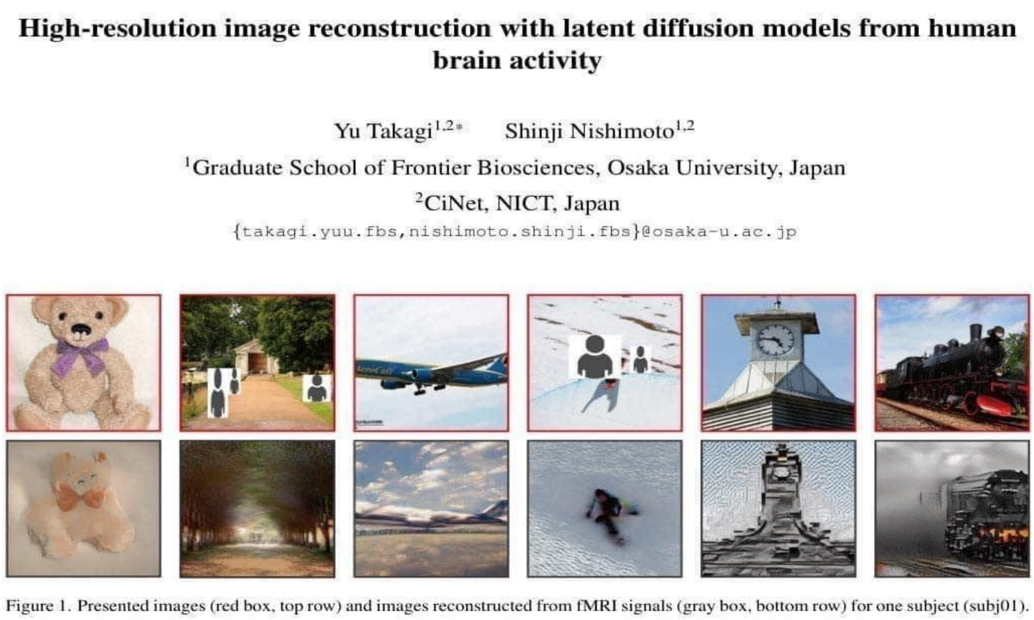
Motor areas: These areas of your cerebral cortex are involved in voluntary muscle movement. Your frontal lobe processes these functions. Functions include:

- Coordination of muscle movement.
- Planning of complex movements.

- Learning through imitation and empathy.

There are other Ideas in BCI created by (open Ai)

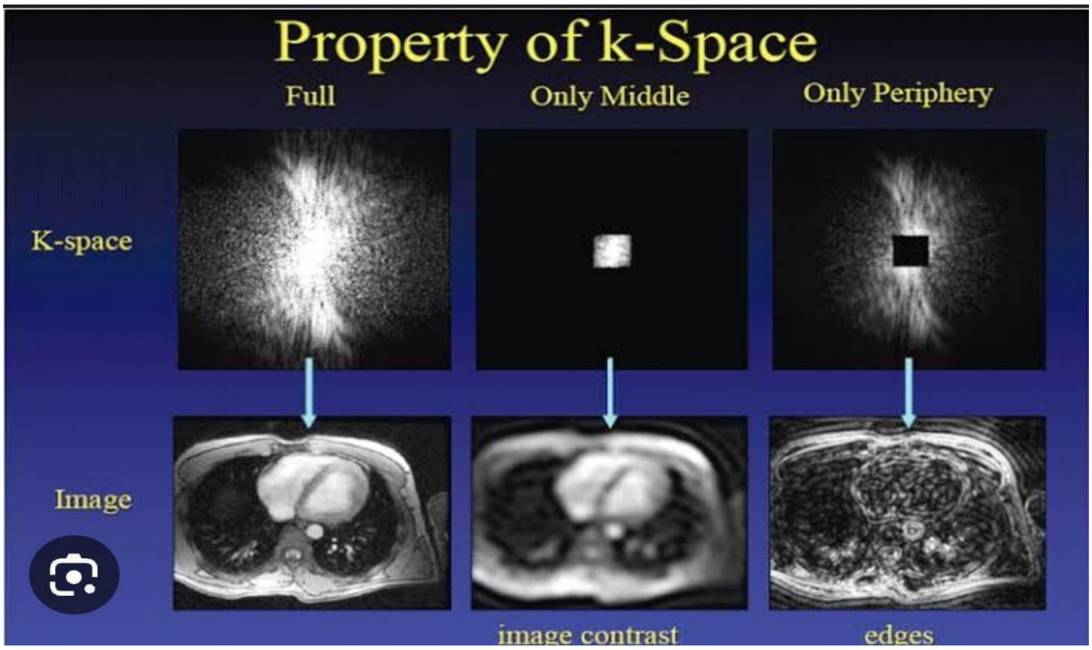
That could take an MRI signal and convert it to the related image that the person thinking about it



High-resolution image reconstruction with latent diffusion models from human brain activity

Yu Takagi[1,2]* Shinji Nishimoto[1,2]
[1]Graduate School of Frontier Biosciences, Osaka University, Japan
[2]CiNet, NICT, Japan
{takagi.yuu.fbs,nishimoto.shinji.fbs}@osaka-u.ac.jp

Figure 1. Presented images (red box, top row) and images reconstructed from fMRI signals (gray box, bottom row) for one subject (subj01).



Limitations of Brain-Computer Interfaces in Virtual Reality

While Brain-Computer Interface (BCI) technology combined with Virtual Reality (VR) has shown promising results in various applications, there are also some limitations to consider. Here are some aspects to keep in mind:

Accuracy

One of the main limitations of BCI with VR is the accuracy of the technology. The accuracy of BCI systems can be affected by various factors, such as the user's mental state, the quality of the EEG signal, and the complexity of the task being performed[1]. Inaccurate readings can result in incorrect actions being taken in the virtual environment, which can be frustrating for the user and limit the effectiveness of the system.

Cost

Another limitation of BCI with VR is the cost of the technology. BCI systems can be expensive to develop and require specialized equipment, such as EEG sensors and signal processing software[1]. Additionally, VR systems can also be costly, especially if high-end equipment is required for a realistic experience. These costs can limit the accessibility of BCI with VR technology to certain populations.

User Training

BCI with VR technology requires users to undergo training to learn how to use the system effectively. This training can be time-consuming and may require the assistance of a trained professional[1]. Additionally, some users may not be able to use BCI with VR technology due to cognitive or physical limitations.

Overall, while BCI with VR has shown potential in various applications, it is important to consider the limitations of the technology, such as accuracy, cost, and user training. These factors can affect the effectiveness and accessibility of BCI with VR technology.

Unfortunately, the search results did not provide much information on the limitations of BCI with VR technology.

Resources

- https://www.researchgate.net/publication/370125547_BRAIN-COMPUTER_INTERFACE_ENHANCED_BY_VIRTUAL_REALITY_TRAINING_FOR_CONTROLLING_A_LOWER-LIMB_EXOSKELETON
- https://drive.google.com/file/d/1atgC-sBJyqpJB3HnfObvfGKOl85hEA7C/view?usp=sharing