# VisionExtract: Isolation from Images using Image Segmentation

**Project Statement:**

The goal of this project is to build a machine learning model capable of automatically extracting the main subject from an image. For any given input picture, the model should output a new image in which only the subject is visible and everything else is rendered completely black. This subject isolation process can be used for automation in photography, digital art, augmented reality, virtual conferencing, and background replacement applications.

**Use Cases:** The project addresses one primary use case:

**Automated Subject Isolation:**

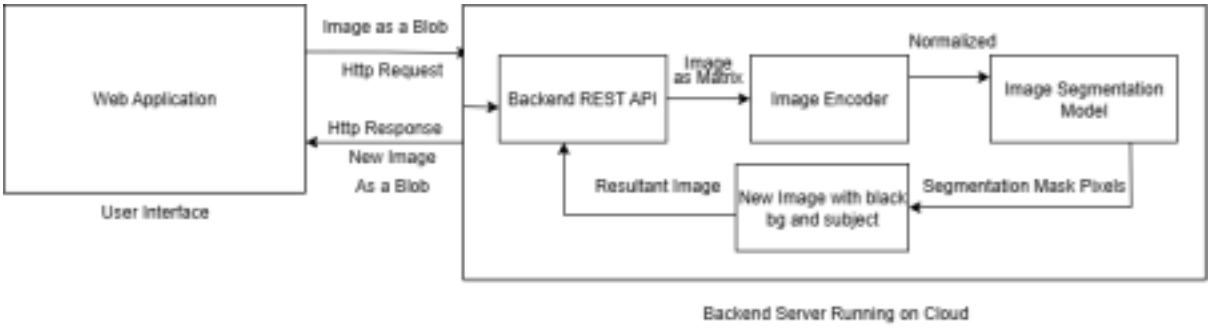
- Description: Automatically detect and extract the main subject from any image. The output will be an image where only the subject is displayed as in the original photo, while the rest of the pixels are set to black. This replicates the "cutout" functionality required in many media editing pipelines.

**Outcomes:**
By the end of this project, students will:

- Understand the principles of semantic segmentation using deep learning models for pixel-wise image tasks.
- Learn and implement data preprocessing strategies for segmentation, including managing mask annotations and input normalization.
- Train, validate, and evaluate an image segmentation model for the specific application of subject isolation.
- Deploy and demonstrate a solution that processes user-uploaded images and returns the desired isolated-subject results.
- Prepare detailed documentation and a presentation on their data pipeline, modelling approaches, and results.

**Dataset: https://www.kaggle.com/datasets/awsaf49/coco-2017-dataset**

**Model Architecture:**



**Modules to be implemented**

**1. Data preprocessing and feature engineering.**

- Raw image and mask data often require significant preparation.
- Important steps include handling image and mask resolution consistency, data augmentation (cropping, flipping, colour perturbation), normalization, and converting masks to binary subject-background format.
- The data pipeline must ensure each input image aligns correctly with its mask and that the resultant mask strictly delineates the subject for accurate training.

**2. Building the Segmentation model**

- The segmentation model will predict a pixel-wise binary mask for each input image, separating the subject from the background.
- During inference, this mask will be used to generate the output image: input pixels are retained for the subject and replaced with black for all background pixels.

**3. Evaluation and Fine-tuning**

- Model performance will be quantitatively evaluated using metrics such as Intersection over Union (IoU), Dice coefficient, precision, recall, and pixel-wise accuracy.
- Qualitative sample output images should also be reviewed to assess visual fidelity of the subject isolation.
- Fine-tuning may include experimenting with hyperparameters or data augmentation strategies to improve accuracy and generalization.

**4. Documentation and Presentation Preparation**

- Document every phase, including dataset handling, pipeline construction, experimental results, and sample input/output images.
- Prepare visuals showcasing before-and-after effects of subject isolation and explain key technical challenges and solutions.

**Week-wise module implementation and high-level requirements with output screenshots**

**Milestone 1:**
**Week 1: Project Initialization and Dataset Acquisition.**

- Define project objectives and expected results.
- Acquire and inspect an appropriate image segmentation dataset with subject mask annotations.
- Explore dataset structure, view example images and masks.

**Week 2: Data Preprocessing and Validation.**

- Develop a pipeline for data preprocessing (resize, normalization, augmentation). • Ensure correct alignment of images and annotated masks.

• Convert multi-class masks (if present) into binary masks identifying only the main subject.

**Milestone 2:**

**Week 3: Initial Model Training.**

- Implement the initial image segmentation network.
- Train the model on the processed data and monitor basic metrics.
- Visualize early predictions to verify subject extraction is proceeding as expected.

**Week 4: Predictions and Fine-tuning.**

- Generate predictions on a validation set.
- Examine output images and compare with ground-truth subject masks.
- Adjust model hyperparameters and perform data augmentation to improve results.

**Milestone 3:**

**Week 5: Improve Data and Experiment with Architectures.**

• Refine preprocessing, explore further augmentations or post-processing for cleaner masks. • Explore additional model architectures or training strategies for improved subject separation. • Compare variant performance and document findings.

**Week 6: Inference**

- Deploy the trained model for inference on new, unseen images.
- Automate the pipeline: for uploaded input, generate and save the subject-isolated output image.
- Test with images outside the original dataset to check robustness.

**Milestone 4:**

**Week 7: Full Pipeline and User Interface**

• Build a basic web application where users can upload an image and receive the result. • Integrate preprocessing, model inference, and output generation into a streamlined pipeline.

**Week 8: Documentation, Presentation, and Demo**

- Compile all technical documentation, including pipeline details, experimental results, and lessons learned.
- Prepare a presentation with before-and-after visual results demonstrating the effectiveness of the model.
- Finalize and conduct a live demonstration, if required.

Evaluation Criteria:

1. **Completion of Milestones:** Assess whether each major project phase was reached within its timeline, including successful data handling, model development, inference, and UI integration (if applicable).
2. **Accuracy of Subject Isolation:** Compare output images to ground-truth masks and visually inspect qualitative quality of separation. Use quantitative segmentation metrics (e.g., IoU, Dice score) for thorough evaluation.
3. **Clarity and Depth of Documentation and Presentation:** Review the completeness and clarity

of technical documentation; assess the ability to effectively communicate process and results during the final presentation and demo, including before/after image examples.

**Model Performance - Quantitative Metrics:**

- **Subject Segmentation**
  - Metric: Intersection over Union (IoU)
  - Description: Measures the overlap between the predicted subject region and the ground-truth mask divided by their union. High IoU indicates better segmentation accuracy.

**Additional Metrics:**

- **Metric:** Dice Coefficient, Pixel-wise Accuracy
- **Description:** Measure the alignment and quality of predicted masks, especially useful for binary separation tasks.

**Example Quantitative Metrics for Evaluation**

**1. Subject Isolation Performance**

- Goal: Achieve high IoU and Dice scores on validation images, indicating the model reliably separates the subject from the background.