**Report Summary:**

by:

* Eden moshe 207730391
* Yaniv Shtein 207656356
* Iris Kanter 211566484
* Bar Elrom 209121573

**High-resolution Iterative Feedback Network (HitNet) for Camouflaged Object Detection:**

**Introduction:**  
Camouflaged Object Detection (COD) involves identifying objects that are visually indistinguishable from their background a challenging task for both humans and machines. Traditional CNN-based models often downsample images, losing crucial edge details necessary for detecting such objects. The paper proposes HitNet, a high-resolution iterative feedback network that captures and refines fine details using transformer-based multi-scale features.

**Project Study Process:**The study began by identifying the challenge of detecting camouflaged objects due to their similarity to the background, which traditional models struggle to segment accurately. The researchers then developed HitNet, a model that iteratively refines low-resolution features using high-resolution information through a feedback mechanism. They designed and implemented three core modules—transformer-based feature extraction, multi-resolution refinement, and iteration-supervised feedback—and trained the model on four benchmark datasets. Extensive experiments, ablation studies, and visual analyses were conducted to validate the model’s effectiveness. Finally, they introduced a cross-domain learning approach to augment training data by generating synthetic camouflaged images from salient ones.

**Project Flow and Methodology:**

The architecture of HitNet is based on three core components:

* Transformer-based Feature Extraction (TFE):  
  Uses Pyramid Vision Transformer (PVT) to extract high-resolution multi-scale features from input images. PVT reduces memory cost while maintaining high performance by applying spatial-reduction attention and progressive shrinking of the image dimensions.
* Multi-resolution Iterative Refinement (RIR):  
  This module refines low-resolution features by incorporating high-resolution feedback through multiple iterations. It ensures that finer object boundaries and textures are preserved and corrected with each iteration.
* Iteration Feature Feedback (IFF):  
  A feedback mechanism imposes segmentation loss at every iteration, encouraging progressive improvement. A weighted loss function ensures deeper iterations have more impact.

The model is trained in a supervised learning setting using four standard COD datasets: COD10K, CHAMELEON, CAMO, and NC4K.

Additionally, a Cross-Domain Learning (CDL) technique is applied to generate synthetic camouflaged training data by converting salient object images, thereby improving training diversity and performance.

**Results:**

The model is benchmarked against 29 state-of-the-art models using standard metrics:

* Weighted F-measure (Fwβ)
* Structure-measure (Sα)
* E-measure (Eφ)
* Mean Absolute Error (MAE)

On the challenging COD10K dataset:

* Fwβ = 0.798, MAE = 0.024
* This outperformed the previous best model by 16.5% in Fwβ and 31.4% in MAE.

The model also showed strong qualitative results in preserving detailed structures like thin antennae, leaf edges, and partially occluded objects.

**Conclusions:**

HitNet introduces a robust architecture that bridges high-resolution feature extraction and iterative refinement, resulting in superior detection of camouflaged objects. Its ability to retain fine details while operating efficiently in real time (up to 39 fps) makes it ideal for real-world applications such as wildlife monitoring, medical imaging, and rescue operations.

Moreover, the Cross-Domain Learning (CDL) pipeline for data augmentation is a valuable contribution that allows broader application and improvement of COD models. Despite a few failure cases involving highly complex structures, HitNet sets a new performance benchmark in the field.