Exploring the Effects of Perturbed and Adversarial Dreaming on Learning Cortical Representations

A Replication Study of 'Learning Cortical Representations Through Perturbed and Adversarial Dreaming'

Thamires de Souza Oliveira
Knowledge Discovery course

Professor: Concetto Spampinato

Article main goals and proposals

- **Proposed Functional Model:** This project introduces a novel functional model of cortical representation learning, suggesting that dreams, particularly their creative development of episodic memories, play a crucial role in the formation of semantic representations during evolution.

- **Inspired by Generative Adversarial Networks (GANs):** The researchers present a cortical architecture inspired by generative adversarial networks (GANs) to simulate the learning process. By leveraging established datasets of natural images, the model is trained to evaluate the quality of the acquired representations.

- **Insights into Brain Processing and Representation:** Through this research, valuable insights into the processing and representation of sensory experiences in the brain are uncovered. These findings have potential implications for understanding how humans and other organisms learn from sensory input, shedding light on the underlying mechanisms of learning and cognition.

The article can be found here: https://elifesciences.org/articles/76384

Visual Pathway: Encoder, Generator and Brain States

1. **Visual Pathway:** The brain's visual pathway consists of different levels or layers through which information flows in a step-by-step manner.
2. **Encoder:** The encoder transforms signals from lower levels to compressed representations ("z") in higher levels, providing simplified summaries of information.
3. **Generator/Feedback Pathway:** This pathway allows information to flow bidirectionally, from higher levels back down to lower levels, facilitating feedback and interaction.
4. **Hippocampal Module:** The hippocampal module stores and replays the compressed representations ("z") to support the encoding and decoding processes.
5. **Brain States:** Wakefulness, non-REM sleep, and REM sleep are three brain states of interest for studying the visual pathway's functional roles.

Main differences between kinds of sleep

REM

Characterized by adversarial dreaming, where the brain generates realistic sensory experiences by combining memories and organizes representations based on object semantics.

Kinds of sleep

NREM

Involves perturbed dreaming, allowing the brain to consolidate and refine learned representations by abstracting away unnecessary details.

Network Structure

Feature 1

Four convolutional layers: 64, 128, 256, and 256 channels.

Feature 2

Utilizes a 4x4 kernel, stride of 2, and LeakyReLU nonlinearity.

ENCODER

Feature 3

Feature size is reduced by half in each layer.

Feature 4

Output of the last layer is denoted as "z."

Network Structure

Feature 1

Includes an additional convolutional layer followed by a sigmoid nonlinearity.

Feature 2

Maps the second-to-last layer to a scalar value "d" for internal/external discrimination.

DISCRIMINATOR

Feature 3

The mapping from input data "x" to "d" is referred to as Ed.

Feature 4

The first three convolutional layers are shared between Ez and Ed.

Network Structure

Feature 1

Takes data from the latent space and maps it back to the pixel space.

Feature 2

Consists of four deconvolutional layers with different numbers of channels: 256, 128, 64, and 3.

GENERATOR

Feature 3

Utilizes a 4x4 kernel, a stride of 2, and either a LeakyReLU or tanh nonlinearity.

Feature 4

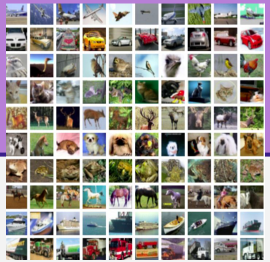
Feature size is doubled in each layer.

Feature 5

The output of the generator represents the reconstructed pixel data.

The Datasets

60.000 images



CIFAR10

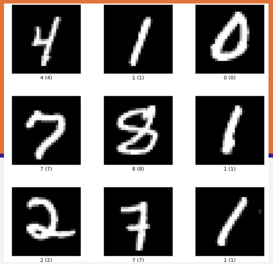
Range of objects, including airplanes, automobiles, birds, cats, deer, dogs, frogs, horses, ships, and trucks.

73.257 images

SVHN

Real-world images of house numbers extracted from Google Street View images.

70.000 images



MNIST

Collection of handwritten digit images sourced from multiple contributors.

70.000 images

FASHION

Images of various fashion items, such as T-shirts, trousers, pullovers and so on.

Training

Optimization algorithm

ADAM Optimizer

Learning rate

0.0002

Mini-batch

size of 64

Loss functions

Condition-specific loss functions

differentiability

Gradients can be computed for all parameters using backpropagation, facilitating efficient optimization of the model's parameters

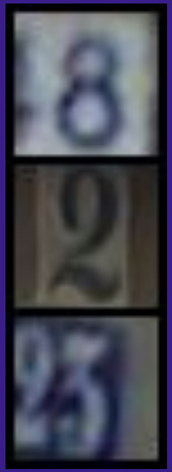
Samples

- The **Generative Adversarial Network** (GAN) model trains a generator to create synthetic images by inputting random noise and optimizing its output to resemble real training data.

- Simultaneously, a discriminator is trained to distinguish between **real and generated images,** providing feedback to the generator, leading to the generation of increasingly realistic and diverse synthetic images.

Sample 1 - svnh

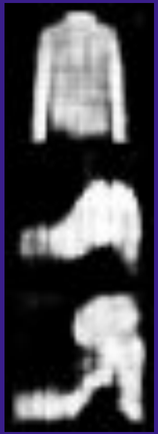
Data during wakeful state



Samples

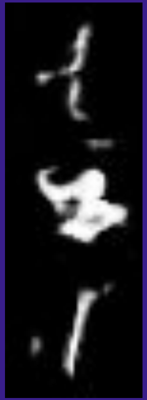
Sample 2 - fashion

Data during NREM state



Sample 3 - mnist

Data during REM state



Evaluation Procedure

Linear Classifier

Trained a linear classifier using latent features (Z) extracted from training images, with a weight matrix (W) projecting features onto label neurons for predictions.

Losses

Used multiclass cross-entropy loss to measure the difference between predicted class probabilities and target class, guiding classifier training.

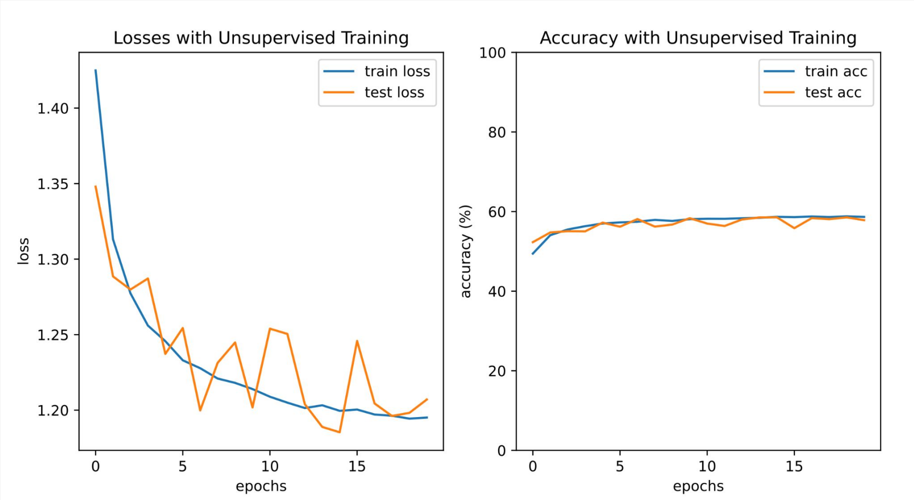
Occluded data

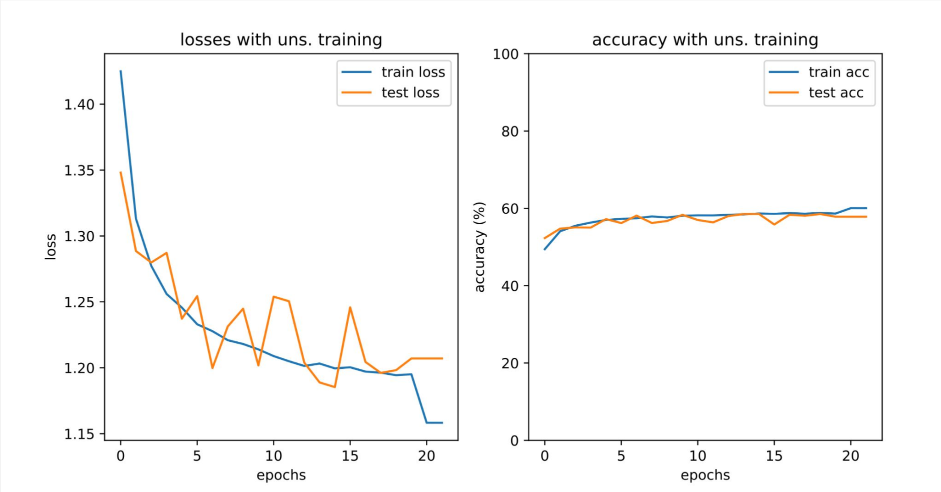
Model performance on occluded data by applying random square occlusion masks to test samples at a fixed size of 4, reporting results for different occlusion probabilities.

Confusion matrix

Evaluated classifier's performance using the confusion matrix, comparing true and predicted labels to gain insights into classification accuracy.

Results - CIFAR10

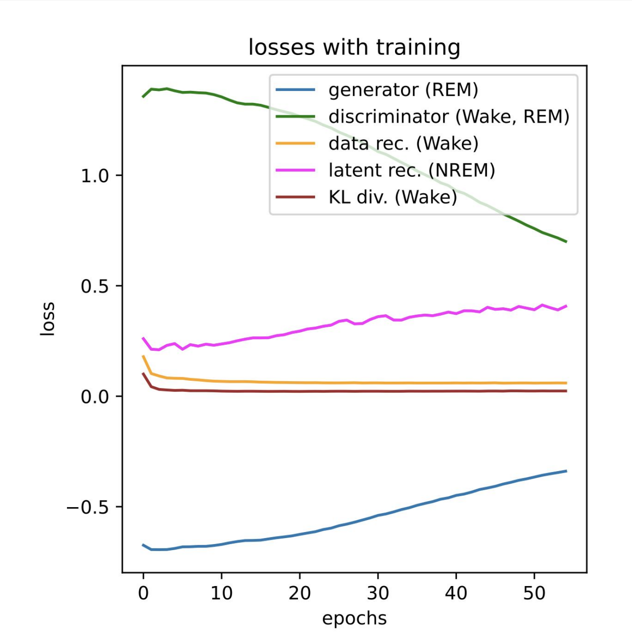


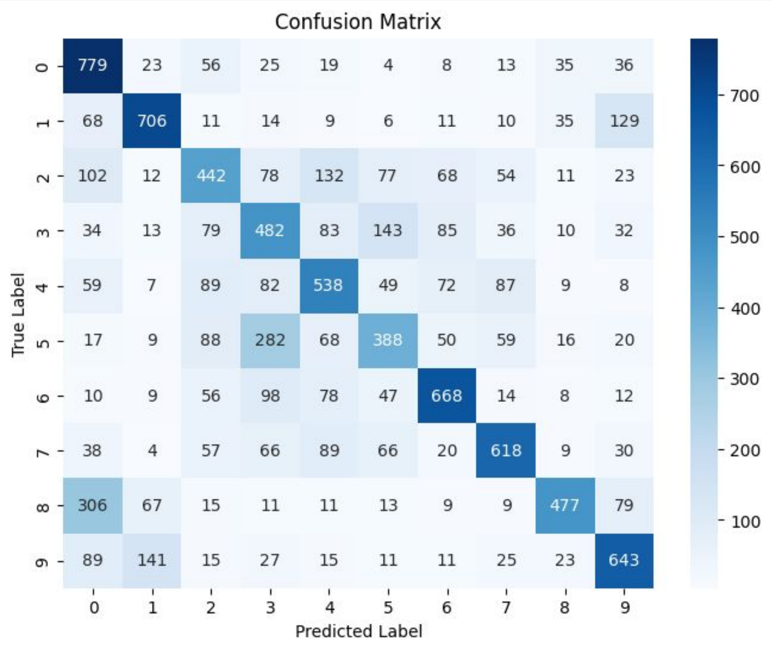


Without occlusions

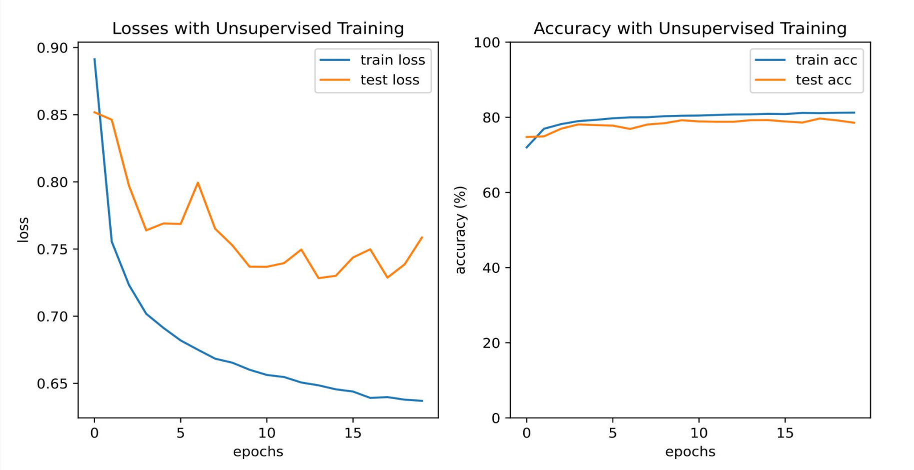
With occlusions

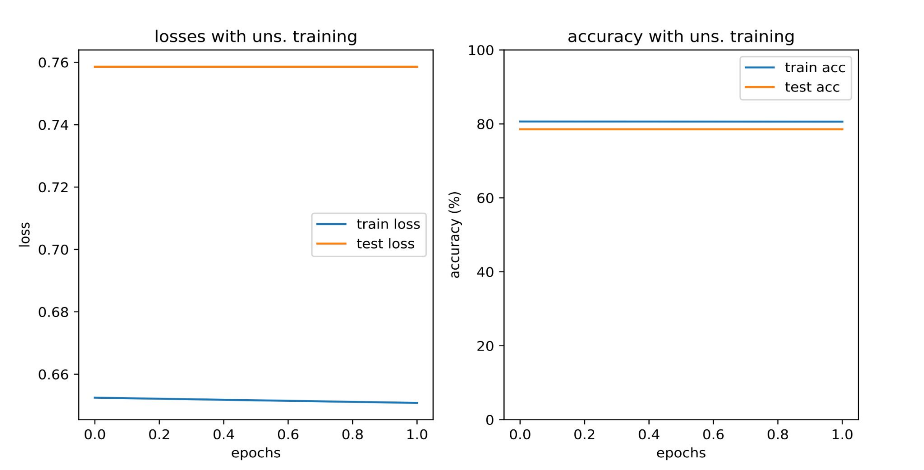
Results - CIFAR10





Results - SVHN

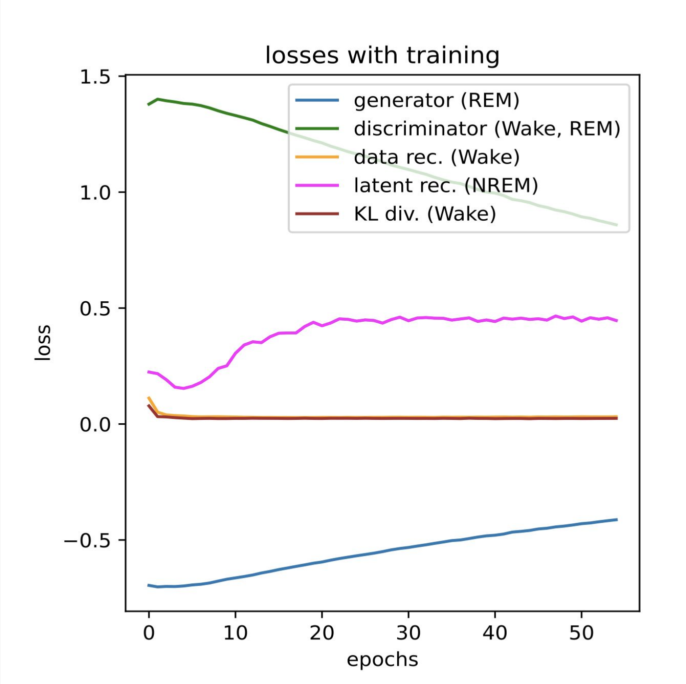


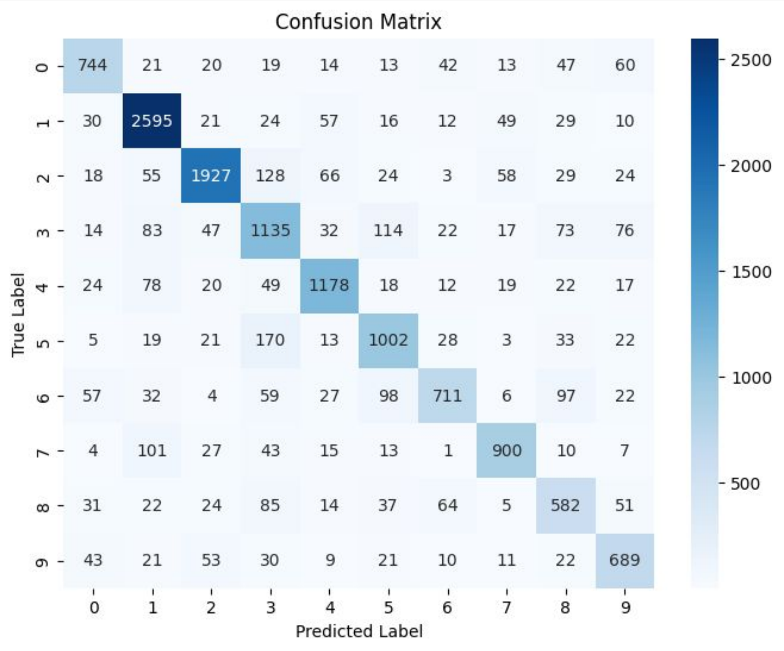


Without occlusions

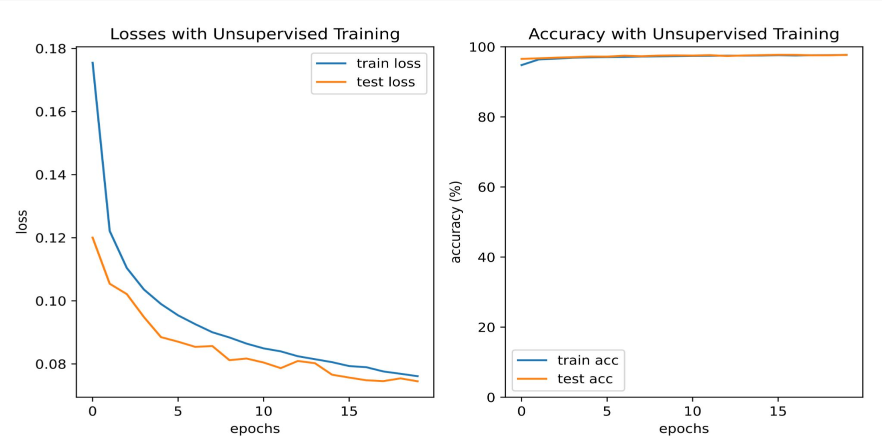
With occlusions

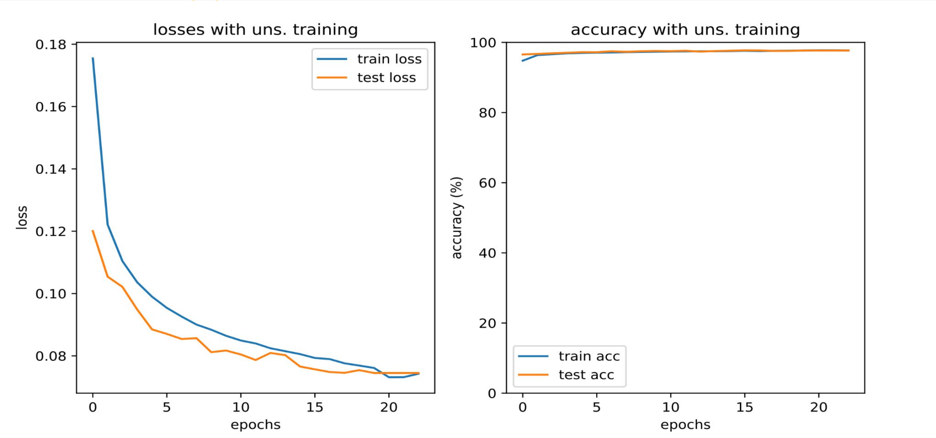
Results - SVHN





Results - mnist

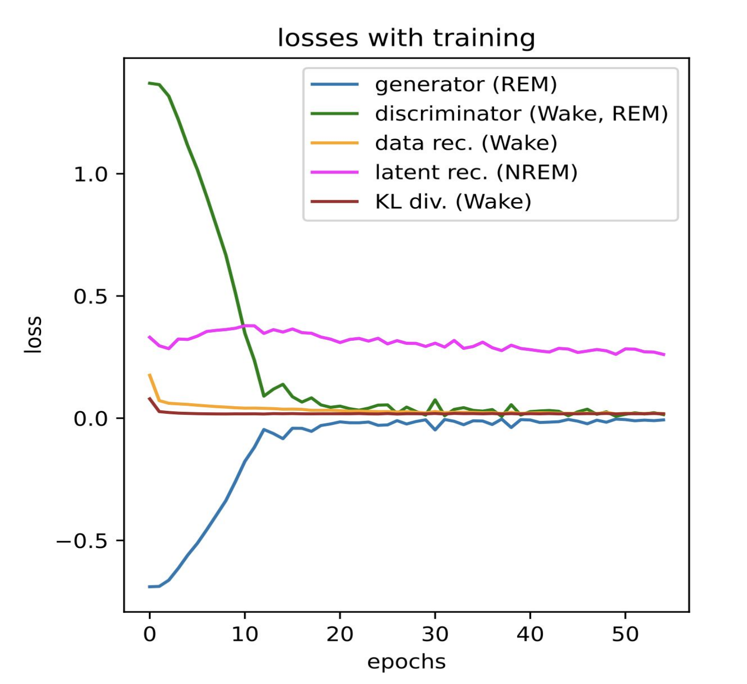


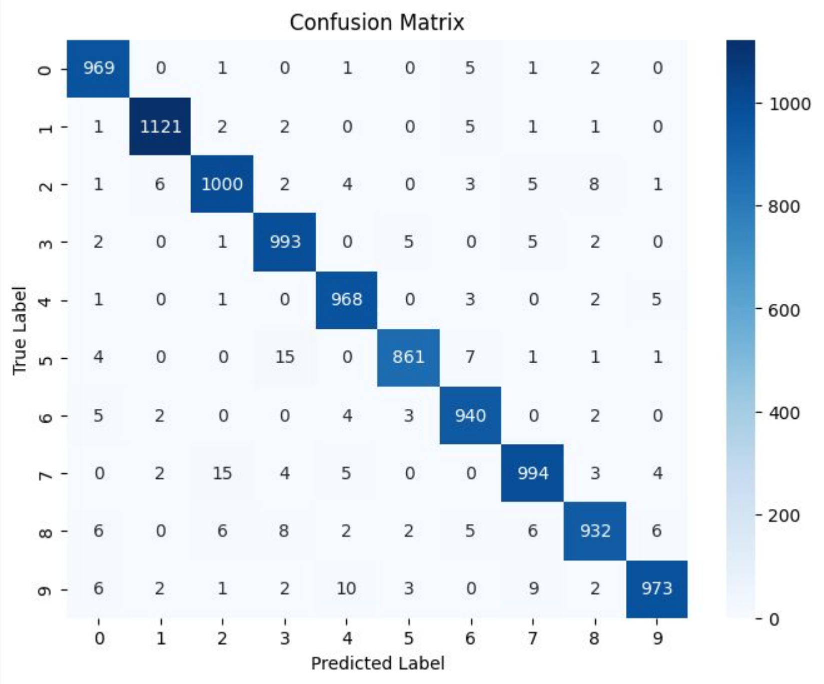


Without occlusions

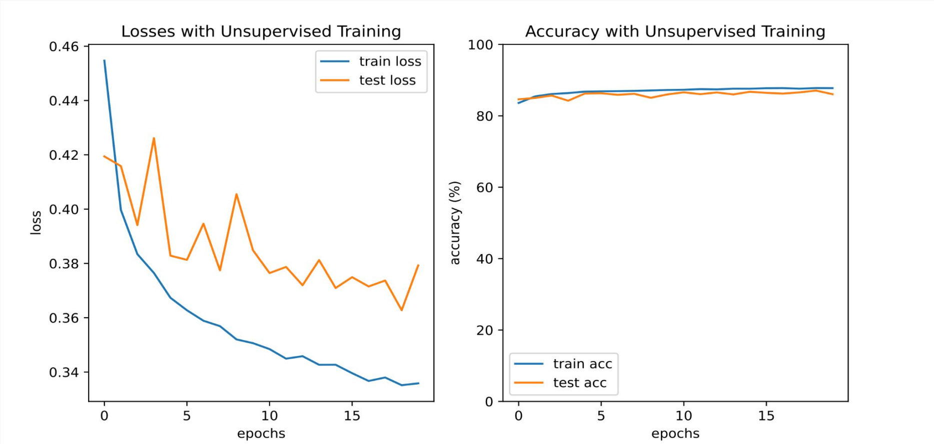
With occlusions

Results - mnist

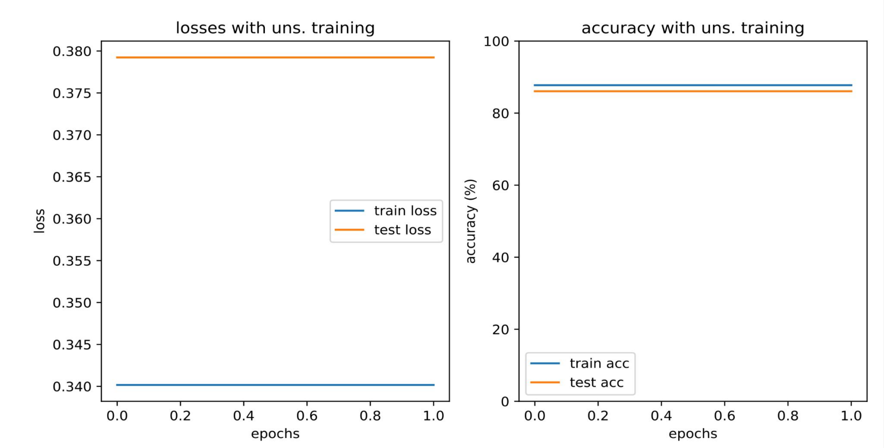




Results - Fashion

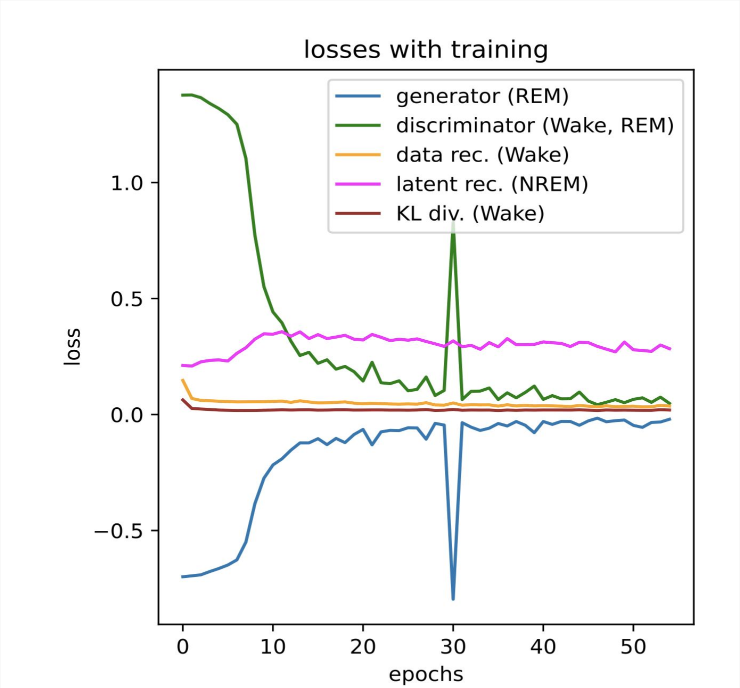


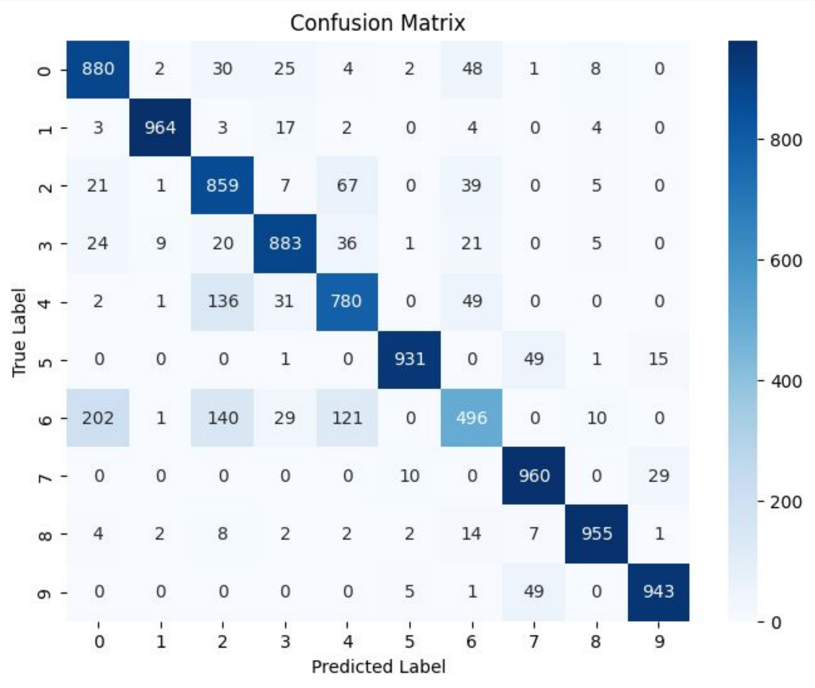
Without occlusions



With occlusions

Results - Fashion





Conclusions

Dreams and episodic memories

Dreams play a fundamental role in forming semantic representations through creative development of episodic memories.

Dream-Induced Enrichment

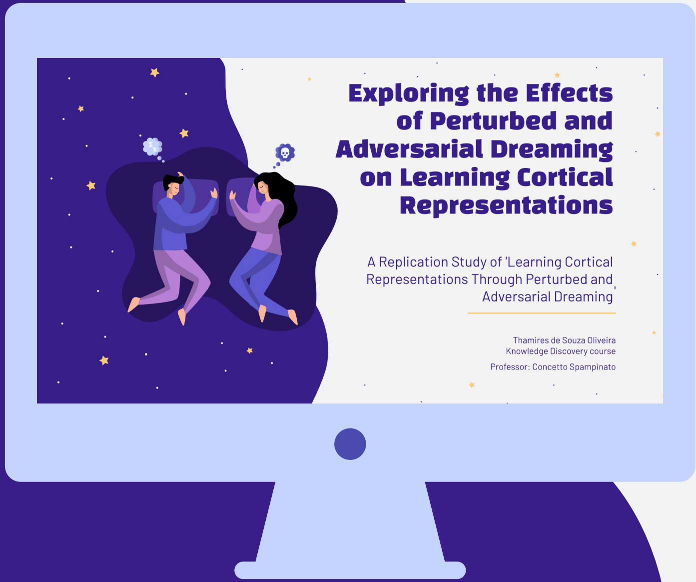
Incorporating dream-induced memories enhances the quality of semantic representations beyond direct sensory input.

Understanding How Dreams Inspire Learning

The model reveals mechanisms of dream-driven learning, replicating the brain's ability to consolidate and generalize sensory experiences.

Implications for Learning

Understanding the role of dreams in forming semantic representations has broader implications for learning and knowledge acquisition in humans and other organisms.



Thank you!

Contact: thasouzaoliv@gmail.com