**Project Proposal: Text-to-Image Synthesis of Dog Breeds Using GANs**

**Problem Statement and Task Definition:**

Our system is designed to generate photorealistic images of dogs from textual descriptions, specifically the dog breed. The input is a text string denoting the breed of a dog, and the output is a corresponding image of a dog that reflects the characteristics of the given breed. The real-world application for such a system ranges from aiding in the visualization for breeders and potential pet owners to assisting in educational materials for veterinary training. The system aligns well within a reasonable scope as it is not as broad as generating any object from text, nor as narrow as creating images of dogs in only one posture or setting.

**Input/Output Behavior:**

The input to the system is a simple text string, such as "Golden Retriever" or "Siberian Husky". The output is an image of a dog that visually matches the breed described by the input text. Utilizing the Stanford Dogs Dataset, which contains numerous images of different breeds, allows us to illustrate the anticipated input and output. For example, an input of "Dachshund" would produce an image of a short-legged, long-bodied dog, characteristic of the breed.

**Evaluation Metric:**

To measure the success of our system, we will employ the Inception Score (IS) and Fréchet Inception Distance (FID) as our primary evaluation metrics. These metrics are suitable because they not only assess the diversity of the generated images (variability within a breed) but also the similarity to authentic images (how convincing the breed characteristics are). We will curate a dataset of input-output pairs where the outputs are actual images from the Stanford Dogs dataset, against which we will compare our generated images.

**Related Works:**

Projects such as "This Cat Does Not Exist" demonstrate the potential of GANs in generating realistic animal images. Research papers such as "Text-to-Image Synthesis Using Generative Adversarial Networks" by Reed et al. and "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks" by Radford et al. provide foundational techniques and insights into the capabilities of GANs in a related context.

**Baseline and Oracle:**

As a baseline, we will implement a system that generates a random image from the dataset, regardless of the input text. This will set a lower bound on performance. Our oracle will be a system that selects an actual image of the requested breed from the dataset, ensuring that it is indistinguishable from a real photograph, setting the upper bound.

**Methodology:**

We will approach this problem by utilizing Generative Adversarial Networks (GANs), which are well-suited for generating complex images. Variational Autoencoders (VAEs) may also be experimented with for comparison. The main challenge is ensuring that the generated images capture the unique features of each breed accurately. To model this task, we will consider the latest techniques in deep learning for image generation, such as StyleGANs or Attention GANs, which have shown promising results in related fields. The tradeoffs between accuracy and efficiency will be considered when choosing specific algorithms, as more complex models may yield better images but require more computational resources.

**Description of the Challenges:**

Challenges include the subtle differentiation between similar dog breeds and the system's ability to interpret the text accurately to modify fine details in the generated images. High-level techniques from the field of computer vision and natural language processing will be necessary to tackle these issues, potentially involving advanced feature extraction and semantic understanding. We will explore the literature and existing technologies to overcome these obstacles effectively.