**Research Plan: Improving Particle Discrimination in WIMP Dark Matter Detection Experiments Using Neural Networks**

**Rationale**

Detection of dark matter is one of the foremost endeavors in modern particle physics. Due to unexpectedly large gravitational forces affecting the rotation of galaxies and the deflection of photons, it is hypothesized that there are a large number of “dark” particles that rarely interact via forces other than gravity. One of the most-researched candidates for such a particle is the Weakly Interacting Massive Particle, or WIMP.

The identification of background radiation events is a major hurdle to overcome in all experiments for WIMP detection. The current practice of manually developing a discriminator function to eliminate background events is challenging and time-consuming when available calibration data is frequently impure and present in limited quantities. Machine learning, particularly in the form of neural networks, has the potential to be a powerful solution, because it permits automation of the process.

However, there are several problems that must be overcome. For instance, impure data can be detrimental to the training of many types of models, and data often comes in formats that are non-trivial to process, such as audio and 3D geometries. It is not always clear which of these formats can be used to produce the most effective discriminator, nor is it obvious what kind of neural network will learn the most efficiently and accurately on a given data format.

If solutions to these issues can be found, it stands to reason that the pace of many dark matter research projects could be considerably improved. There could potentially be a great reduction in the engineering time that is spent on manually deriving a discriminator function every time some aspect of the experiment changes.

**Engineering Goal**

In summary, the goal of this project is to develop machine learning systems that can effectively and efficiently learn to discriminate between signal and background events in the context of the PICO-60 and DEAP-3600 WIMP detection experiments. This will be achieved by experimenting with a variety of different input formats, network architectures, and novel systems for improving data efficiency.

Furthermore, based on the results that come out of this investigation, a secondary goal of this project is to provide some insight on the properties of the input data that are taken advantage of by the neural network discriminators.

**Procedures**

In the case of the PICO-60 project, there are a number of different input formats that can potentially be used as input data for a neural network. The format used in past research on the experiment is a Fourier transform integrated into eight frequency bands. This was used as the input to a conventional discriminator known as the Acoustic Parameter (AP), as well as to a multi-layer perceptron that was applied experimentally as part the original PICO-60 study.

There are three other data formats that will be investigated in the context of PICO-60:

1. A full-resolution analogue to the banded Fourier transform, using all values in the high-resolution curve directly without any integration
2. A full-resolution audio waveform, which includes time domain information (such as how the frequency distribution changes over time) which may be useful
3. Images captured by cameras in the detector, which have been used only for reconstructing the positions of events but are not proven to be inapplicable for background event identification

Using whichever of these data formats is observed to produce the most effective discriminator, the intent is to experiment with a wide variety of neural network hyperparameters, as well as more