**SWS3009 Summer Workshop**

**Assg – Answer Book**

**Submission Deadline: Saturday 7 July 2024 2359**

Name 1: YANG RUNKANG

Name 2: WEN SIJIE

Name 3: HE JUNHANG

Name 4: ZHANG CHENHAN

Introduction

In this project, we aimed to build a remotely piloted vehicle to identify various species of cats by recognizing pictures of cats placed along a maze. The objective was to collect images of specific cat breeds, train a convolutional neural network (CNN) using transfer learning, and deploy the trained model to classify images accurately.

Image Collection

We collected images of five cat breeds from the internet:

* Ragdolls
* Singapura cats
* Persian cats
* Sphynx cats
* Pallas cats

The images were divided into 85% training images and 15% validation images, ensuring an equal number of pictures for each breed. We aimed for a dataset comprising over 1,000 images to achieve a balanced representation for each class.

Selecting and Training the Deep Learning Network

We chose to use a CNN architecture for this task due to its proven effectiveness in image recognition tasks. Specifically, we utilized the InceptionV3 architecture, a well-known model pre-trained on the ImageNet dataset. The choice of InceptionV3 was based on its strong performance in handling diverse and complex image data.

Model Architecture

The architecture of our model involved the following steps:

* **Base Model:** We used the InceptionV3 model without the top classification layer. The model was initialized with weights pre-trained on ImageNet.
* **Global Average Pooling:** The output of the InceptionV3 base model was fed into a global average pooling layer to reduce the dimensionality and aggregate the feature maps.
* **Fully Connected Layers:** A dense layer with 200 neurons and ReLU activation was added, followed by a final dense layer with softmax activation to produce probability distributions for the five cat breeds.

Transfer Learning Process

We implemented transfer learning to adapt the pre-trained InceptionV3 model to our specific task:

* **Freezing Layers:** Initially, all layers of the InceptionV3 base model were set to be non-trainable to preserve the learned weights from ImageNet.
* **Training New Layers:** The newly added dense layers were trained using our cat breed images.
* **Fine-tuning:** After training the dense layers, we fine-tuned the entire model by unfreezing the last three layers of the InceptionV3 base model. This allowed us to adapt the high-level features to our dataset while maintaining the robustness of the pre-trained weights.

Training Process

The training process was implemented in two stages:

* **Initial Training:** We trained the dense layers for 20 epochs using the RMSprop optimizer with categorical cross-entropy loss.
* **Fine-tuning:** The model was further trained for another 20 epochs using the SGD optimizer with a very low learning rate to fine-tune the entire network.

We used data augmentation techniques, such as rescaling, shear range, zoom range, and horizontal flip, to enhance the robustness of our model by generating varied training samples.