DINDER: A DOG MATCHING WEB-BASED APPLICATION USING CONVOLUTIONAL NEURAL NETWORKS WITH EFFICIENTNETB0 TRANSFER LEARNING

I. INTRODUCTION

Dogs have been a great source of companionship to humans for decades, where they were often hailed as human best friends according to Cohut (2018). According to Biswas (2019), dogs provide therapeutic benefits to people, thus improving the quality of their lives. As a result, many dog owners find the companionship of their dog so rewarding that they want to breed their dog, to carry on the bloodline or to keep a puppy. Dog owners commonly find partners for their dogs through posts from dog breed communities online or by reaching out to known dog breeders. However, there were a few acknowledged difficulties with the aforementioned methods. Maxwell (2020) asserted that finding the perfect dog partner through contact with dog breeders required asking each one for an available and preferred dog partner. Moreover, through Facebook groups, one must look thoroughly through each and every post even if one was hardly relevant to an owner's preference. As a result, selecting a preferred dog partner was often viewed as an expensive process that took a lot of time and effort.

Therefore, a better method for matching dogs would be practical. The use of a software application to find a matching dog partner and provide images of their possible offsprings could reduce the time and effort taken by dog owners in searching for a match for their dogs.

Accordingly, the software application in the form of a web application prototype was developed and integrated with a deep learning algorithm which was a type of artificial intelligence (AI) inspired by the structure of the human brain according to Oppermann et al. (2019). In particular, convolutional neural network (CNN), a deep neural network specialized in image recognition or classification as defined by Kim (2017), was used in the prediction or image recognition of dogs' breeds, coat colors, and types.

Accordingly, the web application prototype developed in this study was called Dinder, inspired from a popular application, Tinder, that used personal data for matching. Dinder was built as a dog partner matching application with image display of offsprings relies on dogs' information, primarily their breeds, coat colors, and types as well as other data such as location, gender, etc. to potentially reduce the time and effort needed to find the ideal dog partner.

II. RESEARCH OBJECTIVES

Generally, the study aimed to design and develop a prototype for dog matching in the form of a web application integrated with deep neural networks, particularly transfer learning-based convolutional neural network (CNN) models, with the transfer learning model selected through evaluations of several pre-trained CNN models having distinct numbers of parameters or weights

for the benefit of dog owners in potentially reducing the time and effort taken to search for a preferred dog partner for their dogs.

Specifically, this study sought:

1. To design and develop a dog matching web application prototype with image recognition of breeds, coat colors, and types as well as image display of offsprings for Labrador Retrievers, Pomeranians, Poodles, Pugs, and Shih-Tzus.

2. To conduct model evaluation based on the accuracy, precision, recall, and F1-score of pre-trained CNN models, ResNet50, VGG16, and EfficientNetB0, with distinct numbers of parameters, to determine which transfer learning model to integrate.

3. To examine the performance of the models developed integrated into the prototype through actual testing.

III. LITERATURE OF THE STUDY

In conducting this study, several artificial intelligence (AI) theories were looked into for image classification. But, only one branch out of the six branches of artificial intelligence was used for this study. The determined branch of AI chosen by the researchers was neural networks, which was defined by Mahanta (2017). The reason this was chosen by the researchers was because it was extremely popular for computer vision tasks. Moreover, neural networks were effective at handling complex input, such as images, which was perfect for classification of images of dogs.

In general, a single neural network won't be enough to build an accurate image classifier. A set of neural networks would be needed for this study. Rusk (2016) referred to this as deep learning. Through deep learning, Lecun et al. (2015) stated that features extraction would be done for the image datasets. This would allow the extraction of features, such as edges, colors, etc. Moreover, Fu et al. stated that deep learning would be really necessary because of the advancement of technology.

Technically, a neural network could also be called an artificial neural network (ANN). The reason was because as stated by Mohseni-Dargah et al. (2022), it consisted of input, hidden, and output layers with connected neurons inspired from the human brain. Malik (2019) then explained the responsibility of the input layer, hidden layers, and output layer. For image recognition, Mehta (2022) stated that there were many types of neural networks, but of all the types mentioned, convolutional neural networks were used for this study, since it was the most effective for image classification or recognition.

In this study, convolutional neural networks were used for image recognition. Mishra (2020) stated that it typically has three layers, namely, a convolutional layer, a pooling layer, and a fully connected layer. For convolutional layers, Yamashita (2018) discussed its process. For pooling layers, Seb (2021) explained its importance, including the common types of pooling, and why the global average pooling was used for

this study. In fully connected layers, Arc (2018) explained its process as the last part of a convolutional neural network, including the structure of input it accepts.

For this study, several model architectures of convolutional neural networks were used. The first type of model was ResNet, which was defined by Shakhadri (2021), particularly ResNet50. The reason ResNet was developed was discussed by He et al. (2016). Architecturally, Ji et al. (2021) discussed its composition. In terms of its performance, Ayanzadeh & Vahidnia (2018) conducted a performance evaluation.

The second type of convolutional network used for this study was VGG, which was defined by Boesch (2021), particularly VGG16. In terms of its performance, the study by Tammina (2019) performed testing for its accuracy.

The last type of model architecture used for this study was EfficientNet, which was defined by Kumar (2020), particularly EfficientNetB0. He et al. (2016) then mentioned the advantage of EfficientNet. In terms of its performance, the EfficientNetB0, which was the baseline model for the family of EfficientNets according to Zoph & Le (2017), was evaluated in the study by Stock & Cavey (2021) for accuracy.

But, without the application of transfer learning, which was defined by Donges (2022), this study wouldn't be possible. The design methodology of transfer learning was further discussed by Koehrsen (2018).

In order to evaluate the transfer learning-based models for this study, the accuracy metric which was defined by Agrawal (2021), precision metric which was defined by Fernández et al. (2018), recall metric which was defined by Ma et al. (2013), and F1-score metric which was defined by Fong & Biuk-Aghai (2009), based on the confusion matrix which was defined by Suresh (2020), were used.

Overall, this study was focused on dogs, particularly breeding. King et al. (2012) mentioned that nowadays, it was more important for dogs to be perceived based on their beauty when breeding. This was further discussed by Hirst (2019), stating that when an interesting or useful trait was identified in a dog, owners would breed from that dog in the hope of the trait being passed on.

IV. METHODOLOGY

Research Design

This study was conducted with the use of experimental and developmental research design methods. As stated by Arikunto (2006), an experimental research design was a study that aimed to know if there was an effect on the variable being studied. This method was utilized in order to determine the best-performing pre-trained convolutional neural network model in terms of image classification with respect to the datasets collected.

Furthermore, as defined by Richey (1994), developmental research was the systematic study of designing, developing, and evaluating instructional programs, processes, and products that must meet the

criteria of internal consistency and effectiveness. The researchers used developmental research to determine the features to be developed for the system. It was also used to assess the system's consistency and efficacy during development.

Fact Finding Technique

This section would discuss the thorough review of the technique used to collect important information, crucial to accomplish the objectives of the study. It was discussed in the following:

Online Research

Several methods were conducted through online research. First, preliminary investigations were conducted to establish the research problem, and to determine the scope of breeds, coat colors, and types of dogs. Second, data collection was conducted for image datasets. Last, a thorough knowledge of the development process for the prototype was conducted.

In conducting the preliminary investigations, it was determined that indeed finding a dog breeding partner consumed a lot of time and effort because of the need to look thoroughly through each Facebook post and contact various breeders just to find the preferred one. Before collecting the image datasets, the top five most common breeds in the Philippines as of March 2022 based on the study by Co (2022) was determined as the scope for dog breeds in this study. In terms of their coat colors, two coat colors for each breed were selected for this study. Moreover, in terms of their types, two types for each breed except for pugs, were selected for this study, since pugs were differentiated based on their coat colors only.

In collecting the datasets, the images of the particular dog breeds along with two of their coat colors and two of their types were able to be validated and found with the use of Google Images. Each image was manually downloaded and categorized into respective folders using a file manager application, File Explorer

In acquiring a thorough knowledge of the development process to build the web application prototype, Dinder, the necessary skills and knowledge were studied. The vast number of resources available online enabled the study for deep neural networks, particularly image recognition.

Algorithm Analysis

The use of deep neural networks for image recognition was chosen for this study. However, training a deep neural network from scratch required a significant amount of time, datasets, and memory for an accurate model. Therefore, since the image datasets for this study were small, transfer learning was integrated. It referred to the transfer of knowledge gained from a previously trained deep neural network model. The use of transfer learning made it possible to train even small datasets. Moreover, the training time and memory usage was significantly less.

A popular approach to the utilization of transfer learning was through the Keras library. It was a high-level, deep learning API developed by Google for implementing neural networks. Keras had convolutional neural network models which were trained and validated on the ImageNet dataset that contained 1000 object classes and over 1.2 million images. For the purpose of this study, four pre-trained models with distinct numbers of parameters were considered and evaluated. The said pre-trained models were discussed in the following:

Data Model Generation

This section would enumerate the various methodologies employed in developing the algorithm models. It would include splitting the datasets into train and test sets, image preprocessing, image augmentation, and training the transfer learning-based models

Dataset Splitting

The technique of splitting a dataset, based on Brownlee (2020), was used to evaluate the performance of algorithms when they would be used to make predictions on data that were not yet seen by a trained model

Image Preprocessing

Images come in different shapes and sizes. Image preprocessing, according to Nelson (2020), referred to steps taken to format images before they would be used for training models.

Image Augmentation

The image datasets collected for this study were very small in size. In order to have somewhat of an accurate model, image augmentation was applied to the image datasets. Image augmentation, according to Saxena (2021), was a technique of altering the existing data to create some more data for the model training process in order to increase the size of training data, prevent overfitting, and help the models generalize better. For each iteration in training, an additional of 10 images would be produced.

Training

Generally, in order to train and develop a neural network model, the model architecture had to be given an input layer, hidden layers, and an output layer. In a convolutional neural network, its input layer would be images, its hidden layers would be convolutional layers, and its output layer would be fully connected layers. However, in a transfer learning-based convolutional neural network, its hidden layers would be convolutional layers from pre-trained convolutional neural networks.

In training, the three image datasets: breeds dataset, coat colors dataset, and types dataset were used to train three image classification models: breed classification, coat color classification, and type classification. Each model was a transfer learning-based convolutional neural network model, with the pre-trained model to be determined in model evaluation

Model Evaluation

In order to calculate the performance of the transfer learning-based convolutional neural network models, they would need to be run through a test set. The transfer learning-based model used for evaluation was the dog breed classifier only. Moreover, the test set used for evaluation was from the breed images dataset only. This was due to the fact that testing on the other developed transfer learning-based convolutional neural network models as well would just return the same performance levels.

Since the models were used on the prediction of category labels, the metrics for classification algorithms that were based on the confusion matrix were used. A confusion matrix was simply a table used to define the performance of classification algorithms, where:

- True Positive (TP) referred to the total number of correctly predicted positive data points;
- True Negative (TN) referred to the total number of correctly predicted negative data points;
- False Positive (FP) referred to the total number of incorrectly predicted positive data points;
- False Negative (FN) referred to the total number of incorrectly predicted negative data points;

System Development Methodology

A Software Development Life Cycle (SDLC) model was utilized in the development of the prototype and its neural network models. According to Iqbal &

Idrees (2017), SDLC models were used typically to improve the quality of software and the overall development process.

There were many SDLC models to be used for software development, each with their own strengths and weaknesses, but the Rapid Application Development (RAD) model was chosen for the development of this study.

According to Dopico (2019), RAD was a form of agile software development methodology that prioritizes rapid prototype releases and iterations. As a result, it could deliver developed software and make changes quickly while aiming to deliver a product that more closely fits the needs of the user through consistent testing and tweaking. The process of RAD as shown in the figure above consisted of four (4) main phases that steer the development process.

V. RESULTS AND DISCUSSION

RESEARCH OBJECTIVE 1: To design and develop a dog matching web application prototype with image recognition of breeds, coat colors, and types as well as image display of offsprings for Labrador Retrievers, Pomeranians, Poodles, Pugs, and Shih-Tzus.

In order to achieve the first objective of the study which was to design and develop the application, Dinder, online research was utilized as the initial stage in the Rapid Application Development (RAD) software methodology as indicated in the third chapter of the study to identify the necessary features to develop.

RESEARCH OBJECTIVE 2 To conduct model evaluation based on the accuracy, precision, recall, and F1-score of pre-trained CNN models, ResNet50, VGG16, and EfficientNetB0, with distinct numbers of parameters, to determine which transfer learning model to integrate.

To achieve the second objective of the study, the researchers evaluated the pre-trained models, ResNet50, VGG16, and EfficientNetB0 with different numbers of parameters. Each pre-trained models' number of parameters were identified by the researchers upon observation from the Keras library API (*Keras Applications*, n.d.). The pre-trained models were evaluated based on their accuracy, precision, recall, and F1-score. This evaluation served as the basis for the selection of the pre-trained models to integrate for the transfer learning-based models.

RESEARCH OBJECTIVE 3: To examine the performance of the models developed integrated into the prototype through actual testing.

In order to achieve the last specific objective of the study, prototype actual testing was conducted. The predictions of the image classifiers were assessed based on their reliability. Moreover, the offspring image displays were assessed based on their logicalness.

VI. CONCLUSION

Based on the findings of the study from the objectives, the researchers have reached the following conclusions and generalizations:

1. In developing the prototype, the researchers discovered that the use of Flask as the backend framework for the prototype fit very well with the RAD software development methodology's quick process of development.

2. The researchers discovered that based on Chapter 4, the integration of the parameters or weights of EfficientNetB0 in a CNN model resulted as the best-performing, since the small datasets adapted better with the simpler convolutional neural network with 5.3 million parameters as opposed to other pre-trained models of tens of millions of parameters.

3. Upon conducting the prototype actual testing, the researchers found that once the predictions of breeds and coat colors were correct, the following test cases would be correct as well.

VII. RECOMMENDATIONS

Based on the conclusions, the researchers have reached the following recommendations:

1. The researchers recommend a full development of the prototype with the prototype being open-sourced on a GitHub repository.

2. The researchers recommend the use of simpler pre-trained models that have significantly less parameters or weights for simple datasets that have few images.

3. The researchers recommend the input of blurry or pixelated images to fully test the prototype.