International Journal of Scientific Research & Engineering Trends
Volume 10, Issue 5, Sept-Oct-2024, ISSN (Online): 2395-566X

Malaysian Noodle Images Classification System Using CNN and Transfer Learning

Ibrahim Abba[1], Ubaid Mohammed Dahir[2], Mohammed Shettima[3]
School of Computer Science, Mai Idris Alooma Polytechnic Geidam Yobe State. Nigeria[1]
School of Computer Science, SIMAD University Somalia[2,3]

Abstract- Image Recognition is a term used to describe a set of algorithms and technologies that attempt to analyze images and understand the hidden representations of features behind them and apply these learned representations for different tasks like classifying images into distinct categories automatically, understanding which objects are present and where in an image, etc. These technologies leverage various traditional computer vision methods as well as machine learning and deep learning algorithms to achieve the required results for solving such problems. This paper shows a recognition model for classifying Malaysian Noodle images. Convolutional Neural Network (CNN) algorithms, a deep learning technique extensively applied to image recognition were used for this task. The model uses a deep learning process that was trained on natural images (AlexNet and SqueezeNet dataset) and was fine-tuned to generate the predictive Noodle model, which comprised approximately 4308 images. The dataset was divided into ten groups/categories of Noodles images which include the following: Mee Bee Hoon Goreng, Mee Bee Hoon Sup, Mee Goreng, Mee Koay Teow Goreng, Mee Koay Teow Sup, Mee Laksa Goreng, Mee Laksa Sup, Mee Maggi Goreng, Mee Maggi Sup, Mee Sup. The trained model achieved high accuracy on the test set, demonstrating the feasibility of this approach.

Index Terms- Noodle, Classification, Deep Learning, CNN, AlexNet, SqueezeNet.

I. INTRODUCTION

Noodles are a type of food made from unleavened dough or rice that is rolled flat and cut, stretched, or extruded, into long strips or strings [1]. The dough is being prepared and shape into a certain size and thickness to suit certain types of cooking. Noodles can either be stored in a dried shelf for long term usage or it can also be stored in a refrigerator to keep it fresh for a short-term usage. Noodles have been a significant component of cuisines throughout the world. It is no different in Malaysia, where Malay and Chinese noodles are an integral part of the local culture. Worldwide, there are varieties of noodles, each with a unique shape, flavor, and preparation method. Noodles are a food that is quite well-liked across the world, and they often have a particular or regional flavor to them. Noodle classification is a complex problem because it looks almost same and is difficult to differentiate between humans.

The picture below shows categories of noodles dishes in Malaysia:

1. Mee Goreng
Mee Goreng, or mi goreng, refers to fried noodles in the Malay-archipelago cultures. Mee Goreng is a staple food in places like, the Southeast Asian states of Brunei, Malaysia, and Singapore. It is also a plate of street food-style Mee Goreng.

Figure 1: Mee goreng

2. Laksa
Laksa is a mild-spicy noodle dish popular in Southeast Asia. Laksa consists of thick rice noodles, with toppings such as chicken, prawn, or fish. Laksa noodle is cooked with mackerel paste, and it may come with various add-ons vegetables to lift the taste.

© 2024 IJSRET

Figure 2: Laksa

3. Mee Bee Hoon Goreng

Bee Hoon Goreng, or commonly known as Bihun Goreng is originally a Chinese dish. With mixture of culture and adaptations from Sumatera, Bihun goreng that we have today fits to Malay palates. Bihun noodles is way thinner compared to Mee, and it requires an extra effort to prepare and preserve for cooking. Classic version of Bihun goreng is cooked with prawns, beef balls, and chicken slices.

Figure 3: Mee Bee Hoon Goreng

4. Maggi Goreng

Maggi Goreng is a Malaysian dish made from instant noodle, where the noodle is visibly curly compared to other noodles. Maggi goreng is cooked using wok with several types of seasoning, spices, meats, and vegetables. But usually, the recipe to cook Maggi goreng is quite like Mee goreng, and even the look is quite similar in a first glance. Maggi's curly visible traits and noodle texture is what mostly differentiate them.

Figure 4: Maggi Goreng

5. Mee Bee Hoon Sup

Bihun sup is a traditional noodle soup originating from Malaysia. The dish consists of beef broth, rice vermicelli noodles, and sambal chili paste. The broth is made with a combination of beef, ginger, garlic, shallots, cloves, star anise, cinnamon, and salt. Bee Hoon Sup may come with additional scallop garnishes and sambal on top.

Figure 5: Mee Bihun Sup

6. Mee Koay Teow Goreng

Koay Teow refers to flat rice noodles. It is made from noodles, locally known as kwetiau, which are stir fried in cooking oil with garlic, onion or shallots, beef, chicken, fried prawn, crab or sliced bakso (meatballs), chili, Chinese-cabbage, cabbages, tomatoes, egg, and other vegetables with an ample amount of kecap manis (sweet soy sauce). In Asia, kwetiau is available in two forms, dried and fresh. Its recipe is quite like another Chinese Indonesian favorite, mee goreng, except for replacing yellow wheat noodles with flat rice noodles.

Figure 6: Mee Koay Teow Goreng

7. Mee Koay Teow Sup

Koay Teow sup is a Chinese noodle soup dish that may come with various spicy broths such as chicken broth, beef broth, and pork broth. Koay Teow Sup has a similar visible traits like Bee Hoon Sup, which the dish is prepared with extra garnishes on top with spicy sauce to top it off.

© 2024 IJSRET

Figure 7: Mee Koay Teow Sup

8. Mee Laksa Sup

The soup is made with mackerel and its main distinguishing feature is the Assam or tamarind which gives the soup a sour and appetizing taste. The fish is poached and then flaked. Other ingredients that give Penang Laksa its distinctive flavor include mint, pineapple slices and otak udang.

Figure 8: Mee Laksa Sup

9. Mee Maggi Sup

Mee Maggi Sup is Mee Maggi served with an aromatic broth that captures the essence of long-boiled sup tulang, seasoned with six spices that are familiar by Malaysians, paired with springy noodles made from Australia-sourced wheat grains. Compared to Bee Hoon sup or Mee sup, Maggi sup has a thicker clear broth due to the egg that is cooked within the broth.

Figure 9: Mee Maggi Sup

10. Mee Sup

Mee sup (translated simply as noodle soup) is a traditional Malaysian dish with a spicy broth base. There are different varieties of mee sup that combine either tofu, prawns, pork, or other types of meat with a variable blend of vegetables like chilies, onions, garlic, and shallots.

Figure 10: Mee Sup

Deep learning is used in image recognition and these methods have made important progress which will be able to extract features of noodles. Noodle classification can be classified as an image classification problem. The proposed system has used Convolutional Neural Networks (CNN) for detecting the different noodles information systems from images. The major concern here is that most people are finding it difficult to differentiate noodles by just sighting or being able to recognize noodles that are familiar to their culture. By using people who are familiar with noodles type, the image set will be able to help train Neural Network of Noodle classification system and obtain a high accuracy rather than the related works. The system has provided performance with simplicity and efficiency.

This paper will focus on the following contributions:

- To develop a high-performance Noodle classification system. Noodles images can be used to perform deep learning
- To study the performance of a type of neural network for detecting objects of images.
- To know how to build a Noodle classification system using machine learning approaches based on the neural networks.
- To improve the classification quality by tuning the model based on the discovery from the study of multiple types of neural network models.

II. LITERATURE SURVEY

1. CNN-based Approach for Recognition

Convoluted Neural Network (CNN) deep Learning algorithm widely adopted in image recognition. CNN deep learning mimics the way our human brain functions where it recognizes a certain pattern and communicates through neurons and synapses [2]. In general, CNN has serves multiple purposes such as image processing, segmentation,

© 2024 IJSRET

classification, and more. One of the most used CNN models is image processing solution for food recognition.

In 2014, scholars from Japan have constructed image recognition system by using CNN to identify common food in Japan [3]. The list of the common food goes by ramen, miso soup, natto, salmon, and more Japanese cuisine. Multiple hyperparameters such as layering, data processing, and data training is needed to build CNN. This study interestingly uses local response normalization (LRN) to normalize the pool layer of CNN across the map. With 6-fold cross validation, the CNN model is compared to an SVM-based recognition. Results show that CNN produces 70% accuracy while SVM-based recognition produces only 50% accuracy. This study has perfectly shown that CNN is a better algorithm to perform food recognition.

In 2015, another study from Japan is conducted to help identify food calories to ensure healthy lifestyle by using a CNN model [4]. This study is performed on many types of Japanese food, including instant noodle. According to Figure 11, each of the components of the instant noodle is separated according to its own bounding box, and thus each element is completely separated. Below is a picture of the proposed method to calculate food calories of an instant noodle:

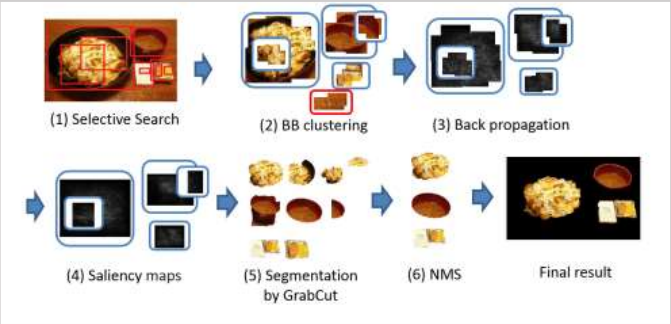

Figure 11: Method breakdown for image processing by using selective box search model

This method is remarkably interesting because each component of the food was break down into small sample size, thus it can help to identify the ingredients better. For types of noodle recognition, this method can be served as one of the ways for image recognition.

2. AlexNet Algorithm

According to Towards Data Science article, AlexNet is one of the most used DCNN algorithms that was introduced back in 2012 due to a competition between ImageNet [5]. AlexNet has grown popular since it utilizes less convolutional layers compared to other commonly used DCNN-models such as ImageNet itself. Basic layers of AlexNet are composed of five convolutional layers, three polling, two local responses, and three fully connected layers as shown in Figure 2 [6]. In the journal written by Yiğit, AlexNet is used to perform recognition to a single type of image such as apple, bread, and so on. This algorithm shows a satisfactory performance with accuracy of 82%, which is better than CaffeNet at 80% [6].

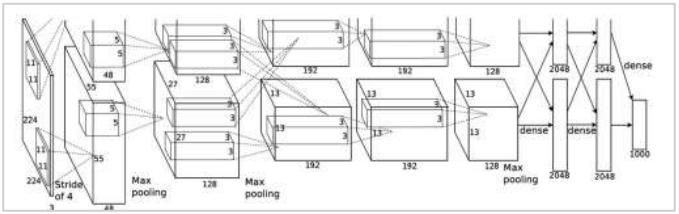

Figure 12: AlexNet architecture

Another similar research was done by a group of students from Indian Institute of Technology, mentioned that AlexNet architecture was chosen as their baseline because computational time is significantly lesser compared to other architecture such as GoogLeNet [7]. In this experiment, AlexNet is trained in a parallel fashion with a depth of two and it returns a good accuracy of 80%. Since our dataset is noticeably big, we believe that AlexNet can help us to perform DCNN in a less computational time with good accuracy.

3. SqueezeNet Algorithm

SqueezeNet was developed with the goal of developing CNN architecture with fewer parameters while maintaining competitive accuracy so that it can be stored efficiently in computer memory and smoothly implemented over a computer network [8].

To achieve the objective, three major strategies were conducted where filters of the CNN were replaced with a smaller size while decreasing the number of input channels. The third step was to down sample late in the network so that convolution layers have large activation maps. In this paper, the authors were comparing SqueezeNet and AlexNet. For perspective, AlexNet has 240MB of parameters while SqueezeNet has 5MB of parameters, yet they produced the same 80.3% Top 5 accuracy on ImageNet classification validation dataset.

The building brick of SqueezeNet were clearly explained by Hassanpour et al. [9]. The SqueezeNet stacks multiple pooling layers and multiple fire modules. Both the squeeze layer and the expand layer maintain the same size for the feature map, however the former decreases the depth while the latter increases it as can be seen in Figure 1.

Squeeze layer is simply a layer of 1x1 convolution maps while the expand layer is the result of integration between 1x1 and 3x3 maps. In neural architectures, the squeezing (bottleneck layer) and expansion behavior is typical. Another typical strategy is to get high level abstract by decreasing feature map size while increasing depth.

© 2024 IJSRET

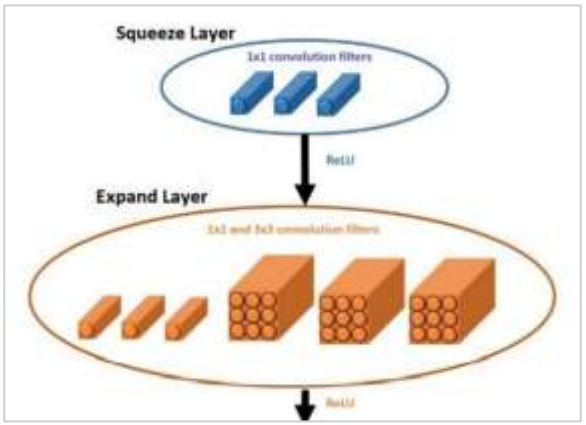

Figure 13: Overall structure of Fire module in SqueezeNet module

In a paper by Yadav et al [10], SqueezeNet architecture was deployed for food image classification. Food images have high intra class variance and low inter class variance thus making it a challenging task for classic machine learning approach such as SVM to classify it however it is a walk in the park for SqueezeNet where it achieved acceptable accuracy at 77.20%. CNNs can automatically extract complex features due to the substantial number of parameters and more depth. The result was attained by performing data augmentation and fine tuning the hyperparameters.

Another prime example of SqueezeNet application can be seen in Vegetable and Fruit Image Classification that has been conducted by BAYGIN [11]. The main process in executing the project were feature extraction, feature selection and classification. In this context SqueezeNet was used to do feature extraction through transfer learning approach. Following the approach, the ”GlobalAveragePooling” layer of the SqueezeNet architecture was used. SqueezeNet was chosen as it has the highest accuracy score compared to other pre trained models such as ResNet101, GoogleNet and VGG16.

III. PROPOSED DEEP LEARNING FOR NOODLE RECOGNITION

In this paper, we proposed four CNN architecture that starts with AlexNet, SqueezeNet, Basic CNN(6 layers), and Basic CNN(7 layers).

1. Proposed CNN Architecture

AlexNet

The 2012 Imagenet large-scale visual recognition competition was won by Alexnet. The approach was put out by Alex Krizhevsky and his colleagues in their 2012 research publication, Imagenet Classification with Deep Convolution Neural Network [12]. There are eight learnable layers in the Alexnet. ReLU activation is used in each of the five levels of the model, with the exception of the output layer, which uses max pooling followed by three fully connected layers.

They discovered that the training process was nearly six times faster when the ReLU was used as an activation function. Additionally, they made use of dropout layers, which stopped their model from overfitting. Moreover, the Imagenet dataset is used to train the model.

The photos with a size of 227*227*3 are used as the model's input and to avoid the size of the feature maps from significantly decreasing, Alex added padding to the model. Figure 14 shows the architecture of Alexnet.

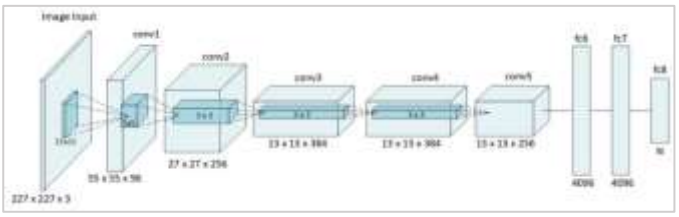
Figure 14: Architecture of Alexnet [12]

Table 1: Architecture of Alexnet [13]

Layer	Filters/ neurons	Filter size	Stride	Padding	Size of feature map	Activation function
Input	-	-	-	-	227*227 *3	-
Conv 1	96	11*1 1	4	-	55*55*9 6	ReLU
Max Pool 1	-	3*3	2	-	27*27*9 6	
Conv 2	256	5*5	1	2	27*27*2 56	ReLU
Max Pool 2	-	3*3	2	-	13*13*2 56	-
Conv 3	384	3*3	1	1	13*13*3 84	ReLU
Conv 4	384	3*3	1	1	13*13*3 84	ReLU
Conv 5	256	3*3	1	1	13*13*2 56	ReLU
Max Pool 3	-	3*3	2	-	6*6*256	-
Dropo ut 1	Rate = 0.5	-	-	-	6*6*256	-

From Table 1, it shows the architecture of Alexnet with filters, filter size, stride, padding, size of feature map and activation function used in each layer. The photos that make up this model's input are 227*227*3 pixels. The first convolution layer is then applied, using 96 filters with a size of 11*11 and a stride of 4. ReLU is the activation function utilized in this

© 2024 IJSRET

layer. The final feature map measures 55*55*96 pixels. The first Maxpooling layer, which is of size 3*3 and stride 2. The final feature map is then produced and has the dimensions 27*27*96.

The second convolution process is then used. This time, the filter size is 5*5, and there are 256 of them. Padding is two and the stride is one.

Once more, ReLU is the activation function utilized. The output size that we now receive is 27*27*256. In addition, we used stride 2 to apply a max-pooling layer in the size 3*3. The resultant feature map has the dimensions 13*13*256.

The third convolution operation is now applied, using 384 filters of size 3*3, stride 1 and padding 1. Once more, ReLU is used as the activation function. The final feature map has the dimensions 13*13*384. The fourth convolution operation, which uses 384 filters of size 3*3. The padding and stride together are 1. Additionally, ReLU is employed as an activation function. The output size is still 13*13*384 as of right now.

Following this, we get the final 3*3 convolution layer with 256 of these filters. The stride, padding, and activation function are all set to one. The final feature map has the dimensions 13*13*256.

The third max-pooling layer, measuring with 3*3, and stride 2 are then applied. Creating the feature map with the dimensions 6*6*256. As we go further, there are more filters being used. As we go more into the architecture, it is thus extracting additional functionality. The original filter was larger, and as we move forward, the filter size is diminishing, which causes a reduction in the form of the feature map.

SqueezeNet

SqueezeNet" that achieves AlexNet-level accuracy on ImageNet with fifty times fewer parameters.[14] SqueezeNet uses three strategies, first replaces 3x3 filters with 1x1 filters to reduce parameters.

Secondly, decrease the number of input channels to 3x3 filter. [14] Thirdly down sample late in the network for convolution layers have large activation map. The overall architecture is summarized in figure 3 and table 2.

Strategy one and strategy two used in SqueezeNet reduces number of parameters in a Convoluted Neural Network while maintaining accuracy.

The third strategy will then complement the previous two strategies to maximizing accuracy from the reduced parameters from strategy one and strategy two.

Table 2: Architecture of SqueezeNet [14]

Layer	Filters/ neurons	Filter size	Stride	Padding	Size of feature map	Activation function
Input	-	-	-	-	224*224*3	-
Conv one	96	7*7	2	-	111*111*96	ReLU
Max Pool 1	-	3*3	2	-	55*55*96	-
Fire2	-	-	-	-	55*55*128	ReLU
Fire3	-	-	-	-	55*55*128	ReLU
Fire4	-	-	-	-	55*55*256	ReLU
Max Pool 4	384	3*3	2	-	27*27*256	-
Fire5	-	-	-	-	27*27*256	ReLU
Fire6	-	-	-	-	27*27*384	ReLU
Fire7	-	-	-	-	27*27*384	ReLU
Fire8	-	-	-	-	27*27*512	ReLU
Max Pool 8		3*3	2	-	13*13*512	-
Fire9	-	-	-	-	13*13*512	ReLU
Conv10	1000	1*1	1	-	13*13*1000	ReLU
AvgPool10				-	1*1*1000	-

Basic-CNN Models (6 layers)

As part of this study, we also created a CNN model to help us categorize the distinct types of noodles in Malaysia like Koay Teow, Maggie, Mee, Laksa and Bee Hoon. We create two basic-CNN models with 6 layers and 7 layers of convolutional which inspired from the Alexnet Architecture.

Table 3: Architecture of Basic-CNN models (6 layers)

Layer	Filters/ neurons	Filter size	Stride	Padding	Size of feature map	Activation function
Input	-	-	-	-	227*227*3	-
Conv 1	96	11*11	4	-	55*55*96	ReLU
Max Pool 1	-	3*3	2	-	27*27*96	
Conv 2	256	5*5	1	2	27*27*256	ReLU
Max Pool 2	-	3*3	2	-	13*13*256	-
Conv 3	384	3*3	1	1	13*13*384	ReLU
Conv 4	384	3*3	1	1	13*13*384	ReLU
Conv 5	256	3*3	1	1	13*13*256	ReLU
Max Pool 3	-	3*3	2	-	6*6*256	-
Conv 6	256	3*3	1	1	13*13*256	ReLU
Max Pool 4	-	3*3	2	-	6*6*256	-
Dropout 1	Rate = 0.5	-	-	-	6*6*256	-

© 2024 IJSRET

For the first basic-CNN models (6 layers), the first 5 layers in the Alexnet is freeze and added an extra convolutional layer using 256 filters and 3*3 filter size with stride 1 and padding 1. The feature map is with the dimension of 13*13*256 and activation function ReLU. Furthermore, the fourth max pooling is added with filter size of 3*3, stride 2 and feature map of 6*6*256. Adding these convolutional layers and max pooling based on Alexnet Architecture is the Basic-CNN models (6 layers) developed by us. Table 3 shows the architecture of Basic-CNN models (6 layers)

Basic CNN-models (7 layers)

The first five Alexnet layers for the initial basic-CNN models (7 layers) were frozen, then an additional convolutional layer was inserted as 6th layers using 256 filters and a 3*3 filter size with stride 1 and padding 1. The 13*13*256-pixel feature map has the activation function ReLU. The addition of the fourth maximum pooling with a filter size of 3*3, stride 2, and a feature map of 6*6*256 is made. The convolutional layers and max pooling will then duplicate again as the 7th layer. This Basic-CNN models (7 layers) that we constructed added with these convolutional layers and max pooling are based on Alexnet Architecture. Table 4 shows the architecture of Basic-CNN models (7 layers)

Table 4: Architecture of Basic-CNN models (7 layers)

Layer	Filters/ neurons	Filter size	Stride	Padding	Size of feature map	Activation function
Input	-	-	-	-	227*227*3	-
Conv 1	96	11*11	4	-	55*55*96	ReLU
Max Pool 1	-	3*3	2	-	27*27*96	
Conv 2	256	5*5	1	2	27*27*256	ReLU
Max Pool 2	-	3*3	2	-	13*13*256	-
Conv 3	384	3*3	1	1	13*13*384	ReLU
Conv 4	384	3*3	1	1	13*13*384	ReLU
Conv 5	256	3*3	1	1	13*13*256	ReLU
Max Pool 3	-	3*3	2	-	6*6*256	-
Conv 6	256	3*3	1	1	13*13*256	ReLU
Max Pool 4	-	3*3	2	-	6*6*256	-
Conv 7	256	3*3	1	1	13*13*256	ReLU
Max Pool 5	-	3*3	2	-	6*6*256	-
Dropout 1	Rate = 0.5	-	-	-	6*6*256	-

2. Dataset and Data Augmentation

A dataset in computer vision is a curated set of digital photographs that developers use to test, train, and evaluate the performance of their algorithms. The algorithm is said to learn from the examples contained in the dataset. In our model, we had gathered total of 13,376 images with five types of noodles.

The types of noodles are Mee, Mee Bee Hoon, Mee Koay Teow, Mee Laksa and Mee Maggi. We also gather the difference type of cooking which make the classification more diversify by Goreng and Sup.

The method of collecting the dataset is by real life photos, google image, Facebook and Instagram. After the collection of the basic dataset, we did some prepossessing of the image by resize to 224 x 224.

Jason Dsoaza mentioned “One of the easiest ways to increase validation accuracy is to add more data.”[16] Hence, we had generate more data image to train the model by mirror, 90 degree rotation and 180 degree rotation. We also standardize the image size by resize to 224x224.

Figures 15 shows the samples image of our dataset after processing.

Figure 15: Dataset Images after processing

Table 5: The total image collected in dataset

Classification	Number of Image
Mee Sup	1968
Mee Goreng	1848
Mee Bee Hoon Sup	1086
Mee Bee Hoon Goreng	1496
Mee Koay Teow Sup	1386
Mee Koay Teow Goreng	1350
Mee Laksa Sup	1560
Mee Laksa Goreng	762
Mee Maggi Sup	588
Mee Maggi Goreng	1332

© 2024 IJSRET

The total image we finally collected are as Table 5. However, due to the high number of augmented images generated. The dataset with 4334 images were used to train the models due to limitation of computation power and time we had. The large dataset which contains more than 13 thousands image will be used as part of the experiment to simulate real world data.

IV. PERFORMANCE RESULT

1. Experimental Setup

Experiments performed based on four models trained using 4334 datasets with ten classes. Experiment one, the model created were evaluated with 500 selected datasets.

The results from the experiment will be used as a basis to evaluate the performance of the four proposed models' ability to classify ten different classes of noodles.

Experiment two, the models created were evaluated with 13 thousand images which is generated from data augmentation. The result of the experiments is to identify the performance of the models from a larger image dataset, which is used to simulate real world data from pseudo-public sets.

Experiment three, the models created were evaluate with the original 4334 dataset to serve as a control experiment. The results are used as a control experiment, ensuring that the model used are using same or similar dataset.

Lastly, performance of models from experiment one are analyzed with confusion matrix and classification reports. From the experiment, the result will provide insight on how to improve the models proposed in future works.

2. Training Models

To train the model, the dataset used are based on table, has 4334 images which are defined by ten different classes of noodles.

The training of the models is completed once accuracy 1.0 and loss less than 0.03 are achieved during the training.

Once, the completion of the training condition is achieved, the bias, weight and the model will be saved as model which has completed training.

The table shows the numbers of epochs are used is less than 20 for all the model used.

Figure 16 and 17 shows the loss and the accuracy over the epoch during the training.

The table 6 shows overall performance of model training and at which epoch the model training is completed.

Table 6: Models trained against Loss and Accuracy

Epochs	Type of CNN							
	Alexnet		Squeeze Net		Basic-CNN (6 layers)		Basic-CNN (7 layers)	
	Loss	Accuracy	Loss	Accuracy	Loss	Accuracy	Loss	Accuracy
1	2.2126	0.3646	2.1141	0.3743	2.1002	0.356	1.941	0.3544
2	1.1308	0.6089	1.247	0.5464	1.3251	0.5238	1.4118	0.4885
3	0.6377	0.7937	0.9093	0.6786	0.9588	0.6652	1.1741	0.5865
4	0.3079	0.919	0.5954	0.8126	0.6487	0.7972	0.9252	0.6721
5	0.1322	0.9746	0.3201	0.91	0.3743	0.9042	0.7149	0.758
6	0.0619	0.9919	0.1502	0.9626	0.2011	0.9587	0.477	0.8602
7	0.0377	0.9958	0.0765	0.9827	0.107	0.9827	0.2949	0.9204
8	0.0212	0.9958	0.0467	0.9912	0.0634	0.9938	0.164	0.9612
9	0.0156	0.9993	0.0439	0.9892	0.0424	0.9958	0.0844	0.9848
10	0.0117	0.9995	0.0201	0.997	0.0259	0.9984	0.0562	0.9903
11	0.0116	0.9993	0.016	0.9982	0.0145	0.9998	0.0265	0.9975
12	0.0098	0.9991	0.0086	0.9995	0.0133	0.9991	0.0173	0.9984

© 2024 IJSRET

13	0.0073	0.9995	0.0074	0.9991	0.0097	0.9995	0.0075	0.9995
14	0.0061	1	0.0056	0.9998	0.0084	1	0.0079	0.9995
15			0.0212	0.997			0.0071	0.9993
16			0.0042	1			0.0061	0.9998
17							0.0046	0.9995
18							0.0029	1

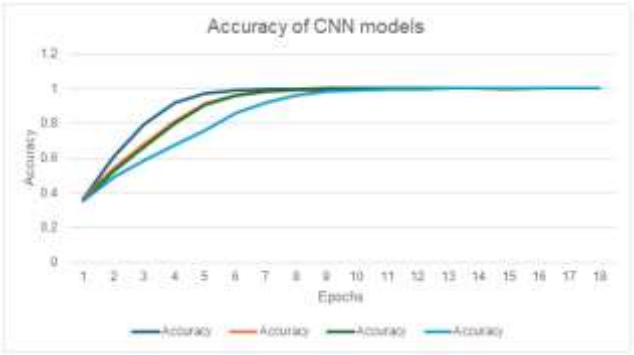

Figure 16: Accuracy of CNN models over numbers of Epochs

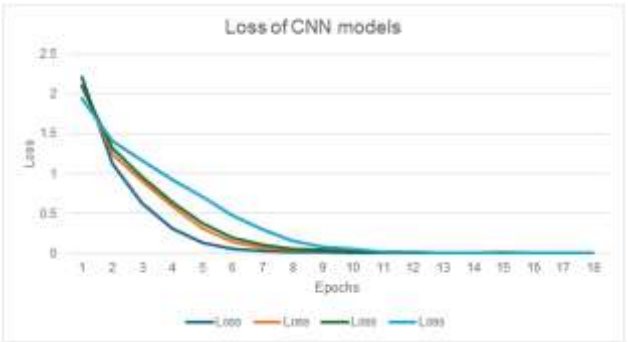

Figure 17: Loss of CNN models over numbers of Epochs

3. Experimental Results

Experiment one result

The results from Experiment 1 based on table 7 and figure 18. The Basic-CNN 7 layers provides the highest accuracy compared to other models.

Table 7: Test set accuracy based on fifty pictures per class

Test Set Accuracy (50 pic per class)			
Alexnet	SqueezeNet	Basic-CNN (6 layers)	Basic-CNN (7 layers)
82.80%	83.00%	83.40%	84.20%

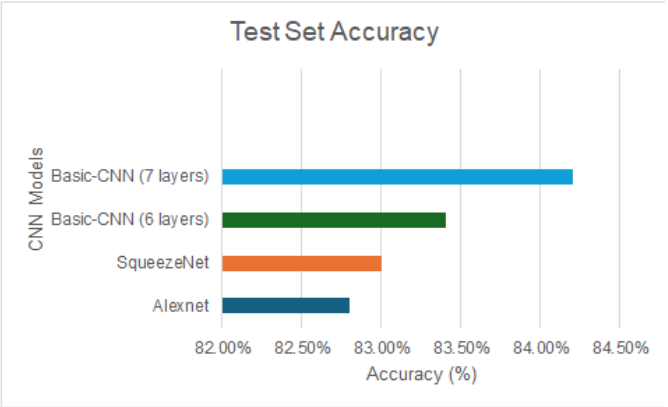

Figure 18: Test Set Accuracy of Experiment 1

4. Experiment Two Result

The results as shown in table 8 and figure 19 from the result of experiment two shows that SqueezeNet model perform better than the rest of the model. By using large dataset of 13000 the experiment shows that the model Basic-CNN (7 layers) has less accuracy than SqueezeNet. The performance of the Basic-CNN(7 layers) performed better than the AlexNet model.

Table 8: Test set accuracy based on 13376 images belonging to ten classes unevenly distributed

Large Test Set Accuracy (13376 images belonging to ten classes unevenly distributed)			
Alexnet	SqueezeNet	Basic-CNN (6 layers)	Basic-CNN (7 layers)
74.61%	80.52%	74.61%	77.11%

Figure 19: Test Set Accuracy of Experiment 2

5. Experiment Three Result

Based on table 9 and figure 20, the model has accuracy more than 99.9%. This validates the dataset used in the models are the same.

© 2024 IJSRET

Table 9: Test set accuracy based on 4334 images used to train the models.

Reuse Training Set Accuracy (4334 images belonging to ten classes unevenly distributed)			
Alexnet	SqueezeNet	Basic-CNN (6 layers)	Basic-CNN (7 layers)
100.00%	100.00%	99.97%	100.00%

Figure 20: Test Set Accuracy of Experiment 3

Model Evaluation

Based on the f1-score from the classifications reports for the models as specified in table 10, table 11, table 12 and table 13. The classification reports that the class "Mee Bee Hoon Goreng" and "Mee Koay Teow Goreng" f1-score is consistently below value of 0.8 across the four models.

The four models used to classify noodles, has shown consistently "Mee Bee Hoon Goreng" is incorrectly label by all the models as other classes. The recall values of the "Mee Bee Hoon Goreng" which is 0.44 in table 10, 0.5 in table 11, 0.6 in table 12, and 0.54 in table 13.

The class or label "Mee Koay Teow Goreng" is in correctly labeled as other classes. The recall values of "Mee Koay Teow Goreng" which is 0.3 in table 10, 0.18 in table 11, 0.24 in table 12, and 0.18 in table 13.

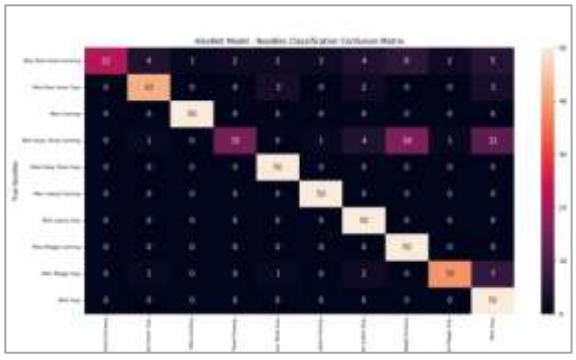

Figure 21: AlexNet Noodles Classification Confusion Matrix

The low recall score for class "Mee Bee Hoon Goreng" and "Mee Koay Teow Goreng" shows all the four models is have deficient performance to identify both "Mee Bee Hoon Goreng" and "Mee Koay Teow Goreng" correctly.

However, the four models proposed are able to classify "Mee Goreng", "Mee Koay Teow Sup", "Mee Laksa Goreng", "Mee Laksa Sup", "Mee Maggi Goreng" and "Mee Sup" has 0.80 in as shown in table 9, table 10, table 11, and table 12.

Table 10: AlexNet Model Classification Report

AlexNet Model - Noodles Classification Report				
	precision	recall	f1-score	support
Mee Bee Hoon Goreng	1	0.44	0.61	50
Mee Bee Hoon Sup	0.88	0.84	0.86	50
Mee Goreng	0.98	1	0.99	50
Mee Koay Teow Goreng	0.88	0.3	0.45	50
Mee Koay Teow Sup	0.89	1	0.94	50
Mee Laksa Goreng	0.94	1	0.97	50
Mee Laksa Sup	0.81	1	0.89	50
Mee Maggi Goreng	0.69	1	0.82	50
Mee Maggi Sup	0.93	0.78	0.85	50
Mee Sup	0.65	1	0.79	50
accuracy			0.84	500
macro avg	0.87	0.84	0.82	500
weighted avg	0.87	0.84	0.82	500

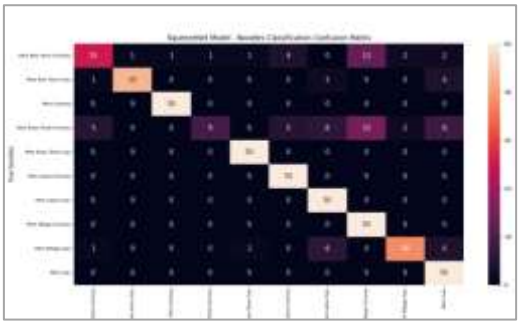

Figure 22: SqueezeNet Noodles Classification Confusion Matrix

Table 11: SqueezeNet Model Classification Report

SqueezeNet Model - Noodles Classification Report				
	precision	recall	f1-score	support
Mee Bee Hoon Goreng	0.78	0.5	0.61	50
Mee Bee Hoon Sup	0.98	0.86	0.91	50
Mee Goreng	0.98	1	0.99	50
Mee Koay Teow Goreng	0.9	0.18	0.3	50
Mee Koay Teow Sup	0.96	1	0.98	50
Mee Laksa Goreng	0.85	1	0.92	50
Mee Laksa Sup	0.78	1	0.88	50
Mee Maggi Goreng	0.64	1	0.78	50
Mee Maggi Sup	0.9	0.76	0.83	50
Mee Sup	0.74	1	0.85	50
accuracy			0.83	500
macro avg	0.85	0.83	0.8	500
weighted avg	0.85	0.83	0.8	500

© 2024 IJSRET

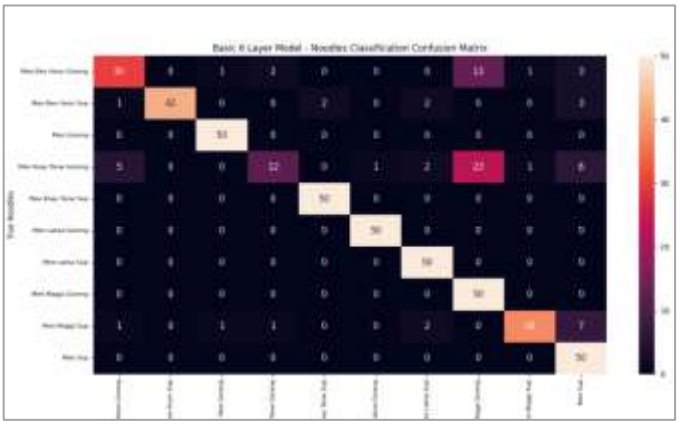

Figure 23: Basic-CNN (6 Layers) Noodles Classification Confusion Matrix

Table 12: Basic-CNN (6 Layers) Model Classification Report

Basic-CNN 6 Layer Model - Noodles Classification Report				
	precision	recall	f1-score	support
Mee Bee Hoon Goreng	0.81	0.6	0.69	50
Mee Bee Hoon Sup	1	0.84	0.91	50
Mee Goreng	0.96	1	0.98	50
Mee Koay Teow Goreng	0.8	0.24	0.37	50
Mee Koay Teow Sup	0.96	1	0.98	50
Mee Laksa Goreng	0.98	1	0.99	50
Mee Laksa Sup	0.89	1	0.94	50
Mee Maggi Goreng	0.58	1	0.74	50
Mee Maggi Sup	0.95	0.76	0.84	50
Mee Sup	0.72	1	0.84	50
accuracy			0.84	500
macro avg	0.87	0.84	0.83	500
weighted avg	0.87	0.84	0.83	500

Figure 24: Basic-CNN (7 Layers) Noodles Classification Confusion Matrix

Table 13: Basic-CNN (7 Layers) Model Classification Report

Basic 7 Layers Model - Noodles Classification Report				
	precision	recall	f1-score	support
Mee Bee Hoon Goreng	0.87	0.54	0.67	50
Mee Bee Hoon Sup	0.93	0.86	0.9	50
Mee Goreng	0.96	1	0.98	50
Mee Koay Teow Goreng	0.9	0.18	0.3	50
Mee Koay Teow Sup	0.94	1	0.97	50
Mee Laksa Goreng	0.86	1	0.93	50
Mee Laksa Sup	0.85	1	0.92	50
Mee Maggi Goreng	0.62	1	0.77	50
Mee Maggi Sup	0.93	0.76	0.84	50
Mee Sup	0.71	1	0.83	50
accuracy			0.83	500
macro avg	0.86	0.83	0.81	500
weighted avg	0.86	0.83	0.81	500

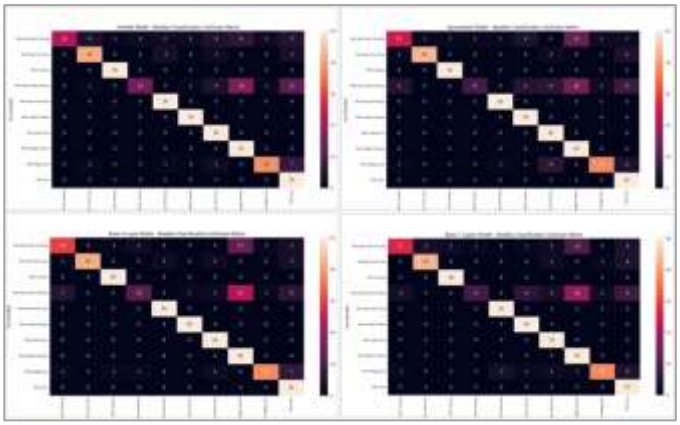

Figure 25: Comparison of confusion matrix of 4 Models

V. CONCLUSION

Figure 25 show the performance of all the four models is similar. However, from the figure 7, with controlled and proper sanitized data, basic-CNN 7 layers can provide the best performance and accuracy despite less complicate with its counterpart the AlexNet and SqueezeNet. When using large datasets to simulate a larger population, SqueezeNet model shows better performance as shown in figure 8. SqueezeNet provides a better accuracy potential. Hence, SqueezeNet can use to classify real data.

Despite the accuracy of Basic-CNN 7 layers and the SqueezeNet has shown from all the experiments. Both models

© 2024 IJSRET

suffer loss of accuracy to identify two classes, which is “Mee Bee Hoon Goreng” and “Mee Koay Teow Goreng.”

The suggestion to improve the model are, first train the model with a larger dataset more than 4334 images with SqueezeNet model. Then, create a transfer learning from the SqueezeNet trained on larger noodles data set, remove the fully connected layers. Freeze the SqueezeNet model, and further tune the new model with additional layers of convolution. As a conclusion, with controlled dataset the Basic-CNN provides accurate prediction and classification of noodles. When a large noodles dataset is used to be classified, SqueezeNet can provide better prediction and classification of noodles. Transfer learning from SqueezeNet is recommended to improve noodles classification deep learning model.

REFERENCES

1. Lu, Z.-H., & Collado, L. S. (2019). Rice noodles. Rice, 557–588. https://doi.org/10.1016/b978-0-12-811508-4.00017-4
2. Walch, K. (2021, September 17). How neural network training methods are modeled after the human brain. SearchEnterpriseAI. Retrieved July 16, 2022, from https://www.techtarget.com/searchenterpriseai/feature/How-neural-network-training-methods-are-modeled-after-the-human-brain
3. Kagaya, H., Aizawa, K., & Ogawa, M. (2014). Food detection and recognition using convolutional neural network. Proceedings of the 22nd ACM International Conference on Multimedia. https://doi.org/10.1145/2647868.2654970
4. Shimoda, W., & Yanai, K. (2015). CNN-based food image segmentation without pixel-wise annotation. New Trends in Image Analysis and Processing -- ICIAP 2015 Workshops, 449–457. https://doi.org/10.1007/978-3-319-23222-5_55
5. Alake, R. (2020, July 10). What alexnet brought to the world of Deep learning. Medium. Retrieved July 16, 2022, from https://towardsdatascience.com/what-alexnet-brought-to-the-world-of-deep-learning-46c7974b46fc
6. Ozsert Yigit, G., & Ozyildirim, B. M. (2017). Comparison of convolutional neural network models for Food Image Classification. 2017 IEEE International Conference on INnovations in Intelligent SysTems and Applications (INISTA). https://doi.org/10.1109/inista.2017.8001184
7. Pandey, P., Deepthi, A., Mandal, B., & Puhan, N. B. (2017). FoodNet: Recognizing foods using ensemble of Deep Networks. IEEE Signal Processing Letters, 24(12), 1758–1762. https://doi.org/10.1109/lsp.2017.2758862
8. F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, “SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size,” Feb. 2016, [Online]. Available: http://arxiv.org/abs/1602.07360
9. M. Hassanpour and H. Malek, “Learning document image features with SqueezeNet convolutional neural network,” International Journal of Engineering, Transactions A: Basics, vol. 33, no. 7, pp. 1201–1207, Jul. 2020, doi: 10.5829/ije.2020.33.07a.05.
10. S. Yadav, Alpana, and S. Chand, “Automated Food image Classification using Deep Learning approach,” in 2021 7th International Conference on Advanced Computing and Communication Systems, ICACCS 2021, Mar. 2021, pp. 542–545. doi: 10.1109/ICACCS51430.2021.9441889.
11. M. BAYĞIN, “Vegetable and Fruit Image Classification with SqueezeNet based Deep Feature Generator,” Turkish Journal of Science and Technology, Mar. 2022, doi: 10.55525/tjst.1071338.
12. Simplified illustration of the AlexNet architecture. (2018). Research Gate. https://www.researchgate.net/figure/Simplified-illustration-of-the-AlexNet-architecture_fig2_329790469
13. Saxena, S. (2021, March 30). Introduction to The Architecture of Alexnet. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2021/03/introduction-to-the-architecture-of-alexnet/
14. SqueezeNet CNN architecture. (2016). Research Gate. https://www.researchgate.net/publication/306003694_A_fast_embedded_implementation_of_a_Convolutional_Neural_Network_for_Image_Recognition/figures?lo=1
15. Forrest N. Iandola, Song Han, Matthew W. Moskewicz, Khalid Ashraf, William J. Dally, Kurt Keutzer (2016, November 4). SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size. https://arxiv.org/abs/1602.07360v4
16. Jason Dsouza, How to Improve the Accuracy of Your Image Recognition Models, NOVEMBER 29, 2021/#DEEP LEARNING https://www.freecodecamp.org/news/improve-image-recognition-model-accuracy-with-these-hacks/

© 2024 IJSRET