Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12987
e-ISSN : 2541-2019
p-ISSN : 2541-044X

PyTorch Deep Learning for Food Image Classification with Food Dataset

Iswahyudi[1)], Djarot Hindarto[2)*], Handri Santoso[3]
[1,3)] Pradita University, Tangerang, Indonesia
[2)] Informatics Study Program, Faculty of Communication and Informatics Technology, Universitas Nasional, Jakarta, Indonesia
[1)] iswahyudi@student.pradita.ac.id, [2)*] djarot.hindarto@civitas.unas.ac.id, [3)] handri.santoso@pradita.ac.id

Submitted : Sep 9, 2023 | **Accepted** : Sep 10, 2023 | **Published** : Oct 1, 2023

Abstract: Classification of food images is crucial in today's increasingly digitally connected world. In the rapidly evolving mobile applications and social media era, the demand for an automated system that can recognize food types from an image is intensifying. This study employs deep learning and the PyTorch framework to develop a dependable and efficient solution for classifying food images. This research is motivated by the growing complexity of food introduction challenges. The primary challenge is improving the accuracy of food type recognition and overcoming variations in the visual presentation of food, such as lighting, shooting angles, and proportional and textural differences. Convolutional Neural Networks (CNN) are effective for image classification and are incorporated into the methods utilized. In addition, we employ ResNet101 transfer learning techniques to capitalize on the knowledge of trained models for large image datasets. The primary objective of this study is to develop a food image classification model that is accurate, training-efficient, and capable of accurately recognizing various types of food. In testing and evaluation, the developed model could realize multiple types of food with satisfactory accuracy. The accuracy of training reached 99.35%, while the accuracy of testing reached 94.65%. This study also reveals how Resnet101 transfer learning is utilized by deep learning technology.

Keywords: Convolutional Neural Networks; Classification; Food Image; PyTorch; Image Datasets

INTRODUCTION

In the increasingly advanced digital era, the recognition and classification of food images have become an essential issue in image processing and artificial intelligence (Ozbayoglu & Yuksel, 2012), (Dhanush et al., 2023), (Zimmermann et al., 2021)—the rapid growth of mobile applications, social media, and image technologies. Modern society frequently shares food experiences via online platforms. However, the issue that often arises is correctly identifying the food type from the uploaded image. In the fields of image processing and artificial intelligence, food image recognition is an integral component. They use algorithms and machine learning models to identify and classify food types in images. Variation in food appearance caused by lighting, shooting angle, decoration, and presentation is one of the most challenging aspects of identifying food from photographs. This issue is significant not only from a technical and scientific standpoint but also from a social perspective. Automatically identifying food types from images can enhance diet management, facilitate calorie counting, and enable more interactive image-based services in culinary applications.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v8i4.12987 p-ISSN : 2541-044X

To address this issue, a significant amount of research has been conducted to develop models that can accurately identify food types. Sometimes, these models employ artificial neural networks (Swift et al., 2023), (Adedeji, 2023), (Czinege & Harangozó, 2023) trained on large and diverse datasets. This training data includes thousands of images of various types of cuisine, including Asian, Western, and fast food. Using this extensive training data, models can discover distinct visual patterns for each food type. In addition, image processing techniques such as feature extraction and object segmentation are employed to aid in identifying food varieties. Food's color, texture, and shape can be extracted from images via feature extraction. Object segmentation enables the model to separate the food from the image background, making identifying it easier.

In addition to social media and image-sharing platforms, this technology has applications in various other fields. In the food industry, for instance, food image recognition can be used to automate inventory management processes and monitor food production quality. This technology can also watch a patient's diet and assist with calorie counting in the health sector. Although significant progress has been made in recognizing and classifying food images, there are still obstacles to overcome. One is when food photographs are taken with poor lighting or odd angles. Moreover, regional differences in food presentation can also be problematic.

Using deep learning (Saputra et al., 2023), focusing on the PyTorch framework, this study aims to overcome this challenge by developing a reliable food image classification model. This research is motivated by the complexity of visual variations in food presentation, which is frequently challenging to overcome using conventional methods. We describe how we overcame this issue in this study using deep learning (Pribanić et al., 2023), (Djenouri et al., 2023) and PyTorch (Pérez-García et al., 2021). Explanation of the methods used to train and evaluate the performance of a food image classification model on multiple food image datasets. The objective is to produce a model capable of recognizing various types of food with high precision and training efficiency.

This study also provides a deeper understanding of PyTorch's capabilities in the context of image processing and image classification, which may inspire future research and development in this area. Descriptions of model training methods, testing, and performance evaluation on diverse food datasets are conducted carefully. The outcomes of this experiment will provide insight into how deep learning technology can be effectively applied to overcome the difficulty of classifying food images with a high degree of visual variation.

The introduction description raises the following research question: Does the performance of food recognition models using public and private food image datasets from specific food companies differ significantly? How can we use deep learning and the Food Image Dataset to improve food type recognition from images?

LITERATURE REVIEW

The recognition and classification of food images have become an essential issue in image processing and artificial intelligence in this increasingly digital era. To overcome these obstacles, using PyTorch, a robust deep learning framework, has emerged as a crucial component in developing models capable of accurately recognizing and classifying food types. This literature review will explore using PyTorch in food image recognition by employing the Food Dataset, a food-centric dataset. This literature review will describe significant advancements in prior research, new techniques, and the challenges and opportunities in the quest for more accurate and efficient food image recognition. Cervical cancer is a significant issue for women's health in Indonesia, where 15,000 cases are reported annually. The accuracy of Pap smear tests and ResNet50/ResNet101 architectures for detecting and classifying cervical cancer was 91% and 89%, respectively (Niswati et al., 2021). However, due to a lack of training data, the complexity of ResNet101 did not produce better results. This study evaluated the model in average, early COVID-19, and severe COVID-19 patients using the COVID-19 CT dataset (Chena et al., 2022). The experimental results demonstrate that Inception-Resnet is more accurate than other image classification methods by 89.13%. This research aims to develop a system for gender identification based on periocular images using a CNN (Hindarto & Santoso, 2021) model trained with a Transfer Learning technique. On the UBIPr dataset, this method achieved average precisions of 98.65%, 98.96%, and 98.99% using the basic models VGG19,

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12987
e-ISSN : 2541-2019
p-ISSN : 2541-044X

ResNet101, and ResNet50, respectively (Kumar & Seeja, 2023). Due to the COVID-19 pandemic, this study aims to address the problem of gender identification when wearing face masks. Osteosarcoma, a type of bone cancer frequently affecting long bones, is caused by DNA mutations. Imaging and X-rays are necessary for early detection. This study automates the time-consuming biopsy procedure using deep learning, achieving an accuracy of 90.36 percent with ResNet101 (Gawade et al., 2023) in predicting bone cancer. The existing body of research about accuracy has yet to yield satisfactory results, thus creating a gap in the current knowledge base. This study aims to enhance the precision of the ResNet101 model, which has previously been unable to attain a high level of accuracy in existing research. Consequently, this study aims to enhance the level of precision.

METHOD

Food Image Dataset

The Food Image Dataset, a substantial compilation of food images, is widely employed in diverse research endeavors and practical applications within image recognition, image processing, and artificial intelligence. The provided dataset plays a crucial role in facilitating the training and evaluation of machine learning models in food image recognition, enabling tasks such as identification, classification, and even prediction of food types. Food image datasets with varying sources can be categorized into two main groups: public and private. Public datasets are accessible to the general public without any cost and are frequently employed in scholarly investigations and the advancement of artificial intelligence models. An example of a widely recognized public dataset is Food-101, comprising over 100,000 food images categorized into 101 distinct food categories. Datasets such as Food-101 offer diverse food categories, enabling models to acquire enhanced knowledge about culinary items from various cultural and gastronomic traditions.

Furthermore, it should be noted that there are other publicly available datasets, namely UEC Food-256 and MIT Food101-201, which possess a substantial quantity of images suitable for comprehensive training purposes. In addition to publicly available datasets, there exist proprietary food datasets. Private datasets are typically under the ownership of specific organizations or companies and are frequently withheld from public access due to commercial or confidential considerations. Food companies and restaurants possess the capacity to curate their own proprietary food image datasets for internal purposes, including menu development and consumer trend analysis. Typically, researchers or developers external to the dataset's owning organization need more access to these datasets.

Fig.1 Food dataset
Source: Kaggle

Fig.1, The significance of possessing extensive and high-caliber private datasets lies in their capacity to contain specific product information that is not accessible within public datasets. This

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12987
e-ISSN : 2541-2019
p-ISSN : 2541-044X

dataset includes significant value for companies, enabling them to enhance customer experience and optimize operational efficiency. Nevertheless, it is imperative to acknowledge that using private datasets must adhere to relevant data privacy and copyright regulations. The Food Image Dataset is pivotal in advancing food recognition technology and processing food images. Public datasets play a crucial role in fostering innovation and facilitating collaboration among researchers on a global scale. Conversely, private datasets serve as valuable resources for companies seeking to enhance their offerings and optimize their services. Integrating these two elements can expedite the progress of food-related applications, encompassing allergen identification and tailored food suggestions. This convergence can enhance the gastronomic encounter and propel the culinary sector.

Food Image Classification

Food Image Classification is a subfield within the image processing domain that endeavors to identify and categorize diverse food types by analyzing the visual content of images or photographs. The primary objective of Food Image Classification is to develop an automated system capable of accurately identifying the various types of food depicted in paintings, eliminating the need for human involvement. The significance of Food Image Classification lies in its diverse range of practical applications. One of the primary applications lies within the culinary sector, wherein restaurants and food delivery platforms frequently leverage this technology to enable users to explore and place orders for food items using visual content. Furthermore, Food Image Classification can be applied to the domains of diet monitoring and food management, facilitating individuals to compute their caloric consumption and identify foods that align with their dietary preferences. Moreover, within the realm of health, Food Image Classification can be employed to identify food items that may contain specific allergens or potentially harmful components.

The process of Food Image Classification encompasses multiple stages. The initial phase involves the acquisition of a dataset comprising images of food items that have been annotated with corresponding categorizations based on their respective types. The addition of this dataset is imperative for training the machine learning model, which will subsequently be employed for the classification of images. Once the dataset has been gathered, the subsequent phase involves preparing the model utilizing diverse machine learning algorithms, including Convolutional Neural Networks (Sze et al., 2022), demonstrating their efficacy in image processing. The training procedure encompasses acquiring knowledge from exemplar images to discern distinctive patterns and unique characteristics that distinguish various food items. Following the completion of the training process, the subsequent phase involves examining and assessing the model. The model undergoes testing using an independent dataset distinct from the one employed for training to evaluate its efficacy in food classification. The test results are utilized to assess the precision and effectiveness of the model. If improvements are deemed necessary, the model can be enhanced through additional training or parameter adjustments. In summary, Food Image Classification is a significant domain within image processing, possessing numerous practical implications in industries such as culinary arts, healthcare, and dietary management. The progress of machine learning technology has led to an enhanced capacity of systems to identify and categorize food items depicted in images. This development has proven beneficial to individuals in multiple domains of their everyday existence that pertain to food.

Deep Learning

Deep learning, a subfield of artificial intelligence (AI), has garnered significant attention in recent years owing to its remarkable capacity to address intricate challenges. Deep understanding is founded upon neural networks that draw inspiration from the complex organization of the human brain. Compared to conventional approaches, one distinguishing characteristic of deep learning is its capacity to autonomously acquire meaningful data representations, progressing from a lower level of features to a higher level. This capability facilitates a more profound comprehension of intricate datasets. Deep learning utilizes artificial neural networks that are composed of multiple interconnected layers. Every layer comprises several artificial neurons or units dedicated to information processing. Every individual neuron within a neural network receives a specific amount of input, subsequently multiplied by the corresponding weight assigned to it. Following this, the neuron undergoes an

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12987
e-ISSN : 2541-2019
p-ISSN : 2541-044X

activation function, resulting in the generation of an output. The layers mentioned above establish a hierarchical structure in which each layer exhibits increased abstraction concerning the underlying data. This implies that the initial processing layer may be dedicated to analyzing essential visual elements such as lines and borders. In contrast, the final layer may be able to comprehend intricate concepts such as human faces or objects depicted in an image.

Deep neural networks (DNN), alternatively referred to as deep feedforward networks, are widely recognized as one of the most prevalent deep learning methodologies. Deep neural networks (DNNs) typically have multiple hidden layers between the input and output layers. While the training of deep neural networks (DNNs) is challenging in practice due to the issue of missing or burst gradients, the development of various techniques, including proper initialization, utilization of non-linear activation functions like Rectified Linear Units (ReLU), and optimization algorithms such as stochastic gradient descent (SGD), has significantly enhanced the effectiveness and efficiency of deep learning.

One of the primary benefits of deep learning is its capacity to effectively handle extensive and intricate datasets, encompassing various forms of information such as images, text, and sound. These advancements have led to notable progress in diverse disciplines, such as computer vision (specifically in image and video recognition), natural language processing (mainly in machine translation and text comprehension), speech recognition, and the emergence of autonomous vehicles. The efficacy of deep learning is primarily attributed to ample data and substantial computational resources, enabling artificial neural networks to acquire intricate and abstract patterns from data. In artificial intelligence and technology, deep learning has ushered in significant transformations. The capacity of this technology to address previously challenging or seemingly insurmountable problems has facilitated the emergence of novel advancements and boundless practical implementations. Consequently, it remains an area of research that is both highly stimulating and experiencing rapid expansion within scientific and industrial spheres.

PyTorch

PyTorch is a framework for deep learning processing that has gained significant popularity and is known for its innovative features. PyTorch, created by Facebook's AI Research (FAIR) in 2016, offers a comprehensive set of tools enabling researchers and engineers to develop, train, and deploy artificial intelligence models quickly and efficiently. PyTorch distinguishes itself from other frameworks through its dynamic tensor computing approach, enabling users to construct neural networks imperatively, akin to conventional Python code composition. Users widely appreciate the flexible workflow of PyTorch as one of its key features. This implies that users can modify the neural network's architecture, conveniently retrieve tensor values, and dynamically compute gradients while constructing their models. This capability proves particularly advantageous in model exploration and rapid prototyping and tackling intricate and ever-changing challenges like text processing or computer vision tasks. In addition to its inherent flexibility, PyTorch is widely recognized for its exemplary documentation and vibrant community engagement. PyTorch offers comprehensive documentation, tutorials, and a wealth of online resources, facilitating the acquisition and proficiency of users across various skill levels. Furthermore, the PyTorch community actively contributes valuable packages and extensions that enhance the functionality of PyTorch for diverse applications. PyTorch is a versatile framework encompassing many artificial intelligence tasks, such as image processing, speech recognition, natural language processing, and other related charges. The platform provides a comprehensive library that includes modules such as PyTorch Vision, PyTorch Audio, and PyTorch Text, which enhance the model development process tailored explicitly for these tasks. Moreover, PyTorch offers robust GPU support, facilitating substantial speedup in computations, thereby rendering it well-suited for training intricate and extensive datasets.

In addition to its application in research, PyTorch is frequently employed in production settings. PyTorch offers a range of resources for transforming trained models into a compatible format suitable for deployment in diverse production settings, encompassing mobile devices and embedded systems. The PyTorch framework enables the execution of applications on various platforms with optimal performance. PyTorch is a robust and adaptable deep-learning framework that benefits from extensive community support. PyTorch remains a popular choice among researchers and engineers in

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v8i4.12987 p-ISSN : 2541-044X

developing artificial intelligence models due to its proficiency in managing intricate tasks and adaptable tensor computing methodology.

Pythorch transfer-learning
Utilizing the ResNet-101 model in PyTorch for transfer learning is a highly effective strategy for addressing a wide range of computer vision tasks encompassing image recognition. ResNet-101 is a highly intricate and efficacious Convolutional Neural Network (CNN) architecture that has undergone extensive training on expansive datasets, such as ImageNet, thereby equipping it with the capability to discern diverse objects within images. To implement transfer learning with ResNet-101, the first step is to load the pre-trained model using the PyTorch framework. The model can be readily obtained using the PyTorch library and importing its pre-trained weights and architecture. Upon loading the model, it is common practice to remove the top layers as they have already undergone training for ImageNet's image classification tasks. Subsequently, the uppermost layer is substituted with layers that are deemed suitable as see Fig. 2.

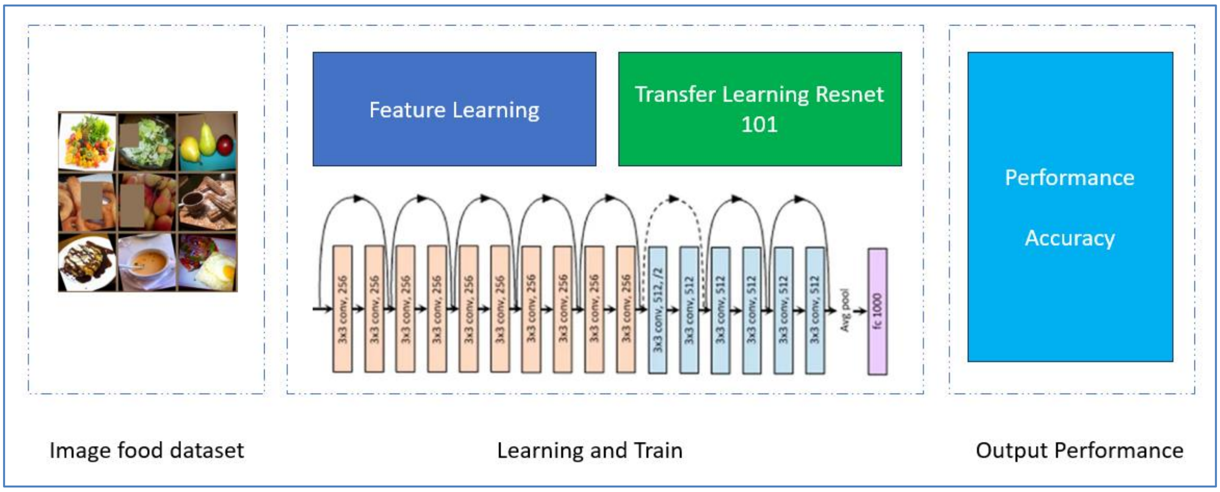

Fig. 2 Process Transfer Learning Resnet101
Source: Researcher Property

Performing image classification involves incorporating Dense (fully connected) layers above the final layer of ResNet-101. The number of units in the last layer corresponds to the desired categories. The inclusion of this component will be integrated into the customized model stack. The dataset should conform to a format deemed acceptable by the ResNet-101 model, typically consisting of images with specific dimensions—the process of partitioning a dataset into distinct subsets for training and testing. Once the dataset has been prepared, the next step involves commencing the model's training process—applying fine-tuning methodologies for training models using datasets. Throughout the training procedure, the model undergoes weight updates to align with the discernible patterns within the dataset while simultaneously preserving the pre-existing knowledge derived from the ResNet-101 that was previously trained. One of the primary benefits of utilizing ResNet-101 and the practice of transfer learning is its capacity to attain exceptional outcomes, even when working with datasets of limited size. ResNet-101 can hierarchically extract significant features from images, effectively handling diverse image types with varying degrees of complexity. Implementing this process using PyTorch is relatively straightforward and yields robust models for computer vision tasks.

RESULT

The present study employs Intel Core i9 hardware, 32GB of RAM, a 500GB SSD, and an Nvidia 3060 GPU. The presence of a CUDA device in this hardware facilitates the efficient training of the ResNet101 pretrain model.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12987
e-ISSN : 2541-2019
p-ISSN : 2541-044X

```
In [5]:  1 torchvision.models.resnet.resnet101(pretrained=True)

Downloading: "https://download.pytorch.org/models/resnet101-63fe2227.pth" to C:\Users\Djarot Hindarto
nts\resnet101-63fe2227.pth

  0%|          | 0.00/171M [00:00<?, ?B/s]

Out[5]: ResNet(
  (conv1): Conv2d(3, 64, kernel_size=(7, 7), stride=(2, 2), padding=(3, 3), bias=False)
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
  (maxpool): MaxPool2d(kernel_size=3, stride=2, padding=1, dilation=1, ceil_mode=False)
  (layer1): Sequential(
    (0): Bottleneck(
      (conv1): Conv2d(64, 64, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv3): Conv2d(64, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace=True)
      (downsample): Sequential(
        (0): Conv2d(64, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      )
    )
```

Fig. 3 Download model Resnet100
Source: Experiment Researcher

Fig. 3, this dataset contains 16643 images of food organized into 11 categories. This dataset includes three splits: evaluation, training, and validation. Each split contains eleven food categories: Bread, Dairy, Sweets, Egg, Fried Food, Meat, Noodles-Pasta, Rice, Seafood, Soup, and Vegetable-Fruit.

The process of training a model for a duration of 30 epochs is considered a crucial step in the fields of machine learning and deep learning. The computer model is provided with a training dataset to enhance its knowledge during each epoch. At the commencement of each epoch, the model generates a prediction by utilizing the provided training data. Subsequently, an error calculation is executed to quantify the disparity between the model's predictions and the target value. The provided information is utilized to compute the necessary gradient, which represents the alteration in the weights and bias of the model. The procedure above is commonly referred to as backpropagation. Subsequently, the optimization algorithm is employed to modify the weights and preferences, and this iterative process is repeated for 30 iterations. Over 30 epochs, the model exhibits a progressive enhancement in its performance, leading to a gradual reduction in the error associated with its predictions. The objective is to attain a notable degree of precision or favorable outcome in a specific undertaking, such as categorizing images or making data-based forecasts, which is heavily contingent upon the dataset and the problem being addressed. The optimal number of epochs may vary depending on the task's complexity and the point at which the model attains the desired outcomes. The observation is depicted in Figure 4.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12987
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Train model

```
In [19]:
%%time

num_epochs = 30
losses = []

for epoch in range(num_epochs):
    for i , (inputs, targets) in enumerate(train_dataloader):
        inputs = inputs.to(device)
        targets = targets.to(device)

        #train model
        outputs = model(inputs)

        #criterion
        loss = criterion(outputs, targets)
        losses.append(loss.item())

        #backward
        optimizer.zero_grad()
        loss.backward()

        #update parameters
        optimizer.step()

        # report
        if (i + 1) % 50 == 0:
            print('Epoch [%2d/%2d], Step [%3d/%3d], Loss: %.4f'
                  % (epoch + 1, num_epochs, i + 1, len(train_dataset) // batch_size, loss.item()))

Epoch [ 1/30], Step [ 50/308], Loss: 0.4860
Epoch [ 1/30], Step [100/308], Loss: 0.3983
Epoch [ 1/30], Step [150/308], Loss: 0.2800
Epoch [ 1/30], Step [200/308], Loss: 0.6852
Epoch [ 1/30], Step [250/308], Loss: 0.4329
Epoch [ 1/30], Step [300/308], Loss: 0.4549
Epoch [ 2/30], Step [ 50/308], Loss: 0.2502
Epoch [ 2/30], Step [100/308], Loss: 0.1076
Epoch [ 2/30], Step [150/308], Loss: 0.1832
```

Fig.4 Training Model 30 epoch
Source: Experiment Researcher

Nvidia graphics processing units (GPUs) exhibit exceptional performance in parallel computing, rendering them well-suited for a range of computationally intensive applications, including deep learning, graphics rendering, and data science. Nvidia graphics processing units (GPUs) provide rapid and effective processing capabilities, thereby enhancing the performance of applications that necessitate computationally demanding calculations. The observation is depicted in Figure 5.

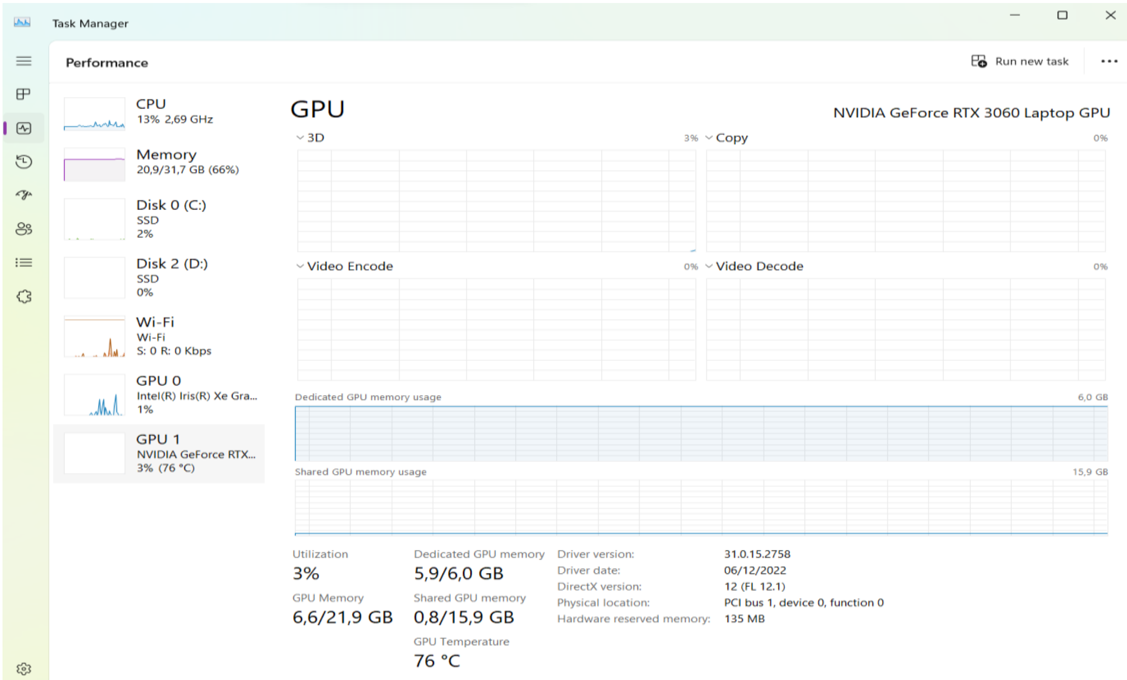

Fig. 5 Performance GPU Nvidia
Source: Experiment Researcher

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12987
e-ISSN : 2541-2019
p-ISSN : 2541-044X

The model's accuracy of 99.35% indicates a significantly high success rate in predicting data. This implies that the model consistently achieves high accuracy in accurately classifying the data. Nevertheless, it is imperative to carefully deliberate upon the implications of such elevated levels of precision. In certain instances, an imbalanced dataset can lead to a situation where one class is disproportionately represented, potentially resulting in high accuracy. Models can readily attain a high level of accuracy by making predictions based on the majority class without necessarily comprehending more intricate problems. Hence, assessing the model using additional metrics such as the confusion matrix or f1-score is crucial to obtain a more comprehensive comprehension of its genuine capability to handle diverse scenarios and categories within the dataset. The observation is depicted in Figure 6.

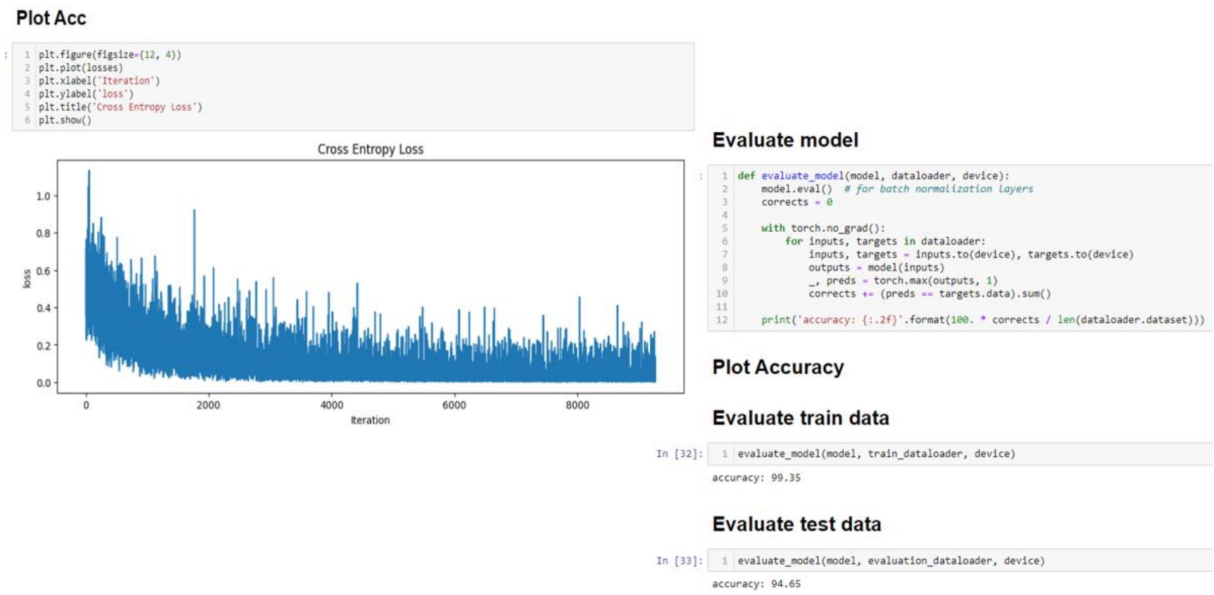

Fig. 6 Loss and Accuracy model 99,35%
Source: Experiment Researcher

DISCUSSIONS

Does the performance of food recognition models using public and private food image datasets from specific food companies differ significantly?

In image processing and artificial intelligence, comparing the performance of food recognition models using public and private food image datasets from a specific food company is an important issue. Public datasets are generally freely available and often used in academic research, while food companies own private datasets for commercial or confidential reasons. The question arises as to whether private datasets from food companies offer significant training advantages for food recognition models.

This research will identify potential quality and diversity differences between these datasets. Public datasets frequently contain foods from multiple sources, enabling models to discover more general patterns. On the other hand, private datasets may concentrate more on specific products or menus owned by the company, which can provide more detailed and specific information about these products. Private datasets can be advantageous in training food recognition models in multiple ways. First, the images in the private dataset are higher quality because more resources were used to shoot the product. Second, private datasets may offer more significant variation in presenting certain foods unavailable in public datasets. This can assist the model in identifying and differentiating products with similar appearances but significant differences.

However, there are challenges associated with using private datasets, particularly regarding data access and privacy. Due to economic and security concerns, food companies do not wish to share their data publicly. In addition, using private datasets can raise bias concerns, as the dataset may only

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12987
e-ISSN : 2541-2019
p-ISSN : 2541-044X

contain products from specific companies and not reflect broader dietary variations. In conclusion, comparing the performance of food recognition models on public and private food image datasets is an intriguing research question. This research will shed light on the potential benefits of private datasets in the context of food recognition. It can aid in developing more accurate models and applications about food. However, it is essential to consider data access and bias challenges when using private datasets from specific food companies.

How can we use deep learning and the Food Image Dataset to improve food type recognition from images?
Using deep learning and food image datasets to improve the recognition of food types from images is a highly effective strategy. An essential first step is collecting a large and diverse data set encompassing various foods from various cultures. The data must then be processed with image normalization and enhancement to ensure consistency and variation. Convolutional Neural Networks (CNNs) are frequently chosen as model architectures due to their effectiveness in image recognition tasks. Depending on the complexity of the dataset, ResNet can be used. During training, the model acquires visual patterns associated with the type of food, and training parameters such as learning rate and batch size must be adjusted with care. After training, evaluation of the model using the test dataset will provide an overview of the extent to which the model can recognize the type of food correctly. Fine-tuning and regularization techniques can be used to enhance performance if necessary.
Moreover, if data is scarce, transfer learning can be used to solve the problem by starting with a model trained on a larger dataset. To improve consumer experience and safety, the trained model can be implemented in various applications, including diet planning applications, restaurant systems, and food safety surveillance systems. With these steps, we can maximize the potential of using deep learning and food image data to recognize food types from images.

CONCLUSION

This study concludes that employing the Transfer Learning approach with the ResNet101 model yields superior results in the task of gender identification using periocular images. ResNet101 effectively extracts relevant features from periocular images by utilizing models pre-trained on visual studies, specifically object recognition on extensive datasets. The experimental findings indicate that ResNet101 attains an average accuracy of 99.35%, surpassing other baseline models such as VGG19 and ResNet50, which achieved 98.65% and 98.96%, respectively. The superior accuracy of ResNet101 can be understood as its capacity to discern gender characteristics in periocular images, even in challenging scenarios where a mask hinders facial visibility. The capability above exhibits significant potential in a diverse range of biometric applications, as well as in the realms of consumer identification and security systems. It is important to note that using ResNet101 may necessitate higher computational resources and longer training duration compared to alternative models. Hence, selecting a model requires careful consideration of the trade-off between accuracy and the resources at hand to achieve the most suitable applicability within a specific context.

REFERENCES

Adedeji, B. P. (2023). Electric vehicles survey and a multifunctional artificial neural network for predicting energy consumption in all-electric vehicles. *Results in Engineering*, *19*(July), 101283. https://doi.org/10.1016/j.rineng.2023.101283

Chena, Y., Lina, Y., Xua, X., Dinga, J., Li, C., Zenga, Y., Liue, W., Xie, W., & Huang, J. (2022). Classification of lungs infected COVID-19 images based on inception-ResNet. *Computer Methods and Programs in Biomedicine*.

Czinege, I., & Harangozó, D. (2023). Application of artificial neural networks for characterisation of formability properties of sheet metals. *International Journal of Lightweight Materials and Manufacture*. https://doi.org/10.1016/j.ijlmm.2023.08.003

Dhanush, G., Khatri, N., Kumar, S., & Shukla, P. K. (2023). A comprehensive review of machine vision systems and artificial intelligence algorithms for the detection and harvesting of agricultural produce. *Scientific African*, *21*(July), e01798.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12987
e-ISSN : 2541-2019
p-ISSN : 2541-044X

https://doi.org/10.1016/j.sciaf.2023.e01798

Djenouri, Y., Belhadi, A., Srivastava, G., & Lin, J. C. W. (2023). Hybrid graph convolution neural network and branch-and-bound optimization for traffic flow forecasting. *Future Generation Computer Systems*, *139*, 100–108. https://doi.org/10.1016/j.future.2022.09.018

Gawade, S., Bhansali, A., Patil, K., & Shaikh, D. (2023). Application of the convolutional neural networks and supervised deep-learning methods for osteosarcoma bone cancer detection. *Healthcare Analytics*, *3*(February), 100153. https://doi.org/10.1016/j.health.2023.100153

Hindarto, D., & Santoso, H. (2021). Plat Nomor Kendaraan dengan Convolution Neural Network. *Jurnal Inovasi Informatika*, *6*(2), 1–12. https://doi.org/10.51170/jii.v6i2.202

Kumar, A., & Seeja, K. R. (2023). Periocular Region based Gender Identification using Transfer Learning. *International Journal of Cognitive Computing in Engineering*, *4*(July), 277–286. https://doi.org/10.1016/j.ijcce.2023.07.003

Niswati, Z., Hardatin, R., Muslimah, M. N., & Hasanah, S. N. (2021). Perbandingan Arsitektur ResNet50 dan ResNet101 dalam Klasifikasi Kanker Serviks pada Citra Pap Smear. *Faktor Exacta*, *14*(3), 160. https://doi.org/10.30998/faktorexacta.v14i3.10010

Ozbayoglu, A. M., & Yuksel, H. E. (2012). Analysis of gas-liquid behavior in eccentric horizontal annuli with image processing and artificial intelligence techniques. *Journal of Petroleum Science and Engineering*, *81*, 31–40. https://doi.org/10.1016/j.petrol.2011.12.008

Pérez-García, F., Sparks, R., & Ourselin, S. (2021). TorchIO: A Python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning. *Computer Methods and Programs in Biomedicine*, *208*. https://doi.org/10.1016/j.cmpb.2021.106236

Pribanić, I., Simić, S. D., Tanković, N., Debeljuh, D. D., & Jurković, S. (2023). Reduction of SPECT acquisition time using deep learning: A phantom study. *Physica Medica*, *111*(January). https://doi.org/10.1016/j.ejmp.2023.102615

Saputra, A. D., Hindarto, D., & Santoso, H. (2023). Disease Classification on Rice Leaves using DenseNet121, DenseNet169, DenseNet201. *Sinkron*, *8*(1), 48–55. https://doi.org/10.33395/sinkron.v8i1.11906

Swift, J. R., Turner, M. A., & Reynolds, J. C. (2023). A rapid dynamic headspace method for authentication of whiskies using artificial neural networks. *Food Chemistry Advances*, *3*(June), 100417. https://doi.org/10.1016/j.focha.2023.100417

Sze, E., Santoso, H., & Hindarto, D. (2022). *Review Star Hotels Using Convolutional Neural Network*. *7*(1), 2469–2477.

Zimmermann, L., Buschmann, M., Herrmann, H., Heilemann, G., Kuess, P., Goldner, G., Nyholm, T., Georg, D., & Nesvacil, N. (2021). An MR-only acquisition and artificial intelligence based image-processing protocol for photon and proton therapy using a low field MR. *Zeitschrift Fur Medizinische Physik*, *31*(1), 78–88. https://doi.org/10.1016/j.zemedi.2020.10.004

*name of corresponding author