International Journal of Integrated Science and Technology (IJIST)
Vol.2, No.5, 2024: 411-419

Comparative Analysis of Convolutional Neural Network (CNN) Architectures in Classification of Cattle and Pig Rambaks

Haryono[1] , Cahya Rahmad[2] , Banni Satria Andoko[3]
Politeknik Negeri Malang

Corresponding Author: Haryono; haryono88@proton.me

ARTICLEINFO

Keywords: Architecture, Classification, CNN, Deep Learning, Rambak

Received : 20, March
Revised : 25, April
Accepted: 29, May

©2024 Haryono, Rahmad, Andoko(s): This is an open-access article distributed under the terms of the Creative Commons Atribusi 4.0 Internasional.

ABSTRACT

Rambak crackers are one of the food ingredients that have the characteristics of expansion and crispy texture. The general public often faces difficulties in distinguishing between pork and beef rambak crackers that have been processed, so it is important to rely on technology, especially artificial intelligence (AI), to help distinguish between them. With its predominantly Muslim population, adheres to Islamic dietary laws (halal) that strictly prohibit the consumption of pork. As such, accurately differentiating between images of pork and beef is essential for ensuring compliance with religious dietary restrictions. This study was conducted to compare the capabilities of several CNN architectures in classifying images of pork and beef rambak crackers. The results of the study showed that the Xception architecture had the highest accuracy rate in classifying pork and beef rambak crackers, with an average accuracy rate of 98.24%.

DOI: https://doi.org/10.59890/ijist.v2i5.1793

ISSN: 3026-4685
https://journal.multitechpublisher.com/index.php/ijist

INTRODUCTION

Rambak crackers are a food ingredient that has flowering properties and a crunchy texture. Rambak crackers are a food product made from animal skin that has undergone a series of processes, including liming, hair removal, boiling, drying and frying (Negara et al., 2023). Apart from that, rambak is also an economic commodity that has strategic value. As a popular source of protein, rambak has become a favorite choice for public consumption to meet the body's nutritional needs. Although beef and pork rambak have a similar appearance, they both have different characteristics. Beef rambak generally has a paler red color and a distinctive aroma and taste, while pork rambak tends to have a color that varies from pale to pink with a specific aroma.

Ordinary people often have difficulty distinguishing between pig rambak and manipulated cow rambak, so technology is important to help identify the differences between the two. One technology that is increasingly being used in various sectors of life today is artificial intelligence (AI). In the context of classifying images, one branch of artificial intelligence (AI) that is useful is deep learning, especially in the field of digital image processing (Dinata et al., 2024). One of the most successful methods for recognizing images is Convolutional Neural Network (CNN). CNN is a well-known and most frequently used method (Zhou, 2020).

This research aims to compare the results of image classification of pig rambak and cow rambak using several CNN architectures including, 'ResNet50V2', 'ResNet152V2', 'InceptionV3', 'Xception', and 'MobileNetV2'.

LITERATURE REVIEW

Convolutional Neural Network (CNN) is a type of neural network architecture inspired by the workings of the visual cortex in the human brain, especially in visual processing (Goodfellow et al., 2016). Computer vision powered by convolutional neural networks (CNNs) has facilitated achievements previously deemed unattainable over the past few centuries (Li et al., 2022). These advancements include but are not limited to, facial recognition systems, autonomous vehicles, self-service supermarkets, and sophisticated medical treatments.The basic concept of CNN is the use of convolution operations to recognize patterns in image data. CNNs can have multiple alternating convolutional and pooling layers. This process allows the network to gradually learn a hierarchical representation of the image, where early layers recognize simple features such as edges and lines, while deeper layers can recognize more complex features such as shapes and objects. Although convolutional neural networks (CNNs) entail a heavier computational load compared to conventional artificial neural networks (ANNs), they offer the advantage of automatically detecting significant features without manual intervention. Consequently, CNNs are often regarded as more potent than traditional ANNs (Sarker, 2021). In the context of Convolutional Neural Network (CNN), "architecture" refers to the structure or arrangement of the various layers that make up the network. An architecture is a guide or plan for how the layers are connected to each other, and how information flows through the network to perform a specific task, such as image classification or object detection.

Research conducted Berliani et al. (2023) using the ResNet-50 and VGG-16 architecture in lung X-ray image classification found that ResNet-50 had better accuracy with an accuracy rate of 95.13%. The findings from Putra et al. (2021) those who obtained the VGG-16 architecture had an accuracy of 95.33% in classifying meat through images. Research Rahman et al. (2024) using the MobileNetV2 architecture obtained an accuracy of 96.93% in classifying pork and beef through images.

Architecture in a Convolutional Neural Network (CNN) refers to the hierarchical structure of neural layers used to process and analyze images. Model architecture is an important factor in improving the performance of various applications (Alzubaidi et al., 2021). The development of this architecture has increased complexity with the addition of depth and more complex pattern recognition through architectures such as VGGNet, GoogLeNet (Inception), and ResNet. These developments also address computational efficiency, as seen in MobileNet, to support applications on mobile and embedded devices. With ongoing research and innovation, CNN architectures continue to evolve to address new challenges in image pattern recognition.

The architectural paradigm of ResNet (Residual Network) represents a significant breakthrough in the evolution of Convolutional Neural Networks (CNNs), initially introduced by (He et al., 2015). The primary conceptual innovation within ResNet lies in its residual blocks, which enable the network to bypass certain layers through "skip connections," thereby facilitating the direct propagation of original information from input to output. This mechanism addresses the vanishing gradient problem encountered in deep networks and facilitates the construction of significantly deeper architectures. Leveraging residual blocks, ResNet achieves depths exceeding 100 layers, thereby engendering more robust and intricate feature representations. The ResNet architecture has exhibited remarkable performance across diverse image recognition tasks, securing victories in various competitions and serving as a foundational cornerstone for subsequent CNN architecture developments.

The architecture of Inception, often referred to as GoogLeNet, constitutes a significant breakthrough within Convolutional Neural Network (CNN) research, pioneered by the Google team in 2014. The core principle underpinning Inception revolves around the utilization of Inception modules, facilitating the network to concurrently learn features across multiple scales by employing various sizes of convolutional filters within a single layer. In an Inception network, these modules are arranged in a stacked manner, occasionally interspersed with max-pooling layers set to a stride of 2, which effectively reduces the grid resolution by half (Szegedy et al., 2014). This modular approach allows the network to extract features more efficiently compared to conventional methods, as it enables the network to select and amalgamate features detected at diverse resolution levels. Consequently, this framework facilitates the construction of deeper and more intricate networks while maintaining a relatively modest parameter count. The Inception architecture has demonstrated its efficacy across a spectrum of image recognition tasks and has served as a wellspring of inspiration for the development of more sophisticated CNN architectures.

Xception architecture represents a sophisticated advancement stemming from the Inception architecture, also developed by the Google team. The term "Xception," derived from "Extreme Inception," introduces the concept of "depthwise separable convolution," which replaces the standard convolution within the Inception module. Xception represents a convolutional neural network design characterized by its exclusive use of depthwise separable convolution layers (Chollet, 2016). Depthwise separable convolution segregates the convolution process into two distinct stages: spatial convolution and depthwise convolution. This approach substantially reduces the parameter count without compromising performance, as it enables the independent learning of spatial and depth features. By leveraging depthwise separable convolution, Xception achieves higher levels of efficiency compared to Inception, attaining equivalent or superior performance in image recognition tasks. Similar to traditional convolution, Xception does not require convolution across all channels. This reduces the number of connections and results in a lighter model (Salim et al., 2023). Xception architecture has emerged as a frontrunner in the development of modern convolutional neural networks and remains an actively researched and developed area within the field.

MobileNet architecture constitutes a convolutional neural network (CNN) developed by Google in 2017, emphasizing computational efficiency and low resource utilization. The core concept of MobileNet revolves around decomposing convolution kernels (Wang et al., 2020). The MobileNet architecture leverages the depthwise separable convolutional structure, wherein spatial convolution and depthwise convolution are decoupled. This approach enables a drastic reduction in the requisite parameter count, rendering it well-suited for deployment on resource-constrained devices such as mobile or embedded systems. MobileNet is a convolutional neural network tailored specifically for mobile phone CPUs. This customization is achieved through a blend of methods, including hardware-conscious network architecture search (NAS) coupled with the NetAdapt algorithm (Howard et al., 2019). MobileNet has demonstrated efficacy across various image recognition tasks on mobile devices, positioning it as a primary choice for applications necessitating high performance coupled with low power consumption. Furthermore, MobileNet has catalyzed further research into the development of computationally efficient CNN architectures.

METHODS

This research focuses on the implementation of the CNN architecture in classifying rambak types. The first stage carried out in this research was collecting images of pig rambak and cow rambak. The next stage is to carry out data processing where what is done is equalizing the size of the image. The image size used is 224 x 224 pixels. Next, data sharing is carried out; 556 images (73%) of the data will be used as *training data,* 176 images (23%) of the images will be used as *testing data* , and 30 images (5%) will be used as validation data. Then augmentation is carried out on the *training data* by *rescaling* , rotating and *shifting* on the horizontal and vertical sides. The next stage is to do modeling with the architecture that has been determined. The final stage is evaluating the results of modeling for each architecture.

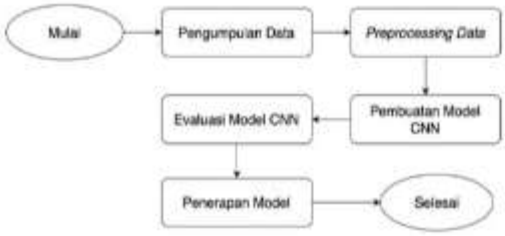

Figure 1. Research Stage

The data used in this research are images of cow rambak and pig rambak taken using the same cellphone for all images. The results of image capture are saved in .jpeg format.

Figure 2. Image of Rambak

The data that has been taken is divided into 3 groups; namely *training data*, *testing data* and validation data.

RESULTS

In this section, the results of the findings and discussion of identifying rambak types using the various architectures used are presented. The CNN architectures used are: 'ResNet50V2', 'ResNet152V2', 'InceptionV3', 'Xception', and 'MobileNetV2'.

In this research, 20 epochs were used to determine the performance of each architecture. Following are the comparison results

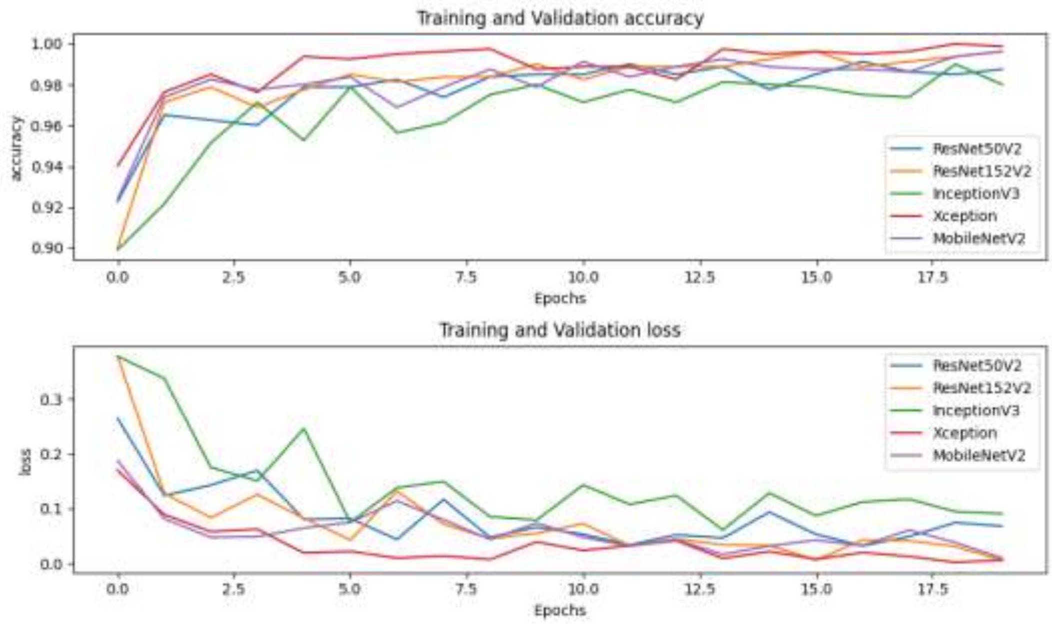

Figure 2. Research Results

From the test results, it was found that for this research the Xception architecture had better performance in terms of loss rate and accuracy.

Table 2. Test Results

Architecture	Epoch					Average
	1	5	10	15	20	
ResNet-50	92.2	97.8	98.5	97.7	98.7	96.98
ResNet-152	90	97.7	99	99.2	99.6	97.1
Inception	89.9	95.2	98	98	98	95.82
Xception	94	99.3	98.7	99.5	99.7	98.24
MobileNet	92.4	98	97.8	98.8	99.6	97.32

Figure 3. Classification Results Using Xception

CONCLUSIONS AND RECOMMENDATIONS

From the research results it was found that the architectural model Xception has the highest accuracy capability in classifying cattle rambak and pig rambak with an average of 98.24%. From these findings, it is hoped that there will be further research to improve the performance of this model

FURTHER STUDY

From these findings, it is hoped that there will be further research to improve the performance of this model .

REFERENCES

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, *8*(1), 53. https://doi.org/10.1186/s40537-021-00444-8

Berliani, T., Rahardja, E., & Septiana, L. (2023). Perbandingan Kemampuan Klasifikasi Citra X-ray Paru-paru menggunakan Transfer Learning ResNet-50 dan VGG-16. *Journal of Medicine and Health*, *5*(2), 123–135. https://doi.org/10.28932/jmh.v5i2.6116

Chollet, F. (2016). *Xception: Deep Learning with Depthwise Separable Convolutions*.

Dinata, I. M. A. M., Gunadi. I Gede Aris, & Sunarya, I. M. G. (2024). Analisis Hyperparameter Pada Klasifikasi Jenis Daging Menggunakan Algoritma Convolutional Neural Network. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, *8*(1), 25–34.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning (Adaptive Computation and Machine Learning series)*. MIT Press.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep Residual Learning for Image Recognition*.

Howard, A., Sandler, M., Chu, G., Chen, L.-C., Chen, B., Tan, M., Wang, W., Zhu, Y., Pang, R., Vasudevan, V., Le, Q. V., & Adam, H. (2019). *Searching for MobileNetV3*.

Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. (2022). A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. *IEEE Transactions on Neural Networks and Learning Systems*, *33*(12), 6999–7019. https://doi.org/10.1109/TNNLS.2021.3084827

Negara, R. G. P., Permadi, H. irawati, Kusuma, C., Anandita, A. S., & Prayitno, C. P. (2023). PENGARUH LAMA PERENDAMAN MENGGUNAKAN LARUTAN Ca(OH)2 DENGAN KONSENTRASI 8%PADA PEMBUATAN KERUPUK RAMBAK TERHADAP KADAR PROTEIN DAN DAYA KEMBANG KULITKAMBING. *Jurnal Pertanian, Peternakan, Perikanan*, *1*. https://doi.org/10.3766/hibrida.v1i1.709

Putra, K. B. V., Bayupati, I. P. A., & Arsa, D. M. S. (2021). Klasifikasi Citra Daging Menggunakan Deep Learning dengan Optimisasi Hard Voting. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *5*(4), 656–662. https://doi.org/10.29207/resti.v5i4.3247

Rahman, A. F. A., Muji, S. Z. M., & Kan, C. U. E. (2024). Beef and Pork Meat Classification using MobileNet V2 Algorithm via Android Smartphone Application. *Evolution in Electrical and Electronic Engineering*, *5*(1), 292–297. https://publisher.uthm.edu.my/periodicals/index.php/eeee/article/view/15319

Salim, F., Saeed, F., Basurra, S., Qasem, S. N., & Al-Hadhrami, T. (2023). DenseNet-201 and Xception Pre-Trained Deep Learning Models for Fruit Recognition. *Electronics*, *12*(14), 3132. https://doi.org/10.3390/electronics12143132

Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, *2*(3), 160. https://doi.org/10.1007/s42979-021-00592-x

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2014). *Going Deeper with Convolutions*.

Wang, W., Hu, Y., Zou, T., Liu, H., Wang, J., & Wang, X. (2020). A New Image Classification Approach via Improved MobileNet Models with Local Receptive Field Expansion in Shallow Layers. *Computational Intelligence and Neuroscience*, *2020*, 1–10. https://doi.org/10.1155/2020/8817849

Zhou, D.-X. (2020). Theory of deep convolutional neural networks: Downsampling. *Neural Networks*, *124*, 319–327. https://doi.org/10.1016/j.neunet.2020.01.018