Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15th June 2025. Vol. 4. No. 3; e-ISSN: 2808-4519

Palm Fruit Ripeness Detection System Using Convolutional Neural Network (CNN) Algorithm

Josua Nainggolan[1]*, Debi Yandra Niska[2], Faridawaty Marpaung[3], Insan Taufik[4], Kana Saputra S[5]

[1,2,4,5] Ilmu Komputer, Universitas Negeri Medan
[3] Matematika, Universitas Negeri Medan
*Josuanainggolan991@gmail.com[1]**

Abstract

Oil Palm Fruit is a valuable natural resource crop in the plantation sector in Indonesia, with promising future growth prospects. To produce the best oil palm fruit, good sorting is needed. With good oil palm fruit, adequate technology is needed to assist in sorting oil palm fruit. Therefore, this study aims to help companies sort oil palm fruit bunches. In this study, CNN was used with the MobileNetV2 algorithm and the training accuracy results reached a peak of 98.20%, while the validation accuracy remained high at 95.00%. This proves that this model is very good and very feasible for further research. This method has proven to be the best choice for achieving high accuracy and low loss, but also minimizing errors in prediction.

Keywords: *Palm, Accuracy, Classification, CNN, MobileNetV2*

1. Introduction

Oil palm is a valuable natural resource crop in the plantation sector in Indonesia, with promising future growth prospects. After coffee and rubber, oil palm natural resources, both raw materials and processed products, are the third largest source of non-oil and gas foreign exchange in Indonesia. Oil produced by oil palm fruit has a number of advantages, including low cholesterol levels. This makes palm oil a reliable source of vegetable oil [1].

Every year, Indonesia's palm oil production continues to experience a significant increase, with an average increase of around 0.55 million tons of CPO per year [2]. Efforts to develop the palm oil industry must be synergized with efforts to realize energy sovereignty in Indonesia. The abundance of crude palm oil production is both an opportunity and a challenge to create Indonesia's energy independence and resilience [3]. Based on data from the Coordinating Ministry for Economic Affairs of the Republic of Indonesia in 2018, 46.68 million tons of palm oil were produced, consisting of 40.57 million tons of Crude Palm Oil (CPO), and 8.11 million tons of Palm Kernel Oil (PKO). These results were obtained from several plantations such as Private Large Plantations of 60%. State Large Plantations of 5% and People's Plantations of 35% [4].

With the high amount of recorded production, and the variety of derivative products from oil palm plants, it proves that oil palm is one of the most sought-after international commodities. Demand will continue to increase, so that oil palm farmers need a more efficient approach in managing the production of Fresh Fruit Bunches (FFB) of oil palm [5]. One of the challenges currently faced is the process of sorting ripeness which is still carried out manually both by oil palm farmers and also sorting officers at the palm oil mill, which is not only time consuming, but also inefficient. Sometimes, the sorters at the company have limited access to a reliable classification system. Where weather factors greatly affect the sorting of oil palm fruit [6].

The result of the detection error resulted in a decrease in the quality of CPO in Indonesia which resulted in high Free Fatty Acid (FFA) content. Handling of oil palm fruit harvest is an important activity in improving CPO quality. Oil palm fruit must be harvested on time, if it is too ripe then the oil produced contains high amounts of FFA (more than 5%) while if it is harvested in an unripe state then the FFA level and oil yield produced will be low [7]. To overcome this problem, many studies have been conducted, including efforts to simplify the maturity assessment process using Artificial Intelligence (AI) technology [8].

Therefore, this study focuses on the development of a lighter CNN to improve the classification capability of oil palm FFB maturity without sacrificing computational efficiency by using the MobileNet method [9]. MobileNet is one of the lightweight and well-known Convolutional Neural Network (CNN) architectures that has been used in various studies involving deep learning. MobileNet was first introduced in 2017 as a CNN specifically designed for mobile applications, with around 4.2 million parameters [10].

2. Research Method

This research consists of five processes (as illustrated in Figure 1):

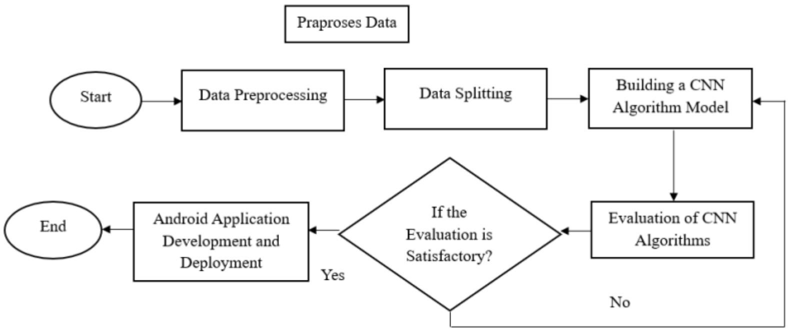

Fig. 1: Research workflows

2.1. Data Preprocessing

The initial stage of this research focused on data collection consisting of various images of Ripe and Unripe Oil Palm Fruit Bunches. This data serves as training data for the model. The data was obtained from direct collection to the company, with a total of 1900 Oil Palm images divided into two categories: Ripe and Unripe. During the data collection process, image quality was checked to ensure that only complete and representative images were used. A total of 300 cropped or low-quality images were removed, leaving 1600 images in the dataset. This removal ensured a balanced number of images in each Ripe and Unripe category, with 800 images per category.

Once the dataset is prepared, data preprocessing is performed to reduce the possibility of errors in the images. This step includes checking the conformity of the images to the specified folder and labeling each data, which is a crucial part of this process [11]. Labels are used to classify data into certain classes. In addition, the collected images are resized to have uniform dimensions, so that consistency between data is maintained. At this stage, data augmentation is also carried out on training images to improve model performance and avoid overfitting[12].

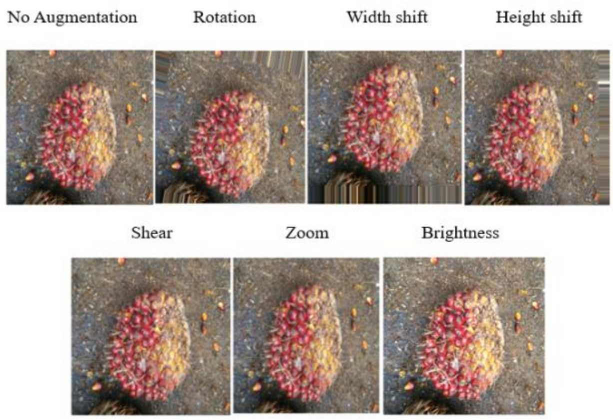

Fig. 2: Results of image augmentation

2.2. Data Splitting

The ratio applied is 60:20:20 (60% training data, 20% validation data, and 20% test data for evaluation). This division is done using the `train_test_split` function from the scikit-learn library to divide the data into training and test sets, and using the `validation_split` parameter when training the model to automatically obtain validation data. The implementation of the Convolutional Neural Network (CNN) for palm oil classification is carried out using the Transfer Learning technique. Researchers implemented the MobileNetV2 model using an image resolution of 224x224. For the division of the dataset, the ratio used was 60:20:20. In addition, testing was carried out using four optimizers, namely Adam [13].

2.3. Building a CNN Algorithm Model

In this step, the CNN model is developed by adding a fully connected layer to the previously selected Transfer Learning architecture. Details of the hyperparameters used in the construction of the CNN model are presented in Table 1.

Table 1: Hyperparameters of the CNN Model

Hyperparameter	Value
Optimization	Adam
Learning Rate	1×10^{-4}
Epoch	50
Batch	32

The optimization algorithm used is Adaptive Momentum Optimization (Adam). To examine the effect of learning rate, experiments were also conducted using a learning rate of 1×10−4. The CNN model was drilled on the data preprocessing using batch sizes of 32 and 50 epochs, in line with previous studies showing that this batch size produces the highest accuracy compared to other batch sizes [14]. The Transfer Learning model used is MobileNetV2 This architecture consists of two components: feature learning and classification. The feature learning stage uses convolutional layers to recognize patterns, while the classification stage uses these patterns to predict classes [15].

2.4. Model Evaluation

This study uses a pre-trained model that has been trained with ImageNet so that the pre-trained model has previous knowledge ranging from brightness, edge, to color, shape and pattern. The pre-trained model provided by Tensorflow has a resolution of 224x224 [16]. Transfer learning uses features that have been extracted from the feature extraction layer in the source domain and applies them to the feature extraction layer in the target domain [17].

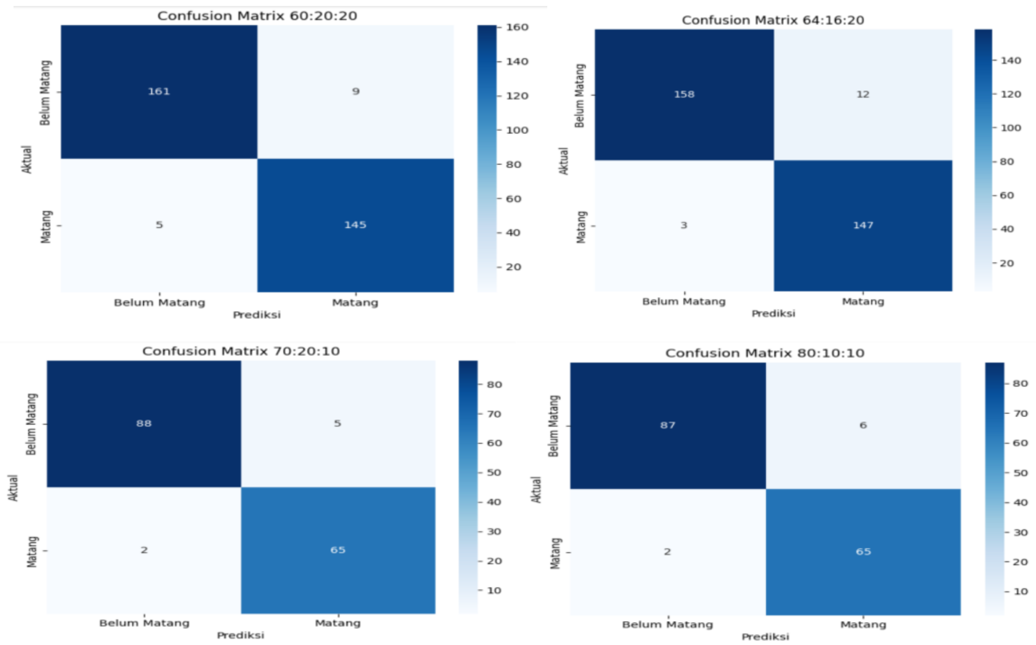

Fig. 3:Confusion Matrix of Various Ratios.

In model evaluation, researchers use training data to assess the model's ability to classify data. Evaluation is carried out using a confusion matrix for various dataset split ratios: 60:20:20, Model Conversion The model that has been built is currently stored in a ".keras" format file. To be used in an Android application, the file needs to be converted to ".tflite" format. The model to be converted is the model with the best performance from all the combinations listed in Table 4.4, namely the model with a 60:20:20 data split configuration, a batch size of 32, and a learning rate of 10^-4 [18].

2.5. Android Application Development and Deployment

Each page view in the application uses fragments, to simplify the process of moving pages and save memory compared to using activities for all pages. In the first stage of creating a UI (XML), then creating an MVVM Hierarchy (creating a folder), creating a BackEnd section (creating a database, DAO, repository), creating a utility section, creating a ViewModel section, Creating a UI, Connection between ViewModel and UI, Testing, and finally testing the application [19].

3. Experiments and Analysis

3.1. Experimental setup

This research was conducted using Google Colab, a cloud platform that provides a Python programming environment and access to hardware such as GPUs and TPUs. The experiments were run using Google Colab's TPUv2 (Tensor Processing Unit), a special hardware designed to accelerate the computational process of machine learning, especially in the field of deep learning. TPUv2 provides high computational capabilities, so that model training can be done more efficiently than using a regular CPU or GPU. By utilizing TPUv2, researchers can perform intensive computational processes without relying on local hardware capabilities.

3.2. Experimental results and analysis

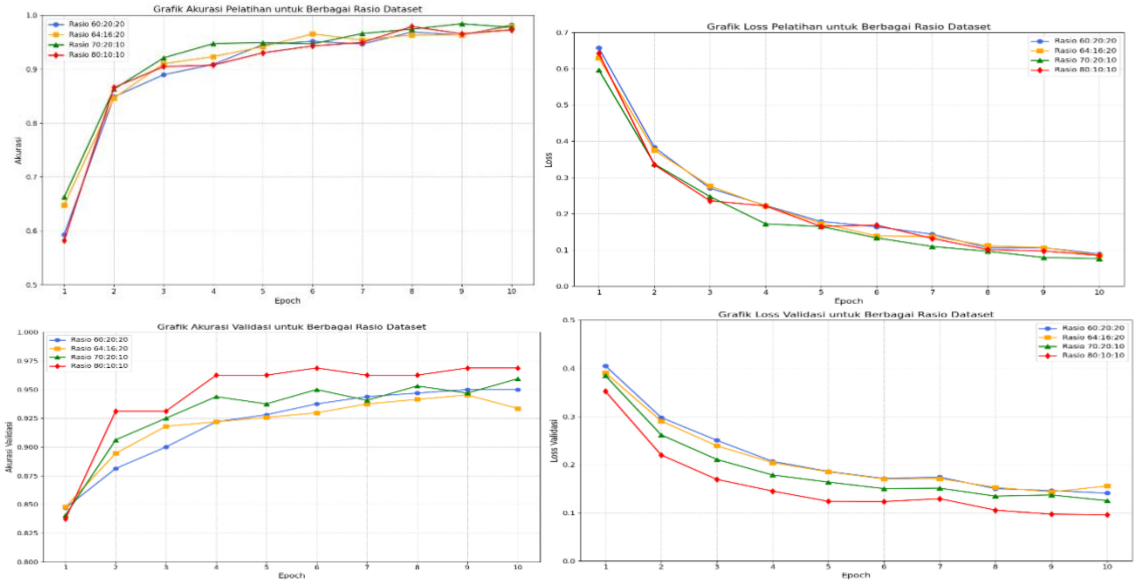

Fig. 4: Training and Validation Accuracy Graph

In the model analysis, the value for training accuracy was obtained. At a ratio of 60:20:20, the training accuracy peaked at 98.20%, while the validation accuracy remained high at 95.00%. After splitting the dataset with this ratio, the researcher conducted a test using the Adam optimizer type with a ratio of 60:20:20. This test aims to evaluate the performance of the model with various optimization algorithms in training.

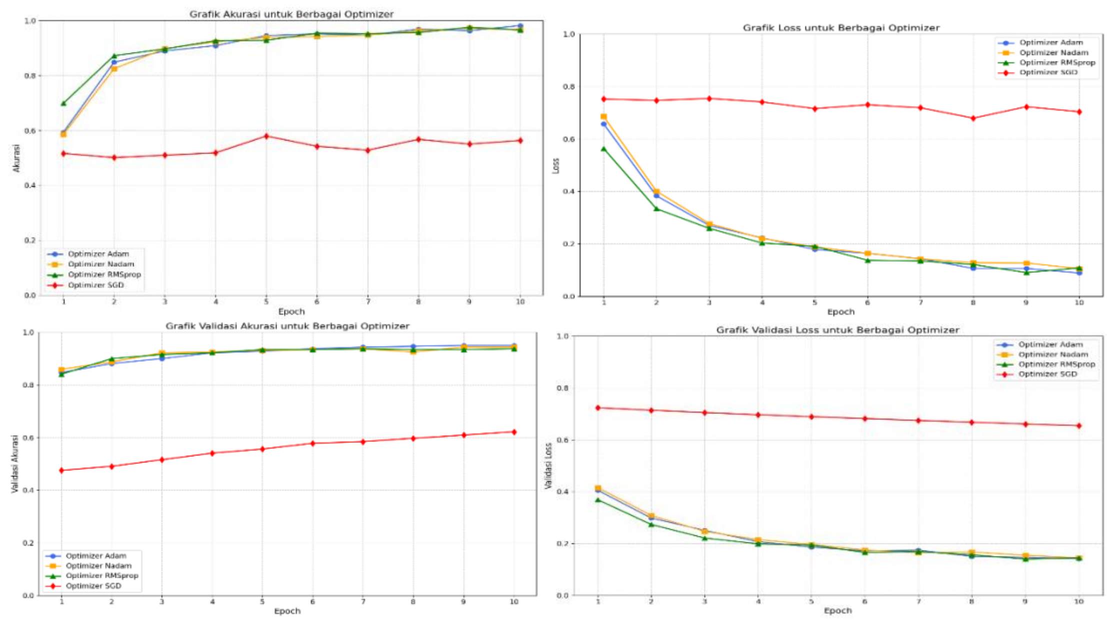

Fig. 5: Adam Optimizer Training and Validation Accuracy Graph

Adam optimizer gives the best results with training accuracy of 98.20% and validation accuracy of 95.00%. This confirms the importance of optimizer selection in the effectiveness of model training, where Adam optimizer is proven to be the best choice to achieve high accuracy and low loss.

3.2. Testing Models

Fig. 6: Testing for each model

After the training process is complete, researchers conduct testing to transmit the model's performance in image classification. Testing is carried out using two types of data: Mature and Immature Data consisting of palm oil images that have never been involved in the training process. In testing using the MobileNetV2 method on the data, it successfully predicted the image correctly, showing a strong ability to recognize the type of palm oil that had been drilled previously. These results indicate that the model shows good generalization ability in distinguishing significantly different visual patterns.

4. Conclusion

This study successfully demonstrated the application of Transfer Learning in classifying ripe and unripe oil palm using CNN architecture. The MobileNetV2 Algorithm model showed the best performance. This model achieved a peak training accuracy of 98.20%, while the validation accuracy remained high at 95.00%. This shows excellent generalization ability in classification tasks. The analysis also revealed a significant difference in model performance, using 60:20:20 data split and Adam Optimizer produced a model accuracy that is worthy of being developed and utilized for the required purposes. indicating excellent model performance. This indicates the efficiency of data use, where the model is able to recognize patterns in new data well without relying on a lot of training data. This ratio is an effective and efficient choice, especially in the context of limited data and training time. This emphasizes the importance of optimizing the effectiveness of model training, where the Adam optimizer is proven to be the best choice to achieve high accuracy and low loss, but also minimize errors in prediction.

Acknowledgement

The authors would like to express their gratitude to everyone who provided valuable feedback and support during the completion of this work

References

[1] Rosmegawati, "Peran Aspek Tehnologi Pertanian Kelapa Sawit Untuk Meningkatkan Produktivitas Produksi Kelapa Sawit," *2021, JURNAL AGRISIA-Vol.13 No.2 Tahun 2302-0091, ISSN : 2302-0091*, vol. 13, no. 2, pp. 1–23, 2016.

[2] A. Alatas, "Trend Produksi dan Ekspor Minyak Sawit (CPO) Indonesia," *AGRARIS: Journal of Agribusiness and Rural Development Research*, vol. 1, no. 2, pp. 114–124, 2015, doi: 10.18196/agr.1215.

[3] Sulardi, *E-book Buku Ajar Budidaya Kelapa Sawit*, Pertama., no. 1 Kali. Medan: PT Dewangga Energi Internasional Anggota IKAPI (403/JBA/2021), 2022. [Online]. Available: https://www.researchgate.net/publication/358981459_E-book_Buku_Ajar_Budidaya_Kelapa_Sawit

[4] Kementrian Koordinator Bidang Perekonomian Republik Indonesia, "Industri Kelapa Sawit Indonesia: Menjaga Keseimbangan Aspek Sosial, Ekonomi, dan Lingkungan - Kementerian Koordinator Bidang Perekonomian Republik Indonesia," *Ekon.Go.Id*, 2021.

[5] A. chandra Saputra and E. D. Oktaviyani, "Rancang Bangun Sistem Deteksi Kematangan Buah Kelapa Sawit Berdasarkan Deteksi Warna Menggunakan Algoritma K-Nn," *Jurnal Teknologi Informasi: Jurnal Keilmuan dan Aplikasi Bidang Teknik Informatika*, vol. 7, no. 2, pp. 222–229, 2023, doi: 10.47111/jti.v7i2.9232.

[6] R. Kurniawan and Nurahman, "Klasifikasi Tingkat Kematangan Buah Sawit Berdasarkan Ekstraksi Fitur RGB dan GLCM Menggunakan Algoritma K-NN," *Jurnal Ilmiah Komputasi*, vol. 4, no. 1, pp. 88–100, 2023, doi: 10.32409.

[7] N. Sabri, Z. Ibrahim, S. Syahlan, N. Jamil, and N. N. A. Mangshor, "Palm Oil Fresh Fruit Bunch Ripeness Grading Udentification using Color Features," *Journal of Fundamental and Applied Sciences*, vol. 9, no. 4S, p. 17, 2017, doi: 10.4314/jfas.v9i4s.3.

[8] M. L. Pura, "PENERAPAN RADIAL BASIS FUNCTION (RBF) UNTUK MENENTUKAN TINGKAT KEMATANGAN BUAH TOMAT MENGGUNAKAN MODEL WARNA HSV," UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU, 2018. [Online]. Available: https://repository.uin-suska.ac.id/16494/

[9] Harsawardana *et al.*, "AI-Based Ripeness Grading for Oil Palm Fresh Fruit Bunch in Smart Crane Grabber," *IOP Conf Ser Earth Environ Sci*, vol. 426, no. 1, 2020, doi: 10.1088/1755-1315/426/1/012147.

[10] A. G. Howard *et al.*, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," 2017, [Online]. Available: http://arxiv.org/abs/1704.04861

[11] MM. Dr. Budi Raharjo, S.Kom., M.Kom., *Deep Learning dengan Python*, Pertama. Semarang: YAYASAN PRIMA AGUS TEKNIK, 2022. [Online]. Available: https://digilib.stekom.ac.id/assets/dokumen/ebook/feb_eab5a7c1f295129ba69a76fee4dff22266879314_1643796893.pdf

[12] Z. Ibrahim, N. Sabri, and D. Isa, "Palm oil fresh fruit bunch ripeness grading recognition using convolutional neural network," *Journal of Telecommunication, Electronic and Computer Engineering*, vol. 10, no. 3–2, pp. 109–113, 2018, [Online]. Available: https://jtec.utem.edu.my/jtec/article/view/4720

[13] O. M. Bensaeed, A. M. Shariff, A. B. Mahmud, H. Shafri, and M. Alfatni, "Oil palm fruit grading using a hyperspectral device and machine learning algorithm," *IOP Conf Ser Earth Environ Sci*, vol. 20, no. 1, 2014, doi: 10.1088/1755-1315/20/1/012017.

[14] Suharjito, Elwirehardja, G. Natanael, Prayoga, and J. Sebastian, "Oil palm fresh fruit bunch ripeness classification on mobile devices using deep learning approaches," *Comput Electron Agric*, vol. 188, 2021, doi: 10.1016/j.compag.2021.106359.

[15] S. Ratna, "Pengolahan Citra Digital Dan Histogram Dengan Phyton Dan Text Editor Phycharm," *Technologia: Jurnal Ilmiah*, vol. 11, no. 3, p. 181, 2020, doi: 10.31602/tji.v11i3.3294.

[16] A. E. Wijaya, W. Swastika, and O. H. Kelana, "Implementasi Transfer Learning Pada Convolutional Neural Network Untuk Diagnosis Covid-19 Dan Pneumonia Pada Citra X-Ray," *Sainsbertek Jurnal Ilmiah Sains & Teknologi*, vol. 2, no. 1, pp. 10–15, 2021, doi: 10.33479/sb.v2i1.125.

[17] Qorry Aina Fitroh and Shofwatul 'Uyun, "Deep Transfer Learning untuk Meningkatkan Akurasi Klasifikasi pada Citra Dermoskopi Kanker Kulit," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 12, no. 2, pp. 78–84, 2023, doi: 10.22146/jnteti.v12i2.6502.

[18] M. Ali, "Pengantar Fungsi Aktivasi di Jaringan Syaraf Tiruan," www.datacamp.com. [Online]. Available: https://www.datacamp.com/tutorial/introduction-to-activation-functions-in-neural-networks

[19] H. Asril, "KLASIFIKASI DOKUMEN TUGAS AKHIR BERBASIS TEXT MINING MENGGUNAKAN METODE NA¨IVE BAYES CLASSIFIER DAN K-NEAREST NEIGHBOR," UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU, 2019. [Online]. Available: https://repository.uin-suska.ac.id/21473/2/Hamdani Asril %28kecuali bab 4%29.pdf