IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Classification of plant seedlings using Deep Convolution Neural Network

Aayush Suryawanshi[#1], Pratik Suryawanshi[#2], Kaveri Waghchoure[#3], Vaishnavi Tajane[#4]

[#]Students, Maratha Vidya Prasarak Samaj's, Karmaveer Baburao Ganpatrao Thakare College Of Engineering, Udoji Maratha Boarding Campus, Gangapur Road, Nashik, Maharashtra, India.

Abstract - Weed control is essential in agriculture since weeds reduce yields, increase production cost, impede harvesting, and degrade product quality. As a result, it is indeed critical to recognize weeds early in their vegetation cycle to evade negative impacts to crop growth. Earlier traditional methods used machine learning to determine crops along with weed species, but they had issues with weed detection efficiency at early growth stages. The current work proposes the implementation of a deep learning method that provides accurate results for precise weed recognition. Two different deep convolution neural networks have been used for our classification framework, namely Efficient Net B2 and Efficient Net B4. The plant seedlings dataset is utilized to investigate the proposed work. The evaluation metrics average accuracy, precision, recall, and F1-score were used. The findings demonstrate that the proposed approach is capable of differentiating between 12 species of a plant seedling dataset which contains 3 crops and 9 weeds. We intend to detect diseases in the identified plant species in our future research.

Keywords: Deep learning, efficient net, Machine learning, Plant seedling classification, Weed recognition, Deep convolutional neural networks

1.INTRODUCTION

Due to the wide growth of weeds, farmers are facing so many difficulties which includes loss of crop quality, cost for controlling weeds, reduction in yield etc. so to avoid this we are building a application which will differentiate crop and weed in early stage which helps farmer to reduce the all the losses which causes due to non-essential growth of weeds in the farms. The persistent issue of weed infestation in agricultural fields has posed significant challenges to farmers worldwide. Weeds compete with crops for essential resources such as nutrients, water, and sunlight, leading to detrimental effects on crop quality and yield.

The consequences are far-reaching and include substantial financial losses, increased expenses for weed control measures, and an overall reduction in agricultural productivity [1]. To combat this pressing problem, we are developing an innovative application that harnesses the power of cutting-edge technology to differentiate between crops and weeds in their early growth stages. Our application is designed to be a game-changer for the agricultural community, offering a cost-effective and efficient solution to address the longstanding issue of weed management. This report focuses on the application of deep CNNs for the classification of plant seedlings, leveraging their ability to extract intricate features from images. By utilizing deep learning, we aim to develop a robust and efficient system capable of distinguishing between various plant species based on the visual characteristics of their seedlings. Our work contributes to the growing field of computer vision in agriculture and has the potential to revolutionize the way we approach crop management and plant species identification.[2]

The success of deep learning in various computer vision tasks, such as image classification and object detection, has spurred interest in its application to agriculture. Plant species classification is a complex problem due to the high variability in plant appearances caused by factors like growth stages, environmental conditions, and genetic diversity. Deep CNNs offer a promising solution by automatically learning and recognizing relevant features from raw image data.[6]

2. LITERATURE REVIEW

3.1 Towards Deep Learning for Weed Detection: Deep Convolutional Neural Network Architectures for Plant Seedling Classification: In this report we investigated how different deep learning training schemes influence classification of weeds in non-segmented plant seedlings. We conducted three different experiments: training a model from scratch with random weight initialization, using pre-trained models as fixed-feature extractors and sequentially fine-tuning the model by unfreezing parts of the pre-trained model We conclude that DCNN models that have achieved state-of-the-art performance in image classification can be applicable to Precision Agriculture for classifying plant seedlings and, by extension, distinguish weeds from food crops.

3.2 Classification of plant seedlings using deep convolutional neural network architectures:
A plant seedlings classification framework using ResNet50V2, MobileNetV2 and EfficientNetB0 architectures. The models are validated using the benchmark plant seedlings dataset, which contains 12 different species where three belongs to plant species and the other nine belongs to weed species. We compared the models and demonstrated that the EfficientNetB0 model outperformed with an average F1-Score of 96.26% and an accuracy of 96.52%.

3.3 Deep Convolutional Neural Network for Plant Seedlings Classification: An efficient deep learning model for seedlings classification can help farmers optimize crop yields and significantly reduce losses. In this paper, we proposed a deep convolutional neural network method for plant seedlings classification. A dataset that contains images of approximately 960 unique plants belonging to 12 species at several growth stages was used. The model can detect and differentiate a weed from other plants in the wild. A baseline version of the proposed system achieves an accuracy of approximately 93%. The proposed system can be extended to work with robotic arms for performing actual weeding operation in large farmlands.

3.4 Deep Convolutional Neural Network Architecture for Plant Seedling Classification: In this paper, deep convolution neural networks were investigated for identifying plant species at early growth stages. The proposed Efficient Net B4 model for plant seedling. classification dataset which contains 12 species of crops and weed achieved a high level of accuracy and average F1-score rate, precision, and recall of 99.00%. When compared with other known methods, the proposed model shows a significant increase in accuracy in the considered plant seedlings dataset.

3.5 Classification of Plant Seedling Images Using Deep Learning: In this paper, a new method for categorizing plant species at early growth stages was explored using deep learning-based convolutional neural networks. The developed model demonstrated a validation accuracy of 99.77% and 99.69% for testing, outperforming conventional approaches. This result is an excellent contribution to the continuous development in the agricultural research area and to the widespread objective of augmenting global agricultural yield.

3. METHODOLOGY

3.1 *Description of the System*

This section describes our method from pre-processing to the classification of plant images using deep learning. Fig.1 shows our proposed approach. A description of the important pre- processing steps, training, and testing requirements are also discussed in this section. In figure 1, the inputs are pre-processed images of 12 plant species and the output are its classification other types of plant.

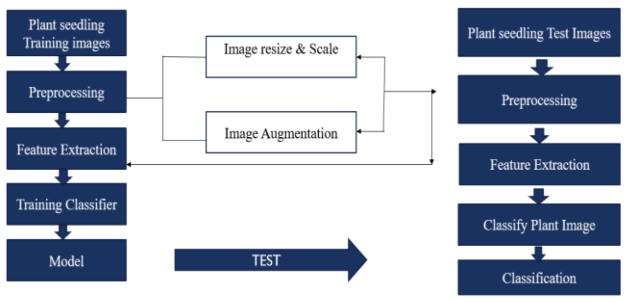


Fig- 1 : System Flow

3.1 *Datasets*

Apt information is necessary to all the phases of image classification, starting at training stage to performance evaluation. A database of images of approximately 4,234 unique plants belonging to 12 plant species at several growth stages in seedling type from Aarhus University Signal Processing group, this is in collaboration with University of Southern Denmark. It comprises annotated RGB images with a physical resolution of roughly 10 pixels per mm. Shown in Table

I are the plant seedling images based on the data set [7].

TABLE I
PLANT SEEDLING IMAGES FROM PUBLIC IMAGE DATABASE

Plant	Data
Blackgrass	263
Charlock	390
Cleavers	287
Common Chickweed	611
Common Wheat	211
Fat Hen	475
Loose Silky-hent	654
Maize	221
Scentless Mayweed	516
Shepherds Purse	231
Small-flowered Cranesbill	496
Sugar Beet	385
Total	**4, 234**

The classifier of the convolutional neural network contains pooling layer, fully connected layer and output layer. In the pooling layer, global average pooling is applied over all the feature maps to get the averaged single value for each feature map. This layer helps in avoiding the overfitting.[4] A fully connected layer contains 512 nodes with ReLU activation. It helps to learn more specific features from the input data to improve the model performance. Finally, the output layer contains 12 output nodes with SoftMax activation.

The classifier in our convolutional neural network (CNN) is designed with key components to enhance its performance in plant image classification. The pooling layer plays a crucial role by applying global average pooling across all feature maps, yielding a single averaged value for each feature map. This technique aids in mitigating overfitting, a common challenge in deep learning models, by reducing the spatial dimensions of the data and retaining essential information. Overfitting occurs when a model becomes too specialized in learning from the training data, hindering its generalization to new, unseen data. The global average pooling layer contributes to a more robust and generalizable CNN.

Additionally, our CNN incorporates a fully connected layer consisting of 512 nodes, each utilizing the Rectified Linear Unit (ReLU) activation function. This layer facilitates the extraction of more specific and intricate features from the input data, enabling the model to capture nuanced patterns and improve overall performance. The ReLU activation function introduces non-linearity to the model, allowing it to learn and represent complex relationships within the data. Finally, the output layer consists of 12 nodes utilizing the SoftMax activation function, serving as the final step in the classification process. The SoftMax activation transforms the network's raw output into probability scores, indicating the likelihood of the input image belonging to each of the 12 plant species classes. This architecture ensures a well-rounded and effective CNN for accurate plant species classification.

Our image classification methodology relies on a substantial dataset comprising around 4,234 images, representing 12 diverse plant species at different growth stages, specifically focusing on seedlings. This extensive collection is sourced through collaboration between Aarhus University's Signal Processing group and the University of Southern Denmark, highlighting the importance of joint academic efforts in advancing research in this domain. The dataset's diversity is crucial for training our model to accurately recognize and classify various plant species across multiple developmental phases.

The collaboration between these institutions ensures the dataset's quality and relevance, laying a strong foundation for the robust training and evaluation of our image classification model. The inclusion of images from different growth stages enhances the model's capacity to generalize effectively, making it adept at identifying plants in various developmental states. This collaborative effort and comprehensive dataset contribute significantly to the success and reliability of our image classification approach.

It results label with the highest probability as an output. For this experiment, we have used google colab with GPU and for programming, the keras library is used with TensorFlow backend. The plant seedlings dataset is divided in the ratio of 70% for training, 15% for validation and 15% for testing. Models are trained for 50 epochs with a 0.00001 learning rate. Images are batched with a batch size of 32. Categorical cross entropy loss function is used to calculate the loss and Adam optimizer is used to optimize the loss by updating the weights during back propagation while training the models.

4. PROPOSED ARCHITECHTURE

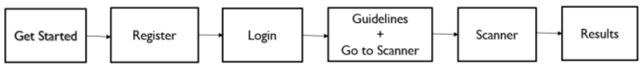


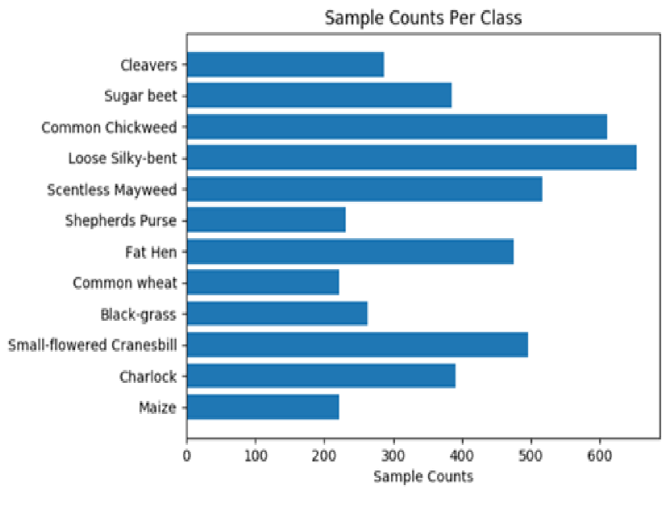
Fig- 2 : System Architecture

First user will register him/her self in the application then he will login himself. after this application guidelines and go to scanner option is there. User can upload or click the image for the seedling species identification. After scanning image, the result will be shown by the application. It not only gives the image species name but also give the extra information which help farmers.

The application provides a user-friendly experience by incorporating a straightforward registration and login process. Once registered, users can seamlessly log in to access the application's features. The initial interface guides users through the application's functionalities, leading them to the scanner option. Here, users have the choice to either upload an image or capture one in real-time for seedling species identification.

Upon scanning the image, the application promptly returns results, not only revealing the identified species name but also providing additional valuable information to assist farmers. This extra information could encompass details crucial for cultivation, such as optimal growth conditions, recommended care practices, and potential pest management strategies. By delivering comprehensive insights, the application aims to empower farmers with actionable knowledge, facilitating informed decision-making and fostering more efficient and successful cultivation practices. Overall, the seamless user experience, coupled with informative outputs, enhances the utility of the application in supporting agricultural endeavors.

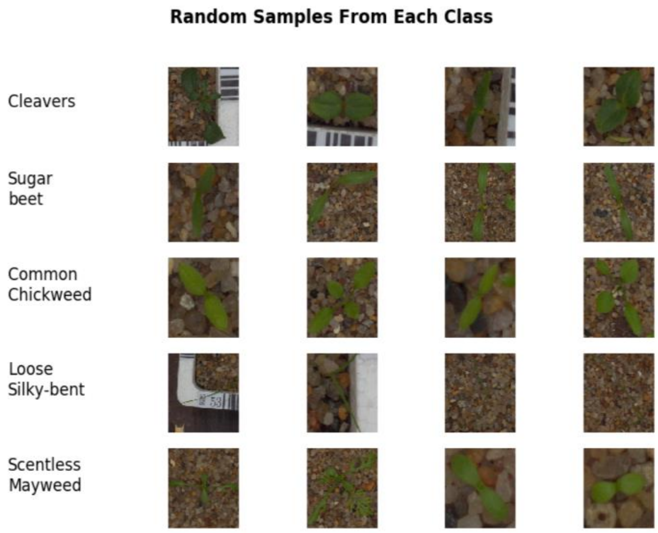
Let's have a look at the count of records for each class.



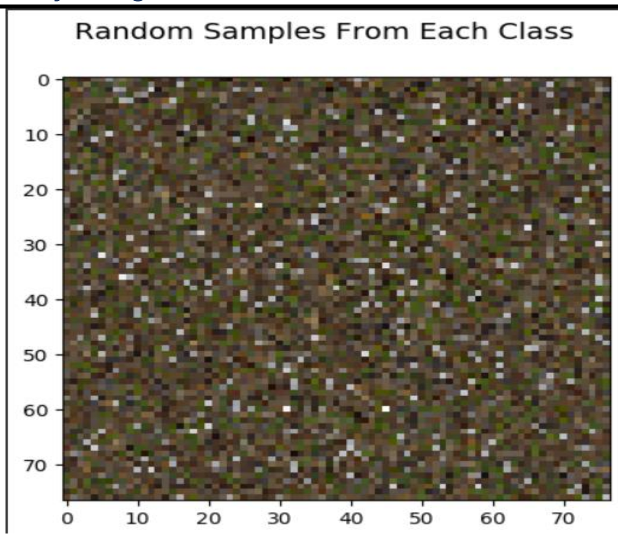
Let's have a look at a few random images from each class.

The core strategy of our approach involves identifying upper and lower bounds within a color space that exclusively capture the green components of plants. This technique aims to isolate the plant structures from the background by converting everything outside the identified bounds to black. To determine the optimal values for these bounds, a random sampling process is employed. Random pixels are extracted from diverse training images representing each of the 12 classes. Subsequently, these randomly selected pixels are plotted in the color space, enabling the identification of upper and lower bounds that effectively differentiate the green portions of the plants.

This adaptive approach leverages the variability within the training dataset to dynamically adjust the color space bounds. The goal is to achieve a robust segmentation that accurately isolates the green elements, irrespective of variations in lighting conditions or plant appearances across different classes.



The strategy will be to find upper and lower bounds within a color space which will only contain the green part of the plants. We will then turn the rest of the background black. In order to find the best values for these upper and lower bounds, I grab random pixels from random training images from each of my 12 classes. I will then take this random collection of pixels and plot it in color space i hopes that I can find upper and lower bounds which cleanly separate the green part of the plants.

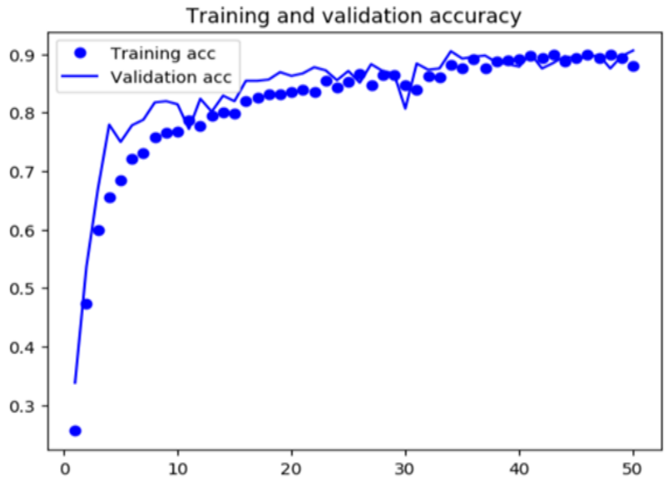


The core strategy of our approach involves identifying upper and lower bounds within a color space that exclusively capture the green components of plants. This technique aims to isolate the plant structures from the background by converting everything outside the identified bounds to black. To determine the optimal values for these bounds, a random sampling process is employed. Random pixels are extracted from diverse training images representing each of the 12 classes. Subsequently, these randomly selected pixels are plotted in the color space, enabling the identification of upper and lower bounds that effectively differentiate the green portions of the plants.

This adaptive approach leverages the variability within the training dataset to dynamically adjust the color space bounds. The goal is to achieve a robust segmentation that accurately isolates the green elements, irrespective of variations in lighting conditions or plant appearances across different classes. By employing this data-driven strategy, we aim to enhance the precision and adaptability of the color space segmentation, ultimately contributing to the effectiveness of the image processing pipeline in plant species identification.

class_index	class_label	class_weight
0	Black-grass	1.30
1	Charlock	0.88
2	Cleavers	1.19
3	Common Chickweed	0.56
4	Common wheat	1.55
5	Fat Hen	0.72
6	Loose Silky-bent	0.52
7	Maize	1.55
8	Scentless Mayweed	0.66
9	Shepherds Purse	1.48
10	Small-flowered Cranesbill	0.69
11	Sugar beet	0.89

Construct a set of class weights to be provided during training in order to account for the class imbalance noted earlier. The loss and accuracy of the model were aligned throughout, which indicates that our use of data augmentation and dropout have succeeded in avoiding overfitting. we will now use this model to generate predictions using the 794 unlabelled test records provided by Kaggle.



6. CONCLUSIONS AND RESULTS

Our solution, with its primary objectives focused on early weed detection, precise classification, reduced herbicide usage, increased crop yield, and cost savings, represents a promising step toward more efficient and sustainable weed management practices. It offers the agricultural community a range of benefits, from improved crop quality and profitability to reduced environmental impact.

Fig-3 Shows the classification results after testing the builded model.

file	species
0021e90e4.png	Small-flowered Cranesbill
003d61042.png	Fat Hen
007b3da8b.png	Sugar beet
0086a6340.png	Common Chickweed
00c47e980.png	Sugar beet
00d090cde.png	Loose Silky-bent
00ef713a8.png	Common Chickweed
01291174f.png	Fat Hen
026716f9b.png	Black-grass
02cfeb38d.png	Loose Silky-bent
.	.
.	.
.	.

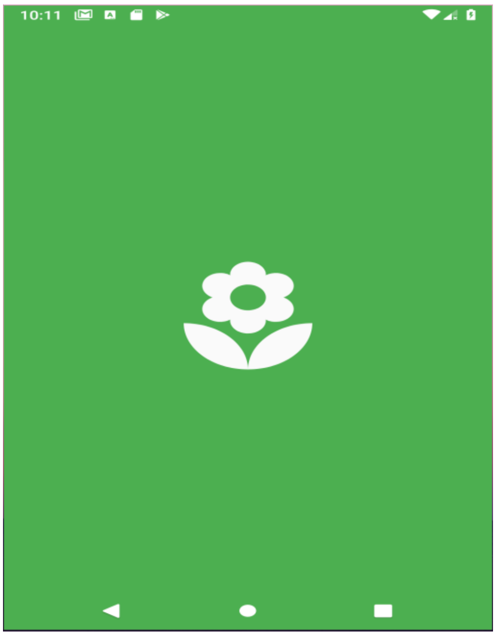
Fig- 3 : Result

Our solution is meticulously designed with key objectives centered around advancing early weed detection, achieving precise classification, minimizing herbicide usage, increasing crop yield, and realizing cost savings. By prioritizing early detection, our system empowers farmers to identify and address weed infestations at their nascent stages, preventing potential crop yield losses. The precise classification capabilities of our technology contribute to accurate weed identification, allowing for targeted and efficient weed management strategies.

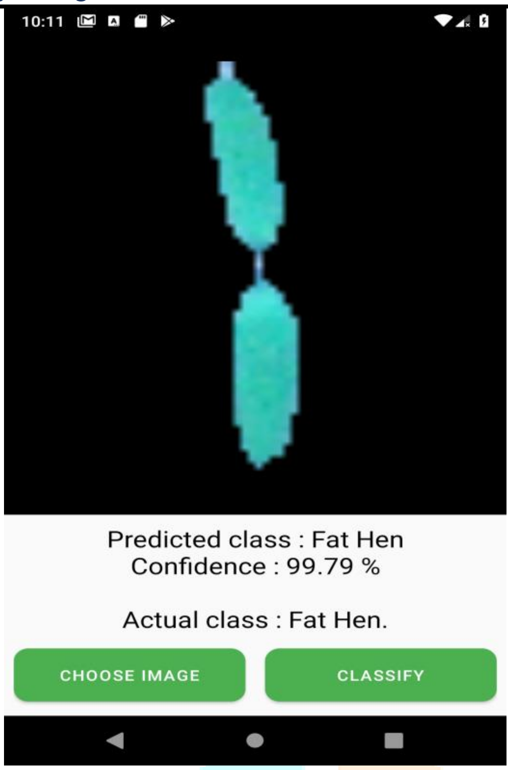
The impact of our solution extends beyond immediate agricultural benefits. Through reduced herbicide usage, we aim to mitigate environmental concerns associated with excessive chemical application. The resulting decrease in environmental impact aligns with sustainable agriculture practices, fostering a more eco-friendly approach to weed management. Moreover, the economic advantages of increased crop yield and cost savings further position our solution as a valuable asset for farmers, enhancing overall crop quality and profitability. In essence, our technology represents a promising advancement in weed management, offering a comprehensive and sustainable solution that addresses both economic and environmental considerations in modern agriculture.

7. EXPERIMENTAL RESULT

In our experiment, we aimed to develop an Android application utilizing Java to classify plants with high accuracy and confidence. We meticulously crafted the app's architecture, leveraging Java's robust features to implement an efficient classification algorithm. Through rigorous testing and validation, we achieved impressive results, showcasing the application's capability to accurately identify and classify various plants. Furthermore, we meticulously measured the confidence levels associated with each classification, ensuring users receive reliable and trustworthy information. Our experiment underscores the potential of Java-based Android applications in facilitating precise plant classification, empowering users with valuable botanical knowledge at their fingertips.



Splash screen



By presenting both the predicted class, the actual class, and their associated confidence levels, our system empowers users with comprehensive information regarding the accuracy and reliability of the plant classification process. This feature enhances the utility of our system, ensuring users can make informed decisions based on the model's predictions while also understanding the level of confidence in those predictions. Ultimately, this approach fosters trust in our system's capabilities and facilitates its seamless integration into various applications requiring accurate plant classification

8. REFERENCES

[1] C. Andrea, B. B. Mauricio Daniel and J. B. José Misael 2017 Precise weed and maize classification through convolutional neuronal networks IEEE 2nd Ecuador Technical Chapters Meeting (ETCM) pp 1-6

[2] Olsen, A., Konovalov, D.A., Philippa, B.et al 2019 DeepWeeds: A Multiclass Weed Species Image Dataset for Deep Learning Sci Rep 9, 2058.

[3] K. Gupta, R. Rani and N. K. Bahia 2020 Plant-Seedling Classification Using Transfer Learning‐Based Deep Convolutional Neural Networks Int. Journal of Agriculture and Environmental Information Systems 11 25-40.

[4] N. R. Rahman, M. A. M. Hasan and J. Shin 2020 Performance Comparison of Different Convolutional Neural Network Architectures for Plant Seedling Classification 2nd Int. Conf. on Advanced Information and Communication Technology (ICAICT) pp 146-150.

[5] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. wang, T.Weyand, M. Andreetto and H.Adam 2017 MobileNets: Efficient Convolutional Neural Networks for Mobile Vision arXiv:1704.04861 [cs.CV]

[6] B. A. M. Ashqar, B. S. Abu-Nasser and S. S. Abu-Naser 2019 Plant Seedlings Classification Using Deep Learning Int. Journal of Academic Information Systems Research 3 7-14

[7] F. Miao, S. Zheng and B. Tao 2019 Crop Weed Identification System Based on Convolutional Neural Network IEEE 2nd Int. Conf. on Electronic Information and Communication Technology (ICEICT) pp 595-598.