MEAT CUTS CLASSIFICATION USING TRANSFER LEARNING & DEEP CONVOLUTIONAL NEURAL NETWORKS

Abhinay Tanguturi
Dublin City University
Dublin, Ireland
abhinay.tanguturi2@mail.dcu.ie

Sahithi Rayidi
Dublin City University
Dublin, Ireland
sahithi.rayidi2@mail.dcu.ie

***Abstract*—The applications of artificial intelligence in the agriculture sector, particularly in the food production sector, have seen a considerable level of growth in recent times. In this study, we propose CNN and deep learning methods to classify meat cuts taken from beef boning primals. The dataset in this paper was collected from the live production environment of the Irish beef industry in a span of 3 weeks. It consists of 8534 meat cut images and 8534 empty weight pan images with corresponding equivalent weights. In this research, CNN and VGG models were implemented along with techniques like data augmentation, stratified sampling, and transfer learning to overcome the class imbalance issue. However, VGGNet outperformed CNN with an accuracy of 98.45%. The highest precision, recall and f1-score obtained are 96% for VGGNet.**

***Index Terms*—Artificial Intelligence, CNN, VGG, Data Augmentation, Stratified Sampling, Transfer learning.**

I. INTRODUCTION

Over the last five years, from 2016 to 2021, data from Euromonitor indicates that the Republic of Ireland is the largest consumer of red meat in the European Union. The Irish industry can also be characterised as one having the highest export demand for red meat due to the industry's high level of international trade. Over the next few years the industry's revenues are expected to increase although there are potential effects like the introduction of custom checks. The EU-UK trade is the one effect for Irish meat industry, as the UK is one of Ireland's largest trading partners in terms of food and drink. Industry profit margins are expected to fluctuate in line with changes in meat prices, production volumes and export demand[10].

The industry, as well as buyers, are at times affected by activities like adulteration, etc., that can potentially lead to loss of income, consumption of meat prohibited in certain religions, unknowingly being exposed to food allergens, and general food safety concerns[4][11]. The meat production system is therefore challenged by the above mentioned activities and it is of great importance that meat products are intensively controlled, monitored, and inspected during their processing, storage, and distribution [1]. To overcome these issues, several analytical methods have been proposed and have subsequently been evaluated. Some of the techniques include the approaches of computer vision, deep learning and transfer learning, where object detection techniques are applied to detect the meat cut from the input image and then classify the meat cut into different categories.

Object detection can be explained as a subset of computer vision, which is an automated method for locating interesting objects in an image with respect to the background. This technique can be employed in the meat industries to reduce the processing time and also to prevent meat adulteration activities or human mistakes in classifying the meat.

A. *Problem & Motivation*

Currently, in meat industries, we can find a meat classification system that is used as a grading system that determines the quality, quantity and compositional attributes present in meat cuts and carcasses. This system grades meat into standard, prime and superior grades based on meat quality irrespective of the meat type[17]. The conceptual basis of why the classification system is used as a grading system is because there is always a different perception and expectations on meat quality from producers, consumers and retailers which also includes even their subsequent eating experience.

However, recent research tells us that the no. of categories of carcass and meat cuts has increased due to the variation between types and breeds of livestock, animal production systems and also due to the usages of different growth-enhancing technologies by different producers. This motivates us to develop an approach to a classification system that can primarily categorise the type of meat cuts based on its weight and some other clearly defined characteristics to ensure the quality, composition of meat, and consumer satisfaction.

B. Contribution

The aim of this research is to investigate a few models and techniques that can deal with large image data when it has a class imbalance issue. As a part of this research, two different CNN models are implemented, with a combination of different techniques like data augmentation, stratified sampling, and transfer learning to overcome the class imbalance issue and classify the images into five different categories. Also, the performance of Deep Convolutional Neural Networks (DCNN) is evaluated in terms of accuracy by varying the different parameters like the number of convolutional layer filters, filter size, and activation function for each layer in classifying the meat cuts based on their appearance and their weights.

C. Paper Structure

This paper is organised as follows: in section 2 an overview of the related work in object detection and classification in various food industries is explained; in section 3 we describe the data and its origin used in our approach; in section 4 we present the proposed transfer learning and deep learning based approaches to classify meat cuts; in section 5 results of proposed algorithms and their evaluation is explained ; we make concluding remarks in section 6 and the scope of future research is explained in section 7.

II. Related Research

Object detection is a popular technology in computer vision and image processing that deals with the detection of semantic objects of different classes in images and videos. Due to the advancement of technology, semantic object detection has seen significant growth in the meat industry. The main objective is to annotate images that contain meat for the purpose of organizing them into different categories. Meat image classification has been an area of interest in many current studies. Researchers have published several approaches to solving this kind of problem[12].

The paper [14] reports two sets of experiments: 1) food/non-food image classification, and 2) food category recognition. Two datasets consisting of 21643 images of food and non-food images were created for this experiment and are further split into three subsets where each one is considered for training, testing, and validation. Another dataset was also created with 4230 food images and 5428 non-food images for evaluating the model. A GoogLeNet model based on deep CNN was fine-tuned and trained using the image data in a deep learning framework - Caffe. The results are evaluated against the sensitivity, specificity, and accuracy of the developed model. However, more accuracy is attained for food/non-food classification than food recognition. The authors of the paper propose recognition of the food in the images with a multi-label approach as a possible research direction. Additionally they suggest the integration of contextual information to improve the accuracy and compare it with different other architectures.

The authors in [3] developed a model for chicken meat freshness identification using a convolutional neural network algorithm. A small image dataset consisting of fresh(6-8 Hrs) and rotten(21-23 Hrs) broiler chicken meat images of different sizes that are taken hours after slaughter were used in this approach. The classification of fresh and rotten chicken meat images is done in 3 main steps. The image acquisition, where the different sized breast pieces of chicken meat samples are gathered in RGB and preprocessed by cropping into different scales 200X200, 300X300, 400X400 and further using OpenCV converted them into grayscale images and then into binary images to calculate the threshold using the Otsu method. Furthermore, these are classified using the Ayam6Net, AlexNet, VGGNet, and GoogLeNet, The outcome of this approach was that Ayam6Net has outperformed all other 3 architectures and also performed better on the 400X400 pixel dataset.

The purpose in [13] was to design a convolutional neural network model to provide a food images collection as an outcome to distinguish the nutrition groups which people take in daily life. Dataset created is a combination of images taken from various datasets from the archive on the web and it consists of 16950 images of 11 groups with at least 1500 images in each group. The dataset is preprocessed by eliminating the irrelevant areas from the images and is then scaled up. Both the pre-trained models AlexNet and CaffeNet were finetuned and a similar structure was trained with the dataset. However, the outcome based on the presented test results is that fine-tuned models provided better results because of their generalization abilities. Besides, the trained structure can also be improved by increasing the number of training examples and can be used as a specific structure for the classification of nutrition groups. As an advancement, the effectiveness of Deep Convolutional Neural Network (DCNN) is examined for food photo recognition tasks in [18][9]. To tackle the problem, the best combinations of DCNN- related techniques are sought such as pre-training with the large-scale ImageNet data, fine-tuning, and activation features extracted from the pre-trained DCNN. Whereas, on the same data CNN, AlexNet and CaffeNet were compared by fine-tuning on the training parameters- momentum and learning rate.

The purpose in [19] is to develop a classification system that is based on the weight of the Cow using the CNN algorithms approach. The data used for this approach has about 40 thin and 40 fat cow images for training and few more images for testing. In this approach, the preprocessing of the images is done using canny edge detection, where the boundaries of the object(i.e Cow) are recognised and this method is also capable of reducing the noise. These preprocessed images are further passed into the CNN with backpropagation. So, the images are trained continuously until the defined standards are met. Based on the result of the training trials the EfficientNet and InceptionV3 performed well. For testing, the following metrics- precision, recall, accuracy, F1 score are compared. Their results show that InceptionV3 outperformed EfficientNet. The authors leave the extension of the model for various other parameters such as learning rate and batch size for future research.

In [5] a novel method is proposed to address the data imbalance problem in medical image datasets by using image complexity. The approach is applied on 4 datasets of medical images (i.e MRI, FFDM, SOKL, HEp-2 cell image) consisting of 63,505 images altogether where the suspicious samples are minority class in every dataset when compared to positive samples. To overcome the issue of minority class features getting neglected, effective perturbing operations are investigated to capture single-class-relevant features in medical images for minority classes in this approach.

The focus in [16] is all about the problem of classification using deep neural networks on 8 imbalanced datasets where 3 are image datasets consisting of 60,000 images that originally belonged to 100 classes but are classified into 20 superclasses for this approach and the remaining 5 documents datasets collected randomly from 20 Newsgroups consisting of approximately 600 documents in each group. In this approach, a novel function called MFE together with its improved version MSFA is proposed for the training of deep neural networks on the taken imbalanced datasets. This method had effectively captured the classification errors from both majority and minority classes equally and better than the conventional methods when compared. However, the authors intend to explore the effectiveness of this proposed loss function on different networks like DBN and CNN, in the future.

The authors in [2] discussed how the combination of different techniques or functions along with CNN can be used on data for a classification experiment that has class imbalance issues. A small dataset of vine leaves which consists of 5 different classes of leaves has been used to test two different approaches. As a first approach a simple CNN model was employed, while during the second approach different deep learning techniques like data augmentation, stratified sampling and transfer learning with state of the art CNN models like VGG are employed. However, CNN in combination with different deep learning techniques yielded better results than when a simple CNN is applied to the data.

In this paper, we use CNN and VGGNet, which are developed recently based on deep convolutional neural networks, in order to classify meat images with several cuts to one of the 5 categories defined in table 1.

III. Dataset

The data was taken from an Irish meat factory and on boning production line. Each individual removes a bone or a piece meat from the original primal product. It is then weighed, inspected and a picture of the product is taken. So, we have a total of 8534 meat images and 8534 empty weight-scale images along with an excel sheet containing attributes like productId, the weight of the meat cuts present in the images etc.

ProductId	No.of Meat Images	Average Weight
20001	1133	6.47
20002	14	8.87
20003	2283	5.77
20004	2241	1.40
20010	2863	7.81

Table 1 : Data

IV. Methodology

A. Background Removal

The extraction of meat cuts from the given two categories of images in grayscale was easier using the cv2.subtract() function. However, extracting the same meat cuts from the background in RGB colour format was very difficult and was not accurate because of the difference between pixel values.

So, here in our approach the “rembg” library was used to extract the meat cuts by eliminating the background from an image without using empty weighing scale images or converting them into grayscale. The ”rembg” library is a user defined library and has an advanced usage of alpha matting which means the extraction of foreground by eliminating the background in the image. This is done by creating soft boundaries to the foreground image more accurately without the loss of any part of the image.

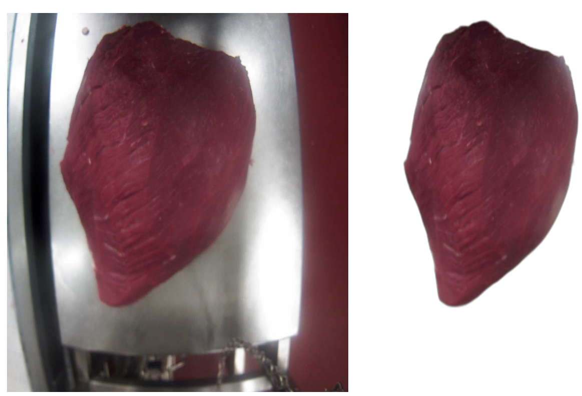

Figure 1. Example of meat image with and without background

B. Data Augmentation

Image data augmentation artificially expands and multiplies the original images present in the dataset to increase their number in the training set when there is a class imbalance issue [6]. Usually deep learning models like CNN are used to train the huge data to obtain a good performance of the model, but our data has a class imbalance issue where the applied CNN model tends to under-fit the training data of minority class and reduce the adaptability of the model to the test sample, which results in the bad performance on the test data.

Our data has about 8534 meat images out of which a class has only 14 images, so when a train-test split is done, there are not many of these minority class images in training data to train the model on their features, so the test results were not that great. So, to overcome the underfitting of these minority class images into the model and also to increase the adaptability and performance of the test set, image data augmentation was performed on the originally available 14 images using some affine transformation techniques like rotation, noise and horizontal flipping. After performing the above operations, originally given 14 images were increased to 151 images. These augmented images will increase the no. of images in the training sample which will lead to better adaptability of the test sample and a smooth training curve.

Figure 2. Example of Data Augmentation

C. Resizing

Usually, images which are fed into the neural network are reduced in dimensions, to decrease the processing time and to avoid the problem of under fitting[7]. Even if we consider an image of 224 x 224 x 3 pixels, when it is converted to one dimensional, it makes an input vector of the size of 150528. However, this 1 dimensional input vector is too large to be fed into the neural network. So, to decrease the processing time and to overcome the issues like underfitting, our data, which is of size 1600 x 1200 x 3, has been resized to 50 x 50 x 3.

D. Stratified Sampling

Stratified sampling is a technique in which the huge population is divided into different groups where data within the group have similar characteristics or features. This approach takes the group size into consideration while sampling to create the best sample that represents the entire population. So, this probability sampling technique can be used to sample the huge data as well as the data that have class imbalance issues. In this way, the data from minority classes cannot be dominated by the majority classes while creating a sample. This approach was used to sample our data at a ratio of 80% and 20% for train and test samples respectively as our data is both huge and has class imbalance issues.

ProductId	No.of Meat Images	Train(80%)	Test(20%)
20001	1133	905	228
20002	151	120	31
20003	2283	1825	458
20004	2241	1793	448
20010	2863	2289	574

Table 2: Results for stratified sampling

E. CNN

There are many machine learning models that can be used for classification tasks. But to classify image data Convolutional Neural Networks have always been famous because of their ability in segmenting the images and extracting the features. They also preserve the spatial information of images to overcome the affine transformation issues or any other effects that are previously performed on the images. As per our requirement convolutional layers can be added before dense layers and also at times reduces the training parameters to increase the speed of the learning process.

CNN's usually consist of convolutional layers, pooling layers, and fully connected layers. Generally, CNN takes images as an input and passes them through different sized kernels where convolutionals are performed to produce the vector as an output which is passed as input to the dense layers. Pooling layers are seen in between convolutional layers and their task is to reduce the size of the input.

In this research, a simple CNN model is applied to our data in two ways. In one, the model is applied to the preprocessed images without passing any meat product weights, whereas

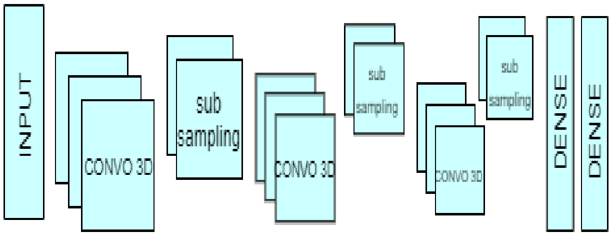

Figure 3. Overview of CNN architecture

in the second approach it is passed with meat weights.

The architecture of CNN without weights is as follows:

1) Three convolutional blocks with 32, 64, and 64 filters.
2) There are 3 max pooling layers placed after every convolutional block.
3) Two dense layers of 512 and 5 units each are stacked at the bottom.

However, the architecture of CNN with weights is similar to that of CNN without weights, but an additional input layer that takes product weights as an input is inserted into the model before the output layer.

F. Transfer Learning

Transfer Learning is mostly used in machine learning, where the approach is all about storing the knowledge gained while solving a problem and applying it to a similar problem[15]. Ideology is to use a model that is pre-trained on a problem to solve another similar kind of problem, mostly by changing the last few layers of the already pre-trained model. [2].

To perform transfer learning, the ImageNet dataset has been used in this approach. This dataset is a large visual dataset containing around 14,197,122 annotated images which can be categorized into more than 20,000 categories [8].

G. VGGNet

VGGNet can be defined as the Visual Geometry Group network which uses convolutional layers on top of each other which are replaced with small sized 3x3 kernels with increasing depth and also uses max-pooling layers in order to reduce the size or resolution of the image. In this network, there are also two dense layers with 4096 neurons in each filled by a soft-max classifier, which is a generalization of logistic regression to support the multi-class probability distribution[2]. Based on the number of weight layers in the net, there are two VGGNets, namely VGG16 and VGG19.

In this approach, the first step was to implement transfer learning by using the ImageNet dataset and then VGG19Net was applied. However, from the literature review, it was understood that training the last layers of VGGNet would provide good results. It is because convolutional layers would capture the low-level information that would be most useful during a classification problem. Whereas, the last layers in the net capture the high-level information that would be data specific. However, the classes related to the ImageNet dataset were removed and a flatten layer along with 5 dense layers each with 1024, 512, 256, 128 and 5 filters were added.

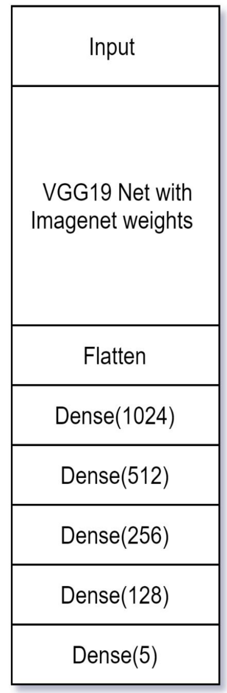

Figure 4. Architecture of VGG19

V. RESULTS

The results show that VGG19 performed well with a test accuracy of 98.45% when compared with CNN, whose test accuracies are 92% and 97%. However, the confusion matrix obtained for VGG19 shows that the product with ID 20004 is classified with 100% accuracy. And, the least classification accuracy of 77% was recorded by a product with Id 20002. This is the same class where the training images were increased by performing affine transformations. However, the remaining products with Ids 20003, 20001 and 20010 were classified with an accuracy of 98%,97% and 99% respectively.

Underfitting was the main problem faced during the early stages of CNN implementation. It was because of the class imbalance issue present and the random train-test split. However, several techniques were explored to choose data

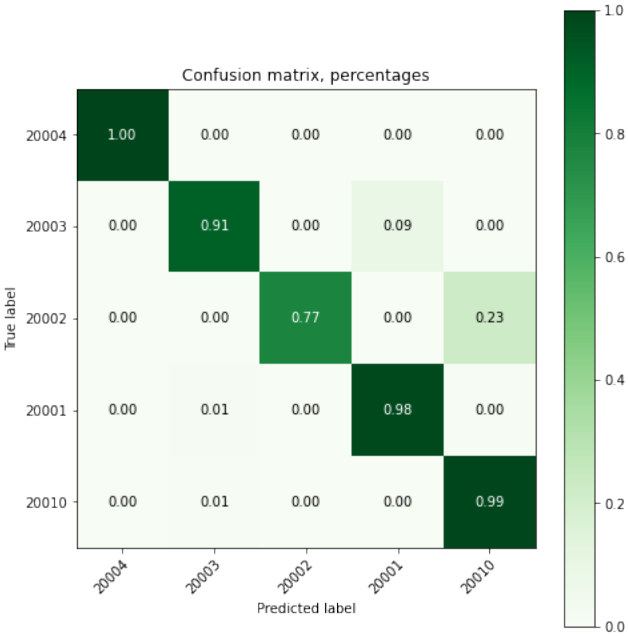

Figure 5. Confusion matrix of VGG19

augmentation and stratified sampling to solve the class imbalance and train-test sampling issues and to overcome the underfitiing of the CNN model. Due to the application of these techniques, the test accuracy of VGG19 is 98.45%. Whereas, the evaluation metrics precision, recall and f1-score are obtained at around 96%.

ProductId	precision	recall	f1-score	support
20004	1.00	1.00	1.00	449
20003	0.91	0.98	0.94	423
20002	0.77	1.00	0.87	24
20001	0.98	0.85	0.91	265
20010	0.99	0.98	0.99	578
accuracy			0.98	1739
macro avg	0.93	0.96	0.94	1739
weighted avg	0.97	0.97	0.97	1739

Figure 6. Precision, Recall and F1-Score of VGG19Net

The metrics in figure 5 are taken into account considering the imbalanced nature of the data to evaluate the model's performance. The precision metric is related to false positive rates in classification, recall is how much good data is missed and the f1-score is the harmonic mean of both the precision and the recall.

VI. Conclusion

In this research, we looked into different techniques to overcome the class imbalance issues which affect the models during classificaion. In experiments, we identified that deep learning models can be performed on the imbalance dataset using certain techniques like data augmentation, transfer learning, and stratified sampling. In this approach, the first experiment using a simple CNN model got an accuracy of around 92% without weights and 97% by passing the weights into the model. Whereas, in the second approach using the VGG19 model, we got an accuracy of 98.45%, which is very good results and the same for other metrics as well. The techniques used in this approach also contributed to the non-occurance of underfitting issues.

VII. Future Scope

In the future, we can also try to implement other imbalance techniques, like generating pertrubed images, or applying MFE (Mean False Error) or MSFE (Mean Square False Error) along with neural networks to this dataset to eliminate the class imbalance issue present. However, other transfer learning models and ensemble techniques can also be investigated and tried on this dataset to achieve more optimised results.

References

[1] Cristina Alamprese et al. "Identification and quantification of turkey meat adulteration in fresh, frozen-thawed and cooked minced beef by FT-NIR spectroscopy and chemometrics". In: *Meat Science* 121 (2016), pp. 175–181. DOI: https://doi.org/10.1016/j.meatsci.2016.06.018.

[2] Amjad Balawi et al. "Classification of a Small Imbalanced Dataset of Vine Leaves Images using Deep Learning Techniques". In: Nov. 2020.

[3] Calvin, G. B. Putra, and E. Prakasa. "Classification of Chicken Meat Freshness using Convolutional Neural Network Algorithms". In: (2020), pp. 1–6. DOI: 10.1109/3ICT51146.2020.9312018.

[4] Donna-Mareè Cawthorn, Harris A. Steinman, and Louwrens C. Hoffman. "A high incidence of species substitution and mislabelling detected in meat products sold in South Africa". In: *Food Control* 32.2 (2013), pp. 440–449. DOI: https://doi.org/10.1016/j.foodcont.2013.01.008.

[5] Long Gao et al. "Handling Imbalanced Medical Image Data: A Deep-Learning-Based One-Class Classification Approach". In: *Artificial Intelligence in Medicine* 108 (Aug. 2020), p. 101935. DOI: 10.1016/j.artmed.2020.101935.

[6] *How to Configure Image Data Augmentation in Keras*. https://machinelearningmastery.com/how-to-configure-image-data-augmentation-when-training-deep-learning-neural-networks/. Accessed: 2019-04-12.

[7] *Image Segmentation Using Convolutional Neural Network*. http://www.ijstr.org/final-print/nov2019/Image-Segmentation-Using-Convolutional-Neural-Network.pdf.

[8] *ImageNet*. https://paperswithcode.com/dataset/imagenet.

[9] Chang Liu et al. *DeepFood: Deep Learning-Based Food Image Recognition for Computer-Aided Dietary Assessment*. 2016. arXiv: 1606.05675 [cs.CV].

[10] *Meat Processing in Ireland - Market Research Report*. https://www.ibisworld.com/ireland/market-research-reports/meat-processing-industry/.

[11] Khadijah Nakyinsige, Yaakob Bin Che Man, and Awis Qurni Sazili. "Halal authenticity issues in meat and meat products". In: *Meat Science* 91.3 (2012), pp. 207–214. DOI: https://doi.org/10.1016/j.meatsci.2012.02.015.

[12] *Object Detection*. https://en.wikipedia.org/wiki/Object_detection.

[13] Gözde Özsert Yiğit and Melis OZYILDIRIM. "Comparison of convolutional neural network models for food image classification". In: *Journal of Information and Telecommunication* 2 (Mar. 2018), pp. 1–11. DOI: 10.1080/24751839.2018.1446236.

[14] Ashutosh Singla, Lin Yuan, and Touradj Ebrahimi. "Food/Non-food Image Classification and Food Categorization using Pre-Trained GoogLeNet Model". In: Oct. 2016, pp. 3–11. DOI: 10.1145/2986035.2986039.

[15] *Transfer Learning*. https://en.wikipedia.org/wiki/Transfer_learning.

[16] Shoujin Wang et al. "Training deep neural networks on imbalanced data sets". In: July 2016, pp. 4368–4374. DOI: 10.1109/IJCNN.2016.7727770.

[17] Edward Webb. "Description of carcass classification goals and the current situation in South Africa". In: *South African Journal of Animal Sciences* 45 (Jan. 2015), pp. 229–233. DOI: 10.4314/sajas.v45i3.1.

[18] Keiji Yanai and Yoshiyuki Kawano. "Food image recognition using deep convolutional network with pre-training and fine-tuning". In: June 2015, pp. 1–6. DOI: 10.1109/ICMEW.2015.7169816.

[19] Army Hudan Zhain et al. "Digital Image Processing to Determine Weight and Classification of Cow Weight with Deep Learning". In: *Psychology and Education Journal* 58 (2021). DOI: 10.17762/pae.v58i1.3752.