p-ISSN: 2460–7223
e-ISSN: 2964-9870

The Application Of The Convolution Neural Network Method Uses A Webcam To Analyze The Facial Expressions Of Problematic Students In The Counseling Guidance Unit
(Case Study At SMAN 1 Penengahan LampungSelatan)

1st Rendi Pratama
Computer Science Faculty
IIB Darmajaya
Lampung, Indonesia
rendi.19110065@mail.daramajaya.ac.id

2nd Rio Kurniawan
Computer Science Faculty
IIB Darmajaya
Lampung, Indonesia
riokurniawan@darmajaya.ac.id

3rd Triowali Rosandi
Computer Science Faculty
IIB Darmajaya
Lampung, Indonesia
triowali.andy@gmail.com

4th Nisar
Computer Science Faculty
IIB Darmajaya
Lampung, Indonesia
nisar@darmajaya.ac.id

Abstract— **Guidance and Counseling is a service provided to students to help them develop their potential optimally . Detecting students' facial expressions in the counseling room plays a crucial role in assisting counselors in understanding the emotional state of students who may need help, such as depression, anxiety, or stress, as they often find it difficult to express their feelings verbally. Therefore, this research will focus on 7 types of facial expressions: Anger, Disgust, Fear, Happiness, Neutral, Sadness, and Surprise. To classify these facial expressions, a Convolutional Neural Network (CNN) technique will be used, which identifies objects based on color and contours in an image. The aim of this research is to create a CNN model that can detect students' facial expressions during counseling sessions. In this study, the machine learning life cycle method is also employed as a stage in building the CNN model, starting with data collection with a total of 618 images, data cleaning, labeling the data, splitting the data into training and testing data with an 80% training data and 20% testing data ratio, creating the CNN architecture, training and evaluating the created model, and finally implementing it using a webcam. The results of this research show that the model achieved an accuracy of 33%. However, the facial expression detection features using the CNN model successfully detected students' facial expressions despite having a low prediction accuracy rate.**

Keywords— Convolutional Neural Network, Facial Expression Detection, Guidance Counseling, Machine Learning, Webcam

I. INTRODUCTION

Convolutional Neural Network (CNN) has become an important foundation in the field of deep learning, capable of solving object detection and object recognition problems with high computational efficiency [1][2]. CNN, specifically, is a method of image recognition and visual processing inspired by the workings of the human visual cortex, where filters are used to recognize patterns in images[3][4]. It has been proven that CNN has been successfully applied in various similar studies, including facial recognition [5], facial expression analysis [6], and emotion monitoring [7].

On the other hand, Guidance and Counseling (BK) is a service that provides encouragement to students to optimize their potential [8]. The success of this service really depends on the competence and professionalism of the guidance and counseling teacher or counselor, who must be able to meet students' needs in various aspects of their lives, including personal, social, academic and career aspects. To achieve this goal, BK services must be well designed in a systematic program [9].

Recent research in the field of facial detection has also had a significant impact. One of its main benefits is its potential application in the BK context. This research aims to analyze students' facial expressions during the guidance and counseling process. It is hoped that this facial expression analysis can help identify the emotions being experienced by students through their facial expressions.

[10] In this study, the Convolutional Neural Network (CNN) method was applied for web-based fugitive face detection. The experimental results demonstrate a detection accuracy rate of 77%, signifying a significant advancement in this particular application.

Another study [11] reported the implementation of CNN for emotion detection through human faces. Although they achieved an accuracy of 81.92% for training and 81.69% for testing, the study also revealed significant error rates. This shows the importance of factors such as proper lighting in the detection process.

[12] Azizi (2021) developed a model for automatic emotion detection using facial expression images. Their study achieved an accuracy rate of 82%, indicating promising potential for emotion recognition through facial expressions.

[13] Azhari (2021) integrated the CNN algorithm to detect human emotions based on facial expressions using a webcam device. The results revealed an accuracy rate within the range of 67-83%, which can be considered quite satisfactory.

[14] Darmawan and Supeno (2022) conducted research on facial expression classification using CNN in the context of online lectures. They succeeded in classifying student enthusiasm in online lecture learning with an accuracy rate of 64%. Although these results suggest accuracy could possibly be improved, this research makes an important contribution to understanding facial expressions in educational contexts.

With increasingly advanced technological developments and research, combining facial expression detection technology with the Convolutional Neural Network method promises great potential to enrich student guidance and counseling services. It is hoped that this research can play an important role in creating an educational environment that is more caring and responsive to students' emotional needs.

Detecting students' facial expressions in counseling sessions has an important role in helping counselors understand students' emotional conditions which may be difficult to express verbally[15]. Students dealing with emotional issues, such as depression, anxiety, or stress, often have difficulty communicating their feelings with words [16]. However, by using facial expression detection technology, counselors can interpret powerful nonverbal language to convey the emotions they feel[17]. The application of this technology has developed rapidly and provides promising potential to support the counseling guidance process. One of the methods used is Convolutional Neural Network (CNN), an intelligent technique that can identify patterns in images with high accuracy. By utilizing CNN, the system can analyze students' facial expressions, providing deep insight into students' emotional states. By combining facial expression detection technology, the Convolutional Neural Network (CNN) method, and the competence of counselors, this research opens up new opportunities to optimize student counseling services and has the potential to make a significant contribution in creating an educational environment that is responsive to students' emotional needs.

I. MATERIALS AND METHOD

Artificial Intelligence refers to systems that give computers the ability to think and make decisions like humans. In this case, AI is able to produce responses or outputs that resemble human intellectual abilities based on the data provided [18]. There are 2 sub-sections of Artificial Intelligence, namely Machine Learning and Deep Learning.

Machine learning is an integral part of AI, it is a method used to teach computers to make decisions without the need to create complex programs. Machine learning is divided into three main types, namely Supervised Learning, which utilizes labeled data; Unsupervised Learning, which looks for patterns in unlabeled data; and Reinforcement Learning, which allows the system to optimize its own algorithm based on previous rewards[19][20].

Additionally, Deep Learning is a way to teach computers using artificial neural networks that imitate human neural networks[21]. Deep learning can be used to develop models with various types of data such as images, text, video and sound [22]. For optimal training results, deep learning often utilizes GPUs. The use of deep learning has now grown rapidly and has been used in various applications, including speech recognition, images, machine translation, and intelligent systems that can perform better than humans[23].

Therefore based on the previous explanation the stages or process flow used as a methodology in this study are using the Machine Learning Life Cycle methodology by Gärtler et al [24]. The entire flow or stages of research can be seen in the following figure 1 along with the explanation.

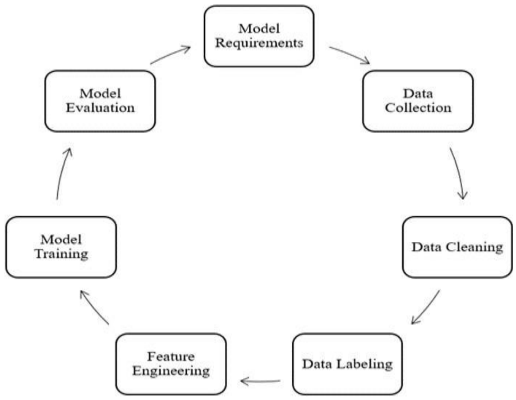

Fig. 1. Machine Leraning Life Cycle

A. *Model Requirements*

In the model requirements phase, CNN architecture selection is carried out, data collection, and the amount of data to be retrieved will then be implemented into the CNN deep learning development.

B. *Data Collection*

At the data collection stage, images of students' faces were collected using a smartphone camera and careful lighting. Each facial image is then annotated with a label containing information about the student's name and facial attributes such as expression. At this stage, data collection is carried out based on the theoretical basis regarding facial expressions where Facial expressions are a powerful way to convey emotions without words. These expressions involve facial movements and positions that reflect a person's feelings. Facial expressions play an important role in nonverbal communication and social interactions [17]. It helps individuals share emotions, build relationships, and convey their wants and needs. Facial expressions can be described in several types:

- **Anger:** Visible through raised eyebrows, tense lips, and narrowed eyes. This indicates anger, frustration, or a desire to confront a threat.
- **Disgust:** Seen by a wrinkled nose, downward curved lips, and a facial expression that reflects discomfort or displeasure with something.
- **Fear:** Visible by raised eyebrows, wide eyes, and an open mouth. It reflects the fear, anxiety, or threat felt by the individual.
- **Happy:** Seen by crinkled or blinking eyes, a smile, and a happy or satisfied facial expression. It reflects joy, satisfaction, or positive feelings.
- **Neutral:** This facial expression is flat and does not convey any special emotion. Even though they appear neutral, individuals may still have feelings or emotions they are experiencing.
- **Sad:** Visible from eyebrows that are pulled down, eyes that are tired or watery, and lips that are hanging or pulled down. It reflects sadness, disappointment, or feelings of sadness.
- **Surprise:** Seen by widened eyes, raised eyebrows, and an open mouth. It reflects feelings of surprise or being unprepared for the unexpected.

C. Data Cleaning

In the Cleaning phase, the data that has been collected is cleaned by performing or applying pixel resizing to 48x48 which is used to change the image pixel value scale to a range between 0 and 1. This is useful in normalizing image data before model training.

D. Data Labeling

The labeling process does not need to be done manually because the data used is already labeled when the dataset is taken or collected. So at this stage what is done is to change the labels which were previously in text or string form into an array, namely ('angry': 0, 'disgust': 1, 'fear': 2, 'happy': 3, 'neutral': 4, 'sad': 5, 'surprise': 6).

E. Feature Engineering

The Feature Engineering process is carried out by breaking down the data into training and testing data. For training data, 80% of the total data is training data and 20% of the total testing data and determines the data used for validation which is different for each class.

F. Model Training

Model training is carried out to train the Neural Network model that has been created. Training is carried out so that the model can work optimally, where in the training process the training data will be divided into two parts, namely the training set and the test set.

However Neural networks are a way to program computers that can operate similarly to the human brain. The main purpose of this neural network is to perform functions that the human brain can perform, such as problem solving and learning [25]. Therefore there are several types of neural networks, including Feedforward Neural Network, Radial Basis Function Neural Network, Recurrent Neural Network, Convolutional Neural Network, Modular Neural Network [26] [27] [28] [29] [30].

CNN is a representation of human neural network architecture and belongs to deep learning algorithms for computer vision [31]. CNNs have proven high performance in various computer vision tasks, such as image classification [32], image segmentation [33], image retrieval [34], object detection [35], image annotation [36], face recognition [37], pose estimation [38], traffic sign recognition [39], voice processing [40], and others. CNNs consist of perceptrons that learn from input and perform self-optimization operations. Apart from that, according to [41] Convolutional Neural Network has 3 layers namely:

- **Convolutional Layer:** This layer plays a key role in CNN. Filters (kernels) are used to learn from input data. Each filter has a small size, such as 3x3, 5x5, or 7x7, and moves across the input image. Depth, stride, and padding are factors that influence the output size of a convolution layer.
- **Pooling Layer:** After the convolution layer, a pooling layer is used to reduce image dimensions, speed up computation, and overcome overfitting. Pooling layers generally use a 2x2 filter and consist of two types: max pooling (taking the maximum value) and average pooling (taking the average value).
- **Fully Connected Layer:** This layer converts data into one dimension or "flatten." In this layer, weights, biases, and activation functions are initialized to calculate the output between the input and the perceptron.

There is also known a Activation Function where used in neural networks to calculate the total number of weights and biases, which can be used to decide whether neurons can be activated or not [42]. Some of the activation functions that commonly used are:

- **ReLU (Rectified Linear Unit):** ReLU are commonly used in hidden layers, especially in Convolutional Neural Networks (CNN) or Deep Learning [43]. The value range of ReLU $\{0. -\infty\}$. If the value of $x < 0$ the resulting output is 0 and if the value of $x > 0$ the output produced is the value of x itself. So to find the value using the ReLU activation function, you can see the following equation.
- **Softmax:** Softmax is a function used in neural networks to calculate probability. The result is a number between 0 and 1, where the total probability is always 1. Softmax is often used to classify many classes and is usually applied at the output layer.

G. Model Evaluation

Model evaluation at this stage will confirm that a model can predict new data quite well. The data that will be included in this evaluation is testing data where the testing data has 20% of all images consisting of seven classes (angry, disgusted, afraid, happy, neutral, sad and shocked). After the model has finished predicting the data, the next step is to use two types of evaluation matrices to see the performance of the model that has been created. The types of matrix are accuracy matrix and confusion matrix to visualizes the performance of classification algorithms using matrix data, by comparing the predicted classification with the actual classification in the form of false positives, true positives, false negatives and true negatives[44].

II. RESULT AND DISCUSSION

The results and discussion of this research illustrate the application of the CNN method to analyze students' facial expressions in the Guidance Counseling unit. In this research, the results of the data collection carried out obtained a total of 618 images. Based on the results of the data collection that has been carried out, the researcher divides the images that will be used for training and testing, where the training data is 494 images and the testing data is 124 images and the number of images is further divided into seven parts or seven classes of facial expressions, namely "angry" facial expressions. /angry" consists of a total of 60 images, with 46 images used for training (train set) and 14 images for testing (test set). The “disgust” class has a total of 58 images, of which 44 images are used for training and 14 images for testing. The “fear” class has a total of 58 images, with 44 images in the training set and 14 images in the testing set. The "happy" class has a total of 95 images, consisting of 72 images in the training set and 23 images in the testing set. The “regular/neutral” class includes a total of 84 images, with 64 images in the training set and 20 images in the test set. The “sad” class has a total of 97 images, with 74 images in the training set and 23 images in the testing set. Finally, the “shock/surprise” class has a total of 66 images, with 50 images in the training set and 16 images in the testing set.

Then the researcher performs data cleaning safely at each iteration, the image is resized to the target size (48x48 pixels), and converted to gray scale and saved in a directory. Apart from that, researchers also apply Data Augmentation techniques to help reduce overfitting in the model, image data is increased in amount with variations produced by augmentation which improves model performance and data normalization which helps process images well in the model.

After the data has been cleaned, the next step is to label the data to give labels to each class of facial expressions (angry, disgust, fear, happy, neutral, sad and surprised), but because the data has been given a label at the time of data collection, this data labeling is carried out by replacing or converting class labels which were previously in the form of text or string data types into labels in the form of numbers or into class labels with array data types (angry': 0, 'disgust': 1, 'fear': 2, 'happy': 3, 'neutral': 4, 'sad': 5, 'surprise': 6).

Furthermore, in modeling the CNN architecture, researchers used Convolutional Layer (Conv2D), MaxPooling Layer (MaxPooling2D), Convolutional Layer (Conv2D), Flatten Layer, Dense Layer (Fully Connected), Dropout Layer and Dense Layer (Output Layer). In creating the Convolutional Neural Network (CNN) model, the following layers were used: Convolutional Layer (Conv2D) with 32 3x3 filters, MaxPooling Layer (MaxPooling2D) to reduce dimensions, additional Convolutional Layer with 64 and 128 filters for more complex feature extraction, Flatten Layer for data vectorization, Dense Layer (Fully Connected) with 512 neurons, Dropout Layer to avoid overfitting, and Dense Layer (Output Layer) with the number of neurons according to the facial expression class. This model is able to identify key features in images and provide results in the form of expression class probabilities that match the input image.

TABLE I. MODELING CNN ARCHITECTURE

Layer (type)	Output Shape	Param #
conv2d(Conv2D)	(None,46,46,32)	320
max_pooling2d(MaxPooling2D)	(None,23,23,32)	0
conv2d_1(Conv2D)	(None,21,21,64)	18496
max_pooling2d_1(MaxPooling2D)	(None,10,10,64)	0
conv2d_2(Conv2D)	(None,8,8,128)	73856
max_pooling2d_2(MaxPooling2D)	(None,4,4,128)	0
flatten(Flatten)	(None,2048)	0
dense(Dense)	(None,512)	1049088
dropout(Dropout)	(None,512)	0
dense_1(Dense)	(None,7)	3591

Therefore, based on the table above, the architecture created has an artificial neural network architecture with three convolution layers (Conv2D) followed by MaxPooling layers, followed by a Flatten layer, then two fully connected layers (Dense), and one dropout layer. The number of parameters in each layer describes how many parameters can be adjusted during model training. Which is then trained to produce the following results from CNN architecture training. The training process for the Convolutional Neural Network (CNN) model with the architecture provided lasts for 20 epochs. Initially, the model experienced low performance with limited accuracy, especially on the validation dataset. During subsequent iterations, a slight improvement in accuracy on the training dataset was seen, but validation results were still volatile. Despite efforts in adjusting the model weights, the performance is still unstable and has not reached a sufficient level on this facial expression classification task. The result of model training can be show from the performance of the learning state from training accuracy and loss accuracy

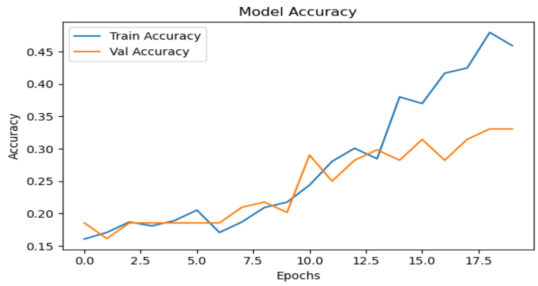

Fig. 2. Model Accuracy

From the training results, the graph in Figure 4.13 shows the trend of changes in training and validation accuracy over 20 epochs. Initially, there were fluctuations in training and validation accuracy, indicating initial adaptation of the model to the data. Then, we see a more steady increase in training accuracy as the epoch progresses. However, validation accuracy has variations and does not always follow the same upward trend as training accuracy. Although training accuracy continues to improve, validation accuracy does not always follow that trend and can experience fluctuations or even stagnation at some epochs. The following table shows a comparison of accuracy between test data and validation data

TABLE II. MODEL TRAINING ACCURACY

Epoch	Training ACC	Validation ACC
1	0.1606	0.1855

2	0.1707	0.1613
3	0.187	0.1855
4	0.1809	0.1855
5	0.189	0.1855
6	0.2053	0.1855
7	0.1707	0.1855
8	0.187	0.2097
9	0.2093	0.2177
10	0.2175	0.2016
11	0.2439	0.2903
12	0.2805	0.25
13	0.3008	0.2823
14	0.2846	0.2984
15	0.3801	0.2823
16	0.3699	0.3145
17	0.4167	0.2823
18	0.4248	0.3145
19	0.4797	0.3306
20	0.4593	0.3306

In the table above, the accuracy evaluation for 20 epochs in facial expression classification analysis is depicted. In the initial five epochs, there were significant fluctuations in training accuracy, indicating early adaptation of the model to the training data. However, validation results in this period still varied. In the 6th to 15th epochs, more consistent improvements in training accuracy were seen, reflecting the model's ability to extract relevant features from the training data. Although training accuracy tends to improve, validation accuracy remains variable. However, at the end of this period, there was a significant improvement in validation accuracy, indicating progress in model generalization. At the end of training (16th to 20th epoch), training accuracy continued to improve, while validation accuracy tended to stagnate, indicating possible overfitting where performance on training data is better than never-before-seen data. Meanwhile, the performance of the model is based on the loss results which can be seen in Figure below:

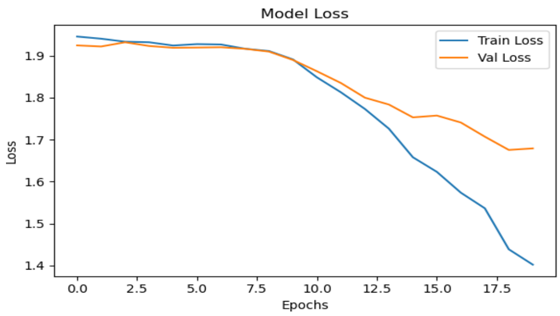

Fig. 3. Model Loss

Based on the image above, the results of model training for 20 epochs are represented through a loss graph which shows changes in training and validation loss values. At the beginning of training, there are fluctuations in training and validation loss values, reflecting the initial adaptation of the model to the data. Furthermore, as the epoch progresses, it can be seen that the training loss value tends to decrease steadily, indicating an increase in the effectiveness of the model in reducing prediction errors on the training data. However, as is the case with accuracy, validation loss values do not always follow the same trend as training loss. The following table shows a comparison of loss between test data and training data.

TABLE III. MODEL LOSS ACCURACY

Epoch	Training Loss	Validation Loss
1	1.9455	1.9245
2	1.9404	1.9218
3	1.9333	1.9319
4	1.9319	1.9231
5	1.9242	1.9187
6	1.9275	1.9193
7	1.9266	1.9200
8	1.9165	1.9163
9	1.9110	1.9097
10	1.8911	1.8897
11	1.8482	1.8628
12	1.8126	1.8348
13	1.7730	1.7999
14	1.7260	1.7835
15	1.6579	1.7530
16	1.6231	1.7573
17	1.5735	1.7407
18	1.5364	1.7070
19	1.4387	1.6754
20	1.4021	1.6790

From the table above, it can be seen that during the first five epochs, there were significant fluctuations in training loss and validation loss, indicating the initial adaptation stage of the model to the training data. However, as time goes by, the training loss consistently decreases, reflecting the model's ability to reduce prediction errors on the training data. Although validation loss also shows a decreasing trend, it is not always parallel to training loss, indicating the complexity of pattern recognition may be more difficult on validation data and the potential for overfitting. Overall, the analysis of these tables provides important insights into the training dynamics of the model and potential improvements in facial expression classification predictions.

From Table 1 and Table 2, it can be seen that there is an increase in training accuracy as the epoch progresses, and training loss consistently decreases. However, validation accuracy and validation loss do not always follow the same trend, highlighting the challenges of generalizing models to never-before-seen data. These variations may indicate potential overfitting or difficulty in fitting validation data. Therefore, this analysis emphasizes the importance of focusing on model generalization and preventing overfitting in subsequent development. After analyzing the results of the CNN architectural modeling that has been created, the next step is to evaluate the performance of the model that has been created using the confusion matrix which can be seen in the following image.

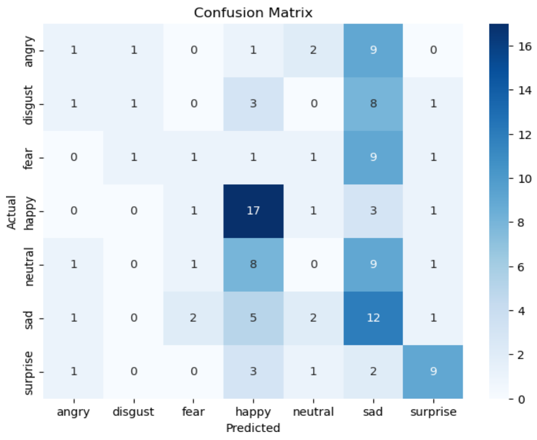

Fig. 4. Model Accuracy

In a detailed analysis of the model results using the confusion matrix, the researchers present the classification performance for each class of facial expressions. In the "Angry" class, there was 1 case that was predicted correctly (True Positive), 1 case that should have been "Angry" but was predicted as another class (False Negative), and 13 cases that were incorrectly predicted as "Angry" (False Positive), the "Disgust" class, there is 1 True Positive case, 3 False Negative cases, and 11 False Positive cases, the "Fear" class has 1 True Positive case, 1 False Negative case, and 12 False cases Positive, class "Happy", there are 17 cases of True Positive, 6 cases of True Negative, and 21 cases of False Positive, class "Sad" (Sad) has 2 cases of True Positive, 5 cases of False Negative, and 5 cases of False Positive , the "Surprise" class includes 9 True Positive cases, 3 False Negative cases, and 18 False Positive cases and the "Neutral" class has 1 True Positive case, 8 False Negative cases, and 10 False Positive cases. Based on the confusion matrix above, a classification report is obtained which can be seen in the following table

TABLE IV. CLASSIFICATION REPORT

Class	precision	recall	f1-score
angry	0.2	0.07	0.11
disgust	0.33	0.07	0.12
fear	0.2	0.07	0.11
happy	0.45	0.74	0.56
neutral	0	0	0
sad	0.23	0.52	0.32
surprise	0.64	0.56	0.6

The classification evaluation results show the model performance on the dataset. Overall accuracy is about 33%. More detailed results for each emotion class show that the model has varying degrees of success. The model performs well in recognizing “happy” expressions with an accuracy rate of around 45%, but has lower performance in recognizing “neutral” expressions with an accuracy of 0%. Likewise, performance in recognizing the emotions "angry" and "fear" was also low with an accuracy of around 20%, while the emotion "disgust" had an accuracy of 33%. The emotions “sad” and “surprise” had an accuracy of about 23% and 64% respectively. In the overall analysis, there was variation in the model's ability to recognize various facial expressions.

Another evaluation is to test the model using new data to find out whether the model is able to carry out classification based on the model that has been created as in the image below.

Fig. 5. New Data Testing

From the results of model testing using 30 test data which produces different accuracy. To see more clearly, see the following table

TABLE V. NEW DATA TESTING

No	ACC	Prediction	Actual
1	43.79%	Sad	Disgust
2	29.47%	Surprise	Disgust
3	50.40%	Neutral	Disgust
4	50.30%	Sad	Disgust
5	22.08%	Disgust	Disgust
6	69.26%	Neutral	Surprise
7	39.76%	Sad	Surprise
8	63.81%	Angry	Surprise
9	39.08%	Angry	Surprise
10	60.03%	Happy	Surprise
11	29.09%	Sad	Fear
12	33.35%	Surprise	Fear
13	48.19%	Sad	Fear
14	35.87%	Sad	Fear
15	79.23%	Sad	Fear
16	24.33%	Fear	Neutral
17	62.34%	Neutral	Neutral
18	27.09%	Sad	Neutral
19	22.62%	Neutral	Neutral
20	27.70%	Angry	Neutral
21	35.06%	Sad	Happy
22	69.78%	Neutral	Happy
23	25.59%	Sad	Happy
24	42.65%	Sad	Happy
25	54.53%	Neutral	Happy
26	31.41%	Sad	Sad
27	68.15%	Sad	Sad
28	57.66%	Disgust	Sad
29	40.78%	Angry	Sad
30	37.43%	Sad	Sad

The results of testing using 30 new data as in the table above are 5 data that were predicted to be correct, namely images 5, 17, 19, 26, 27 and 30 and the remaining 24 data were predicted to be incorrect.

III. CONCLUSIONS AND RECOMMENDATIONS

This research applies Convolutional Neural Network (CNN) to analyze students' facial expressions in Guidance and Counseling (BK) sessions with the aim of understanding their feelings and reactions. The results show a number of problems, including unbalanced emotion classes, overfitting of the model, and unsatisfactory model performance with a test accuracy of around 33%. Some emotional classes also face challenges. Therefore, improvements are needed through adding data, selecting more complex models, and tuning training parameters. Collaboration with experts in psychology or human behavior science can improve students' emotional recognition in the guidance and counseling context. This has the potential to increase the effectiveness of technology such as CNN in understanding and managing students' emotions during guidance and counseling sessions.

Acknowledgment

Thank you to the entire ICITB committee, especially to LP2M Darmajaya who has published these proceedings.

References

[1] L. Jiao *et al.*, "A survey of deep learning-based object detection," *IEEE Access*, vol. 7, hal. 128837–128868, 2019, doi: 10.1109/ACCESS.2019.2939201.

[2] A. Bouguettaya, A. Kechida, dan A. M. Taberkit, "A Survey on Lightweight CNN-Based Object Detection Algorithms for Platforms with Limited Computational Resources," *Int. J. Informatics Appl. Math.*, vol. 2, no. 2, hal. 28–44, 2019.

[3] J. Dapello, T. Marques, M. Schrimpf, F. Geiger, D. D. Cox, dan J. J. DiCarlo, "Simulating a primary visual cortex," *NeurIPS*, no. NeurIPS 2020, hal. 1–15, 2020.

[4] G. W. Lindsay, "Convolutional neural networks as a model of the visual system: Past, present, and future," *J. Cogn. Neurosci.*, vol. 33, no. 10, hal. 2017–2031, 2021, doi: 10.1162/jocn_a_01544.

[5] Y. Hartiwi, E. Rasywir, Y. Pratama, dan P. A. Jusia, "Eksperimen Pengenalan Wajah dengan fitur Indoor Positioning System menggunakan Algoritma CNN," *J. Khatulistiwa Inform.*, vol. 22, no. 2, hal. 109–116, 2020.

[6] P. P. Kusdiananggalih, E. Rachmawati, dan R. Risnandar, "Pengenalan Ekspresi Wajah Dari Cross Dataset Menggunakan Convolutional Neural Network (CNN)," *eProceedings Eng.*, vol. 8, no. 2, 2021.

[7] P. A. Nugroho, I. Fenriana, dan R. Arijanto, "Implementasi Deep Learning Menggunakan Convolutional Neural Network (Cnn) Pada Ekspresi Manusia," *Algor*, vol. 2, no. 1, hal. 12–21, 2020.

[8] Y. Primawati dan Y. Yunita, "Optimizing the Role of Guidance and Counseling Teachers in the Covid-19 Pandemic," *J. Ilm. Az-Ziqri …*, hal. 14–23, 2021, [Daring]. Tersedia pada: http://www.journal.iai-agussalimmetro.ac.id/index.php/AISE/article/download/393/149

[9] D. Ramadani dan H. Herdi, "Studi Kepustakaan Mengenai Kinerja Guru Bimbingan Dan Konseling Dalam Perencanaan Program Bimbingan Dan Konseling Di Sekolah," *J. EDUKASI J. Bimbing. Konseling*, vol. 7, no. 1, hal. 42–52, 2021.

[10] C. Widi Wiguna, J. Dedy Irawan, dan M. Orisa, "Penerapan Metode Convolutional Neural Network Pada Aplikasi Deteksi Wajah Buronan Berbasis Web," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 6, no. 2, hal. 1051–1058, 2023, doi: 10.36040/jati.v6i2.5438. "189-Article Text-1645-2-10-20230601."

[12] F. Azizi, "Deteksi Emosi Menggunakan Citra Ekspresi Wajah Secara Otomatis," *Informatics Eng.*, hal. 1–44, 2021, [Daring]. Tersedia pada: https://dspace.uii.ac.id/handle/123456789/35834 "1562."

[14] D. Darmawan dan H. Supeno, "Klasifikasi Ekspresi Wajah Menggunakan Metode Convolutional Neural Network (Studi Kasus Kuliah Daring)," *Pasinformatik*, vol. 1, no. 1, hal. 1–8, 2022, [Daring]. Tersedia pada: https://journal.unpas.ac.id/index.php/pasinformatik

[15] A. G. Aishwarya Gowda, H. K. Su, dan W. K. Kuo, "Monitoring Panic Situations in Students and Alerting Using Artificial Intelligence," *Sensors Mater.*, vol. 34, no. 9, hal. 3503–3512, 2022, doi: 10.18494/SAM3927.

[16] L. Santl, L. Brajkovic, dan V. Kopilaš, "Relationship between Nomophobia, Various Emotional Difficulties, and Distress Factors among Students," *Eur. J. Investig. Heal. Psychol. Educ.*, vol. 12, no. 7, hal. 716–730, 2022, doi: 10.3390/ejihpe12070053.

[17] F. M. Talaat, "Real-time facial emotion recognition system among children with autism based on deep learning and IoT," *Neural Comput. Appl.*, vol. 35, no. 17, hal. 12717–12728, 2023, doi: 10.1007/s00521-023-08372-9.

[18] T. Markiewicz dan J. Zheng, "Getting Started with Artificial Intelligence," 2020.

[19] X. Luo, "Efficient English text classification using selected Machine Learning Techniques," *Alexandria Eng. J.*, vol. 60, no. 3, hal. 3401–3409, 2021, doi: 10.1016/j.aej.2021.02.009.

[20] X. Zhou, H. Liu, F. Pourpanah, T. Zeng, dan X. Wang, "A Survey on Epistemic (Model) Uncertainty in Supervised Learning: Recent Advances and Applications," *Neurocomputing*, vol. 489, hal. 449–465, 2022, doi: 10.1016/j.neucom.2021.10.119.

[21] A. Goel, A. K. Goel, dan A. Kumar, "The role of artificial neural network and machine learning in utilizing spatial information," *Spat. Inf. Res.*, vol. 31, no. 3, hal. 275–285, 2023, doi: 10.1007/s41324-022-00494-x.

[22] A. J. Moshayedi, A. S. Roy, A. Kolahdooz, dan Y. Shuxin, "Deep Learning Application Pros And Cons Over Algorithm," *EAI Endorsed Trans. AI Robot.*, vol. 1, hal. 1–13, 2022, doi: 10.4108/airo.v1i.19.

[23] I. Hussain, R. Ahmad, S. Muhammad, K. Ullah, H. Shah, dan A. Namoun, "PHTI: Pashto Handwritten Text Imagebase for Deep Learning Applications," *IEEE Access*, vol. 10, no. October, hal. 113149–113157, 2022, doi: 10.1109/ACCESS.2022.3216881.

[24] M. Gärtler, V. Khaydarov, B. Klöpper, dan L. Urbas, "The Machine Learning Life Cycle in Chemical Operations – Status and Open Challenges," *Chemie-Ingenieur-Technik*, vol. 93, no. 12, hal. 2063–2080, 2021, doi: 10.1002/cite.202100134.

[25] M. T. Suwindra, E. Ernawati, dan A. Erlansari, "Analisis Kemiripan Jenis Burung Menggunakan Siamese Neural Network," *Rekursif J. Inform.*, vol. 9, no. 2, hal. 193–205, 2021.

[26] J. Rezaeenour, M. Ahmadi, H. Jelodar, dan R. Shahrooei, "Systematic review of content analysis

algorithms based on deep neural networks," *Multimed. Tools Appl.*, vol. 82, no. 12, hal. 17879–17903, 2023, doi: 10.1007/s11042-022-14043-z.

[27] J. Jiao, J. Liu, F. Shao, X. Xie, dan X. Chen, "Numerical and experimental investigation on longitudinal drift motion of a novel regular tetrahedron closure device in casting process," *Ain Shams Eng. J.*, vol. 13, no. 6, hal. 101783, 2022, doi: 10.1016/j.asej.2022.101783.

[28] O. Iparraguirre-Villanueva *et al.*, "Text prediction recurrent neural networks using long short-term memory-dropout," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 29, no. 3, hal. 1758–1768, 2023, doi: 10.11591/ijeecs.v29.i3.pp1758-1768.

[29] A. Shah *et al.*, "A comprehensive study on skin cancer detection using artificial neural network (ANN) and convolutional neural network (CNN)," *Clin. eHealth*, vol. 6, hal. 76–84, 2023, doi: 10.1016/j.ceh.2023.08.002.

[30] M. K. Dahouda dan I. Joe, "Neural Architecture Search Net-Based Feature Extraction with Modular Neural Network for Image Classification of Copper/ Cobalt Raw Minerals," *IEEE Access*, vol. 10, no. June, hal. 72253–72262, 2022, doi: 10.1109/ACCESS.2022.3187420.

[31] J. Ganesan, A. T. Azar, S. Alsenan, N. A. Kamal, B. Qureshi, dan A. E. Hassanien, "Deep Learning Reader for Visually Impaired," *Electron.*, vol. 11, no. 20, hal. 1–22, 2022, doi: 10.3390/electronics11203335.

[32] J. Huertas-Tato, A. Martín, J. Fierrez, dan D. Camacho, "Fusing CNNs and statistical indicators to improve image classification," *Inf. Fusion*, vol. 79, hal. 174–187, 2022, doi: 10.1016/j.inffus.2021.09.012.

[33] M. Heidari *et al.*, "HiFormer: Hierarchical Multi-scale Representations Using Transformers for Medical Image Segmentation," *Proc. - 2023 IEEE Winter Conf. Appl. Comput. Vision, WACV 2023*, hal. 6191–6201, 2023, doi: 10.1109/WACV56688.2023.00614.

[34] A. Naeem, T. Anees, K. T. Ahmed, R. A. Naqvi, S. Ahmad, dan T. Whangbo, "Deep learned vectors' formation using auto-correlation, scaling, and derivations with CNN for complex and huge image retrieval," *Complex Intell. Syst.*, vol. 9, no. 2, hal. 1729–1751, 2023, doi: 10.1007/s40747-022-00866-8.

[35] K. Zeng, Q. Ma, J. W. Wu, Z. Chen, T. Shen, dan C. Yan, "FPGA-based accelerator for object detection: a comprehensive survey," *J. Supercomput.*, vol. 78, no. 12, hal. 14096–14136, 2022, doi: 10.1007/s11227-022-04415-5.

[36] J. Ma, Y. Ushiku, dan M. Sagara, "The Effect of Improving Annotation Quality on Object Detection Datasets: A Preliminary Study," *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, vol. 2022-June, hal. 4849–4858, 2022, doi: 10.1109/CVPRW56347.2022.00532.

[37] A. J. A. Albdairi *et al.*, "Face Recognition Based on Deep Learning and FPGA for Ethnicity Identification," *Appl. Sci.*, vol. 12, no. 5, 2022, doi: 10.3390/app12052605.

[38] Q. Dang, J. Yin, B. Wang, dan W. Zheng, "Deep learning based 2D human pose estimation: Asurvey," *Tsinghua Sci. Technol.*, vol. 24, no. 6, hal. 663–676, 2019, doi: 10.26599/TST.2018.9010100.

[39] N. Youssouf, "Traffic sign classification using CNN and detection using faster-RCNN and YOLOV4," *Heliyon*, vol. 8, no. 12, 2022, doi: 10.1016/j.heliyon.2022.e11792.

[40] S. Dua *et al.*, "Developing a Speech Recognition System for Recognizing Tonal Speech Signals Using a Convolutional Neural Network," *Appl. Sci.*, vol. 12, no. 12, 2022, doi: 10.3390/app12126223.

[41] T. Bezdan dan N. Bačanin Džakula, "Convolutional Neural Network Layers and Architectures," no. January, hal. 445–451, 2019, doi: 10.15308/sinteza-2019-445-451.

[42] V. S. Tomar, "A critical evaluation of activation functions for autoencoder neural networks Vikalp Singh Tomar".

[43] J. Feng dan S. Lu, "Performance Analysis of Various Activation Functions in Artificial Neural Networks," *J. Phys. Conf. Ser.*, vol. 1237, no. 2, 2019, doi: 10.1088/1742-6596/1237/2/022030.

[44] L. E. Pomme, R. Bourqui, R. Giot, dan D. Auber, "Relative Confusion Matrix: Efficient Comparison of Decision Models," *Proc. Int. Conf. Inf. Vis.*, vol. 2022-July, hal. 98–103, 2022, doi: 10.1109/IV56949.2022.00025.