© July 2025 | IJIRT | Volume 12 Issue 2 | ISSN: 2349-6002

Emotion Recognition using Deep Learning for Behavioral Analysis in human and Decision-Making

Abhishek Kumar
Institute of Engineering and Technology, Ayodhya

***Abstract*—A comprehensive understanding of the factors influencing risk aversion is essential in fields such as economics, psychology, and finance. Despite extensive research, the specific mechanisms that drive risk aversion are not yet fully elucidated. This work explores the relationship between facial expressions and risk aversion in decision-making to better understand the factors influencing risk-taking behavior. By examining individuals' facial expressions as they make decisions involving various levels of risk, we aim to reveal the potential role of nonverbal cues in risk aversion and decision-making processes. The findings from this research will offer important insights across various fields and advance our understanding of human behavior and decision-making. Specifically, I aim to identify the emotional states reflected in these expressions and assess their correlation with observed risk preferences.**

I. INTRODUCTION

Decision-making is a fundamental aspect of human life, with individuals constantly navigating choices involving varying degrees of risk. Risk aversion, the tendency to favor certain outcomes over potentially higher-yielding but uncertain ones, plays a crucial role in these decisions. A comprehensive understanding of the factors influencing risk aversion is essential across diverse fields, including psychology, economics, and finance. While risk aversion has been extensively studied, the underlying mechanisms remain incompletely understood. One promising area of investigation is the role of nonverbal behavior, particularly facial expressions. Facial expressions are reliable indicators of emotions and internal states, suggesting they may also correlate with an individual's level of risk aversion. This study investigates the relationship between risk aversion and facial expressions during decision-making. By examining participants' facial expressions as they make choices involving varying levels of risk, we aim to identify correlations between observed expressions and self-reported risk aversion levels. Furthermore, we will evaluate the potential of facial expressions as predictors of risk aversion. Our approach combines self-report measures with advanced facial expression analysis. Participants will engage in a series of decision-making tasks with varying risk levels, during which their facial expressions will be recorded and analyzed. The findings of this study will offer valuable insights into the role of nonverbal cues in decision-making processes. Potential applications span across fields such as finance—in developing more effective investment strategies—and psychology—in creating novel treatments for anxiety and related conditions that affect decision-making.

Introduction to the Problem

Decision-making is an integral aspect of human life, and individuals are constantly faced with choices that involve varying levels of risk. Risk aversion refers to an individual's tendency to avoid taking risks and to prefer more certain outcomes. Understanding the factors that influence risk aversion is therefore critical for many areas of study, including psychology, economics and finance.

Risk aversion has been widely studied, and various theories have been developed to explain this phenomenon. However, the exact mechanisms that influence risk aversion are still not fully understood. One possible factor that may impact risk aversion is nonverbal behaviour, specifically facial expressions. Facial expressions have been shown to be a reliable indicator of an individual's emotions and internal states and it is possible that these expressions may also be related to an individual’s level of risk aversion.

The purpose of this study is to investigate the relationship between risk aversion and facial expressions in decision-making. The study will examine the facial expressions of individuals while they are making decisions that involve varying levels of risk and will aim to identify any correlations between the observed facial expressions and the individual's self-reported risk aversion levels. In addition, the study will evaluate the accuracy of using facial expressions as a predictor of risk aversion. The study will employ a combination of self-report measures and facial expression analysis to assess risk aversion and decision-making. Participants will be asked to make a series of decisions that involve varying levels of risk and their facial expressions will be recorded during these decisions.

The findings of this study will provide valuable insights into the role of nonverbal cues in decision-making and the potential applications for using these cues in fields such as psychology, finance, and marketing. For example, in the field of finance, a better understanding of risk aversion and decision-making could be used to develop more effective investment strategies. In the field of psychology, a deeper understanding of the relationship between risk aversion and facial expressions could be used to develop new treatments for anxiety and other conditions that impact decision-making.

Machine learning is a type of artificial intelligence that allows computers to learn and make decisions based on data. When it comes to facial detection, machine learning algorithms are trained on large sets of facial images to help the computer understand what a face looks like and how to find it. This training process is what makes facial detection with machine learning so effective. The learning capacity of CNN has significantly improved over the years by exploiting depth and other structural modifications.

Convolutional Neural Networks (CNNs) are commonly used for detecting facial expressions in images. One important advantage of convolutional neural networks (CNNs) is their ability to automatically learn feature representation from raw pixels, eliminating the need for hand-designed procedures. (Wang et al. 2019) The key idea behind using a CNN for facial expression detection is to learn the features that are relevant for recognizing facial expressions. CNNs are some of the best learning algorithms to understand image features and have shown exemplary performance in image segmentation, classification, detection, and retrieval-related tasks. (Ciresan et al. 2012; Liu et al. 2019).

In a typical CNN architecture for facial expression detection, the input is an image of a face and the output is a label indicating the facial expression. The network consists of multiple layers, including convolutional layers, activation functions, pooling layers, and fully connected layers.

The convolutional layer is responsible for detecting local features in the image, such as edges, corners, and textures. The activation function is used to introduce non-linearity into the model, allowing it to capture complex patterns in the data. The pooling layer reduces the spatial resolution of the image, making the model more robust to small variations in the image.

The fully connected layer combines the features learned by the convolutional and pooling layers to classify facial expressions. During training, the network adjusts the weights of the layers to minimize the error between the predicted and actual labels. The output features that are mapped from the final convolution or pooling layer are typically flattened, i.e., these are transformed into a one-dimensional (1D) array or vector of numbers, and connected to many fully connected layers, which are also known as dense layers, in which each input is connected to the output by a learnable weight.

Once the model is trained, it can be used to predict the facial expression in new images by processing the input through the layers and using the learned weights to make a prediction. The result is a label indicating the facial expression present in the image.

In conclusion, this study will contribute to our understanding of the relationship between risk aversion and facial expressions in decisionmaking. By examining the facial expressions of individuals while they are making decisions that involve varying levels of risk, the study will provide insights into the potential role of nonverbal cues in this process. The findings of this study will have important implications for a variety of fields and will help to advance our understanding of human behavior and decision-making.

II. LITERATURE REVIEW

Studies exploring the relationship between risk aversion and facial expressions have been conducted in recent years, with a growing interest in understanding how nonverbal cues can influence decision-making under uncertainty. The following is an expanded summary of key research papers in this area:

The paper by Nguyen et al. [2014] introduces a new methodology, called "face reading" to economics, to investigate the correlation between facial expressions and the level of risk aversion exhibited by an individual. The study presents lotteries to participants and analyzes their facial expressions using psychological constructs such as anger, fear, happiness, and surprise. The key findings of the study are: Positive emotional valence is associated with greater risk tolerance: Individuals who exhibit more positive emotions, such as happiness, tend to exhibit higher levels of risk tolerance. Stronger emotions are associated with more risk-averse decision-making: The study finds that emotions such as fear, anger, happiness, and surprise are all significant correlates of riskaverse decisions. The individuals who exhibit stronger emotions tend to exhibit more risk-averse behavior.

Kahyaoglu et al. [2017] in their paper examined the impact of emotions on risk aversion using the game of DoND and the face reading method. This is one of the first studies to look into the connection between emotions and risk aversion in the context of DoND. The game allows for emotions to be revealed endogenously as the distribution of opened and unopened boxes changes over time. Despite this, the results show that emotions, including anger, happiness, sadness, neutrality, contempt, and surprise, may play a role in the level of risk aversion and the decisions made. These findings align with previous research (Nguyen & Noussair, 2014) that used the face-reading method to explore the relationship between emotions and risk aversion.

Evers et al. [2015] examined the influence of facial expressions on risk perception and decision-making. Participants who were exposed to happy or neutral facial expressions were more likely to perceive risks as lower and make riskier choices. The type of facial expression influenced risk perception and decision-making. Participants who were exposed to angry expressions perceived risks as higher, resulting in more risk-averse decisions.

The impact of facial expressions on risk perception and decisionmaking is not limited to individual differences in personality or risk attitude. This indicates that facial expressions can have a universal influence on risk perception and decision-making. The results highlight the potential of facial expressions as a tool for influencing risk perception and decision-making in various contexts, such as finance and marketing. We believe the reasoning for the unique structure of local facial parts(e.g. eyes, nose, mouth) is key to addressing face detection in an unconstrained environment. (Yang and Luo et al. 2015)

Kuppens et al. [2008] examined the impact of facial expressions on risk taking behavior. Facial feedback, or the physical act of changing one's facial expression, influences risk-taking behavior. Participants who smiled, scowled, or held a neutral expression while playing a game of dice showed different levels of risk-taking behavior. Smiling participants showed higher levels of risk-taking behavior compared to participants who scowled or showed a neutral expression. These findings suggest that the physical act of changing one's facial expression can impact our perception of risk and subsequent decision-making, highlighting the important role of facial feedback in emotional regulation.

The study by Campos-Vazquez et al. [2014] explores the role of emotions on risk aversion using a Prospect Theory experiment. The experiment measures how emotions affect the decision-making process of individuals in regard to risky investments. The study found that emotions play a significant role in shaping an individual's risk-aversion behavior. Negative emotions, such as fear and sadness, increase risk aversion while positive emotions, such as hope and happiness, lead to a decrease in risk aversion. The findings suggest that emotions are an important factor to consider when analyzing an individual's behavior toward risk-taking in investment decisions.

These studies provide evidence that nonverbal cues, such as facial expressions, play a role in risk perception and decision-making. For example, viewing happy faces can lead to more risky choices, while regulating one's emotional state or smiling can increase risk-taking behavior. The impact of other nonverbal cues, such as body language and vocal cues, also warrants further investigation. Understanding the influence of nonverbal cues on risk aversion has important implications for fields such as psychology, economics, and finance.

Data

The data-set was obtained from Kaggle and included a variety of facial expression images, which were then divided into two categories: strong emotions and positive emotions. The strong emotions folder contained expressions such as fear, anger, and sadness, while the positive emotions folder included expressions like happiness and hope. There are 13034 images in the strong emotions folder and 8989 images in the positive emotions folder. Based on the findings from previous research, it can be concluded that individuals who experience positive emotions tend to exhibit risk-seeking behavior, whereas those who experience strong emotions tend to exhibit risk averse behavior.

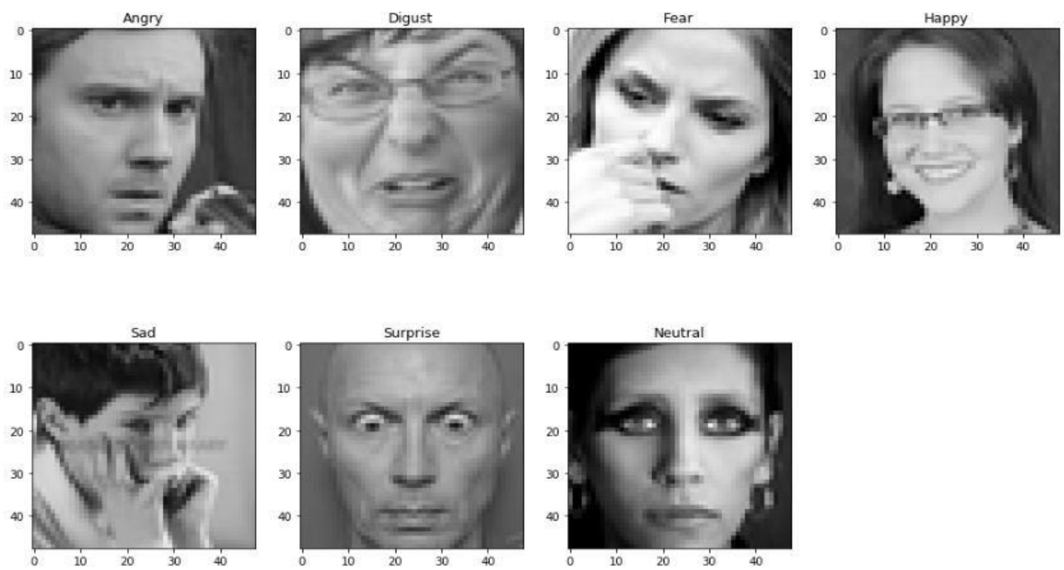

Fig. 3.1: Sample pictures from Data-set

These sample pictures from the data-set denote the variety of images available for testing and training our model. The range of expressions include: angry, disgust, fear, happy, sad, surprise and neutral. The model labels the input in these categories for further evaluation.

III. METHODOLOGY

CNN

Albawi et al. [2017] explained about CNNs in their paper. Convolutional Neural Networks are a type of neural network that is designed to process and analyze images. They are inspired by the structure and function of the human brain's visual cortex, which is responsible for detecting and recognizing visual patterns. The authors of a study demonstrated that numerous neurons in the visual cortex possess a limited local receptive field reacting only to visual stimuli within a confined area of the visual field. These receptive fields may overlap, collectively covering the entire visual field. Additionally, the study found that certain neurons react only to horizontal lines, while others respond to lines of different orientations. Some neurons have more extensive receptive fields and respond to intricate patterns that are a combination of simpler ones. Based on these observations, it was hypothesized that the higher-level neurons are constructed using the outputs of adjacent lowerlevel neurons, forming a robust architecture that can detect various complex patterns across the entire visual field. This architecture is formed by connections between each neuron and only a few neurons from the previous layer. Convolutional Neural Networks (CNN) are some of the most popular architectures used to solve computer vision problems because of their reduced number of parameters compared to fully connected models and intuitive structure, which allows the network to learn translation invariant features. In addition, convolution is a “cheap” operation that can be computed efficiently. (Li and Liu et al. 2014) We can see this complete phenomenon in the below image:

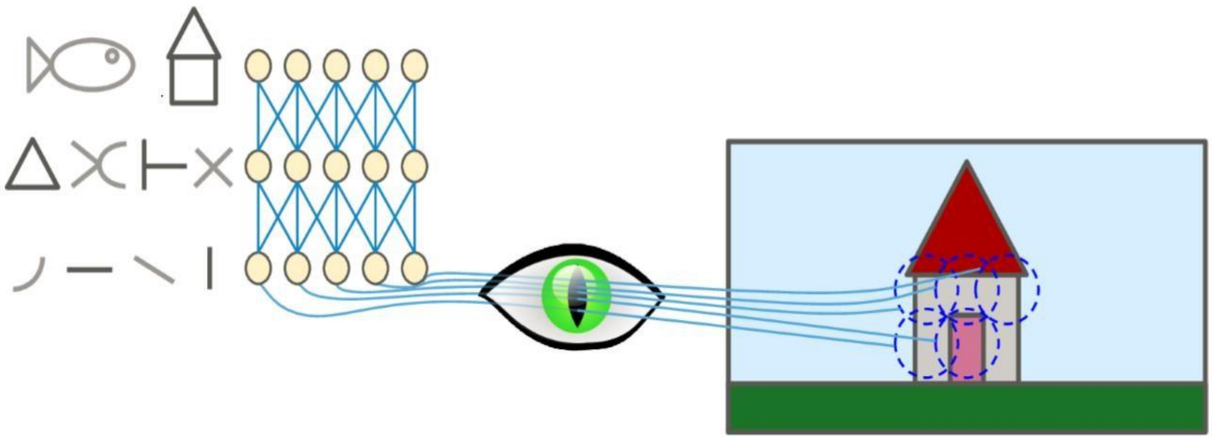

Fig. 4.1: Local Receptive Fields in the visual cortex

Solving segmentation and detection tasks requires applying the network to every patch contained in the image, which is prohibitively expensive when implemented in a naive, straightforward way. (Giusti et al. 2013) The main idea behind CNNs is to use a set of filters, or "convolutional kernels," to scan the input image in small overlapping patches. Each filter is responsible for detecting a specific feature or pattern in the image, such as edges, corners, or textures. The filter "slides" over the image and computes the dot product between its weights and the pixels of the current patch. This process creates a feature map that highlights where in the image the feature was detected. The output of each calibration stage is used to adjust the detection window position for input to the subsequent stage. (Li and Lin et al. 2015)

These feature maps are then passed through a nonlinear activation function, such as ReLU, which adds non-linearity to the network and helps to improve its ability to recognize complex patterns. The resulting feature maps are then further processed by pooling layers, which reduce the spatial dimensions of the feature maps while preserving their important features. It helps to lessen the number of parameters in a network and prevent over-fitting.

Indolia et al. [2018] also note that the use of Rectified Linear Unit (ReLU) has proved itself better than the sigmoid activation function because the calculation of the partial derivative of ReLU is easy and it does not allow gradients to disappear. Despite its depth, one of the key characteristics of modern deep learning systems is to use a non-saturated activation function (e.g. ReLU) to replace its saturated counterpart (e.g. sigmoid, tanh). (Xu, Wang and Chen et al. 2015)

Finally, the output of the last pooling layer is flattened and passed through many fully connected layers, that perform the final classification or regression task. The weights of the entire network are learned through back propagation, which adjusts the weights of each layer based on the error between the predicted output and the actual output.

In summary, CNNs use a set of filters to scan an input image, extract important features, and classify the image based on those features. This makes them particularly useful for image recognition, classification, and other visual tasks. The layers within the CNN are comprised of neurons organized into three dimensions, the spatial dimensionality of the input(height and width) and the depth. Also, CNNs require a large amount of training data to train the model. (Srinivas et al. 2016) When building a Convolutional Neural Network (CNN), it can be helpful to increase the number of layers gradually, one at a time, to see the effect on the model's accuracy and performance. The problem is that for higher layers, the invariances are extremely complex so are poorly captured by a simple quadratic approximation. (Zeiler and Fergus et al. 2013)

CNNs have a large number of parameters and they can produce different classification accuracies for the same tasks based on diverse parameters including input window size, filter size, number of layers, and number of neurons. (Sinha and Verma et al. 2018) Adding more layers to a CNN can increase its complexity, and may improve its ability to learn more complex features and patterns in the input data. However, adding too many layers can also make the model more difficult to train and prone to over-fitting, where it becomes too specialized to the training data and performs poorly on new, unseen data. By increasing the number of layers one at a time, the model builder can monitor the effect on the model's accuracy and performance, and determine the optimal number of layers for the specific task at hand. Convolutional Neural Network models have been developed with varying numbers of convolutional layers, with some models having as few as two layers and others having as many as six layers. There are currently three major techniques that successfully employ CNNs for image classification: training the CNN from scratch, using off-theshelf pre-trained CNN features, and conducting unsupervised
CNN pre training with supervised fine-tuning. (Shin and Roth et al. 2016)

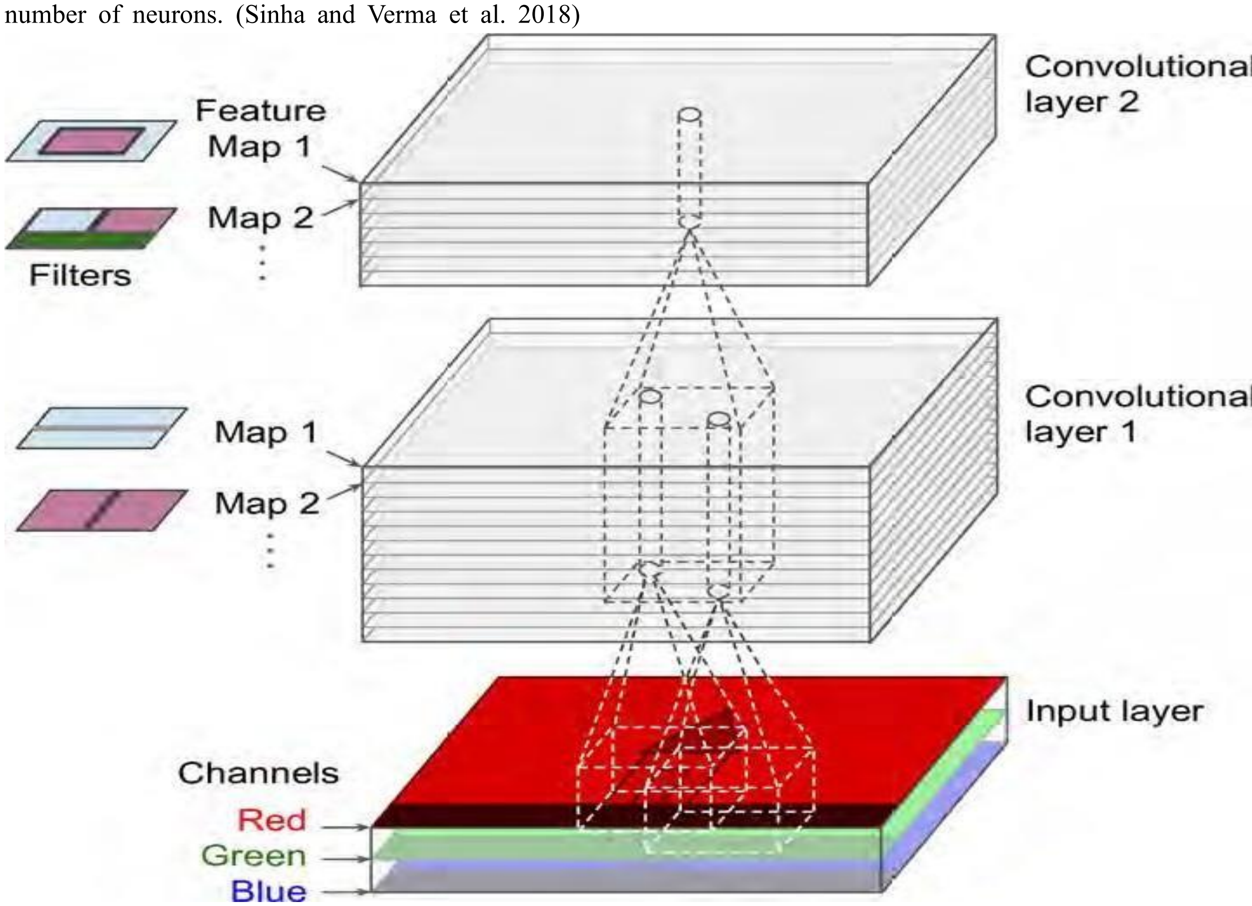

Fig. 4.2: CNN Architecture

To calculate the output of a neuron in a convolutional layer, a specific computation is performed:

$$z_{i,j,k} = b_k + \sum_{u=1}^{fh} \sum_{v=1}^{fw} \sum_{k'=1}^{fn'} x_{i',j',k'} \cdot w_{u,v,k',k} \text{ with } i' = u.sh + fh - 1 \text{ and } j' = v.sw + fw - 1$$

--
The output of the neuron situated at row i and column j in feature map k of the convolutional layer (layer l) is represented by $z_{i,j,k}$.

-- The height and width of the receptive field are represented by f_h and f_w, respectively. The vertical and horizontal strides are indicated by s_h and s_w. The previous layer (layer l - 1) had $f_{n'}$ feature maps.

-- The output of the neuron in feature map k' (or channel k' if the previous layer is the input layer) located at row i' and column j' in layer l - 1 is represented by $x_{i',j',k'}$.

-- The bias term for feature map k (in layer l) is represented by b_k .

-- $W_{u,v,k',k}$ represents the weight of the connection between a neuron in feature map k of layer l and its input, which is situated at row u and column v (relative to the neuron's receptive field), as well as feature map k'.

Transfer Learning

Hussain et al. [2019] explore transfer learning on image classification in their paper. Transfer learning is a machine learning technique that involves leveraging knowledge learned from one task or domain to improve performance on another task or domain. It enables a pre-trained model to be used as a starting point for training a new model, typically with less data or resources required than training from scratch. Transfer learning has become a popular approach in many areas of machine learning, including computer vision, natural language processing (NLP), and speech recognition, due to its ability to accelerate model training, improve model generalization, and achieve better performance. The idea behind transfer learning is that some features learned from one task can be useful for solving another related task. These features may be related to low-level visual features, higher-level abstract features, or even the complete architecture of a model. By reusing these learned features, we can save time and computational resources, and even improve the performance of the model on the new task.

Transfer learning can be understood in three main steps:

1. Pre-training: A model is trained on a large data set from a source domain, typically with a related or similar task. This pre-trained model learns general features or representations from the data, which capture useful patterns, structures, and representations.

2. Feature Extraction: The pre-trained model's learned features are then used as a starting point to extract features from the data in the target domain. These features are representations of the input data that capture higher-level patterns specific to the target domain. This step is often referred to as feature extraction or fine-tuning, as the pre-trained model's weights are updated with a smaller dataset from the target domain.

3. Fine-Tuning: The extracted features are then used as inputs to train a new model, often referred to as the target model or the fine-tuned model. This new model is trained using a smaller dataset from the target domain, which may not be large enough to train a high performing model from scratch. The fine-tuning process updates the pre trained model's weights based on the target domain data, allowing the model to adapt and specialize to the specific characteristics of the target domain.

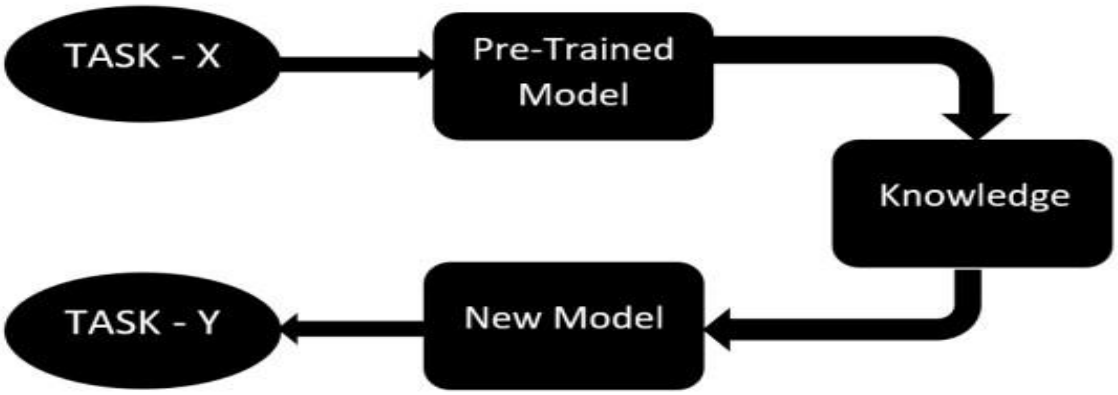

Fig. 4.3: Transfer Learning

Procedure

Initially, I divided the dataset into two parts: strong emotions (risk averse) and positive emotions (risk seeking). Then I divided the overall experimental setup into three significant steps.

In the First step, I built CNN (Convolutional Neural Networks) layer models ranging from 2 to 6 Convolutional and Max-Pooling layers. I observed that as I increased the number of layers (complexity of the model), the accuracy increased significantly, and there was no overfitting. So, I moved towards more complex architecture.

In the second step of my experiment, I extracted features from the image using various pre-trained models such as MobileNet, ResNet50, ResNet, VGG16, VGG19, MobileNetV2, DenseNet169, and DenseNet121. I then passed these features through a global average pooling layer and added a classifier. I employed several machine-learning techniques for the final classification, including XGBoost, Random Forest, SVM, kNN, and AdaBoost.

In the final step of my experiment, I trained pre-built architectures (mentioned above) from scratch, providing maximum task accuracy.

Metrics Used

1. **Accuracy:** It is the ratio of the number of correct predictions made by the model to the total number of predictions made. In other words, it measures the degree of closeness between the predicted and actual values.

2. Precision: It is calculated by dividing the number of true positives by the sum of true positives and false positives. In other words, precision measures the model's ability to accurately identify positive instances without including false positives.

$$\text{Precision} = \frac{TP}{TP + FP}$$

3. **Recall:** It is calculated by dividing the number of true positives by the sum of true positives and false negatives. In other words, recall measures the model's ability to correctly identify positive instances without missing any.

$$\textbf{\textit{Recall}} = \frac{TP}{TP + FN}$$

4. F1 Score: It is a way to balance the trade-off between precision and recall, as these two metrics can often have an inverse relationship. The F1 score is calculated as the harmonic mean of precision and recall, and ranges from 0 to 1, with a higher score indicating better performance. A high F1 score indicates that the model has a high precision and a high recall as well.

$$F1 = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} = 2 \times \frac{Precision \ x \ Recall}{Precision \ + Recall} = \frac{TP}{TP + \frac{FN + FP}{2}}$$

5. ROC Score: The ROC score is the area under the ROC curve, which is a graphical representation of the TPR and FPR values for different classification thresholds. A higher ROC score indicates better model performance, with a score of 0.5 indicating a random model and a score of 1.0 indicates a perfect classifier.

IV. RESULTS

Step 1 (CNN Layers)

The results in Table 5.1 shows that 2 Convolutional Layers model gives us an Accuracy of 0.829 and F1 score of 0.8271**.** The results are good for the preliminary model, but could be improved significantly by increasing the complexity of the model. The results for 3 Convolutional Layers model gives us an Accuracy of 0.8799 and F1 score of 0.8791.

The results improved significantly as compared to the previous model. The results for 4 Convolutional Layers model gives us an Accuracy of **0.8756** and F1 score of 0.8741. The results are almost similar to the previous 3 Convolutional Layers model. The results for 5 Convolutional Layers model gives us an Accuracy of 0.889 and F1 score of 0.8891. The results improved again to quite a good level. Finally, the results for 6 Convolutional Layers give us an Accuracy of 0.8941 and F1 score of 0.8944. The results for this model are very good and we improved the overall accuracy of the model by increasing the complexity of the model.

Table 5.1: Results of Convolutional Layers

S.No.	Model	Accuracy	Precision	Recall	F1 Score	ROC Score
1	2 Conv Layers	0.829	0.8284	0.829	0.8271	0.8123
2	3 Conv Layers	0.8799	0.8796	0.8799	0.8791	0.8692
3	4 Conv Layers	0.8756	0.8766	0.8756	0.8741	0.8603
4	5 Conv Layers	0.889	0.8893	0.889	0.8891	0.8855
5	6 Conv Layers	0.8941	0.8949	0.8941	0.8944	0.892

Step 2 (Feature Extraction)

Tables 5.2 - 5.9 show the result for the feature extraction or transfer learning technique. I extracted features from the image using various pretrained models and then passed these features through a global average pooling layer and added a classifier. The results were not promising in this step and the model behaves poorly on the test data-set. The best result which was obtained in this step was when features were extracted from DenseNet121 and the classifier was XGBoost. The combination of DenseNet121 and XGBoost gave an accuracy of **0.726** and F1 score of **0.7262.**

Table 5.2: Results of Feature Extraction from DenseNet121

DenseNet121						
S.No.	Model	Accuracy	Precision	Recall	F1 Score	ROC Score
1	SVM	0.63	0.63	0.63	0.63	0.6299
2	Random Forest	0.7167	0.7179	0.7167	0.7165	0.7171
3	AdaBoost	0.6973	0.6973	0.6973	0.6973	0.6973
4	KNN	0.624	0.6246	0.624	0.6228	0.623
5	XGBoost	0.726	0.7264	0.726	0.726	0.7262

Table 5.3: Results of Feature Extraction from DenseNet169

DenseNet169						
S.No.	Model	Accuracy	Precision	Recall	F1 Score	ROC Score
1	SVM	0.636	0.6392	0.636	0.6344	0.6365
2	Random Forest	0.6907	0.6907	0.6907	0.6907	0.6907
3	AdaBoost	0.694	0.6941	0.694	0.694	0.6941
4	KNN	0.616	0.6217	0.616	0.6122	0.6168
5	XGBoost	0.7167	0.7167	0.7167	0.7167	0.7167

Table 5.4: Results of Feature Extraction from MobileNet

MobileNet						
S.No.	Model	Accuracy	Precision	Recall	F1 Score	ROC Score
1	SVM	0.512	0.542	0.512	0.4446	0.5209
2	Random Forest	0.532	0.5328	0.532	0.5319	0.5324
3	AdaBoost	0.5407	0.5417	0.5407	0.5405	0.5413
4	KNN	0.55	0.5503	0.55	0.5501	0.55
5	XGBoost	0.5347	0.5358	0.5347	0.5344	0.5353

Table 5.5: Results of Feature Extraction from MobileNetV2

MobileNetV2						
S.No.	Model	Accuracy	Precision	Recall	F1 Score	ROC Score
1	SVM	0.5727	0.5748	0.5727	0.5723	0.5738
2	Random Forest	0.668	0.6698	0.668	0.6679	0.6689
3	AdaBoost	0.6613	0.6615	0.6613	0.6614	0.6612
4	KNN	0.58	0.5794	0.58	0.5788	0.5783
5	XGBoost	0.694	0.6947	0.694	0.6941	0.6943

Table 5.6: Results of Feature Extraction from ResNet

ResNet						
S.No.	Model	Accuracy	Precision	Recall	F1 Score	ROC Score
1	SVM	0.6733	0.6735	0.6733	0.6732	0.6733
2	Random Forest	0.686	0.6862	0.686	0.6859	0.686
3	AdaBoost	0.6627	0.6627	0.6627	0.6627	0.6627
4	KNN	0.5993	0.6001	0.5993	0.5986	0.5993
5	XGBoost	0.7007	0.7007	0.7007	0.7007	0.7007

Table 5.7: Results of Feature Extraction from ResNet50

		ResNet50				
S.No.	**Model**	**Accuracy**	**Precision**	**Recall**	**F1 Score**	**ROC Score**
1	SVM	0.6493	0.6493	0.6493	0.6493	0.6493
2	Random Forest	0.6933	0.6935	0.6933	0.6932	0.6933
3	AdaBoost	0.702	0.702	0.702	0.702	0.702

4	KNN	0.662	0.6674	0.662	0.6594	0.6622
5	XGBoost	0.716	0.7167	0.716	0.7157	0.7159

Table 5.8: Results of Feature Extraction from VGG16

VGG16						
S.No.	Model	Accuracy	Precision	Recall	F1 Score	ROC Score
1	SVM	0.5447	0.5548	0.5447	0.5142	0.5413
2	Random Forest	0.592	0.5936	0.592	0.5912	0.5926
3	AdaBoost	0.5673	0.5676	0.5673	0.5673	0.5675
4	KNN	0.5667	0.5667	0.5667	0.5667	0.5666
5	XGBoost	0.576	0.5763	0.576	0.576	0.5761

Table 5.9: Results of Feature Extraction from VGG19

VGG19						
S.No.	Model	Accuracy	Precision	Recall	F1 Score	ROC Score
1	SVM	0.4833	0.2336	0.4833	0.315	0.5
2	Random Forest	0.5947	0.5942	0.5947	0.5939	0.5932
3	AdaBoost	0.5733	0.5738	0.5733	0.5734	0.5733
4	KNN	0.5793	0.5814	0.5793	0.5791	0.5804
5	XGBoost	0.594	0.5942	0.594	0.5941	0.5938

Step 3 (Train Pre-Built Networks)

As in step 2, I didn't get promising results by using the pre-trained weights, so I decided to train these pre-built architectures from scratch using my data set. It turned out to be a good strategy as the results improved significantly as compared to the previous 2 steps. We can see from Table 5.10 that almost all the architectures performed very well on the task, but Densenet121 and DenseNet169 were the top 2 best models which obtained an accuracy of 0.9173 and 0.904 respectively. The F1 scores obtained by these models were 0.9173 and 0.904 respectively.

Table 5.10: Results of Training Pre-Built Architecture

S.No.	Model	Accuracy	Precision	Recall	F1 Score	ROC Score
1	MobileNet	0.836	0.8361	0.836	0.836	0.836
2	ResNet50	0.888	0.888	0.888	0.888	0.888
3	ResNet	0.8967	0.8967	0.8967	0.8967	0.8967
4	VGG19	0.8867	0.8872	0.8867	0.8866	0.8867
5	VGG16	0.4987	0.2487	0.4987	0.3319	0.5
6	MobileNetV2	0.8613	0.8617	0.8613	0.8613	0.8614
7	DenseNet169	0.904	0.904	0.904	0.904	0.904
8	DenseNet121	0.9173	0.9176	0.9173	0.9173	0.9174

Images Classified by the Algorithm

Strong Emotions

Fig. 5.1: Strong Emotions Image 1

Fig. 5.2: Strong Emotions Image 2

Positive Emotions

Fig. 5.3: Positive Emotions Image 1

Fig. 5.4: Positive Emotions Image 2

V. CONCLUSION

Applications

Financial Services: This project can be utilized in the financial industry to assess the risk aversion of investors, traders, or clients. By analyzing their facial expressions during risk-related tasks or decision-making scenarios, financial institutions can gain insights into the risk preferences of their customers. This information can be used to tailor investment strategies, financial products, or advisory services to better suit the risk tolerance of individual clients.

Behavioral Economics Research: Researchers in the field of behavioral economics can utilize this project to study risk aversion and decisionmaking behavior. By analyzing facial expressions during controlled experiments, researchers can collect valuable data to understand how emotions and risk aversion influence economic choices. This can provide insights into human behavior, and decision-making processes, and help develop models or theories related to risk-taking behavior.

Human Resources and Hiring: This project can be employed in the field of human resources and talent acquisition to assess job candidates' risk aversion. During interviews or assessment exercises, facial expression analysis can provide additional information on candidates' emotional responses to risk related scenarios, which can be useful in evaluating their suitability for roles that require risk-taking, such as leadership positions or sales roles.

Risk Management: This project can be utilized in the field of risk management in various industries, such as insurance, healthcare, or aviation, to assess employees' risk aversion and their ability to handle risky situations. For example, in the aviation industry, pilots' facial expressions during simulated emergency situations can provide insights into their decision-making processes and risk aversion levels, which can be useful in training and evaluating their performance in critical situations.

Psychology and Mental Health: This project can be used in psychological research and clinical settings to assess individuals' risk aversion in relation to mental health conditions, such as anxiety disorders or addiction. By analyzing facial expressions during risk related tasks, clinicians and researchers can gain insights into the emotional responses of individuals with mental health conditions, which can help in developing tailored treatment plans or interventions.

Gaming and Virtual Reality: This project can be applied in the gaming and virtual reality industries to enhance the user experience by incorporating realtime facial expression analysis to determine players' risk aversion levels. This information can be used to dynamically adjust game difficulty, challenges, or rewards based on players' emotional responses, providing a more personalized and engaging gaming experience.

Marketing and Advertising: This project can be utilized in marketing and advertising research to assess consumers' emotional responses to risky decisions or advertisements. By analyzing facial expressions, marketers and advertisers can gain insights into consumers' emotional engagement, preference, or aversion towards different marketing stimuli, which can be useful in designing more effective marketing campaigns or products.

Societal Impact

Improved Financial Decision-Making**:** By analyzing facial expressions during risk-related tasks, financial institutions can gain insights into the risk preferences of their customers, leading to more tailored investment strategies and financial products that align with the risk tolerance of individual clients. This can potentially lead to improved financial decision-making, resulting in better financial outcomes for individuals and society.

Safer Work Environments: The use of facial expression analysis in risk management industries, such as aviation or healthcare, can provide insights into employees' risk aversion levels, allowing for more targeted training and evaluation of their performance in critical situations. This can contribute to safer work environments and better outcomes for employees and society.

Better Mental Health Treatment: Facial expression analysis can be used in psychological research and clinical settings to assess individuals' risk aversion in relation to mental health conditions, leading to tailored treatment plans and interventions that address emotional responses to risk-related scenarios. This can potentially lead to better mental health outcomes and improved quality of life for individuals and society.

More Effective Marketing: The use of facial expression analysis in marketing and advertising research can provide insights into consumers' emotional responses and preferences towards different marketing stimuli, leading to more effective marketing campaigns and products that align with consumers' needs and preferences.

Ethical Considerations: The use of facial expression analysis raises ethical considerations related to privacy and data protection. The development and implementation of ethical frameworks and guidelines can ensure that the use of this technology is conducted with proper consent and adherence to relevant laws and regulations, contributing to a more responsible and trustworthy use of facial expression analysis.

Overall, the societal impact of the project on determining risk aversion using facial expressions can lead to better outcomes for individuals and society in various domains, while also considering the ethical implications of its use.

VI. DISCUSSION

In conclusion, the project on determining risk aversion using facial expressions has the potential to make significant contributions to various fields, including finance, research, human resources, risk management, psychology, gaming, and marketing. By analyzing facial expressions during risk related tasks or decision-making scenarios, valuable insights can be gained into individuals' emotional responses and risk preferences, leading to more tailored strategies, products, or interventions. I used three different strategies for doing this task - Building Convolutional Layers, Feature Extraction, and Training Pre-Built architectures. The best result came from training Prebuilt architectures where DenseNet121 gave an accuracy of 0.9173 on the test data set. However, ethical considerations must be taken into account, and the use of facial expression analysis must be conducted with proper consent and adherence to relevant laws and regulations.

Future research in the field can focus on improving the accuracy and reliability of facial expression analysis, exploring the validity and generalizability of findings, and investigating the practical applications of the technology in real world settings. The societal impacts of the project can include improved financial decision-making, safer work environments, better mental health treatment, more effective marketing, and the development of ethical frameworks and guidelines.

Overall, the project on determining risk aversion using facial expressions has the potential to contribute to a better understanding of human behavior and decision-making processes, leading to more tailored and effective interventions and outcomes for individuals and society.

REFERENCES

[1] Albawi, S. and Mohammed, T.A., [2017]. In Understanding of a Convolutional Neural Network. U 2017 International Conference on Engineering and Technology (ICET), stranice pp. 1-6.

[2] Campos-Vazquez, R.M. and Cuilty, E., [2014]. The role of emotions on risk aversion: a prospect theory experiment. Journal of Behavioral and Experimental Economics, 50, pp. 1-9.

[3] Evers, C., Alshehri, O., & DeCremer, D., [2015]. The Impact of Facial Expressions on Risk Perception and Risky Choices. Journal of Behavioral Decision Making, 28(2), pp. 123-130.

[4] Hussain, M., Bird, J.J. and Faria, D.R., [2019]. A study on cnn transfer learning for image classification. In Advances in Computational Intelligence Systems: Contributions Presented at the 18th UK Workshop on Computational Intelligence, September 5-7, 2018, Nottingham, UK pp. 191-202.

[5] Kahyaoglu, M.B. and Ican, O., [2017]. Risk aversion and emotions in DoND. International Journal of Economics and Finance, 9(1), pp. 32-46.

[6] Kuppens, P., Tuerlinckx, F., & De Houwer, J. [2008]. Facial Feedback Modulates Risk Taking in a Game of Dice. Emotion, 8(3), pp. 361-365.

[7] Nguyen, Y. and Noussair, C.N., [2014]. Risk aversion and emotions. Pacific economic review, 19(3), pp. 296- 312.

[8] O'Shea, K.T. and Nash, R., [2015]. An Introduction to Convolutional Neural Networks. Aberystwyth University.

[9] Yamashita, R., Nishio, M., Do, R.K.G. et al. Convolutional Neural Networks: an overview and application in radiology. Insights Imaging 9, pp. 611-629 [2018].

[10] Indolia, S., Goswami, A.K., Mishra, S.P., Asopa, P., [2018]. Conceptual Understanding of Convolutional Neural Network: A Deep Learning Approach, Procedia Computer Science, 132, pp. 679-688.

[11] Carpenter, S.M. and Niedenthal, P.M., [2020]. Disrupting Facial Action Increases Risk Taking. Emotion, Sep 20(6), pp. 1084-1092.

[12] Liu X, Deng Z, Yang Y (2019) Recent Progress in Semantic Image Segmentation. Artif Intell Rev 52:1089- 1106. doi: 10.1007/s10462-0189641-3.

[13] Ciresan D, Giusti A, Gambardella LM, Schmidhuber J (2012) Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images. In: Advances in Neural Information Processing Systems. Pp. 2843-2851.

[14] Khan, A., Sohail, A., Zahoora, U. et al. A survey of the recent architectures of deep convolutional

neural networks. Artif Intell Rev 53, 5455- 5516 (2020).

[15] Zeiler MD, Fergus R (2013) Visualizing and Understanding Convolutional Networks. arXiv Prepr arXiv13112901v3 30:225-231. doi: 10.1111/j.1475-4932. 1954.tb03086.x

[16] Yang S, Luo P, Loy C-C, Tang X (2015) From facial parts responses to face detection: A deep learning approach. In: Proceedings of the IEEE International Conference on Computer Vision. Pp.3676-3684.

[17] Xu B, Wang N, Chen T, Li M (2015a) Empirical Evaluation of Rectified Activations in Convolutional Networks. J Foot Ankle Res 1: O22. doi: 10.1186/1757-1146-1-S1-O22.

[18] Wang Y., Wang L., Wang H., Li P (2019) End-to-end Image Super Resolution via Deep and Shallow Convolutional Networks. IEEE Access 7:31959-31970. doi: 10.1109/ACCESS.2019.2903582.

[19] Srinivas S, Sarvadevabhatla RK, Mopuri KR, et al (2016) A Taxonomy of Deep Convolutional Neural Nets for Computer Vision. Front Robot AI 2:1- 13. doi: 10.3389/frobt.2015.00036.

[20] Sinha T, Verma B, Haidar A (2018) Optimization of Convolutional Neural Network Parameters for Image Classification 2017 IEEE Symp Ser Comput Intell SSCI 2017 - Proc 2018 - Janua: 1-7. doi: 10.1109/SSCI.2017.8285338.

[21] Shin H-CC, Roth HR, Gao M, et al. (2016) Deep Convoutional Neural Networks for Computer Aided Detection : CNN Architectures, Data-set Characteristics and Transfer Learning. IEEE Trans Med Imaging 35:1285-1298. doi: 10.1109/TMI.2016.2528162.

[22] Li S, Liu Z-Q, Chan AB (2014) Heterogeneous multi-task Learning for Human Pose Estimation with Deep Convolutional Neural Network. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops. IEEE, pp 488-495.

[23] Li H, Lin Z, Shen X, et al (2015) A convolutional neural network cascade for face detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Pp 5325-5334.

[24] Giusti A, Ciresan DC, Masci J, et al (2013) Fast image scanning with deep max-pooling convolutional neural networks. In: 2013 IEEE International Conference on Image Processing. IEEE, pp 4034-4038.