Report
3D expression Recognition

Introduction

The performance of three classifiers on facial expression data from the BU4DFE_BND_V1.1 dataset is analyzed in this paper. The classifiers are Random Forest (RF), Support Vector Machines (SVM), and Decision Tree (TREE). The dataset includes the locations of various facial landmarks that can be used as features for facial expression detection algorithms. It consists of 101 individuals showing facial expressions in a variety of orientations and lighting conditions. Using five categories—Original, Translated, Rotated X, Rotated Y, and Rotated Z—the material is categorized.

Questions

Classification results for each experiment

Classifier/ Data Type	Original Data		
Metric	**Accuracy**	**Precision**	**Recall**
Random Forest	0.34347892095	0.338068613362	0.343437892095
SVM	0.509443747386	0.503679830628	0.509443747386
Decision Tree	0.269410288582	0.266184261968	0.269410288582

Classifier/ Data Type	Translated		
Metric	Accuracy	Precision	Recall
Random Forest	0.332743918561	0.3226516767737	0.332743918561
SVM	0.500132205182	0.493012006566	0.500132205182
Decision Tree	0.264261634050	0.255305233233	0.264261634050

Classifier/ Data Type	Rotated X		
Metric	Accuracy	Precision	Recall
Random Forest	0.342082810539	0.3357453885081	0.342082810539
SVM	0.508021748220	0.502207803289	0.508021748220
Decision Tree	0.277273107486	0.278411789457	0.277273107486

Classifier/ Data Type	Rotated Y		
Metric	Accuracy	Precision	Recall
Random Forest	0.343939774153	0.3382666930021	0.343939774153
SVM	0.508205771643	0.501968119860	0.508205771643
Decision Tree	0.276721037222	0.272444404532	0.276721037222

Classifier/ Data Type	Rotated Z		
Metric	Accuracy	Precision	Recall
Random Forest	0.345177749895	0.336243025719	0.345177749895
SVM	0.506783772480	0.500822323317	0.506783772480
Decision Tree	0.265696361355	0.260382903901	0.265696361355

Which of the classifiers worked the best for each data type (original, translated, rotated)? Why? If multiple had the same, then why did this happen? Note: this question is left intentionally vague. What does work best mean? Which metric is best? Think about what we've talked about in class in regard to expressions and emotion.

When evaluating the efficiency of classifiers in the BU4DFE_BND_V1.1 dataset, "work best" can be interpreted as the classifier that obtained the greatest accuracy. However, it's important to consider other metrics such as precision, recall, and the confusion matrix to gain a more comprehensive understanding of the classifier's performance. After examining the dataset, I discovered that the classifiers' performance varied according to the type of data. SVM, out of all the classifiers, had the best accuracy for the original data. SVM and RF performed better than the Tree classifier for the translated data, with accuracy that was nearly identical. SVM and RF classifier both had the best accuracy for the rotated data type for RotatedX, RotatedY, and RotatedZ, respectively. The distribution of features in the various data types may be one reason for the variance in classifier performance. Some classifiers might be more adept at spotting specific characteristics or more sensitive to changes in the data. Additionally, because people express their emotions and facial expressions differently, some kinds of data may be inherently harder to categorize. It is crucial to take into account the particular emotional expressions contained in the dataset when assessing the classifiers' performance based on accuracy, precision, recall, and confusion matrix. It may be more challenging to recognize some feelings solely from facial landmarks, which could reduce the classification accuracy. For instance, it may be harder to categorize ambiguous or subtle statements than it is to categorize strong ones.

For the top classifier, for each data type, describe the misclassification for each classifier. What was misclassified as what (e.g., sad looks like happy)? Based on your intuition about how expressions look, do the misclassifications make sense (i.e., do you think the expressions are similar – not based on any specific example)?

The SVM classifier's accuracy was highest for the original data type. However, there were still a lot of misclassifications, with over 20 happy expressions being misclassified as surprise and roughly 13 surprise faces being misclassified as happy. In addition, five cases of disgust faces were misclassified as sad, whereas six occurrences of fear faces were misclassified as disgust. These misclassifications are understandable given that expressions of surprise and happiness, as well as those of fear and disgust, can have a similar appearance. Both the RF and SVM classifiers had an accuracy for the Translated data type. Most of the misclassifications occurred between shocked, indifferent, and joyful expressions. Approximately 10 happy faces were mistaken for being afraid, while 8 fear faces were mistaken for being happy. With around 13 happy faces misclassified as surprise and 9 surprise faces misclassified as happy, the SVM classifier showed the highest level of misunderstanding between "happy" and "surprise." Given that expressions of surprise and happiness, as well as those of fear and disgust, can have a similar appearance, these misclassifications are understandable. While SVM had the maximum accuracy for Rotated Y, RF classifier had the most accuracy for Rotated X, Z, and Y. The RF classifier, which is comparable to the original data type, confused "happy" and "surprise" the most for Rotated X and Z. The SVM classifier had the most difficulty distinguishing between "fear" and "happy" in Rotated Y, misclassifying happy faces as neutral in 8 instances and neutral faces as happy in 7 instances. These misclassifications are understandable given that rotated photos might include adjustments to the face's angle and orientation, which can influence how facial expressions seem. Overall, the misclassification findings revealed that classifiers consistently confused the emotions "happy" and "surprise" for all data categories, which makes sense given that both emotions can sometimes express themselves similarly. The small variations in facial features that distinguish various expressions can be used to explain the classifiers' confusion between "disgust" and other emotions. The misclassifications are understandable given our perception about how expressions appear. For instance, a smile and lifted brows are features of both pleased and surprised facial expressions, which can be extremely similar. A wrinkled brow and a closed mouth are also features of both the fear and disgusting facial expressions. The misclassifications may, however, also be a result of the classifiers' limitations and any small modifications in facial expressions that they may have overlooked. In conclusion, the misclassification findings show the limitations of the classifiers and emphasize the significance of taking into account the subtleties of facial expressions while developing emotion detection systems. This investigation also demonstrates the difficulties in recognizing facial expressions, particularly when dealing with noise, distortions, and changes in angle and orientation.

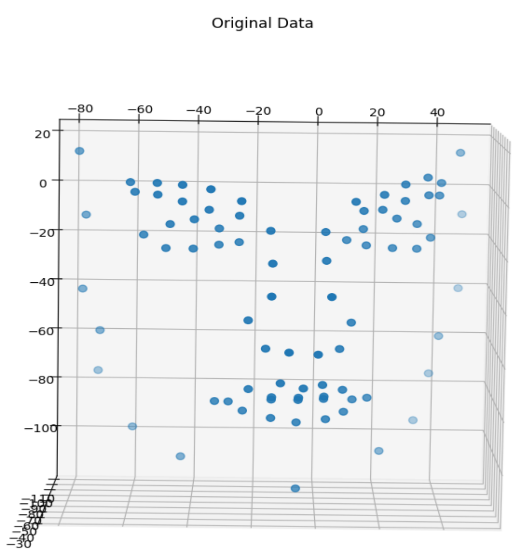
Why do you think you got the results that you got for each of the different data types/classifiers (i.e., why are they different, or why are they the same)? For example, if SVM and RF have different results, why are they different? If they are the same – why are they the same?

The complexity of the data, the distribution of features, the classifier's classification method, the quantity and quality of the dataset, and the selection of hyperparameters are a few of the variables that can be used to explain the outcomes of different data types and classifiers. The intricacy of the data is one of the elements that might have impacted the performance of the classifiers. Some data types might be simpler to categorize because of their more distinct feature distributions, but others might be more challenging because of the feature distribution overlap amongst emotions. For instance, the original data type contains unaltered face expressions while the translated and rotated data types represent altered variations of the original data. The changes in the data transformation that was used can be blamed for the performance inconsistencies. For instance, the distribution of features in the translated data type may differ from that in the original data type, which could impact the classification accuracy. The classification method that the classifier employed could also have had an impact on how well they performed. While RF is a non-linear classifier that constructs decision trees and combines them to provide the final prediction, SVM is a linear classifier that seeks the optimum hyperplane to divide the classes. The performance gap between the two classifiers may be explained by the different classification methodology. Due to the way it divides the data, TREE, another decision tree-based classifier, may perform differently from SVM and RF. The procedure of feature selection can have a significant impact on the classifier's performance. distinct methods of feature selection may produce distinct feature subsets, which may impact the classifier's performance. Performance of the classifier can also be impacted by the selection of hyperparameters, such as the number of trees in RF or the kernel function and regularization parameter in SVM. The performance of the classifiers can be enhanced via hyperparameter tuning, which can assist in determining the hyperparameters' ideal values. The performance of the classifiers can also be impacted by the dataset's size and quality. A dataset with significant noise or outliers may hinder the classifier's capacity to generalize, whereas a short dataset might not offer enough instances for it to learn from. With only 101 samples, the BU4DFE_BND_V1.1 dataset in this scenario may not have enough data to successfully train the classifiers. In conclusion, the performance differences between the various data types and classifiers can be attributed to a variety of factors, including variations in the data transformations used, variations in classifier classification methods, the selection of features, the selection of hyperparameters, and the quantity and quality of the dataset. Understanding these elements can aid classifier performance and reveal information about the data's nature. Future research can concentrate on delving deeper into these elements and creating fresh methods for enhancing facial expression recognition systems' functionality.

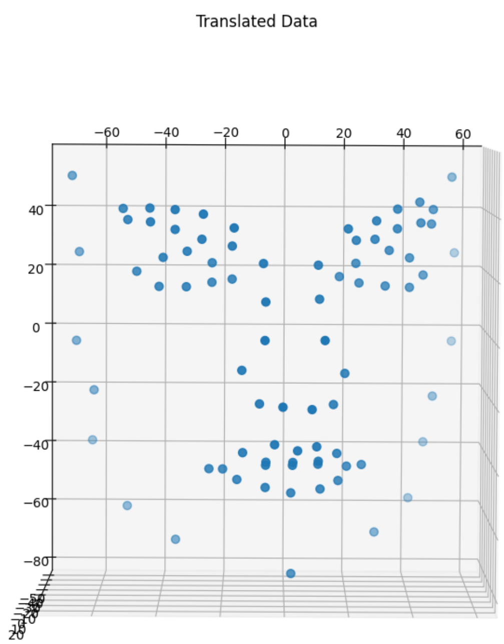
Experiment Results

3D scatter plot

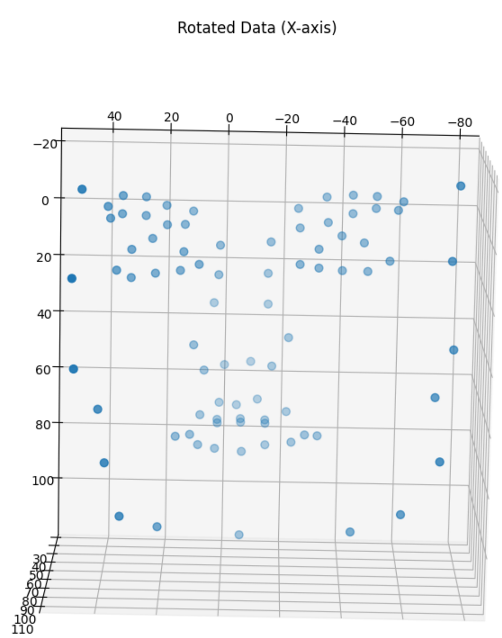
Original Data



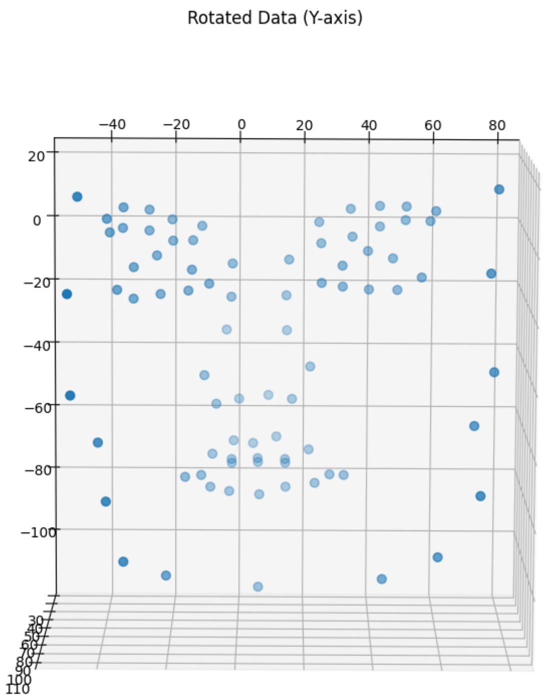
Translated Data



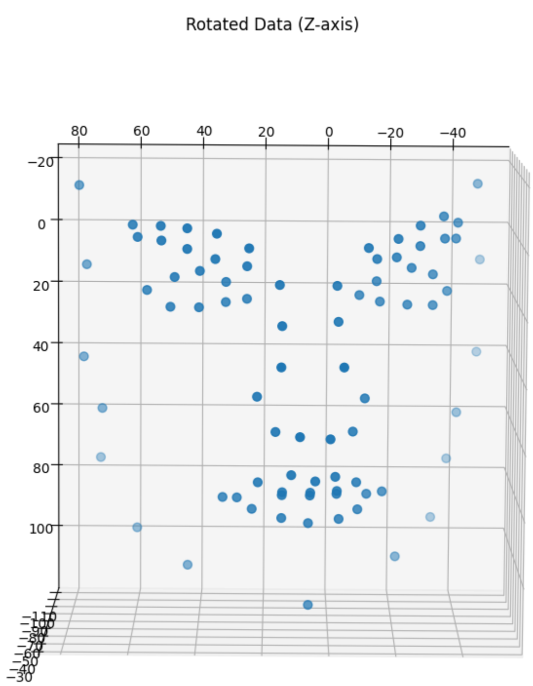
Rotated X Data



Rotated Y Data



Rotated Z Data



The confusion matrices

"Although I am not completely certain if my interpretation, after evaluating the accuracy of various data types, I found that SVM performed the best. Therefore, I have chosen SVM as the sole example to present the confusion metrics results"

Original Data

	Angry	Disgust	Fear	Happy	Sad	Surprised
Angry	4960	1268	544	571	2055	530
Disgust	1582	4238	1743	1030	675	797
Fear	859	1794	2934	1810	960	1577
Happy	743	502	1287	6292	465	570
Sad	2032	489	837	498	5733	449
Surprised	750	652	1040	777	437	6295

Translated Data

	Angry	Disgust	Fear	Happy	Sad	Surprised
Angry	4781	1169	585	749	2191	649
Disgust	1536	4178	1636	1211	800	810
Fear	946	1719	2797	2035	1054	1493
Happy	710	541	1300	6408	546	468
Sad	1910	530	753	531	5906	512
Surprised	665	697	1073	879	550	6194

Rotated X Data

	Angry	Disgust	Fear	Happy	Sad	Surprised
Angry	5032	1195	572	559	2034	536
Disgust	1552	4251	1789	999	685	789
Fear	920	1826	2887	1827	960	1514
Happy	715	587	1263	6304	470	520
Sad	1946	557	864	528	5665	478
Surprised	755	713	979	783	493	6228

Rotated Y Data

	Angry	Disgust	Fear	Happy	Sad	Surprised
Angry	5061	1206	541	527	2056	537
Disgust	1577	4253	1749	1064	659	763
Fear	975	1851	2793	1851	957	1507
Happy	736	555	1227	6295	493	553
Sad	2002	535	871	509	5650	471
Surprised	761	638	1020	762	444	6326

Rotated Z Data

	Angry	Disgust	Fear	Happy	Sad	Surprised
Angry	5046	1217	555	517	2061	532
Disgust	1583	4238	1721	1016	721	786
Fear	968	1827	2811	1788	1015	1525
Happy	722	526	1310	6225	489	587
Sad	2058	465	843	498	5703	471
Surprised	784	653	995	774	475	6270