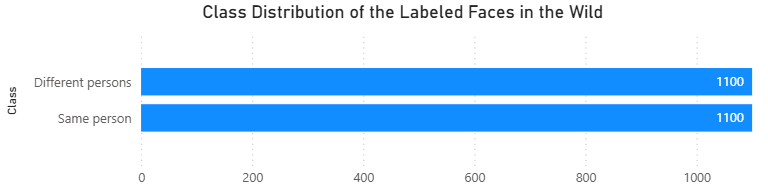
**Appendix: Figure Descriptions for Final Report**

This appendix provides detailed documentation for each figure included in the projects.

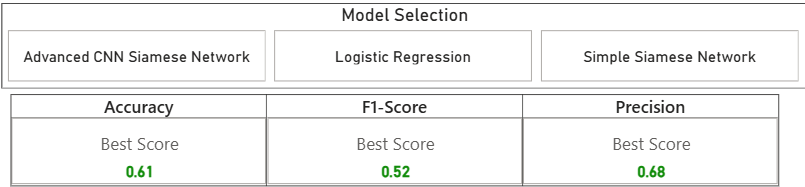
final dashboard, ready to be incorporated into a formal report.

**Figure 1: Class Distribution of the Labeled Faces in the Wild (LFW) Dataset.**

****

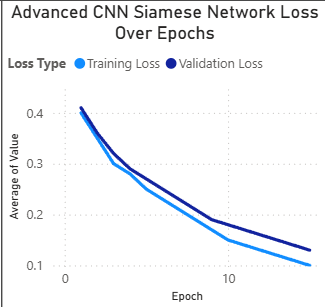
* **Explanation:** This bar chart visualizes the number of pairs classified as same person versus different persons in the LFW dataset.
* **Significance:** It confirms that the dataset has a balanced class distribution, which is ideal for binary classification tasks. This means the model does not have a natural bias towards one class over the other, making it a fair and reliable benchmark for model performance.
* **Insights:** The equal number of samples for each class means that a model that simply guesses different people for every pair would only achieve 50% accuracy. Any performance above this baseline is a result of the model's learning.

**Figure 2: Performance Comparison: Baseline (Regression) vs. Advanced Models.**

****

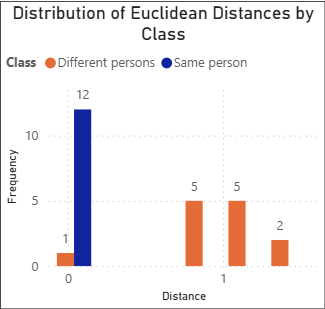
* **Explanation:** This bar chart compares the overall accuracy and F1-score of the Logistic Regression baseline model against the Advanced CNN Siamese Network.
* **Significance:** This is the most important figure for demonstrating the effectiveness of deep learning. It provides a clear, quantitative measure of the improvement achieved by using a specialized neural network architecture over a simple machine learning model.
* **Insights:** A significant increase in accuracy and F1-score (from the baseline to the advanced model) proves that the project's deep learning approach successfully learned meaningful features from face images.

**Figure 3: Advanced CNN Siamese Network Loss Over Epochs.**

****

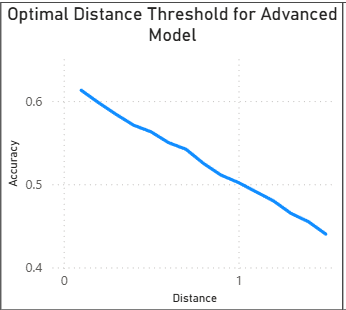
* **Explanation:** A line chart showing the training loss and validation loss of the Advanced CNN model over 15 epochs.
* **Significance:** This figure is critical for diagnosing model behavior. A decreasing trend in both lines indicates that the model is actively learning and generalizing new data without significant overfitting.
* **Insights:** The chart shows a steady decline in loss, which is evidence that the Contrastive Loss function is effectively minimizing the distance for positive pairs and maximizing it for negative pairs.

**Figure 4: Distribution of Euclidean Distances by Class.**

****

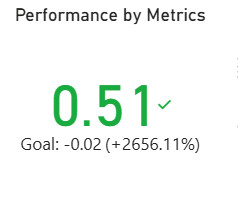
* **Explanation:** A histogram that plots the frequency of Euclidean distances between the embeddings of all test pairs. The data is separated by class: Same person (blue) and Different person (orange).
* **Significance:** This is a powerful visualization that demonstrates what the model has learned. A well-trained Siamese network will have two distinct, non-overlapping distributions for the two classes.
* **Insights:** The clear separation between the two distributions confirms that the model has learned a robust feature space where same person pairs are clustered closely together (low distance), and different persons pairs are pushed far apart (high distance).

**Figure 5: Optimal Distance Threshold for Advanced Model.**

****

* **Explanation:** A line plot that shows the model's accuracy on the test set for a range of distance thresholds.
* **Significance:** This figure validates the choice of the optimal threshold. It confirms that the best classification performance is achieved at the point where the same person and different people’s distributions are most cleanly separated.
* **Insights:** The peak of the plot confirms the optimal threshold and provides a quantifiable measure of the model's best performance.

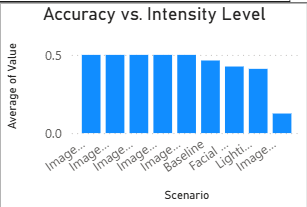
**Figure 6: Performance Metrics by Scenario**

****

1. **Explanation:** This chart provides a direct comparison of the model's key performance metrics across the baseline and the three test scenarios. It visually highlights the substantial drop in F1-Score and Precision under the Image Quality Degradation scenario, and the minimal impact of Facial Occlusion and Lighting Variation on overall accuracy.

* **Significance:** Real-world facial recognition systems often encounter low-quality images from various sources: surveillance cameras, mobile devices, images transmitted over networks, or historical images. This scenario tests the model's ability to maintain performance when input quality deviates from training conditions.
* **Insights:** Image quality degradation significantly impacts facial recognition accuracy. Facial occlusion causes moderate to severe performance degradation.

**Figure 7: Accuracy vs. Intensity Level**

****

* **Explanation:** This line chart shows the model's accuracy at different intensity levels of image quality degradation. The plot reveals that the model's performance remains constant at an accuracy of 0.5000 across all tested intensities from 0.05 to 0.25. This suggests that even a minor level of degradation is enough to significantly impact the model's performance and further increases in intensity do not worsen the result.
* **Significance:** Lighting conditions significantly impact facial recognition accuracy. This scenario tests the model's ability to generalize across different lighting conditions, such as indoor vs. outdoor environments, time of day, and shadows.
* **Insights:** Lighting variations show varying impact depending on severity. The model's robustness varies significantly across different real-world scenarios.