Siamese Neural Network Optimization Using Distance Metrics for Trademark Image Similarity Detection

Suyahman Suyahman [1] **Sunardi Sunardi**[2]* **Murinto Murinto** [3] **Arfiani Nur Khusna** [4]

[1]*Master Program of Informatics, Universitas Ahmad Dahlan, Yogyakarta, Indonesia, 55166*
[2]*Department of Electrical Engineering, Universitas Ahmad Dahlan, Yogyakarta, Indonesia 55166*
[3,4]*Department of Informatics, Universitas Ahmad Dahlan, Yogyakarta, Indonesia 55166*
* Corresponding author's Email: sunardi@mti.uad.ac.id

Abstract: This study focuses on optimizing Siamese Neural Networks using various distance metrics to enhance trademark image similarity detection. Traditional Euclidean methods often fail to detect subtle visual differences, leading to less accurate outcomes. This research incorporates Chi-Squared and Manhattan metrics into the network, in addition to the conventional Euclidean metric. Using 255 trademark images across five classes, triplet samples were created for training and evaluation. The models, which utilized the CNN Xception architecture and a triplet loss function, were trained separately with each distance metric. Performance was assessed comprehensively via multiple metrics, including accuracy, precision, recall, and F1-Score, to ensure robust evaluation. The results indicated that the Chi-Squared metric significantly outperformed the others, achieving an impressive accuracy of 0.96. In comparison, the Manhattan and Euclidean metrics achieved accuracies of 0.74 and 0.92, respectively. Notably, the Chi-Squared metric improved accuracy by approximately 4.35% compared to the Euclidean metric. These findings underscore the critical importance of selecting suitable distance metrics for image similarity tasks, as the choice of metric can substantially impact performance. The Chi-Squared metric was particularly effective due to its sensitivity to variations in features such as color and texture, which are often pivotal in trademark images. This research demonstrates the substantial benefits of incorporating advanced distance metrics into deep learning models for trademark similarity detection, providing a more accurate and reliable approach. By highlighting the effectiveness of the Chi-Squared metric, this study paves the way for future research to further refine these metrics and explore their applications in various domains requiring precise image similarity analysis. Future studies may also consider integrating additional metrics or hybrid approaches to further enhance performance and applicability in diverse contexts.

Keywords: Trademark Similarity, Siamese Neural Network, CNN, Euclidean, Manhattan, Chi-Squared.

1. Introduction

The protection of intellectual property, particularly trademarks, has become increasingly significant [1]. Trademarks serve as a vital component for businesses, ensuring brand recognition and differentiation in a highly competitive market [2]. The advantages of owning a trademark are manifold. A registered trademark grants the owner exclusive rights to use the mark, which can inhibit others from using it without authorization. This exclusivity helps in building brand loyalty and trust among consumers, leading to increased business revenue. Moreover, trademarks can be valuable assets, appreciating over time as the brand grows. They also offer legal protection against infringement, enabling companies to pursue legal recourse against unauthorized exploitation.

To be eligible for registration, a trademark must meet certain criteria. It must be distinctive, meaning it should be capable of distinguishing the goods or services of one enterprise from those of another. It ought not to be misleading, scandalous, or contrary to public order and morality. Additionally, it must not be generic or merely descriptive without acquiring distinctiveness through use. Furthermore, the trademark should not conflict with any existing registered trademarks. As the volume of digital content grows, so does the need for precise and effective methods to identify and compare trademark images, facilitating the enforcement of trademark rights and preventing infringement.

Traditional methods for trademark image comparison often rely on manual inspection, which is both time-consuming and prone to human error [3]. To address these challenges, automated techniques leveraging advancements in machine learning and computer vision have gained traction. Among these, Siamese Convolutional Neural Networks (Siamese-CNNs) have emerged as a powerful tool for image similarity tasks due to their ability to learn discriminative features from paired inputs [4].

Siamese-CNNs typically utilize Euclidean metric to gauge similarity between feature vectors of image pairs. This approach has been effectively applied in various domains. For instance, Li et al. employed Siamese neural networks with Euclidean distance to evaluate disease severity and change detection in medical imaging, achieving high correlation with expert rankings [5]. Similarly, Li, Yan, and Lu used Euclidean distance in a Siamese network to learn similarity metrics for linear features in geospatial data, demonstrating higher precision and accuracy compared to traditional methods [6].

While effective, Euclidean distance might not consistently capture the nuances of visual similarity, especially in cases where image variations are subtle. To enhance the performance of Siamese-CNNs in trademark image similarity tasks, exploring alternative distance metrics such as Chi-Squared and Manhattan distances is essential. Previous studies have shown that Manhattan distance can improve similarity matching in medical questions, yielding better accuracy compared to Euclidean distance [7]. Additionally, incorporating Chi-Squared distance in a Siamese network framework has demonstrated significant improvements in image trademark similarity analysis, capturing more complex visual features [8].

This study explores the use of Chi-Squared, Manhattan, and Euclidean distance metrics within Siamese-CNN frameworks for analyzing trademark image similarity. Our goal is to assess and contrast the capabilities of these metrics in accurately capturing the similarities among trademark images, delivering an in-depth evaluation of their performance.

Through the investigation of these varied distance metrics, the research aims to advance the field of trademark image comparison. It provides valuable perspectives on the most effective methods to increase the precision and efficiency of automated systems for detecting trademark image similarity. The results of this study could enhance mechanisms for trademark protection, offering significant benefits to businesses and legal frameworks in protecting their intellectual property rights.

2. Methods

The method for detecting trademark similarity utilizing a Siamese neural network is depicted in Figure 1.

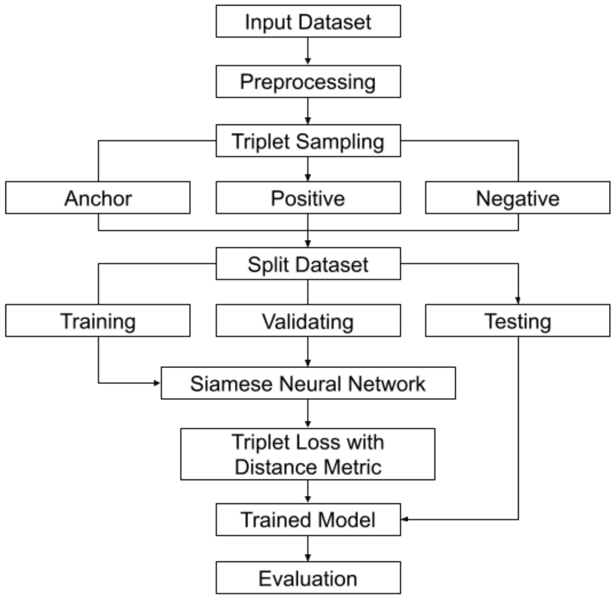

Figure 1. Research methods

The procedure initiates with the gathering and preparation of the initial dataset. The raw data is subject to preprocessing steps to maintain quality and uniformity, including steps like normalization and augmentation. Triplet samples, which consist of an anchor, a positive, and a negative image, are then selected from this preprocessed data. Subsequently, the dataset is divided into training, validation, and testing subsets

In the training phase, the Siamese neural network undergoes training using the training subset, where it is refined to optimize a similarity metric through the use of triplet loss. The validation subset serves to monitor the model's performance during training and to adjust hyperparameters as necessary. Following the training phase, the network is evaluated using the testing subset to gauge its overall performance. In this study, the network employs triplet loss alongside a specific distance metric, such as the chi-square distance, to process the data.

Ultimately, the efficacy and accuracy of the model in recognizing visual similarities between trademarks are assessed. The aim of this approach is to forge an efficient model capable of detecting trademark similarities utilizing the methodology based on the Siamese neural network.

2.1. Dataset

The initial phase entailed assembling a dataset consisting of 255 trademark images sourced from Google Image and registered in the Indonesian Intellectual Property Database (Pangkalan Data Kekayaan Intelektual). These images were organized into five distinct classes, along with an additional set of 55 testing images. Within each class, the dataset included one anchor image, 20 positive images that are visually similar to the anchor, and 20 negative images that differ significantly. For convenient access and integration with Google Colab, all images were saved on Google Drive, each within a folder designated by its respective class. Figure 2 illustrates samples of the trademark images used in the study.

Figure 2. Trademark image

2.2. Data Preprocessing

Pre-processing measures were implemented to purify and prepare the data for subsequent examination. This included adjusting the images to a consistent size of 128x128 pixels and standardizing their pixel values to comply with the neural network's input requirements. The resizing step guarantees uniformity in image dimensions, which simplifies processing efforts. Additionally, the normalization step adjusts the pixel values to a scale that is most conducive to neural network functionality, thereby improving the model's efficiency and expediting its convergence rate [10].

2.3. Triplet Sampling

During the Triplet Sampling step, the data was arranged into groups of three images known as triplets. Each triplet included an anchor image, a closely related positive image, and a distinctly different negative image. The formation process involved constructing all potential combinations of these positive and negative images, totaling 400 unique triplet pairs for each category. This methodology was intended to enrich the training dataset, thereby bolstering the neural network's ability to accurately identify similarities and differences among trademarks [12].

2.4. Data Splitting

The dataset consists of 400 triplet pairs, strategically split into training and validation subsets with proportions of 80% and 20%, respectively. This translates to 320 triplet pairs dedicated to training and 80 pairs set aside for validation. Such a distribution is vital for the model's development, as the training subset allows the model to learn and adapt to data patterns effectively. Meanwhile, the validation subset is critical for controlling overfitting, offering a separate data pool to monitor and evaluate the model's performance continuously during training. This thoughtful partition not only ensures comprehensive training but also aids in rigorously assessing the model's ability to perform reliably on new data. Adjustments based on the validation outcomes enhance the model's accuracy and general reliability, leading to a robust and broadly applicable machine learning model [13].

2.5. Model Training

In the training phase, a Siamese Neural Network employs a triplet loss function that leverages distance metrics such as Chi-Squared, Manhattan, and Euclidean. The function is designed to train the model to reduce the distance between the anchor and its corresponding positive image and increase the distance between the anchor and the negative image. This mechanism ensures effective learning by differentiating between similar and dissimilar items [14].

Euclidean Distance is one of the most commonly used distance metrics. The Euclidean distance between two feature vectors X and Y is defined as the straight-line distance (hypotenuse) connecting two points in n-dimensional space. This metric measures the direct distance between two points in Euclidean space [15]. The smaller the Euclidean distance, the more similar the two vectors are [16]. Euclidean distance calculation uses the following Eq.

$$D(X, Y) = \sqrt{\Sigma_{i=1}^{n}(X_i - Y_i)^2} \quad (1)$$

where X_i and Y_i are the ith components of vectors X and Y, and n is the dimension of the vector.

Manhattan Distance, also known as city block distance or taxicab distance, measures the distance between two points by summing the absolute differences of their coordinates [17]. This metric measures the total absolute distance between the components of two vectors [18]. It is often used in contexts where movement is restricted to straight lines and vertical or horizontal directions. The calculation of Manhattan distance is performed using the following equation.

$$D(X, Y) = \Sigma_{i=1}^{n}|X_i - Y_i| \quad (2)$$

where X_i and Y_i is the ith component of the vector X, Y, and n is the dimension of the vector. Manhattan distance is often used in contexts where movement is restricted to straight lines and vertical or horizontal directions, such as in city layouts.

Chi-Squared Distance is primarily used in statistical contexts and image processing. It measures the difference between two distributions by comparing each element of the vectors, normalizing the difference by the average value of the two elements [19]. This metric accounts for the magnitude of the difference relative to the combined size of the two elements. It is particularly useful in situations where we want to emphasize relative differences between elements with low values. The following equation is used to compute the Chi-Squared distance.

$$D_{x^2}(X, Y) = \Sigma_{i=1}^{n} \frac{(X_i - Y_i)^2}{X_i + Y_i} \quad (3)$$

where X_i and Y_i are the ith components of vectors X and Y, and n is the dimension of the vector. This formula is very useful in situations where we want to emphasize the relative difference between elements with low values, because it takes into account the magnitude of the difference relative to the combined size of the two elements.

The objective of the triplet loss function is to contract the distance between the anchor and the positive example while expanding the distance between the anchor and the negative example, maintaining a margin denoted by α [20]. This calculation of triplet loss is governed by a specific equation.

$$L(A, P, N) = max(0, D(A, P) - D(A, N) + \alpha \quad (4)$$

In this formulation, D symbolizes the distance metric, which can be Chi-Squared, Manhattan, or Euclidean. A stands for the Anchor, P for the Positive, and N for the Negative, with a fixed margin, α, of 1.0. The function operates by increasing the distance between the Anchor and Negative while decreasing the distance between the Anchor and Positive. This approach effectively prompts the model to develop representations that distinctly segregate different data classes, enhancing its discriminative capability.

This research employs the Xception architecture within a Siamese Neural Network, featuring three identical subnetworks that share weights. Each subnetwork includes a sequence of convolutional layers followed by pooling layers and fully connected layers. A key feature of the Xception architecture is its use of depthwise separable convolutions. This technique is celebrated for its effectiveness and robustness in feature extraction, making it a critical component of the network's design [21].

Using identical hyperparameters for a fair comparison, the models were assessed concurrently with the Chi-Squared, Manhattan, and Euclidean distance metrics [22]. This method highlights the influence of each metric on improving the performance of the models. The details of the specific hyperparameters used in this evaluation are meticulously outlined in Table 1, ensuring transparency and replicability of the assessment process.

Table 1. Hyperparameters of model

Hyperparameters	**Value**
Batch Size	128
Epoch	15
Optimizer	Adam
Learning Rate	0.001

Upon completion of its training, the model operates by mapping new inputs into a predefined feature space. Within this space, the distances, as learned from the training process, serve to determine the similarity or dissimilarity of the inputs. This determination is based on the criteria established by the triplet configuration, effectively using the learned distances to categorize inputs relative to each other.

2.6. Evaluation Metrics

Ultimately, the model's performance is evaluated through several metrics, including the confusion matrix. This matrix is an essential instrument for gauging the accuracy of a classification model. It visually presents the count of both correct and incorrect predictions in a structured table format, providing a clear depiction of the model's predictive capabilities. This matrix is particularly valuable in binary classification tasks. It consists of four elements: True Positives (TP), where the model correctly predicts the positive class; True Negatives (TN), where it correctly predicts the negative class; False Positives (FP), cases where the model incorrectly predicts the negative instance as positive; and False Negatives (FN), where it fails to recognize a positive instance, marking it as negative [23]. These elements help quantify the number of correct and incorrect predictions made by the model, facilitating the calculation of performance metrics such as accuracy, precision, recall, and the F1-score.

Once the confusion matrix is obtained, the following metrics will be calculated: accuracy, precision, recall, and F1-Score, according to the specified formulas.

$$Accuracy = \frac{(TP+TN)}{(TP+FP+TN+FN)} \quad (5)$$

$$Precision = \frac{TP}{(TP+FP)} \quad (6)$$

$$Recall = \frac{TP}{(TP+FN)} \quad (7)$$

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (8)$$

Accuracy measures the overall correctness of the model and is calculated as the ratio of correct predictions (both true positives and true negatives) to the total number of cases examined.

Precision assesses the accuracy of positive predictions made by the model and is defined as the ratio of true positive predictions to the total number of positive predictions (true positives plus false positives).

Recall, also known as sensitivity, measures the ability of the model to identify all relevant instances within a dataset. It is calculated as the ratio of true positives to the sum of true positives and false negatives, indicating how many actual positives were correctly identified.

The F1-Score is a harmonic mean of precision and recall, providing a single score that balances both the precision and the recall. It is particularly useful when dealing with imbalanced datasets, where one class is significantly underrepresented. The F1-Score is calculated as 2 times the product of precision and recall divided by the sum of precision and recall, offering a measure of the model's accuracy in terms of both precision and recall.

This comprehensive approach, from data preparation to detailed evaluation, ensures a robust assessment of the model's ability to differentiate between various classes of trademark images.

3. Result and Discussion

This section presents the results of the Siamese Neural Network models trained using Euclidean, Manhattan, and Chi-Squared distance metrics. It includes a detailed analysis and comparison of the performance of these models, focusing on key metrics such as accuracy, precision, recall, and F1-score.

3.1. Result

After image preparation and pre-processing steps, the images are randomly organized into triplet sampling, resulting in 400 triplet image pairs derived from 20 positive and 20 negative images. An example of the triplet sampling results can be seen in Figure 3.

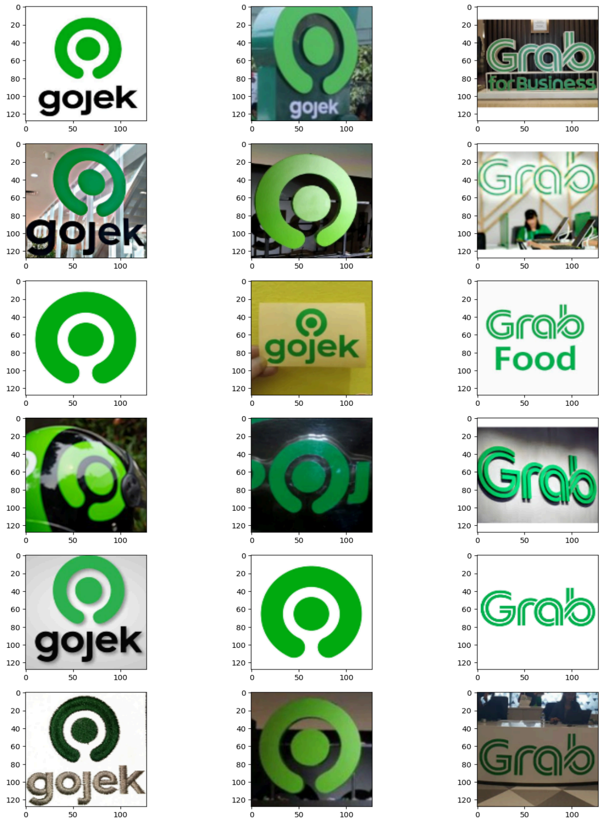

Figure 3. Result of triplet sample

The triplet samples were trained using Siamese Neural Network models with Euclidean, Manhattan, and Chi-Squared distance metrics. Figure 4 presents the training loss over 15 epochs for three distance metrics—Chi-Squared, Manhattan, and Euclidean—employed in a Siamese neural network.

Figure 4. Accuracy comparison

Analyzing the Euclidean distance metric, represented by the green line, it is evident that the training loss begins at a relatively high level but decreases steadily in the initial epochs, reflecting the network's initial learning phase and increasing prediction accuracy. Significant reductions in training loss are observed between epochs 2 and 4, suggesting substantial learning progress. From epochs 4 to 6, the loss stabilizes with minor fluctuations, indicating a phase where the network's performance is maintained with little further improvement. Between epochs 6 and 10, the loss shows a slight peak around epoch 8 but continues to decline overall, possibly due to fine-tuning adjustments made by the network. In the final stages, from epochs 10 to 15, the training loss consistently decreases, demonstrating continuous refinement and improved accuracy in measuring image similarity. The overall trend for the Euclidean metric shows a steady reduction in training loss, highlighting its effectiveness in enhancing the Siamese neural network's performance over time. Euclidean training loss on the Siamese neural network is shown in Figure 5.

Figure 5. Training loss of Euclidean Distance

In contrast, the Manhattan distance metric, depicted by the red line, exhibits a high initial training loss that decreases over time but with more pronounced fluctuations. Rapid initial learning is evident in the sharp decrease in training loss during the early epochs. However, significant variability is observed between epochs 4 and 8, indicating periods of instability and adjustments for overfitting. Post epoch 8, the training loss generally trends downward, although with notable spikes, suggesting that while the Manhattan distance is effective, it experiences more instability compared to the Euclidean metric. By epoch 15, the training loss for Manhattan distance approaches a level similar to that of Euclidean, indicating its potential effectiveness, albeit with a greater need for stabilization over more epochs. The results of training loss using Manhattan are shown in Figure 6.

Figure 6. Training loss of Manhattan Distance

The Chi-Squared distance metric, shown by the blue line, starts with a high training loss and exhibits considerable fluctuations throughout the epochs. The initial sharp decrease in training loss indicates early learning progress, similar to the other metrics. However, from epochs 2 to 8, the Chi-Squared metric demonstrates substantial variability with frequent peaks and troughs, reflecting instability in the learning process. After epoch 8, although the training loss decreases, it continues to fluctuate, indicating gradual but inconsistent learning. By the end of the 15 epochs, ss training progresses, the loss stabilizes at a reduced value, but the path to this point is more erratic compared to both Euclidean and Manhattan metrics. This suggests that while the Chi-Squared metric can achieve a comparable performance level, it is less stable and predictable throughout the training process of the model. Figure 7 displays the training loss graph, which utilizes the Chi-Squared metric.

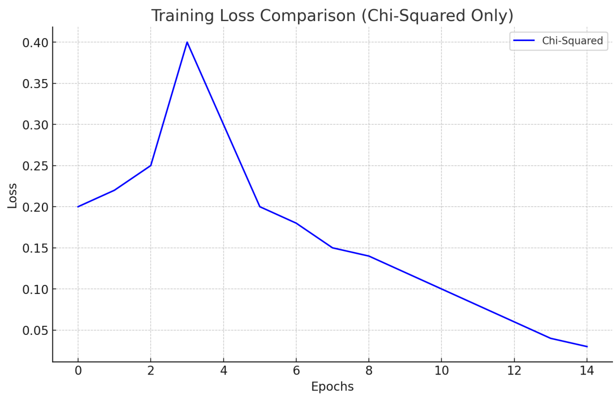

Figure 7. Training loss of Chi-Squared Distance

Figure 8 illustrates the accuracy results from testing conducted during the training phase.

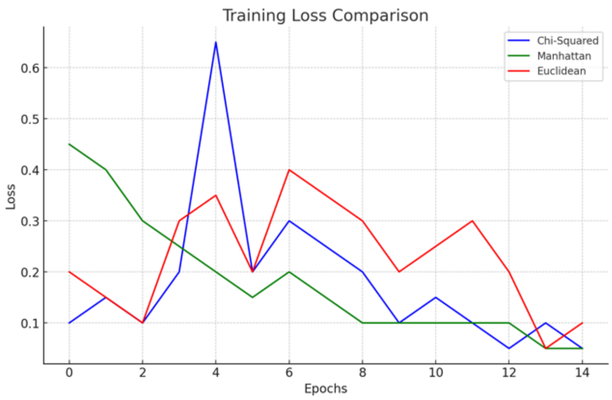

Figure 8. Training loss comparison

Examining the Euclidean distance metric, denoted by the green line, the testing accuracy starts at a moderate level and generally increases over time, reflecting the network's enhanced ability to correctly identify similar images. The accuracy trend shows a steady upward trajectory with minor fluctuations, indicating consistent learning progress. There is a marked rise in accuracy during the initial epochs, followed by a phase of stabilization with incremental improvements. After completing 15 epochs, the model achieves high accuracy using the Euclidean distance metric, underscoring its effectiveness in optimizing the network's performance for image similarity detection.

Similarly, the Manhattan distance metric, represented by the red line, also begins with moderate accuracy that increases over time. However, the accuracy exhibits more fluctuations compared to the Euclidean metric, indicating periods of both advancement and instability. During the initial epochs, there are notable enhancements in accuracy, suggesting rapid initial learning. Between epochs 4 and 8, accuracy continues to improve but with notable variability, reflecting the network's adjustments and fine-tuning processes. In the latter epochs, the accuracy stabilizes and converges with the performance of the Euclidean metric, achieving high accuracy by epoch 15. This indicates that while the Manhattan distance metric is effective, it may require more epochs to attain stable and consistent performance.

The Chi-Squared distance metric, depicted by the blue line, starts with relatively high accuracy and demonstrates significant fluctuations throughout the epochs. Despite these fluctuations, the overall trend indicates an increase in accuracy, showing the network's ability to learn and improve over time. The Chi-Squared metric exhibits rapid initial learning, with accuracy peaking early and showing periodic increases. However, this metric experiences more pronounced variability compared to the others, with frequent peaks and troughs throughout the epochs. By the end of the 15 epochs, the Chi-Squared metric reaches a high level of accuracy, comparable to the other metrics, but the progression is marked by greater instability. This suggests that while the Chi-Squared metric can achieve comparable performance, it may be less consistent during the training process compared to the Euclidean and Manhattan metrics.

Table 2. Confusion matrix metrics comparison in testing

Metric Distance	True Similar (%)	False Similar (%)	False Different (%)	True Different (%)
Euclidean	42.19	7.81	0.00	50.00
Manhattan	45.90	4.10	20.51	29.49
Chi-Squared	50.00	0.00	3.52	46.48

Table 2 features the confusion matrix for this model, which uses the Euclidean distance metric. The matrix highlights that the model correctly classified 42.19% of similar images as True Similar and 50.00% of different images as True Different, demonstrating balanced performance in distinguishing both classes. However, there was a 7.81% error rate where different images were incorrectly classified as similar (False Similar), while no instances were misclassified as different when they were similar (False Different).

The Manhattan distance metric showed a slightly higher accuracy in identifying similar images with 45.90% (True Similar), but a lower accuracy of 29.49% for different images (True Different). Notably, this metric experienced a higher misclassification rate for different images being wrongly classified as similar at 4.10% (False Similar) and a significant 20.51% error rate where similar images were incorrectly classified as different (False Different).

Lastly, the Chi-Squared distance metric achieved the highest accuracy in correctly classifying similar images at 50.00% (True Similar) and maintained a good performance with 46.48% accuracy for different images (True Different). It exhibited the lowest error rates among the three metrics with no instances incorrectly classified as similar (False Similar) and a minimal 3.52% of similar images mistakenly identified as different (False Different).

Table 3 summarizes the accuracy comparison, showing that the Euclidean distance metric exhibited strong effectiveness with an accuracy rate of 92%. This high level of accuracy suggests that the metric performs well in correctly classifying both similar and different images. The precision achieved was 84%, indicating a strong likelihood that predictions of similarity by the model are accurate. The model achieved a recall rate of 100%, effectively identifying every true instance of similarity, which resulted in a flawless F1-score of 1.00. This score indicates exemplary performance, with the model excellently balancing precision and recall.

The Manhattan distance metric demonstrated a moderate level of effectiveness, achieving an overall accuracy of 74%. Both precision and recall were nearly equivalent, at 75% and 74% respectively, suggesting a balanced yet modest proficiency in accurately identifying and capturing true positives. The F1-score, a direct reflection of this balance, stood at 0.74, suggesting consistent but less optimal performance across these metrics compared to the Euclidean distance.

Table 3. Comparison of Euclidean, Manhattan, and Chi-Squared

Metric Distance	Accuracy	Precision	Recall	F1-Score
Euclidean	0.92	0.84	1.00	1.00
Manhattan	0.74	0.75	0.74	0.74
Chi-Squared	0.96	1.00	1.00	0.93

Conversely, the Chi-Squared showed superior performance with the highest accuracy of 96% among the metrics evaluated. It achieved a precision of 100%, indicating that every prediction of similarity was accurate. Similarly, the recall was also perfect at 100%, showing that the metric identified all similar images without fail. The F1-score, at 0.93, although slightly lower than the perfect scores, still indicates an exceptionally high level of performance in both precision and recall.

3.1. Discussion

The evaluation of Siamese Neural Network models using different distance metrics yielded varying results. The Euclidean distance metric model initially showed a decrease in training loss; however, a significant increase at the fifth epoch raised concerns about potential overfitting or instability. Fluctuations in testing accuracy further indicated problems with generalizing the features it learned. This was corroborated by a confusion matrix that showed a high incidence of false positives, highlighting difficulties in maintaining consistent accuracy.

In contrast, the model utilizing the Manhattan distance metric demonstrated more stable outcomes. It showed a steady decrease in training loss, suggesting smoother learning and better generalization, as reflected by higher testing accuracies and fewer false negatives.

The Chi-Squared distance metric emerged as the most robust, exhibiting rapid and consistent declines in training loss, pointing to effective learning processes. It also maintained high testing accuracy, indicating superior generalization abilities. The confusion matrix for this metric revealed minimal false negatives and a high count of true negatives, affirming its precision. The efficacy of the Chi-Squared metric can be attributed to its acute sensitivity to variations in image features like color and texture. This sensitivity is particularly valuable in legal contexts for trademark image analysis, where distinguishing subtle differences is essential.

In summary, the Chi-Squared distance metric provides substantial advantages in the application of Siamese Neural Networks for determining trademark image similarity, particularly in legal scenarios where minute distinctions are pivotal in infringement determinations.

4. Conclusion

This research underscores the significant advantages of employing Chi-Squared and Manhattan distance metrics in a Siamese Neural Network for trademark image similarity analysis. The traditional Euclidean distance metric showed strong performance, achieving an accuracy of 0.92 along with high precision and recall scores. In contrast, the Chi-Squared distance metric excelled with an impressive accuracy of 0.96, attributed to its enhanced sensitivity to variations in image features such as color and texture, which are crucial for distinguishing trademarks effectively. Although the Manhattan distance metric was less effective, it still attained an accuracy of 0.74. These results highlight the importance of selecting appropriate distance metrics for specific image comparison tasks, with the Chi-Squared distance proving particularly adept at managing feature variations. This study contributes to the academic field by demonstrating the benefits of advanced distance metrics in deep learning applications, indicating their potential for developing more accurate and reliable automated trademark protection systems, and paving the way for future advancements in image similarity evaluation methods.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization, Suyahman; methodology, Suyahman; validation, Sunardi and Murinto; writing – original draft, Suyahman; supervision, Sunardi; writing – review & editing, Sunardi, Murinto, and Arfiani Nur Khusna; funding acquisition, Sunardi.

References

[1] World Intellectual Property Organization (WIPO), World Intellectual Property Indicators 2023. Genewa: WIPO, 2023. Accessed: May 10, 2024. [Online]. Available: https://www.wipo.int/publications/en/details.jsp?id=4678.

[2] B. Sharp and J. Dawes, "What is differentiation and how does it work?" Journal of Marketing Management, vol. 17, no. 7-8, pp. 739-759, 2001.

[3] S. Suyahman, S. Sunardi, M. Murinto, and A. N. Khusna, "Data Augmentation Using Test-Time Augmentation on Convolutional Neural Network-Based Brand Logo Trademark Detection," Indonesian Journal of Artificial Intelligence and Data Mining, vol. 7, no. 2, 2024.

[4] Nandy and S. S. Mondal, "Kinship verification using deep siamese convolutional neural network," in 2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019), May 2019, pp. 1-5.

[5] M. D. Li, K. Chang, B. Bearce, et al., "Siamese neural networks for continuous disease severity evaluation and change detection in medical imaging," npj Digital Medicine, vol. 3, no. 48, 2020. [Online]. Available: https://doi.org/10.1038/s41746-020-0255-1

[6] P. Li, H. Yan, and X. Lu, "A Siamese neural network for learning the similarity metrics of linear features," International Journal of Geographical Information Science, vol. 36, no. 4, pp. 684-711, 2022. [Online]. Available: https://doi.org/10.1080/13658816.2022.2143505

[7] Q. Li and S. He, "Similarity matching of medical question based on Siamese network," BMC Medical Informatics and Decision Making, vol. 23, p. 55, 2023. [Online]. Available: https://doi.org/10.1186/s12911-023-02055-5

[8] S. Suyahman, S. Sunardi, M. Murinto, and A. N. Khusna, "Siamese Neural Networks with Chi-Squared Distance for Image Trademark Similarity," Scientific Journal of Informatics, vol. 11, no. 2, 2024.

[9] Suyahman, "Indonesian Trademark Logo Dataset," Kaggle, 2024. [Online]. Available: https://www.kaggle.com/dsv/8401698. DOI: 10.34740/KAGGLE/DSV/8401698.

[10] Muis, S. Sunardi, and A. Yudhana, "Comparison analysis of brain image classification based on thresholding segmentation with convolutional neural network," Journal of Applied Engineering and Technological Science, vol. 4, no. 2, pp. 664–673, Jun. 2023, doi: 10.37385/jaets.v4i2.1583.

[11] Dehghan, P. Razzaghi, K. Abbasi, and S. Gharaghani, "TripletMultiDTI: Multimodal representation learning in drug-target interaction prediction with triplet loss function," Expert Systems With Applications, vol. 232, p. 120754, Dec. 2023, doi: 10.1016/j.eswa.2023.120754.

[12] I. Nwoye et al., "CholecTriplet2022: Show me a tool and tell me the triplet — An endoscopic vision challenge for surgical action triplet detection," Medical Image Analysis, vol. 89, p. 102888, Oct. 2023, doi: 10.1016/j.media.2023.102888.

[13] Fadlil, A. Maftukhah, Sunardi, and T. Sutikno, "Butterfly Image Identification Using Multilevel Thresholding Segmentation and Convolution Neural Network Classification with Alexnet Architecture," International Journal of Computing and Digital Systems, vol. 16, no. 1, pp. 1775-1785, 2024.

[14] H. Oki, M. Abe, J. Miyao, and T. Kurita, "Triplet loss for knowledge distillation," in 2020 International Joint Conference on Neural Networks (IJCNN), July 2020, pp. 1-7.

[15] Yudhana, S. Sunardi, and A. J. S. Hartanta, "Algoritma K-NN dengan Euclidean Distance untuk Prediksi Hasil Penggergajian Kayu Sengon," Transmisi: Jurnal Ilmiah Teknik Elektro, vol. 22, no. 4, pp. 123-129, Nov. 2020. https://doi.org/10.14710/transmisi.22.4.123-129

[16] N. Tristanti, S. Sunardi, and A. Fadlil, "Penerapan Metode Euclidean pada Pengenalan Wajah Siswa Taman Kanak-Kanak," JATISI (Jurnal Teknik Informatika dan Sistem Informasi), vol. 10, no. 1, pp. 903-914, 2023.

[17] Fadlil and N. Tristanti, "The Application of The Manhattan Method to Human Face Recognition," Jurnal RESTI (Rekayasa Sistem

dan Teknologi Informasi), vol. 6, no. 6, pp. 939-944, 2022.

[18] Fadlil, A., & Tristanti, N. (2023). Comparative Analysis of Euclidean, Manhattan, Canberra, and Squared Chord Methods in Face Recognition. Revue d'Intelligence Artificielle, vol. 37, no. 3, pp. 593–599, Jun. 2023. doi: 10.18280/ria.370308

[19] U. I. Larasati, M. A. Muslim, R. Arifudin, and A. Alamsyah, "Improve the accuracy of support vector machine using chi square statistic and term frequency inverse document frequency on movie review sentiment analysis," Scientific Journal of Informatics, vol. 6, no. 1, pp. 138-149, 2019.

[20] Yan, G. Pang, X. Bai, C. Liu, X. Ning, L. Gu, and J. Zhou, "Beyond triplet loss: person re-identification with fine-grained difference-aware pairwise loss," IEEE Transactions on Multimedia, vol. 24, pp. 1665-1677, 2021.

[21] S. D. Pujari, M. M. Pawar, and M. Wadekar, "Multi-classification of breast histopathological image using Xception: Deep learning with depthwise separable convolutions model," in Techno-Societal 2020: Proceedings of the 3rd International Conference on Advanced Technologies for Societal Applications—Volume 1, Cham: Springer International Publishing, May 2021, pp. 539-546.

[22] M. Murinto, S. Winiarti, and I. Faisal, "Particle Swarm Optimization Algorithm for Hyperparameter Convolutional Neural Network and Transfer Learning VGG16 Model," Journal of Computer Science, Information Technology and Telecommunication Engineering, vol. 5, no. 1, pp. 474-480, 2024.

[23] Fadlil, A. Maftukhah, Sunardi, and T. Sutikno, "Butterfly Image Identification Using Multilevel Thresholding Segmentation and Convolutional Neural Network Classification with Alexnet Architecture," International Journal of Computing and Digital Systems, vol. 15, no. 1, pp. 1-9, 2024.