

# Image Based Machine Learning for Component Identification for Remanufacturing

Abu Islam<sup>1</sup>, Sri Priya Das<sup>1</sup>, Nenad Nenadic<sup>1</sup>

Everardo FriasRios<sup>2</sup>, Audra Stapella<sup>2</sup>, Sebastian Przybylski<sup>2</sup>

<sup>1</sup>Rochester Institute of Technology, 1 Lomb Memorial Dr, Rochester, NY 14623

<sup>2</sup>CoreCentric Solutions, Inc, 191 E North Ave., Carol Stream, IL 60188

## Abstract

Remanufacturing durable goods presents a significant opportunity to reduce landfill waste and decrease reliance on energy-intensive recycling methods. The first critical step in this process involves sorting and inspecting returned components, referred to as "cores" within the remanufacturing industry. These operations are currently performed manually, making them labor-intensive, error-prone, and ergonomically inefficient. The challenge is further compounded in environments with low-volume, high-variability product mixes, where rapid operator training and adaptation to unpredictable returns and diverse physical conditions are especially difficult. To address these limitations and alleviate staffing constraints, an automated image-based part identification system has been developed. This solution leverages component images to train a neural network using Siamese architecture, incorporating a pre-trained ResNet-18 model as the feature extractor and a triplet loss function. This configuration enables effective training with limited datasets and allows the system to distinguish between visually similar components. A complementary mobile application prototype has been created to streamline the identification process and enable early customer engagement. Users can capture an image of a product using a smartphone and receive immediate identification results. They can approve the result, or modify it, and both correct and incorrect detection is saved with accompanied dataset to automatically create a valuable dataset for future model improvements, and extensions to additional parts. The system architecture is modular, separating end-user interaction from model training and updates, thereby enhancing scalability and maintainability. The concept is currently undergoing demonstrations and validation at an industry partner's small appliance remanufacturing facility. The paper provides details on image acquisition techniques, model development workflow, testing protocols, and performance metrics. It will also examine the challenges and opportunities associated with automated defect detection, particularly in products with highly reflective surfaces.

**Keywords:** Computer vision, Siamese networks, Remanufacturing, Object identification, Cores, Neural Networks

## Introduction and Motivation

Remanufacturing, a vital component of the circular economy, involves restoring used products to like new condition rather than discarding or recycling them (1). Considering the wide disparity in the quality of returned used products, commonly referred to as "cores" (2), one of the critical steps of remanufacturing is to assess the condition of these cores. Hence, "inspection and disposition" is one of the key components of a remanufacturing business. But the inspection and disposition processes are often manual, making them labor-intensive, time-consuming, prone to errors, and ergonomically challenging (3). Overarching goal of the project with an industry partner CoreCentric is to address these inefficiencies and mitigate staffing shortages by automating the part identification, inspection and sorting systems.

CoreCentric is a remanufacturing and supply-chain solutions company specializing in repairing, refurbishing and providing sustainable replacement parts for consumer and commercial electronics and appliances. They process over 600,000 returned consumer appliances annually across various product lines. The average throughput yield for re-manufacturable products is approximately 50%–60%, with the remainder either recycled or sent to landfill. These tasks demand extensive training to ensure accurate evaluation of each returned part. Apart from that CoreCentric manages over 30 Stock Keeping Units (SKU) with unpredictable product lines and varying physical conditions of the returned products. Hence, the proposed automated solution needs to be accurate, flexible and adapt to handle low-volume, high SKU environments.

As a first step to tackle some of these challenges a computer vision system that can automatically identify appliances as they move along the remanufacturing line is developed. Given the unpredictable nature of the returned appliances and the time-intensive process of gathering images, the proposed solution should be a) trainable on small image data of selected appliances and b) extendable to new or unseen appliances with minimal or no retraining. A Siamese network (SN) (4) with its one-shot learning capability was selected as it meets both the above-mentioned criteria *viz.*, it can be trained on relatively small datasets and can be extended to unseen data. After a successful implementation and demonstration of SN based appliance identification model on test datasets, a mobile application was developed for real-world testing. The mobile app was deployed at CoreCentric's remanufacturing facility, allowing operators to capture images of returned products and receive instant product identification.

The following sections provide more details on related research, data collection, SN implementation, testing protocols, results, mobile app development, and mobile app performance. Challenges and potential for automated defect detection in returned appliances are also discussed.

## Review of Related Work

Most computer vision-based techniques in remanufacturing focus on defect detection after disassembly, with relatively less attention given to product identification. Product identification, the first step in the remanufacturing process, determines how items are processed in subsequent stages such as defect detection, disassembly, inspection, and repair. For practical deployment on a remanufacturing line, product identification must be fast and computationally efficient. However, many current automation solutions still rely on manual identification, with limited research addressing this step.

In (5) Schlüter et al., proposed a Convolutional Neural Network (CNN) based method for product identification using Logic.Cube, a vision-based inspection system combining five color cameras and a scale to capture images alongside intrinsic attributes such as weight, volume, and dimensions. Their approach employed CNN architectures (LeNet, AlexNet) to classify objects from a database of 146 items, achieving up to 98.6% accuracy using image data alone, and 96% when incorporating weight, volume, and rule-based identifiers. Building on this work, an AI-enhanced system that integrates machine learning with visual recognition to automate identification, inspection, and sorting of returned parts was developed in (6). In a pilot study on automotive components, this ResNet50-based CNN achieved 96% identification accuracy from images, improving throughput, accuracy, and decision reproducibility while reducing manual effort. Similarly, (7) proposed an automated core-sorting system utilizing multi-camera conveyors and controlled lighting. Their deep learning models, including transfer learning and YOLO detection, achieved over 95% accuracy within 10 seconds, with multi-camera probabilistic fusion enhancing robustness in ambiguous cases. (8) further advanced this domain by presenting an ensemble approach that combines YOLOv5 and DeepLabV3+, with a decision tree merging outputs from both models to classify parts at 93.33% accuracy.

While these approaches are effective, they often rely on specialized lighting and equipment for data acquisition and require large datasets with hundreds of images per class for accurate performance. This is where SNs offer an advantage. Unlike traditional classification and object detection methods, SNs learn to find similarities between images by generating discriminative embeddings, enabling identification from a small set of reference images (4). Additionally, SNs support few-shot learning, allowing the model to generalize or scale to new classes without retraining. In (9) Othniel et al., demonstrated SN performance on the AT&T face dataset (400 grayscale images of 40 subjects), achieving 0.81 accuracy, 0.9375 recall, 0.9 precision, and 0.732 F1-score, highlighting strong identification of similar image pairs. In medical imaging, (10) implemented a SN using a VGG-16 network with contrastive loss to extract 4096-dimensional feature vectors for content-based image retrieval. Their two-phase approach includes K-means clustering followed by SN-based ranking of top-k similar images. It was demonstrated that combining deep metric learning with clustering effectively captures high-level semantic features for accurate and efficient retrieval. (11) applied SNs for few-shot object recognition in mobile robots operating in office and safety-critical environments. Using a ResNet50 backbone with L2 normalization, their approach leveraged ImageNet-based transfer learning to achieve 84% accuracy and 0.87 F1-score on novel classes. While prior applications of SNs highlight their strength in learning from small datasets and distinguishing visually similar items, their potential for remanufacturing product identification remains underexplored. Building on these successes, an SN tailored to appliance identification is implemented, providing fast, scalable, and reliable results.

## Technology Approach

### Data Collection and Preprocessing

As the SN estimates object similarities by comparing feature vectors, the input image data should capture as much distinctive visual information as possible. For the appliances in this study, key identifying features include control panel, placement of the group head & steam wand, and the overall shape, and size of the unit. To represent these features effectively, the front side of each appliance was imaged from four angles as shown in Figure 1.



Figure 1 Data collection by capturing front side of the appliance at various angles

Image data were collected at CoreCentric facilities as returned products enter the remanufacturing line. An overview of an initial dataset used in this work is presented Figure 2a. The dataset is noticeably imbalanced, as some appliance models contain far more images than others due to unpredictable number of returns. The dataset of 108 front-view images was split into training and validation sets with 64 and 34 respectively (Figure 2b). Because Siamese architecture compares image pairs, only appliance models with at least four images or more were included in training to ensure effective learning.

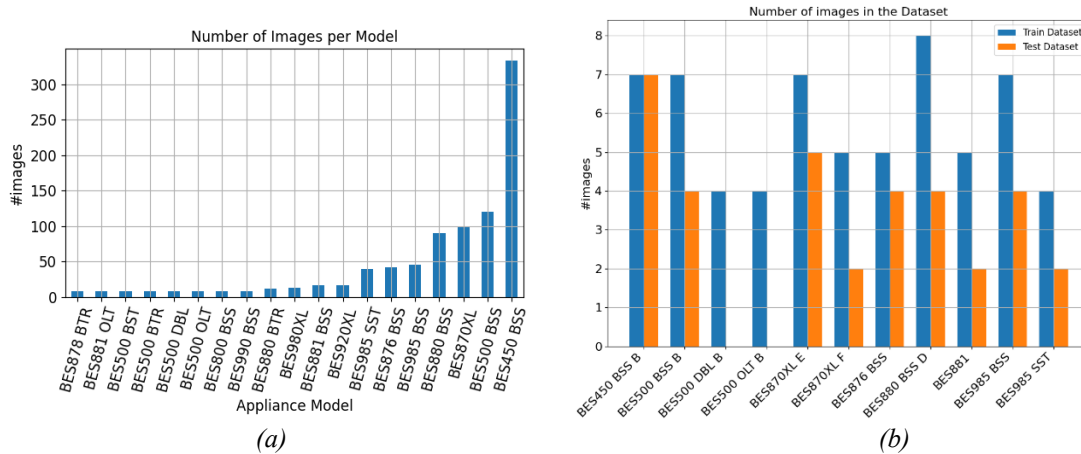


Figure 2: (a) Initial set of images captured for appliance classification

Since the images were captured in an industrial environment, the backgrounds varied considerably, adding noise and inconsistency during feature extraction. To address this, each image was preprocessed using U<sup>2</sup>-Net deep learning-based background removal method (12), which isolated the appliance from its surroundings as shown in Figure 3a.

### Siamese Network Implementation

A SN illustrated in Figure 3b determines the similarity or difference between image pairs based on a loss function rather than absolute classification (4). During training, the network computes distance between feature embeddings of paired images. If the two images belong to the same model, the weights are updated to minimize their embedding distance; if they belong to different models, the distance is maximized. Because the SN learns feature relationships through embeddings, it can be effectively trained using a relatively small number of images. Once trained, the same network can be used to generate reference embedding for new appliance models without retraining.

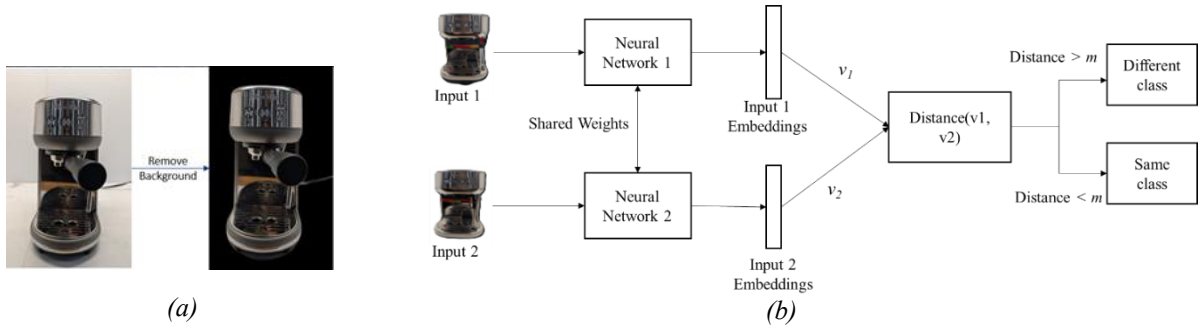


Figure 3: (a) Background removal (b) Typical Siamese Architecture

## Loss Functions and Input Generation

An essential aspect of metric learning in Siamese architecture is the generation of image pairs for training which in turn depends on the chosen loss function. The most common loss functions used in pair based metric learning are contrastive loss, triplet loss and N-pair loss. They differ in how the samples are structured, contrastive loss compares pairs of images, triplet loss operates on triplets and N-pair loss extends the triplet concept to multiple negative samples. In this study, both contrastive and triplet loss functions are implemented. The following sections describe their formulation, implementation and results.

## Contrastive Loss

The objective of contrastive loss is to bring embeddings of similar inputs closer together in the feature space and push dissimilar embeddings apart by at least a defined margin  $m$ , as shown in Figure 4. During training, the distances between the anchor image and all other embeddings are computed. Distances between the anchor and same-class samples are minimized, while distances between anchor and different-class samples are maximized beyond the margin threshold. The contrastive loss function adapted from (13), is shown in Equation 1.

$$L = (1 - y) \cdot \frac{1}{2} D^2 + y \cdot \frac{1}{2} [\max(0, m - D)]^2 \quad (1)$$

where  $D$  is distance between the embeddings,  $y = 0$  for similar and  $y = 1$  for dissimilar pairs,  $m$  is the margin.

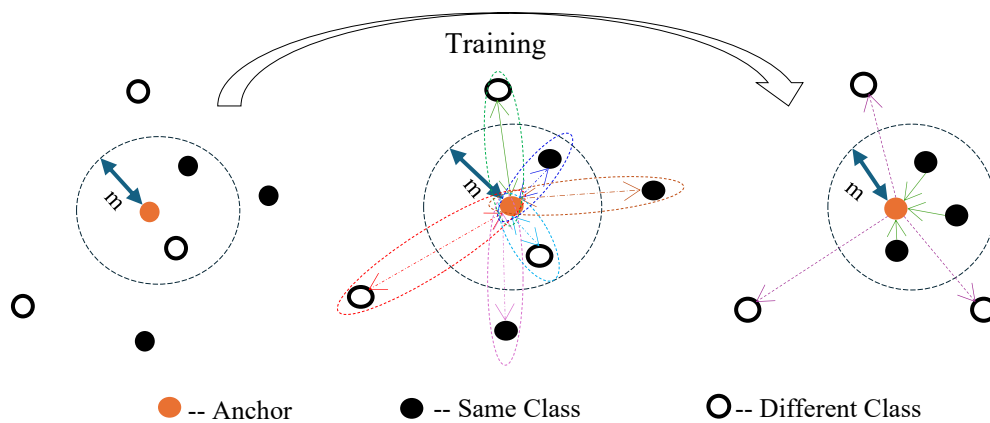


Figure 4: Contrastive loss training

Figure 5a illustrates three examples of input pairs, where label '0' indicates that the pair belongs to the same class and label '1' indicates different classes. The SN for this implementation used a simple CNN backbone with a

128-dimensional embedding space (Figure 5b). The network was trained for 50 epochs using a learning rate  $\eta = 10^{-4}$ , the Adam optimizer, and the margin of 1.

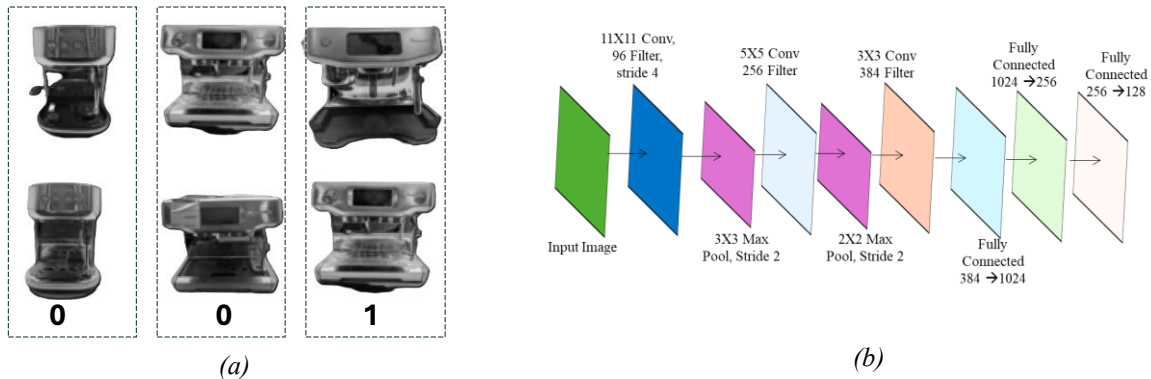
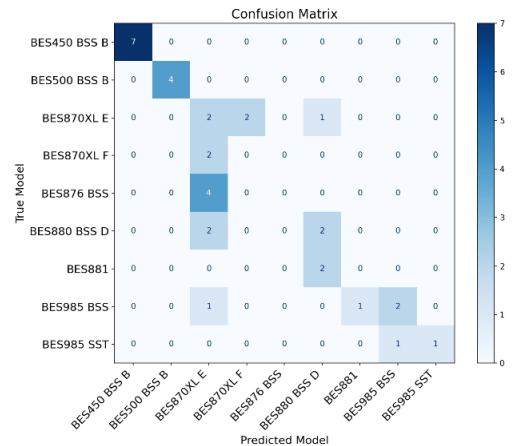
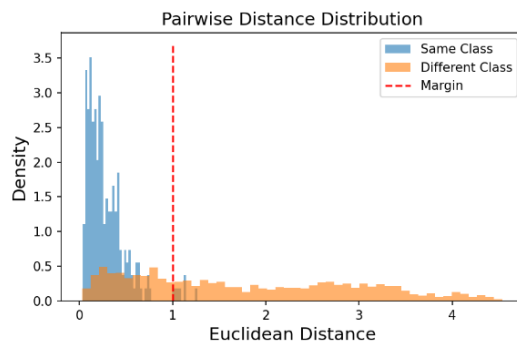


Figure 5 (a) Contrastive loss input pairs (b) CNN architecture Backbone

After training, embeddings for all training and test images were generated. For each test image, the Euclidean distance between its embedding and those of the training images was computed, and the label corresponding to the smallest distance was assigned as the ‘Predicted Model’. Results for the contrastive loss model are shown in Figure 6.

Figure 6a depicts the pairwise Euclidean distance distribution for the test dataset. The network effectively separated the same class images, as most of their distances were below the margin. However, some dissimilar pairs also had distances below the margin, indicating difficulty distinguishing certain models.

The confusion matrix in Figure 6b shows that models ‘BES450BSS B’ and ‘BES500BSS B’ were correctly classified, while results for other models were mixed. For example, multiple models were predicted as ‘BES870XL E’. Despite these limitations, performance was better than expected given the small training set. Encouraged by these results, a second SN was



developed using a Resnet-18 backbone and trained with triplet loss function as described in the next section.

Figure 6: (a) Pairwise distance distribution (b) Confusion Matrix of test dataset

### Triplet Loss

Similar to contrastive loss, the triplet loss function aims to minimize distances between embeddings of similar samples while maximizing those between dissimilar samples. However, instead of pairs, it uses triplets of images - an anchor, a positive (same class), and a negative (different class) as shown in Figure 7(14). During training, the distance between anchor and the positive sample is minimized below the margin, while the distance between the anchor and the negative is increased above it. A sample training dataset is shown in Figure 8a, with corresponding labels for anchor, positive and negative samples.

(a)

(b)

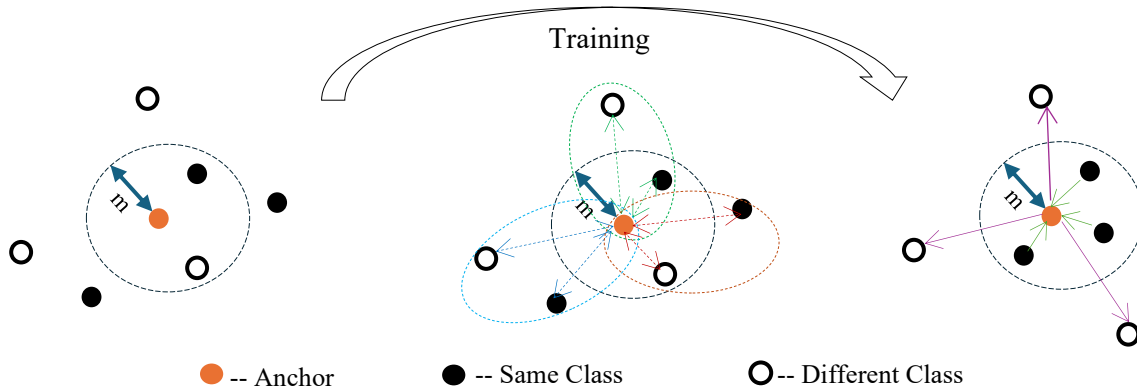


Figure 7: Triplet loss

Due to the limited representational capacity of the simple CNN used earlier (Figure 5b), a pretrained ResNet-18 architecture was adopted to extract more discriminative feature embeddings. ResNet-18 architecture, with a 128-dimensional embedding space is shown Figure 8b.

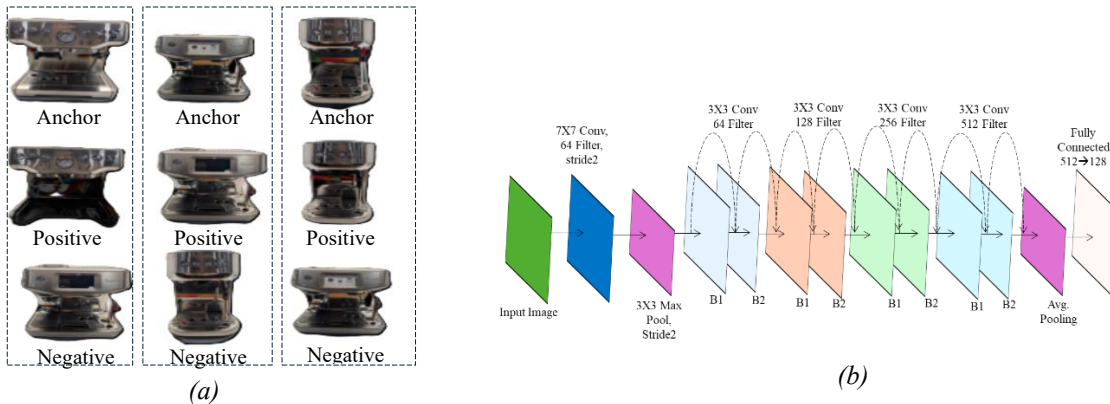


Figure 8: (a) Sample inputs for triplet loss and Figure (b) ResNet-18 backbone architecture

The same training and test datasets, parameters and inference methodology as the contrastive loss experiment were used. The results shown in Figure 9 demonstrate improved performance with the ResNet-18 triplet loss model. The distance distributions show clearer separation between same and different class embeddings, and the confusion matrix reveals accurate predictions for most models. Misclassifications primarily occurred between models differing only in color, e.g. *BSS* vs *SST*, suggesting that future work should incorporate color-based features into the model.

### Contrastive loss vs triplet loss

Table 1 presents a quantitative comparison between the SN trained with contrastive loss and triplet loss. Evaluation metrics include the ROC-AUC, intra/inter-class distance ratio, Recall@1, Recall@5 and Mean Average Precision(mAP).

- ROC-AUC measures the overall class separability (1 indicates perfect ranking)
- Intra/inter-class distance ratio evaluates embedding space compactness; lower values indicate tighter, more distinct clusters
- Recall@1 and Recall@5 measure the proportion of correct matches within the top 1 and top 5 ranked predictions, respectively
- mAP assesses retrieval accuracy by averaging precision across all recall levels

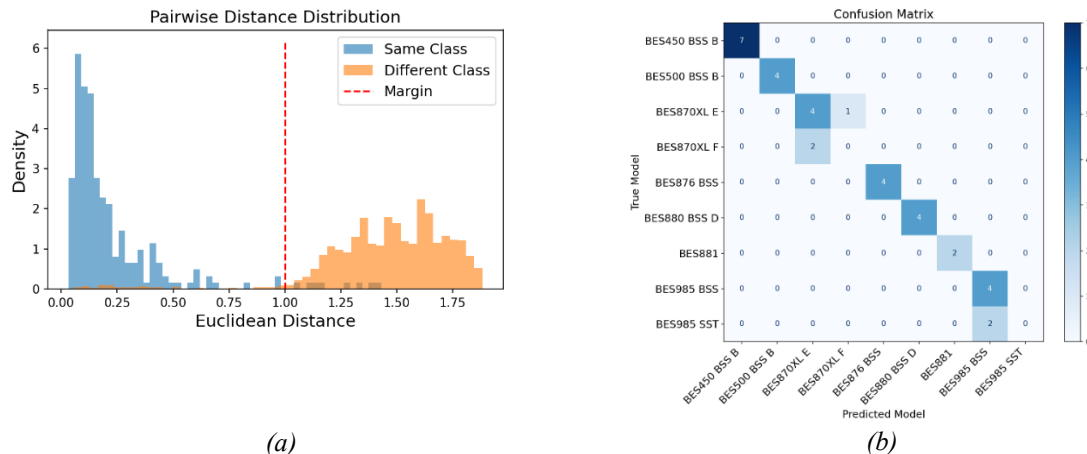


Figure 9: (a) Class separation of test embeddings (b) Confusion Matrix of the test dataset

Metric	Contrastive Loss	Triplet Loss
ROC-AUC Score	0.91	0.98
Intra/Inter Ratio	0.17	0.21
Recall@1	0.5	0.82
Recall@5	0.8	0.91
mAP	0.6	0.87

Table 1: Overall performance comparison of contrastive loss and triplet loss

As shown in Table 1, the triplet loss with ResNet-18 backbone consistently outperformed the contrastive loss model for the same number of training samples per class. Both achieved high ROC-AUC values and low intra/inter ratios, suggesting good global separation. However, in the case of contrastive loss, clusters were close together and occasionally overlapped, reducing classification confidence. This was reflected in the Recall@1 value of 0.5, indicating that only half of the top-ranked predictions were correct. In contrast, the triplet loss model formed well-separated clusters, resulting in higher accuracy and a mAP of 0.87, indicating fewer false negatives across all classes. Motivated by this performance, the SN trained with triplet loss was deployed for real-world testing at CoreCentric’s facility via a mobile application.

## Mobile App Development

The SN with triplet loss function was implemented as a mobile application to enable appliance identification directly within the remanufacturing workflow. As shown in Figure 10a Streamlit-based architecture was adopted. This set up allows the host (RIT) to continue updating and retraining the model, while operators at CoreCentric can use mobile devices to capture images and receive identification results instantly. In addition to providing on-site identification, the app serves as a data collection tool, gathering new field images that more accurately represent real-world variations. The interface features a simple web-based GUI that allows users to confirm or correct predictions using dropdown menus, as shown in. Figure 10 (b,c). Each image, along with its file path, predicted label, correction (if any), and verification result, is saved on the server for ongoing model improvement.

## App Performance

The mobile app was used to identify a total of 81 appliances. Of these 54 (66.66%) were correctly identified and 27 (33.33%) were misclassified. This outcome was expected, as the real-world conditions differ significantly from controlled test environments. An analysis of the misclassifications, summarized in Table 2, revealed two main error categories: development-related and image acquisition-related. The largest portion of the errors stemmed from unseen data, appliance models not included in training. As the network compares query embedding with reference embedding of the training data, if an appliance model is not in the training set then the reference embedding is not available for comparison. These will be addressed in future iterations by extending the training dataset and refining color sensitivity.

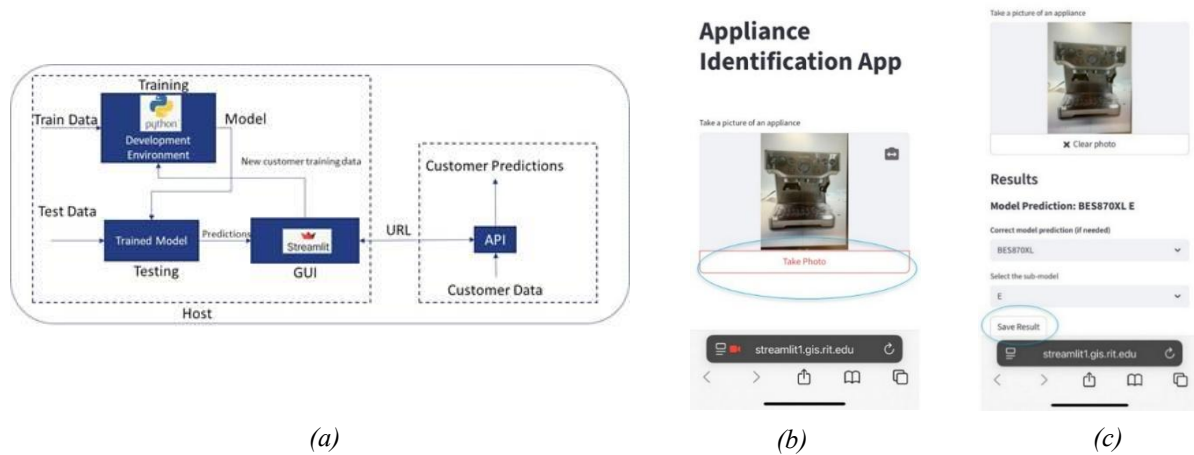


Figure 10: (a) Streamlit architecture for a mobile App for appliance identification (b) App GUI (c) Steps to accept Prediction

Error type	Count	Relative error (%)	Notes
Color mismatch	5	6.1	Same appliance model but in different colors
Unseen/new data	10	12.3	Models that were not used in training
Incorrect Orientation	4	5.00	Orientation different from images in the training data
Background	2	2.	Too close shots very background is not visible
Other	6	7.4	Tags, packing material, app testing etc.
Total	27	33.34	

Table 2: Breakdown of incorrect identifications of the mobile app

Image acquisition issues, such as poor orientation or background clutter will be mitigated through operator training on best practices for image capture. A key insight from field deployment is that many returned appliances include tags or packing material. To better represent real-world conditions, future training datasets will include such variations.

## Discussion

The contrastive loss network with a custom CNN backbone effectively separated appliance models with pronounced global differences, such as overall shape or width, demonstrating that coarse visual features were captured even with limited training data. However, it struggled with appliances differing only in fine-grained details like button or knob layouts, reflecting the low capacity of the custom CNN and the small dataset. Consequently, embedding clusters were well separated for distinct models but overlapped for visually similar classes, causing retrieval confusion consistent with prior studies showing contrastive loss can underperform on small, fine-grained datasets.

To address this, a pretrained ResNet-18 backbone was trained using triplet loss, leveraging prior feature learning and achieving stronger intra-class compactness and inter-class separation. This improved retrieval performance, with roughly 80% of queries correctly identified. Remaining errors mainly involved visually similar appliances differing only in color or minor design elements, highlighting dataset limitations in capturing subtle differences.

The SN's integration into a mobile application confirmed its practical feasibility for remanufacturing. However, generalization to new or unseen models can be achieved by capturing a reference image and generating its embedding, a one-time process that extends adaptability without retraining. These results demonstrate that SNs can provide fast, reliable, real-time appliance identification in operational settings.

## Conclusions & Recommendations

This study explored the use of SNs for product identification in remanufacturing, addressing industry needs without requiring large datasets or specialized setups. Leveraging few-shot learning, SNs generate discriminative embeddings that enable identification from small reference sets. A custom CNN-based SN with contrastive loss effectively captured coarse visual features but struggled with fine-grained differences. Replacing it with a pretrained ResNet-18 and triplet loss improved intra-class compactness and inter-class separation, achieving approximately 80% retrieval accuracy. Integration into a mobile application demonstrated practical feasibility, enabling rapid identification. In the short term, efforts will focus on reducing errors caused by visually similar appliances and color variations and extending the functionality of the mobile app to new appliance models without retraining. In the long term, product identification will be combined with defect detection to provide a comprehensive system for analysis and classification of incoming products in remanufacturing.

## Acknowledgements

This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Advanced Materials & Manufacturing Technologies Office (AMMTO), Award Number DE-EE0007897 awarded to the REMADE Institute, a division of the Sustainable Manufacturing Innovation Alliance Corp (SMIA), and by the New York State Department of Economic Development (DED).

The opinions, results, findings and/or interpretations of data contained herein are the responsibility of the author and do not necessarily represent the opinions, interpretations or policy of the State of New York; furthermore, the views expressed herein do not necessarily represent the views of the U.S. Department of Energy, the United States Government, SMIA, or the REMADE Institute.

## References

1. Nasr N. Remanufacturing and the Circular Economy. In: *A New Dynamic 2: Effective Systems in a Circular Economy*. Ellen MacArthur Foundation; 2016.
2. Lund RT. Remanufacturing: the experience of the United States and implications for developing countries (English). 1985.
3. Steinhilper R. Remanufacturing: the ultimate form of recycling. Fraunhofer IRB Verlag; 1998. 108 p.
4. Koch G, Zemel R, Salakhutdinov R. Siamese neural networks for one-shot image recognition. In: *International Conference on Machine Learning Workshop*. 2015. p. 1–30.
5. Schlüter M, Niebuhr C, Lehr J, Krüger J. Vision-based Identification Service for Remanufacturing Sorting. *Procedia Manuf*. 2018; 21:384–91.
6. Schlüter M, Lickert H, Schweitzer K, Bilge P, Briese C, Dietrich F, et al. AI-enhanced Identification, Inspection and Sorting for Reverse Logistics in Remanufacturing. *Procedia CIRP*. 2021; 98:300–5.
7. Islam, A., Jain, S., Thurston, M., Nenadic, N, Moss, B., Greenberg J. Image Based Machine Learning in Automotive Used Parts Identification for Remanufacturing. In: *Proceedings of REMADE Circular Economy Conference*, Washington DC. 2024.
8. Saiz FA, Alfaro G, Barandiaran I. An Inspection and Classification System for Automotive Component Remanufacturing Industry Based on Ensemble Learning. *Information*. 2021 Nov 23;12(12):489.
9. Clark Gngangby Aaron Othniel, Choukpin Adoto Mignonkoun Sourou Yannick. Advancing Image Retrieval Through Similarity Measures using Siamese Neural Networks. *International Journal of Engineering Research & Technology (IJERT)*. 2025 Jan;14(1).

10. Vu VH, Ta QH. The two-phases model combining Siamese network and clustering improves semantic distance in medical image retrieval. *Multimed Tools Appl.* 2025 Mar 28;84(33):40937–57.
11. Chiatti A, Bardaro G, Bastianelli E, Tiddi I, Mitra P, Motta E. Task-Agnostic Object Recognition for Mobile Robots through Few-Shot Image Matching. *Electronics (Basel)*. 2020 Feb 25;9(3):380.
12. Qin X, Zhang Z, Huang C, Dehghan M, Zaiane OR, Jagersand M. U2-Net: Going deeper with nested U-structure for salient object detection. *Pattern Recognit.* 2020 Oct;106:107404.
13. Hadsell R, Chopra S, LeCun Y. Dimensionality Reduction by Learning an Invariant Mapping. In: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2 (CVPR'06). IEEE; p. 1735–42.
14. Kim H, Kim H, Ko B, Shim J, Hwang E. Two-stage person re-identification scheme using cross-input neighborhood differences. *J Supercomput.* 2022 Feb 23;78(3):3356–73.

## About the Author(s)

- **Abu Islam** a Research Associate Professor at RIT. received his Ph.D. in Mechanical Engineering from Rensselaer Polytechnic Institute in 1995. His research interests include the use of computer vision, data analytics and artificial intelligence in remanufacturing and health monitoring of machines. He holds a Black Belt in Design for Lean Six Sigma and 26 US Patents.
- **Sri Priya Das** is a Data Science Engineer at the Golisano Institute for Sustainability at RIT. She specializes in data-driven solutions using machine learning, deep learning and computer vision for complex challenges in remanufacturing, diagnostics and engineering applications. She holds a Ph.D. in Electrical Engineering from RIT and has authored several publications in leading scientific journals and conferences.
- **Nenad Nenadic** is a Research Associate Professor at RIT. He received Ph.D. in Electrical and Computer Engineering from the University of Rochester in 2001. His research interests include sensors, battery systems, automation, and machine learning.
- **Everardo FriasRios** received his B.S. in Computer Engineering Technology from DeVry University in 2014. He joined CoreCentric Solutions Inc. the same year, where he is currently an Engineering Manager. He is responsible for leading teams in the design and development of remanufacturing procedures for consumer electronics.
- **Sebastian Przybylski** received his B.S. in Computer Science from Lewis University in 2025. He joined CoreCentric in 2023 as an intern and now works as a full-time associate engineer. He is responsible for implementing new machine vision technologies into production and updating IT/Computer systems for data I/O processes.