2 LITERATURE REVIEW

This section aims to provide a background of existing research in the loss functions of the Deep Metric Learning. It explains the importance of the research and activity in the development of novel loss functions. Similar survey of methods and loss functions has been done also for Deep Reinforcement Learning [109]. Results of the findings for Deep Reinforcement Learning have been included in Appendix C.

2.1 Methodolgy of Literature Review

The methodology of SLR (Systematic Literature Review) presented in this document is based on a systematic mapping study [80] [43]. The results of SLR contain the map of clusters based on the origins of loss functions and methods, as well as a qualitative review based on research questions. The results also include a list of limitations identified for loss functions and methods used in the reviewed papers.

The method for selecting and evaluating papers contains the steps listed in Fig. 3. Initially the most well-known publications [7], [23] in the field of deep metric learning (DML) have been selected. Additionally, the following keywords were used for the initial search of papers: triplet loss, contrastive loss, ranking loss, deep metric learning, representation learning, one-shot learning, zero-shot learning, product re-identification task, signature re-identification, face re-identification task. Then the publications have been thoroughly analyzed and documented to check if publications match the field of DML loss function research. Then matching to Quality Assessment criteria has been evaluated. If at least single assessment criterion has been met, a publication was added to the main list. In addition, if answers to research questions have been found in selected publications, then those were documented. The references and citations of this publication have been found. For each of the relevant publications, their citation count has been found and divided by years passed since publishing. Those with the highest value of influence were analyzed first.

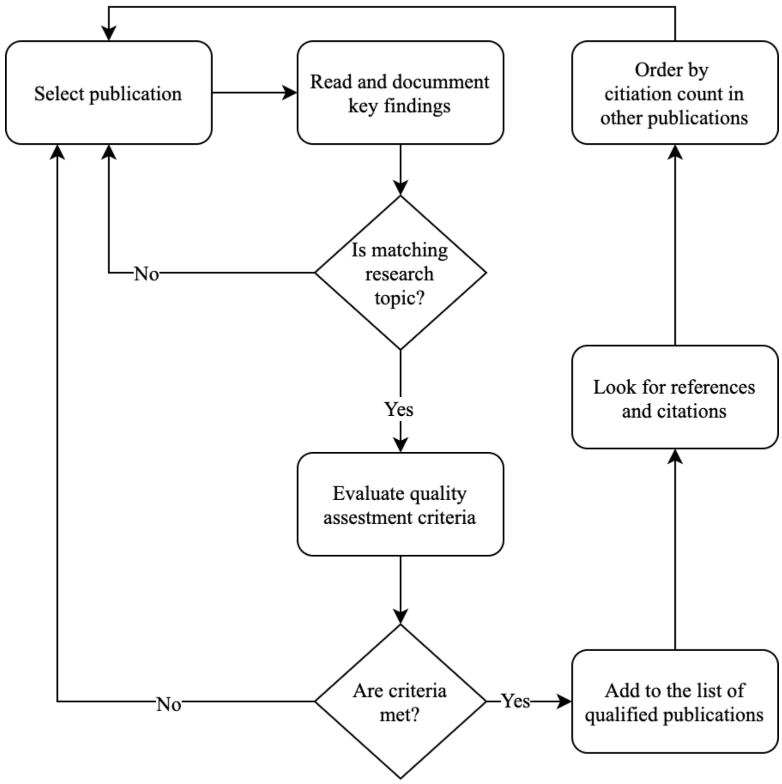

Figure 3: The methodology of SLR.

To find a valid direction of further research, few research questions (RQ) were selected. The research questions addressed by this study are:

- RQ1: What kinds of functions have been studied similar to Triplet Loss functions?
- RQ2: Do the novel loss functions achieve significantly better results than previous functions?
- RQ3: Do the novel loss functions have theoretical grounding, or are they purely empirical?
- RQ4: What are the limitations of novel loss functions?

2.2 Results of the Literature Review on Deep Metric Learning

The results of SLR regarding DML are mapped in multiple tables depending on the relevant properties extracted from papers. Information about authors, affiliation, country of origin, and conferences regarding DML have been listed in Table 1.

Publications have been ordered by the year of publishing, and the numbering of publications has been maintained also in the following tables.

Table 1:

Authors and conferences on studies regarding DML.

No	Title	Authors	Affiliation	Country	Year	Conference / Journal
1	Signature Verification Using A "Siamese" Time Delay Neural Network [7]	J. Bromley, J. Bentz, L. Bottou, I. Guyon, Y. LeCun, C. Moore, E. Sckinger, R. Shah	AT&T Bell laboratories	USA	1993	INT J PATTERN RECOGN
2	Neighbourhood Components Analysis [27]	J. Goldberger, S. Roweis, G. Hinton, R. Salakhutdinov	AT&T Bell laboratories	Canada	2004	NIPS
3	Learning a Similarity Metric Discriminatively, with Application to Face Verification [14]	S. Chopra, R. Hadsell, Y. LeCun	NYU	USA	2005	CVPR
4	Distance metric learning for large margin nearest neighbor classification [127]	K. Q. Weinberger, L. Saul	Yahoo!, University of California	USA	2005	NIPS
5	Large scale metric learning from equivalence constraints [46]	M. KŽstinger, M. Hirzer, P. Wohlhart, P. Roth, H. Bischof	Graz University of Technology	Austria	2012	CVPR
6	Quadruplet-Wise Image Similarity Learning [52]	M. Law, N. Thome, M. Cord	Sorbonne University	France	2013	ICCV
7	Reidentification by Relative Distance Comparison [135]	W. Zheng, S. Gong, T. Xiang	College of Electronic and Information, South China University of Technology	China	2013	TPAMI
8	Deep Metric Learning for Practical Person Re-Identification [132]	D. Yi, Z. Lei, S. Li	IEEE	China	2014	ArXiv
9	FaceNet: A unified embedding for face recognition and clustering [23]	F. Schroff, D. Kalenichenko, J. Philbin	Google	USA	2015	CVPR
10	Improved Deep Metric Learning with Multi-class N-pair Loss Objective [98]	K. Sohn	NEC	USA	2016	NIPS
11	A Discriminative Feature Learning Approach for Deep Face Recognition [128]	Y. Wen, K. Zhang, Z. Li, Y. Qiao	SIAT	China	2016	ECCV

12	Deep Metric Learning via Lifted Structured Feature Embedding [100]	H. O. Song, Y. Xiang, S. Jegelka, S. Savarese	Stanford University, MIT	USA	2016	CVPR
13	Deep clustering: Discriminative embeddings for segmentation and separation [34]	J. Hershey, Z. Chen, J. Le Roux, S. Watanabe	Mitsubishi, Columbia University	USA	2016	ICASSP
14	Learning Deep Embeddings with Histogram Loss [113]	E. Ustinova, V. Lempitsky	Skoltech	Russia	2016	NIPS
15	Local Similarity-Aware Deep Feature Embedding [37]	C. Huang, C. C. Loy, X. Tang	The Chinese University of Hong Kong, SenseTime Group Limited	China	2016	NIPS
16	Metric Learning with Adaptive Density Discrimination [86]	O. Rippel, M. Paluri, P. DollĞr, L. D. Bourdev	Facebook	USA	2016	ICLR
17	L2-constrained Softmax Loss for Discriminative Face Verification [84]	R. Ranjan, C. D. Castillo, R. Chellappa	UMIACS	USA	2017	ArXiv
18	In Defense of the Triplet Loss for Person Re-Identification [33]	A. Hermans, L. Beyer, B. Leibe	RWTH	Germany	2017	ArXiv
19	Deep Metric Learning with Angular Loss [117]	J. Wang, F. Zhou, S. Wen, X. Liu, Y. Lin	Baidu	China	2017	ICCV
20	No Fuss Distance Metric Learning Using Proxies [70]	Y. Movshovitz-Attias, A. Toshev, T. Leung, S. Ioffe, S. Singh	Google	USA	2017	ICCV
21	Sampling Matters in Deep Embedding Learning [64]	R. Manmatha, C. Y. Wu, A. Smola, P. KrŁhenb§hl	UT Austin, Amazon	USA	2017	ICCV
22	Deep Metric Learning via Facility Location [99]	H. O. Song, S. Jegelka, V. Rathod, K. Murphy	Google	USA	2017	CVPR
23	Deep spectral clustering learning [53]	M. Law, R. Urtasun, R. Zemel	University of Toronto	Canada	2017	ICML
24	Hard-Aware Deeply Cascaded Embedding [133]	Y. Yuan, K. Yang, C. Zhang	MOE, Peking University, DeepMotion, Microsoft Research	China	2017	ICCV
25	PPFNet: Global Context Aware Local Features for Robust 3D Point Matching [17]	H. Deng, T. Birdal, S. Ilic	TMU, NUDT	Germany, China	2018	CVPR
26	Ranked List Loss for Deep Metric Learning [121]	X. Wang, Y. Hua, E. Kodirov, G. Hu, R. Garnier, N. Robertson	Anyvision, QueenÕs University Belfast	UK	2019	CVPR
27	Multi-Similarity Loss with General Pair Weighting for Deep Metric Learning [119]	X. Wang, Xintong Han, W. Huang, D. Dong, M. Scott	Malong Technologies	China	2019	CVPR
28	A Simple and Effective Framework for Pairwise Deep Metric Learning [82]	Q. Qi, Y. Yan, Z. Wu, X. Wang, T. Yang	The Chinese University of Hong Kong	China	2019	ECCV

29	Deep Metric Learning Meets Deep Clustering: An Novel Unsupervised Approach for Feature Embedding [73]	B. X. Nguyen, B. D. Nguyen, G. Carneiro, E. Tjiputra, Q. D. Tran, T. T. Do	AIOZ	Singapore	2020	ArXiv
30	Exponential triplet loss [110]	E. Urtans, A. Nikitenko, V. Vecins	RTU	Latvia	2020	ICCDA

In, Table 3 information about novel loss functions and their properties regarding DML have been listed. Embedding space refers to normalization or measurement methods between two or more vectors in a latent space. Each of the embedding vectors has been produced by a deep learning based model for the data point. Then two or more embedding vectors have been processed using the loss function and a deep learning based model weights are calculated using the back-propagation algorithm. In addition, for many of these papers sample mining methods are used to select the best training samples to improve the results and speed of the training.

Table 3:

Novel loss functions of studies regarding DML.

No	Title	Year	Embedding space	Sample Mining	Loss function
1	Signature Verification Using A "Siamese" Time Delay Neural Network [7]	1993	Euclidean	None	Contrastive loss
2	Neighbourhood Components Analysis [27]	2004	Euclidean, Mahalanobis	None	NCA Loss
3	Learning a Similarity Metric Discriminatively, with Application to Face Verification [14]	2005	L1, Euclidean	None	Contrastive Loss
4	Distance metric learning for large margin nearest neighbor classification [127]	2005	Euclidean, Mahalanobis	None	Triplet Hinge Loss
5	Large scale metric learning from equivalence constraints [46]	2012	Mahalanobis	None	KISS-BCE Loss
6	Quadruplet-Wise Image Similarity Learning [52]	2013	Qwise	None	Quadruplet Hinge Loss
7	Reidentification by Relative Distance Comparison [135]	2013	RDC	None	RDC Loss
8	Deep Metric Learning for Practical Person Re-Identification [132]	2014	Cosine distance	Hard	Binomial Deviance Loss
9	FaceNet: A unified embedding for face recognition and clustering [23]	2015	L2, Euclidean	Hard, Semi-Hard	Triplet Loss, Harmonic Triplet Loss
10	Improved Deep Metric Learning with Multi-class N-pair Loss Objective [98]	2016	L2, Cosine distance	N Hard Mining	multi-class N-pair loss
11	A Discriminative Feature Learning Approach for Deep Face Recognition [128]	2016	Cosine distance	None	Center loss

12	Deep Metric Learning via Lifted Structured Feature Embedding [100]	2016	L2, Euclidean	Mining positives	Lifted Structured Loss, Lifted Struct
13	Deep clustering: Discriminative embeddings for segmentation and separation [34]	2016	L2, Euclidean	None	Pairwise metric Loss
14	Learning Deep Embeddings with Histogram Loss [113]	2016	Cosine distance	None	Histogram Loss
15	Local Similarity-Aware Deep Feature Embedding [37]	2016	PDDM	Hard mining	PDDM - Double Header Hinge Loss
16	Metric Learning with Adaptive Density Discrimination [86]	2016	Euclidean	Neighbourhood Sampling	Magnet Loss
17	L2-constrained Softmax Loss for Discriminative Face Verification [84]	2017	Cosine distance	None	L2 constrained Softmax Loss
18	In Defense of the Triplet Loss for Person Re-Identification [33]	2017	L2, Euclidean	None	Batch All Triplet Loss
19	Deep Metric Learning with Angular Loss [117]	2017	Angle	None	Angular loss
20	No Fuss Distance Metric Learning Using Proxies [70]	2017	L2, Euclidean	None	Proxy Ranking Loss, Proxy NCA Loss
21	Sampling Matters in Deep Embedding Learning [64]	2017	L2, Euclidean	Distance weighted sampling	Triplet Loss, Contrasitve Loss
22	Deep Metric Learning via Facility Location [99]	2017	L2, Euclidean	None	Struct Clust, Clustering Loss
23	Deep spectral clustering learning [53]	2017	L2, Euclidean	None	Spectral Clustering Loss
24	Hard-Aware Deeply Cascaded Embedding [133]	2017	Euclidean	Model-based	Any / Contrastive loss
25	PPFNet: Global Context Aware Local Features for Robust 3D Point Matching [17]	2018	L2, Euclidean	None	N-Tuple loss
26	Ranked List Loss for Deep Metric Learning [121]	2019	Euclidean	Hard	Ranked List Loss
27	Multi-Similarity Loss with General Pair Weighting for Deep Metric Learning [119]	2019	Cosine distance	Hard	Multi-Similarity Loss
28	A Simple and Effective Framework for Pairwise Deep Metric Learning [82]	2019	Euclidean	TopK Loss mining	DRO-TopK Loss
29	Deep Metric Learning Meets Deep Clustering: An Novel Unsupervised Approach for Feature Embedding [73]	2020	L2, Euclidean	None	Unsupervised UDML Loss
30	Exponential triplet loss [110]	2020	Unit-Range	Hard	Exponential Triplet Loss

In Fig. 4 relationship of DML Loss functions has been summarized. Colours denote similar groups of loss functions by their origin and methodology. It is possible to observe that most of the loss functions come from the seminal works of Contrastive Loss [7], NCA Loss [27], and Triplet Loss [127]. Most of the functions are extensions of simple Hinge Loss [127]. As seen in Table 3, most of the loss functions use sample mining methods, because they are trained only using a few data samples per training iteration. Some methods like Histogram Loss [113] or Quadruplet Hinge Loss [52] use more samples per training iteration, but their results on benchmark datasets are not significantly better than other methods as seen in Table 5.

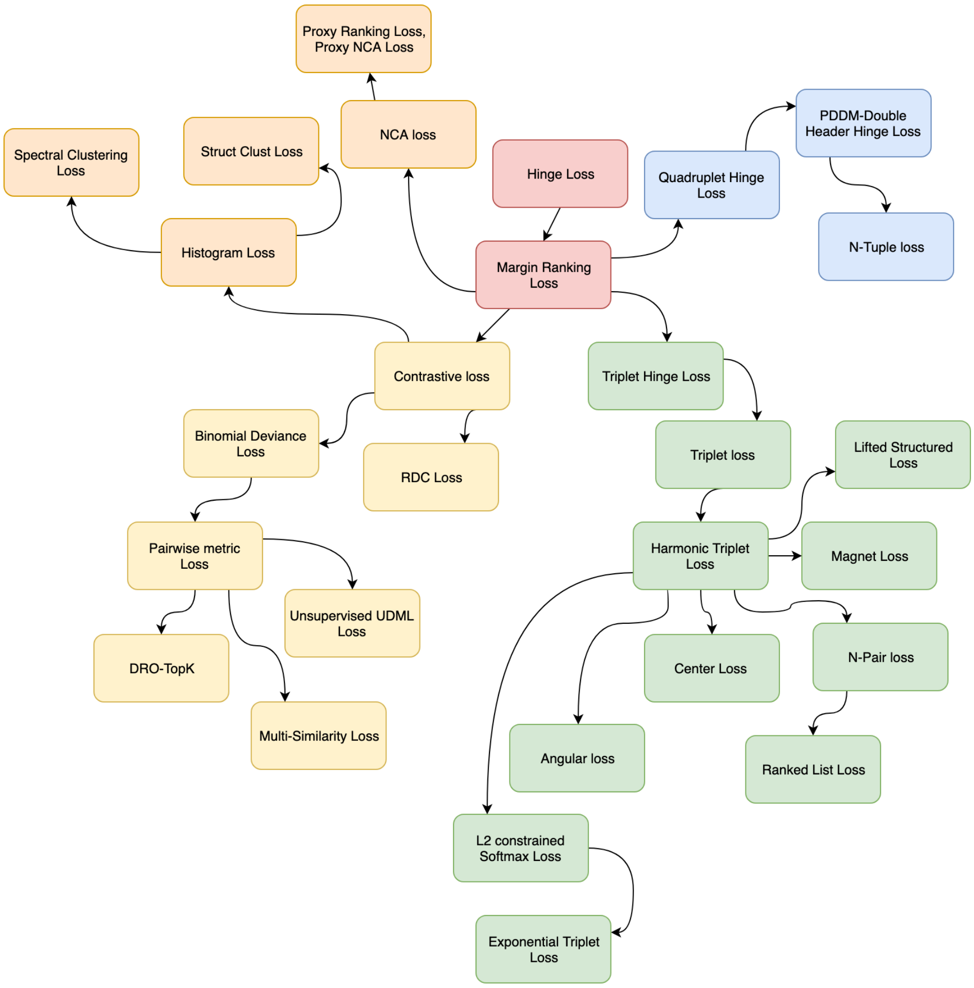

Figure 4: Relationship of DML Loss functions. Colours denote similar groups of loss functions by their origin and methodology.

Table 5 lists practical applications for each of DML loss functions that have been studied, as well as their benchmark datasets and the best results on those datasets. Where applicable, Top-1 accuracy has been selected for the best results on each of the datasets. As seen in the listings, most of the practical applications and datasets have been used for face and product re-identification.

Table 5:

Practical applications and best results for every dataset.

No	Title	Year	Practical application	Dataset / Top-1 Acc.
1	Signature Verification Using A "Siamese" Time Delay Neural Network [7]	1993	Signature re-identification	Signatures: 97%
2	Neighbourhood Components Analysis [27]	2004	Handwriting identification, Face re-identification	USPS: 85% FERET-B
3	Learning a Similarity Metric Discriminatively, with Application to Face Verification [14]	2005	Face Re-identification	AT&T: 92.5%
4	Distance metric learning for large margin nearest neighbor classification [127]	2005	Handwriting identification, text classification	MNIST: 98.8% Letters: 96.3% 20news: 92% Isolet: 96.6% YaleFaces: 93.9%
5	Large scale metric learning from equivalence constraints [46]	2012	Face Re-identification, Image Re-idenification	LFW: 80.5% VIPeR: 22%
6	Quadruplet-Wise Image Similarity Learning [52]	2013	Product or image retrieval	OSR: 74.6% Pubfig: 77.6%
7	Reidentification by Relative Distance Comparison [135]	2013	Face Re-identification	ETHZ: 61.58% i-LIDS: 32.60% VIPeR: 9.12%
8	Deep Metric Learning for Practical Person Re-Identification [132]	2014	Face Re-identification	VIPER: 34.49%
9	FaceNet: A unified embedding for face recognition and clustering [23]	2015	Face Re-identification	LFW: 99.63% YTF: 95.12%
10	Improved Deep Metric Learning with Multi-class N-pair Loss Objective [98]	2016	Product image retrieval, Face Re-identification	LFW: 98.33% SOP: 28.19% CAR-196: 33.5% CUB-200: 27.24%
11	A Discriminative Feature Learning Approach for Deep Face Recognition [128]	2016	Face Re-identification	LFW: 99.28% YTF: 94.9% MegaFace: 76.5%
12	Deep Metric Learning via Lifted Structured Feature Embedding [100]	2016	Product or image retrieval	CUB200: 55%, CARS196: 48%, SOP: 62%
13	Deep clustering: Discriminative embeddings for segmentation and separation [34]	2016	Speaker diarization, seperation	WSJ0: 2.74 dB (SDR)
14	Learning Deep Embeddings with Histogram Loss [113]	2016	Product or image retrieval	CUHK03: 65.7% CUB-200: 51% Market-1501: 59.47% SOP: 65%
15	Local Similarity-Aware Deep Feature Embedding [37]	2016	Product or image retrieval	CARS196: 57.4% CUB-200: 58.3% ImageNet: 48.2%
16	Metric Learning with Adaptive Density Discrimination [86]	2016	Image classification, Face Re-identification	Stanford Dogs: 75.1% Flowers-102: 91.4% Oxford-IIIT Pet: 89.4% ImageNet: 84.1%
17	L2-constrained Softmax Loss for Discriminative Face Verification [84]	2017	Image classification, Face Re-identification	LFW: 99.33% YTF: 99.78% MNIST: 99.05% IJB-A: 97.5%
18	In Defense of the Triplet Loss for Person Re-Identification [33]	2017	Product image retrieval, Face Re-identification	MARS: 90.53%, Market-1501: 79.8%, CUHK03: 87.58%

19	Deep Metric Learning with Angular Loss [117]	2017	Product or image retrieval	CAR-196: 71.4%, CUB-200: 54.7%, SOP: 70.9%
20	No Fuss Distance Metric Learning Using Proxies [70]	2017	Product or image retrieval	CARS196: 73:22% CUB200: 73.22% SOP: 73.73%
21	Sampling Matters in Deep Embedding Learning [64]	2017	Product or image retrieval, Face Re-identification	CARS196: 86.9% CUB200: 63.9% SOP: 72.7%
22	Deep Metric Learning via Facility Location [99]	2017	Product or image retrieval	CARS196: 58.11% CUB200: 48.18% SOP: 67.02%
23	Deep spectral clustering learning [53]	2017	Product or image retrieval	CARS196: 73.07% CUB200: 43.22% SOP: 67.59%
24	Hard-Aware Deeply Cascaded Embedding [133]	2017	Product or image retrieval	CARS196: 83.8% CUB-200: 60.7% In-shop: 62.1 % SOP: 70.1%
25	PPFNet: Global Context Aware Local Features for Robust 3D Point Matching [17]	2018	3D Point Cloud matching	SUN3D: 71%
26	Ranked List Loss for Deep Metric Learning [121]	2019	Product or image retrieval	CARS196: 82.1% CUB-200: 61.3% SOP: 79.8%
27	Multi-Similarity Loss with General Pair Weighting for Deep Metric Learning [119]	2019	Product or image retrieval	CARS196: 77.3% CUB-200: 65.7% In-Shop: 78.2%
28	A Simple and Effective Framework for Pairwise Deep Metric Learning [82]	2019	Product or image retrieval	In-shop: 91.3% CARS-196: 86.2% CUB-200: 68.1%
29	Deep Metric Learning Meets Deep Clustering: An Novel Unsupervised Approach for Feature Embedding [73]	2020	Product or image retrieval	CUB200: 47.5%, Car196: 42.6%
30	Exponential triplet loss [110]	2020	Face Re-identification, Image Re-idenification	VGGFace2: 85.7% EMNIST: 86% FMNIST: 93.1% CIFAR10: 87.3% MNIST: 99.6%

The Quality Assessment (QA) criteria are as follows:

- QA1: Does the publication provide open-source implementation of a novel loss function or methodology?
- QA2: Has a publication achieved state-of-the-art results on the datasets it studies?
- QA3: Does the publication provide a theoretical proof of a novel loss function or methodology?
- QA4: Does the publication include an ablation study to test effects on the results of functional parts one by one?
- QA5: Does a publication have over 100 citations?

Table 7:

Evaluation of quality of publications baset on criteria.

No	Title	QA1	QA2	QA3	QA4	QA5	Total
11	A Discriminative Feature Learning Approach for Deep Face Recognition [128]	Yes	Yes	No	Yes	Yes	4
18	In Defense of the Triplet Loss for Person Re-Identification [33]	Yes	Yes	No	Yes	Yes	4
21	Sampling Matters in Deep Embedding Learning [64]	No	Yes	Yes	Yes	Yes	4
23	Deep spectral clustering learning [53]	No	Yes	Yes	Yes	Yes	4
28	A Simple and Effective Framework for Pairwise Deep Metric Learning [82]	Yes	Yes	Yes	Yes	No	4
3	Learning a Similarity Metric Discriminatively, with Application to Face Verification [14]	No	Yes	Yes	No	Yes	3
5	Large scale metric learning from equivalence constraints [46]	No	Yes	Yes	No	Yes	3
9	FaceNet: A unified embedding for face recognition and clustering [23]	No	Yes	No	Yes	Yes	3
16	Metric Learning with Adaptive Density Discrimination [86]	Yes	Yes	No	No	Yes	3
17	L2-constrained Softmax Loss for Discriminative Face Verification [84]	No	Yes	No	Yes	Yes	3
19	Deep Metric Learning with Angular Loss [117]	No	Yes	Yes	No	Yes	3
20	No Fuss Distance Metric Learning Using Proxies [70]	No	Yes	Yes	No	Yes	3
22	Deep Metric Learning via Facility Location [99]	No	Yes	Yes	No	Yes	3
25	PPFNet: Global Context Aware Local Features for Robust 3D Point Matching [17]	No	Yes	No	Yes	Yes	3
27	Multi-Similarity Loss with General Pair Weighting for Deep Metric Learning [119]	Yes	No	No	Yes	Yes	3
1	Signature Verification Using A "Siamese" Time Delay Neural Network [7]	No	Yes	No	No	Yes	2
4	Distance metric learning for large margin nearest neighbor classification [127]	No	No	Yes	No	Yes	2
6	Quadruplet-Wise Image Similarity Learning [52]	No	No	Yes	Yes	No	2
7	Reidentification by Relative Distance Comparison [135]	No	Yes	No	No	Yes	2
8	Deep Metric Learning for Practical Person Re-Identification [132]	No	No	Yes	Yes	Yes	2
10	Improved Deep Metric Learning with Multi-class N-pair Loss Objective [98]	No	No	No	Yes	Yes	2
12	Deep Metric Learning via Lifted Structured Feature Embedding [100]	No	Yes	No	No	Yes	2

14	Learning Deep Embeddings with Histogram Loss [113]	Yes	No	No	No	Yes	2
24	Hard-Aware Deeply Cascaded Embedding [133]	Yes	No	No	No	Yes	2
26	Ranked List Loss for Deep Metric Learning [121]	No	Yes	No	Yes	No	2
29	Deep Metric Learning Meets Deep Clustering: An Novel Unsupervised Approach for Feature Embedding [73]	No	No	Yes	Yes	No	2
30	Exponential triplet loss [110]	Yes	Yes	No	No	No	2
2	Neighbourhood Components Analysis [27]	No	No	No	No	Yes	1
13	Deep clustering: Discriminative embeddings for segmentation and separation [34]	No	No	No	No	Yes	1
15	Local Similarity-Aware Deep Feature Embedding [37]	No	No	No	No	Yes	1

After reviewing over 30 publications in the field of DML, the following answers have been found to the research questions (RQ):

- RQ1: In this study, 27 types of loss functions for DML have been identified. They have been categorized and listed in their historical order in Fig. 4. All the DML loss functions originate from Margin Ranking Loss, which itself is a variant of earlier Hinge Loss functions [127]. Then most of the newer loss functions originate from Contrastive loss [7], Triplet Loss [23], Histogram Loss [113], and Quadruplet Hinge Loss [52]. For most of the publications included in the study research subject is either the loss function itself or the sample mining methodology.

- RQ2: Latest loss functions and sample mining strategy achieve significantly better results than the previous functions as seen in Table 5. Also, datasets used in experiments have changed over time, but practical applications like image re-identification have not.

- RQ3: Most of the novel loss functions do not have theoretical explanations or derivations of the novel loss functions used in the model, but nonetheless some loss functions like Contrastive Loss [14], Triplet Hinge Loss [127], KISS-BCE Loss [46], Quadruplet Hinge Loss [52] and Binomial Deviance Loss [132] do have theoretical proof. Most of the other loss functions discussed in this study have their grounding in empirical experiments.

- RQ4: A number of significant limitations of DML loss functions and methods have been found in this study. Most of the loss functions require hyper-paramter α that is, a margin between clusters, but in realistic datasets this might not be equal for all classes. Some classes might have more variance than others. Some efforts have been made to resolve the issue like Proxy NCA Loss [70], but even this loss function requires hyper-parameter tuning and prior knowledge of class distributions.

  Intra-class similarities are also a significant problem. Most of DML loss functions ignore the fact that the same class samples also have their own distributions of similarities. Some works address this problem, but it still not fully solved [86] [52] [121].

  Sample mining strategies also are a major problem as they require significant computing resources dedicated just for selecting the best samples to train the model and apply the loss function. Multiple sampling strategies have been developed like Hard [23], Semi-Hard [23], N Hard Mining [98], Neighborhood Sampling [86], Distance weighted sampling [64] and others, but the problem is still not yet fully solved.

  Choice of the number of dimensions of embedding vectors and their embedding space also is a problem that still needs more studies. Publications differ in suggestions, how many dimensions to choose, and what normalization methods to apply to embeddings. Typical method is to use Euclidean distance with L2 normalized embeddings with high dimensionality of at least 128 dimensions [23], but some of the latest papers propose also alternative embedding space normalization [37] [117] and lower number of dimensions per embedding [110].

  Another significant limitation is the computing resources required to reach higher accuracy in re-identification tasks, some earlier works from 2015 required over 2000 CPU hours to reach the highest accuracy on face re-identification tasks [23]. Latest works have been using GPUs to accelerate and parallelize training, but even nowadays as datasets grow larger that requires expensive GPU hardware [110].

2.3 Conclusions of Literature Review

The results of the literature review indicate that the loss function design is an active research topic in DML. Many of the DML loss functions have been designed using empirical experiments instead of the derivation of classical mathematical theories.

DML loss functions come from Hinge loss and Ranking loss functions that have been around since the early days of computing. Then most of the modern loss functions originated from two major approaches, either by using triplets and Triplet Hinge Loss [127] or by using pairs and Contrastive Loss [7]. There have also been studies to use quadruplets or even more permutations of samples, but none has been as effective as Contrastive Loss and Triplet Loss based methods. The last group of DML methods are related to NCA [27] and other dimension reduction methods that produce similar outcomes like DML, but would not work as well for zero-shot learning settings.

Other significant areas of research have been identified for sample mining methods and embedding space normalization functions. Sample mining methods are necessary so that the model would use only difficult data samples in a loss function to improve the convergence speed of a loss function. The most common sampling methods are the Hard sample mining and the Semi-Hard sample mining [23]. Also, normalization of the embedding space is important, because the embeddings should be bounded to a predictable range of values to fill evenly the embedding space with clusters of different known and unknown classes. The most common method for normalization is to use Euclidean distance with L2 normalization, but some studies also have used cosine distances, Mahalanobis distance, angle in degrees, etc. DML is applicable to different types of practical solutions like re-identification of faces, speakers, signatures, handwriting, and product images.

3 IMPROVING THE PERFORMANCE OF FUNCTIONS USING DEEP LEARNING MODELS

This section of the Thesis introduces a novel deep learning-based approach to optimize the performance of well-studied functions. As a practical application, the Value Iteration algorithm has been used. It finds the shortest path from any position in the map to the target position in the map. Its performance degrades exponentially with the larger input size of the map as it cannot be executed in parallel, but novel iterative deep learning-based models can produce comparable results using parallelized architectures and achieve higher performance on larger maps.

The problem domain of Value Function and Value Iteration algorithm learning has been described in Subsection 3.1, then the existing Deep Learning based methods that can be used to model Value Function are shown in Subsection 3.2 and in Subsection 3.3. Finally, a novel method to model Value Function has been presented in Subsection 3.4. The results of these methods have been shown in Subsection 6.1.

3.1 Value Iteration Algorithm

Value Iteration Algorithm (VI) is used in classical reinforcement learning tasks to find an optimal policy for any problem within a fully observable environment. It can take into account the state transition model when the transition is uncertain [88]. VI is often used for finding the optimal path in maps with discrete states. A path finding task formalizes and discretizes a natural terrain and obstacles of the environment. Often this information is gathered using sensors that are attached to the mobile robot. These sensors might include LIDAR (Light Detection and Ranging), ultrasonic sensors for distance measurement, or IMU (Internal Measurement Units), etc. Discretization of a map is usually done by generating an occupancy grid.

VI is an iterative algorithm that repeatedly applies the same Value function of Equation (3) over all positions of a map to find the cumulative value of each cell position, as shown in Fig. 5.

Then the gradient between the values of these positions gives the policy of the optimal path. The policy of the optimal path enables an agent to find its way from any state in a discretized map to a positive terminal state. For